The Pennsylvania State University
The Graduate School

IDENTIFICATION, ESTIMATION AND TESTING IN EMPIRICAL GAMES

A Dissertation in
Economics
by
Nianqing Liu

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2013

The dissertation of Nianqing Liu was reviewed and approved* by the following:

Quang H. Vuong
Professor of Economics
Dissertation Advisor, Co-Chair of Committee

Sung Jae Jun
Associate Professor of Economics
Co-Chair of Committee

Isabelle M. Perrigne
Professor of Economics
Special Member

Paul Grieco
Assistant Professor of Economics

Runze Li
Professor of Statistics

Robert Marshall
Professor of Economics
Head of the Department of Economics

*Signatures are on file in the Graduate School.

Abstract

This dissertation consists of three chapters.

CHAPTER 1: Rationalization and Identification of Discrete Games with Correlated Types

This paper studies the rationalization and identification of discrete games where players have correlated private information. Our approach is fully nonparametric. First, under monotone pure strategy Bayesian Nash Equilibrium, we characterize all the restrictions if any on the distribution of players' choices imposed by the game-theoretic model as well as restrictions associated with three assumptions that have been frequently used in the empirical analysis of discrete games. Namely, we consider additive separability of the private information in the payoffs, exogeneity of the payoff shifters relative to the private information, and mutual independence of the private information conditional on the payoff shifters. Second, we study the nonparametric identification of the payoff functions and types distribution under exclusion restrictions and rank conditions. In particular, we show that our structural model is identified up to a location-scale normalization in the separable case. Third, without imposing exclusion restrictions, we characterize the sharp identification region for the payoff functions and types' distribution. Lastly, we discuss possible estimation and testing procedures.

CHAPTER 2: Semiparametric analysis of Binary Games of Incomplete Information

This paper studies the semiparametric identification and estimation of binary games with arbitrary finite number of players under incomplete information. Our approach allows private types to be correlated across players. By focusing on the monotone pure strategy Bayesian Nash Equilibrium (BNE), we show that, in our semiparametric model with linear payoffs, the equilibrium strategies can be represented as a single–agent binary response model. Under weak restrictions, we identify the joint distribution of private information nonparametrically, and the payoff functions in a linear–index setup. Following Klein and Spady (1993), we propose a three–stage procedure for estimating the payoff coefficients and show that our estimator is $\sqrt{n}$–consistent, asymptotically normally distributed. A Monte Carlo experiment shows that our estimator has good properties in moderately sized samples.

CHAPTER 3: Nonparametric Test of Monotonicity of Bidding Strategy in First-price Auctions

This paper develops nonparametric tests of monotonicity of bidding strategy in first price auctions. The monotonicity testing problem is shown to be equivalent to a convexity testing problem, and a root–N consistent test statistic, which measures a distance of integral of inverse bidding strategy from convexity, is proposed. We obtain two types of critical values: one of them is given by the asymptotic distribution, and the other one is given through bootstrap approach. We also show that our testing procedure has the correct size and is consistent.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am deeply indebted to my advisor Quang Vuong for his invaluable advice, guidance and encouragement. He has generously shared his time and expertise with me. I cannot describe how much I have learned from Quang over the past years.

I am very grateful to the members of my committee, Isabelle Perrigne, Sung Jae Jun, Paul Grieco, and Runze Li for their helpful insights, comments and suggestions. Insightful suggestions from Edward Green, Neil Wallace, Kalyan Chatterjee and Joris Pinkse are also gratefully acknowledged.

I would like to thank all of my classmates, senior and junior colleagues for helpful discussions, joint learning and support. In particular, I would like to thank my friend and coauthor Haiqing Xu for his passion and advice on research. I have benefited much from the conversation with him in the past. I would also like to thank Yao Luo, Hongsong Zhang, Huihui Li and Jicheng Liu.

I am most thankful to my wife for her encouragement and sacrifice during my graduate study in Penn State. Her comfort helped me really much during the days of disappointment and stress.

I have been much blessed by the State College Chinese Alliance Church during my stay in State College. It has been a great spiritual home for me. Such home strengthens me when I am weak, gives me rest when I am tired. Through the fellowship with brothers and sisters there, I have experienced much love from God, without whose grace

I cannot finish my doctoral study with so many difficulties.

Dedication

To my wife, Zhihui, and my daughter, Joy.

Chapter 1

Rationalization and Identification of Discrete Games with Correlated Types

1.1 Introduction

Over the last decades, games with incomplete information have been much successful to understand the strategic interactions among agents in the analysis of various economic and social situations. A leading example is auctions with e.g. Vickrey (1961), Riley and Samuelson (1981), Milgrom and Weber (1982) for the theoretical side and Porter (1995), Guerre, Perrigne and Vuong (2000) and Athey and Haile (2002) for the empirical component. In this paper, we study the identification of static binary games of incomplete information where players have correlated types.[1] We also characterize all the restrictions if any imposed by such models on the observables, which are the players' choices. Following the work by Laffont and Vuong (1996) and Athey and Haile (2007) for auctions, our approach is fully nonparametric.

The empirical analysis of discrete games is almost thirty years old. In particular, the range of applications includes, among others, labor force participation (e.g. Bjorn

[1] To simplify, we focus on binary games in this paper. We left the extension of our approach to general discrete games for future research.

and Vuong, 1984, 1985; Kooreman, 1994; Soetevent and Kooreman, 2007), firms' entry decisions (e.g. Bresnahan and Reiss, 1990, 1991; Berry, 1992; Berry and Tamer, 2006; Ciliberto and Tamer, 2009; Jia, 2008). These papers deal with discrete games under complete information. More recently, discrete games under incomplete information have been used to analyze social interactions by Brock and Durlauf (2001a, 2007); Xu (2011); Kline (2012) among others, firm entry and location choices by Seim (2006b), timing choices of radio stations commercials by Sweeting (2009), stock market analysts' recommendations by Bajari, Hong, Krainer, and Nekipelov (2010), capital investment strategies by Aradillas-Lopez (2010) and local grocery markets by Grieco (2011).

Our paper contributes to this literature in several aspects. First, we focus on monotone pure strategy Bayesian Nash equilibria (BNE) throughout. Monotonicity is a desirable property in many applications for both theoretical and empirical reasons. For instance, White, Xu, and Chalak (2011) show that monotone strategies are never worse off than non–monotone strategies in a private value auction model. Xu (2010) shows that in an entry game the only equilibrium is a monotone pure strategy BNE when the strategic effects are reasonably small. Moreover, the recent literature on nonseparable models heavily relies on the monotonicity of the structural functions in latent variables for identification analysis (see, e.g., Matzkin, 2003; Chesher, 2003, 2005; Imbens and Newey, 2009; Jun, Pinkse, and Xu, 2011). On theoretical grounds, Athey (2001) provided seminal results on the existence of a monotone pure strategy BNE whenever a Bayesian game obeys a Spence–Mirlees single–crossing restriction. Relying on the powerful notion of contractibility, Reny (2011) has extended Athey's results and related results by McAdams (2003) to give weaker conditions ensuring the existence of a monotone pure strategy BNE. Using Reny's results, we show the existence of a monotone pure strategy BNE under a high–level assumption in our setting. In particular, such a high–level condition is satisfied if the game is of strategic complement and private information are positively regression dependent.

Second, we allow players' private information (i.e. types) to be correlated. Allowing correlated private information is motivated primarily by empirical concerns. In oligopoly entry, for example, the correlation among types allows us to see "whether entry occurs because of unobserved profitability that is independent of the competition effect" (Berry and Tamer, 2006). In contrast, mutual independence of private information has been widely assumed in the empirical literature. See, e.g., Brock and Durlauf (2001a); Pesendorfer and Schmidt-Dengler (2003); Seim (2006b); Aguirregabiria and Mira (2007); Sweeting (2009); Bajari, Hong, Krainer, and Nekipelov (2010); Tang (2010); De Paula and Tang (2012); Lewbel and Tang (2012); exceptions include Aradillas-Lopez (2010) and Wan and Xu (2010). Such an independence of types is a convenient assumption for identification, but imposes strong restrictions such as the mutual independence of players' choices, a property which can be invalidated by the data.[2] On the other hand, when private information are correlated, the BNE solution concept requires that each player's beliefs on rivals' choices depend on her private information, which invalidates the usual two–step identification/estimation procedure, see, e.g., Bajari, Hong, Krainer, and Nekipelov (2010). With such type–dependent beliefs, Wan and Xu (2010) establish some upper/lower bounds of beliefs in a semiparametric setting with linear–index payoffs. They nonparametrically estimate these bounds in the first step and then apply a modified maximum score estimator approach (see Manski and Tamer, 2003) to construct a set estimator for the payoff coefficients. Alternatively, Aradillas-Lopez (2010) adopts a different equilibrium concept suggested by Aumann (1987), in which each player's equilibrium beliefs do not rely on her private information.

Third, our analysis is fully nonparametric in the sense that players' payoffs and the joint distribution of the players' private information are subject to some mild smoothness conditions only. As far as we know, with the exception of De Paula and Tang (2012) and Lewbel and Tang (2012), every paper analyzing empirical discrete games has imposed parametric restrictions on the payoffs and/or the distribution of private information. For instance, Brock and Durlauf (2001a); Seim (2006b); Sweeting (2009) and Xu (2010) have specified both payoffs and the private information distribution parametrically. In a semiparametric context, Aradillas-Lopez (2010); Tang (2010) and Wan and Xu (2010), among others, parameterize players' payoffs, while Bajari, Hong, Krainer, and Nekipelov (2010) parameterize the private information distribution. On the other hand,

[2] A model with unobserved heterogeneity and independent private information also generates dependence among players' choices conditional on observed regressors (see Grieco, 2011).

De Paula and Tang (2012) and Lewbel and Tang (2012) do not introduce any parameter but impose some restrictions on the payoffs' functional form. In particular, they impose multiplicative separability in the strategic effect and assume that it is a known function of the other players' choices. In addition to being fully nonparametric, we do not require either that players' private information enter additively in the payoffs. Consequently, our baseline discrete game is the most general one and closest to that considered in game theory. We show that such a model imposes essentially no restrictions on the distribution of players' choices. In other words, monotone pure strategy BNE can explain almost all observed choice probabilities in discrete games.

In view of the preceding results, we consider three assumptions that have been frequently used in the empirical analysis of discrete games. First, we consider the assumption that private information enters additively in the player's payoff. To the best of our knowledge, such an assumption has been made in every paper analyzing discrete games empirically. We show again that the resulting model imposes essentially no restrictions on the distribution of players' choices. We also show that the players' payoffs and the joint distribution of the players' private information are not identified nonparametrically whether the private information are additively separable or not.

A second assumption that has been frequently imposed in empirical work is the exogeneity of some variables shifting the players' payoffs relative to players' private information. Papers using such an assumption are, e.g., Brock and Durlauf (2001a); Seim (2006b); Sweeting (2009); Aradillas-Lopez (2010); Bajari, Hong, Krainer, and Nekipelov (2010); De Paula and Tang (2012) and Lewbel and Tang (2012). We show that the resulting model restricts the distribution of players' choices conditional upon the payoff shifters and we characterize all those restrictions. Specifically, the exogeneity assumption restricts the joint choice probability to be a monotone function of the corresponding marginal choice probabilities. With the exogeneity assumption, we show that one can identify the equilibrium belief of the player at the margin under a mild support condition. We then characterize the partially identified set of payoffs and the distribution of private information under the exogeneity assumption and the support condition. In particular, the partially identified region is unbounded and quite large unless we impose additional restrictions on the payoffs' functional form.

Further, we consider some identifying restrictions, namely some exclusion restrictions and rank conditions, to achieve identification in both separable and non–separable cases. We show that the copula function of the types' distribution is identified on an appropriate support in the nonseparable case. Then, the players' payoffs are identified up to scale for each fixed value of the exogenous state variables, as well as up to the marginal distributions of players' private information. With additive separability in payoffs, however, we show that both the players' payoffs and distribution of private information are identified up to a scale-location normalization. Our identification results can be viewed as an extension to game theoretic models of the nonparametric analysis in traditional threshold–crossing models considered by, e.g., Matzkin (1992). An important difference is that the discrete game setup allows us to exploit the exclusion restrictions for identification. Such restrictions have been used frequently in the empirical analysis of discrete games. See, e.g., Aradillas-Lopez (2010); Bajari, Hong, Krainer, and Nekipelov (2010); Wan and Xu (2010); Lewbel and Tang (2012); exceptions include De Paula and Tang (2012).

For completeness, we consider a third assumption, namely the mutual independence of players' private information. Specifically, we characterize all the restrictions imposed by exogeneity and mutual independence as considered by Brock and Durlauf (2001a); Seim (2006b); Sweeting (2009); Bajari, Hong, Krainer, and Nekipelov (2010); Lewbel and Tang (2012), among others. We show that the only restriction is that the players' choices are mutually independent conditionally on the payoff shifters. In particular, we show that the restrictions imposed by mutual independence are stronger than those imposed by exogeneity. In other words, exogeneity is redundant in terms of explaining players' choices as soon as mutual independence is imposed.

The paper is arranged as follows. We introduce our baseline model in Section 2. We define and establish the existence of a monotone pure strategy BNE. We also characterize such equilibrium strategies under additive separability of the private information. In Section 3, we study the restrictions imposed by the baseline model, whether the private information are additively separable or not. We also derive all the restrictions imposed by the exogeneity and mutual independence assumptions. In Section 4, we establish the identification of the belief of the player at the margin under exogeneity only. We then establish the nonparametric identification of the model primitives for the additively nonseparable and separable cases under some exclusion restriction and rank conditions. In Section 5, without exclusion restrictions, we study the partial identification of the payoffs under both additive separability and non–separability in private information, respectively. Section 6 concludes with a discussion of the case when the same monotone pure strategy BNE is not played across identical games. We also discuss briefly estimation and testing in the nonseparable and separable cases.

1.2 Monotone Pure Strategy BNE

We consider a discrete game of incomplete information. There is a finite number of players, indexed by $i = 1, 2, \cdots, I$. Each player simultaneously chooses a binary action $Y_i \in \{0, 1\}$. Let $\mathcal{A} = \{0, 1\}^I$ be the space of possible actions for all players and $Y = (Y_1, \cdots, Y_I) \in \mathcal{A}$ be an action profile. Following the convention, let $\mathcal{A}_{-i}$ and Y_{-i} denote the action space and a profile of actions for all players except i, respectively. Let $X \in \mathscr{S}_X \subset \mathbb{R}^d$ be a payoff relevant random vector, which is publicly observed by all players and also by the researcher.[3] For instance, X can include individual characteristics of the players as well as specific variables for the game. For each player i, we further assume that the random variable $U_i \in \mathbb{R}$ is her private information which is not observed by other players. Let $U = (U_1, \cdots, U_I)$ and $F_{U|X}(\cdot|\cdot)$ be the conditional distribution function of U given the state variable X. The conditional distribution $F_{U|X}(\cdot|\cdot)$ is assumed to be common knowledge.[4]

The payoff for player i is described as

$$\Pi_i(Y, X, U_i) = \begin{cases} \pi_i(Y_{-i}, X, U_i), & \text{if } Y_i = 1, \\ 0, & \text{if } Y_i = 0, \end{cases}$$

[3] Grieco (2011) discusses the case with unobserved heterogeneity in publicly observed state variables.

[4] For a standard notion of common knowledge in game theory, see, e.g., Fudenberg and Tirole (1991), Chapter 14.

where π_i is a structural function in our model. The zero payoff for action $Y_i = 0$ is a standard payoff normalization in binary response models.

Following the literature on Bayesian games, player i's decision rule is a function $Y_i = \delta_i(X, U_i)$, where $\delta_i(\cdot,\cdot) : \mathbb{R}^d \times \mathbb{R} \to \{0, 1\}$ maps all the information that she knows to the binary choice set. For $i = 1, \cdots, I$, let $a_{-i} = (a_1, \cdots, a_{i-1}, a_{i+1}, \cdots, a_I) \in \mathcal{A}_{-i}$. Given a strategy profile $\delta = (\delta_1, \cdots, \delta_I)$, we denote by $\sigma^{\delta}_{-i}(a_{-i}|x, u_i)$ the conditional probability of others choosing $a_{-i} \in \mathcal{A}_{-i}$, i.e.,

$$\begin{aligned}\sigma^{\delta}_{-i}(a_{-i}|x,u_i) &\equiv \mathbb{P}_{\delta}\left(Y_{-i} = a_{-i}|X = x, U_i = u_i\right)\\ &= \int_{\mathbb{R}^{I-1}} \left[\prod_{j\neq i} \mathbf{1}\left\{\delta_j(x,u_j) = a_j\right\}\right] \times f_{U_{-i}|X,U_i}(u_{-i}|x,u_i)du_{-i}\end{aligned}$$

where $\mathbf{1}\{\cdot\}$ is the indicator function and $f_{U_{-i}|X,U_i}(\cdot|\cdot,\cdot)$ is the conditional density function of U_{-i} given X and U_i. Here $\mathbb{P}_{\delta}$ denotes the (conditional) probability measure under the strategy profile δ.

The equilibrium concept we adopt is the pure strategy Bayesian Nash equilibrium (BNE). Mixed strategy equilibria are not considered hereafter, since with probability one, each player has a unique best response in our case. Fix $X = x \in \mathscr{S}_X$. We now characterize BNEs in our discrete game. In equilibrium, player i with $U_i = u_i$ chooses action 1 if and only if her expected payoff is greater than zero, i.e.,

$$\delta^*_i(x,u_i) = \mathbf{1}\left[\sum_{a_{-i}} \pi_i(a_{-i}, x, u_i)\sigma^*_{-i}(a_{-i}|x,u_i) \geq 0\right], \tag{1.1}$$

where $\delta^* = (\delta^*_1, \cdots, \delta^*_I)$ is the equilibrium strategy profile and $\sigma^*_{-i}(a_{-i}|x,u_i)$ is a short notation for $\sigma^{\delta^*}_{-i}(a_{-i}|x,u_i)$. Note that σ^*_{-i} depends on δ^*_{-i}. Hence, eq. (2.2) for $i = 1, \cdots, I$ defines a simultaneous equation system in δ^* referred to as "mutual consistency". A pure strategy BNE is a fixed point δ^* of such a system, which holds for all $u = (u_1, \ldots, u_I)$ in the support $\mathscr{S}_{U|X=x}$. Ensuring equilibrium existence in Bayesian games is a complex and deep subject in the literature. It is well known that a solution of such an equilibrium generally exists in a broad class of Bayesian games (see, e.g., Vives, 1990).

Recently, much attention has focused on monotone pure strategy BNEs, since monotonicity is a natural property in many applications such as auctions, entry, and global games. A monotone pure strategy BNE is defined as follows:

Definition 1 (Monotone pure strategy BNE)**.** *Fix $x \in \mathscr{S}_X$. A pure strategy profile $\delta^*(x) \equiv (\delta_1^*(x,\cdot),\cdots,\delta_I^*(x,\cdot))$, where $\delta_i^*(x,\cdot) : \mathscr{S}_{U_i|X=x} \to \{0,1\}$, is a monotone pure strategy BNE if $\delta^*(x)$ is a BNE and $\delta_i^*(x,\cdot)$ is a monotone function on $\mathscr{S}_{U_i|X=x}$ for every $i = 1,\cdots,I$.*

Monotone pure strategy BNEs are easier to characterize than general BNEs. Fix $X = x$. In a monotone pure strategy BNE, players' strategies can be explicitly defined by a threshold profile $(u_1^*(x),\cdots,u_I^*(x))$ (recall that δ_i^* can take only two values, 0 or 1.) Formally, a monotone pure strategy BNE can be represented by a profile of cutoff values: $u^*(x) \equiv (u_1^*(x),\cdots,u_I^*(x)) \in \mathscr{S}_{U|X=x}$, where $u_i^*(\cdot) : \mathscr{S}_X \to \mathscr{S}_{U_i}$, such that $\delta_i^*(x,u_i) = \mathbf{1}[u_i \le u_i^*(x)]$, or $\delta_i^*(x,u_i) = \mathbf{1}[u_i > u_i^*(x)]$.[5] Without loss of generality, we restrict our attention to monotone decreasing pure strategy (m.d.p.s.) BNEs hereafter. This serves as a normalization. To see this, suppose in a structure $[\pi; F_{U|X}]$, player i's equilibrium strategy is a monotone increasing function for some fixed $x \in \mathscr{S}_X$, i.e. $\delta_i^*(x,\cdot) = \mathbf{1}[\cdot \ge u_i^*(x)]$. We can then construct an observationally equivalent structure $[\tilde{\pi}; \tilde{F}_{U|X}]$ by letting $\tilde{\pi}_i(\cdot,x,u_i) = \pi_i(\cdot,x,-u_i)$ for all $u_i \in \mathbb{R}$ with $\tilde{\pi}_j = \pi_j$ ($j \ne i$) and $\tilde{F}_{U|X}(\cdot|x) = F_{\widetilde{U}|X}(\cdot|x)$, where $\widetilde{U}$ differs from U only in its i–th argument: $\widetilde{U}_i = -U_i$. It can be shown that in equilibrium for the constructed structure, i's strategy is monotone decreasing, i.e. $\tilde{\delta}_i^*(x,\cdot) = \mathbf{1}[\cdot \le -u_i^*(x)]$.

Given equilibrium m.d.p.s. profile of the form $\delta_i^*(x,u_i) = \mathbf{1}[u_i \le u_i^*(x)]$ for $i = 1,\cdots,I$, the mutual consistency defined by eq. (2.2) for a BNE solution requires that $\forall x \in \mathscr{S}_X$,

$$u_i \le u_i^*(x) \Longleftrightarrow \mathbb{E}_{\delta^*}\left[\pi_i(Y_{-i},X,U_i)\middle|X = x, U_i = u_i\right] \ge 0, \tag{1.2}$$

where $\mathbb{E}_{\delta^*}$ denotes the (conditional) expectation under the strategy profile δ^*. Without causing any confusion, we will suppress the subscript δ^* when it is an equilibrium strategy profile. In eq. (1.2), the conditional distribution of Y_{-i} given $X = x$ and $U_i = u_i$,

[5] The left–continuity of strategies considered hereafter is not restrictive given our assumptions below.

i.e. $\mathbb{P}(Y_{-i} = a_{-i}|X = x, U_i = u_i)$, can be written as:

$$\sigma^*_{-i}(a_{-i}|x, u_i) = \mathbb{P}\left[\forall a_j = 1, U_j \leq u^*_j(X); \forall a_j = 0, U_j > u^*_j(X) \middle| X = x, U_i = u_i\right]. \quad (1.3)$$

Under Assumption R below, the $u^*_j(x)$s, if not on the support boundary, are defined by the set of simultaneous equations:

$$\sum_{a_{-i}} \pi_i(a_{-i}, x, u^*_i(x))\sigma^*_{-i}(a_{-i}|x, u^*_i(x)) = 0 \quad (1.4)$$

for $i = 1, \ldots, I$.

The seminal work on the existence of a monotone pure strategy BNE in games of incomplete information was first provided by Athey (2001) in both *supermodular* and *logsupermodular* games, and later extended by McAdams (2003) and Reny (2011). Applying Reny (2011) Theorem 4.1, we establish the existence of monotone pure strategy BNEs in our binary game under some weak regularity assumptions.

Assumption R (Conditional Radon–Nikodym Density). *For every $x \in \mathscr{S}_X$, the conditional distribution of U given $X = x$ is absolutely continuous w.r.t. Lebesgue measure and has a continuous positive conditional Radon–Nikodym density $f_{U|X}(\cdot|x)$ a.e. over the nonempty interior of its hypercube support $\mathscr{S}_{U|X=x}$.*

Assumption R allows the support of U given $X = x$ to be bounded, namely of the form $\times_{i=1,\ldots,I}[\underline{u}_i(x), \overline{u}_i(x)]$ for some finite $\underline{u}_i(x)$ and $\overline{u}_i(x)$ as frequently used when U_i is i's private information, or unbounded as when $\mathscr{S}_{U|X=x} = \mathbb{R}^I$ used typically in binary models. Assumption R can be greatly weakened as shown by Reny (2011).

Assumption M (Monotone Decreasing Expected Payoff). *Fix an arbitrary $x \in \mathscr{S}_X$. For any monotone decreasing pure strategy (m.d.p.s.) profile δ, $\mathbb{E}_\delta\left[\pi_i(Y_{-i}, X, U_i) \middle| X = x, U_i = u_i\right]$ is a monotone decreasing function in $u_i \in \mathscr{S}_{U_i|X=x}$.*

Assumption M guarantees that the best response function is monotone decreasing in u_i if all other players adopt monotone decreasing pure strategies.

Lemma 1. *Suppose that Assumptions R and M hold. For any $x \in \mathscr{S}_X$, there exists an m.d.p.s. BNE.*

Proof. See Appendix A.1.1 □

Lemma 17 in Appendix A.1.2 provides some sufficient primitive conditions for Assumption M. Specifially, we assume positive regression dependence across U_is given X, strategic complementarity of players' actions and non-increasing payoffs in the U_is. Thus, m.d.p.s. BNEs generally exist in a large class of binary games. As far as we know, with the only exception of Aradillas-Lopez and Tamer (2008) and Xu (2010), every paper analyzing empirical discrete games has imposed certain restrictions to guarantee monotone pure strategy BNEs.

Note that Lemma 1 is silent about the existence of BNEs with non–monotone strategies. In a parametric setup, Xu (2010) propose a method to nonparametrically identify subsets of the space of covariates, for which all BNEs are monotone pure strategy equilibria. Lemma 1 does not ensure either that the m.d.p.s. BNE is unique. Throughout our analysis, we assume that only a single m.d.p.s. BNE is played in the data. In the Conclusion, we discuss the case when such an assumption is relaxed.

An assumption made in every paper in the empirical discrete game literature is the additive separability of the error terms in the payoffs.

Assumption S (Additive Separability)**.** *We have* $\pi_i(a_{-i}, x, u_i) = \pi_i(a_{-i}, x) - u_i$ *for every* i, a_{-i}, x *and* u_i.

In Assumption S, the negative sign in front of u_i is only for notational convenience. Assumption S allows us to represent equilibrium strategies as a semi–linear–index binary response model as shown in the following lemma.

Lemma 2. *Suppose that Assumptions R, M and S hold. If an m.d.p.s. BNE is being played, i.e.,* $\delta^* = (\delta_1^*, \cdots, \delta_I^*)$ *where* δ_i^* *is a monotone decreasing function on* $\mathscr{S}_{U_i|X}$*, then equilibrium choices can be written as a semi–linear–index binary response model:*

$$Y_i = \mathbf{1}\left[U_i \leq \sum_{a_{-i}} \pi_i(a_{-i}, X)\sigma^*_{-i}(a_{-i}|X, u_i^*(X))\right], \tag{1.5}$$

Proof. See Appendix A.1.3. □

In particular, when $0 < \mathbb{E}(Y_i|X = x) < 1$ for all i, the profile of thresholds $u^*(x) = (u_1^*(x), \ldots, u_I^*(x))$ is a solution in $\times_{i=1}^{I} \mathscr{S}_{U_i|X=x}$ of the system of I equations:

$$\sum\nolimits_{a_{-i}} \pi_i(a_{-i}, x)\sigma^*_{-i}(a_{-i}|x, u_i^*(x)) = u_i^*(x),\ \forall i = 1, \cdots, I, \tag{1.6}$$

which is a special case of eq. (1.4).[6] The representation in eq. (2.3) of the equilibrium strategies as a semi–linear–index binary response model relates to single-agent binary threshold crossing models studied e.g. by Matzkin (1992). We will show in Section 4 that the belief of the player at the margin σ^*_{-i} in eq. (1.4) and (1.6) can be nonparametrically identified under additional weak conditions. The player at the margin is the one that receives a private information equal to the threshold $u_i = u_i^*(x)$ so that she is indifferent between action 1 and 0.

Example 1. *Consider a 2–by–2 game of incomplete information with the following payoffs,*

$$\Pi_i = \begin{cases} 2 - 4Y_{-i} - U_i & \text{if } Y_i = 1 \\ 0 & \text{otherwise} \end{cases}$$

where the private information $U = (U_1, U_2) \in \mathbb{R}^2$ conforms to a joint normal distribution with mean zero, unit variances and correlation parameter $\rho_0 \in [0, \overline{\rho}]$, $0 < \overline{\rho} < 1$.

An m.d.p.s. BNE characterized by $(u_1^*, u_2^*) \in \mathbb{R}^2$ satisfies the following conditions,

$$u_1 \leq u_1^* \Leftrightarrow 2 - 4 \times \mathbb{P}(U_2 \leq u_2^*|U_1 = u_1) - u_1 \leq 0, \tag{1.7}$$

$$u_2 \leq u_2^* \Leftrightarrow 2 - 4 \times \mathbb{P}(U_1 \leq u_1^*|U_2 = u_2) - u_2 \leq 0. \tag{1.8}$$

From the above conditions, we can derive equations for (u_1^*, u_2^*)

$$2 - 4 \times \Phi\left(\frac{u_2^* - \rho_0 u_1^*}{\sqrt{1 - \rho_0^2}}\right) - u_1^* = 0, \quad 2 - 4 \times \Phi\left(\frac{u_1^* - \rho_0 u_2^*}{\sqrt{1 - \rho_0^2}}\right) - u_2^* = 0.$$

[6] In eq. (1.6) it is understood that $u_i^*(x) = \underline{u}_i(x)$ and $u_i^*(x) = \overline{u}_i(x)$ if $\mathbb{E}_{\delta^*}\left[\pi_i(Y_{-i}, X)\middle|X = x, U_i = \underline{u}_i(x)\right] < \underline{u}_i(x)$ and $\mathbb{E}_{\delta^*}\left[\pi_i(Y_{-i}, X)\middle|X = x, U_i = \overline{u}_i(x)\right] > \overline{u}_i(x)$, respectively.

It follows that

$$2-2\rho_0-4\Phi\left(\frac{u_2^*-\rho_0u_1^*}{\sqrt{1-\rho_0^2}}\right)+4\rho_0\Phi\left(\frac{u_1^*-\rho_0u_2^*}{\sqrt{1-\rho_0^2}}\right)-(u_1^*-\rho_0u_2^*)=0,$$
$$2-2\rho_0-4\Phi\left(\frac{u_1^*-\rho_0u_2^*}{\sqrt{1-\rho_0^2}}\right)+4\rho_0\Phi\left(\frac{u_2^*-\rho_0u_1^*}{\sqrt{1-\rho_0^2}}\right)-(u_2^*-\rho_0u_1^*)=0.$$

Let $b_1^*=\Phi\left(\frac{u_1^*-\rho_0u_2^*}{\sqrt{1-\rho_0^2}}\right)$ *and* $b_2^*=\Phi\left(\frac{u_2^*-\rho_0u_1^*}{\sqrt{1-\rho_0^2}}\right)$. *By definition,* (b_1^*,b_2^*) *is a one–to–one mapping of* (u_1^*,u_2^*). *Moreover, we*

$$2-2\rho_0-4b_2^*+4\rho_0b_1^*-\sqrt{1-\rho_0^2}\Phi^{-1}(b_1^*)=0, \tag{1.9}$$
$$2-2\rho_0-4b_1^*+4\rho_0b_2^*-\sqrt{1-\rho_0^2}\Phi^{-1}(b_2^*)=0. \tag{1.10}$$

Note that the solutions to eqs. (1.9) *and* (1.10) *are necessary but not sufficient for an equilibrium and we apply eqs.* (1.7) *and* (1.8) *to verify each solution candidate. Table 1 describes the set of m.d.p.s. BNE in terms of* (b_1^*,b_2^*). *We can see that the solution* $(0.5,0.5)$ *is not an equilibrium for* $\rho_0=0.6,0.8$.[7] *Indeed, this equilibrium will be "lost" when* $\rho_0\geq\approx 0.54$.

Table 1: the set of m.d.p.s. BNEs at different ρ_0*s*

	$\rho_0=0$	0.2	0.4	0.6	0.8
e_1	(0.0302, 0.9698)	(0.0092, 0.9908)	(0.0026, 0.9974)	(0.0007, 0.9993)	(0.0002,0.9998)
e_2	(0.5, 0.5)	(0.5, 0.5)	(0.5, 0.5)	—	—
e_3	(0.9698, 0.0302)	(0.9908, 0.0092)	(0.9974, 0.0026)	(0.9993, 0.0007)	(0.9998, 0.0002)

1.3 Rationalization

In this section we will study the baseline model defined by Assumptions R and M as well as three other models obtained by imposing additional assumptions frequently made in the empirical game literature such as Assumption S. Specifically, we will characterize all the restrictions imposed on the distribution of observables (Y,X) by each of these models. We will say that a distribution of the observables is rationalized by a model if

[7] By definition, $(b_1^*,b_2^*)=(0.5,0.5)$ corresponds to $(u_1^*,u_2^*)=(0,0)$.

and only if it satisfies all the restrictions of the model. In other words, a distribution of the observables is rationalized if and only if there is a structure (not necessarily unique) in the model that generates such a distribution. In particular, rationalization logicallly precedes identification as the latter, which is addressed in Sections 5 and 6, makes sense only if the data generating process (DGP) can be rationalized by the model under consideration.

Besides Assumption S introduced above, we consider two additional assumptions. The first is the exogeneity of X relative to U, an assumption that has been frequently made in the empirical discrete game literature. See, e.g., Bajari, Hong, Krainer, and Nekipelov (2010); De Paula and Tang (2012) and Lewbel and Tang (2012).

Assumption E (Exogeneity)**.** *X and U are independent of each other.*[8]

Another assumption called as mutual independence has been widely used in the literature. For examples, see an extensive list of references in two recent surveys: Bajari, Hong, and Nekipelov (2010) and de Paula (2012).

Assumption I (Mutual Independence)**.** *$U_1, \cdots, U_I$ are mutually independent conditional on X.*

Let $S \equiv [\pi; F_{U|X}]$. We consider the following models (classes of structures):

$$\begin{aligned}
\mathcal{M}_1 &\equiv \{S : \text{Assumptions R and M hold and} \\
&\qquad\qquad \text{a single m.d.p.s. BNE is played}\}, \\
\mathcal{M}_2 &\equiv \{S \in \mathcal{M}_1 : \text{ Assumption S holds}\}, \\
\mathcal{M}_3 &\equiv \{S \in \mathcal{M}_2 : \text{ Assumption E holds}\}, \\
\mathcal{M}_4 &\equiv \{S \in \mathcal{M}_3 : \text{ Assumption I holds}\}.
\end{aligned}$$

The last requirement in $\mathcal{M}_1$ is not restrictive when there is a unique m.d.p.s. BNE. When this is not the case, we follow most of the literature by assuming that the same equilibrium is played in the DGP for a given x. Relaxing such a requirement has

[8] By conditioning on W, our results can be easily extended to the weaker assumption that Z and U are independent from each other conditional on W, where $X = (W, Z)$.

been addressed in recent work and will be discussed in the Conclusion. Note that $\mathcal{M}_1 \supsetneq \mathcal{M}_2 \supsetneq \mathcal{M}_3 \supsetneq \mathcal{M}_4$.

We introduce some notation for our following analysis. Fix a structure $S \in \mathcal{M}_1$, let $\alpha_i(x) \equiv F_{U_i|X}(u_i^*(x)|x)$. Because equilibrium strategies are monotone decreasing, it is straightforward that $\alpha_i(x) = \mathbb{E}(Y_i|X = x)$. For every $p = 2, \cdots, I$, and $1 \leq i_1 < \cdots < i_p \leq I$, let $C_{U_{i_1},\cdots,U_{i_p}|X}(\cdot, \ldots, \cdot|\cdot)$ be the conditional copula function of $(U_{i_1}, \cdots, U_{i_p})$ given X, i.e. $\forall(\alpha_{i_1}, \cdots, \alpha_{i_p}) \in [0,1]^p$ and $x \in \mathscr{S}_X$,

$$C_{U_{i_1},\cdots,U_{i_p}|X}(\alpha_{i_1}, \cdots, \alpha_{i_p}|x) \equiv F_{U_{i_1},\cdots,U_{i_p}|X}\left(F^{-1}_{U_{i_1}|X}(\alpha_{i_1}|x), \cdots, F^{-1}_{U_{i_p}|X}(\alpha_{i_p}|x)\Big| x\right).$$

The next proposition determines distributions of Y given X that can be *rationalized* by a structure in $\mathcal{M}_1$.

Proposition 1. *A conditional distribution $F_{Y|X}(\cdot|\cdot)$ is rationalized by a structure in $\mathcal{M}_1$ if for every $x \in \mathscr{S}_X$ and $a \in \mathcal{A}$, $\mathbb{P}(Y = a|X = x) = 0$ implies that $\mathbb{P}(Y_i = a_i|X = x) = 0$ for some i.*

Proof. See Appendix A.2.1 □

Thus, $\mathcal{M}_1$ rationalizes all distributions for Y given X that belong to the interior of the simplex in $\mathbb{R}^{2^{I-k}}$ $(0 \leq k \leq I)$. In particular, $\mathcal{M}_1$ rationalizes all the distributions with strictly positive choice probabilities. Proposition 1 also indicates that the only possible distributions that cannot be rationalized by $\mathcal{M}_1$ satisfy $\mathbb{P}(Y = a|X = x) = 0$ for some $a \in \mathcal{A}$, i.e distributions for which there are "structural zeros." As a matter of fact, it is possible to characterize all those distributions that cannot be rationalized by $\mathcal{M}_1$.[9] They arise because of Assumption R. As noted earlier, one can replace the latter by Reny (2011)'s weaker assumptions. Lemma 18 in Appendix A.2.2 then shows that any distribution for Y given X can be rationalized. In other words, our binary game-theoretical model $\mathcal{M}_1$ imposes essentially no restrictions on the distribution of observables.

We now turn to the rationalization of $\mathcal{M}_2$. To do so, we establish the observational equivalence between $\mathcal{M}_1$ and $\mathcal{M}_2$ despite $\mathcal{M}_1 \supsetneq \mathcal{M}_2$.

[9] Such a result is available upon request to the authors.

Lemma 3. *For any given structure $S \equiv [\pi; F_{U|X}] \in \mathcal{M}_1$, there always exists an observationally equivalent structure $\tilde{S} \equiv [\tilde{\pi}; \tilde{F}_{U|X}] \in \mathcal{M}_2$.*

Proof. See Appendix A.2.3 □

The next proposition follows immediately from Lemma 3.

Proposition 2. *A conditional distribution $F_{Y|X}(\cdot|\cdot)$ is rationalized by a structure in $\mathcal{M}_1$ if and only if it is rationalized by a structure in $\mathcal{M}_2$.*

In particular, additive separability of private information in the payoffs (Assumption S) does not impose any additional restrictions relative to model $\mathcal{M}_1$. Moreover, it follows from Proposition 1 that $\mathcal{M}_2$ can still rationalize all distributions for Y given X that belong to the interior of the simplex in $\mathbb{R}^{2^{I-k}}$ with $0 \leq k \leq I$. In other words, $\mathcal{M}_2$ imposes essentially no restrictions on the distribution of observables.

Next, we consider the rationalization of $\mathcal{M}_3$. To do so, we first provide a necessary and sufficient condition for two structures in $\mathcal{M}_2$ to be observationally equivalent.

Lemma 4. *Two structures $S \equiv [\pi; F_{U|X}]$ and $\widetilde{S} \equiv [\tilde{\pi}; \tilde{F}_{U|X}]$ in $\mathcal{M}_2$ are observationally equivalent if and only if for every $x \in \mathscr{S}_X$,*

(i) $\forall\, i = 1, \cdots, I$, *we have* $\alpha_i(x) = \widetilde{\alpha}_i(x)$.

(ii) $\forall\, p = 2, \cdots, I$, *and* $1 \leq i_1 < \cdots < i_p \leq I$, *we have*

$$C_{U_{i_1},\cdots,U_{i_p}|X}\left(\alpha_{i_1}(x), \cdots, \alpha_{i_p}(x)|x)\middle|x\right) = \widetilde{C}_{U_{i_1},\cdots,U_{i_p}|X}\left(\widetilde{\alpha}_{i_1}(x), \cdots, \widetilde{\alpha}_{i_p}(x)|x)\middle|x\right).$$

Proof. See Appendix A.2.4. □

Lemma 4 decomposes the observational equivalence of two structures in $\mathcal{M}_2$ in two aspects: Condition (i) relates the conditional marginal distributions of U_i given X and the payoffs in the two structures; Condition (ii) equates the conditional dependence among the U_is given X between the two structures using the conditional copula function.

Based on Lemma 4, we can now give a necessary and sufficient condition for a structure in $\mathcal{M}_2$ to be observationally equivalent to a structure in $\mathcal{M}_3$.

Lemma 5. *For any given structure $S \in \mathcal{M}_2$, there exists an observationally equivalent structure $\widetilde{S} \in \mathcal{M}_3$ if and only if $\forall p = 2, \cdots, I$, and $\forall\, 1 \leq i_1 < \cdots < i_p \leq I$*

(i) $\forall x \in \mathscr{S}_X$, *we have* $C_{U_{i_1},\cdots,U_{i_p}|X}(\alpha_{i_1}(x), \cdots, \alpha_{i_p}(x)|x) = m_p(\alpha_{i_1}(x), \cdots, \alpha_{i_p}(x))$ *where*

$$
\begin{aligned}
&m_p(\alpha_{i_1}(x), \cdots, \alpha_{i_p}(x)) \\
&\quad \equiv \mathbb{E}\left[C_{U_{i_1},\cdots,U_{i_p}|X}\left(\alpha_{i_1}(X), \cdots \alpha_{i_p}(X)\middle|X\right)|\alpha_{i_1}(X) = \alpha_{i_1}(x), \cdots, \alpha_{i_p}(X) = \alpha_{i_p}(x)\right]
\end{aligned}
$$

(ii) $m_p(\cdot, \cdots, \cdot)$ *is monotone strictly increasing on* $\mathscr{S}_{\alpha_{i_1}(X),\cdots,\alpha_{i_p}(X)}$ *except at values for which some coordinates are zero.*

(iii) $m_p(\cdot, \cdots, \cdot)$ *is continuously differentiable in the interior of* $\mathscr{S}_{\alpha_{i_1}(X),\cdots,\, \alpha_{i_p}(X)}$.

Proof. See Appendix A.2.5. □

Lemma 5 shows that condition (i) in Lemma 4 does not bind for a structure in $\mathcal{M}_2$ to be observationally equivalent to a structure in $\mathcal{M}_3$. It is due to the fact that the marginal choice probabilities generated by any structure in $\mathcal{M}_2$ can always be matched by I single-agent binary threshold crossing models. Specifically, for any $S \equiv (\pi; F_{U|X}(\cdot|x)) \in \mathcal{M}_2$ with cutoff values of $(u_1^*(x), \cdots, u_I^*(x))$, we can let $\tilde{\pi}_i(a_{-i}, x, u_i) = F_{U_i|X}(u_i^*(x)|x) - u_i$ and $\tilde{F}_{U_i|X}(\cdot|x)$ be the cdf of a uniform distribution on $[0, 1]$. It can be shown that the structure $\tilde{S} \equiv (\tilde{\pi}; \tilde{F}_{U|X})$ generates the same marginal choice probabilities as the given structure S. Thus, Lemma 5 requires only conditions (i)–(iii) on the copula of the structure in $\mathcal{M}_2$. Such conditions arise as $m_p(\cdot, \cdots, \cdot)$ can be viewed as a copula in a model with exogenous payoff shifters, i.e. a model $\mathcal{M}_3$. For instance, conditions (ii)-(iii) follow from the properties of a copula.

We can now characterize all the restrictions imposed on the distribution of observables by model $\mathcal{M}_3$. Because model $\mathcal{M}_2$ does not impose any restriction by Proposition 1 and Proposition 2, these restrictions are due to Assumption E only. [10] Essentially, the following proposition translates conditions (i)–(iii) in Lemma 5 in terms of observables.

[10] It also follows that the restrictions imposed by the set of structures with nonseparable private information and exogenous payoff shifters $\mathcal{M}_3' \equiv \{S \in \mathcal{M}_1 : \text{ Assumption E holds}\}$ is given by Proposition 3.

Proposition 3. *A conditional distribution $F_{Y|X}(\cdot|\cdot)$ rationalized by a structure in $\mathcal{M}_2$ is also rationalized by a structure in $\mathcal{M}_3$ if and only if $\forall p = 2, \cdots, I$ and $\forall\, 1 \leq i_1 < \cdots < i_p \leq I$,*

(R1): $\forall\, x \in \mathscr{S}_X$, *we have*

$$\mathbb{E}\left(\prod_{j=1}^{p} Y_{i_j}\Big|X = x\right) = \mathbb{E}\left(\prod_{j=1}^{p} Y_{i_j}\Big|\alpha_{i_1}(X) = \alpha_{i_1}(x), \cdots, \alpha_{i_p}(X) = \alpha_{i_p}(x)\right).$$

(R2): $\mathbb{E}\left(\prod_{j=1}^{p} Y_{i_j}|\alpha_{i_1}(X) = \cdot, \cdots, \alpha_{i_p}(X) = \cdot\right)$ *is monotone strictly increasing on* $\mathscr{S}_{\alpha_{i_1}(X), \cdots, \alpha_{i_p}(X)}$ *except at values for which some coordinates are zero.*

(R3): $\mathbb{E}\left(\prod_{j=1}^{p} Y_{i_j}|\alpha_{i_1}(X) = \cdot, \cdots, \alpha_{i_p}(X) = \cdot\right)$ *is continuously differentiable in the interior of* $\mathscr{S}_{\alpha_{i_1}(X), \cdots, \alpha_{i_p}(X)}$.

Proof. See Appendix A.2.6. □

Proposition 3 shows that the joint choice probabilities rationalized by $\mathcal{M}_3$ are monotone strictly increasing and continuously differentiable functions of the marginal choice probabilities. Moreover, note that $\alpha_i(x) \equiv F_{U_i|X}(u_i^*(x)|x) = \mathbb{E}(Y_i|X = x)$ by monotone decreasing pure strategy BNE. Therefore, $\alpha_i(\cdot)$ is observable. It follows that the restrictions (R1)-(R3) are testable in principle. This is discussed further in the Conclusion.

For completeness, we study the restrictions on the distribution of observables imposed by $\mathcal{M}_4 \equiv \{S \in \mathcal{M}_3 : \text{ Assumption I holds}\}$. It should be noted that Assumption M is implied by Assumptions S and I, which means that the restriction of monotone decreasing expected payoff is redundant in $\mathcal{M}_4$. Special cases of this model have been considered by several researchers using parametric or functional form restrictions, see, e.g., Bajari, Hong, Krainer, and Nekipelov (2010) and Lewbel and Tang (2012). We first give a necessary and sufficient condition for a structure in $\mathcal{M}_3$ to be observationally equivalent to a structure in $\mathcal{M}_4$.

Lemma 6. *For an arbitrary given structure $S \in \mathcal{M}_3$, there exists an observationally equivalent structure $\widetilde{S} \in \mathcal{M}_4$ if and only if $\forall x \in \mathscr{S}_X$, $\forall p = 2, \cdots, I$, and $\forall\, 1 \leq i_1 < \cdots < i_p \leq I$, we have*

$$C_{U_{i_1}, \cdots, U_{i_p}|X}\left(\alpha_{i_1}(x), \cdots, \alpha_{i_p}(x)\big|x\right) = \prod_{j=1}^{p} \alpha_{i_j}(x). \tag{1.11}$$

Proof. See Appendix A.2.7. □

We note that as in Lemma 5, condition (1.11) involves only the copula but neither the marginal distributions nor the payoffs of the structure S. We also note that condition (1.11) is stronger than conditions (i)-(iii) of Lemma 5 together.

We can now characterize all the restrictions imposed on the distribution of observables by model $\mathcal{M}_4$. Because model $\mathcal{M}_2$ does not impose any restriction by Proposition 1 and Proposition 2, these restrictions are due to Assumptions E and I only. Essentially, the following proposition translates the restrictions in Lemma 5 and Lemma 6 in terms of observables. Since the restrictions in Lemma 5 are weaker than that in Lemma 6, as noted above, only the latter binds.

Proposition 4. *A conditional distribution $F_{Y|X}(\cdot|\cdot)$ rationalized by a structure in $\mathcal{M}_2$ is also rationalized by a structure in $\mathcal{M}_4$ if and only if $Y_1, \cdots, Y_I$ are conditionally independent given X.*

Proof. See Appendix A.2.8. □

As a matter of fact, a proof similar to that of Lemma 6 shows that condition (1.11) is also a necessary and sufficient condition for a structure in $\mathcal{M}_2$ to be observationally equivalent to a structure with separable and mutually independent private information, i.e. to a structure in $\mathcal{M}_4' \equiv \{S \in \mathcal{M}_2 : \text{ Assumption I holds}\}$. Thus it follows that Proposition 4 holds with $\mathcal{M}_4 \equiv \{S \in \mathcal{M}_2 : \text{ both Assumptions E and I hold}\}$ replaced by $\mathcal{M}_4'$. We summarize such discussion in the following corollary.

Corollary 1. *Model $\mathcal{M}_4$ imposes the same restrictions on the distribution of players' choices as $\mathcal{M}_4'$, and both models are observationally equivalent.*

This is a surprising result. Because $\mathcal{M}_4 \subsetneq \mathcal{M}_4'$, exogeneity of the payoff shifters (Assumption E) becomes redundant in terms of restrictions on the observables as soon as mutual independence of private information conditional on X (Assumption I) is imposed. [11]

[11] It also follows that the restrictions on distribution of observables imposed by the set of structures with nonseparable and mutually independent private information $\mathcal{M}_4'' \equiv \{S \in \mathcal{M}_1 : \text{ Assumption I holds}\}$ is given by Proposition 4.

1.4 Nonparametric Identification

In this section we study the nonparametric identification of the baseline game-theoretic model $\mathcal{M}_1$, and its special cases $\mathcal{M}_2$, $\mathcal{M}_3$ and $\mathcal{M}_4$. As a preliminary to the identification of $\mathcal{M}_3$, we also consider the identification of $\mathcal{M}_3' \equiv \{S \in \mathcal{M}_1 : \text{ Assumption E holds}\}$ which is the nonseparable extension of $\mathcal{M}_3$, and which is of interest by itself. The recent literature has focused on the parametric or semiparametric identification of structures in $\mathcal{M}_4$, see, e.g., Brock and Durlauf (2001a); Seim (2006b); Sweeting (2009); Bajari, Hong, Krainer, and Nekipelov (2010), and Tang (2010). As far as we know, Lewbel and Tang (2012) is the only paper that studies the nonparametric identification of a model in $\mathcal{M}_4$ under additional restrictions on the payoffs' functional form.

In our context, the identification of each model is equivalent to the identification of the payoffs π_i, the marginal distribution function $F_{U_i|X}$ and the copula function $C_{U|X}$ of the joint distribution of private information. For all $x \in \mathscr{S}_X$, let $Q_{U_i|X}(\cdot|x) \equiv F^{-1}_{U_i|X}(\cdot|x)$ be the conditional quantile function of the conditional marginal distribution of U_i. Under assumption R, $Q_{U_i|X}$ is well defined. Throughout, the identification of models discussed here refers to the identification of the triple $\{\pi_i, Q_{U_i|X}, C_{U|X}\}$.

In the following analysis, it is shown that without exogeneity (Assumption E) neither $\mathcal{M}_1$ nor $\mathcal{M}_2$ is identified. For $\mathcal{M}_3'$, we show that additional identifying restrictions, namely, some exclusion restrictions and rank conditions can help achieve identification of some features of the underlying structure: Namely, the copula function of the private information distribution, and how the payoffs vary across different rivals' action profiles. Further, we exploit the additive separability restriction in $\mathcal{M}_3$ to establish the identification of both the marginal distribution of private information and the payoffs up to a location–and–scale normalization. Finally, we show that the identification of $\mathcal{M}_4$ requires weaker support restrictions than those for $\mathcal{M}_3$, though the difference in the approaches are not essential.

For the identification of $\mathcal{M}_3'$ and $\mathcal{M}_3$, we require different rank conditions that depend on whether there are strategic interactions in the underlying structure. The lack of strategic effects is essentially a degenerate case of the game–theoretic structural model to I single–agent binary response models, which can be detected by some testable conditions.

1.4.1 Nonidentification of $\mathcal{M}_1$ and $\mathcal{M}_2$

We begin with the most general model $\mathcal{M}_1$ and $\mathcal{M}_2$, in which we show that the structure is typically unidentified.

Proposition 5. *Neither $\mathcal{M}_1$ nor $\mathcal{M}_2$ is identified nonparametrically.*

The non–identification of $\mathcal{M}_2$ follows directly from the observational equivalence between any structure S in $\mathcal{M}_2$ and a collection of I–single-agent binary responses models: Let $\tilde{S} \equiv (\tilde{\pi}; \tilde{F}_{U|X})$ in which $\tilde{\pi}_i(\cdot, x) \equiv \tilde{\pi}_i(x)$ is arbitrarily chosen, and a distribution function $\tilde{F}_{U|X}$ satisfies Assumption R with $\tilde{F}_{U|X}(\tilde{\pi}(x)|x) = F_{U|X}(u^*(x)|x)$ for all $x \in \mathscr{S}_X$. Thus, $\tilde{S}$ and S are observationally equivalent thereby establishing the non–identification of $\mathcal{M}_2$. The non–identification of $\mathcal{M}_1$ follows immediately from Proposition 2.

Next, we turn to the identification of $\mathcal{M}_3'$, $\mathcal{M}_3$ and $\mathcal{M}_4$, for which the identification analysis proceeds in two steps: first, we establish the identification of the copula function C of the types' distribution and the equilibrium belief of the player at the margin, i.e. $\sigma^*_{-i}(\cdot|x, u_i^*(x))$. We then examine the identification of the payoffs π_i and the joint distribution of private information under additional identifying conditions for $\mathcal{M}_3'$, $\mathcal{M}_3$ and $\mathcal{M}_4$, respectively.

1.4.2 Nonparametric identification of C and $\sigma^*_{-i}(a_{-i}|x, u_i^*(x))$

In $\mathcal{M}_3' \supsetneq \mathcal{M}_3 \supsetneq \mathcal{M}_4$, we maintain Assumption E, i.e. that X and U are independent of each other. Because $\alpha_i(x) = F_{U_i|X}(u_i^*(x)|x)$, it follows that $\alpha_i(x) = F_{U_i}(u_i^*(x))$. Let $\alpha(X) \equiv (\alpha_1(X), \cdots, \alpha_I(X))$.

Without imposing additional restrictions, the copula function C can be nonparametrically identified in $\mathcal{M}_3'$ on an appropriate domain which is essentially the support of the F_U–quantile associated with $\alpha(X)$. To see this, let $p = 2, \cdots, I$ and $1 \leq i_1 < \cdots < i_p \leq I$. It follows that for all $(\alpha_{i_1}, \cdots, \alpha_{i_p}) \in \mathscr{S}_{(\alpha_{i_1}(X), \cdots, \alpha_{i_p}(X))}$,

$$\mathbb{E}\left(\prod_{j=1}^{p} Y_{i_j}|\alpha_{i_1}(X)=\alpha_{i_1},\cdots,\alpha_{i_p}(X)=\alpha_{i_p}\right)$$
$$=\mathbb{P}\left\{U_{i_1}\leq F_{i_1}^{-1}(\alpha_{i_1});\cdots;U_{i_p}\leq F_{i_p}^{-1}(\alpha_{i_p})\big|\alpha_{i_1}(X)=\alpha_{i_1},\cdots,\alpha_{i_p}(X)=\alpha_{i_p}\right\}$$
$$=\mathbb{P}\left\{U_{i_1}\leq F_{i_1}^{-1}(\alpha_{i_1});\cdots;U_{i_p}\leq F_{i_p}^{-1}(\alpha_{i_p})\right\}=C_{U_{i_1},\cdots,U_{i_p}}(\alpha_{i_1},\cdots,\alpha_{i_p}). \quad (1.12)$$

Therefore, the copula function C is identified by the players' conditional choice probabilities controlling for the corresponding marginal choice probabilities $\alpha(X)$. Key among those conditions for the nonparametric identification of C is the restriction of a single m.d.p.s. BNE for the data generating process. Under the additional Assumption I, De Paula and Tang (2012) developed a nonparametric testing procedure for such a restriction.

Next, we make the following assumption for the identification of σ_{-i}^*.

Assumption RC (Rank Condition)**.** *The support $\mathscr{S}_{\alpha(X)}$ of $\alpha(X)$ is a convex and compact subset of $[0,1]^I$ with dim$\left(\mathscr{S}_{\alpha(X)}\right)=I$.*

Fix $x\in\mathscr{S}_X$. Because a single equilibrium generates the observed distribution $F_{Y|X}(\cdot|x)$, it follows that $\alpha_i(x)=\mathbb{E}(Y_i|X=x)$ and $\alpha_i(x)$ is identified. Hence, Assumption RC is verifiable. It essentially requires that the payoff shifters X have a rich support. In the next lemma, we establish the identification of $\sigma_{-i}^*(\cdot|x,u_i^*(x))$ under Assumptions E and RC.

Lemma 7. *Let $S\in\mathcal{M}_1$. Suppose that Assumptions E and RC hold. Fix $x\in\mathscr{S}_X$. Then the equilibrium beliefs $\sigma_{-i}^*(\cdot|x,u_i^*(x))$ is identified, i.e. for all $a_{-i}\in\mathcal{A}_{-i}$,*

$$\sigma_{-i}^*(a_{-i}|x,u_i^*(x))=\left.\frac{\partial\mathbb{P}\left(Y_i=1;Y_{-i}=a_{-i}|\alpha(X)=\alpha\right)}{\partial\alpha_i}\right|_{\alpha=\alpha(x)}. \quad (1.13)$$

Proof. See Appendix A.3.1. □

In Lemma 7, the rank condition is needed for taking the derivative in eq. (1.13). In contrast, neither assumption S nor I is required. When Assumption I holds, $\mathbb{P}(Y_i=1;Y_{-i}=a_{-i}|\alpha(X)=\alpha)=\alpha_i\mathbb{P}\left(Y_{-i}=a_{-i}|\alpha(X)=\alpha\right)$ so that the RHS of eq. (1.13) becomes

$\mathbb{P}\left(Y_{-i} = a_{-i} | \alpha(X) = \alpha\right)$, which is consistent with the literature, e.g., Bajari, Hong, Krainer, and Nekipelov (2010); Lewbel and Tang (2012).

We use $\Sigma^*_{-i}(X) \in \mathbb{R}^{2^{I-1}}$ to denote the random vector $\sigma^*_{-i}\left(\cdot | X, u^*_i(X)\right)$. By definition, $\sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma^*_{-i}\left(a_{-i} | x, u^*_i(x)\right) = 1$, then $\Sigma^*_{-i}(X)$ is distributed on a hyperplane in $\mathbb{R}^{2^{I-1}}$. By Lemma 7, we treat $\Sigma^*_{-i}(X)$ as observable hereafter.

1.4.3 Identification of $\mathcal{M}'_3$

Before we proceed, it should be noted that by Lemma 5 the payoffs of a structure in $\mathcal{M}_3$ (or $\mathcal{M}'_3$) are unidentified in general, unless we impose some additional identifying restrictions. We use some "exclusion restrictions". In the context of discrete games, the identification power of exclusion restrictions was first demonstrated in Pesendorfer and Schmidt-Dengler (2003), and later was discussed by Bajari, Hong, Krainer, and Nekipelov (2010) under a general setting with nonparametric payoffs.

Assumption ER (Exclusion Restriction). *Let $X = (X_1, \cdots, X_I)$. For all i, a_{-i}, x and u_i, we have $\pi_i(a_{-i}, x, u_i) = \pi_i(a_{-i}, x_i, u_i)$.* [12]

Under Assumption ER, the payoff function π_i does not vary with x_{-i}. Further, we make the following assumption.

Assumption C. *The payoff function $\pi_i(a_{-i}, x_i, u_i)$ is continuous in u_i for every i, a_{-i} and $x_i \in \mathscr{S}_{X_i}$.*

When π_i is additively separable in u_i, i.e. $\pi_i(a_{-i}, x_i, u_i) = \pi_i(a_{-i}, x_i) - u_i$, assumption C is trivially satisfied. Moreover, for each $(x_i, \alpha_i) \in \mathscr{S}_{X_i, \alpha_i(X)}$, we define a 2^{I-1} by 2^{I-1} matrix $\mathcal{R}_i$ as follows,

$$\mathcal{R}_i(x_i, \alpha_i) \equiv \mathbb{E}\left[\Sigma^*_{-i}(X)\Sigma^*_{-i}(X)^\top \middle| X_i = x_i, \alpha_i(X) = \alpha_i\right].$$

Conditional on $X_i = x_i$ and $\alpha_i(X) = \alpha_i$, suppose there is "no multicollinearity" among the variables in $\Sigma^*_{-i}(X)$, then the matrix $\mathcal{R}_i(x_i, \alpha_i)$ has full rank 2^{I-1}.

[12] As a matter of fact, X_is can have some common variables. To simplify, we assume that X_is partition X.

As we have mentioned above, the largest possible rank of $\mathcal{R}_i(x_i, \alpha_i)$ depends on whether there are strategic interactions in the underlying structure. In the degenerate case, i.e. $\pi_i\left(a_{-i}, x_i, u_i^*(x)\right) = \pi_i\left(a'_{-i}, x_i, u_i^*(x)\right)$ for all $a_{-i} \neq a'_{-i}$ and some x, it is possible to control for $X_i = x_i$ and $\alpha_i(X) = \alpha_i(x)$, while letting $\mathcal{R}_i(x_i, \alpha_i(x))$ have full rank by varying X_{-i}.

Proposition 6. *Let $S \in \mathcal{M}'_3$. Suppose that Assumptions RC, ER and C hold. Fix $x_i \in \mathscr{S}_{X_i}$ and $\alpha_i \in \mathscr{S}_{\alpha_i(X)|X_i=x_i} \cap (0,1)$. If $\mathcal{R}_i(x_i, \alpha_i)$ has full rank 2^{I-1}, then $\mathscr{S}_{\alpha_i(X)|X_i=x_i}$ must be a singleton $\{\alpha_i\}$ and $\pi_i\big(\cdot, x_i, Q_{U_i}(\alpha_i)\big)$ is point identified by $\pi_i\big(a_{-i}, x_i, Q_{U_i}(\alpha_i)\big) = 0$ for all $a_{-i} \in \mathcal{A}_{-i}$. If $\mathcal{R}_i(x_i, \alpha_i)$ has rank $2^{I-1} - 1$, then $\pi_i\big(\cdot, x_i, Q_{U_i}(\alpha_i)\big)$ is identified up to scale. In addition, if there exists $\alpha' \in \mathscr{S}_{\alpha(X)|X_i=x_i} \cap (0,1)$ such that $\alpha_i \neq \alpha'_i$ and $(\alpha_i, \alpha'_{-i}) \in \mathscr{S}_{\alpha(X)}$,[13] then the sign of $\pi_i\big(a_{-i}, x_i, Q_{U_i}(\alpha_i)\big)$ is also identified for all $a_{-i} \in \mathcal{A}_{-i}$.*

Proof. See Appendix A.3.3. □

Proposition 6 shows that fixing x_i, the nonseparable payoffs are identified, or identified up to scale, as well as up to the marginal distributions of players' private information on an appropriate domain, which is essentially the support of the F_{U_i}-quantile associated with $\mathbb{E}(Y_i|X)$ controlling for $X_i = x_i$. If $\mathcal{R}_i(x_i, \alpha_i)$ has full rank 2^{I-1} for some $\alpha_i \in (0,1)$, then the domain has to be a singleton. If $\mathcal{R}_i(x_i, \alpha_i)$ has rank $2^{I-1} - 1$ for all $\alpha_i \in (0,1)$, however, the more variations in $\mathbb{E}(Y_i|X)$ when X_{-i} varies, the larger will be this domain.[14]

Suppose there is no strategic effect for some $\alpha_i \in \mathscr{S}_{\alpha_i(X)|X_i=x_i} \bigcap (0,1)$, i.e. $\pi_i(a_i, x_i, Q_{U_i}(\alpha_i)) = \pi_i(a'_i, x_i, Q_{U_i}(\alpha_i)) = 0$ for all $a_{-i} \neq a'_{-i}$. Intuitively, under Assumption M there should be no variation in $\mathbb{E}(Y_i|X)$ conditional on $X_i = x_i$. In other words, the full rank of $\mathcal{R}_i(x_i, \alpha_i)$ for some $\alpha_i \in \mathscr{S}_{\alpha_i(X)|X_i=x_i} \bigcap (0,1)$ implies that the support $\mathscr{S}_{\alpha_i(X)|X_i=x_i}$ is a singleton, which can be easily verified from the data. On the other hand, the lack of strategic effects in equilibrium can be detected by the full rank conditions on $\mathcal{R}_i(x_i, \alpha_i(x))$. In particular, suppose $I = 2$. Then the full rank condition implies that in

[13] The existence of such an α' can be guaranteed if for some $a_{-i} \in [0,1]^{I-1}$, (α_i, α_{-i}) is an interior point in $\mathscr{S}_{\alpha(X)}$ and (α_i, α_{-i}) is a limit point of $\mathscr{S}_{\alpha(X)|X_i=x_i}$.

[14] This domain excludes the boundaries $Q_{U_i}(0)$ and $Q_{U_i}(1)$ for technical reasons.

equilibrium player i doesn't response to the change of her rival's move. It follows that $\pi_i(0, x_i, Q_{U_i}(\alpha_i)) = \pi_i(1, x_i, Q_{U_i}(\alpha_i))$.

When there exist strategic effects, i.e. $\pi_i(a_{-i}, x_i, Q_{U_i}(\alpha_i)) \neq \pi_i(a'_{-i}, x_i, Q_{U_i}(\alpha_i))$ for some $a_{-i} \neq a'_{-i}$, then $2^{I-1} - 1$ is the largest possible rank that the matrix $\mathcal{R}_i(x_i, \alpha_i)$ can have, which is due to eq. (1.4) — the existence of a non–zero solution for $\pi_i(\cdot, x_i, Q_{U_i}(\alpha_i))$ requires that $\mathcal{R}_i(x_i, \alpha_i)$ should not have full rank. In such a non–degenerate case, Proposition 6 shows that $\pi_i(\cdot, x_i, Q_{U_i}(\alpha_i))/\|\pi_i(\cdot, x_i, Q_{U_i}(\alpha_i))\|$ can be point identified when both $\Sigma^*_{-i}(X)$ and $\alpha(X)$ have sufficient variations through X_{-i}. However, there is no essential restriction on the scale term $\|\pi_i(\cdot, x_i, Q_{U_i}(\alpha_i))\|$, and we are "almost" free to pick a positive function for the scale as long as Assumption M is satisfied. In this sense, the payoffs as function of x_i are not identified.

By the up–to–scale and sign identification of $\pi_i(\cdot, x_i, Q_{U_i}(\alpha_i))$, it follows that that the signs of the strategic effects, i.e. $\pi_i(a_{-i}, x_i, Q_{U_i}(\alpha_i)) - \pi_i(a'_{-i}, x_i, Q_{U_i}(\alpha_i))$ for all $a_{-i} \neq a'_{-i}$, are also identified. De Paula and Tang (2012) developed a different approach for nonparametrically identifying the signs of the strategic effects under a different setup. By exploiting the identification power of multiple equilibria, their approach does not require Assumption E or ER, but relies on the conditions that private information are conditionally independent, multiple equilibria have been played in the data[15], and the payoffs functions have some particular form.

Next, we turn to the identification of Q_{U_i}. Essentially, the marginal distributions of private information are not identified. This is actually expected in view of Matzkin (2003) results for nonseparable models. Indeed, we are "almost" free to choose the quantile functions Q_{U_i} and the only restriction comes from Assumption R, which implies that Q_{U_i} is a strictly increasing and differentiable function.

1.4.4 Identification of $\mathcal{M}_3$

Now we consider model $\mathcal{M}_3$ in which $\pi_i(\cdot, x, u_i) = \pi_i(\cdot, x_i) - u_i$ under Assumption S and the identifying Assumption ER. Because $\mathcal{M}_3$ is a subclass of $\mathcal{M}'_3$, Proposition 6

[15]De Paula and Tang (2012) also established a nonparametric test of the existence of multiple equilibria in the data.

applies directly to $\mathcal{M}_3$. Under the additional Assumption S, however, we can achieve stronger identification results, namely, the identification of the payoffs function $\pi_i(\cdot,\cdot)$ and the marginal quantile function Q_{U_i} up to a location–scale normalization under weaker identifying conditions.

The additive separability of private information has been widely imposed in the discrete game literature, see Bajari, Hong, and Nekipelov (2010); de Paula (2012). Combined with the exclusion restriction, this condition implies that the equilibrium conditions eq. (1.4) can be written as follows,

$$\sum_{a_{-i}\in\mathcal{A}_{-i}} \pi_i(a_{-i}, x_i) \times \sigma^*_{-i}(a_{-i}|x, u^*_i(x)) = u^*_i(x). \tag{1.14}$$

In particular, this expression allows us to treat the unobserved thresholds $u^*_i(x)$ as fixed effects, which can be controlled by the marginal choice probability $\alpha_i(x)$, and to compare payoffs under different values of x_i.

Let $\overline{\Sigma}^*_{-i}(X) \equiv \Sigma^*_{-i}(X) - \mathbb{E}\left[\Sigma^*_{-i}(X)|X_i, \alpha_i(X)\right]$ and $\overline{\mathcal{R}}_i(x_i) = \mathbb{E}\left[\overline{\Sigma}^*_{-i}(X)\overline{\Sigma}^*_{-i}(X)^\top\middle|X_i = x_i\right]$. Because $\sum_{a_{-i}\in\mathcal{A}_{-i}} \sigma^*_{-i}(a_{-i}|x, u^*_i(x)) = 1$, then $\iota'\overline{\Sigma}^*_{-i}(X) = 0$ a.s. where $\iota \equiv (1,\cdots,1)' \in \mathbb{R}^{2^{I-1}}$. Therefore, $\overline{\Sigma}^*_{-i}(X)$ consists of a set of linearly dependent variables and the largest possible rank of the matrix $\overline{\mathcal{R}}_i(x_i)$ is $2^{I-1}-1$. As we will see from the next proposition, the $2^{I-1}-1$ rank occurs only if there is no strategic effects on i given $X_i = x_i$. We first present a proposition similar to Proposition 6, which gives some identification results on players' payoffs for each fixed $x_i \in \mathscr{S}_{X_i}$.

Proposition 7. *Let $S \in \mathcal{M}_3$. Suppose that Assumptions RC and ER hold. Fix $x_i \in \mathscr{S}_{X_i}$ such that $\alpha_i(x) \in (0,1)$ for some $x \in \mathscr{S}_{X|X_i=x_i}$. If $\overline{\mathcal{R}}_i(x_i)$ has the largest possible rank $2^{I-1}-1$, then $\mathscr{S}_{\alpha_i(X)|X_i=x_i}$ must be a singleton $\{\alpha_i(x)\}$ and $\pi_i(\cdot, x_i)$ is identified up to the α_i–quantile of U_i's distribution: $\pi_i(a_{-i}, x_i) = Q_{U_i}(\alpha_i)$ for all $a_{-i} \in \mathcal{A}_{-i}$. If $\overline{\mathcal{R}}_i(x_i)$ has rank $2^{I-1}-2$, then $\pi_i(\cdot, x_i)$ is identified up to location and scale, or equivalently, $\pi_i(\cdot, x_i) - \pi_i(a^0_{-i}, x_i)$ is identified up to scale. In addition, if there exist $\alpha, \alpha' \in \mathscr{S}_{\alpha(X)|X_i=x_i}$ such that $\alpha_i \neq \alpha'_i$ and $(\alpha_i, \alpha'_{-i}) \in \mathscr{S}_{\alpha(X)}$, then the sign of $\pi_i(a_{-i}, x_i) - \pi_i(a^0_{-i}, x_i)$ is also identified for all $a_{-i} \in \mathcal{A}_{-i}\backslash\{a^0_{-i}\}$.*

Proof. See Appendix A.3.4. □

Proposition 7 shows that fixing x_i, the payoffs are identified, or identified up to scale, as well as up to location under two different possible rank conditions of $\overline{\mathcal{R}}_i(x_i)$, respectively.[16] If $\overline{\mathcal{R}}_i(x_i)$ has the largest possible rank $2^{I-1}-1$, then $\mathscr{S}_{\alpha_i(X)|X_i=x_i}$ has to be a singleton $\{\alpha_i\}$. This case could happen only if there is no strategic interactions from player i's rivals' actions on i's payoffs. If $\overline{\mathcal{R}}_i(x_i)$ has rank $2^{I-1}-2$, then holding $X_i = x_i$ and changing X_{-i} could cause variations in $\mathscr{S}_{\alpha_i(X)|X_i=x_i}$ when there are strategic effects, i.e. $\pi_i(a_{-i}, x_i) \neq \pi_i(a'_{-i}, x_i)$ for some $a_{-i} \neq a'_{-i}$. In both cases, note that the rank conditions for identification require that: $\dim\left(\mathscr{S}_{\alpha_i(X)|X_i=x_i}\right) + \text{Rank}\left(\overline{\mathcal{R}}_i(x_i)\right) = 2^{I-1}-1$. Note also that the rank conditions in Proposition 6 are stronger than those in Proposition 7. To see this, fix $(x_i, \alpha_i) \in \mathscr{S}_{X_i,\alpha_i(X)}$. It can be shown that the rank of $\mathcal{R}_i(x_i, \alpha_i)$ equals r, $1 \leq r \leq 2^{I-1}$, if and only if $\overline{\mathcal{R}}_i(x_i)$ has rank $r-1$. The decrease in rank by one comes from the fact that $\iota'\Sigma^*_{-i}(X) = 1$ almost surely.

When there exist strategic effects, i.e. $\pi_i(a_{-i}, x_i) \neq \pi_i(a'_{-i}, x_i)$ for some $a_{-i} \neq a'_{-i}$, then $2^{I-1}-2$ is the largest possible rank of the matrix $\overline{\mathcal{R}}_i(x_i)$. This observation comes from the following two facts: (i) $\iota'\overline{\Sigma}^*_{-i}(X) = 0$ for $\iota \equiv (1, \cdots, 1)' \in \mathbb{R}^{2^{I-1}}$; and (ii) by eq. (1.14), the linear equations

$$c' \times \overline{\Sigma}^*_{-i}(x) = 0, \quad \text{for all } x \in \mathscr{S}_{X|X_i=x_i}$$

have a non-constant solution, which is $c = \pi_i(\cdot, x_i)$. In this case, Proposition 7 shows that $\pi_i(\cdot, x_i) - \pi_i(a^0_{-i}, x_i)$ is identified up to scale and the sign of $\pi_i(\cdot, x_i) - \pi_i(a^0_{-i}, x_i)$ is also identified when $\overline{\Sigma}^*_{-i}(X)$ and $\alpha_i(X)$ have sufficient variations through X_{-i}.

When $\pi_i(\cdot, x_i)$ is identified up to location and scale, Q_{U_i} can also be identified up–to–location–scale on an appropriate domain which is essentially the support of $\mathbb{E}(Y_i|X)$ controlling for $X_i = x_i$. Under Assumptions S and ER, the equilibrium conditions (1.4) can be written as

$$Q_{U_i}(\alpha_i(x)) = \sum_{a_{-i} \in \mathcal{A}_{-i}} \pi_i(a_{-i}, x_i) \times \sigma^*_{-i}(a_{-i}|x, u^*_i(x)),$$

[16] Based on Proposition 7, Lemma 20 in Appendix A.3.6 provides a necessary and sufficient condition for two structures in $\mathcal{M}_3$ to be observationally equivalent.

from which $Q_{U_i}(\alpha_i(x))$ is identified up to the same location and scale normalization for $\pi_i(\cdot, x_i)$. More importantly, Q_{U_i} can further help pin down the location and scale for other values $x_i' \neq x_i$.

The discussion above shows that, for normalization purposes, we are free to choose one location–scale normalization for each player to identify both π_i and Q_{U_i}. One convention is to normalize the scale of payoffs and one quantile of the marginal distribution of U_i. For notational simplicity, however, we employ a slightly different normalization scheme.

Assumption N (Payoff Normalization). *For some given $x_i^* \in \mathscr{S}_X$ satisfying (i) $\overline{\mathcal{R}}_i(x_i^*)$ has rank $2^{I-1}-2$; and (ii) $\mathscr{S}_{\alpha(X)|X_i=x_i}$ contains two different elements α and α' such that $0 < \alpha_i \neq \alpha_i' < 1$ and $(\alpha_i, \alpha_{-i}') \in \mathscr{S}_{\alpha(X)}$, let $\pi_i(a_{-i}^0, x_i^*) = 0$ and $\|\pi_i(\cdot, x_i^*)\| = 1$.*[17]

Let $S \in \mathcal{M}_3$. Suppose that Assumptions RC, ER and N hold. By Proposition 7, $\pi_i(\cdot, x_i^*)$ is point identified. It follows that $Q_{U_i}(\alpha_i)$ is identified on the support $\mathscr{S}_{\alpha_i(X)|X_i=x_i^*} \bigcap (0,1)$ by

$$Q_{U_i}(\alpha_i) = \sum_{a_{-i} \in \mathcal{A}_{-i}} \pi_i(a_{-i}, x_i^*) \mathbb{E}\left[\sigma_{-i}^*(a_{-i}|X, u_i^*(X)) \middle| X_i = x_i^*, \alpha_i(X) = \alpha_i\right].$$

Further, for all $x_i \in \mathscr{S}_{X_i}$ such that $\mathscr{S}_{\alpha_i(X)|X_i=x_i} \bigcap \mathscr{S}_{\alpha_i(X)|X_i=x_i^*} \bigcap (0,1)$ contains at least two elements and the rank of $\overline{\mathcal{R}}_i(x_i)$ equals to $2^{I-1}-2$, $\pi_i(\cdot, x_i)$ is also point identified by Proposition 7. Repeating such an argument, we can show that $\pi_i(\cdot, x_i)$ can be point identified for all x_is in a collection, denoted as $\mathscr{X}_i^\infty$.

Definition 2. *$\mathscr{X}_i^\infty$ is a subset of $\mathscr{S}_{X_i}$ defined by the following iterative scheme. Let $\mathscr{X}_i^0 = \{x_i^*\}$. Then, for all $t \geq 0$, $\mathscr{X}_i^{t+1}$ consists of all elements $x_i \in \mathscr{S}_{X_i}$ such that at least one of the following conditions is satisfied: (i) $x_i \in \mathscr{X}_i^t$; (ii) $\overline{\mathcal{R}}_i(x_i)$ has rank $2^{I-1}-2$ and there exists an $x_i' \in \mathscr{X}_i^t$ such that $\mathscr{S}_{\alpha_i(X)|X_i=x_i} \bigcap \mathscr{S}_{\alpha_i(X)|X_i=x_i'} \bigcap (0,1)$ contains at least two different elements; and (iii) $\overline{\mathcal{R}}_i(x_i)$ has rank $2^{I-1}-1$ and there exists an $x_i' \in \mathscr{X}_i^t$ such that $\mathscr{S}_{\alpha_i(X)|X_i=x_i} \subsetneq \mathscr{S}_{\alpha_i(X)|X_i=x_i'} \bigcap (0,1)$.*

[17] Because $\mathscr{S}_{\alpha(X)|X_i=x_i^*}$ is not a singleton, then there exist strategic effects in player i's payoffs, i.e. $\|\pi_i(\cdot, x_i^*) - \pi_i(a_{-i}^0, x_i^*)\| \neq 0$.

In Definition 2, condition (ii) is the key to effectively expand $\mathscr{X}_i^\infty$ in an iterative manner by enlarging $\cup_{x_i\in\mathscr{X}_i^t}\mathscr{S}_{\alpha_i(X)|X_i=x_i}$ to $\cup_{x_i\in\mathscr{X}_i^{t+1}}\mathscr{S}_{\alpha_i(X)|X_i=x_i}$.

Proposition 8. *Let $S\in\mathcal{M}_3$. Suppose that assumptions RC, ER and N hold. Then $\pi_i(a_{-i},x_i)$ is point identified for all $a_{-i}\in\mathcal{A}_{-i}$ and $x_i\in\mathscr{X}_i^\infty$.*

Proof. See Appendix A.3.5 □

Proposition 8 shows that the payoffs and marginal distributions of private information are point identified under our additional location–and–scale normalization on an appropriate domain which essentially depends on how $\mathscr{S}_{\alpha_i(X)|X_i=x_i}$ varies across x_i. Regarding the choice of the starting point x_i^*, intuitively, we should choose x_i^* in a way such that $\mathscr{X}_i^\infty$ is the largest. For instance, suppose that $\cup_{x_i\in\mathscr{X}_i^\infty}\mathscr{S}_{\alpha_i(X)|X_i=x_i}=\mathscr{S}_{\alpha_i(X)}$ and either one of the two rank conditions is satisfied, then $Q_{U_i}(\cdot)$ and $\pi_i(\cdot,\cdot)$ are point identified on the support $\mathscr{S}_{\alpha_i(X)}$ and $\mathcal{A}_{-i}\times\mathscr{S}_{X_i}$, respectively. Note that our normalization does not apply to a nonparametric single–agent binary response model (see, e.g., Matzkin, 1992). Due to lack of strategic effects, the support of the marginal choices probability $\mathscr{S}_{\alpha_i(X)|X_i=x_i^*}$ is a singleton in a binary response model. In an interaction model, in contrast, we can use the variation of X_{-i} controlling for X_i to identify a set of quantiles of the latent variable's distribution.

Alternatively, we could identify the payoffs in model $\mathcal{M}_3$ using the single–index structure suggested in Lemma 11: Let $h_i(a_{-i},x_i)=\pi_i(a_{-i},x_i)-\pi(a_{-i}^0,x_i)$. Then

$$\mathbb{E}\left(Y_i|X_i=x_i,\Sigma_{-i}^*(X)=\Sigma_{-i}^*(x)\right)\\ =F_{U_i}\left(\pi_i\left(a_{-i}^0,x_i\right)+\sum_{a_{-i}\in\mathcal{A}/\{a_{-i}^0\}}h_i(a_{-i},x_i)\sigma_{-i}^*\left(a_{-i}|x,u_i^*(x)\right)\right) \quad (1.15)$$

in which $\Sigma_{-i}^*(x)$ is identified by Lemma 7. Similarly to Powell, Stock, and Stoker (1989), we could identify $h_i(\cdot,x_i)$ up to scale by differentiating eq. (1.15) with respect to $\sigma_{-i}^*(\cdot|x,u_i^*(x))$. Thus, the payoff functions $\pi_i(\cdot,x_i)$ are identified up to location and scale. This strategy, however, involves an additional support condition on $\mathscr{S}_{\Sigma_{-i}^*(X)|X_i=x_i}$.

1.4.5 Identification of $\mathcal{M}_4$

The identification of $\mathcal{M}_4$ does not essentially differ from that of $\mathcal{M}_3$: Assumption I only relaxes the rank condition for identification of σ^*_{-i} in Lemma 7.

Lemma 8. *Let $S \in \mathcal{M}_4$. Fix $x \in \mathscr{S}_X$. Then $\sigma^*_{-i}(\cdot|x, u^*_i(x))$ is identified by*

$$\sigma^*_{-i}(a_{-i}|x, u^*_i(x)) = \mathbb{P}\left(Y_{-i} = a_{-i}|X = x\right). \tag{1.16}$$

The proof is straightforward, hence omitted. As noted after Lemma 7, we also have $\sigma^*_{-i}(a_{-i}|x, u^*_i(x)) = \mathbb{P}\left(Y_{-i} = a_{-i}|\alpha(X) = \alpha(x)\right)$ under Assumption I. Similar results can be found in Aguirregabiria and Mira (2007); Bajari, Hong, Krainer, and Nekipelov (2010), among others. Further, the identification of π_i and Q_{U_i} in $\mathcal{M}_4$ follows Propositions 7 and 8.

1.5 Partial Identification

In this section we examine the partial identification of structures in $\mathcal{M}'_3$ and $\mathcal{M}_3$. Partial identification is closely related to rationalization and identification studied in Sections 1.3 and 1.4. Given a probability distribution of observables, i.e. $F_{Y|X}$, we are interested in which structure or which set of structures in model $\mathcal{M}$ can generate the observed distribution. For each $S \in \mathcal{M}$, let $\phi : \mathcal{M} \to \mathscr{P}$ be a correspondence mapping a structure to a set of conditional distributions of observables that could be generated from the given structure, and let ϕ^{-1} be the inverse correspondence defined as

$$\phi^{-1}(F_{Y|X}) = \{S \in \mathcal{M} : F_{Y|X} \in \phi(S)\}.$$

Regarding the set $\phi^{-1}(F_{Y|X})$, there could be two distinct possibilities

Case I: if $\phi^{-1}\left(F_{Y|X}\right)$ is an empty set, then we say that $F_{Y|X}$ is **unrationalizable** by $\mathcal{M}$.

Case II: if $\phi^{-1}\left(F_{Y|X}\right)$ is a non–empty set, then we say $F_{Y|X}$ is **rationalized** by $\mathcal{M}$. Further, if $\phi^{-1}(F_{Y|X})$ is a singleton, we say $\mathcal{M}$ is **point identified**, or **identified** given $F_{Y|X}$; otherwise we say $\mathcal{M}$ is **partially identified** given $F_{Y|X}$.

Our rationalization results in Section 1.3 characterize wether the set $\phi^{-1}\left(\mathbb{P}_{Y|X}\right)$ is empty or not. For instance, given the characterization of the set of rationalizable distributions in Section 1.3, it follows that those probability distributions which violate the conditions in Propositions 1 and 3 have an empty identified set. In other words, those distributions are not generated by any structure in $\mathcal{M}_3$ or $\mathcal{M}_3'$, respectively. On the other hand, our Section 1.4 studies point identification under some identifying assumptions such as Assumption ER. In this section, we relax the latter assumption and completely characterize the (non–empty) set of structures in $\mathcal{M}_3$ or $\mathcal{M}_3'$ which could generate the observed distribution.

Before proceeding, it is worth emphasizing that partial identification of a structure is not due in our case to multiple equilibria, but to the lack of identifying restrictions. This is similar to, e.g., Shaikh and Vytlacil (2011) who study partial identification of the average structural function in a triangular model without imposing a restrictive support condition. Moreover, when there are multiple monotone pure strategy BNEs, we still maintain the assumption of a single equilibrium being played for generating the distribution of observables. Such an assumption is realistic if the equilibrium selection rule is actually governed by some game invariant factors, like culture, social norm, etc. See, e.g., de Paula (2012) for a detailed discussion.

1.5.1 Partial Identification of $\mathcal{M}_3$ and $\mathcal{M}_3'$

We first characterize the set of structures in $\mathcal{M}_3$ that could generate a given rationalized distribution for the observables. We begin with the identification region for the copula function C. Given the point–identification of C on the support $\mathscr{S}_{\alpha(X)}$ discussed in subsection 4.2, by Billingsley (2012, Theorem 14.1) we can obtain a set of quantile functions by extending the support of the identified part to the whole domain $[0,1]^I$. Let $\mathscr{C}$ be the set of strictly increasing (on the interior of its support) and differentiable copula functions mapping $[0,1]^I$ to $[0,1]$, and

$$\mathscr{C}_I = \Big\{ \tilde{C} \in \mathscr{C} : \tilde{C}_{U_{i_1},\cdots,U_{i_p}}(\alpha_{i_1},\cdots,\alpha_{i_p}) = C_{U_{i_1},\cdots,U_{i_p}}(\alpha_{i_1},\cdots,\alpha_{i_p})$$
$$\text{for all } p = 2,\cdots,I,\ 1 \le i_1 < \cdots < i_p \le I \text{ and } \alpha_{i_1},\cdots,\alpha_{i_p} \in \mathscr{S}_{\alpha_{i_1}(X),\cdots,\alpha_{i_p}(X)} \Big\}.$$

Next, we turn to the partial identification of the marginal distribution of types. Under Assumptions R and E, $Q_{U_i|X} = Q_{U_i}$ and Q_{U_i} belongs to the set of strictly increasing and differentiable functions from $[0,1]$ to $\mathbb{R}$, denoted as $\mathscr{Q}_i$. Let $\mathscr{Q} = \times_{i=1}^{I} \mathscr{Q}_i$. The next lemma says that $\mathcal{M}_3$ imposes no restrictions on Q_{U_i} and its identification region is $\mathscr{Q}_i$.

Lemma 9. *Let $S \in \mathcal{M}_3$ (or $\mathcal{M}'_3$). For any $\left(\tilde{Q}_{U_1}, \cdots, \tilde{Q}_{U_I}\right) \in \mathscr{Q}$, then there exists an observationally equivalent structure $\tilde{S} \in \mathcal{M}_3$ (therefore, $\tilde{S} \in \mathcal{M}'_3$) with marginal quantile functions $\left(\tilde{Q}_{U_1}, \cdots, \tilde{Q}_{U_I}\right)$.*

Proof. See Appendix A.4.1 □

We can now focus on the sharp identification region for the payoff function $\pi_i(\cdot, x)$ for all $x \in \mathscr{S}_X$. Let $\mathscr{G}_i$ be the set of real functions mapping $\mathcal{A}_{-i} \times \mathscr{S}_X$ to $\mathbb{R}$. Let $\mathscr{G} = \times_{i=1}^{I} \mathscr{G}_i$, $\pi = (\pi_1, \cdots, \pi_I) \in \mathscr{G}$ and $Q_U = (Q_{U_1}, \cdots, Q_{U_I}) \in \mathscr{Q}$.

Proposition 9. *Let $S \in \mathcal{M}_3$. Suppose that Assumption RC holds. Then the sharp identification region for (π, Q_U, C) is given by $\left\{(\tilde{\pi}, \tilde{Q}_U, \tilde{C}) : \tilde{Q}_U \in \mathscr{Q},\ (\tilde{\pi}, \tilde{C}) \in \Theta_I(\tilde{Q}_U)\right\}$, in which*

$$\begin{aligned}\Theta_I(\tilde{Q}_U) \equiv \Big\{&(\tilde{\pi}, \tilde{C}) \in \mathscr{G} \times \mathscr{C}_I : (a) \text{ for all } x \in \mathscr{S}_X \text{ and } i,\\ &\tilde{Q}_{U_i}(\alpha_i(x)) = \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\pi}_i(a_{-i}, x) \times \sigma^*_{-i}(a_{-i}|x, u^*_i(x)); (b) \text{ for any m.d.p.s. profile } \delta,\\ &\quad \tilde{\mathbb{E}}_\delta\left[\tilde{\pi}_i\left(Y_{-i}, X\right)|X = x, U_i = \tilde{Q}_{U_i}(\alpha_i)\right] - \tilde{Q}_{U_i}(\alpha_i) \text{ is decreasing in } \alpha_i \in (0,1)\Big\}.\end{aligned}$$

Proof. See Appendix A.4.2. □

In the definition of $\Theta_I(\tilde{Q}_U)$, restriction (a) only requires that $\pi_i(\cdot, x)$ belong to a hyperplane, for which the slopes are given by $\sigma^*_{-i}(\cdot|x, u^*_i(x))$. By Lemma 7, the latter is identified. Condition (b) is weak as it does not impose much restriction on the structural parameters. In particular, it always holds if $\tilde{\pi}_i(a_{-i}, x)$ equals $\tilde{Q}_{U_i}(\alpha_i(x))$ for all a_{-i}. Moreover, given $\tilde{Q}_U \in \mathscr{Q}$, it can be shown that $(\tilde{\pi}, C) \in \Theta_I(\tilde{Q}_U)$ for any linear transformation $\tilde{\pi}$ of π, i.e. $\tilde{\pi}_i(\cdot, x) = a_i(x) + b_i(x) \times \pi_i(\cdot, x)$, where b_i is an arbitrary non-negative function and $a_i(x) = \tilde{Q}_{U_i}(\alpha_i(x)) - b_i(x) \times Q_{U_i}(\alpha_i(x))$. Hence, $\Theta_I(\tilde{Q}_U)$ is not empty.[18]

[18]It should also be noted that $\Theta_I(\tilde{Q}_U)$ is a convex set.

The set $\Theta_I(\tilde{Q}_U)$ is unbounded and quite large. To narrow down the identification region, additional restrictions need to be introduced. An alternative approach to imposing our exclusion restrcitions (Assumption ER) is to make assumptions on the functional form of payoffs. For instance, De Paula and Tang (2012) specify the payoffs as $\pi_i(a_{-i}, x) = \pi_i^*(x) + g_i(a_{-i}) \times h_i^*(x)$, where g_i is a function known to all players as well as to the econometrician, and $(\pi_i^*, h_i^*(x))$ are the new structural parameters.

When $\mathscr{S}_{\alpha(X)} = [0,1]^I$, $\mathscr{C}_I$ degenerates to the singleton $\{C\}$. In this case, the identification region of structures in $\mathcal{M}_3$ is defined by a set of payoffs π and marginal quantile functions Q_U. Specifically, the sharp identification region can be expressed in a more straightforward manner as

$$\Theta_I^* = \Big\{(\tilde{\pi}, \tilde{Q}_U) \in \mathscr{G} \times \mathscr{Q} : (a') \text{ for all } x \in \mathscr{S}_X \text{ and } i,$$
$$\tilde{Q}_{U_i}(\alpha_i(x)) = \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\pi}_i(a_{-i}, x) \times \sigma_{-i}^*(a_{-i}|x, u_i^*(x)); (b') \text{ and for all } \alpha_{-i} \in [0,1]^{I-1},$$
$$\sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\pi}_i(a_{-i}, x) \times \sigma_{-i}^C(a_{-i}|\alpha_{-i}, \alpha_i) - \tilde{Q}_{U_i}(\alpha_i) \text{ is decreasing in } \alpha_i \in (0,1)\Big\},$$

where $\sigma_{-i}^C(a_{-i}|\alpha_{-i}, \alpha_i) \equiv \mathbb{P}_C(\forall a_j = 1, V_j \leq \alpha_j, \ \forall a_j = 0, V_j > \alpha_j | V_i = \alpha_i)$, in which V conforms to a distribution with the copula function C as its cdf.

Similarly, we can characterize the sharp identification region for the structural parameters in $\mathcal{M}_3'$. Let $\mathscr{G}_i'$ be the set of functions mapping $\mathcal{A}_{-i} \times \mathscr{S}_X \times \mathbb{R}$ to $\mathbb{R}$ and $\mathscr{G}' = \bigtimes_{i=1}^I \mathscr{G}_i'$.

Proposition 10. *Let $S \in \mathcal{M}_3'$. Suppose that Assumption RC holds. Then the sharp identification region for (π, Q_U, C) is given by $\{(\tilde{\pi}, \tilde{Q}_U, \tilde{C}) : \tilde{Q}_U \in \mathscr{Q}, \ (\tilde{\pi}, \tilde{C}) \in \Theta_I'(\tilde{Q}_U)\}$, in which*

$$\Theta_I'(\tilde{Q}_U) = \Big\{(\tilde{\pi}, \tilde{C}) \in \mathscr{G}' \times \mathscr{C}_I : \textit{ for all } x \in \mathscr{S}_X \textit{ and } i,$$
$$0 = \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\pi}_i\left(a_{-i}, x, \tilde{Q}_{U_i}(\alpha_i(x))\right) \times \sigma_{-i}^*(a_{-i}|x, u_i^*(x)); \textit{ and for any m.d.p.s. profile } \delta$$
$$\tilde{\mathbb{E}}_\delta\left[\tilde{\pi}_i\left(Y_{-i}, X, \tilde{Q}_{U_i}(\alpha_i)\right) | X = x, U_i = \tilde{Q}_{U_i}(\alpha_i)\right] \textit{ is decreasing in } \alpha_i \in (0,1)\Big\}.$$

The argument is similar to that for Proposition 9, and therefore omitted.

1.5.2 Applying Chesher and Rosen (2011, CR) to incomplete information games

In the context of simultaneous–equation models with discrete outcome variables, CR proposed a general method for the set identification under four possible combinations of coherence and completeness. First, we should note that the sharp identification region discussed in Propositions 9 and 10 involves an additional restriction beyond the range of CR's paper: a single equilibrium played for the data generating process. To discuss the identification region provided by CR's method in our setting, here we look at the following model $\mathcal{M}_3''$,

$$\mathcal{M}_3'' \equiv \left\{S : \text{Assumptions R, M, S, and E hold, and equilibrium concept is m.d.p.s. BNE}\right\}.$$

Following CR's criteria, $\mathcal{M}_3''$ is coherent but not complete.

Moreover, our structural relationship can be written as $h(Y, X, U) = 0$, where

$$h = \sum_{i=1}^{I} \Big| Y_i - \mathbf{1}\left\{U_i \leq u_i^*(X)\right\} \Big| = 0,$$

in which $u^*(x) = (u_1^*(x), \cdots, u_I^*(x))$ solves eq. (1.2).

In our setting, CR's approach does not provide a sharp identification region, since there are additional restrictions due to the BNE solution concept. To see this, it is helpful to introduce a general equilibrium selection rule which complements the specification of the structural relation and yields an econometrics model that is complete. It should be noted that the equilibrium selection rule itself does not impose any additional restrictions to our model and it only describes which equilibrium get played in the data generating process.

Fix $X = x$ and the structural parameter $\theta \equiv \{\pi_i, Q_{U_i}, C\}$. Let $\Delta(x, \theta)$ be the set of equilibria and $s^{[k]}$ be the k–th m.d.p.s. BNE,

$$\Delta(x, \theta) = \left\{s^{[1]}(x), \cdots, s^{[K(x,\theta)]}(x)\right\}$$

where $K(x, \theta)$ is the number of equilibria and $s^{[k]}(x)$ is the k–th equilibrium strategy profile — a set of type–contingent plans. Further, let $\mathscr{E}$ is the space of sets of equilibria. Now we define the equilibrium selection rule as a random function $\psi : \mathscr{E} \times \Omega \to \mathbb{N}_+$. Given each w, it maps a set of equilibria to a choice of equilibrium. A proposer equilibrium selection rule should satisfy $1 \leq \psi(\Delta, w) \leq \#\Delta$, almost surely. Thus, players' equilibrium responses can be written as

$$Y_i = s_i^{[Z]}(X, U_i), \quad \text{for all } i,$$
$$Z = \phi(\Delta(X, \theta), w).$$

By definition, $\Delta(x, \theta)$ is a (nonrandom) collection of equilibria, but Z is a random variable distributed on support $\{1, \cdots, K(x, \theta)\}$ given $X = x$.

The BNE solution implies that Z has to be conditionally independent with U given X, which implies that the fact that a particular equilibrium selected to play does not provide any additional information to players on the joint distribution of types. By definition, the prior distribution $F_{U|X}$ in the setting are "common knowledge", which does not change according to which equilibrium gets played, i.e.,

$$F_{U|X,Z} = F_{U|X}. \tag{1.17}$$

Equation (1.17) is the additional restriction to CR's setting. We now characterize the set of the distribution of Y could be generated from the structure with and without this restriction which is a dual problem of the characterization of the identification region of θ.

Let $P^{[k]}(x; \theta)$ be the corresponding probability distribution of Y given $X = x$ and $Z = k$, which depends on the equilibrium strategy profile $s^{[k]}(x)$ and $F_{U|X=x}$. When eq. (1.17) is taken into account, any rationalizable distribution of Y given $X = x$ can be written as

$$P^*(x; \theta) = \sum_{k=1}^{K(x,\theta)} \lambda_k(x) P^{[k]}(x; \theta)$$

in which $\lambda_k(x) \geq 0$ is the probability of equilibrium k chosen by equilibrium selection rule given $X = x$. By definition, $\sum_{k=1}^{K(x,\theta)} \lambda_k(x) = 1$.

Now we turn to the rationalizable distributions without restriction eq. (1.17). Suppose Z depends on X and U in an arbitrary way. Let $G_{Z|X,U}$ be the conditional cdf of Z given (X,U). Thus,

$$F_{U|X,Z} = \frac{G_{Z|X,U} \times F_{U|X}}{\lambda_Z(X)}.$$

The distribution of Y given X could be

$$P^{CR}(x;\theta) = \sum_{k=1}^{K(x,\theta)} \lambda_k(x) Q^{[k]}(x;\theta)$$

where $Q^{[k]}(x,\theta)$ is the corresponding probability distribution of Y given $X = x$ and $Z = k$, which depends on the equilibrium strategy profile $s^{[k]}(x)$ and $F_{U|X=x,Z=k}$. Because $Q^{[k]}$ is in a larger space of distributions, which contains $P^{[k]}$, it is straightforward that the set of distributions $P^{CR}(x;\theta)$ generated without the additional independence restriction is much larger than the set of distributions $P^*(x;\theta)$ under restriction eq. (1.17). Accordingly, the identification region will not be sharp if we don't take eq. (1.17) into account in set identification.

1.6 Conclusion

This paper addresses the rationalization and identification of discrete games with correlated types in a fully nonparametric way. We show that our baseline game-theoretical model does not impose any essential restriction on observables. This implies that binary Bayesian games are not testable in view of players' choices only. We also characterize all the restrictions on players' choices imposed by three assumptions frequently made in the empirical analysis of discrete games. We then exploit exclusion restrictions to identify our structural model nonparametrically in both nonseparable and separable cases. These restrictions take the form of excluding part of a player's payoff shifters from all other players' payoffs as frequently assumed in the empirical discrete game literature. We also characterize the sharp identification region of the structural parameters without the exclusion restrictions.

We require that the same m.d.p.s. BNE is played in the DGP for a given x. The nonparametric analysis relaxing such a requirement clearly needs to be developed. In particular, all of our rationalization and identification results will be weakened in that situation. In particular, $\mathcal{M}_1$ and $\mathcal{M}_2$ still impose no essential restrictions on observables whether or not a single monotone pure strategy BNE is played. On the other hand, $\mathcal{M}_3$ and $\mathcal{M}_4$ will impose weaker restrictions than those in Proposition 3 and Proposition 4, respectively. Location/scale identification of these two models, as established in Propositions 6 and 7, will be lost. Thus, point identification of the primitives would require additional identifying assumptions. Alternatively, one could follow the set–identification approach initiated by Tamer (2003) in the presence of multiple equilibria. A recent review on set identification in discrete games can be found in de Paula (2012).

A second line of research, which needs to be developed, concerns model testing. Our Proposition 3 and Proposition 4 become especially useful as they characterize all the restrictions in terms of observables imposed by $\mathcal{M}_3$ and $\mathcal{M}_4$. Thus such restrictions are in principle testable for the purpose of model specifications. In particular, some tests can be relatively easy to develop as they only involve some nonparametric regressions. For instance, the restriction given in Proposition 4 can be tested by using conditional independence tests developed in statistics and econometrics (see, e.g. Su and White, 2007, 2008). It is also worthnoting that such tests do not rely on identification of the model and consequently on the assumptions used to identify the primitives.

Lastly, a third line of research deals with the nonparametric estimation of the various models. In a semiparametric setup, Liu and Xu (2012) propose an estimation procedure for our model $\mathcal{M}_3$ with linear payoff, and establish the root-N consistency of the linear payoff coefficients. A fully nonparametric estimation, however, deserves future investigation. A strategy could rely on the identification results and propose a sample-analog type of estimators for the players' payoffs and the joint distribution of private information. Establishing the asymptotic properties of such an estimation procedure is left for future research. The main difficulty relies on the generated covariates, namely the belief of the player at the margin which appears in the expected payoff. Such a problem could be addressed by using the most recent literature on nonparametric regression with nonparametrically generated covariates (see, e.g. Mammen, Rothe, and Schienle, 2012).

Chapter 2

Semiparametric analysis of Binary Games of Incomplete Information

2.1 Introduction

In this paper, we study the semiparametric identification and estimation of static binary games of incomplete information with correlated private types. The range of applications of binary games includes, among others, models of entry (Bresnahan and Reiss, 1990, 1991; Berry, 1992; Jia, 2008; Seim, 2006a), couples' retirement decisions (Banks, Blundell, and Casanova Rivas, 2010; Casanova, 2010), labor force participation (Bjorn and Vuong, 1984; Soetevent and Kooreman, 2007)), stock market analysts' recommendations (Bajari, Hong, Krainer, and Nekipelov, 2010), advertising (Sweeting, 2009), and social interactions (Brock and Durlauf, 2001a,b; Xu, 2011).

To simply our exposition, we formally consider throughout this paper the equilibrium solution that can be represented by the following structural equations (i.e., best responses): for $i = 1, \cdots, I$,

$$Y_i = \mathbf{1}\Big\{ X_i'\beta_i + \sum_{j \neq i} \alpha_{ij} \mathbb{P}(Y_j = 1 | X, U_i) - U_i \geq 0 \Big\}, \tag{2.1}$$

where subscript i is an index of players in the game; X_i is a vector of exogenous payoff relevant variables, while the error term U_i is i's private information, which is not observed by other players; We allow $U = (U_1, \cdots, U_I)$ to be correlated with each other under an unknown form. This model is a natural extension of Manski (1975, 1985)'s binary threshold crossing model in the single–agent setup to a structural model with strategic interactions.

This paper contributes to the existing discrete game literature in several aspects. First, we do not require the (conditional) independence of private payoff shocks across players, which is widely adopted by most of the literature, e.g., Aguirregabiria and Mira (2007); Bajari, Hong, Krainer, and Nekipelov (2010); De Paula and Tang (2010); Grieco (2011); Pesendorfer and Schmidt-Dengler (2003) and Lewbel and Tang (2011); exceptions include Aradillas-Lopez (2010); Wan and Xu (2010) and Xu (2010).[1]

Allowing correlated private signals is motivated primarily by empirical concerns. The (conditional) independence assumption of U is convenient but imposes strong restrictions — players' choices must be conditionally independent, which could be invalidated by the data.[2] Moreover, in the social interaction framework, the correlation among players' private payoff shock represents the "homophily" effects in social behaviors, which is caused by the unobserved "similarity" in players' preference. In contrast, the peer effects are purely the strategic effects caused by interactions with other group members. Both effects accounting for the "herding" behavior in a society group can be identified and distinguished with each other in our model.

Second, we make no parametric assumptions on the joint distribution of private payoff shocks, which distinguishes our paper from Xu (2010). We establish nonparametric identification results for the copula function of private payoff shocks, from which we can derive equilibrium belief function. In a similar semiparametric setup, Wan and Xu (2010) establish partial identification of payoff coefficients when types are positively regression dependent, and further achieve point identification under an additional support condition on regressors. The maximum score type estimator they suggested converges at $\sqrt[3]{n}$–rate. In this paper, we establish point identification of structural parameters under

[1]The novel approach developed in Aradillas-Lopez (2010) assumes that players do not have exact knowledge about the distributions involved and then using an equilibrium concept defined in Aumann (1987).

[2]A model featured with unobserved heterogeneity and independent private signals also generates dependence among players' choices conditional on observed regressors (see Grieco, 2011).

weak conditions. In addition, Our Klein–Spady type estimator in this paper is shown to be $\sqrt{n}$–consistent.

The key in our semiparametric identification approach is to focus on the class of monotone pure strategy BNEs. Athey (2001) provided the seminar result that a monotone pure–strategy BNE exists whenever a Bayesian game obeys a Spence–Mirlees single–crossing restriction. McAdams (2003) and Reny (2011) extended Athey (2001)'s results. Applying Reny (2011) in our setup, we show that a monotone strategy BNE generally exists under mild conditions.

Third, we propose a Klein–Spady type pseudo maximum likelihood estimator for the structural parameter, which is shown to be $\sqrt{n}$–consistent. In the proposed estimation procedure, we estimate the belief component nonparmaetrically. Then, following Klein and Spady (1993), we construct a pseudo loglikelihood function using the estimated beliefs as part of covariates. Monte–Carlo evidence indicates that there is only modest efficiency losses relative to the semiparametric estimation when the belief component is known to researchers.

The rest of the paper is arranged as follows. We introduce the setup of our game model in Section 2 and establish the existence of monotone pure strategy BNE in Section 3. In Section 4, We discuss the semiparametric identification of the structural model. We then propose a Klein–Spady type estimator in a two–player setup in Section 5. Section 6 provides Monte–Carlo simulations.

2.2 Model

We consider a static binary game of incomplete information, commonly referred to as a Discrete Bayesian game. There are a finite number of players, indexed by $i \in \mathcal{I} \equiv \{1, 2, \cdots, I\}$, and each player i simultaneously chooses an action $Y_i \in \{0, 1\}$.[3] Define $\mathcal{A} = \{0, 1\}^I$ as the action space of the game and let $y = (y_1, \cdots, y_I) \in \mathcal{A}$ be a generic element of $\mathcal{A}$. Following the convention, let $\mathcal{A}_{-i}$ and y_{-i} denote the action space and a profile of actions for all players but excluding player i, respectively.

[3]For notational simplicity, we restrict players to make binary decisions and all of our results could be generalized to the case where the choice set for each player is finite.

For each player i, $X_i \in \mathbb{R}^{d_i}$ is a vector of payoff relevant random variables, which are publicly observed by all players. Define $X = (X_1, \cdots, X_I) \in \mathbb{R}^p$, where $p = \sum_{i=1}^I d_i$, as all the publicly observed information in the game. Player i's payoff shock U_i is i's private information, which is not observed by other players. Let $U = (U_1, \cdots, U_I)$ and F_{XU} be the cumulative distribution function (c.d.f.) of (X, U). The joint distribution F_{XU} is assumed to be common knowledge among all players.

The payoff for player i is described as follows,

$$\pi_i(y, x_i, u_i) = \begin{cases} x_i'\beta_i + \sum_{j\neq i} \alpha_{ij} y_j - u_i, & \text{if } y_i = 1, \\ 0, & \text{if } y_i = 0, \end{cases}$$

where $\beta_i \in \mathbb{R}^{d_i}$ and $\alpha_{ij} \in \mathbb{R}$ ($j \neq i$) are the parameters of interest. α_{ij} ($j \neq i$) are strategic interaction parameters, which measure the *ceteris paribus* effects on i's payoff from j's choice. Our payoff function here is similar to the parametric case in Bajari, Hong, Krainer, and Nekipelov (2010).[4] The zero payoff for action $y_i = 0$ is a standard normalization.

Regarding the payoff shock U, our analysis involves neither (conditional) independence restrictions between U_i and U_j nor parametric assumptions in its joint distribution. This is different from most papers in the static discrete game literature (e.g., Bajari, Hong, Krainer, and Nekipelov, 2010); only exceptions include Aradillas-Lopez (2010), Liu, Vuong, and Xu (2012), and Wan and Xu (2010).

Following the literature on Bayesian games, player i's decision rule is a function $Y_i = s_i(X, U_i)$, where $s_i : \mathscr{S}_X \times \mathbb{R} \to \{0, 1\} \in \Delta_i$ maps all the information that player i knows to a binary response and Δ_i is her strategy space. Note that X_{-i} also enters player i's decision rule s_i, since the opponents' decisions have effects on i's response through the strategic interactions.

Fix $x \in \mathscr{S}_X$. For any strategy profile $s = (s_1, \cdots, s_I) \in \times_{i=1}^I \Delta_i$ and $j \neq i$, we let $\sigma_{ij}^s(x, u_i)$ be the conditional probability $\mathbb{P}\left\{s_j(X, U_j) = 1 | X = x, U_i = u_i\right\}$, i.e.,

$$\sigma_{ij}^s(x, u_i) = \int_{\mathbb{R}} \mathbf{1}\left\{s_j(x, v) = 1\right\} f_{U_j|X,U_i}(v|x, u_i) dv$$

[4] Aradillas-Lopez (2010), Lewbel and Tang (2011), and Wan and Xu (2010), among others, have also studied binary games with the same payoff structure but under a two–player framework.

where $\mathbf{1}[\cdot]$ is the indicator function and $f_{U_j|X,U_i}$ is the conditional probability density function of U_j given X and U_i. Hence, the term $\sigma_{ij}^s(x, u_i)$ is player i's belief on the event $Y_j = 1$, given i's information (x, u_i) and the specified decision rule s.

The equilibrium concept we adopt is the pure strategy Bayesian Nash equilibrium (BNE). Similar to Bajari, Hong, Krainer, and Nekipelov (2010), the mixed strategy equilibrium is not considered hereafter, since with probability one, each player has a unique best response when the private information U is continuously distributed (as assumed in Assumption A later). Let $s^* = (s_1^*, \cdots, s_I^*)$ be the equilibrium strategy profile and $\sigma_{ij}^*(\cdot,\cdot)$ be a short notation for $\sigma_{ij}^{s^*}(\cdot,\cdot)$. In equilibrium, player i's strategy satisfies a "mutual consistency" requirement, i.e.

$$s_i^*(x, u_i) = \mathbf{1}\left[x_i'\beta_i + \sum_{j \neq i} \alpha_{ij}\sigma_{ij}^*(x, u_i) - u_i \geq 0\right], \quad i = 1, \ldots, I. \tag{2.2}$$

Equation (2.2) are indeed a simultaneous equation system, since player i's equilibrium beliefs σ_{ij}^* on the right-hand side depend on $s_j^*(x,\cdot)$, and vice versa. Therefore, s^* is defined as a fixed point to eq. (2.2). Although ensuring equilibrium existence in Bayesian games is a complex and deep subject in the literature, it is well known that a solution of such an equilibrium generally exists in a broad class of Bayesian games including the binary games under discussion (see, e.g., Vives, 1990).

2.3 Monotone pure strategy BNE

Monotone pure strategy BNEs, in which equilibrium strategies are monotone functions in private types, are desirable in many applications in auctions, entry, social interactions and global games, for example. The seminar work on the existence of a monotone pure strategy BNE in games of incomplete information was provided by Athey (2001) in both *supermodular* and *logsupermodular* games, and later extended by McAdams (2003) and Reny (2011).

To apply Theorem 4.1 in Reny (2011), we make the following assumption.

Assumption A. *Let the conditional distribution of U given X be absolutely continuous w.r.t.*

the Lebesgue measure and have positive and continuous conditional Radon–Nikodym densities $f_{U|X}$ *a.e. over* $\mathbb{R}^I$.

Assumption A requires the conditional c.d.f. $F_{U|X}$ to be twicely differentiable and have a full support on the Euclidean space.

Assumption B (Monotone Best Response Functions)**.** *For all* $x \in \mathscr{S}_X$, $i \in \mathcal{I}$, *and* $v \in \mathbb{R}^I$, *we have* $1 - \sum_{j \neq i} \left\{ \alpha_{ij} \times \partial F_{U_j|X,U_i}(v_j|x, v_i)/\partial u_i \right\} \geq 0$.

Note that Assumption B is trivially satisfied if U are mutually independent. It also holds if $\alpha_{ij} \leq 0$ and U_i and U_j are positively regression dependent for all $i \neq j$.

Lemma 10. *Suppose that Assumptions A and B hold, then there exists at least one monotone pure strategy BNE in our binary discrete games.*

Proof. See Lemma 1 in Liu, Vuong, and Xu (2012). □

It should be noted that we are silent about the existence of non–monotone strategy BNEs under Assumptions A and B in Lemma 10. Xu (2010) shows that non–monotone strategy BNEs can be ruled out under further restrictions on the correlation between private signals. Lemma 10 does not ensure either the uniqueness of monotone pure strategy BNE. Throughout our analysis, we assume that under Assumptions A and B, only one monotone pure strategy BNE is played.

With a monotone pure strategy BNE, player i's equilibrium strategy is a weakly monotone function of her private signal and can be characterized by a threshold function, i.e., fix $x \in \mathscr{S}_X$,

$$s_i^*(x, u_i) = \mathbf{1}\{u_i \leq u_i^*(x)\},$$

where $u_i^* : \mathscr{S}_X \to \mathbb{R}$. Further, the mutual consistency condition for BNEs requires that for all i,

$$u_i \leq u_i^*(x) \Longleftrightarrow x_i'\beta_i + \sum_{j \neq i} \alpha_{ij} \times F_{U_j|X,U_i}\big(u_j^*(x)|x, u_i\big) - u_i \geq 0.$$

In a monotone pure strategy BNE, we can represent the equilibrium strategies as a semi–linear–index binary response model. For all $x \in \mathscr{S}_X$, let $\varphi_{ij}(x) = F_{U_j|X,U_i}\big(u_j^*(x)|x, u_i^*(x)\big)$

and $P_{ij} = \varphi_{ij}(X)$. Let further $P_i = [P_{ij}]_{j \neq i}$ and $\alpha_i = [\alpha_{ij}]_{j \neq i}$ be an $I-1$–dimensional random and deterministic vector, respectively.

Lemma 11. *Suppose that Assumptions A and B hold and that a monotone pure strategy BNE, $s^* = (s_1^*, \cdots, s_I^*)$, is played. Then the structural model can be represented as follows,*

$$Y_i = \mathbf{1}\left[U_i \leq X_i'\beta_i + P_i'\alpha_i\right], \tag{2.3}$$

Proof. See Lemma 2 in Liu, Vuong, and Xu (2012). □

Lemma 11 shows that the equilibrium choices of a monotone pure strategy BNE can be represented as a semi–linear–index binary response model. In such a model, both the player i's payoff relevant vector X_i and the belief P_i of the player at the margin are treated as regressors.

2.4 Identification

In this section, we discuss the semiparametric identification of the structural parameters — α_i, β_i and $F_{U|X}$. The definition of identification of parameters in a structural model follows Hurwicz (1950) and Koopmans and Reiersol (1950), i.e. given the conditional distribution $\mathbb{P}_{Y|X}$ of observables that is generated from a structure with parameter θ_0, the structural parameter θ_0 is identified if there exists a function $\mathscr{G}$ such that $\theta_0 = \mathscr{G}(\mathbb{P}_{Y|X})$.

Our identification strategy takes two steps: first, we establish nonparametric identification of the function φ_{ij} and the (conditional) copula function of the distribution of U; second, we identify (α_i, β_i) and F_{U_i} under an additional location–scale normalization of the payoff function. To proceed, we first make the following assumptions.

Assumption C. *Let $X_i = (W_i, Z_i) \in \mathbb{R}^{d_{W_i}} \times \mathbb{R}^{d_{Z_i}}$ where $d_{W_i} + d_{Z_i} = d_i$. Conditional on $W = (W_1, \cdots, W_I)$, U and $Z = (Z_1, \cdots, Z_I)$ are independent of each other.*

Assumption C assumes the conditional independence between U and Z given W, which has been frequently made in the empirical discrete game literature. See, e.g. Aradillas-Lopez (2010), Bajari, Hong, Krainer, and Nekipelov (2010), and Lewbel and Tang (2011).

Fix $W = w$. For any $i \neq j$ and $(v_i, v_j) \in [0,1]^2$, define a copula function $C_{ij}(\cdot; w)$: $[0,1]^2 \to [0,1]$ as follows:

$$C_{ij}(v_i, v_j; w) = \mathbb{P}\left(U_i \le F_{U_i}^{-1}(v_i), U_j \le F_{U_j}^{-1}(v_j) | W = w\right).$$

By definition, $C_{ij}(v; w) = C_{ji}(v'; w)$, where v' is the vector generated by interchanging the two elements of vector $v \in [0,1]^2$. Let further $V_i = \mathbb{E}(Y_i|X)$. Note that $C_{ij}(\cdot; w)$ can be identified on the support $\mathscr{S}_{V_iV_j|W=w}$ by

$$C_{ij}(v_i, v_j; w) = \mathbb{E}(Y_iY_j|V_i = v_i, V_j = v_j, W = w).$$

where $(v_i, v_j) \in \mathscr{S}_{V_iV_j|W=w}$.

Assumption D. *For some $w \in \mathscr{S}_W$, the support $\mathscr{S}_{V_iV_j|W=w}$ is convex and compact subset of $[0,1]^2$, and has full rank, i.e., dim $\left(\mathscr{S}_{V_iV_j|W=w}\right) = 2$.*

The second half of Assumption D is a restriction similar to the exclusion restriction which requires a rich support for Z conditional on X (see, e.g. Bajari, Hong, Krainer, and Nekipelov, 2010). The first part is restrictive, but can be relaxed significantly. For the brevity of notation, we will not pursue this direction. Please note, however, that the support of (V_i, V_j) given W needs not to be $[0,1]^2$ and, as a consequence, the conditional distribution $F_{U_i|W}(\cdot|w)$ of U_i given $W = w$ can be only disclosed on a subset of $[0,1]$. It should also be noted that the support restriction on (V_i, V_j) given W is only required for some w in the support, instead of the whole support of W.

Assumptions C and D allow us to identify φ_{ij} on the support $\mathscr{S}_{X|W=w}$ as shown in the following lemma.

Lemma 12. *Suppose that Assumptions A to D hold and that a monotone pure strategy BNE, $s^* = (s_1^*, \cdots, s_I^*)$, is played. Then for any $i \neq j$, $\varphi_{ij}(\cdot)$ is identified on the support $\mathscr{S}_{X|W=w}$.*

Proof. See Appendix B.1.1 □

The identification of (α_i, β_i) is similar to the single agent binary response model. By

Lemma 11,

$$F^{-1}_{U_i|W}(V_i|W) = X_i'\beta_i + P_i'\alpha_i \tag{2.4}$$

Let $T_i = [X_i' - \mathbb{E}(X_i'|V_i, W), P_i' - \mathbb{E}(P_i'|V_i, W)]'$. Thus we can define a hyperplane in terms of T_i and payoff coefficients (α_i, β_i):

$$T_i' \times \begin{pmatrix} \beta_i \\ \alpha_i \end{pmatrix} = 0,$$

from which we identify (α_i, β_i) under a scale normalization and a rank condition. Moreover, given the identification of (α_i, β_i) and $\varphi_{ij}(\cdot)$ on support $\mathscr{S}_{X|W=w}$, we can identify $F_{U_i|W}(\cdot|w)$ using the fact that $F_{U_i|W}(X_i'\beta_i + P_i'\alpha_i|W) = \mathbb{E}(Y_i|X)$.

Assumption E. $\|\beta_i\| = 1$.

Assumption E normalizes the scale of β_i only, instead of (α_i, β_i), because in Section 5 we will estimate β_i up to scale in the first stage, therefore this normalization will simplify our estimation analysis.

Assumption F. *For some $w \in \mathscr{S}_W$ satisfying Assumption D, the matrix $\mathbb{E}(T_iT_i'|W = w)$ has a rank of $d_i + I - 2$.*

In addition to Assumption D, Assumption F is another rank condition, which implicitly excludes the constant term in X_i and serves as a location normalization. Assumption F is not a primitive restriction because P_i obtains from the equilibrium. Please note, however, it's not difficult to view that a full rank condition on $X_i' - \mathbb{E}(X_i'|V_i, W)$ and a rich support of $X_{-i}'\beta_{-i}$ given X_i will imply Assumption F.

Theorem 1. *Suppose that Assumptions A to F hold and that a monotone pure strategy BNE, $s^* = (s_1^*, \cdots, s_I^*)$, is played. Then (α_i, β_i) is identified. Moreover, $F_{U_i|W}(\cdot|w)$ is also identified on $\mathscr{S}_{X_i'\beta_i + P_i'\alpha_i|W=w}$.*

The proof of Theorem 1 is straightforward from the above discussion and, therefore, omitted.

2.5 Semiparametric Estimation of Index Payoffs

In this section, we discuss the estimation of coefficients (α_i, β_i) in the payoff function and leave the private type distribution $F_{U|X}$ as a nuisance parameter. For the brevity of notation, we illustrate our method in a two–player setup, i.e. $I = 2$. Our estimation procedure takes three steps: First, we estimate β_i up to scale at a $\sqrt{N}$ rate. Next, we estimate the belief function φ_i at a uniform non–parametric rate using kernel method. Finally, we propose a simple estimator for α_i and show that $\widehat{\alpha}_i$ converges at a $\sqrt{N}$ rate. We also establish asymptotic distributions for $\widehat{\beta}_i$ and $\widehat{\alpha}_i$.

Without causing any confusion, we denote by subscript n (or ℓ, alternatively) the index of observation in a sample and by N the sample size. In contrast, we use subscript i (or j, k, alternatively) to denote the index of player. Let $X_n = (X_{1n}, X_{2n})$ and $Y_n = (Y_{1n}, Y_{2n})$. Following most papers in the empirical game literature, we adopt the i.i.d. assumption in the data generating process as follows.

Assumption G. *Let $\{(X_n, Y_n) : n = 1, \cdots, N\}$ be an i.i.d. random sample.*

2.5.1 Estimation of β_i

In a two–player game, the payoff function for player i becomes

$$\pi_i(y, x_i, u_i) = \begin{cases} x_i'\beta_i + \alpha_i y_{-i} - u_i, & \text{if } y_i = 1, \\ 0, & \text{if } y_i = 0, \end{cases}$$

where the strategic effect coefficient is a scale. Suppose that the conditions in Lemma 10, i.e. Assumptions A and B, hold and that the equilibrium played is a monotone pure strategy BNE, (s_1^*, s_2^*), where $s_i^*(x, u_i) = \mathbf{1}\{u_i \leq u_i^*(x)\}$. Then the mutual consistency restriction requires that

$$x_1'\beta_1 + \alpha_1 \mathbb{P}\left(U_2 \leq u_2^* | X = x, U_1 = u_1^*\right) - u_1^* = 0, \tag{2.5}$$

$$x_2'\beta_2 + \alpha_2 \mathbb{P}\left(U_1 \leq u_1^* | X = x, U_2 = u_2^*\right) - u_2^* = 0. \tag{2.6}$$

Note that there could be multiple solution to eqs. (2.5) and (2.6), but we assume that only one solution contributes to the equilibrium played. We also maintain the following assumption throughout this section, which strengthens Assumption C.

Assumption H. *Let X and U be independent of each other.*

Under Assumption H, we have $F_{U|X} = F_U$, and consequently, $u_i^*(x) = u_i^*(x_1'\beta_1, x_2'\beta_2)$ from eqs. (2.5) and (2.6). Therefore, $\mathbb{E}(Y_i|X) = G_i(X_1'\beta_1, X_2'\beta_2)$, where $G_i(t_1, t_2) = F_{U_i}(u_i^*(t_1, t_2))$. Following the literature on the index models, β_i can be estimated up to scale at a $\sqrt{N}$ rate, which is well discussed (see, e.g. Bierens, 2011; Ichimura, 1993; Klein and Spady, 1993; Powell, Stock, and Stoker, 1989). For example, here we simply describe a procedure to estimate β by following Klein and Spady (1993).

Let $\beta = (\beta_1', \beta_2')'$ and B be the parameter space for β such that Assumption E is satisfied for all its elements. For $y \in \mathcal{A}$, $x \in \mathscr{S}_X$ and $b \in B$, let $P(y|x;b) = \mathbb{E}(\mathbf{1}(Y = y) \mid X_1'b_1 = x_1'b_1, X_2'b_2 = x_2'b_2)$ and $\tilde{P}(y|x_n;b)$ be a Kernel estimator for the conditional probabilities $P(y|x_n;b)$ given the n–th observation $X_n = x_n$, i.e.

$$\tilde{P}(y|x_n;b) = \frac{\sum_{\ell \neq n} \mathbf{1}(Y_\ell = y) K_p\left(\frac{X_{1\ell}'b_1 - x_{1n}'b_1}{h_p}, \frac{X_{2\ell}'b_2 - x_{2n}'b_2}{h_p}\right) + \tilde{\delta}_{1n}(b)}{\sum_{\ell \neq n} K_p\left(\frac{X_{1\ell}'b_1 - x_{1n}'b_1}{h_p}, \frac{X_{2\ell}'b_2 - x_{2n}'b_2}{h_p}\right) + \tilde{\delta}_n(b)},$$

where $K_p(\cdot) : \mathbb{R}^2 \to \mathbb{R}$ denotes a Parzen–Rosenblatt kernel function and h_p is a bandwidth, and $\tilde{\delta}_{1n}$ and $\tilde{\delta}_n$ are trimming sequences introduced for technical reasons, see Klein and Spady (1993) for more details.

Therefore, we define a Klein–Spady type estimator as follows:

$$\tilde{\beta} = \text{argmax}_{b \in B} \sum_{n=1}^{N} (\tilde{\tau}_n / 2) \left\{ \sum_{y \in \mathcal{A}} \left[\mathbf{1}\left\{Y_n = y\right\} \ln \tilde{P}^2(y|X_n;b) \right] \right\},$$

in which $\tilde{\tau}_n$ is a trimming sequence. Given the rich literature on the asymptotic properties of such kind of index estimators, in the following analysis, we simply assume a pilot $\sqrt{N}$–consistent estimator $\tilde{\beta} = \beta + O_p(N^{-1/2})$.

2.5.2 Estimation of Belief Function φ_i

Now we establish a nonparametric estimator for the equilibrium belief function $\varphi_i(\cdot)$. Rather than following the identification strategy in Section 2.4, we derive a different expression for (φ_1, φ_2) here. For $t \in \mathbb{R}^2$ and $i = 1, 2$, let $m_i(t) = \mathbb{E}(Y_i | X_1'\beta_1 = t_1, X_2'\beta_2 = t_2)$. Let further $M(t) = \mathbb{E}(Y_1 Y_2 | X_1'\beta_1 = t_1, X_2'\beta_2 = t_2)$. Then

$$\varphi_1(x) = \frac{\frac{\partial M(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \times \frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} - \frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \times \frac{\partial M(x_1'\beta_1, x_2'\beta_2)}{\partial t_2}}{\frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \times \frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} - \frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \times \frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_2}}, \tag{2.7}$$

$$\varphi_2(x) = \frac{\frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \times \frac{\partial M(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} - \frac{\partial M(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \times \frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_2}}{\frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \times \frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} - \frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \times \frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_2}}, \tag{2.8}$$

which come from the fact that

$$\frac{\partial M(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} = \varphi_1(x) \times \frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} + \varphi_2(x) \times \frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_1},$$
$$\frac{\partial M(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} = \varphi_1(x) \times \frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} + \varphi_2(x) \times \frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_2}.$$

Therefore, we estimate $\varphi_i(X_n)$ for each observation X_n by plugging into the leave–one–out Nadaraya–Watson estimator for each term in eqs. (2.7) and (2.8).

Let

$$\hat{f}_X(x_n) = \sum_{\ell \neq n} K_\varphi \left(\frac{(X_{1\ell} - x_{1n})'\tilde{\beta}_1}{h_\varphi}, \frac{(X_{2\ell} - x_{2n})'\tilde{\beta}_2}{h_\varphi} \right) \Big/ (N-1)h_\varphi^2,$$
$$\hat{q}_i(x_n) = \sum_{\ell \neq n} Y_{i\ell} K_\varphi \left(\frac{(X_{1\ell} - x_{1n})'\tilde{\beta}_1}{h_\varphi}, \frac{(X_{2\ell} - x_{2n})'\tilde{\beta}_2}{h_\varphi} \right) \Big/ (N-1)h_\varphi^2,$$
$$\hat{Q}(x_n) = \sum_{\ell \neq n} Y_{1\ell} Y_{2\ell} K_\varphi \left(\frac{(X_{1\ell} - x_{1n})'\tilde{\beta}_1}{h_\varphi}, \frac{(X_{2\ell} - x_{2n})'\tilde{\beta}_2}{h_\varphi} \right) \Big/ (N-1)h_\varphi^2,$$

where $K_\varphi(\cdot) : \mathbb{R}^2 \to R$ denotes a Parzen–Rosenblatt kernel function and h_φ is a bandwidth. Thus, $M(X_{1n}'\beta_1, X_{2n}'\beta_2)$ and $m_i(X_{1n}'\beta_1, X_{2n}'\beta_2)$ can be estimated by $\hat{Q}(X_n)/\hat{f}_X(X_n)$ and $\hat{q}_i(X_n)/\hat{f}_X(X_n)$, respectively. For notational brevity, we denote $\hat{M}(X_n) = \hat{Q}(X_n)/\hat{f}_X(X_n)$ and $\hat{m}_i(X_n) = \hat{q}_i(X_n)/\hat{f}_X(X_n)$.

Moreover, let

$$\hat{a}_i(x_n) = -\frac{1}{(N-1)h_\varphi^3}\sum_{\ell\neq n}^{N} K_{\varphi i}\left(\frac{(X_{1\ell}-x_{1n})'\tilde{\beta}_1}{h_\varphi}, \frac{(X_{2\ell}-x_{2n})'\tilde{\beta}_2}{h_\varphi}\right),$$
$$\hat{b}_{ji}(x_n) = -\frac{1}{(N-1)h_\varphi^3}\sum_{\ell\neq n}^{N} Y_{j\ell}\times K_{\varphi i}\left(\frac{(X_{1\ell}-x_{1n})'\tilde{\beta}_1}{h_\varphi}, \frac{(X_{2\ell}-x_{2n})'\tilde{\beta}_2}{h_\varphi}\right),$$
$$\hat{c}_i(x_n) = -\frac{1}{(N-1)h_\varphi^3}\sum_{\ell\neq n}^{N} Y_{1\ell}Y_{2\ell}\times K_{\varphi i}\left(\frac{(X_{1\ell}-x_{1n})'\tilde{\beta}_1}{h_\varphi}, \frac{(X_{2\ell}-x_{2n})'\tilde{\beta}_2}{h_\varphi}\right),$$

where $K_{\varphi i}(\cdot)$ denotes the partial derivative of kernel $K_\varphi(\cdot)$ with respect to the i-th argument. Thus we estimate $\partial M(X'_{1n}\beta_1, X'_{2n}\beta_2)/\partial t_i$ by $\hat{f}_X^{-2}(X_n)\left[\hat{c}_i(X_n)\hat{f}_X(X_n) - \hat{a}_i(X_n)\times\hat{Q}(X_n)\right]$, and $\partial m_j(X'_{1n}\beta_1, X'_{2n}\beta_2)/\partial t_i$ by $\hat{f}_X^{-2}(X_n)\left[\hat{b}_{ji}(X_n)\hat{f}_X(X_n) - \hat{a}_i(X_n)\times\hat{q}_j(X_n)\right]$. Hence, we obtain an estimator for $\varphi_i(x_n)$ as follows,

$$\widehat{\varphi}_i(x_n) = \frac{\hat{A}_i(x_n)}{\hat{A}(x_n)}, \tag{2.9}$$

in which for $j \neq i$

$$\begin{aligned}\hat{A}_i(x_n) \equiv & \left[\hat{c}_i(x_n)\hat{f}_X(x_n) - \hat{a}_i(x_n)\hat{Q}(x_n)\right]\times\left[\hat{b}_{jj}(x_n)\hat{f}_X(x_n) - \hat{a}_j(x_n)\hat{q}_j(x_n)\right]\\ & - \left[\hat{c}_j(x_n)\hat{f}_X(x_n) - \hat{a}_j(x_n)\hat{Q}(x_n)\right]\times\left[\hat{b}_{ji}(x_n)\hat{f}_X(x_n) - \hat{a}_i(x_n)\hat{q}_j(x_n)\right],\\ \hat{A}(x_n) \equiv & \left[\hat{b}_{11}(x_n)\hat{f}_X(x_n) - \hat{a}_1(x_n)\hat{q}_1(x_n)\right]\times\left[\hat{b}_{22}(x_n)\hat{f}_X(x_n) - \hat{a}_2(x_n)\hat{q}_2(x_n)\right]\\ & - \left[\hat{b}_{12}(x_n)\hat{f}_X(x_n) - \hat{a}_2(x_n)\hat{q}_1(x_n)\right]\times\left[\hat{b}_{21}(x_n)\hat{f}_X(x_n) - \hat{a}_1(x_n)\hat{q}_2(x_n)\right].\end{aligned}$$

To guarantee a uniform convergence, we further impose a convenient assumption that restricts the denominator in eq. (2.9) to be bounded away from zero almost surely.

Assumption I. *There exists a constant $c_0 > 0$ such that*

$$\inf_{x\in\mathscr{S}_X}\left|\frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_1}\times\frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} - \frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_2}\times\frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_1}\right| \geq c_0.$$

Notice that Assumption I can be avoided by introducing trimming adjustments to the denominator of the belief estimator $\widehat{\varphi}_i(\cdot)$ (see, e.g., Klein and Spady (1993)).

To establish the uniform convergence of our belief estimator $\widehat{\varphi}_i(\cdot)$, we impose some additional assumptions, which are standard for the uniform convergence of kernel estimator, as follows.

Assumption J. *Let $R \geq 1$. For some $\delta > 0$ and all $\beta^\delta \in \{b \in \mathbb{B} : \|b - \beta\| \leq \delta\}$, $f_{X_1'\beta_1^\delta, X_2'\beta_2^\delta}(\cdot)$ is (R+1)–times continuously differentiable on $\mathbb{R}^2$ with bounded (R+1)th–partial derivatives on $\mathbb{R}^2$. Further, $\mathbb{E}(Y_i|(X_1'\beta_1^\delta, X_2'\beta_2^\delta) = \cdot)$ and $\mathbb{E}\left(Y_1 Y_2|(X_1'\beta_1^\delta, X_2'\beta_2^\delta) = \cdot\right)$ are (R+1)–times continuously differentiable on $\mathbb{R}^2$ with bounded (R+1)th–partial derivatives on $\mathbb{R}^2$.*

In particular, $f_{X_1'\beta_1^\delta, X_2'\beta_2^\delta}(\cdot)$ is uniformly continuous on $\mathbb{R}^2$ and integrable. Thus $f_{X_1'\beta_1^\delta, X_2'\beta_2^\delta}(\cdot)$ is bounded, i.e., $\sup_{t \in \mathbb{R}^2} f_{X_1'\beta_1^\delta, X_2'\beta_2^\delta}(t) < \infty$. Moreover, a similar argument also applies to functions $\mathbb{E}(Y_i|(X_1'\beta_1^\delta, X_2'\beta_2^\delta) = \cdot)$ and $\mathbb{E}(Y_1 Y_2|(X_1'\beta_1^\delta, X_2'\beta_2^\delta) = \cdot)$.

Assumption K. *Let $\kappa_N \propto N^\iota$ for some $\iota > 0$ and $\inf_{x \in \mathscr{S}_X} f_X(x) > 0$.*

Let $\eta_N \equiv \inf_{\|x\| \leq \kappa_N} f_X^4(x)$. Note that we can let η_N go to zero at an arbitrary slow rate by choosing small ι. We will derive the uniform convergence of $\widehat{\varphi}_i(x)$ with respect to the compact sub–support $\{x : \|x\| \leq \kappa_N\}$. If the second half condition in Assumption K does not hold, then the observations in the compact sub–support with $f_X^4(x) \leq \eta_N$ need to be trimmed.

Assumption L. *Let $\mathbb{E}|X| < \infty$.*

Assumptions K and L could be replaced by the simpler conditions that the support of X is compact and f_X is bounded away from zero.

Assumption M. *$K_\varphi(u) : \mathbb{R}^2 \to \mathbb{R}$ is (R+1)–continuously differentiable on $\mathbb{R}^2$ with bounded (R+1)th–partial derivatives on $\mathbb{R}^2$. The support of $K_P(\cdot)$ is a convex subset of $\mathbb{R}^2$ with nonempty interior, with the origin as an interior point. $K_\varphi(u)$ satisfies*

$$\begin{aligned} \int u_1^{r_1} u_2^{r_2} K_\varphi(u) dx &= 0 \quad \text{if } r_1 + r_2 = R, \\ &< \infty \quad \text{if } r_1 + r_2 = R + 1. \end{aligned}$$

Assumption N. *Setting $h_\varphi = (\ln N/N)^{1/(2R+4)}$.*

Proposition 11. *Suppose that $\tilde{\beta} = \beta + O_p(N^{-1/2})$. If Assumption G through N hold, then*

$$\sup_{\|x_n\| \leq \kappa_N} \|\widehat{\varphi}_i(x_n) - \varphi_i(x_n)\| = O_p\left(\eta_N^{-1} \left(\frac{\ln N}{N}\right)^{R/(2R+4)}\right).$$

Proof. See Appendix B.2.1 □

Note that our choice of h_φ implies over smoothing for the nonparametric estimation of functions m_i and M and would be sub–optimal in this sense. However, this sub–optimality will not affect the fact that $\widehat{\varphi}_i$ converges uniformly at the best possible rate, which mainly relies on the optimal convergence rate for the derivate estimator of functions m_i and M.

2.5.3 Estimation of Strategic Component α_i

Our final step is to estimate α_i (together with β_i) at a $\sqrt{N}$–convergence rate. Since, in equilibrium, $Y_i = \mathbf{1}\{U_i \leq X_i'\beta_i + \alpha_i\varphi_i(X)\}$ is a single index model on $(X_i, \varphi_i(X))$, we then replace $\varphi_i(X)$ by its estimator $\widehat{\varphi}_i(X)$ and simply follow the approach proposed by Klein and Spady (1993), whose estimator achieves the semi–parametric efficiency bound, to estimate (α_i, β_i).[5] To simplify our discussion and the notation, we use the marginal distribution of Y_i to derive the quasi–likelihood function indexed by $(a_i, b_i) \in A_i \times B_i$, instead of employing the joint distribution of (Y_1, Y_2). Thus, our estimator is defined by

$$(\widehat{\alpha}_i, \widehat{\beta}_i) = \underset{(a_i,b_i)\in A_i\times B_i}{\operatorname{argsup}} \hat{L}_i(a_i, b_i; \hat{\tau}), \tag{2.10}$$

where

$$\hat{L}_i(a_i, b_i; \hat{\tau}) \equiv \sum_{n=1}^{N} (\hat{\tau}_n/2) \left\{ Y_{in} \ln \left[\hat{P}_i(X_n; a_i, b_i)\right]^2 + (1 - Y_{in}) \ln \left[1 - \hat{P}_i(X_n; a_i, b_i)\right]^2 \right\},$$

and

$$\hat{P}_i(X_n; a_i, b_i) = \frac{\sum_{\ell\neq n} \left[Y_{i\ell} \times K_P\left(\frac{(X_{i\ell}-X_{in})'b_i + a_i[\widehat{\varphi}_i(X_\ell) - \widehat{\varphi}_i(X_n)]}{h_P}\right)\right] + \hat{\delta}_{1n}(a_i, b_i)}{\sum_{\ell\neq n} K_P\left(\frac{(X_{i\ell}-X_{in})'b_i + a_i[\widehat{\varphi}_i(X_\ell) - \widehat{\varphi}_i(X_n)]}{h_P}\right) + \hat{\delta}_n(a_i, b_i)},$$

and $\hat{\tau}_n$, $\hat{\delta}_{1n}$ and $\hat{\delta}_n$ are trimming sequences (for details, see Klein and Spady (1993)).

Note that the only difference between our estimator and the one defined in Klein and

[5] Because P_i is bounded between $[0, 1]$ and has positive density close to the boundary, which violate the conditions in Powell, Stock, and Stoker (1989).

Spady (1993) is that we replace the unobserved belief $\varphi_i(X)$ with the belief estimator $\hat{\varphi}_i(X)$. By Proposition 11 and with a similar argument as in Klein and Spady (1993), we can show that $(\hat{\alpha}_i', \hat{\beta}_i')$ is a $\sqrt{N}$–consistent estimator of (α_i', β_i').

Assumption O. *The parameter vector (α_i', β_i') lies in the interior of a compact space $A_i \times B_i \subseteq \mathbb{R} \times \mathbb{R}^{d_i}$.*

Assumption P. *Let X be distributed in a compact support, and η_N be a strict positive constant by setting $\kappa_N \equiv \sup_{x \in \mathscr{S}_X} \|x\|$. Let $f_{X_{i1|o}}(x_{i1})$ be the density for some continuous variable, denoted as X_{i1}, conditioned on the remaining exogenous variables (including X_{-i}), and U. This conditional density is smooth in that for all $x \in \mathscr{S}_X$, there exists a constant $c_1 \in \mathbb{R}_+$ such that*

$$\left|D_{x_{i1}}^r f_{X_{i1}|o}(x_{i1})\right| < c_1, \quad (r = 1, 2, 3, 4).$$

Assumption Q. *With $h_P \to 0$, the trimming function employed to down weight observations has the form*

$$\tau(t, \varepsilon) \equiv \left\{1 + \exp\left[(h_P^{\varepsilon/5} - t)/h_P^{\varepsilon/4}\right]\right\}^{-1},$$

where $\varepsilon > 0$ and t is to be interpreted as a density estimator (e.g. $\hat{f}_{X_i' b_i + a_i \varphi_i(X)}$.) Let

$$\hat{\delta}_{dn} \equiv \tau(\hat{g}_{idn}(\widetilde{a}_{iP}, \widetilde{b}_{iP}), \varepsilon), \quad \text{for } d = 0, 1,$$
$$\text{and} \qquad \hat{\delta}_n \equiv \hat{\delta}_{0n} + \hat{\delta}_{1n},$$

where for $d = 0, 1$,

$$\hat{g}_{idn}(\widetilde{a}_{iP}, \widetilde{b}_{iP}) \equiv \sum_{\ell \neq n}^{N} \frac{\mathbf{1}(Y_{i\ell} = d)}{h_P} K_P\left(\frac{(X_{in} - X_{i\ell})'\widetilde{b}_{iP} + \widetilde{a}_{iP}\left[\hat{\varphi}_i(X_n) - \hat{\varphi}_i(X_\ell)\right]}{h_P}\right) \Big/ (N-1),$$

and $(\widetilde{a}_{iP}, \widetilde{b}_{iP})$ is a preliminary consistent estimator for which $\|(\widetilde{a}_{iP}, \widetilde{b}_{iP}) - (a_i, b_i)\|$ is $O_p(N^{-1/3})$.

Assumption R. *The kernel function, $K_P(\cdot) : \mathbb{R} \to \mathbb{R}$, is a symmetric function that integrates to one, has bounded third moment, and for some $c_2 > 0$,*

$$\max\left\{ \left|D_u^r K_P(u)\right|, \int \left|D_u^r K_P(u)\right| du \right\} < c_2, \quad (r = 0, 1, 2, 3, 4),$$
$$\int u^2 K_P(u) du = 0.$$

Moreover, let h_P be a bandwidth sequence satisfying (i) $N^{-R/(2R+4)} \times h_P^{-2} \to 0$; (ii) $N^{-1/4} < h_P < N^{-1/8}$.

Note that we apply a stronger result of uniform convergence in Hansen (2008), which modifies the lower bound of h_P from $N^{-1/6}$ in Klein and Spady (1993) to $N^{-1/4}$ in our Assumption R, (ii). Assumption R implies that $R > 2$, a restriction to the order of kernel in our first–step estimation.

Assumption S. *For $i = 1, 2$, there exists no proper linear subspace of $\mathbb{R}^d$ having probability 1 under $\mathbb{P}_X$.*

Theorem 2. *Suppose that $\sup_x \|\hat{\varphi}_i(x) - \varphi_i(x)\| = O_p\left((\ln N/N)^{-R/(2R+4)}\right)$ for some $R \geq 1$. If Assumption G through S hold. Then*

$$\sqrt{N}\begin{pmatrix} \hat{\alpha}_i - \alpha_i \\ \hat{\beta}_i - \beta_i \end{pmatrix} \xrightarrow{d} N(0, \Sigma),$$

where

$$\Sigma \equiv \mathbb{E}\left\{\frac{f^2_{U_i}\left(u_i^*(X)\right) \times (\varphi_i(X), X_i')'(\varphi_i(X), X_i')}{F_{U_i}\left(u_i^*(X)\right)\left[1 - F_{U_i}\left(u_i^*(X)\right)\right]}\right\}^{-1}.$$

Proof. See Appendix B.3 □

2.5.4 A sketch of semiparametric estimation in I–player games

Now we consider a discrete game with I players. In the setup specified in Section 2.2, the equilibrium strategy can be written as

$$Y_i = \mathbf{1}\left\{U_i \leq X_i'\beta_i + \sum_{j \neq i} \alpha_{ij} \mathbb{P}\left(U_j \leq u_j^*(X) | U_i = u_i^*(X)\right)\right\}.$$

In our first–step estimation, similar to the two player case, we estimate β by $\tilde{\beta}$ in an I–index model. Second, let $\varphi_{i,j}(x) = \mathbb{P}\left(U_j \leq u_j^*(x) | U_i = u_i^*(x)\right)$, and similar to eqs. (2.7) and (2.8), we can derive an expression for player i's belief φ_{ij} in I–player case as follows,

$$\varphi_{ij}(X)=\frac{\frac{\partial \mathbb{E}(Y_iY_j|X_1'\beta_1,\cdots,X_I'\beta_I)}{\partial t_i}\times\frac{\partial \mathbb{E}(Y_j|X_1'\beta_1,\cdots,X_I'\beta_I)}{\partial t_j}-\frac{\partial \mathbb{E}(Y_j|X_1'\beta_1,\cdots,X_I'\beta_I)}{\partial t_i}\times\frac{\partial \mathbb{E}(Y_iY_j|X_1'\beta_1,\cdots,X_I'\beta_I)}{\partial t_j}}{\frac{\partial \mathbb{E}(Y_i|X_1'\beta_1\cdots,X_I'\beta_I)}{\partial t_i}\times\frac{\partial \mathbb{E}(Y_j|X_1'\beta_1,\cdots,X_I'\beta_I)}{\partial t_j}-\frac{\partial \mathbb{E}(Y_j|X_1'\beta_1,\cdots,X_I'\beta_I)}{\partial t_i}\times\frac{\partial \mathbb{E}(Y_i|X_1'\beta_1,\cdots,X_I'\beta_I)}{\partial t_j}}. \tag{2.11}$$

Hence, we obtain a nonparametric estimator $\widehat{\varphi}_{ij}$ by plugging into the leave–one–out Nadaraya–Watson estimator for each term on the RHS of equation (2.11). By a similar argument as that for Proposition 11, it can be shown that under similar set of conditions, there is

$$\sup_x \left\|\widehat{\varphi}_{ij}(x)-\varphi_{ij}(x)\right\| = O_p\left(\eta_N^{-1}\left(\frac{\ln N}{N}\right)^{R/(2R+I+2)}\right).$$

Finally, by an analogous analysis, we follow Klein and Spady (1993) to obtain a $\sqrt{N}$–consistent estimator for (α_i', β_i') under a similar set of conditions, for which we require $R > 1 + I/2$.

2.6 Monte Carlo Simulations

In this section, we use a numerical experiment to illustrate the performance of our estimator in a finite–size sample. Let $I = 2$, $d_1 = d_2 = 2$ and $X_1 = (X_{11}, X_{12})$ and $X_2 = (X_{21}, X_{22})$, where $X \equiv (X_1, X_2)$ has a mean zero normal distribution with identity covariance matrix. Let U_1 and U_2 be independent of X and conform to a joint mean zero normal distribution with unit variances and correlation parameter $\rho = 0.5$.

Moreover, let $\beta_1 = \beta_2 = (1,1)'$, $\alpha_1 = \alpha_2 = 1$. It can be shown that a (unique) monotone strategy BNE exists under this design, i.e., for each x, there exist cutoff values $u_1^*(x)$ and $u_2^*(x)$, such that player j chooses 1 whenever her private signal $u_j \leq u_j^*(x)$. We compute $u_j^*(x)$ by solving the following equations for each X_n in the sample:

$$u_1^* = \beta_{11}x_{11} + \beta_{12}x_{12} + \alpha_1\Phi\left(\frac{u_2^*-\rho u_1^*}{\sqrt{1-\rho^2}}\right), u_2^* = \beta_{21}x_{21} + \beta_{22}x_{22} + \alpha_2\Phi\left(\frac{u_1^*-\rho u_2^*}{\sqrt{1-\rho^2}}\right).$$

where $\Phi(\cdot)$ is the c.d.f of standard normal distribution.

Table 2.1 shows the composition of one random sample with $N = 500$. In our first–

Table 2.1. Sample composition

Choice profile	Percentage
$Y = (1,1)$	46.0%
$Y = (1,0)$	15.8%
$Y = (0,1)$	17.8%
$Y = (0,0)$	20.4%

step estimation, β_i obtains by the recipe of Klein and Spady (1993). Specifically, we use second order biweight kernel and choose bandwidth according to rule of thumb. Table 2.2 reports summary statistics for $\tilde{\beta}_1$,[6] including the sample mean(MEAN), the median (MEDIAN), the standard deviation (SD), and the root–mean–squared–error (RMSE).

Table 2.2. Finite–Sample Behavior of $\tilde{\beta}_1$

N	TRUE	MEAN	MEDIAN	SD	RMSE
250	1.00	1.0109	0.9969	0.1739	0.1742
500	1.00	1.0063	0.9984	0.1160	0.1161
1000	1.00	1.0038	0.9987	0.0829	0.0830
2000	1.00	1.0037	1.0018	0.0547	0.0548

For the estimation of φ_i, we employ the fourth order biweight product kernel, i.e., $K_\varphi(u_1, u_2) = k(u_1) \cdot k(u_2)$ where $k(u_i) = \frac{7}{4}(1 - 3u_i^2) \cdot \frac{15}{16}(1 - u_i^2)^2 \cdot \mathbf{1}(|u_i| \leq 1)$ and choose $h_\varphi = 4.40 \cdot \widehat{\sigma} \cdot \big(N/log(N)\big)^{-1/10}$ where $\widehat{\sigma}$ is the estimated standard error of the regressor.

Figure 2.1 plots φ_1, φ_2 and their kernel estimates. For presentation purpose, we fix $x_1 = (0,0)$, but a similar pattern holds for other values of x_1. The upper panel shows functions φ_1 and φ_2 and their estimates. The lower–left panel shows the estimate of φ_1 and the infeasible estimate of φ_1 when (β_1, β_2) are known. Further, the lower–right panel shows the the marginal distribution of $\varphi_1(X)$, $f_{\varphi_1(X)}$, and its estimate.

In the last step, we use second order biweight kernel and rule of thumb bandwidth again to implement the Klein and Spady (1993) estimation procedure.

Table 2.3 reports the finite sample performance for estimating α_1 by our three-step estimation procedure. The case of estimating α_2 has similar result. There are five numbers reported for each type of estimator with a certain sample size. The first number refers to the true value of the parameter, the second number refers to the mean (MEAN), the

[6] We normalize $\beta_{11} = 1$, so Table 2.2 actually reports the results for estimating β_{12}.

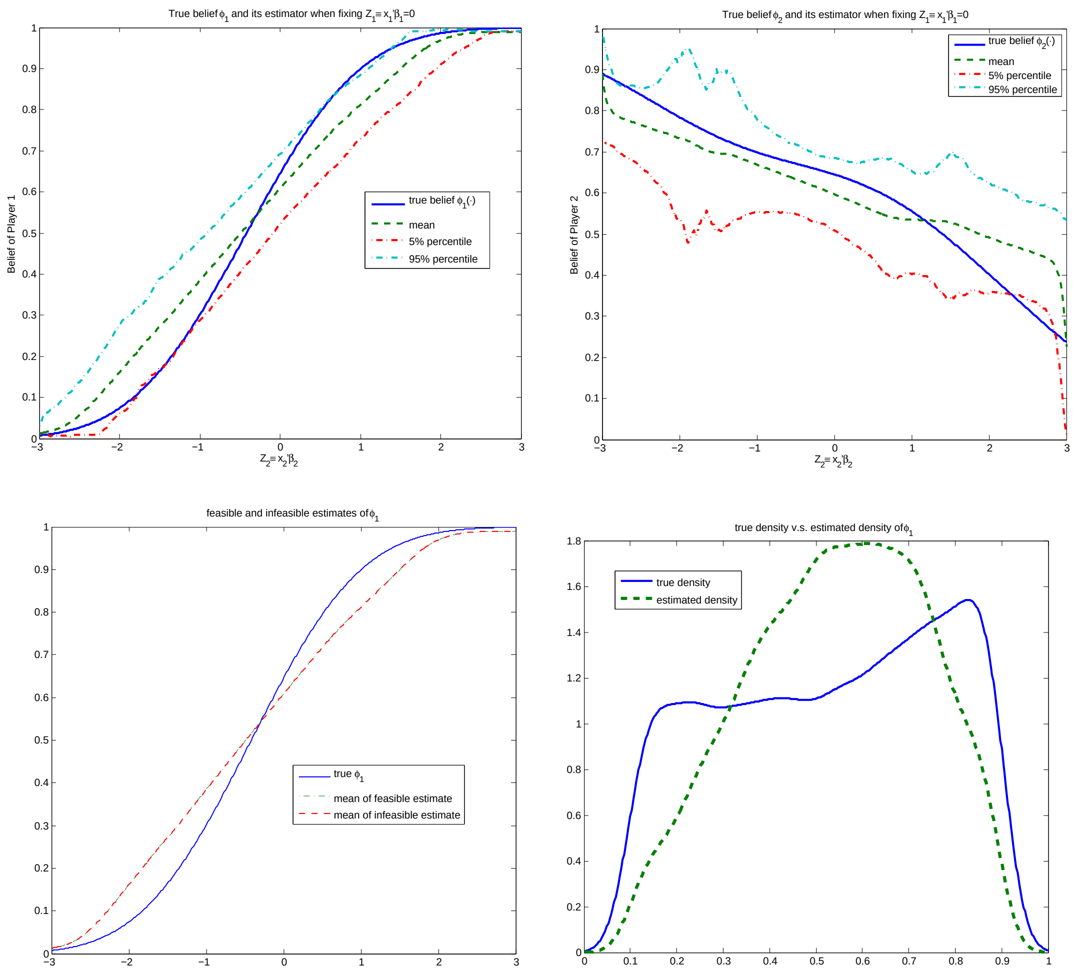

Figure 2.1. Kernel estimates of φ_1, φ_2 and $f_{\varphi_1(X)}$

third one refers to the median (MEDIAN), the fourth one refers to Standard Deviation (SD), and the last one refers to the Root Mean Square Error (RMSE).

Table 2.4 reports the finite sample performance for estimating β_1 in the last step of our estimation procedure.[7] The case of estimating β_2 yields similar result. Similar to Table 2.3, there are five numbers reported in the table.

[7] We normalize $\beta_{11} = 1$, so Table 2.4 actually reports the results for estimating β_{12}.

Table 2.3. Mean, median, SD and RMSE for estimating α_1

Sample size	Our Estimator					Infeasible Estimator				
	true	mean	median	SD	RMSE	true	mean	median	SD	RMSE
250	1.00	0.946	0.926	0.4314	0.4346	1.00	0.988	0.984	0.3347	0.3348
500	1.00	0.988	0.988	0.3022	0.3022	1.00	1.0103	1.0168	0.2366	0.2367
1000	1.00	0.984	0.979	0.2072	0.2078	1.00	1.0032	1.0050	0.1628	0.1628
2000	1.00	0.993	0.994	0.1425	0.1426	1.00	0.999	0.995	0.1067	0.1067

Table 2.4. Mean, median, SD and RMSE for estimating β_1 in last step

Sample size	Our Estimator					Infeasible Estimator				
	true	mean	median	SD	RMSE	true	mean	median	SD	RMSE
250	1.00	1.0197	0.9996	0.1853	0.1861	1.00	1.0163	0.9963	0.1646	0.1652
500	1.00	1.0045	1.0048	0.1161	0.1161	1.00	1.0049	0.9968	0.1114	0.1114
1000	1.00	0.9970	0.9942	0.0826	0.0826	1.00	0.9953	0.9902	0.0774	0.0775
2000	1.00	1.0008	1.0017	0.0557	0.0557	1.00	1.0003	1.0007	0.0518	0.0518

2.7 Conclusion

This paper addresses semiparametric identification and estimation of binary games with arbitrary finite number of players whose private types can be correlated. The part of the model that is not specified is the joint distribution of players' private types. Under exclusion restrictions and rank conditions, we identify the joint distribution of private types nonparametrically and the payoff functions in a linear–index setup. We then propose a three–stage estimation procedure for the payoff coefficients by following Klein and Spady (1993). We also obtain the root-N asymptotic normality of our payoff coefficients estimator. A Monte Carlo experiment provides some evidence for good properties of our estimator in moderately sized samples.

A research line, which needs to be developed, is a semiparmetric estimation procedure for payoff coefficients in a model with the same specification as ours except that players can choose from more than two alternatives. This new model is more complicated than ours, its estimation procedure, however, can follow a strategy similar to ours: In the first step, the players' beliefs are recovered; in the second step, the payoff coefficients are estimated by treating the players' beliefs as regressors and following the semiparametric estimation procedure of multinomial discrete-choice models. The asymptotic properties of such an estimation procedure need to be determined, and are left for future research. The main difficulty relies on the estimated regressors, i.e. the opponents' beliefs which are recovered in the first step.

Chapter 3

Nonparametric Test of Monotonicity of Bidding Strategy in First-price Auctions

3.1 Introduction

The empirical analysis of auctions has become very prominent in the past two decades. First, there has been a large development in the theoretical side since Vickrey (1961). Second, there are huge available data due to the large implementation of auctions in both private sector (e.g. ebay) and public sector for government procurement. Third, the rapid development of econometrics method, especially the nonparametric one, has boosted the empirical analysis of auction data. See, e.g. Paarsch (1992); Porter (1995); Guerre, Perrigne, and Vuong (2000); Athey and Haile (2002). In this paper, we study the nonparametric tests of monotonicity of bidding strategy in first–price auctions. We propose a root–N consistent test statistic and obtain two types of critical values: one of them is given by the asymptotic distribution, and the other one is given through bootstrap approach. We also show that our testing procedure has the correct size and is consistent.

There is a large literature in empirical analysis of first–price auctions. In particular, the range of applications includes, among others, procurement auctions for treeplanting contracts (e.g. Paarsch (1992)), timber auctions (e.g. Elyakime, Laffont, Loisel, and Vuong (1994); Athey, Levin, and Seira (2011)), eggplant sales (e.g. Laffont, Ossard, and Vuong (1995)), procurement auctions of highway repair contracts (e.g. Bajari (1997); Bajari and Ye (2003); Bajari, Houghton, and Tadelis (2011)) and Outer Continental Shelf wildcat auctions (e.g. Li, Perrigne, and Vuong (2000)).

Our paper contributes to the literature in several aspects. First, it is the first one to test monotonicity of bidding strategy non–parametrically. Monotonicity of biding strategy is important to establish the one to one mapping between the unobservables and bids, and is therefore imposed by most papers in empirical analysis of first–price auctions. There is, however, a little discussion on testing monotonicity of bidding strategy in the literature. To the best of our knowledge, our paper is the first one to explore the non-parametric tests of monotonicity of bidding strategy in first-price auctions. In addition, The monotonicity of bidding strategy is, as shown by Guerre, Perrigne, and Vuong (2000), the essential restriction on bids imposed by the theoretical first–price auction models. Consequently, our test is also the essential model specification test of the first–price auctions in the sense that whether the observed bids can be explained by the first–price auction model.

Second, our test statistic is shown to be root–N consistent in spite of the non–parametric nature of the test. The root–N consistency of the test statistic makes our test behaving well in moderately sized samples. Such property is obtained by constructing a test on an integration of the inverse bidding strategy so that the density of bids in the inverse bidding strategy disappears. As a matter of fact, our test is a convexity test of an Integrated Conditional Moment (ICM) function which relates our test to the ICM tests studied by e.g. Bierens and Ploberger (1997).

The rest of the paper is organized as follows. We introduce our monotonicity test problem in Section 2. We then propose a root–N consistent test statistic and establish its asymptotic properties in Section 3. In Section 4, we provide two types of critical values: one is from the asymptotic distribution and the other is given by bootstrap. We show that those two types of critical values have the correct size and is consistent. Section

5 concludes with a discussion of extending our test to the case where auction specific heterogeneities are available. We also discuss briefly how to adjust our test and apply it to other applications in the Industrial Organization literature.

3.2 The Monotonicity Test Problem

In this section, we will give the hypothesis of interest first, and then rewrite it equivalently so that the new hypothesis is easier to test.

3.2.1 Hypothesis of Interest

In structural analysis of auction data, we usually need to specify some theoretical auction model to explain the observed bids. Some aspects of the auction model, e.g. pricing rule such as first–price or second–price, are directly available from the data. Other aspects such as private values or common value among bidders, however, are rarely available from the data. It is therefore empirically relevant for us to see whether there is enough evidence in the data to support the chosen theoretical model. To do so, we need to obtain the restrictions on the observed bids imposed by the chosen theoretical model and then test whether they hold or not. Throughout this paper, we want to see whether a symmetric first–price auction model with Independent Private Value (IPV) can explain a given bids dataset. As shown by Guerre, Perrigne, and Vuong (2000), the restrictions imposed on the bids by such an auction model are given as follows.

Lemma 13. *Let $I \geq 2$. Let $\mathbf{G}(\cdot)$ belong to the set $\mathcal{P}^I$ with support $[\underline{b}, \bar{b}]^I$. There exists a distribution of bidders' private values $F(\cdot) \in \mathcal{P}$ such that $\mathbf{G}(\cdot)$ is the distribution of the equilibrium bids in a first-price sealed-bid auction with independent private values and a non-binding reservation price if and only if:*

C1: $\mathbf{G}(b_1, \ldots, b_I) = \Pi_{p=1}^{I} G(b_p)$.

C2: The function $\xi(\cdot, G, I)$ defined below is strictly increasing on $[\underline{b}, \bar{b}]$ and its inverse is differentiable on $[\underline{v}, \bar{v}] = [\xi(\underline{b}, G, I), \xi(\bar{b}, G, I)]$.

$$\xi(b, G, I) \equiv b + \frac{1}{I-1} \cdot \frac{G(b)}{g(b)}$$

Moreover, when $F(\cdot)$ exists, it is unique with support $[\underline{v}, \bar{v}]$ and satisfies $F(v) = G(\xi^{-1}(v, G, I))$ for all $[\underline{v}, \bar{v}]$. In addition, $\xi(\cdot, G, I)$ is the quasi inverse of the equilibrium strategy in the sense that $\xi(b, G, I) = s^{-1}(b, F, I)$ for all $b \in [\underline{b}, \bar{b}]$.

Lemma 13 is the same as Theorem 1 of Guerre, Perrigne, and Vuong (2000), so its proof is omitted. Lemma 13 says that the restriction of first-price sealed-bid auction model under the symmetry and IPV paradigm is the strict monotonicity of $\xi(\cdot, G, I)$ on $[\underline{b}, \bar{b}]$ if both of $G(\cdot)$ and $g(\cdot)$ are differentiable and there is no binding reservation price. Consequently, under some reguality conditions, we only need to test the strict monotonicity of $\xi(\cdot, G, I)$ if we want to see whether the bids can be rationalized by the first-price auction model within the IPV paradigm. Hereafter, we will omit the dependence of inverse bid function $\xi(\cdot, G, I)$ on G and I, i.e. $\xi(\cdot, G, I)$ will be abbreviated as $\xi(\cdot)$.

We will maintain the following assumption throughout the whole paper

Assumption T. *(i). The bid distribution $G(\cdot)$ is absolutely continuous with a density $g(\cdot)$, and has a support of $\left[\underline{b}, \bar{b}\right]$.*

(ii). For any $b \in \left[\underline{b}, \bar{b}\right]$, $g(b) \geq c_g > 0$.

Part (i) of assumption T restricts the distribution of the observed bids to have a density and a bounded support. Such requirement is mainly for technical convenience. Part (ii) requires the bids density function $g(\cdot)$ to be bounded away from zero in its entire support. This guarantees that the inverse bidding strategy can be defined on the whole support of bids distribution.

From Lemma 13, we ultimately want to construct a hypothesis test as follows:

$$\begin{aligned} &\tilde{H}_0 : \xi(\cdot) \text{ is strictly increasing in } [\underline{b}, \bar{b}] \\ \text{v.s.} \quad &\tilde{H}_1 : \text{otherwise} \end{aligned}$$

where

$$\xi(b) = b + \frac{1}{I-1} \cdot \frac{G(b)}{g(b)}, \quad b \in [\underline{b}, \bar{b}] \tag{3.1}$$

Most papers in the statistical testing literature, however, only consider testing monotonicity rather than testing strict monotonicity. Consequently, we will follow that convention and only consider testing the monotonicity of the inverse bid function $\xi(\cdot)$. In other words, we will consider the following testing problem:

$$\begin{aligned} & \breve{H}_0 : \xi(\cdot) \text{ is (weakly) increasing in } \left[\underline{b}, \bar{b}\right] \\ \text{v.s.} \quad & \breve{H}_1 : \text{otherwise} \end{aligned} \tag{3.2}$$

We remark that the upper lower bounds of the bids support can be normalized to 0 and 1, respectively, by the transformation of $b_i' = \frac{b_i - \underline{b}}{\bar{b} - \underline{b}}$. Such normalization, however, does not affect the monotonic property of the inverse bidding strategy $\xi(\cdot)$. Thus, we can normalize

$$\begin{aligned} \underline{b} &= 0 \\ \bar{b} &= 1 \end{aligned}$$

without loss of generality.

In the next subsection, we are going to rewrite the hypothesis of (3.2) and get an equivalent one.

3.2.2 Equivalent Hypothesis

We rewrite the inverse bid function as follows:

$$\begin{aligned} \xi(b) &= b + \frac{1}{I-1} \cdot \frac{G(b)}{g(b)} \\ &= b + \frac{H(b)}{h(b)} \end{aligned}$$

where $H(b) = G(b)^{I-1}$ is the winning probability for any bidder to bid b, and $h(b) = H'(b) = (I-1) G(b)^{I-2} g(b)$ is the density of $H(b)$.

Note that

$$\begin{aligned} & \xi\left(b\right) = b + \frac{H\left(b\right)}{h\left(b\right)} \text{ is increasing in } b \\ \iff & \xi\left(H^{-1}\left(\tilde{t}\right)\right) \text{ is increasing in } \tilde{t} \in [0,1] \\ \iff & D\left(t\right) \equiv \int_0^t \xi\left(H^{-1}\left(\tilde{t}\right)\right) d\tilde{t} \text{ is convex in } t \in [0,1] \end{aligned}$$

In addition, we also have:

$$\begin{aligned} D\left(t\right) & \equiv \int_0^t \xi\left(H^{-1}\left(\tilde{t}\right)\right) d\tilde{t} \\ & = \int_0^t \left[H^{-1}\left(\tilde{t}\right) + \frac{\tilde{t}}{h\left(H^{-1}\left(\tilde{t}\right)\right)}\right] d\tilde{t} \\ & = \int_0^t d\left[\tilde{t} \cdot H^{-1}\left(\tilde{t}\right)\right] \\ & = t \cdot H^{-1}\left(t\right) \end{aligned} \tag{3.3}$$

Consequently, monotonicity test of inverse bid function $\xi\left(\cdot\right)$ is also equivalent to the following test [1]

$$\begin{aligned} & H_0 : \quad D\left(\cdot\right) \text{ is convex on } [0,1] \\ \text{v.s.} \quad & H_1 : \quad D\left(\cdot\right) \text{ is not convex on } [0,1] \end{aligned} \tag{3.4}$$

where $D\left(\cdot\right)$ is defined by (3.3).

We remark that the idea to get the above equivalent test will become obvious if we rewrite $\xi\left(H^{-1}\left(\tilde{t}\right)\right)$ in the following way:

$$\xi\left(H^{-1}\left(\tilde{t}\right)\right) = H^{-1}\left(\tilde{t}\right) + \tilde{t} \cdot \frac{d\left(H^{-1}\left(\tilde{t}\right)\right)}{d\tilde{t}}.$$

[1] Delgado and Escanciano (2012) use a similar idea to test the stochastic dominance.

3.3 The Test Statistic

3.3.1 Some Definitions

First, let us define the concept of greatest convex minorant (g.c.m.) which will be used frequently later.

Definition 3. *$\underline{\mathcal{C}}_X [U] (x)$ is a g.c.m. of a function $U (x)$ with a domain of X if:*

i. $\underline{\mathcal{C}}_X [U] (x)$ is convex in X

ii. For any $x \in X$, $\underline{\mathcal{C}}_X [U] (x) \leq U (x)$

iii. For any convex function $V (x)$, if $V (x) \leq U (x)$ for any $x \in X$, then $V (x) \leq \underline{\mathcal{C}}_X [U] (x)$ for any $x \in X$.

The existence and uniqueness of g.c.m. of a function have been established under some reguralityconditions in the literature. In addition, this operator is continuous under the sup-norm.

One important implication of the above definition of g.c.m. is that:

$$D (\cdot) \text{ is convex in } [0,1] \iff D (t) = \underline{\mathcal{C}}_{[0,1]} [D] (t) , \quad \forall t \in [0,1]$$

Such implication gives the basic idea to construct the test statistic as follows: the distance between $D (\cdot)$ and $\underline{\mathcal{C}}_{[0,1]} [D] (\cdot)$ measures how far away the function $D (\cdot)$ is from convexity. In principle, a consistent estimator of $D (\cdot)$ should not be far away from its g.c.m. under the null assumption of convexity as the sample size is large enough. Consequently, that distance could be used to form the test statistic for the convexity test of function $D (\cdot)$.

Second, we want to define some estimators being used quite often later. Let $\hat{G}^{-1} (\cdot)$ be the sample quantile function of c.d.f. $G (\cdot)$, i.e.,

$$\hat{G}^{-1} (t) \quad = \quad \inf \left\{ b : \hat{G} (b) \geq t \right\} ,$$

where

$$\hat{G}(b) \quad = \quad \frac{1}{I \cdot L} \sum_{p=1}^{I} \sum_{l=1}^{L} \mathbf{1}\left(B_{pl} \le b\right). \tag{3.5}$$

3.3.2 Test Statistic

We define the test statistic as

$$t_n = \sup_{t \in [0,1]} \sqrt{n} \cdot \left[\hat{D}(t) - \underline{\mathcal{C}}_{[0,1]}\left(\hat{D}\right)(t)\right] \tag{3.6}$$

where $n = I \cdot L$, $\underline{\mathcal{C}}_{[0,1]}$ is the greatest convex minorant (g.c.m.) operator and

$$\hat{D}(t) \quad = \quad t\hat{G}^{-1}\left(t^{\frac{1}{I-1}}\right)$$

which is obtained by the fact of $H^{-1}(t) = G^{-1}(t^{\frac{1}{I-1}})$.

As the sample size is large enough, the sample quantile function $\hat{G}^{-1}(\cdot)$ would be very close to the true quantile function $G^{-1}(\cdot)$, and by functional continuous mapping theorem, $\hat{D}(\cdot)$ and $\underline{\mathcal{C}}_{[0,1]}\left(\hat{D}\right)(\cdot)$ will be very close to $D(\cdot)$ and $\underline{\mathcal{C}}_{[0,1]}(D)(\cdot)$, respectively. Consequently, $\hat{D}(\cdot)$ and $\underline{\mathcal{C}}_{[0,1]}\left(\hat{D}\right)(\cdot)$ are very close to each other under the null H_0, i.e. test statistic t_n should be very small under the null H_0. On the other hand, the test statistic t_n should be very larger under the alternative hypothesis H_1. Hence, the rejection region of t_n should be

$$\left\{t_n \ge c_n^{1-\alpha}\right\}$$

Lemma 14. $\sqrt{n}\left[\hat{D}(t) - D(t)\right]$ *weakly converges to the following process,*

$$\frac{-t \cdot \mathbb{B}(t^{\frac{1}{I-1}})}{g(G^{-1}(t^{\frac{1}{I-1}}))}$$

.

Proof. See Appendix C.1.1. □

The following proposition gives us an up-bound for the test statistic t_n defined in

(3.6) and its asymptotic distribution under the null H_0.

Proposition 12. *Under H_0, we have $t_n \leq \tilde{t}_n$ where*

$$\tilde{t}_n \equiv \sup_{t\in[0,1]} \left\{ \sqrt{n} \left[\hat{D}(t) - D(t) \right] - \sqrt{n} \cdot \underline{\mathcal{C}}_{[0,1]} \left(\hat{D} - D \right)(t) \right\}$$

Besides, $\tilde{t}_n \xrightarrow{d} \sup_{t\in[0,1]} \left\{ \mathbb{K}(t) - \underline{\mathcal{C}}_{[0,1]}(\mathbb{K})(t) \right\}$ where

$$\mathbb{K}(t) := \frac{-t \cdot \mathbb{B}(t^{\frac{1}{I-1}})}{g(G^{-1}(t^{\frac{1}{I-1}}))}, \quad \forall t \in [0,1] \tag{3.7}$$

Proof. See Appendix C.1.2. □

Remark 1. *Notice that $\mathbb{K}(0) = \mathbb{K}(1) = 0$, and there is no singularity issue on $\mathbb{K}(t)$ for any $t \in (0,1)$.*

Remark 2. *Actually, proposition 12 has also provided a way to form a test: we can use the asymptotic distribution of up-bound $\tilde{t}_n$ to determine an asymptotic $(1-\alpha)$ critical value for test statistic t_n. However, that test may be a conservative test in some cases.*

Remark 3. *Notice that we can also form a test statistic based on another estimator of $D(t)$, namely $\tilde{D}(t) = t \cdot \widehat{H}^{-1}(t)$, where $\widehat{H}$ is the empirical c.d.f. of the maximum bids in each auction, and $\widehat{H}^{-1}$ is the corresponding sample quantile function. Using this idea, we can form the following test statistic*

$$t'_n = \sup_{t\in[0,1]} \sqrt{n} \cdot \left[\tilde{D}(t) - \underline{\mathcal{C}}_{[0,1]}[\tilde{D}](t) \right]$$

With such test statistic, we can obtain its up-bound statistic $\bar{t}'_n$ as follows

$$\bar{t}'_n := \sup_{t\in[0,1]} \left\{ \sqrt{n} \left[\tilde{D}(t) - D(t) \right] - \sqrt{n} \cdot \underline{\mathcal{C}}_{[0,1]} \left(\tilde{D} - D \right)(t) \right\}$$

Also, we can get the limiting distribution of $\bar{t}'_n$ as $\sup_{t\in[0,1]} \left[\mathbb{K}'(t) - \underline{\mathcal{C}}_{[0,1]}[\mathbb{K}'](t) \right]$ where $\mathbb{K}'(t)$ is defined as follows

$$\mathbb{K}'(t) := \frac{1}{I-1} \cdot t^{\frac{1}{I-1}} \cdot \frac{-\mathbb{B}(t)}{g(G^{-1}(t^{\frac{1}{I-1}}))}, \quad t \in [0,1]$$

3.3.3 Affine Case of $D(t)$

By definition, affine form of $D(t)$ means that

$$D(t) = \alpha + \beta \cdot t, \qquad t \in [0,1]$$

We can determine the values of α and β from $D(0)$ and $D(1)$ as follows:

$$\begin{cases} 0 = D(0) = \alpha \\ H^{-1}(1) = D(1) = \alpha + \beta \end{cases}$$

whose solution is $\begin{cases} \alpha = 0 \\ \beta = H^{-1}(1) = \bar{b} \end{cases}$. Consequently, the affine form of $D(t)$ is

$$D(t) = \bar{b} \cdot t, \qquad t \in [0,1]$$

which corresponds to $H^{-1}(t) = \bar{b}$ for any $t \in (0,1]$.

In other words, the affine form of $D(t)$ is corresponding to

$$H(b) = \begin{cases} 0, & b \in [\underline{b}, \bar{b}) \\ 1, & b = \bar{b} \end{cases}$$

which is not in the maintained assumption M.

3.4 The Critical Values

Now, the remaining part of the above equivalent testing problem is to determine the critical value of the test statistic t_n. There are two usual ways to determine the critical values of hypothesis tests in the literature: the first one is using the asymptotic distribution of test statistic under the null hypothesis, and the second one is using bootstrap method. In section 3.4.1, we will use the asymptotic distributions of the up-bounds of test statistic, namely $\tilde{t}_n$, to get the asymptotic critical value for test statistic t_n. Then we will also provide the bootstrap critical value in section 3.4.2.

3.4.1 Asymptotic Critical Value

As mentioned earlier, proposition 12 has provided a way to determine some (probably conservative) asymptotic critical values for test statistic t_n through the asymptotic distributions of its up-bounds, namely $\tilde{t}_n$

$$c_{1-\alpha}\left(G^{-1}(\cdot), g(\cdot)\right) \equiv (1-\alpha) \text{ quantile of } \sup_{t\in[0,1]}\left\{\mathbb{K}(t) - \underline{\mathcal{C}}_{[0,1]}[\mathbb{K}](t)\right\}$$

Then we can form the rejection region of t_n as follows

$$\text{rejection region of } t_n: \quad \left\{t_n : t_n \geq c_{1-\alpha}\left(\hat{G}^{-1}(\cdot), \hat{g}(\cdot)\right)\right\}$$

where $\hat{G}^{-1}(\cdot)$ is a quantile function estimator based on empirical c.d.f. $\hat{G}(\cdot)$, and $\hat{g}(\cdot)$ is a kernel density estimator.

Proposition 13. *(i). (Asymptotic size) For any $G(\cdot) \in H_0$, we have*

$$\lim_{n\to\infty} Pr_G\left(t_n \geq c_{1-\alpha}\left(\hat{G}^{-1}(\cdot), \hat{g}(\cdot)\right)\right) \leq \alpha$$

(ii). (Consistency) For any $G(\cdot) \in H_1$, we have

$$\lim_{n\to\infty} Pr_G\left(t_n \geq c_{1-\alpha}\left(\hat{G}^{-1}(\cdot), \hat{g}(\cdot)\right)\right) = 1$$

Proof. See Appendix C.1.3. □

3.4.2 Bootstrap Critical Value

Denote $\tilde{t}_n$ as $R_n(b^{(n)}, G)$ where G is defined as before, and its distribution as $J_n(\cdot, G)$. Besides, let $S_n(b^{(n)}, G) = \sqrt{n}\cdot[\hat{D}(\cdot) - D(\cdot)]$ and $L_n(G)$ be its distribution.

Lemma 15. *The bootstrap satisfies*

$$\rho(L_n(\hat{G}), L_n(G)) \to 0 \textit{ with probability one}$$

where ρ is the distance associated with sup-norm.

Proof. The result follows from Theorem 5.1 of Bickel and Freedman (1981) and its remarks. □

Based on the above result, we can get the following result for the validity of bootstrapping $\tilde{t}_n$:

Lemma 16. *The bootstrap satisfies*

$$\rho(J_n(\cdot,\hat{G}), J_n(\cdot,G)) \to 0 \text{ with probability one}$$

Proof. See Appendix C.1.4. □

Lemma 16 actually establishes the (point-wise) consistency of bootstrapping $\tilde{t}_n$, and consequently, establishes the validity of the test. Denote the $1-\alpha$ quantile of $J_n(\cdot,G)$ as $J_n(G,1-\alpha)$, then the following proposition summarizes the above discussion.

Proposition 14. *(i). (Asymptotic size) For any $G(\cdot) \in H_0$, we have*

$$\lim_{n\to\infty} Pr_G\left(t_n \geq J_n\big(\hat{G},1-\alpha\big)\right) \leq \alpha$$

(ii). (Consistency) For any $G(\cdot) \in H_1$, we have

$$\lim_{n\to\infty} Pr_G\left(t_n \geq J_n\big(\hat{G},1-\alpha\big)\right) = 1$$

Proof. See Appendix C.1.5. □

3.5 Conclusion

This paper develops nonparametric tests of monotonicity of bidding strategy in first price auctions. The original testing problem is to see whether the inverse bid function $\xi(\cdot)$ is increasing. We find that it is equivalent to see whether some transformation of inverse bidding strategy $\xi(\cdot)$, namely $D(\cdot)$, is convex. Consequently, we equivalently change the monotonicity testing problem to a convexity testing problem. With the help of greatest convex minorant operator, we propose a test statistic which is essentially a measure of distance between $D(\cdot)$ and convexity. Such test statistic is shown to be root–N consistent. We then establish the asymptotics of such a test statistic, and propose two types of critical values: one of them is given by the asymptotic distribution, and the other one is given through bootstrap approach. We also show that our testing procedure has the correct size and is consistent.

Due to the celebrating result of Guerre, Perrigne, and Vuong (2000), Our test of monotonicity of bidding strategy is the essential test to see whether the auction data can be rationalized by the symmetric first price auction model under the IPV paradigm, i.e. the monotonicity test is essentially a model specification test of the symmetric first price auction model with IPV.

We leave several extensions, which clearly need to be developed, for future research. The first one is establishing the uniform validity of bootstrap in our test. Our results show that the test is valid with fixed data generating process. The Data Generating Process (DGP) may vary with different sample sizes in some data scenarios, so we need to check whether the testing procedure is robust to such drifting DGP, i.e. whether the test is the uniformly valid.

The second extension is to incorporate auction specific heterogeneity in our test. We can observe auction specific heterogeneities, e.g. engineer's estimate of cost in procurement auctions of highway repair contracts (see Bajari and Ye (2003); Bajari, Houghton, and Tadelis (2011)), in some auction data, so such an extension comes with empirical concern. One straightforward way is to apply our test by replacing the unconditional distribution function and density of bids with the conditional ones. The resulting test, however, will suffer the curse of dimensionality due to the conditioning auction specific heterogeneity.

Third, we might apply our test to other applications in Industrial Organization beyond the first–price auctions considered here. Li, Perrigne, and Vuong (2000) obtains some restrictions, which are similar to those of Guerre, Perrigne, and Vuong (2000), imposed on bids by Conditionally Independent Private Value (CIPV) model, thus our test might be applicable to such context after some adjustment. In a nonlinear pricing framework, Perrigne and Vuong (2011) shows that the monotonicity relationship between the private type of a buyer θ and the buyer's purchased quantity q is the essential restriction on observables imposed by the theoretical model. In addition, Hickman (2010) shows that the monotonicity relationship between the marginal utility cost of a student and her achievement is main restriction of his model, and their relationship is expressed similarly to eq. (3.1). Due to the similarity, our test might also apply to these applications in Perrigne and Vuong (2011) and Hickman (2010).

Appendix **A**

Appendix to Chapter 1

A.1 Proofs of Results in Section 1.2

A.1.1 Proof of Lemma 1

First, Assumptions G1–G6 of Reny (2011) are satisfied in our discrete game under Assumption R. Moreover, by Assumption M, when other players employ monotone decreasing pure strategies, player i's best response is also a monotone decreasing pure strategy, which is unique. By Reny (2011, Theorem 4.1), the conclusion follows. □

A.1.2 Existence of monotone pure strategy BNEs under primitive conditions

Definition 4. *a set $A \subseteq \mathbb{R}^d$ is upper if and only if its indicator function is non–decreasing, i.e., for any $x, y \in \mathbb{R}^d$, $x \in A$ and $x \leq y$ imply $y \in A$, where $x \leq y$ means $x_i \leq y_i$ for $i = 1, \cdots, d$.*

Assumption U (Positive regression dependence). *For any $x \in \mathscr{S}_X$ and any upper set $A \subseteq \mathbb{R}^{I-1}$, the conditional probability $\mathbb{P}\left(U_{-i} \in A | X = x, U_i = u_i\right)$ is non–decreasing in $u_i \in \mathscr{S}_{U_i|X=x}$.*

Assumption V (Strategic complement). *For any $x \in \mathscr{S}_X$ and $u_i \in \mathscr{S}_{U_i|X=x}$, suppose $a_{-i} \leq a'_{-i}$, then $\pi_i(a_{-i}, x, u_i) \leq \pi_i(a'_{-i}, x, u_i)$.*

Assumption W (Non-increasing Payoffs). *$\forall\ i = 1, \cdots, I$ and $\forall\ (a_{-i}, x) \in \mathcal{A}_{-i} \times \mathscr{S}_X$, $\pi_i(a_{-i}, x, \cdot)$ are non-increasing functions in $u_i \in \mathscr{S}_{U_i|X=x}$.*

Lemma 17. *Suppose that assumptions R, A, B and C hold. For any $x \in \mathscr{S}_X$, there exists an m.d.p.s. BNE.*

Proof. By Lemma 1, it suffices to show that Assumption M holds. Fix $x \in \mathscr{S}_X$. Given an arbitrary m.d.p.s. profile: $\delta_i(x, u_i) = \mathbf{1}[u_i \leq u_i(x)]$ for $i = 1, \cdots, I$, where $u_i(x) \in \mathscr{S}_{U_i|X=x}$. By assumptions U and V, and Lehmann (1955), for any $u_i < u_i'$ in $\mathscr{S}_{U_i|X=x}$, we have

$$\mathbb{E}_\delta \left[\pi_i(Y_{-i}, X, u_i)|X = x, U_i = u_i'\right] \leq \mathbb{E}_\delta \left[\pi_i(Y_{-i}, X, u_i)|X = x, U_i = u_i\right].$$

Further, by assumption W,

$$\mathbb{E}\delta \left[\pi_i(Y_{-i}, X, u_i')|X = x, U_i = u_i'\right] \leq \mathbb{E}_\delta \left[\pi_i(Y_{-i}, X, u_i)|X = x, U_i = u_i'\right].$$

Thus, $\mathbb{E}_\delta \left[\pi_i(Y_{-i}, X, U_i)|X = x, U_i = u_i\right]$ is a non–increasing function of u_i. □

A.1.3 Proof of Lemma 11

Fix $X = x$. By Assumption S, there is $\mathbb{E}_{\delta^*} \left[\pi_i(Y_{-i}, X, U_i)\middle|X = x, U_i = u_i\right] = \mathbb{E}_{\delta^*}[\pi_i(Y_{-i}, X)$ $\middle|X = x, U_i = u_i] - u_i$. Because $\mathbb{E}_{\delta^*} \left[\pi_i(Y_{-i}, X)\middle|X = x, U_i = u_i\right]$ is a linear combination of $\sigma^*_{-i}(a_{-i}|x, u_i)$ for all $a_{-i} \in \mathcal{A}_{-i}$, which are continuous in u_i under Assumption R, then $\mathbb{E}_{\delta^*} \left[\pi_i(Y_{-i}, X, U_i)\middle|X = x, U_i = u_i\right]$ is a continuous and monotone decreasing function in u_i under Assumption M. Hence, the cutoff value $u_i^*(x)$ defining player i's equilibrium strategy satisfies: if $\underline{u}_i(x) < u_i^*(x) < \overline{u}_i(x)$, we have

$$\mathbb{E}_{\delta^*} \left[\pi_i(Y_{-i}, X)\middle|X = x, U_i = u_i^*(x)\right] - u_i^*(x) = 0,$$

which implies that: conditional on $\underline{u}_i(X) < u_i^*(X) < \overline{u}_i(X)$, there is

$$Y_i = \mathbf{1}\left[U_i \leq u_i^*(X)\right] = \mathbf{1}\left[U_i \leq \sum_{a_{-i}} \pi_i(a_{-i}, X)\sigma^*_{-i}\left(a_{-i}|X, u_i^*(X)\right)\right].$$

If $u_i^*(x) = \overline{u}_i(x)$, then $\mathbb{E}_{\delta^*} \left[\pi_i(Y_{-i}, X)\middle|X = x, U_i = \overline{u}_i(x)\right] - \overline{u}_i(x) \geq 0$, which implies that: conditional on $u_i^*(x) = \overline{u}_i(x)$, there is

$$Y_i = \mathbf{1}\left[U_i \leq \overline{u}_i(X)\right] \leq \mathbf{1}\left[U_i \leq \sum_{a_{-i}} \pi_i(a_{-i}, X)\sigma^*_{-i}\big(a_{-i}|X, \overline{u}_i(X)\big)\right].$$

Because $\mathbf{1}\left[U_i \leq \overline{u}_i(X)\right] = 1$ a.s., thus

$$Y_i = \mathbf{1}\left[U_i \leq \overline{u}_i(X)\right] = \mathbf{1}\left[U_i \leq \sum_{a_{-i}} \pi_i(a_{-i}, X)\sigma^*_{-i}\big(a_{-i}|X, \overline{u}_i(X)\big)\right] \quad a.s..$$

Similar arguments hold for the case $u_i^*(x) = \underline{u}_i(x)$. □

A.2 Proofs of Results in Section 1.3

A.2.1 Proof of Proposition 1

Fix $x \in \mathscr{S}_X$. First, we assume $\mathbb{P}(Y = a|X = x) > 0$ for all $a \in \mathcal{A}$. Now we construct a structure in $\mathcal{M}_1$ to rationalize $F_{Y|X}(\cdot|x)$. Let $\pi_i(a_{-i}, x, u_i) = \mathbb{E}(Y_i|X = x) - u_i$ for $i = 1, \cdots, I$. Note that there is no strategic effect by construction and Assumption M is satisfied. Now we construct $F_{U|X}(\cdot|x)$. Let $F_{U_i|X}(\cdot|x)$ be uniformly distributed on $[0, 1]$. So it suffices to construct the copula function $C_{U|X}(\cdot|x)$. Let $C_{U|X}(\alpha_1, \cdots, \alpha_I|x) = 0$ if $\alpha_i = 0$ for some i. Then only restriction left for constructing such a copula is: on the support $\{\mathbb{E}(Y_1|X = x), 1\} \times \cdots \times \{\mathbb{E}(Y_I|X = x), 1\}$, let $C_{U|X}(\alpha_1, \cdots, \alpha_I|x) = \mathbb{E}(\prod_{j=1}^{p} Y_{i_j}|X = x)$ where $i_1, \cdots, i_p$ are all the indexes such that $\alpha_{i_j} = \mathbb{E}(Y_{i_j}|X = x)$. On this sub–support, it is straightforward that $C_{U|X}(\cdot|x)$ is monotone increasing in each index by the fact that $\mathbb{P}(Y = a|X = x) > 0$ for all $a \in \mathcal{A}$. Thus it is straightforward that we can extend $C_{U|X}(\cdot|x)$ to the whole support $[0, 1]^I$ such that $C_{U|X}(\cdot|x)$ is monotone increasing and smooth. It can be shown that the given conditional distribution of Y given $X = x$ can be rationalized by this constructed structure in $\mathcal{M}_1$: $\mathbb{P}(Y_{i_1} = 1, \cdots, Y_{i_p} = 1|X = x) = C_{U_{i_1},\cdots,U_{i_p}|X}(\mathbb{E}(Y_{i_1}|X = x), \cdots, \mathbb{E}(Y_{i_p}|X = x))$ for any subset index $\{i_1, \cdots, i_p\}$.

When $\mathbb{P}(Y = a|X = x) = 0$ for some a's in $\mathcal{A}$. By the condition in Proposition 1, the conditional distribution of Y given $X = x$ is degenerated in some indexes. W.l.o.g., let $\{1, \cdots, k\}$ be set of indexes such that $\mathbb{P}(Y_i = 1|X = x) = 0$ or 1; and $\{k+1, \cdots, I\}$ satisfying $0 < \mathbb{P}(Y_i = 1|X = x) < 1$. Then let $\pi_i(a_{-i}, x, u_i) = \mathbb{E}(Y_i|X = x) - u_i$ for $i = 1, \cdots, I$. For player $i = k+1, \cdots, I$, we can construct a copula function $C_{U_{k+1},\cdots,U_I|X}(\cdot|x)$ as described above such that $C_{U_{k+1},\cdots,U_I|X}(\cdot|x)$ is monotone increasing and smooth. Similarly, the constructed structure rationalizes the given conditional distribution of Y given $X = x$. □

A.2.2 Rationalizing All Probability Distributions

Suppose we replace Assumption R with the following conditions in Reny (2011): For every $x \in \mathscr{S}_X$,

G.2. The distribution $F_{U_i|X}(\cdot|x)$ on $\mathscr{S}_{U_i|X=x}$ is atomless.

G.3. There is a countable subset $\mathscr{S}^0_{U_i|X=x}$ of $\mathscr{S}_{U_i|X=x}$ such that every set in $\mathscr{S}_{U_i|X=x}$ assigned positive probability by $F_{U_i|X}(\cdot|x)$ contains two points between which lies a point in $\mathscr{S}^0_{U_i|X=x}$.

Note that it is straightforward that Assumptions G.1 and G.4 through G.6 in Reny (2011) are all satisfied in our discrete game because the action space $\mathcal{A}$ is finite and the conditional distribution of U given $X = x$ has a hypercube support in $\mathbb{R}^I$. Thus, the conclusion in Lemma 1 still holds (i.e., existence of a monotone pure strategy BNE) under Assumptions G.2, G.3 and M. Moreover, if we define

$$\mathcal{M}'_1 \equiv \{S : \text{G.2, G.3 and M hold, and a single m.d.p.s. BNE is played}\}$$

Then we can show the following result:

Lemma 18. *For any $x \in \mathscr{S}_X$, $F_{Y|X}(\cdot|x)$ can be rationalized by a structure $S \equiv [\pi_1, \cdots, \pi_I; F_{U|X}]$ $\in \mathcal{M}'_1$.*

Proof. Fix $x \in \mathscr{S}_X$. Now we construct a structure in $\mathcal{M}'_1$ to rationalize $F_{Y|X}(\cdot|x)$. Let $\pi_i(a_{-i}, x, u_i) = \alpha_i(x) - u_i$ for $i = 1, \cdots, I$. Note that there is no strategic effect by construction and Assumption M is satisfied. Now we construct $F_{U|X}(\cdot|x)$. Let $[0,1]^I$ be the support of the distribution and partition it into 2^I disjoint events:

$\bigotimes_{i=1}^{I}\{[0,\alpha_i(x)),[\alpha_i(x),1]\}$ [1]. For each event B_j, ($j=1,\cdots,2^I$), we define a conditional distribution $F_{U|X=x,U\in B_j}$ as a uniform distribution on B_j, where B_j is the j–th event in the partition of the support. Moreover, let $\mathbb{P}(U\in B_j|X=x)=\mathbb{P}(Y=a(j)|X=x)$ where $a(j)\in\mathcal{A}$ and satisfies $a_i(j)=0$ if the i–th argument of event B_j is $[\alpha_i(x),1]$, and $a_i(j)=1$ if the i–th argument is $[0,\alpha_i(x))$. With such construction, the marginal distribution of U_i given $X=x$ is a uniform distribution on $[0,1]$ which satisfies Assumptions G.2 and G.3. Thus, the constructed structure $S\equiv[\pi_1,\cdots,\pi_I;F_{U|X}]\in\mathcal{M}_1'$ and it rationalizes $F_{Y|X}(\cdot|x)$ by construction. □

A.2.3 Proof of Lemma 3

Fix $x\in\mathscr{S}_X$. For any structure $S\in\mathcal{M}_1$, let $\tilde{F}_{U|X}(\cdot|x)=F_{U|X}(\cdot|x)$ and $\tilde{\pi}_i(a_{-i},x,u_i)=u_i^*(x)-u_i$ where $\left(u_1^*(x),\ldots,u_I^*(x)\right)$ is the equilibrium cut-off value profile under structure S. It is easy to see that $\tilde{S}\equiv[\tilde{\pi};\tilde{F}_{U|X}]\in\mathcal{M}_2$, and $\tilde{S}$ is observationally equivalent to the given structure S. □

A.2.4 Proof of Lemma 4

The only if part is straightforward: Note that $\alpha_i(x)=\mathbb{E}(Y_i|X=x)=\mathbb{P}(Y_i=1|X=x)$. Thus, given two observationally equivalent structures $S\equiv[\pi,F_{U|X}]$ and $\widetilde{S}\equiv[\tilde{\pi};\tilde{F}_{U|X}]$ in $\mathcal{M}_2$, condition (i) requires that both structures lead to the same value for $\mathbb{P}(Y_i=1|X=x)$, while condition (ii) requires to have the same value for $\mathbb{P}(Y_{i_1}=1,\cdots,Y_{i_p}=1|X=x)$.

For the if part, it suffices to verify that for every $x\in\mathscr{S}_X$, the conditional equilibrium choice probability $\mathbb{P}\left(Y=a|X=x\right)$ induced by the structure S can also be generated by another structure $\widetilde{S}$ that satisfies the two conditions in Lemma 4. We verify this for $a=(1,\cdots,1)$ only. By the definition of monotone pure strategy BNE, we have $\mathbb{P}_{\tilde{\delta}^*}\left(Y_1=\cdots=Y_I=1|X=x\right)=\tilde{F}_{U|X}\left(\tilde{u}_1^*(x),\cdots,\tilde{u}_I^*(x)|x\right)$. By condition (i), $\tilde{u}_i^*(x)=\tilde{F}_{U_i|X}^{-1}\left(\alpha_i(x)|x\right)$. Thus

$$\mathbb{P}_{\tilde{\delta}^*}\left(Y_1=\cdots=Y_I=1|X=x\right) \quad = \quad \tilde{F}_{U|X}(\tilde{F}_{U_1|X}^{-1}(\tilde{\alpha}_1(x)|x),\cdots,\tilde{F}_{U_I|X}^{-1}(\tilde{\alpha}_I(x)|x)|x)$$

[1]To have meaningful partition, it is understood that $\{[0,\alpha_i(x)),[\alpha_i(x),1]\}$ becomes $\{\{0\},(0,1]\}$ when $\alpha_i(x)=0$.

$$
\begin{aligned}
&= \tilde{C}_{U|X}\left(\tilde{\alpha}_1(x), \cdots, \tilde{\alpha}_I(x) \middle| x\right) \\
&= C_{U|X}\left(\alpha_1(x), \cdots, \alpha_I(x) \middle| x\right) \\
&= \mathbb{P}_{\delta^*}\left(Y_1 = \cdots = Y_I = 1 | X = x\right)
\end{aligned}
$$

where the third equality follows from (ii). □

A.2.5 Proof of Lemma 5

We first show the "only if" part. Suppose that the observationally equivalent structure is $\widetilde{S} = [\tilde{\pi}; \tilde{F}_{U|X}] \in \mathcal{M}_3$. Then $\tilde{F}_{U|X} = \tilde{F}_U$. Condition (ii) of Lemma 4 implies that

$$
C_{U_{i_1},\cdots,U_{i_p}|X}\left(\alpha_{i_1}(x), \cdots, \alpha_{i_p}(x) \middle| x\right) = \widetilde{C}_{U_{i_1},\cdots,U_{i_p}}\left(\alpha_{i_1}(x), \cdots, \alpha_{i_p}(x)\right) \tag{A.1}
$$

for every $x \in \mathscr{S}_X$, every $p = 2, \cdots, I$ and every i_js such that $1 \leq i_1 < \cdots < i_p \leq I$. Consequently, we have

$$
\begin{aligned}
& m_p(\alpha_{i_1}(x), \cdots, \alpha_{i_p}(x)) \\
\equiv\ & \mathbb{E}\left[C_{U_{i_1},\cdots,U_{i_p}|X}\left(\alpha_{i_1}(X), \cdots \alpha_{i_p}(X) \middle| X\right) | \alpha_{i_1}(X) = \alpha_{i_1}(x), \cdots \alpha_{i_p}(X) = \alpha_{i_p}(x)\right] \\
=\ & \mathbb{E}\left[\widetilde{C}_{U_{i_1},\cdots,U_{i_p}}\left(\alpha_{i_1}(X), \cdots, \alpha_{i_p}(X)\right) | \alpha_{i_1}(X) = \alpha_{i_1}(x), \cdots, \alpha_{i_p}(X) = \alpha_{i_p}(x)\right] \\
=\ & \widetilde{C}_{U_{i_1},\cdots,U_{i_p}}\left(\alpha_{i_1}(x), \cdots, \alpha_{i_p}(x)\right)
\end{aligned} \tag{A.2}
$$

for every $x \in \mathscr{S}_X$. Condition (i) is then established by eqs. (A.1) and (A.2).

Notice that eq. (A.2) also implies that $m_p(\alpha_{i_1}, \cdots, \alpha_{i_p}) = \widetilde{C}_{U_{i_1},\cdots,U_{i_p}}(\alpha_{i_1}, \cdots, \alpha_{i_p})$ for every $(\alpha_{i_1}, \cdots, \alpha_{i_p})$ in the support of $(\alpha_{i_1}(X), \cdots, \alpha_{i_p}(X))$. In addition, $\tilde{C}_{U_{i_1},\cdots,U_{i_p}}$ is monotone strictly increasing and continuously differentiable by Assumption R. Thus, condition (ii) holds due to the strict monotonicity of $\tilde{C}_{U_{i_1},\cdots,U_{i_p}}$, and condition (iii) can be obtained by the continuous differentiability of $\widetilde{C}_{U_{i_1},\cdots,U_{i_p}}$.

For the if part, our proof is constructive. For any $x \in \mathscr{S}_X$, let $\tilde{\pi}_i(a_{-i}, x, u_i) = \alpha_i(x) - u_i$ for $i = 1, \cdots, I$, where $\alpha_i(x) = F_{U_i|X}(u_i^*(x)|x)$. We further construct $\tilde{F}_U$. Let $\tilde{F}_{U_i}$ be uniformly distributed on $[0, 1]$; And for all $1 \leq i_1 < \cdots < i_p \leq I$, $(\alpha_{i_1}, \cdots, \alpha_{i_p}) \in$

$\mathscr{S}_{\alpha_{i_1}(X),\cdots,\alpha_{i_p}(X)}$ and $x \in \mathscr{S}_X$, define $\widetilde{F}_{U_{i_1},\cdots,U_{i_p}}(\cdot,\cdots,\cdot)$ as follows

$$\widetilde{F}_{U_{i_1},\cdots,U_{i_p}}(\alpha_{i_1},\cdots,\alpha_{i_p}) = \mathbb{E}[C_{U_{i_1},\cdots,U_{i_p}|X}(\alpha_{i_1}(X),\cdots,\alpha_{i_p}(X)|X)\big|\alpha_{i_1}(X)=\alpha_{i_1},\cdots,\alpha_{i_p}(X)=\alpha_{i_p}],$$

which is monotone strictly increasing on $\mathscr{S}_{\alpha_{i_1}(X),\cdots,\alpha_{i_p}(X)}$ by condition (ii) and continuously differentiable in the interior of $\mathscr{S}_{\alpha_{i_1}(X),\cdots,\alpha_{i_p}(X)}$ by condition (iii). Thus we can extend the distribution $\tilde{F}_U(\cdot)$ to the whole support $[0,1]^I$ such that (1) $\tilde{F}_U(\cdot)$ is monotone strictly increasing and continuously differentiable on $[0,1]^I$; (2) $\tilde{F}_U(0,\cdots,0)=0$ and $\tilde{F}_U(1,\cdots,1)=1$. Thus $\tilde{F}_U(\cdot)$ is a proper distribution function and yields a positive and continuous conditional Radon–Nikodym density on $[0,1]^I$. By construction, $[\tilde{\pi};\tilde{F}_U(\cdot)] \in \mathcal{M}_3$, and by Lemma 4 it is observationally equivalent to the structure S, because $\tilde{F}_{U_i}(\alpha_i(x)) = \alpha_i(x) = F_{U_i|X}(u_i^*(x)|x)$ where $(\alpha_1(x),\cdots,\alpha_I(x))$ is an equilibrium cutoff value profile under constructed structure $[\tilde{\pi};\tilde{F}_U(\cdot)]$, and

$$C_{U_{i_1},\cdots,U_{i_p}|X}(\alpha_{i_1}(x),\cdots,\alpha_{i_p}(x)|x) = \tilde{F}_{U_{i_1},\cdots,U_{i_p}}(\alpha_{i_1}(x),\cdots,\alpha_{i_p}(x))$$

which is implied by condition (i). This completes the proof. □

A.2.6 Proof of Proposition 3

Proof. Note that $C_{U_{i_1},\cdots,U_{i_p}|X}(\alpha_{i_1}(x),\cdots,\alpha_{i_p}(x)|x) = F_{U_{i_1},\cdots,U_{i_p}|X}\left(u_{i_1}^*(x),\cdots,u_{i_p}^*(x)|x\right) = \mathbb{E}\left(\prod_{j=1}^p Y_{i_j}\Big|X=x\right)$. In addition,

$$\begin{aligned}
& m_p(\alpha_{i_1}(x),\cdots,\alpha_{i_p}(x)) \\
= \;& \mathbb{E}\left[C_{U_{i_1},\cdots,U_{i_p}|X}\left(\alpha_{i_1}(X),\cdots\alpha_{i_p}(X)\big|X\right)|\alpha_{i_1}(X)=\alpha_{i_1}(x),\cdots\alpha_{i_p}(X)=\alpha_{i_p}(x)\right] \\
= \;& \mathbb{E}\left[\mathbb{E}\left(\prod_{j=1}^p Y_{i_j}|X\right)|\alpha_{i_1}(X)=\alpha_{i_1}(x),\cdots\alpha_{i_p}(X)=\alpha_{i_p}(x)\right] \\
= \;& \mathbb{E}\left[\prod_{j=1}^p Y_{i_j}|\alpha_{i_1}(X)=\alpha_{i_1}(x),\cdots\alpha_{i_p}(X)=\alpha_{i_p}(x)\right]
\end{aligned}$$

Thus, conditions (R1)-(R3) follow from (i)-(iii) in Lemma 5. □

A.2.7 Proof of Lemma 6

Proof. The only if part follows directly from Assumption I. It suffices to show the if part.

We use a constructive approach to show the if part. Fix an arbitrary structure $[\pi; F_{U|X}] \in \mathcal{M}_3$ for which eq. (1.11) is satisfied. Fix arbitrarily $x \in \mathscr{S}_X$. Let $\tilde{F}_{U_i|X}(\cdot|x) = F_{U_i}(\cdot)$ and $\tilde{F}_{U|X}(\cdot|x) = \prod_{i=1}^{I} F_{U_i}(\cdot)$. Moreover, for any $x \in \mathscr{S}_X$, let $\tilde{\pi}_i(a_{-i}, x, u_i) = u_i^*(x) - u_i$ where $(u_1^*(x), \ldots, u_I^*(x))$ is the equilibrium cut-off value profile under the given structure $[\pi; F_{U|X}]$ and $X = x$. By construction, $[\tilde{\pi}; \tilde{F}_{U|X}]$ satisfies Assumptions M, S, E, and I. Regarding to Assumption R, it suffices to show that $F_{U_i}(\cdot)$ is absolutely continuous w.r.t. Lebesgue measure and has a continuous conditional Radon–Nikodym density $f_{U_i}(\cdot)$, which is true due to the fact $[\pi; F_{U|X}] \in \mathcal{M}_3$. Hence, $[\tilde{\pi}; \tilde{F}_{U|X}] \in \mathcal{M}_4$.

By construction, it is straightforward to show that the conditions (i) and (ii) in Lemma 4 are satisfied by the constructed structure $[\tilde{\pi}; \tilde{F}_{U|X}]$ and the given structure S, which ensures the observational equivalence between the two structures. □

A.2.8 Proof of Proposition 4

Proof. Note that the condition given in Lemma 6 is stronger than those in Lemma 5, so the necessary and sufficient condition for a structure in $\mathcal{M}_2$ to be observationally equivalent to a structure in $\mathcal{M}_4$ is the same as the one in Lemma 6. In addition, we have $C_{U_{i_1},\cdots,U_{i_p}|X}\left(\alpha_{i_1}(x), \cdots, \alpha_{i_p}(x) \middle| x\right) = F_{U_{i_1},\cdots,U_{i_p}|X}\left(u_{i_1}^*(x), \cdots, u_{i_p}^*(x) \middle| x\right) = \mathbb{E}\left(\prod_{j=1}^{p} Y_{i_j} \middle| X = x\right)$. Thus condition (1.11) becomes

$$E\left(\prod_{j=1}^{p} Y_{i_j} | X = x\right) = \prod_{j=1}^{p} E(Y_{i_j} | X = x)$$

for $p = 2, \cdots, I$, which implies that $Y_1, \cdots, Y_I$ are conditionally independent given X. □

A.3 Proofs of Results in Section 1.4

A.3.1 Proof of Lemma 7

Proof. Our proof is similar to that in Darsow, Nguyen, and Olsen (1992), who use a copula argument. Fix $X = x$. Because,

$$\begin{aligned}
&\mathbb{P}\left(Y_i = 1; Y_{-i} = a_{-i} | \alpha(X) = \alpha\right) \\
&\quad = \mathbb{P}\left(U_i \leq F_{U_i}^{-1}(\alpha_i); \forall a_j = 1,\ U_j \leq F_{U_j}^{-1}(\alpha_j);\ \forall a_j = 0,\ U_j > F_{U_j}^{-1}(\alpha_j)\right) \\
&\quad = \int_{-\infty}^{F_{U_i}^{-1}(\alpha_i)} \mathbb{P}\left(\forall a_j = 1,\ U_j \leq F_{U_j}^{-1}(\alpha_j);\ \forall a_j = 0,\ U_j > F_{U_j}^{-1}(\alpha_j) \middle| U_i = u_i\right) f_{U_i}(u_i) du_i,
\end{aligned}$$

in which the last step comes from the law of iterated expectation. Then

$$\begin{aligned}
&\left.\frac{\partial \mathbb{P}\left(Y_i = 1; Y_{-i} = a_{-i} | \alpha(X) = \alpha\right)}{\partial \alpha_i}\right|_{\alpha = \alpha(x)} \\
&\quad = \mathbb{P}\left(\forall a_j = 1,\ U_j \leq F_{U_j}^{-1}(\alpha_j(x));\ \forall a_j = 0,\ U_j > F_{U_j}^{-1}(\alpha_j(x)) \middle| U_i = F_{U_i}^{-1}(\alpha_i(x))\right) \\
&\quad = \mathbb{P}\left(\forall a_j = 1,\ U_j \leq u_j^*(x);\ \forall a_j = 0,\ U_j > u_j^*(x) \middle| U_i = u_i^*(x)\right) \\
&\quad = \mathbb{P}\left(\forall a_j = 1,\ U_j \leq u_j^*(x);\ \forall a_j = 0,\ U_j > u_j^*(x) \middle| X = x, U_i = u_i^*(x)\right) \\
&\quad = \mathbb{P}\left(Y_{-i} = a_{-i} | X = x, U_i = u_i^*(x)\right) = \sigma_{-i}^*(a_{-i} | x, u_i^*(x)).
\end{aligned}$$

Note that the RHS of the first equation need to multiply with an additional term $f_{U_i}\left(F_{U_i}^{-1}(\alpha_i)\right) \times \frac{1}{f_{U_i}\left(F_{U_i}^{-1}(\alpha_i)\right)}$, which equals to one, hence, omitted. □

A.3.2 Model restrictions imposed by Assumption ER

Let $\mathcal{M}_5 \equiv \{S \in \mathcal{M}_3 : \text{Assumption ER holds}\}$.

Lemma 19. *Suppose $\mathscr{S}_{\alpha(X)} = [0,1]^I$. Thus, for any given structure $S \in \mathcal{M}_3$, there exists an observationally equivalent structure $\widetilde{S} \in \mathcal{M}_5$ if and only if for each $i = 1, \cdots, I$ there exist a function $\beta_i : \mathcal{A}_{-i} \times \mathscr{S}_{X_i} \to \mathbb{R}$ and a strictly monotone increasing and differentiable function $\varphi_i : [0,1] \to \mathbb{R}$ such that*

(i) *for all* $x \in \mathscr{S}_X$, *there is*

$$\varphi_i(\alpha_i(x)) = \sum_{a_{-i} \in \mathcal{A}_{-i}} \beta_i(a_{-i}, x_i) \times \sigma^*_{-i}(a_{-i}|x, u_i^*(x)).$$

(ii) *for all* $x_i \in \mathscr{S}_{X_i}$ *and* $\alpha_{-i} \in [0,1]^{I-1}$, *the function* $\sum_{a_{-i} \in \mathcal{A}_{-i}} \beta_i(a_{-i}, x_i) \times \sigma^U_{-i}(a_{-i}|\alpha_{-i}, \alpha_i) - \varphi_i(\alpha_i)$ *is monotone decreasing in* α_i, *where* $\sigma^U_{-i}(a_{-i}|\alpha_{-i}, \alpha_i)$ *is the conditional probability* $\mathbb{P}\big(\forall a_j = 1, F_{U_j}(U_j) \le \alpha_j; \forall a_j = 0, F_{U_j}(U_j) > \alpha_j \big| F_{U_i}(U_i) = \alpha_i\big)$.

Proof. For the only if part. Suppose $[\pi; F_U] \in \mathcal{M}_3$ and $[\tilde{\pi}; \tilde{F}_U] \in \mathcal{M}_5$ are observationally equivalent. Then $C_U = \tilde{C}_U$ by Lemma 4. Let $\varphi_i = \tilde{F}^{-1}_{U_i}$ and $\beta_i(a_{-i}, x_i) = \tilde{\pi}_i(a_{-i}, x_i)$. It is straightforward that condition (i) is satisfied by eq. (1.6). For condition (ii), it should be noted that the monotonicity of $\sum_{a_{-i} \in \mathcal{A}_{-i}} \beta_i(a_{-i}, x_i) \times \sigma^U_{-i}(a_{-i}|\alpha_{-i}, \alpha_i) - \varphi_i(\alpha_i)$ in α_i is equivalent to the monotonicity of $\sum_{a_{-i} \in \mathcal{A}_{-i}} \beta_i(a_{-i}, x_i) \times \sigma^U_{-i}\big(a_{-i}|\alpha_{-i}, \varphi_i^{-1}(u_i)\big) - u_i$ in u_i, given the strict monotonicity of φ_i.

For the if part, it suffices to construct a structure $\tilde{S} \in \mathcal{M}_5$: Let $\tilde{C}_{U_1 \cdots U_I}(\alpha) = C_{U_1 \cdots U_I}(\alpha)$ for all $\alpha \in [0,1]^I$, $\tilde{F}_{U_i}(u_i) = \varphi_i^{-1}(u_i)$ for all $u_i \in [\varphi_i(0), \varphi_i(1)]$ and $\tilde{\pi}_i(a_{-i}, x_i, u_i) = \beta_i(a_{-i}, x_i) - u_i$ for all $(a_{-i}, x_i, u_i) \in \mathcal{A}_{-i} \times \mathscr{S}_X \times [\varphi_i(0), \varphi_i(1)]$. By construction, $\tilde{S} \in \mathcal{M}_5$. For any fixed $x \in \mathscr{S}_X$, we can verify that $\mathbf{1}\left\{\cdot \le \tilde{F}^{-1}_{U_i}(\alpha_i(x))\right\}$ for $i = 1, \cdots, I$ is an equilibrium solution under $\tilde{S}$, then $\tilde{S}$ is observationally equivalent to S. $\square$

Proposition 15. *Suppose* $\mathscr{S}_{\alpha(X)} = [0,1]^I$. *Thus, a conditional distribution* $\mathbb{P}_{Y|X}$ *generated from a structure in* $\mathcal{M}_3$ *can also be rationalized by a structure in* $\mathcal{M}_5$ *if and only if*

(i) *Fix each* i *and every* $x_i \in \mathscr{S}_{X_i}$. *If* $\mathscr{S}_{\alpha_i(X)|X_i=x_i}$ *is not a singleton, then for every* $\alpha_i \in \mathscr{S}_{\alpha_i(X)|X_i=x_i}$, *there exist* $(\bar{\beta}_i, \gamma_i) \in \mathbb{R}^{2^{I-1}} \times \mathbb{R}$, $\|\bar{\beta}_i\| = 1$, *such that for all* $x \in \mathscr{S}_{X|X_i=x_i, \alpha_i(X)=\alpha_i}$

$$\sum_{a_{-i} \in \mathcal{A}_{-i}} \bar{\beta}_{i,a_{-i}} \times \sigma^*_{-i}(a_{-i}|x, u_i^*(x)) = \gamma_i.$$

The hyperplanes for different $\alpha_i \in \mathscr{S}_{\alpha_i(X)|X_i=x_i}$ *are parallel to each other, i.e., the slope term* $\bar{\beta}_i$ *is constant across* α_i, *denoted as* $\bar{\beta}_i(x_i)$; *The intercept term, denoted as* $\gamma_i(x_i, \alpha_i)$, *can* *be represented by*

$$\gamma_i(x_i, \alpha_i) = \rho_i(x_i) \times \bar{\varphi}_i(\alpha_i) + \varrho_i(x_i)$$

where $\bar{\varphi}_i : \mathscr{S}_{\alpha_i(X)} \to \mathbb{R}$ *is a monotone increasing and differentiable function,* $\rho_i : \mathscr{S}_{X_i} \to \mathbb{R}^{++}$ *is a nonnegative function and* $\varrho_i : \mathscr{S}_{X_i} \to \mathbb{R}$.

(ii) For each i, $x_i \in \mathscr{S}_{X_i}$, *and* $\alpha_{-i} \in [0,1]^{I-1}$, *the function* $\sum_{a_{-i} \in \mathcal{A}_{-i}} \beta_{i,a_{-i}}(x_i) \times \sigma^U_{-i}(a_{-i}|\alpha_{-i}, \alpha_i) - \gamma_i(x_i, \alpha_i)$ *is monotone decreasing in* α_i.

Proof. It directly follows from Lemma 19. □

Remark 4. *Condition (i) in Proposition 15 implies that*

a. *Condition on* X_i, $\sigma^*_{-i}(a_{-i}|X, u^*_i(X))$, $a_{-i} \in \mathcal{A}_{-i}$, *and* $\varphi_i(\alpha_i(X))$ *are perfectly linearly dependent, i.e., the latter can be linearly represented as a function of the former variables.*

b. *Condition on* X_i, $\overline{\sigma}^*_{-i}(a_{-i}, X)$, $a_{-i} \in \mathcal{A}_{-i}/\{a^0_{-i}\}$, *are perfectly linearly dependent, where* $\overline{\sigma}^*_{-i}(a_{-i}, X) = \sigma^*_{-i}(a_{-i}|X, u^*_i(X)) - \mathbb{E}\left[\sigma^*_{-i}(a_{-i}|X, u^*_i(X))\middle|X_i, \alpha_i(X)\right]$.

A.3.3 Proof of Proposition 6

Proof. When $\alpha_i \in \mathscr{S}_{\alpha_i(X)|X_i=x_i} \cap (0,1)$, there exists some $x \in \mathcal{S}_X$ such $\alpha_i(x) = \alpha_i$. Because $\mathbb{E}_{\delta^*}\left[\pi_i(Y_{-i}, X_i, U_i)\middle|X = x, U_i = u_i\right]$ is a continuous and monotone function in u_i by Assumptions R, M, and C, we then have

$$0 = \mathbb{E}_{\delta^*}\left[\pi_i(Y_{-i}, X_i, U_i)\middle|X = x, U_i = u^*_i(x)\right] = \sum_{a_{-i}} \pi_i(a_{-i}, x_i, u^*_i(x)) \times \sigma^*_{-i}(a_{-i}|x, u^*_i(x)).$$

Since $u^*_i(x) = Q_{U_i}(\alpha_i)$, then we have

$$\sum_{a_{-i} \in \mathcal{A}_{-i}} \pi_i\big(a_{-i}, x_i, Q_{U_i}(\alpha_i)\big) \times \sigma^*_{-i}(a_{-i}|x, u^*_i(x)) = 0. \tag{A.3}$$

Thus, if $\mathcal{R}_i(x_i, \alpha_i)$ has full rank 2^{I-1}, eq. (A.3) only admits a unique solution: $\pi_i(\cdot, x_i, Q_{U_i}(\alpha_i)) = 0$. Moreover, by Assumption M, for any $u_i \in \mathscr{S}_{U_i}$ and m.d.p.s. profile δ we have: if $u_i > Q_{U_i}(\alpha_i)$,

$$\mathbb{E}_\delta\left[\pi_i(Y_{-i}, X_i, U_i)\middle|X = x, U_i = u_i\right] < \mathbb{E}_\delta\left[\pi_i(Y_{-i}, X_i, U_i)\middle|X = x, U_i = Q_{U_i}(\alpha_i)\right]$$

and if $u_i < Q_{U_i}(\alpha_i)$,

$$\mathbb{E}_\delta\left[\pi_i(Y_{-i}, X_i, U_i)\middle|X = x, U_i = u_i\right] > \mathbb{E}_\delta\left[\pi_i(Y_{-i}, X_i, U_i)\middle|X = x, U_i = Q_{U_i}(\alpha_i)\right].$$

Therefore, $\mathcal{S}_{\alpha_i(X)|X_i=x_i}$ must be a singleton $\{\alpha_i\}$.

Further, if $\mathcal{R}_i(x_i, \alpha_i)$ has rank $2^{I-1} - 1$, then conditioning on $X_i = x_i$ and $\alpha_i(X) = \alpha_i$, the random vector $\Sigma^*_{-i}(X)$ consists of a set of random variables which can be linearly independent if we exclude some variable. By excluding, w.l.o.g., $\sigma^*_{-i}\left(a^0_{-i}|X, u^*_i(X)\right)$ from $\Sigma^*_{-i}(X)$, the random variables left in Σ^*_{-i} are conditionally linearly independent given $X_i = x_i$ and $\alpha_i(X) = \alpha_i$. Because

$$\sum_{a_{-i}\in\mathcal{A}_{-i}\setminus\{a^0_{-i}\}} \pi_i\big(a_{-i}, x_i, Q_{U_i}(\alpha_i)\big) \times \sigma^*_i(a_{-i}|x, u^*_i(x)) = -\sigma^*_i(a^0_{-i}|x, u^*_i(x))\pi_i\big(a^0_{-i}, x_i, Q_{U_i}(\alpha_i)\big),$$

the rank condition implies that the $\pi_i(a_{-i}, x_i, Q_{U_i}(\alpha_i))$ are identified for all $a_{-i} \in \mathcal{A}_{-i}\setminus\{a^0_{-i}\}$ up to the scale $\pi_i\big(a^0_{-i}, x_i, Q_{U_i}(\alpha_i)\big)$.

Now we show the identification of the sign of $\pi_i\big(\cdot, x_i, Q_{U_i}(\alpha_i)\big)$. W.L.O.G., let $\alpha'_i < \alpha_i$. Let further $x' \in \mathscr{S}_X$ such that $\alpha(x') = \alpha'$. Then by Assumption M, there is

$$\sum_{a_{-i}\in\mathcal{A}_{-i}} \pi_i\big(a_{-i}, x_i, Q_{U_i}(\alpha_i)\big) \times \sigma^*_i(a_{-i}|x', Q_{U_i}(\alpha_i))$$
$$< \sum_{a_{-i}\in\mathcal{A}_{-i}} \pi_i\big(a_{-i}, x_i, Q_{U_i}(\alpha'_i)\big) \times \sigma^*_i(a_{-i}|x', Q_{U_i}(\alpha'_i)) = 0$$

By the proof of Lemma 7, $\sigma^*_i(a_{-i}|x', Q_{U_i}(\alpha_i))$ is identified by taking derivative of $\mathbb{P}\big(Y_i = 1; Y_{-i} = a_{-i}|\alpha(X) = (\alpha_i, \alpha'_{-i})\big)$ with respect to α_i at (α_i, α'_{-i}). Moreover, since $\pi_i\big(\cdot, x_i, Q_{U_i}(\alpha_i)\big)$ are identified up to the unknown scale $\pi_i\big(a^0_{-i}, x_i, Q_{U_i}(\alpha_i)\big)$, let

$$\pi_i\big(\cdot, x_i, Q_{U_i}(\alpha_i)\big) = k_i\big(\cdot, x_i, Q_{U_i}(\alpha_i)\big) \times \pi_i\big(a^0_{-i}, x_i, Q_{U_i}(\alpha_i)\big)$$

where $k_i \in \mathbb{R}^{2^{I-1}}$ is known. By definition, $k_i\left(a^0_{-i}, x_i, Q_{U_i}(\alpha_i)\right) = 1$. Thus

$$\left\{\sum_{a_{-i}\in\mathcal{A}_{-i}} k_i\left(a_{-i}, x_i, Q_{U_i}(\alpha_i)\right) \times \sigma_i^*(a_{-i}|x', Q_{U_i}(\alpha_i))\right\} \times \pi_i\left(a^0_{-i}, x_i, Q_{U_i}(\alpha_i)\right) < 0,$$

from which we identify the sign of $\pi_i\left(a^0_{-i}, x_i, Q_{U_i}(\alpha_i)\right)$ and hence the sign of $\pi_i\left(a_{-i}, x_i, Q_{U_i}(\alpha_i)\right)$.

□

A.3.4 Proof of Proposition 7

With additive separability in the payoffs as well as Assumption ER, we can obtain the following equilibrium condition by Lemma 11,

$$\sum_{a_{-i}\in\mathcal{A}_{-i}} \pi_i(a_{-i}, x_i)\sigma^*_{-i}\left(a_{-i}|x, u_i^*(x))\right) = Q_{U_i}(\alpha_i(x)) \tag{A.4}$$

for all $x \in \mathscr{S}_X$ such that $\alpha_i(x) \in (0,1)$. It follows that

$$\sum_{a_{-i}\in\mathcal{A}_{-i}} \pi_i(a_{-i}, x_i)\mathbb{E}\left[\sigma^*_{-i}\left(a_{-i}|X, u_i^*(X))\right)|X_i = x_i, \alpha_i(X) = \alpha_i(x)\right] = Q_{U_i}(\alpha_i(x)) \tag{A.5}$$

The difference between eq. (A.4) and (A.5) yields

$$\sum_{a_{-i}\in\mathcal{A}_{-i}} \pi_i(a_{-i}, x_i) \times \overline{\sigma}^*_{-i}\left(a_{-i}, x\right) = 0 \tag{A.6}$$

where $\overline{\sigma}^*_{-i}\left(a_{-i}, x\right) \equiv \sigma^*_{-i}\left(a_{-i}|x, u_i^*(x)\right) - \mathbb{E}\left[\sigma^*_{-i}\left(a_{-i}|X, u_i^*(X)\right)|X_i = x_i, \alpha_i(X) = \alpha_i(x)\right]$.

When x_i is fixed, we can identify $\pi_i(\cdot, x_i)$ as coefficients by varing $\overline{\sigma}^*_{-i}(a_{-i}, x)$ through x_{-i}. Thus, if $\overline{\mathcal{R}}_i(x_i)$ has rank $2^{I-1} - 1$, by a similar argument as that in the proof for Proposition 6, $\pi_i(\cdot, x_i)$ are identified up to scale. Note that by definition $\sum_{a_{-i}\in\mathcal{A}_{-i}} \overline{\sigma}^*_{-i}\left(a_{-i}, x\right) = 0$. Hence, $\pi_i(a_{-i}, x_i)$ equals to the same constant for all $a_{-i} \in \mathcal{A}_{-i}$. Moreover, similarly to Proposition 6, $\mathscr{S}_{\alpha_i(X)|X_i=x_i}$ has to be a singleton set, i.e. $\{\alpha_i(x)\}$ for some $x \in \mathscr{S}_{X|X_i=x_i}$. Therefore, $\pi_i(a_{-i}, x_i) = Q_{U_i}(\alpha_i(x))$ for all a_{-i}.

Further, suppose $\overline{\mathcal{R}}_i(x_i)$ has rank $2^{I-1} - 2$. Then there exists a known vector $k_i(\cdot, x_i) \in$

$\mathbb{R}^{2^{I-1}}$, $k_i \neq (1, \cdots, 1)$, such that $k_i(\cdot, x_i)$ and $(1, \cdots, 1)$ satisfy

$$\sum_{a_{-i} \in \mathcal{A}_{-i}} k_i(a_{-i}, x_i) \times \overline{\sigma}^*_{-i}(a_{-i}, x) = 0,$$
$$\sum_{a_{-i} \in \mathcal{A}_{-i}} 1 \times \overline{\sigma}^*_{-i}(a_{-i}, x) = 0.$$

Then, by linear algebra, $\pi_i(\cdot, x_i)$ can be written as

$$\pi_i(\cdot, x_i) = c_i(x_i) + b_i(x_i) \times k_i(\cdot, x_i)$$

where $c_i(x_i), b_i(x_i) : \mathscr{S}_{X_i} \to \mathbb{R}$. Hence, π_i are identified up to location (c_i) and scale (b_i).

The identification of the sign of $\pi_i(a_i, x_i) - \pi_i(a_i^0, x_i)$, it is similar to that in proof for Proposition 6: let $x, x' \in \mathscr{S}_{X|X_i = x_i}$, $\alpha(x) = \alpha$, $\alpha(x') = \alpha'$, and w.l.o.g., $\alpha_i' < \alpha_i$. Then

$$\sum_{a_{-i} \in \mathcal{A}_{-i}} \pi_i(a_{-i}, x_i)\sigma^*_{-i}\big(a_{-i}|x', Q_{U_i}(\alpha_i)\big) < Q_{U_i}(\alpha_i(x)) = \sum_{a_{-i} \in \mathcal{A}_{-i}} \pi_i(a_{-i}, x_i)\sigma^*_{-i}\big(a_{-i}|x, Q_{U_i}(\alpha_i)\big),$$

from which we have

$$b_i(x_i) \times \sum_{a_{-i} \in \mathcal{A}_{-i}} k_i(a_{-i}, x_i) \times \left[\sigma^*_{-i}\big(a_{-i}|x', Q_{U_i}(\alpha_i)\big) - \sigma^*_{-i}\big(a_{-i}|x, Q_{U_i}(\alpha_i)\big)\right] < 0.$$

Thus we identify the sign of $b_i(x_i)$. It follows that the sign of $\pi_i(a_{-i}, x_i) - \pi_i(a'_{-i}, x_i) = b_i(x_i) \times \left[k_i(a_{-i}, x_i) - k_i(a'_{-i}, x_i)\right]$ is also identified. □

A.3.5 Proof of Proposition 8

Proof. It suffices to show that suppose the identification of $\pi_i(\cdot, x_i)$ for all $x_i \in \mathscr{X}_i^t$ implies that $\pi_i(\cdot, x_i)$ is point identified for all $x_i \in \mathscr{X}_i^{t+1}$.

Let $x_i \in \mathscr{X}_i^{t+1}$. The claim is straightforward if $x_i \in \mathscr{X}_i^t$. Hence, let $x \in \mathscr{X}_i^{t+1} / \mathscr{X}_i^t$. Then it could be Case (ii) or (iii).

Suppose that Case (ii) occurs, i.e. $\overline{\mathcal{R}}_i(x_i)$ has rank $2^{I-1} - 2$ and there exists $x_i' \in \mathscr{X}_i^t$ such that $\mathscr{S}_{\alpha_i(X)|X_i = x_i} \bigcap \mathscr{S}_{\alpha_i(X)|X_i = x_i'} \bigcap (0, 1)$ contains at least two different elements, $0 < \alpha_i' < \alpha_i < 1$. Because $x_i' \in \mathscr{X}_i^t$, then by assumption $\pi_i(\cdot, x_i')$ are identified. Then both

$Q_{U_i}(\alpha_i)$ and $Q_{U_i}(\alpha_i')$ are identified by the equilibrium condition.

Further, because $\overline{\mathcal{R}}_i(x_i)$ has a rank $2^{I-1}-2$, then by Proposition 7 $\pi_i(\cdot,x_i)$ is identified up to location and scale, i.e. $\exists\, a_i(x_i), b_i(x_i) \in \mathbb{R}$ and a known vector $k_i(\cdot,x_i) \in \mathbb{R}^{2^{I-1}}$, such that $\pi_i(\cdot,x_i) = a_i(x) + b_i(x) \times k_i(\cdot,x_i)$. Moreover, because $\alpha_i, \alpha_i' \in \mathscr{S}_{\alpha_i(X)|X_i=x_i} \bigcap (0,1)$, then we have

$$\sum_{a_{-i}\in\mathcal{A}_{-i}} \pi_i(\cdot,x_i) \times \sigma^*_{-i}(a_{-i}|x,u_i^*(x)) = Q_{U_i}(\alpha_i),$$
$$\sum_{a_{-i}\in\mathcal{A}_{-i}} \pi_i(\cdot,x_i) \times \sigma^*_{-i}(a_{-i}|x',u_i^*(x')) = Q_{U_i}(\alpha_i').$$

It follows that

$$a_i(x_i) + b_i(x_i) \times \sum_{a_{-i}\in\mathcal{A}_{-i}} k_i(\cdot,x_i) \times \sigma^*_{-i}(a_{-i}|x,u_i^*(x)) = Q_{U_i}(\alpha_i),$$
$$a_i(x_i) + b_i(x_i) \times \sum_{a_{-i}\in\mathcal{A}_{-i}} k_i(\cdot,x_i) \times \sigma^*_{-i}(a_{-i}|x',u_i^*(x')) = Q_{U_i}(\alpha_i').$$

Note that $Q_{U_i}(\alpha_i), Q_{U_i}(\alpha_i'), k_i(\cdot,x_i), \sigma^*_{-i}(a_{-i}|x,u_i^*(x))$ and $\sigma^*_{-i}(a_{-i}|x',u_i^*(x'))$ are all known terms and $Q_{U_i}(\alpha_i') < Q_{U_i}(\alpha_i)$. Then we can identify $a_i(x_i)$ and $b_i(x_i)$ from above equations. Therefore, $\pi_i(\cdot,x_i)$ are identified.

Suppose that Case (iii) occurs, i.e. $\overline{\mathcal{R}}_i(x_i')$ has a rank $2^{I-1}-1$ and there exists $x_i' \in \mathscr{X}_i^t$ such that $\mathscr{S}_{\alpha_i(X)|X_i=x_i} \subseteq \mathscr{S}_{\alpha_i(X)|X_i=x_i'} \bigcap (0,1)$. Because of the rank condition, $\pi_i(\cdot,x_i)$ is identified by $Q_{U_i}(\alpha_i)$, which is known by the fact $\mathscr{S}_{\alpha_i(X)|X_i=x_i} \subseteq \mathscr{S}_{\alpha_i(X)|X_i=x_i'} \bigcap (0,1)$. $\square$

A.3.6 Necessary and sufficient condition for two structures in $\mathcal{M}_3$ to be observationally equivalent

Lemma 20. *Suppose that two structures $S \equiv [\pi; F_U]$ and $\widetilde{S} \equiv [\tilde{\pi}; \tilde{F}_U]$ in $\mathcal{M}_3$ satisfy assumptions RC and ER for all $x \in \mathscr{S}_X$. In addition, for every $i = 1,\cdots,I$ and every $x_i \in \mathscr{S}_{X_i}$, let $\overline{\mathcal{R}}_i(x_i)$ has rank $2^{I-1}-2$ and there exist $\alpha, \alpha' \in \mathscr{S}_{\alpha(X)|X_i=x_i}$ such that $\alpha_i \neq \alpha_i'$ and $(\alpha_i, \alpha'_{-i}) \in \mathscr{S}_{\alpha(X)}$. Then S and $\widetilde{S}$ are observationally equivalent if and only if (a) $C_U = \tilde{C}_U$ on $\mathscr{S}_{\alpha(X)}$; (b) there exist $\beta \in \mathbb{R}^I$ and $\gamma \in \mathbb{R}^I_+$, such that for all $i, a_{-i} \in \mathcal{A}_{-i}$ and $x_i \in \mathscr{S}_{X_i}$,*

$\tilde{\pi}_i(a_{-i}, x_i) = \beta_i + \gamma_i \times \pi_i(a_{-i}, x_i)$ *and* $\tilde{F}_{U_i}(\cdot) = F_{U_i}(\frac{\cdot - \beta_i}{\gamma_i})$ *on the support* $\mathscr{S}_{\beta_i + \gamma_i \times u_i^*(X)}$.

Proof. The if part is straightforward, its proof is therefore omitted.

Now we show the only if part. Condition (a) follows immediately from condition (ii) of Lemma 4, so it suffices to show condition (b). First, we construct an observationally equivalent structure $\left[\pi_1^e(\cdot,\cdot), \cdots, \pi_I^e(\cdot,\cdot); F_U^e(\cdot)\right]$ of structure S, such that $F_U^{e-1}(\tau_{i1}) = 0$ and $F_U^{e-1}(\tau_{i2}) = 1$ for $\tau_{i1} < \tau_{i2} \in \mathscr{S}_{\alpha_i(X)|X_i = x_i}$ and some $x_i \in \mathscr{S}_{X_i}$. Let $C^e_{U_{i_1}, \cdots, U_{i_p}} = C_{U_{i_1}, \cdots, U_{i_p}}$ for all $p \geq 2$, $F^e_{U_i}(t) = F_{U_i}\left([t - \beta_i^e]/\gamma_i^e\right)$ and $\pi_i^e = \beta_i^e + \gamma_i^e \pi_i$, where

$$\beta_i^e = \frac{-F_{U_i}^{-1}(\tau_{i1})}{F_{U_i}^{-1}(\tau_{i2}) - F_{U_i}^{-1}(\tau_{i1})}, \quad \gamma_i^e = \frac{1}{F_{U_i}^{-1}(\tau_{i2}) - F_{U_i}^{-1}(\tau_{i1})} > 0.$$

Then, by Lemma 4, it is straightforward that $\left[\pi_1^e(\cdot,\cdot), \cdots, \pi_I^e(\cdot,\cdot); F_U^e(\cdot)\right]$ and structure S are observationally equivalent to each other . Moreover, $F^e_{U_i}(0) = \tau_{i1}$ and $F^e_{U_i}(1) = \tau_{i2}$.

Similarly, let $C^{\tilde{e}}_{U_{i_1}, \cdots, U_{i_p}} = \tilde{C}_{U_{i_1}, \cdots, U_{i_p}}$ for all $p \geq 2$, $F^{\tilde{e}}_{U_i}(t) = \tilde{F}_{U_i}\left([t - \beta_i^{\tilde{e}}]/\gamma_i^{\tilde{e}}\right)$ and $\pi_i^{\tilde{e}} = \beta_i^{\tilde{e}} + \gamma_i^{\tilde{e}} \tilde{\pi}_i$, where

$$\beta_i^{\tilde{e}} = \frac{-\tilde{F}_{U_i}^{-1}(\tau_{i1})}{\tilde{F}_{U_i}^{-1}(\tau_{i2}) - \tilde{F}_{U_i}^{-1}(\tau_{i1})}, \quad \gamma_i^{\tilde{e}} = \frac{1}{\tilde{F}_{U_i}^{-1}(\tau_{i2}) - \tilde{F}_{U_i}^{-1}(\tau_{i1})} > 0.$$

Then $\left[\pi_1^{\tilde{e}}(\cdot,\cdot), \cdots, \pi_I^{\tilde{e}}(\cdot,\cdot); F_U^{\tilde{e}}(\cdot)\right]$ is an observationally equivalent structure of structure $\tilde{S}$ such that $F^{\tilde{e}}{}_U^{-1}(\tau_{i1}) = 0$ and $F^{\tilde{e}}{}_U^{-1}(\tau_{i2}) = 1$. Thus, $\left[\pi_1^{\tilde{e}}(\cdot,\cdot), \cdots, \pi_I^{\tilde{e}}(\cdot,\cdot); F_U^{\tilde{e}}(\cdot)\right]$ is also observationally equivalent to structure S because of observational equivalence between structures S and $\tilde{S}$.

There is, however, only one observationally equivalent structure S^e of structure S satisfying $F_{U_i}^{e-1}(\tau_{i1}) = 0$ and $F_{U_i}^{e-1}(\tau_{i2}) = 1$. The reason is given as follows: By Proposition 7, $\pi_i^e(\cdot, x_i)$ is identified up to location and scale; And for some $x_{-i}, x'_{-i} \in \mathscr{S}_{X_{-i}|X_i = x_i}$, there is $\alpha_i(x) = \tau_{i1}$ and $\alpha_i(x') = \tau_{i2}$. Then

$$\sum_{a_{-i}} \pi_i^e(a_{-i}, x_i) \times \sigma^*_{-i}(a_{-i}|X = x, U_i = u_i^*(x)) = F_{U_i}^{e-1}(\tau_{i1}) = 0,$$
$$\sum_{a_{-i}} \pi_i^e(a_{-i}, x_i) \times \sigma^*_{-i}(a_{-i}|X = x', U_i = u_i^*(x')) = F_{U_i}^{e-1}(\tau_{i2}) = 1,$$

from which we obtain a unique location and scale for $\pi_i^e(\cdot, x_i)$.

Consequently, $\pi_i^e(a_{-i}, x_i) = \beta_i^e + \gamma_i^e \pi_i(a_{-i}, x_i) = \beta_i^{\tilde{e}} + \gamma_i^{\tilde{e}} \tilde{\pi}_i(a_{-i}, x_i) = \pi_i^{\tilde{e}}(a_{-i}, x_i)$ for all a_{-i} and x_i, from which we have

$$\tilde{\pi}_i(\cdot, x_i) = \beta_i + \gamma_i \pi_i(\cdot, x_i).$$

where $\beta_i = \frac{\beta_i^e - \beta_i^{\tilde{e}}}{\gamma_i^{\tilde{e}}}, \gamma_i = \frac{\gamma_i^e}{\gamma_i^{\tilde{e}}}$. Besides, in equilibrium, we can obtain the following conditions for structure S and $\tilde{S}$, respectively.

$$\sum_{a_{-i}} \pi_i(a_{-i}, x_i) \cdot \sigma_{-i}^*(a_{-i}|x, u_i^*(x)) = u_i^*(x)$$
$$\sum_{a_{-i}} \tilde{\pi}_i(a_{-i}, x_i) \cdot \tilde{\sigma}_{-i}^*(a_{-i}|x, \tilde{u}_i^*(x)) = \tilde{u}_i^*(x)$$

which imply $\tilde{u}_i^*(x) = \beta_i + \gamma_i \cdot u_i^*(x)$ since we have $\tilde{\pi}_i(\cdot, x_i) = \beta_i + \gamma_i \pi_i(\cdot, x_i)$ obtained earlier and $\sigma_{-i}^*(\cdot|x, u_i^*(x)) = \tilde{\sigma}_{-i}^*(\cdot|x, \tilde{u}_i^*(x))$ from condition (a) and Lemma 7. We then get $\tilde{F}_{U_i}(u_i) = F_{U_i}(\frac{u_i - \beta_i}{\gamma_i})$ for every $u_i \in \mathscr{S}_{\beta_i + \gamma_i \cdot u_i^*(X)}$ from the observational equivalence condition $F_{U_i}(u_i^*(x)) = \tilde{F}_{U_i}(\tilde{u}_i^*(x))$, which completes the proof. □

A.4 Proofs of Results in section 1.5

A.4.1 Proof of Lemma 9

Proof. First, we construct a structure $\tilde{S} \in \mathcal{M}_3$ (which implies that $\tilde{S} \in \mathcal{M}_3'$) such that (1) $\tilde{S}$ has the marginal quantile functions $\left(\tilde{Q}_{U_1}, \cdots, \tilde{Q}_{U_I}\right)$; (2) $\tilde{C}(\cdot) = C(\cdot)$ on $[0,1]^I$; (3) for any $x \in \mathscr{S}_X$, i, and $a_{-i} \in \mathcal{A}_{-i}$, let $\tilde{\pi}_i(a_{-i}, x) = \tilde{Q}_{U_i}(\mathbb{E}(Y_i|X = x))$. By construction, it is straightforward that Assumptions R, M, S and E are satisfied.

Now it suffices to verify the observational equivalence between the two structures. Fix $x \in \mathscr{S}_X$. Note that in the structure $\tilde{S}$ there is no strategic effects, then the equilibrium is: $\mathbf{1}\left\{u_i \leq \tilde{\pi}_i(a_{-i}^0, x)\right\}$ for $i = 1, \cdots, I$. Then similar to the proof for Lemma 7, we only verify the observational equivalence for action profile $(1, \cdots, 1)$ and other action profiles follows similarly:

$$\tilde{\mathbb{P}}(Y_1 = 1; \cdots; Y_I = 1 | X = x) = \tilde{C}\left(\mathbb{E}(Y|X = x)\right)$$
$$= C\left(\mathbb{E}(Y|X = x)\right) = \mathbb{P}(Y_1 = 1; \cdots; Y_I = 1 | X = x). \quad \square$$

A.4.2 Proof of Proposition 9

Proof. It is straightforward that $(\pi, Q_U, C) \in \Theta_I(Q_U)$. For sharpness, it suffices to show that for any $(\pi', C') \in \Theta_I(Q'_U)$, then $S' \equiv (\pi', Q'_U, C') \in \mathcal{M}_3$ and the two structures S and S' are observationally equivalent.

By construction, it is straightforward that Assumptions R, M, S and E are satisfied. Therefore, $S' \in \mathcal{M}_3$. Fix $X = x \in \mathscr{S}_X$. For observationally equivalence, it suffices to show that $\delta^* = \left(\mathbf{1}\left\{u_1 \leq \tilde{Q}_{U_1}(\alpha_1(x))\right\}, \cdots, \mathbf{1}\left\{u_I \leq \tilde{Q}_{U_I}(\alpha_I(x))\right\}\right)$ is a BNE solution to the game. By the proof for Lemma 7, $\tilde{\mathbb{P}}_{\delta^*}\left\{Y_{-i} = a_{-i} | X = x, U_i = \tilde{Q}_{U_i}(\alpha_i(x))\right\} = \sigma^*_{-i}(a_{-i}|x, u^*_i(x))$. Then $\mathbf{1}\left\{u_i \leq \tilde{Q}_{U_i}(\alpha_i(x))\right\}$ is a best response to δ^*_{-i} under assumption M and the condition (a) in Proposition 9. Thus δ^* is a BNE. $\square$

Appendix B

Appendix to Chapter 2

B.1 Proofs of Identification Results

B.1.1 Proof of Lemma 12

Proof. Let $\nu_k(x) = \mathbb{E}(Y_k|X = x)$ for $k \in \mathcal{I}$. By definition and Assumption C,

$$\varphi_{ij}(x) = \mathbb{P}\left(U_j \leq u_j^*(x)|U_i = u_i^*(x), X = x\right)$$
$$= \mathbb{P}\left(U_j \leq u_j^*(x)|U_i = u_i^*(x), W = w\right).$$

Then, it follows from Darsow, Nguyen, and Olsen (1992) that

$$\mathbb{P}(U_j \leq u_j^*(x)|U_i = u_i^*(x), W = w) = \frac{\partial C_{ij}(v_i, v_j; w)}{\partial v_i}\Big|_{v_i = F_{U_i|W}(u_i^*(x)|w), v_j = F_{U_j|W}(u_j^*(x)|w)}$$
$$= \frac{\partial C_{ij}(v_i, v_j; w)}{\partial v_i}\Big|_{v_i = \mathbb{E}(Y_i|X=x), v_j = \mathbb{E}(Y_j|X=x)}.$$

which is identified by the fact that $C_{ij}(\cdot; w)$ can be identified on the support for all $(v_i, v_j) \in \mathscr{S}_{V_i V_j|W=w}$.

□

B.2 Proofs of Statistical Properties

B.2.1 Proof of Proposition 11

Proof. Our proofs follow Guerre, Perrigne, and Vuong (2000). For the notation brevity, here we ignore the difference cased by leaving–one–observation–out in the estimator $\widehat{\varphi}_i$. Moreover, let

$$a_{iN}(x_n) = \frac{1}{Nh_P^3} \sum_{\ell \neq n}^{N} \partial K_\varphi \left(\frac{(X_{1\ell} - x_{1n})'\beta_1}{h_\varphi}, \frac{(X_{2\ell} - x_{2n})'\beta_2}{h_\varphi} \right) \Big/ \partial t_i$$

be the (infeasible) nonparametric estimator of the derivative

$$\frac{\partial}{\partial t_1} \left\{ \mathbb{E}(Y_1 Y_2 | (X_1'\beta_1, X_2'\beta_2) = t) \times f_{X_1'\beta_1, X_2'\beta_2}(t) \right\} \Big|_{t = x_n}$$

using the true parameters β. Similarly, we define b_{iN}, c_{jiN}, q_{iN}, Q_N and f_{XN}. By plugging these infeasible estimators, we define our infeasible estimator $A_{iN}(x_n)$ and $A_N(x_n)$. Further, let $\varphi_{iN}(X_n) = A_{iN}(X_n) / A_N(X_n)$ and

$$A_1(x) = f^4_{X_1'\beta_1, X_2'\beta_2} \left[\frac{\partial M(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} - \frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \frac{\partial M(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} \right],$$

$$A_2(x) = f^4_{X_1'\beta_1, X_2'\beta_2} \left[\frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \frac{\partial M(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} - \frac{\partial M(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} \right],$$

$$A(x) = f^4_{X_1'\beta_1, X_2'\beta_2} \left[\frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} - \frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} \right].$$

Note that

$$\hat{\varphi}_i(x) = \frac{\hat{A}_i(x)}{\hat{A}(x)} = \frac{\hat{A}_i(x)/A(x)}{\hat{A}(x)/A(x)}$$

$$= \frac{A_{iN}(x)/A(x) + \left[\hat{A}_i(x) - A_{iN}(x)\right]/A(x)}{1 + \left[A_N(x)/A(x) - 1\right] + \left[\hat{A}(x) - A_N(x)\right]/A(x)}.$$

Hence it suffices to show

$$\sup_{\|x\|\le\kappa_N}\|A_{iN}(x)/A(x)-A_i(x)/A(x)\| = O_p\left(\eta_N^{-1}\left(\frac{\ln N}{N}\right)^{R/(2R+4)}\right), \tag{B.1}$$

$$\sup_{\|x\|\le\kappa_N}\|A_N(x)/A(x)-1\| = O_p\left(\eta_N^{-1}\left(\frac{\ln N}{N}\right)^{R/(2R+4)}\right), \tag{B.2}$$

$$\sup_{\|x\|\le\kappa_N}\left\|\hat{A}_i(x)/A(x)-A_{iN}(x)/A(x)\right\| = o_p\left(\eta_N^{-1}\left(\frac{\ln N}{N}\right)^{R/(2R+4)}\right), \tag{B.3}$$

$$\sup_{\|x\|\le\kappa_N}\left\|\hat{A}(x)/A(x)-A_N(x)/A(x)\right\| = o_p\left(\eta_N^{-1}\left(\frac{\ln N}{N}\right)^{R/(2R+4)}\right). \tag{B.4}$$

Equations (B.1) and (B.2) are satisfied under Lemmas 22 and 23. We illustrate the argument for eq. (B.3), and then eq. (B.4) is proved analogously. By Lemma 22, it suffice to show

$$\sup_{\|x\|\le\kappa_N}\left\|\hat{A}_i(x)-A_{iN}(x)\right\| = o_p\left(\left(\frac{\ln N}{N}\right)^{R/(2R+4)}\right).$$

We will show that $\sup_{\|x\|\le\kappa_N}|a_{iN}(x)-\hat{a}_i(x)|$, $\sup_{\|x\|\le\kappa_N}\left|b_{iN}(x)-\hat{b}_{ij}(x)\right|$, $\sup_{\|x\|\le\kappa_N}|c_{iN}(x) -\hat{c}_i(x)|$, $\sup_{\|x\|\le\kappa_N}|q_{iN}(x)-\hat{q}_i(x)|$, $\sup_{\|x\|\le\kappa_N}\left|Q_N(x)-\hat{Q}(x)\right|$ and $\sup_{\|x\|\le\kappa_N}|f_{XN}(x)- \hat{f}_X(x)|$ all converge to zero at the $\sqrt{N}$ rate. Since the arguments for all other terms are quite similar to or simpler than those for $\sup_{\|x\|\le\kappa_N}|c_{iN}(x)-\hat{c}_i(x)|$, here we only provide a detailed proof for the latter.

Because $\tilde{\beta}_i=\beta_i+O_p(N^{-1/2})$, then for any fixed $\epsilon>0$, the following inequality holds with probability approaching to 1,

$$\begin{aligned}
|c_{iN}(x_n)-\hat{c}_i(x_n)| = &\left|\frac{1}{Nh_\varphi^3}\sum_{\ell\neq n}^{N}Y_{1\ell}Y_{2\ell}\times\left\{\partial K_\varphi\left(\frac{(X_{1\ell}-x_{1n})'\tilde{\beta}_1}{h_\varphi},\frac{(X_{2\ell}-x_{2n})'\tilde{\beta}_2}{h_\varphi}\right)\Big/\partial t_i\right.\right.\\
&\left.\left.-\,\partial K_\varphi\left(\frac{(X_{1\ell}-x_{1n})'\beta_1}{h_\varphi},\frac{(X_{2\ell}-x_{2n})'\beta_2}{h_\varphi}\right)\Big/\partial t_2\right\}\right|\\
\le \sup_{\|\beta^\dagger-\beta\|\le\epsilon}&\left|\frac{1}{Nh_\varphi^4}\sum_{\ell\neq n}^{N}Y_{1\ell}Y_{2\ell}\times\partial^2K_\varphi\left(\frac{(X_{1\ell}-x_{1n})'\beta_1^\dagger}{h_\varphi},\frac{(X_{2\ell}-x_{2n})'\beta_2^\dagger}{h_\varphi}\right)\Big/\partial t_i\partial t_1\times(X_{1\ell}-x_{1n})'(\tilde{\beta}_1-\beta_1)\right|\\
+\sup_{\|\beta^\dagger-\beta\|\le\epsilon}&\left|\frac{1}{Nh_P^4}\sum_{\ell\neq n}^{N}Y_{1\ell}Y_{2\ell}\times\partial^2K_\varphi\left(\frac{(X_{1\ell}-x_{1n})'\beta_1^\dagger}{h_\varphi},\frac{(X_{2\ell}-x_{2n})'\beta_2^\dagger}{h_\varphi}\right)\Big/\partial t_i\partial t_2\times(X_{2\ell}-x_{2n})'(\tilde{\beta}_2-\beta_2)\right|.
\end{aligned}$$

By lemma 21, we have

$$\sup_x |a_{iN}(x) - \hat{a}_i(x)| \leq \left\|\tilde{\beta}_1 - \beta_1\right\| \times O_p(1) + \left\|\tilde{\beta}_2 - \beta_2\right\| \times O_p(1) = O_p(N^{-1/2}). \quad \square$$

Lemma 21. *Suppose that Assumptions G, H, J, L and M hold. Thus,*

$$\sup_{\|x\|\leq\kappa_N} \sup_{\|b-\beta\|\leq\delta} \left\| \frac{1}{Nh_P^4} \sum_{\ell=1}^{N} Y_{1\ell}Y_{2\ell} \times \right.$$
$$\left. \partial^2 K_\varphi\left(\frac{(X_{1\ell}-x_1)'b_1}{h_\varphi}, \frac{(X_{2\ell}-x_2)'b_2}{h_\varphi}\right) \Big/ \partial t_i \partial t_j \times (X_{j\ell} - x_j)' \right\| = O_p(1).$$

Proof. Fix i, j. Let $S_\ell(x,b) = \frac{1}{h_\varphi^4} Y_{1\ell}Y_{2\ell} \times \partial^2 K_\varphi\left(\frac{(X_{1\ell}-x_1)'b_1}{h_\varphi}, \frac{(X_{2\ell}-x_2)'b_2}{h_\varphi}\right) \Big/ \partial t_i \partial t_j \times (X_{j\ell} - x_j)'$ as a random vector indexed by x and b. Let further $\psi_{x,b}(t) = \mathbb{E}[\mathbb{E}(Y_1 Y_2|X) \times (X_j - x_j')| (X_1'b_1, X_2'b_2) = t]$ and $\phi_{x,b}(t) = \psi_{x,b}(t) \times f_{X_1'b_1, X_2'b_2}(t)$.[1] Then we have

$$\begin{aligned}
\sup_{\|x\|\leq\kappa_N} \sup_{\|b-\beta\|\leq\delta} \left\| \frac{1}{N}\sum_{\ell=1}^N S_\ell(x,b) \right\| &\leq \sup_{\|x\|\leq\kappa_N} \sup_{\|b-\beta\|\leq\delta} \left\| \frac{1}{N}\sum_{\ell=1}^N S_\ell(x,b) - \mathbb{E}S_\ell(x,b) \right\| \\
&+ \sup_{\|x\|\leq\kappa_N} \sup_{\|b-\beta\|\leq\delta} \left\| \frac{1}{N}\sum_{\ell=1}^N \mathbb{E}S_\ell(x,b) - \partial^2\phi_{x,b}(x_1'b_1, x_2'b_2)/\partial t_1 \partial t_2 \right\| \\
&+ \sup_{\|x\|\leq\kappa_N} \sup_{\|b-\beta\|\leq\delta} \left\| \partial^2\phi_{x,b}(x_1'b_1, x_2'b_2)/\partial t_1 \partial t_2 \right\|.
\end{aligned}$$

By Theorem 1 in Andrews (1992), the first term of the RHS is $o_p(1)$; and by Assumptions J and L, the last term is $O_p(1)$.

Moreover, for the second term, we have

$$\begin{aligned}
\mathbb{E}S_\ell(x,b) &= \frac{1}{h_P^4}\mathbb{E}\left[\psi(X_{1\ell}'b_1, X_{2\ell}'b_2) \times \partial^2 K_\varphi\left(\frac{(X_{1\ell}-x_1)'\beta_1}{h_\varphi}, \frac{(X_{2\ell}-x_2)'\beta_2}{h_\varphi}\right)\right] \\
&= \frac{1}{h_\varphi^2}\int_{\mathbb{R}^2} \phi_{x,b}(x_1'b_1 - h_\varphi u_1, x_2'b_2 - h_\varphi u_2) \times \partial^2 K_\varphi(u)/\partial t_1 \partial t_2 du \\
&= \int_{\mathbb{R}^2} \partial^2 \phi_{x,b}(x_1'b_1 - h_\varphi u_1, x_2'b_2 - h_\varphi u_2)/\partial t_1 \partial t_2 \times K_\varphi(u)\, du.
\end{aligned}$$

[1] Note that we suppress a subscript j in the notation for $\psi_{x,b}$ and $\phi_{x,b}$.

By Taylor expansion of order 2 with integral remainder,

$$\partial^2\phi_{x,b}(x_1'b_1 - h_\varphi u_1, x_2'b_2 - h_\varphi u_2)/\partial t_1\partial t_2 = \partial^2\phi_{x,b}(x_1'b_1, x_2'b_2)/\partial t_1\partial t_2$$
$$- h_\varphi \sum_{k=1}^{2} \frac{\partial^3\phi_{x,b}(x_1'b_1, x_2'b_2)}{\partial t_1\partial t_2\partial t_k} u_k + \frac{h_\varphi^2}{2}\int_0^1 (1-s) \sum_{k_1=1}^{2}\sum_{k_2=1}^{2} \frac{\partial^4\phi_{x,b}(x_1'b_1 - th_\varphi u_1, x_2'b_2 - th_\varphi u_2)}{\partial t_1\partial t_2\partial t_{k_1}\partial t_{k_2}} u_{k_1}u_{k_2} ds.$$

By Assumption J, $\phi_{x,b}$ has bounded fourth order derivative. Thus, uniformly over x and b

$$\left\|\mathbb{E}S_\ell(x,b) - \partial^2\phi_{x,b}(x_1'b_1, x_2'b_2)/\partial t_1\partial t_2\right\|$$
$$= \frac{h_\varphi^2}{2}\left\|\int_{\mathbb{R}^2}\int_0^1 (1-t) \sum_{k_1=1}^{2}\sum_{k_2=1}^{2} \frac{\partial^4\phi_{x,b}(x_1'b_1 - th_\varphi u_1, x_2'b_2 - th_\varphi u_2)}{\partial t_1\partial t_2\partial t_{k_1}\partial t_{k_2}} u_{k_1}u_{k_2} K_\varphi(u)\, dtdu\right\| \leq C \times \frac{h_\varphi^2}{2},$$

for some $C \in \mathbb{R}_+$. Thus the second term is $o_p(1)$. □

Lemma 22. *Suppose that Assumptions I and K hold. Then*

$$\inf_{\|x\|\leq\kappa_N} \|A(x)\| = O(\eta_N).$$

Proof. By Assumptions I and K

$$|A(x)| = f^4_{X_1'\beta_1, X_2'\beta_2}(x_1'\beta_1, x_2'\beta_2)$$
$$\times \left|\frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_1} \times \frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} - \frac{\partial m_1(x_1'\beta_1, x_2'\beta_2)}{\partial t_2} \times \frac{\partial m_2(x_1'\beta_1, x_2'\beta_2)}{\partial t_1}\right| \geq c_0\eta_N. \quad \square$$

Lemma 23. *Suppose that assumptions in Proposition 11 hold. Then*

$$\sup_{\|x\|\leq\kappa_N} \|A_{iN}(x) - A_i(x)\| = O_p\left(\left(\frac{\ln N}{N}\right)^{R/(2R+4)}\right),$$
$$\sup_{\|x\|\leq\kappa_N} \|A_N(x) - A(x)\| = O_p\left(\left(\frac{\ln N}{N}\right)^{R/(2R+4)}\right),$$

Proof. We only illustrate the argument for $\sup_{\|x\|\leq\kappa_N} \|A_{1N}(x) - A_1(x)\|$; other results can be established analogously. It suffices to show that

$$\sup_{\|x\|\leq\kappa_N}\left\|c_{1N}(x)f_{XN}(x)-a_{1N}(x)Q_N(x)-f^2_{X_1'\beta_1,X_2'\beta_2}\times\frac{\partial M(x_1'\beta_1,x_2'\beta_2)}{\partial t_1}\right\|=O_p\left(\left(\frac{\ln N}{N}\right)^{\frac{R}{2R+4}}\right),$$

$$\sup_{\|x\|\leq\kappa_N}\left\|b_{22N}(x)f_{XN}(x)-a_{2N}(x)q_{2N}(x)-f^2_{X_1'\beta_1,X_2'\beta_2}\times\frac{\partial m_2(x_1'\beta_1,x_2'\beta_2)}{\partial t_2}\right\|=O_p\left(\left(\frac{\ln N}{N}\right)^{\frac{R}{2R+4}}\right),$$

$$\sup_{\|x\|\leq\kappa_N}\left\|c_{2N}(x)f_{XN}(x)-a_{2N}(x)Q_N(x)-f^2_{X_1'\beta_1,X_2'\beta_2}\times\frac{\partial M(x_1'\beta_1,x_2'\beta_2)}{\partial t_1}\right\|=O_p\left(\left(\frac{\ln N}{N}\right)^{\frac{R}{2R+4}}\right),$$

$$\sup_{\|x\|\leq\kappa_N}\left\|b_{21N}(x)f_{XN}(x)-a_{1N}(x)q_{2N}(x)-f^2_{X_1'\beta_1,X_2'\beta_2}\times\frac{\partial m_2(x_1'\beta_1,x_2'\beta_2)}{\partial t_2}\right\|=O_p\left(\left(\frac{\ln N}{N}\right)^{\frac{R}{2R+4}}\right).$$

Again, we provide a detailed proof only for the first term due to the similarity. Let $Q(t)=M(t)\times f_{X_1'\beta_1,X_2'\beta_2}(t)$. Because

$$\frac{\partial Q(t)}{\partial t_1}=\frac{\partial M(t)}{\partial t_1}\times f_{X_1'\beta_1,X_2'\beta_2}(t)+\frac{\partial f_{X_1'\beta_1,X_2'\beta_2}(t)}{\partial t_1}\times M(t),$$

then

$$f^2_{X_1'\beta_1,X_2'\beta_2}\times\frac{\partial M(x_1'\beta_1,x_2'\beta_2)}{\partial t_1}=\frac{\partial Q(x_1'\beta_1,x_2'\beta_2)}{\partial t_1}\times f_{X_1'\beta_1,X_2'\beta_2}(x_1'\beta_1,x_2'\beta_2)-\frac{\partial f_{X_1'\beta_1,X_2'\beta_2}(x_1'\beta_1,x_2'\beta_2)}{\partial t_1}\times Q(x_1'\beta_1,x_2'\beta_2).$$

Thus, it suffices to show that

$$\sup_{\|x\|\leq\kappa_N}\left\|c_{1N}(x)-\frac{\partial Q(x_1'\beta_1,x_2'\beta_2)}{\partial t_1}\right\|=O_p\left(\left(\frac{\ln N}{N}\right)^{\frac{R}{2R+4}}\right),\tag{B.5}$$

$$\sup_{\|x\|\leq\kappa_N}\left\|f_{XN}(x)-f_{X_1'\beta_1,X_2'\beta_2}(x_1'\beta_1,x_2'\beta_2)\right\|=O_p\left(\left(\frac{\ln N}{N}\right)^{\frac{R}{2R+4}}\right),\tag{B.6}$$

$$\sup_{\|x\|\leq\kappa_N}\left\|a_{1N}(x)-\frac{\partial f_{X_1'\beta_1,X_2'\beta_2}(x_1'\beta_1,x_2'\beta_2)}{\partial t_1}\right\|=O_p\left(\left(\frac{\ln N}{N}\right)^{\frac{R}{2R+4}}\right),\tag{B.7}$$

$$\sup_{\|x\|\leq\kappa_N}\left\|Q_N(x_1'\beta_1,x_2'\beta_2)-Q(x_1'\beta_1,x_2'\beta_2)\right\|=O_p\left(\left(\frac{\ln N}{N}\right)^{\frac{R}{2R+4}}\right).\tag{B.8}$$

Equations (B.6) and (B.7) directly follows Hansen (2008), Theorem 6, and by following its proof, eqs. (B.5) and (B.8) also obtain, which is straightforward, and hence omitted here.

□

B.3 Proof of Theorem 2

Our proof follows Klein and Spady (1993). Throughout appendix B.3, we introduce some notation, which is consistent with Klein and Spady (1993).

Let $v_i(X; a_i, b_i) \equiv X_i' b_i + a_i \varphi_i(X)$ and $\bar{v}_i(X; a_i, b_i) \equiv X_i' b_i + a_i \widehat{\varphi}_i(X)$. Through, we suppress the subscript for player i in v_i and $\hat{v}_i$, i.e., we use $v(x; a_i, b_i)$ and $\bar{v}(x; a_i, b_i)$ to denote $v_i(x; a_i, b_i)$ and $\bar{v}_i(x; a_i, b_i)$, respectively. Similarly, we will suppress subscript i in the following discussion. Let $v_n(a_i, b_i) \equiv v(X_n; a_i, b_i)$ and $\bar{v}_n(a_i, b_i) \equiv \bar{v}(X_n; a_i, b_i)$. Similarly, by replacing $\hat{\varphi}_i$ with the underlying belief φ_i, we can define $\bar{\tau}_n$, $\bar{\tau}_{0n}$, $\bar{\tau}_{1n}$, $\bar{\delta}_n$, $\bar{\delta}_{0n}$, $\bar{\delta}_{1n}$. Let $g_v(v_n; a_i, b_i)$ be the density of $v_n(a_i, b_i)$. Moreover, for $d = 0, 1$ let $g_{dv}(v_n; a_i, b_i) \equiv \mathbb{P}(Y_i = d | v(a_i, b_i) = v_n(a_i, b_i)) g_v(v_n; a_i, b_i)$ and for $d = 0, 1$

$$\bar{g}_{dv}(v_n; a_i, b_i) \equiv \sum_{\ell \neq n}^{N} \frac{\mathbf{1}(Y_{i\ell} = d)}{h_P} K\left(\frac{v_\ell - v_n}{h_P}\right) \Big/ (N-1),$$
$$\hat{g}_{dv}(v_n; a_i, b_i) \equiv \sum_{\ell \neq n}^{N} \frac{\mathbf{1}(Y_{i\ell} = d)}{h_P} K\left(\frac{\bar{v}_\ell - \bar{v}_n}{h_P}\right) \Big/ (N-1),$$

Let further

$$\bar{L}_N(a_i, b_i; \bar{\tau}) \equiv \sum_{n=1}^{N} (\bar{\tau}_n/2) \left\{ Y_{in} \ln\left[\bar{P}_i(v_n; a_i, b_i)^2\right] + (1 - Y_{in}) \ln\left[1 - \bar{P}_i(v_n; a_i, b_i)\right]^2 \right\} \Big/ N$$

and

$$\bar{P}_i(v_n; a_i, b_i) \equiv \left[\bar{g}_{i1v}(v_n; a_i, b_i) + \bar{\delta}_{1n}(v_n; a_i, b_i)\right] \Big/ \left[\bar{g}_{iv}(v_n; a_i, b_i) + \bar{\delta}_n(v_n; a_i, b_i)\right],$$
$$P(v_n; a_i, b_i) \equiv g_{1v}(v_n; a_i, b_i) / g_v(v_n; a_i, b_i).$$

Also, we define the r–th order derivative of any function g with respect to z by

$$D_z^r[g] = \begin{cases} g, & r = 0, \\ \partial^r g / (\partial z)^r, & r = 1, 2, \cdots. \end{cases}$$

Further, we use $\|\cdot\|$ to denote the Euclidean norm.

Let

$$\hat{G}(\alpha_i, \beta_i) \equiv \left[\partial \hat{L}_i / \partial(a_i, b_i)\right] \Big|_{(a_i,b_i)=(\alpha_i,\beta_i)} = \sum_{n=1}^{N} \hat{\tau}_n \hat{r}_n \hat{w}_n / N,$$

where

$$\hat{r}_n \equiv \left[Y_{in} - \hat{P}_i(X_n; \alpha_i, \beta_i)\right] \Big/ \hat{c}_n, \ \ \hat{c}_n \equiv \hat{g}_v(v_n; \alpha_i, \beta_i) \left[\hat{P}_i(X_n; \alpha_i, \beta_i)(1 - \hat{P}_i(X_n; \alpha_i, \beta_i))\right],$$
$$\hat{w}_n \equiv \hat{g}_v(v_n; \alpha_i, \beta_i) \left[\partial \hat{P}_i(X_n; \alpha_i, \beta_i) / \partial(a_i, b_i)\right].$$

Let further

$$G_N(\alpha_i, \beta_i) \equiv [\partial L_N / \partial(a_i, b_i)] \Big|_{(a_i,b_i)=(\alpha_i,\beta_i)} = \sum_{n=1}^{N} \tau_n r_n w_n / N,$$

where

$$r_n \equiv [Y_{in} - P_i(v_n; \alpha_i, \beta_i)] / c_n, \ \ c_n \equiv [g_v(v_n; \alpha_i, \beta_i) + \delta_n(v_n; \alpha_i, \beta_i)]$$
$$\times [P_i(v_n; \alpha_i, \beta_i)(1 - P_i(v_n; \alpha_i, \beta_i))], \quad w_n \equiv g_v(v_n; \alpha_i, \beta_i) \left[\partial P(v_n; \alpha_i, \beta_i) / \partial(a_i, b_i)\right]$$

B.3.1 Proof for Theorem 2

Proof. The consistency simply follows the uniform convergence of $\widehat{\varphi}_i$ to φ_i and the proof for Theorem 3 in Klein and Spady (1993), which is omitted here.

For asymptotic normality, it suffices to show $N^{1/2}\widehat{G}(\alpha_i, \beta_i) - N^{1/2} G_N(\alpha_i, \beta_i) = o_p(1)$, and all the left simply follows Klein and Spady (1993), Theorem 4.

$$\begin{aligned} N^{1/2}\hat{G}(\alpha_i, \beta_i) &- N^{1/2} G_N(\alpha_i, \beta_i) \\ &= N^{-1/2} \sum_{n=1}^{N} \tau_n (\hat{r}_n \hat{w}_n - r_n w_n) + N^{-1/2} \sum_{n=1}^{N} (\hat{\tau}_n - \tau_n) r_n w_n \\ &\quad + N^{-1/2} \sum_{n=1}^{N} (\hat{\tau}_n - \tau_n)(\hat{r}_n \hat{w}_n - r_n w_n). \qquad \text{(B.9)} \end{aligned}$$

For the first term in equation (B.9), denoted as **A**,

$$\mathbf{A} = N^{-1/2}\sum_{n=1}^{N}\tau_n(\hat{r}_n - r_n)w_n + N^{-1/2}\sum_{n=1}^{N}\tau_n(\hat{r}_n - r_n)(\hat{w}_n - w_n) + N^{-1/2}\sum_{n=1}^{N}\tau_n r_n(\hat{w}_n - w_n). \tag{B.10}$$

For the first term, similar to the arguments for $\mathbf{A}_1$ in Lemma 6 of Klein and Spady (1993), it is $o_p(1)$. For the second term, because

$$\left| N^{-1/2}\sum_{n=1}^{N}\tau_n(\hat{r}_n - r_n)(\hat{w}_n - w_n)\right| \le N^{1/2}\sup|\tau_n(\hat{r}_n - r_n)|\sup|\tau_n(\hat{w}_n - w_n)|$$

By definition

$$\hat{r}_n - r_n = \frac{Y_{in}}{\hat{g}_{1vn}} - \frac{Y_{in}}{g_{1vn}} + \frac{1-Y_{in}}{\hat{g}_{0vn}} - \frac{1-Y_{in}}{g_{0vn}}.$$

By Lemma 24, we have

$$\tau_n\frac{Y_{in}}{\hat{g}_{1vn}} = \tau_n\frac{Y_{in}/g_{1vn}}{\hat{g}_{1vn}/g_{1vn}} = \tau_n\frac{Y_{in}/g_{1vn}}{1+(\hat{g}_{1vn}-g_{1vn})/g_{1vn}} = \tau_n\frac{Y_{in}}{g_{1vn}} + \tau_n\frac{O_p\left(\sqrt{\ln N/Nh_P}\vee h^2\right)}{g_{1vn}},$$

then

$$\sup\left|\tau_n\frac{Y_{in}}{\hat{g}_{1vn}} - \tau_n\frac{Y_{in}}{g_{1vn}}\right| = O_p\left(\sqrt{\ln N/Nh_P}\vee h^2\right).$$

Similarly,

$$\sup\left|\tau_n\frac{1-Y_{in}}{\hat{g}_{0vn}} - \tau_n\frac{1-Y_{in}}{g_{0vn}}\right| = O_p\left(\sqrt{\ln N/Nh_P}\vee h^2\right).$$

Then we have $\sup|\tau_n(\hat{r}_n - r_n)| = O_p\left((\ln N/N)^{2/(2p+3)}\right)$. By a similar argument, $\sup|\tau_n(\hat{w}_n - w_n)| = O_p\left(\sqrt{\ln N/Nh_N^3}\vee h^2\right)$. Further, by the condition (ii) in assumption R,

$$N^{1/2}\sup|\tau_n(\hat{r}_n - r_n)|\sup|\tau_n(\hat{w}_n - w_n)| = o_p(1).$$

For the last term in the RHS of equation (B.10), denoted by $\mathbf{A}_3$, we have

$$\mathbb{E}\left(\mathbf{A}_3^2\right) = \sum_{n=1}^{N}\mathbb{E}\left[\tau_n^2 r_n^2(\hat{w}_n - w_n)^2\right]\Big/N + \mathbb{E}\sum_{\ell\neq n} r_n r_\ell \tau_n\tau_\ell(\hat{w}_n - w_n)(\hat{w}_\ell - w_\ell)\Big/N.$$

By lemma 24 and Chung(1974, Thm. 4.5.2), the first term is $o_p(1)$, Note that the second term is more complicated than the corresponding part in Klein and Spady (1993). Recall that, by definition, $\widehat{\varphi}_i(X_n)$ is estimated by leaving out one observation Y_n. Similarly, we define $\overline{\varphi}_i(X_n;\ell)$ by leaving out two observations Y_n and Y_ℓ. Thus we can define $\overline{w}_n$ by replacing $\widehat{\varphi}_i(X_k)$ with $\overline{\varphi}_i(X_k;n)$ for all $k \neq n$ and $\widehat{\varphi}_i(X_n)$ with $\overline{\varphi}_i(X_n;\ell)$ in $\widetilde{w}_n$. Note that $\overline{w}_n$ depends neither on Y_{in} and $Y_{i\ell}$, then by a similar argument as in Klein and Spady (1993), Lemma 6, we have $\mathbb{E}\sum_{\ell\neq n} r_n r_\ell \tau_n \tau_\ell (\overline{w}_n - w_n)(\overline{w}_\ell - w_\ell)\Big/N = o_p(1)$. It should also be noted that $\overline{w}_n - \widetilde{w}_n = O(N^{-1})$ uniformly over x, since $\overline{\varphi}_i(X_n;k) - \widehat{\varphi}_i(X_n) = O_p(N^{-1})$ uniformly. Therefore the second term in the RHS of above equation is also $o_p(1)$.

Turning to the second term in (B.9) above, under a similar argument used to analysis $\mathbf{A}_3$, it is $o_p(1)$. The proof for the last term in equation (B.9) being $o_p(1)$ simply follows the corresponding part of the arguments in Klein and Spady (1993).

□

Lemma 24. *Suppose that assumptions in Theorem 2 hold. Then for $y = 0, 1$,*

$$\sup\left|\widetilde{g}_{yv}(v_n;\alpha_i,\beta_i) - g_{yv}(v_n;\alpha_i,\beta_i)\right| = O_p\left(\sqrt{\ln N/Nh_P} \vee h^2\right)$$
$$\sup\left|D^1_{(a_i,b_i)}\widetilde{g}_{yv}(v_n;\alpha_i,\beta_i) - D^1_{(a_i,b_i)}g_{yv}(v_n;\alpha_i,\beta_i)\right| = O_p\left(\sqrt{\ln N/Nh_N^3} \vee h^2\right).$$

Proof. First, let $\mathbb{P}_N\widetilde{g}_{yv}(v_n;\alpha_i,\beta_i) \equiv \int \frac{\mathbf{1}(Y_i=y)}{h_P} K\left(\frac{\widehat{v}(x;\alpha_i,\beta_i)-\widehat{v}(X;\alpha_i,\beta_i)}{h_P}\right) dF_{XY}$ and $\mathbb{P}_N g_{yv}(v_n;\alpha_i,\beta_i) \equiv \int \frac{\mathbf{1}(Y_i=y)}{h_P} K\left(\frac{v(x;\alpha_i,\beta_i)-v(X;\alpha_i,\beta_i)}{h_P}\right) dF_{XY}$. By triangular inequality,

$$\begin{aligned}
&\sup_x\left|\widetilde{g}_{yv}(v_n;\alpha_i,\beta_i) - \mathbb{P}_N\widetilde{g}_{yv}(v_n;\alpha_i,\beta_i)\right| \\
&\leq \sup_x \sup_{\|\widehat{\varphi}-\varphi\|\downarrow 0}\left|\widetilde{g}_{yv}(v_n;\alpha_i,\beta_i) - \mathbb{P}_N\widetilde{g}_{yv}(v_n;\alpha_i,\beta_i) - \left[\widehat{g}_{yv}(v_n;\alpha_i,\beta_i) - \mathbb{P}_N g_{yv}(v_n;\alpha_i,\beta_i)\right]\right| \\
&\qquad + \sup_x\left|\widehat{g}_{yv}(v_n;\alpha_i,\beta_i) - \mathbb{P}_N g_{yv}(v_n;\alpha_i,\beta_i)\right|,
\end{aligned}$$

where the first term is $o_p(N^{-1/2})$, referred as the stochastic equicontinuity condition, by Theorem 11.16 in Kosorok (2008).

Next, $\sup\left|D^r_{(a_i,b_i)}\widehat{g}_{yv}(v_n;\alpha_i,\beta_i) - \mathbb{P}_N g_{yv}(v_n;\alpha_i,\beta_i)\right| = O_p\left(\sqrt{\ln N/Nh_P^{1+2r}}\right)$ by Hansen

(2008), Theorem 8. Hence,

$$\sup\left|\widetilde{g}_{yv}(v_n;\alpha_i,\beta_i)-g_{yv}(v_n;\alpha_i,\beta_i)\right| \leq \\ \sup\left|\mathbb{P}_N\widetilde{g}_{yv}(v_n;\alpha_i,\beta_i)-g_{yv}(v_n;\alpha_i,\beta_i)\right|+O_p\left((\ln N/Nh_P)^{1/2}\right)$$

Let $\Delta(x)=(\widehat{\varphi}_i(x)-\varphi_i(x))/h_P$. Note that, uniformly on x

$$\begin{aligned}\mathbb{P}_N\widetilde{g}_{yv}(v_n;\alpha_i,\beta_i) &= \int_{\mathbb{R}^{2d}}\frac{1}{h}K\left(\frac{v(x;\alpha_i,\beta_i)-v(t;\alpha_i,\beta_i)}{h_P}+\Delta(x)-\Delta(t)\right)g(v(t;\alpha_i,\beta_i))dt \\ &= \int K(u)\,g\Big[(v(t;\alpha_i,\beta_i)-(u-\Delta(x)+\Delta(t))h_P\Big]dt \\ &= g\left[v(t;\alpha_i,\beta_i)\right]+O_p\left((\ln N/N)^{R/(2R+4)}\right)+O(h_N^2).\end{aligned}$$

By assumption R,

$$\sup\left|\widetilde{g}_{yv}(v_n;\alpha_i,\beta_i)-g_{yv}(v_n;\alpha_i,\beta_i)\right|=O_p\left(\sqrt{\ln N/Nh_P}\vee h^2\right).$$

Similarly,

$$\sup\left|D^1(a_i,b_i)\widetilde{g}_{yv}(v_n;\alpha_i,\beta_i)-D^1(a_i,b_i)g_{yv}(v_n;\alpha_i,\beta_i)\right|=O_p\left(\sqrt{\ln N/Nh_N^3}\vee h^2\right).$$

□

Appendix C

Appendix to Chapter 3

C.1 Proofs of Lemmas and Propositions

C.1.1 Proof of Lemma 14

Proof. Consider the following map,

$$\phi: \ \theta(\cdot) \to \cdot \times \theta(\cdot^{\frac{1}{I-1}})$$

We can verify that ϕ is linear and Hadamard-differentiable, and

$$\phi'_\theta(h)(t) = t \cdot h(t^{\frac{1}{I-1}})$$

In addition, we know that $\sqrt{n}\left(\hat{G}^{-1}(t) - G^{-1}(t)\right) \Rightarrow \frac{-\mathbb{B}(t)}{g(G^{-1}(t))}$ for all $t \in [0,1]$ because the density $g(\cdot)$ is bounded away from zero in the whole support and an argument of Example 3.9.24 on P387 of Van Der Vaart and Wellner (1996) applies. By functional delta method, we have

$$\sqrt{n}\left[\hat{D}(t) - D(t)\right] \Rightarrow \frac{-t \cdot \mathbb{B}(t^{\frac{1}{I-1}})}{g(G^{-1}(t^{\frac{1}{I-1}}))}$$

where such weak convergence holds for all $t \in [0,1]$.

□

C.1.2 Proof of Proposition 12

Proof. Notice that

$$\begin{aligned} t_n &= \sup_{t\in[0,1]} \sqrt{n}\cdot\left[\hat{D}(t) - \underline{\mathcal{C}}_{[0,1]}\left(\hat{D}\right)(t)\right] \\ &= \sup_{t\in[0,1]} \left\{\sqrt{n}\left[\hat{D}(t) - D(t)\right] - \sqrt{n}\left[\underline{\mathcal{C}}_{[0,1]}\left(\hat{D}\right)(t) - D(t)\right]\right\} \end{aligned} \tag{C.1}$$

By the definition of g.c.m. operator $\underline{\mathcal{C}}_{[0,1]}$, we have

$$\begin{aligned} & \underline{\mathcal{C}}_{[0,1]}\left[\hat{D} - D\right](t) + D(t) \\ \leq\ & \hat{D}(t) - D(t) + D(t) \\ =\ & \hat{D}(t) \qquad \text{for} \quad \forall t \in [0,1] \end{aligned} \tag{C.2}$$

Under H_0, $D(t)$ is convex in $[0,1]$, thus $\underline{\mathcal{C}}_{[0,1]}\left[\hat{D} - D\right](t) + D(t)$ is convex in $[0,1]$ under $\forall D(\cdot) \in H_0$. Consequently, $(C.2)$ implies that for any $D(\cdot) \in H_0$,

$$\underline{\mathcal{C}}_{[0,1]}\left[\hat{D} - D\right](t) + D(t) \quad \leq \quad \underline{\mathcal{C}}_{[0,1]}\left[\hat{D}\right](t), \qquad \text{for} \quad \forall t \in [0,1]$$

i.e.

$$\underline{\mathcal{C}}_{[0,1]}\left[\hat{D} - D\right](t) \quad \leq \quad \underline{\mathcal{C}}_{[0,1]}\left[\hat{D}\right](t) - D(t), \qquad \text{for} \quad \forall t \in [0,1] \tag{C.3}$$

where the equality holds when $D(t)$ is affine.

Combining $(C.1)$ and $(C.3)$, we can get that for any $D(\cdot) \in H_0$,

$$\begin{aligned} t_n &\leq \sup_{t\in[0,1]} \left\{\sqrt{n}\left[\hat{D}(t) - D(t)\right] - \sqrt{n}\cdot\underline{\mathcal{C}}_{[0,1]}\left(\hat{D} - D\right)(t)\right\} \\ &\equiv \tilde{t}_n \end{aligned}$$

where equality holds when $D(t)$ is affine.

In the rest of this proof, we will show that $\tilde{t}_n \xrightarrow{d} \sup_{t\in[0,1]}\left\{\mathbb{K}(t) - \underline{\mathcal{C}}_{[0,1]}(\mathbb{K})(t)\right\}$ where process $\mathbb{K}(t)$ is defined by (3.7).

First of all, Lemma 14 shows that $\sqrt{n}\left(\hat{D}(t) - D(t)\right) \Rightarrow \mathbb{K}(t)$. By applying functional continuous mapping theorem, we could get

$$\begin{aligned}
\tilde{t}_n &\equiv \sup_{t\in[0,1]} \left\{\sqrt{n}\left[\hat{D}(t) - D(t)\right] - \sqrt{n}\cdot \underline{\mathcal{C}}_{[0,1]}\left(\hat{D} - D\right)(t)\right\} \\
&\overset{d}{\to} \sup_{t\in[0,1]} \left\{\mathbb{K}(t) - \underline{\mathcal{C}}_{[0,1]}(\mathbb{K})(t)\right\}
\end{aligned}$$

□

C.1.3 Proof of Proposition 13

Proof. To prove (i), it suffices to show that

$$\lim_{n\to\infty} Pr_G\left(\tilde{t}_n \geq c_{1-\alpha}\left(\hat{G}^{-1}(\cdot), \hat{g}(\cdot)\right)\right) = \alpha$$

for any $G(\cdot) \in H_0$, due to the fact of $t_n \leq \tilde{t}_n$ under the null.

Note that $c_{1-\alpha}\left(G^{-1}(\cdot), g(\cdot)\right)$ is continuous in $G^{-1}(\cdot)$ and $g(\cdot)$. Consequently,

$$\begin{aligned}
&\lim_{n\to\infty} Pr_G\left(\tilde{t}_n \geq c_{1-\alpha}\left(\hat{G}^{-1}(\cdot), \hat{g}(\cdot)\right)\right) \\
=& \lim_{n\to\infty} Pr_G\left(\tilde{t}_n \geq c_{1-\alpha}\left(G^{-1}(\cdot), g(\cdot)\right)\right) \\
=& \alpha
\end{aligned}$$

To prove (ii), we note that $\hat{D}(\cdot) \to D(\cdot)$ as sample size $n \to \infty$. By continuous mapping theorem, $\underline{\mathcal{C}}_{[0,1]}(\hat{D})(\cdot) \to \underline{\mathcal{C}}_{[0,1]}(D)(\cdot)$ as $n \to \infty$. Consequently, as $n \to \infty$,

$$\sup_{t\in[0,1]} \left[\hat{D}(t) - \underline{\mathcal{C}}_{[0,1]}(\hat{D})(t)\right] \to \sup_{t\in[0,1]} \left[D(t) - \underline{\mathcal{C}}_{[0,1]}(D)(t)\right]$$

In addition, we have $\sup_{t\in[0,1]} \left[D(t) - \underline{\mathcal{C}}_{[0,1]}(D)(t)\right] > 0$ under H_1. Thus, under the alternative, $\sqrt{n}\cdot\sup_{t\in[0,1]} \left[D(t) - \underline{\mathcal{C}}_{[0,1]}(D)(t)\right] \to \infty$ as $n \to \infty$ which implies that

$$t_n \equiv \sqrt{n}\cdot \sup_{t\in[0,1]} \left[\hat{D}(t) - \underline{\mathcal{C}}_{[0,1]}(\hat{D})(t)\right] \to \infty, \quad n \to \infty$$

under H_1. Also, we have $c_{1-\alpha}\left(\hat{G}^{-1}(\cdot),\hat{g}(\cdot)\right) \to c_{1-\alpha}\left(G^{-1}(\cdot),g(\cdot)\right)$ as $n \to \infty$. Part (ii) then follows.

□

C.1.4 Proof of Lemma 16

Proof. Notice that $\tilde{t}_n = R_n(b^{(n)}, G)$ is actually a continuous function of $S_n(b^{(n)}, G)$. From continuous mapping theorem, we can conclude that

$$\rho(J_n(\cdot,\hat{G}), J_n(\cdot,G)) \to 0 \text{ with probability one}$$

based on the result of Lemma 15. □

C.1.5 Proof of Proposition 14

Proof. To prove part (i), it suffices to show the following statement is true: For any $G(\cdot) \in H_0$, we have

$$\lim_{n\to\infty} Pr_G\left(\tilde{t}_n \geq J_n(\hat{G}, 1-\alpha)\right) = \alpha$$

It can be verified that $J_n(G, 1-\alpha)$ is continuous in G , and $\| \hat{G}(\cdot) - G(\cdot) \|_\infty \to 0$ as $n \to \infty$ due to the Glivenko-Cantelli theorem. Thus we can conclude that

$$\begin{aligned} & Pr_G\left(\tilde{t}_n \geq J_n(\hat{G}, 1-\alpha)\right) \\ = \; & Pr_G\left(\tilde{t}_n \geq J_n(G, 1-\alpha)\right) + o(1) \end{aligned}$$

which implies that $\lim_{n\to\infty} Pr_G\left(\tilde{t}_n \geq J_n(\hat{G}, 1-\alpha)\right) = \alpha$.

To prove part (ii), notice that $J_n(\hat{G}, 1-\alpha) \to J_\infty(G, 1-\alpha)$ as $n \to \infty$. The rest of proof follows a similar argument of proving part (ii) of proposition 13. □

Bibliography

AGUIRREGABIRIA, V., AND P. MIRA (2007): "Sequential estimation of dynamic discrete games," *Econometrica*, 75(1), 1–53.

ANDREWS, D. (1992): "Generic uniform convergence," *Econometric Theory*, 8(02), 241–257.

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

ARADILLAS-LOPEZ, A., AND E. TAMER (2008): "The identification power of equilibrium in simple games," *Journal of Business & Economic Statistics*, 26(3), 261–283.

ATHEY, S. (2001): "Single crossing properties and the existence of pure strategy equilibria in games of incomplete information," *Econometrica*, 69(4), 861–889.

ATHEY, S., AND P. A. HAILE (2002): "Identification of standard auction models," *Econometrica*, 70(6), 2107–2140.

ATHEY, S., J. LEVIN, AND E. SEIRA (2011): "Comparing open and Sealed Bid Auctions: Evidence from Timber Auctions*," *The Quarterly Journal of Economics*, 126(1), 207–257.

AUMANN, R. (1987): "Correlated equilibrium as an expression of Bayesian rationality," *Econometrica: Journal of the Econometric Society*, pp. 1–18.

BAJARI, P., H. HONG, J. KRAINER, AND D. NEKIPELOV (2010): "Estimating Static Models of Strategic Interactions," *Journal of Business and Economic Statistics*, 28(4), 469–482.

BAJARI, P., H. HONG, AND D. NEKIPELOV (2010): "Game Theory and Econometrics: A Survey of Some Recent Research," in *Econometric Society World Congress in Shanghai*.

BAJARI, P., S. HOUGHTON, AND S. TADELIS (2011): "Bidding for Incomplete Contracts: An Empirical Analysis of Adaptation Costs," Working paper.

BAJARI, P., AND L. YE (2003): "Deciding between competition and collusion," *Review of Economics and Statistics*, 85(4), 971–989.

BAJARI, P. L. (1997): "The First Price Sealed Bid Auction with Asymmetric Bidders: Theory and Applications," Ph.D. thesis, University of Minnesota.

BANKS, J., R. BLUNDELL, AND M. CASANOVA RIVAS (2010): "The dynamics of retirement behavior in couples: reduced-form evidence from England and the US," *working Paper*.

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

BICKEL, P., AND D. FREEDMAN (1981): "Some asymptotic theory for the bootstrap," *The Annals of Statistics*, 9(6), 1196–1217.

BIERENS, H. (2011): "The Hilbert Space Theoretical Foundation of Semi-Nonparametric Modeling," *working paper*.

BIERENS, H. J., AND W. PLOBERGER (1997): "Asymptotic theory of integrated conditional moment tests," *Econometrica*, pp. 1129–1151.

BILLINGSLEY, P. (2012): *Probability and measure*, vol. 939. Wiley.

BJORN, P., AND Q. VUONG (1984): "Simultaneous equations models for dummy endogenous variables: a game theoretic formulation with an application to labor force participation," *working Papers*.

BJORN, P. A., AND Q. H. VUONG (1985): "Econometric modeling of a Stackelberg game with an application to labor force participation," *California Institute of Technology, Division of the Humanities and Social Sciences Working Papers*.

BRESNAHAN, T., AND P. REISS (1990): "Entry in Monopoly Market," *The Review of Economic Studies*, 57(4), 531.

BRESNAHAN, T. F., AND P. C. REISS (1991): "Empirical models of discrete games," *Journal of Econometrics*, 48(1-2), 57–81.

BROCK, W., AND S. DURLAUF (2001a): "Discrete choice with social interactions," *The Review of Economic Studies*, 68(2), 235–260.

——— (2001b): "Interactions-based models," *Handbook of econometrics*, 5, 3297–3380.

——— (2007): "Identification of binary choice models with social interactions," *Journal of Econometrics*, 140(1), 52–75.

CASANOVA, M. (2010): "Happy together: A structural model of couples' joint retirement choices," Discussion paper, Mimeo, UCLA.

CHESHER, A. (2003): "Identification in nonseparable models," *Econometrica*, 71(5), 1405–1441.

——— (2005): "Nonparametric identification under discrete variation," *Econometrica*, 73(5), 1525–1550.

CHESHER, A., AND A. ROSEN (2011): "Simultaneous equations models for discrete outcomes: coherence, completeness, and identification," *working paper*.

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

DARSOW, W., B. NGUYEN, AND E. OLSEN (1992): "Copulas and Markov processes," *Illinois Journal of Mathematics*, 36(4), 600–642.

DE PAULA, Á. (2012): "Econometric analysis of games with multiple equilibria," *Annual Review of Economics*, 5(1).

DE PAULA, A., AND X. TANG (2010): "Inference of signs of interaction effects in simultaneous games with incomplete information, second version," *forthcoming, Econometrica*.

——— (2012): "Inference of signs of interaction effects in simultaneous games with incomplete information," *Econometrica*, 80(1), 143–172.

DELGADO, M. A., AND J. C. ESCANCIANO (2012): "Distribution-free tests of stochastic monotonicity," *Journal of Econometrics*, 170(1), 68–75.

ELYAKIME, B., J. J. LAFFONT, P. LOISEL, AND Q. VUONG (1994): "First-price sealed-bid auctions with secret reservation prices," *Annales d'Economie et de Statistique*, pp. 115–141.

FUDENBERG, D., AND J. TIROLE (1991): "Game Theory, 1991," .

GRIECO, P. (2011): "Discrete Games with Flexible Information Structures: An Application to Local Grocery Markets," *working paper*.

GUERRE, E., I. PERRIGNE, AND Q. VUONG (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68(3), 525–574.

HANSEN, B. (2008): "Uniform convergence rates for kernel estimation with dependent data," *Econometric Theory*, 24(3), 726.

HICKMAN, B. (2010): "Effort, Race Gaps and Affirmative Action: a Structural Policy Analysis of US College Admissions," Working paper, University of Chicago.

HURWICZ, L. (1950): "Generalization of the concept of identification," *Statistical Inference in Dynamic Economic Models (T. Koopmans, ed.). Cowles Commission, Monograph*, 10, 245–257.

ICHIMURA, H. (1993): "Semiparametric least squares (SLS) and weighted SLS estimation of single-index models," *Journal of Econometrics*, 58(1-2), 71–120.

IMBENS, G. W., AND W. K. NEWEY (2009): "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," *Econometrica*, 77(5), 1481–1512.

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

JUN, S. J., J. PINKSE, AND H. XU (2011): "Tighter bounds in triangular systems," *Journal of Econometrics*, 161(2), 122–128.

KLEIN, R., AND R. SPADY (1993): "An efficient semiparametric estimator for binary response models," *Econometrica*, 61(2), 387–421.

KLINE, B. (2012): "Identification of complete information games," *working paper*.

KOOPMANS, T., AND O. REIERSOL (1950): "The identification of structural characteristics," *The Annals of Mathematical Statistics*, 21(2), 165–181.

KOOREMAN, P. (1994): "Estimation of Econometric Models of Some Discrete Games," *Journal of Applied Econometrics*, pp. 255–268.

KOSOROK, M. (2008): *Introduction to empirical processes and semiparametric inference*. Springer Verlag.

LAFFONT, J.-J., H. OSSARD, AND Q. VUONG (1995): "Econometrics of first-price auctions," *Econometrica*, pp. 953–980.

LEHMANN, E. (1955): "Ordered families of distributions," *The Annals of Mathematical Statistics*, 26(3), 399–419.

LEWBEL, A., AND X. TANG (2011): "Identification and estimation of discrete Bayesian games with multiple equilibria using excluded regressors," Discussion paper, Working paper, Boston College.

——— (2012): "Identification and Estimation of Discrete Bayesian Games with Multiple Equilibria Using Excluded Regressors," Discussion paper, working paper, Boston College.

LI, T., I. PERRIGNE, AND Q. VUONG (2000): "Conditionally independent private information in OCS wildcat auctions," *Journal of Econometrics*, 98(1), 129–162.

LIU, N., Q. VUONG, AND H. XU (2012): "Nonparametric analysis of binary games of incomplete information," *Working paper*.

LIU, P. N., AND H. XU (2012): "Semiparametric analysis of binary games of incomplete information," *working Paper*.

MAMMEN, E., C. ROTHE, AND M. SCHIENLE (2012): "Nonparametric regression with nonparametrically generated covariates," *The Annals of Statistics*, 40(2), 1132–1170.

MANSKI, C. (1975): "Maximum score estimation of the stochastic utility model of choice," *Journal of Econometrics*, 3(3), 205–228.

——— (1985): "Semiparametric analysis of discrete response:: Asymptotic properties of the maximum score estimator," *Journal of Econometrics*, 27(3), 313–333.

MANSKI, C., AND E. TAMER (2003): "Inference on regressions with interval data on a regressor or outcome," *Econometrica*, 70(2), 519–546.

MATZKIN, R. (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.

——— (2003): "Nonparametric estimation of nonadditive random functions," *Econometrica*, 71(5), 1339–1375.

MCADAMS, D. (2003): "Isotone equilibrium in games of incomplete information," *Econometrica*, 71(4), 1191–1214.

PAARSCH, H. (1992): "Deciding between the common and private value paradigms in empirical models of auctions," *Journal of Econometrics*, 51(1-2), 191–215.

PERRIGNE, I., AND Q. VUONG (2011): "Nonlinear pricing in yellow pages," Working paper, Pennsylvania State University.

PESENDORFER, M., AND P. SCHMIDT-DENGLER (2003): "Identification and estimation of dynamic games," Discussion paper, National Bureau of Economic Research.

PORTER, R. H. (1995): "The Role of Information in US Offshore Oil and Gas Lease Auction," *Econometrica*, pp. 1–27.

POWELL, J., J. STOCK, AND T. STOKER (1989): "Semiparametric estimation of index coefficients," *Econometrica: Journal of the Econometric Society*, pp. 1403–1430.

RENY, P. (2011): "On the Existence of Monotone Pure-Strategy Equilibria in Bayesian Games," *Econometrica*, 79(2), 499–553.

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

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

SHAIKH, A., AND E. VYTLACIL (2011): "Partial identification in triangular systems of equations with binary dependent variables," *Econometrica*, 79(3), 949–955.

SOETEVENT, A., AND P. KOOREMAN (2007): "A discrete-choice model with social interactions: with an application to high school teen behavior," *Journal of Applied Econometrics*, 22(3), 599–624.

SU, L., AND H. WHITE (2007): "A consistent characteristic function-based test for conditional independence," *Journal of Econometrics*, 141(2), 807–834.

——— (2008): "A nonparametric Hellinger metric test for conditional independence," *Econometric Theory*, 24(4), 829.

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.

TANG, X. (2010): "Estimating simultaneous games with incomplete information under median restrictions," *Economics Letters*, 108(3), 273–276.

VAN DER VAART, A., AND J. WELLNER (1996): *Weak convergence and empirical processes*. Springer Verlag.

VICKREY, W. (1961): "Counterspeculation, auctions, and competitive sealed tenders," *The Journal of finance*, 16(1), 8–37.

VIVES, X. (1990): "Nash equilibrium with strategic complementarities," *Journal of Mathematical Economics*, 19(3), 305–321.

WAN, Y., AND H. XU (2010): "Semiparametric Estimation of Binary Decision Games of Incomplete Information with Correlated Private Signals," *working paper*.

WHITE, H., H. XU, AND K. CHALAK (2011): "Causal Discourse in a Game of Incomplete Information," *working paper*.

XU, H. (2010): "Estimation of Discrete Games with Correlated Types," *working paper*.

——— (2011): "Social interactions: a game theoretic approach," *working paper*.

Vita

Nianqing Liu

Education

Ph.D. Economics, The Pennsylvania State University, August 2013.

M.A. Economics, Peking University, June 2007.

B.S. Mathematics, Nankai University, June 2004.

Research Fields

Primary: Econometrics, Applied Econometrics.

Secondary: Industrial Organization, Game Theory.

Working Papers

1. "Rationalization and Nonparametric Identification of Discrete Games with Correlated Types," joint with Quang H. Vuong and Haiqing Xu, 2012.
2. "Semi-parametric Analysis of Binary Games of Incomplete Information," joint with Haiqing Xu, 2012.
3. "Nonparametric Test of Monotonicity of Bidding Strategy in First-price Auctions," joint with Quang H. Vuong, 2011.

Work in Progress

1. "Nonparametric Identification and Estimation of k-Double Auctions," joint with Huihui Li.
2. "Nonparametric Test of Exogeneity of Participation in First Price Auctions," joint with Yao Luo.
3. "The Distribution of the Maximal Difference between a Weighted Brownian Bridge and its Concave Majorant," joint with Quang H. Vuong.