NBER WORKING PAPER SERIES

IDENTIFICATION OF POTENTIAL GAMES AND DEMAND MODELS FOR BUNDLES

Jeremy T. Fox
Natalia Lazzati

Working Paper 18155
http://www.nber.org/papers/w18155

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2012

Thanks to Tilman Borgers and Stephan Lauermann for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Identification of Potential Games and Demand Models for Bundles
Jeremy T. Fox and Natalia Lazzati
NBER Working Paper No. 18155
June 2012
JEL No. C35,L13

ABSTRACT

This paper studies nonparametric identification in binary choice games of complete information. We allow for correlated unobservables across players. We propose conditions under which the binary choice game is a so-called potential game and impose that the selected equilibrium maximizes its associated potential function. Our framework is formally equivalent to a multinomial choice demand model where a consumer can elect to purchase any bundle of products. We present a separate identification result for two-player games that does not rely on equilibrium selection.

Jeremy T. Fox
Economics Department
University of Michigan
238 Lorch Hall
611 Tappan Ave
Ann Arbor, MI 48109
and NBER
jeremyfox@gmail.com

Natalia Lazzati
Economics Department
University of Michigan
238 Lorch Hall
611 Tappan Ave
Ann Arbor, MI 48109
nlazzati@umich.edu

1 Introduction

Our paper presents identification results for two classes of models that we show are formally equivalent. For clarity, we discuss the two models separately in the introduction.

1.1 Discrete Choice Games

We study identification of static, binary choice games of complete information with two or more players. Such models include games of product adoption with social interactions, entry games, and labor force participation games, among many others.[1] In these models the interaction effects capture the strategic interdependence among the decisions of players and are therefore a fundamental target of estimation. Interaction effects are often hard to distinguish from unobserved preferences that are correlated across players, as these two sources of interdependence have similar observable implications. For example, similar smoking behavior among friends may be due to peer effects or to correlated tastes for smoking. Likewise, firms may tend to enter into geographic markets in clusters because the competitive effects of entry are small or because certain markets are profitable for unobservable reasons. Though it is well known that not taking the correlation of unobservables into account can lead to a serious bias in the estimates, much of the empirical literature assumes some kind of independence or conditional independence. We introduce the class of potential games to model strategic interactions and propose an identification approach that allows for correlated unobservables across players.

A game is said to be a potential game if it admits a so-called exact potential function, which is a real-valued function defined on the space of pure strategy profiles such that the change in any player's payoffs from a unilateral deviation equals the change in the associated potential. In an influential paper, Monderer and Shapley (1996) show that any maximizer of the potential function is a Nash equilibrium of the associated game (for exact potential games) and that this function is uniquely defined up to an additive constant. Thus, the set of maximizers does not depend on the specific potential function that we use and the potential offers an equilibrium refinement. Under mild conditions, in our model the maximizer is

[1]See, e.g., Heckman (1978), Bjorn and Vuong (1984), Bresnahan and Reiss (1991), Berry (1992), Mazzeo (2002), Ciliberto and Tamer (2009), Hartmann (2009), Bajari et al. (2010) and Card and Giuliano (2012).

unique with probability 1. As we will discuss in more detail below, one necessary requirement for a game to admit a potential function is that interaction effects, but not standalone utilities from actions, be groupwise symmetric.

Recent theoretical and experimental work has provided economic justification to the potential function refinement. Most importantly for our approach, Ui (2001) shows that if a unique Nash equilibria maximizes the potential function, then that equilibrium is robust in the sense of Kajii and Morris (1997a, 1997b), particularly the 1997b paper. Roughly speaking, a Nash equilibrium of a complete information game is robust if every incomplete information game with payoffs almost always given by the complete information game has an equilibrium that generates behavior close to the original equilibrium. Ui does not show that all robust equilibria maximize a potential function in a potential game, but, at a minimum, the maximizer of the potential selects one such robust equilibrium. There is also experimental evidence remarkably consistent with this equilibrium selection rule in the minimum effort (stag hunt) game (Van Huyck et al. 1990, Goeree and Holt 2005, and Chen and Chen 2011). An advantage of using any equilibrium selection rule is that the model can make unique counterfactual predictions. As structural analysis is often motivated by a desire to estimate policy counterfactuals, some selection rule will increase the precision (and ease of replication) of certain conclusions in a typical empirical paper estimating a game theoretic model.

This paper explores the power of the potential function selection mechanism from an econometric perspective.[2] As we explain below, the equilibrium selection rule based on potential maximizers not only completes the model, but it also links observed choices to a single maximization problem that has an alternative economic interpretation: under this refinement the game is observationally equivalent to a single agent discrete choice model for bundles of goods.

The identification strategy that we pursue relies on exclusion restrictions at the level of each player. These restrictions yield observed regressors that affect the payoff of one player but not the others and enter payoffs additively separably. These exclusion restrictions have been called special regressors in, for example, the literature on discrete choice (Lewbel 1998, 2000, Matzkin 2007, Berry and Haile 2010, Fox and Gandhi 2012). Our identification

[2]Deb (2009) uses potential game to provide testable implications of a model of consumption with externalities.

strategy also requires a condition under which the inequalities implied by choice data can be inverted in the unobservables. This condition calls for a vector of choices such that whenever individual deviations are not profitable, then that vector of choices maximizes the potential function associated to the game. We prove that this restriction is always satisfied when there are two players, when the interaction effects are negative (i.e., games of strategic substitutes), and when the interaction terms have arbitrary signs but are not too large in absolute value (i.e., concave games in the sense of Ui 2008). Our restrictions are enough to nonparametrically identify the entire structure of the game: players' subutilities of choosing each action as a function of the regressors and the actions of other players as well as the distribution of potentially correlated unobservables.

The parametric identification and estimation of static, discrete choice games of complete information has been recently studied by Bajari et al. (2010), Beresteanu et al. (2011), Ciliberto and Tamer (2009), and Galichon and Henry (2010). Our paper departs from this literature in that it is nonparametric and demonstrates point rather than set identification. Given the trade-offs, our approach is more restrictive in other respects: we focus on the binary actions case (enter or not, smoke or not) and assume a specific equilibrium selection rule. Unlike some work on identification for complete information games, our proofs do not rely on identification at infinity. We do require large support on regressors to identify the joint distribution of the unobservables in the tails, just as in the binary and multinomial choice literature (e.g., Manski 1988, Thompson 1989, Matzkin 1993, Lewbel 2000). Binary games of incomplete information have received a lot of attention (e.g., Brock and Durlauf 2001, 2007, De Paula and Tang 2011, and Lewbel and Tang 2011). As Bajari et al. (2010) explain, the challenges involved in complete information games are quite different.

For the special case of two players, we offer an alternative proof of identification that does not rely on equilibrium selection and the symmetry restriction on the interaction effects needed for the game to admit a potential function. Building on Tamer (2003), Berry and Tamer (2006, Result 4) study identification in two-player, submodular games where the key special regressors enter multiplicatively instead of additively. Both results, theirs and ours, require knowledge of the signs of the interactions terms. The advantage of our identification argument with respect to these two papers is that we do not rely on identification at infinity to recover the interaction terms.

1.2 Discrete Demand for Bundles

Under the equilibrium selection rule based on potential maximizers, we show that the binary game is formally equivalent to a multinomial choice demand model where a consumer can elect to purchase any bundle of products. Therefore, our paper also investigates identification in multinomial choice models for bundles of goods. When applied to these models, we identify the stand-alone payoff of each choice, the terms representing complementarities or substitutabilities among products, and the joint distribution of the unobservables for the individual goods. Here again, we allow for correlated tastes for alternative goods, a distinct explanation for purchasing bundles of goods from true interdependences in consumption. For example, consumers could be seen to often purchase both cable television service and large-screen televisions because there are complementarities in the consumption of these two goods or because consumers with a high stand-alone utility for cable television tend to also have a high stand-alone utility for large-screen televisions. Our identification approach, relying on regressors that affect the stand-alone utility of one product but not the other, can distinguish complementarities in utility from correlated unobservable preferences.

Gentzkow (2007a) studies a parametric demand model for interdependent goods. He focuses on demand for print and online newspapers, allowing the marginal utility of consuming one good to depend on whether the other good is consumed as well. Athey and Stern (1998) study a model where firms have the possibility to combine individual activities in order to achieve their objective, and these activities are complements. To our knowledge, we are the first to give conditions for nonparametric identification in models such as Gentzkow's and Athey and Stern's. The sufficient conditions we propose to identify the distribution of the unobservables are economically similar to some of the requirements imposed in tangentially related identification papers. For example, Berry, Gandhi and Haile (2011) provide a sufficient condition under which a system of demand functions (with continuous outcomes such as shares) can be inverted in unobservables. Their condition applies to goods that are (connected) substitutes, and is loosely related to our result on games of strategic substitutes. Matzkin (2007) studies the inversion of simultaneous equations in models with continuous outcomes, imposing quasi-concavity. As we explained earlier, we suggest a discrete version of her requirement.

1.3 Estimation

We do not formally explore estimation because the estimation problem is relatively straightforward conceptually once identification is established. One could apply simulated maximum likelihood with sieve-based approaches for modeling the distribution of the unobservables and the also infinite-dimensional standalone utilities and interaction effects (Chen 2007).

The rest of the paper is as follows. Section 2 describes the game. Section 3 defines potential games, provides a precise condition under which the games we study admit a potential function, and sheds light on the mechanism by which the potential selects among multiple equilibria. Section 4 provides the identification results for potential games. Section 5 shows identification in two player games without relying on an equilibrium selection rule or on symmetry in interaction effects. Section 6 presents identification results for models of discrete demand for bundles of goods. Section 7 concludes. All proofs are collected in an appendix.

2 The General Game

We consider a simultaneous discrete choice game with complete information. The set of players is $\mathcal{N} = \{1, ..., n\}$. Each player $i \in \mathcal{N}$ chooses an action a_i from two possible alternatives $\{0, 1\}$. We denote by $X \in \mathcal{X} \subseteq \mathbb{R}^k$ a vector of observable state variables (with x the realization of the random variable) and let $\varepsilon \equiv (\varepsilon_i)_{i \leq n} \in \mathbb{R}^n$ indicate a vector of random terms that are observed by the players but not by the econometrician. The random vector ε is distributed according to the cumulative distribution function (CDF) $F_{\varepsilon|x}$. We allow the unobservables to be correlated even after conditioning on the observable state variables. The payoff of player i from choosing action 1 is

$$U_{1i}(a_{-i}, x, \varepsilon_i) \equiv u_i(x) + v_i(a_{-i}, x) + \varepsilon_i, \tag{1}$$

while the return from action 0, $U_{0i}(a_{-i}, x, \varepsilon_i)$, is normalized to 0. The first element of (1), $u_i(x)$, is the stand-alone value of action 1, and $v_i(a_{-i}, x)$ captures the effect that the choices of other players have on player i. We denote this game by $\Gamma(x,\varepsilon)$.[3]

A vector of decisions $a^* \equiv (a_i^*)_{i \le n}$ is a pure strategy Nash equilibrium if, for all $i \in \mathcal{N}$,

$$a_i^* = \begin{cases} 1, & \text{if } u_i(x) + v_i\left(a_{-i}^*, x\right) + \varepsilon_i > 0 \\ 0, & \text{if } u_i(x) + v_i\left(a_{-i}^*, x\right) + \varepsilon_i < 0 \\ 1 \text{ or } 0, & \text{otherwise.} \end{cases}$$

We write $\mathcal{D}(x,\varepsilon)$ for the equilibrium set (in pure strategies) of $\Gamma(x,\varepsilon)$. The same conditions that will facilitate identification of the game guarantee that $\mathcal{D}(x,\varepsilon)$ is non-empty.

The purpose of our analysis is to recover the structure of the game from available data on choices and covariates. Before formalizing our objective, we introduce potential games.

3 Potential Games

In game theory, a potential function—or, more precisely, an exact potential—is a real-valued function defined on the space of pure strategy profiles such that the change in any player's payoffs from a unilateral deviation is equal to the gain in the potential function. When a game admits such a function it is called a potential game.[4] This concept was first used in economics as a way to prove existence of Nash equilibrium in pure strategies.[5] The reason is that the set of maximizers of the potential function corresponds to a subset of equilibria of the related game. In finite games, the potential function has a finite set of values and always has a maximizer. It follows that the equilibrium set (in pure strategies) of any finite potential game is guaranteed to be non-empty. We will argue below that the distribution of unobservables, ϵ, in our model implies that a unique pure strategy equilibrium will maximize the potential

[3] The additive separability of $u_i(x)$ and $v_i(a_{-i},x)$ in $U_{1i}(a_{-i},x,\varepsilon_i)$ does not impose any restriction on the model; it is just a convenient way to write the payoffs. On the other hand, the additive separability of ε_i in $U_{1i}(a_{-i},x,\varepsilon_i)$ is a restriction in the sense that it eliminates the possibility of interaction effects between the unobservables of player i and the choices of other players in the game.

[4] Monderer and Shapley (1996) also define ordinal and weighted potential games. We study identification for exact potential games. Thus, throughout, when we say potential, we mean exact potential. This follows the recent theoretical literature on equilibrium refinements, which focuses on exact potentials (e.g., Ui 2000, 2001).

[5] According to Monderer and Shapley (1996), this concept appeared for the first time in Rosenthal (1973) to prove equilibrium existence in congestion games.

function with probability 1 (over ϵ). Therefore, a potential maximizing equilibrium always exists and is unique with probability 1.

Monderer and Shapley (1996) show that when a game admits an exact potential function, this function is uniquely defined up to an additive constant. Thus, the set of maximizers does not depend on the specific potential function we use and the potential offers an equilibrium refinement. Important work has been done to address whether this selection rule is economically meaningful. Ui (2001) shows that if a unique Nash equilibrium maximizes the potential function, then that equilibrium is generically robust in the sense of Kajii and Morris (1997a, 1997b), particularly the definition in the 1997b paper. Roughly speaking, a Nash equilibrium of a complete information game is said to be robust if every incomplete information game with payoffs almost always given by the complete information game has an equilibrium that generates behavior close to the original equilibrium. See Ui for a formal definition. All robust equilibria do not necessarily maximize a potential function in a potential game, but, at a minimum, the maximizer of the potential selects one such robust equilibrium. Ui writes that "It is an open question when robust equilibria are unique, if they exist."[6]

In lab experiments studying the so-called minimum effort game, observed choice data were shown to be consistent with the maximization of objects close to the potential function associated to the game (Van Huyck et al. 1990, Goeree and Holt 2005, and Chen and Chen 2011). For example, Van Huyck et al. find that subjects converge to high and low effort levels according to the prediction of the potential maximizer refinement.[7] While we do not restrict ourselves to only minimum effort games, it is interesting that the potential maximizer refinement can explain the subtle experimental evidence in this literature.

From an econometric perspective, finite potential games are attractive for at least two

[6]In a certain two-player, exact potential game, Blume (1993) states that a log-linear strategy revision process selects potential maximizers. See also Morris and Ui (2005) for a discussion of so-called generalized potential functions and robust sets of equilibria. Weinstein and Yildiz (2007) present a deep critique of all refinements of rationalizability, including Nash equilibrium. They write about the robustness definition in Kajii and Morris (1997a): "Then the key difference between our notions of perturbation is that they focus on small changes to prior beliefs without regard to the size of changes to interim beliefs, while our focus is the reverse. Their approach is appropriate when there is an ex ante stage along with well understood inference rules and we know the prior to some degree."

[7]See Monderer and Shapley (1996, Section 5) for the original discussion of how Van Huyck et al.'s experimental evidence relates to potential games.

reasons. First, by offering a meaningful equilibrium selection rule, the potential function completes the model. Second, this function links observed choices to a single maximization problem that relates to other relevant economic problems (such as demand for bundles) and aids identification.

We next specify the definition of a potential function and then provide a necessary and sufficient condition for $\Gamma(x,\varepsilon)$ to be a potential game. In what follows, $U_i(a_i, a_{-i}, x, \varepsilon_i) \equiv U_{0i}(a_{-i}, x, \varepsilon_i) 1(a_i = 0) + U_{1i}(a_{-i}, x, \varepsilon_i) 1(a_i = 1)$.

Potential: A function $V : \{0,1\}^n \times \mathcal{X} \times \mathbb{R}^n \to \mathbb{R}$ is a **potential function** for Γ if, for all $i \leq n$, all $a_i, a_i' \in \{0,1\}$, and all $a_{-i} \in \{0,1\}^{n-1}$,

$$V(a_i, a_{-i}, x, \varepsilon) - V(a_i', a_{-i}, x, \varepsilon) = U_i(a_i, a_{-i}, x, \varepsilon_i) - U_i(a_i', a_{-i}, x, \varepsilon_i).$$

Γ is called a **potential game** if it admits a potential function.

Monderer and Shapley (1996) show that Γ is a potential game if (and only if) cross effects on payoffs are symmetric for every pair of players. Ui (2000) provides an alternative characterization of this class of games, which has the additional advantage of describing the potential function. The next result derives from Ui (2000, Theorem 3). We write $S(a) \subset \mathcal{N}$ for the set of players who select action 1 in a,

$$\mathcal{S}(a) = \{S \subseteq S(a) \mid |S| \geq 2\}, \text{and}\, \mathcal{S}(a,i) = \{S \subseteq S(a) \mid |S| \geq 2, i \in S\}.$$

Proposition 1. Γ *is a potential game if and only if there exists a function*

$$\{\Phi(S,x) \mid \Phi(S,x) : S \to \mathbb{R}, S \subset \mathcal{N}, |S| \geq 2\}$$

such that, for all $a \in \{0,1\}^n$ *and all* $i \in \mathcal{N}$,

$$U_i(a, x, \varepsilon_i) = (u_i(x) + \varepsilon_i) 1(a_i = 1) + \sum\nolimits_{S \in \mathcal{S}(a,i)} \Phi(S,x).$$

A potential function is given by

$$V(a,x,\varepsilon)=\sum\nolimits_{i\in\mathcal{N}}(u_i(x)+\varepsilon_i)\,1(a_i=1)+\sum\nolimits_{S\in\mathcal{S}(a)}\Phi(S,x).$$

If ε has an atomless support, then the maximizer of the potential function in Proposition 1 is unique with probability 1. By assuming that observed choices correspond to the potential maximizer, we therefore guarantee there is a unique equilibrium in pure strategies in the data generating process.[8]

The next example applies Proposition 1 to the case of two players. It clarifies the conditions under which the game is a potential game and sheds light on the mechanism by which the potential selects an equilibrium when the model makes multiple predictions.

Example 1. (Two-Player Game) Let $\mathcal{N}=\{1,2\}$. The payoff of player i can be written as

$$(u_i(x)+\varepsilon_i+v_i(x)\,1(a_{-i}=1))\,1(a_i=1)\ \text{ for } i=1,2.$$

There is only one set S to consider, i.e., $\{1,2\}$. Interaction terms are given by

$$v_1(x)\,1(a_1=1)\,1(a_2=1)\ \text{ and } v_2(x)\,1(a_1=1)\,1(a_2=1).$$

Thus, by Proposition 1, Γ is a potential game if (and only if) $v_1(x)=v_2(x)=\Phi(\{1,2\},x)$, in which case the potential function can be written as

$$V(a,x,\varepsilon)=(u_1(x)+\varepsilon_1)\,1(a_1=1)+(u_2(x)+\varepsilon_2)\,1(a_2=1)+\Phi(\{1,2\},x)\,1(a_1=1,a_2=1).$$

Therefore, interaction terms have to be the same for both players for the game to admit a potential function. We write $\Phi(x)=\Phi(\{1,2\},x)$ to simplify the exposition.

Because Γ is a finite potential game, it always has at least one equilibrium in pure strategies. When the equilibrium is unique, it is always a potential maximizer. Under multiple equilibria, the criterion by which the potential function selects one of them depends

[8] Monderer and Shapley (1996) also discuss mixed strategy equilibria. We focus on only pure strategies not because mixed strategies are inconsistent with the potential function, but because mixed strategies complicate identification by reintroducing multiple equilibria.

on the sign of the interaction effect, as we explain next. We assume no tie events in the analysis below.

When $\Phi(x) \leq 0$ and the game has multiple equilibria, then the equilibrium set is $\mathcal{D} = \{(0,1),(1,0)\}$. In this case, $(0,1)$ maximizes the potential if

$$u_2(x) + \varepsilon_2 > u_1(x) + \varepsilon_1$$

and the maximizer is $(1,0)$ otherwise. Thus, the equilibrium selection rule induced by the potential function predicts that the player choosing action 1 shall be the one with the highest stand-alone value.

Alternatively, the equilibrium set is $\mathcal{D} = \{(0,0),(1,1)\}$ when $\Phi(x) \geq 0$ and the game has multiple equilibria. It can be easily checked that $(1,1)$ maximizes the potential if

$$(-u_1(x) - \varepsilon_1 - v(x))(-u_2(x) - \varepsilon_2 - v(x)) > (u_1(x) + \varepsilon_1)(u_2(x) + \varepsilon_2).$$

The maximizer is $(0,0)$ otherwise. That is, players coordinate on $(1,1)$ if the product of deviation losses from selecting action 0 as compared to action 1 while the other player selects action 1 are lower than the product of deviation looses from selecting action 1 instead of action 0 when the other player selects action 0. In this case, the potential maximizer corresponds to the less risky equilibrium of Harsanyi and Selten (1988). □

Example 2. (Three-Player Game) Let $\mathcal{N} = \{1,2,3\}$. Each player has its own stand-alone utility. The interaction terms are $\Phi(\{1,2\},x)$, $\Phi(\{1,3\},x)$, $\Phi(\{2,3\},x)$ and $\Phi(\{1,2,3\},x)$. Then, for example, the utility of choosing action 1 by player 1 is

$$u_1(x)+\varepsilon_1+\Phi(\{1,2\},x)1(a_2=1)+\Phi(\{1,3\},x)1(a_3=1)+\Phi(\{1,2,3\},x)1(a_2=1,a_3=1).$$

□

These two examples shed light on the restrictions embedded in Proposition 1 for a game to admit a potential function. Each interaction term $\Phi(S,x)$ needs to be the same across all players in $S \subset \mathcal{N}$. While interaction terms have to respect this groupwise symmetry, the stand-alone values are allowed to deeply differ across individuals. Consider the example of entry by discount realtors in the United States (e.g., Ellickson et al. 2010, Jia 2008). The assumption of groupwise symmetry rules out that Walmart's entry in a geographic market lowers the profits of Kmart more than Kmart's entry lowers the profits of Walmart. We do allow the monopoly profit function of Walmart to vary considerably from that of Kmart, and we impose no restrictions on the joint distribution of the unobservables entering the profits of Walmart and Kmart. In the three-player case, the interaction term for both Kmart and Walmart entering can differ from the interaction term for both Kmart and Target entering, for example. There is a separate term $\Phi\left(\{1,2,3\},x\right)$ for all three players (Kmart, Target, Walmart) entering.

Many empirical applications impose that the Pareto optimal equilibrium is always selected in coordination games (see, e.g., Gowrisankaran and Stavins 2004, and Hartmann 2009). Under positive spillovers, this amounts to selecting the largest equilibrium (which always exists in this set-up). The criterion based on the potential maximizer makes predictions that are more in line with the promoters of coordination failures (see, e.g., Cooper and John 1988). Which selection device is more appropriate depends on the specific application we are dealing with; more research needs to be done to address this question. We should remark here that our results in Section 5 on the special case of two players offer a natural environment for performing a nonparametric test of equilibrium selection.

4 Identification in Potential Games

4.1 Assumptions and Main Result

For our identification purpose, we will assume that Γ admits a potential function and that observed actions correspond to a potential maximizer. Indeed, we consider games where observed choice behavior derives from the maximization problem

$$\max_a \left\{ V\left(a,x,\varepsilon\right) \equiv \sum\nolimits_{i\in\mathcal{N}} \left(u_i\left(x\right)+\varepsilon_i\right) 1\left(a_i=1\right) + \sum\nolimits_{S\in\mathcal{S}(a)} \Phi\left(S,x\right) \mid a\in\{0,1\}^n \right\}. \quad (2)$$

We assume the econometrician observes the vector of covariates and choices, $\left(x, (a_i)_{i\leq n}\right)$, for a cross section of independent n-player games that share the same structure $\Pi = \left((u_i)_{i\leq n}, \Phi, F_{\varepsilon|x}\right)$. These independent games can be different markets or different groups of friends, according to the particular application. The objective of the analysis is to combine initial assumptions with data to learn about Π.

We define

$$W(a,x) \equiv \sum\nolimits_{i\in\mathcal{N}} u_i(x)\, 1(a_i = 1) + \sum\nolimits_{S\in\mathcal{S}(a)} \Phi(S,x).$$

By our initial normalization, $W(a = (0,0,...,0), x) = 0$. For expositional ease, we order the elements of $\{0,1\}^n$ in terms of the lexicographic order so that $a^1 = (0,0,...,0)$, $a^2 = (1,0,...,0)$, ..., and $a^{2^n} = (1,1,...,1)$.

We indicate by $\triangle W(a',x)$ and $\triangle\varepsilon(a')$ the $(2^n - 1)-$dimensional vectors

$$\left(W(a^j,x) - W(a',x)\right)_{j\leq 2^n, a^j\neq a'} \text{ and } \left(\sum\nolimits_{i\in\mathcal{N}} \varepsilon_i 1(a'_i = 1) - \sum\nolimits_{i\in\mathcal{N}} \varepsilon_i 1\left(a^j_i = 1\right)\right)_{j\leq 2^n, a^j\neq a'}$$

respectively. Thus, for each a',

$$P\left(\triangle W(a',x) \mid x; F_{\varepsilon|x}\right) \equiv \Pr\left(\triangle\varepsilon(a') \leq \triangle W(a',x) \mid x\right)$$

captures the probability of observing choice vector a' conditional on $X = x$, or in simpler notation, $\Pr(a' \mid x)$.[9] The researcher can identify $\Pr(a' \mid x)$ directly from the data.

We now introduce our assumptions and then discuss how these restrictions are satisfied in alternative environments.

Assumption 1: Let $X \equiv \left(X', \widetilde{X}\right)$ with $\widetilde{X} \equiv \left(\widetilde{X}_i\right)_{i\leq n}$. We assume $u_i(x) \equiv u_i(x') + \widetilde{x}_i$,

---

[9] More formally, let us define

$$\triangle^j\varepsilon(a') \equiv \sum\nolimits_{i\in\mathcal{N}} \varepsilon_i 1(a'_i = 1) - \sum\nolimits_{i\in\mathcal{N}} \varepsilon_i 1\left(a^j_i = 1\right) \text{ and } \triangle^j W(a',x) \equiv W(a^j,x) - W(a',x).$$

Thus $\triangle\varepsilon(a')$ is the vector $\left(\triangle^j\varepsilon(a')\right)_{j\leq 2^n, a^j\neq a'}$, which has a distribution $dF_{\triangle\varepsilon(a')|x'}$. Then,

$$P\left(\triangle W(a',x) \mid x; F_{\varepsilon|x}\right) \equiv \int ... \int 1\left(\triangle^1\varepsilon(a') \leq \triangle^1 W(a',x)\right) ... 1\left(\triangle^{2^n}\varepsilon(a') \leq \triangle^{2^n} W(a',x)\right) dF_{\triangle\varepsilon(a')|x}.$$

and $\Phi(S,x) \equiv \Phi(S,x')$ for all $S \subset \mathcal{N}$ such that $|S| \geq 2$.
ASSUMPTION 2: The conditional distribution of each player's $\widetilde{X}_i$ given the other covariates has support on all of $\mathbb{R}$.
ASSUMPTION 3: (I) ε is independent of $\widetilde{X}$ and we write $F_{\varepsilon|x'}$ as a consequence; and (II) $\mathrm{E}(\varepsilon \mid x') = \mathrm{E}(\varepsilon) = 0$.
ASSUMPTION 4: ε has an everywhere positive Lebesgue density (on all $\mathbb{R}^n$) conditional on x'.
ASSUMPTION 5: For each $\widehat{x}'$ there exists a known vector $\widehat{a} \in \{0,1\}^n$ with the following properties. For all $\widetilde{x}$, where $\widehat{x} = (\widehat{x}', \widetilde{x})$, and for all ε, $V(a, \widehat{x}, \varepsilon)$ is maximized at $\widehat{a}$ if $V(\widehat{a}_i, \widehat{a}_{-i}, \widehat{x}, \varepsilon) \geq V(a_i, \widehat{a}_{-i}, \widehat{x}, \varepsilon)$ for all $i \in \mathcal{N}$, where $a_i = 1 - \widehat{a}_i$.

Assumption 1 requires exclusion restrictions at the level of each player, meaning we have n excluded regressors. It assumes $X \equiv \left(X', \widetilde{X}\right)$ includes a subvector $\widetilde{X}$ of player-specific factors that enter payoffs in an additively fashion. These are the special regressors familiar from the literature on binary and multinomial choice models without bundles (e.g., Manksi 1988 and Lewbel 2000). Exclusion restrictions are key to identifying $F_{\varepsilon|x'}$ separately from the interaction effects. The intuition is simple: $\widetilde{X}_i$ only affects player j's action through interaction effects. So interaction effects must be present if changes in the realization $\widetilde{x}_i$ correspond to changes in the marginal probability of player j's action.

Assumption 2 is a standard large support restriction on the special regressors. Without such a restriction, one cannot identify the tails of the distribution of $F_{\varepsilon|x'}$, just as in the literature on binary and multinomial choice. This use of large support is not what is informally called in the literature identification at infinity, which in our model would involve using only covariate values $\widetilde{x}$ that set the probability of all but one agent taking the binary action to be 1 or 0. This in effect turns a multi-player game into a single-agent choice model. Identification at infinity would never allow the identification of the joint distribution of the vector ε, which is key to our separate identification of correlated unobservables and interaction effects. Tamer (2003) did use an identification at infinity argument to establish parametric identification of the interaction effects in two-player games. Berry and Tamer (2006, Result 4) also looked at two-player games where special regressors enter multiplicatively and identified the distribution of unobservables without identification at infinity but used identification at infinity to identify the other parameters. The lack of a player-specific coefficient on $\widetilde{x}_i$ is just a scale normalization: utility is measured in the units of $\widetilde{x}_i$ for each player.[10]

Assumption 3(i) is necessary to recover the distribution of unobservables ε from variation in the special regressors $\widetilde{x}$. Assumption 3(ii) provides a location normalization. For any x', the mean of ε is not separately identified from $u_i(x')$, which subsumes the role of any intercept.[11]

Assumption 4 gives probability zero to tie events. It also insures that $P\left(\triangle W(a',x) \mid x; F_{\varepsilon|x'}\right)$ is continuous and strictly monotone in some of its arguments in the vector $\triangle W(a',x)$ (e.g., Matzkin 1993, Assumptions 2.5 and 2.6). We revisit the qualification "some of" below.

Assumption 5 is a key economic restriction that is new to our paper. It requires that there exists a vector of choices (known by the econometrician) that maximizes the potential function when individual deviations are non-profitable. In the next subsection, we provide three sufficient conditions for Assumption 5 to hold.

The structure $\Pi = \left((u_i)_{i\leq n}, \Phi, F_{\varepsilon|x'}\right)$ is said to be identified whenever for two values, Π_1 and Π_2, there exist values a and x where $\Pr(a \mid x; \Pi_1) \neq \Pr(a \mid x; \Pi_2)$. Here $\Pr(a \mid x; \Pi)$ is the probability of observing the actions a given covariates x when the true structure is Π. By Assumptions 1–4, establishing identification for one point x will extend identification to a set of x with positive measure. Our main identification result follows.

Theorem 1. *Under Assumptions 1–5,* $\Pi = \left((u_i)_{i\leq n}, \Phi, F_{\varepsilon|x'}\right)$ *is identified.*

The proof of the theorem is the direct consequence of two lemmas, which we state separately in order to give some intuition for the steps of the identification argument.

Lemma 1. *If Assumptions 1–3 and 5 are satisfied, then* $F_{\varepsilon|x'}$ *is identified.*

In this lemma's proof, we leverage Assumption 5 and the special regressors $\widetilde{X}$ to trace the distribution $F_{\varepsilon|x'}$. Given Assumption 5, this step is inspired by arguments in the literature on binary and multinomial choice without bundles.

[10] We have written $\widetilde{x}_i$ as entering additively. By changing its units (multiplying by -1), the true underlying regressor could enter negatively as well.

[11] In the context of binary choice, Magnac and Maurin (2007) show that identification under the mean independence assumption is sensitive to the large support assumption. They suggest alternative assumptions, which we could presumably adapt here. More generally, many advances from the binary and multinomial choice literatures could be used in the discrete games setting.

Lemma 2. *If Assumptions 1–3 and 4 are satisfied and if $F_{\varepsilon|x'}$ is identified, then the stand-alone utility functions and interaction terms $\left((u_i)_{i\le n}, \Phi\right)$ are also identified.*

The vector $\triangle W(a', x)$ affects the value of equilibrium probabilities $P\left(\triangle W(a', x) \mid x; F_{\varepsilon|x'}\right)$. The second lemma argues that different parameter values give different $\triangle W(a', x)$ and, by Assumption 4, different probabilities $\Pr(a \mid x)$. On its surface, this lemma is also analogous to arguments in the literature on binary and multinomial choice without bundles (e.g., Matzkin 1993). However, the arguments in the case of discrete games (and later, multinomial choice over bundles) are more complex for one key reason. In multinomial choice without bundles, under Assumption 4, the equivalent to $P\left(\triangle W(a', x) \mid x; F_{\varepsilon|x'}\right)$ is strictly monotone in *each* element of the equivalent to the vector $\triangle W(a', x)$. In our setting, this does not necessarily hold, which can be seen best by example.

Example 3. Consider the two-player game from the previous example. Let $\Phi(x') < 0$ for all x'. Say we are calculating the probability of the outcome $a' = (0, 0)$. This occurs when the following inequalities hold

$$
\begin{aligned}
u_1(x') + \widetilde{x}_1 + \varepsilon_1 &< 0 \\
u_2(x') + \widetilde{x}_2 + \varepsilon_2 &< 0 \\
u_1(x') + \widetilde{x}_1 + \varepsilon_1 + u_2(x') + \widetilde{x}_2 + \varepsilon_2 + \Phi(x') &< 0.
\end{aligned}
$$

Because $\Phi(x') < 0$, when the first two inequalities hold then the last one is always true. Therefore, the probability of the outcome $a = (0, 0)$ will not be strictly monotone in one of the elements of $\triangle W((0,0), x)$, i.e., $u_1(x') + \widetilde{x}_1 + \varepsilon_1 + u_2(x') + \widetilde{x}_2 + \varepsilon_2 + \Phi(x') - 0$. In this case, the empirical moments $\Pr((0,0) \mid x)$ for all x will not identify the interaction effects. The proof of Theorem 1 states that some other action profile a can be found such that $\Pr(a \mid x)$ identifies the interaction effects. □

By Assumption 4, $P\left(\triangle W(a', x) \mid x; F_{\varepsilon|x'}\right)$ is strictly monotone in the set of inequalities that are not implied by other inequalities. Our identification proof exploits the fact that this set of inequalities depends only on the values of the interaction effects $\Phi(x')$, not on the realization of the unobservables.

This paper does not explore estimation. However, $P\left(\triangle W\left(a',x\right) \mid x; F_{\varepsilon|x'}\right)$ is the key ingredient into the likelihood as a function of the infinite-dimensional objects in the structure $\Pi = \left(\left(u_i\right)_{i\leq n}, \Phi, F_{\varepsilon|x'}\right)$. In likelihood-based estimation, sieve approximations would likely be necessary for each of the infinite-dimensional objects (Chen 2007). More practically, given a limited sample size a researcher might use a flexible functional form for only the important unknown functions in the model. The most important unknown functions could be the interaction effects and the distribution of unobservables, if distinguishing these two explanations for correlated actions in the data is a key empirical objective.

In some cases, players will be anonymous, meaning that player indices i have no common meaning across markets. In a study of smoking peer effects, we might have five friends in each group of such friends. In an entry application without a focus on chains, we might have two potential entrants. In these cases, it is reasonable to impose that the function $u_i\left(x'\right)$ actually does not vary by the index of the agent i. Further, the interaction effects $\Phi\left(S,x\right)$ should depend only on the number of players taking the action, not the identity of the other players. Finally, the random vector ε should be exchangeable, conditional on x', in the indices of the players. All these restrictions are special cases of the general result in Theorem 1.

In a similar model (demand for bundles), Gentzkow (2007b, Proposition 2) exploits the linearity of utilities in the special regressors in order to analyze the counterfactual comparative statics of action probabilities in covariates. The same approach to comparative statics can be used here upon identification of the structure Π. The next section proceeds in the opposite direction: we use monotone comparative statics to facilitate identification. More precisely, in the two-player case we use how action probabilities vary with the special regressors to learn the sign of the interaction effect.

4.2 Sufficient Conditions for Assumption 5

4.2.1 Two-Player Games

The next lemma states that Assumption 5 always holds in two-player games.

Lemma 3. *If Assumptions 1–4 are satisfied and* $\mathcal{N} = \{1,2\}$*, then (i) the sign of* $\Phi\left(x'\right)$ *is* *identified at all* x'*; and (ii) Assumption 5 holds with* $\widehat{a} = (1, 0)$ *or* $\widehat{a} = (0, 1)$ *when* $\Phi(\widehat{x}') \geq 0$ *and* $\widehat{a} = (0, 0)$ *or* $\widehat{a} = (1, 1)$ *when* $\Phi(\widehat{x}') \leq 0$.

The proof of Lemma 3 in the appendix involves two steps. We first show (using monotone comparative statics techniques) that the sign of $\Phi(x')$ is identified from data. We then prove that this information is sufficient for Assumption 5 to hold. To shed light on this idea, recall that by Example 1, if Γ is a two-player game we can write its potential function as

$$V(a, x, \varepsilon) = (u_1(x) + \varepsilon_1) 1(a_1 = 1) + (u_2(x) + \varepsilon_2) 1(a_2 = 1) + \Phi(x') 1(a_1 = 1, a_2 = 1).$$

If $\Phi(x) \geq 0$, then V is supermodular in (a_1, a_2). Under Assumption 1, V has increasing differences in both $(a_1, \widetilde{x}_1)$ and $(a_2, \widetilde{x}_1)$. Then, by Topkis' theorem (see the appendix),

$$\arg\max \left\{ V(a, x, \varepsilon) \mid a \in \{0, 1\}^2 \right\}$$

increases in $\widetilde{x}_1$ (with respect to the strong set order). By Assumption 3(i), ε is independent of $\widetilde{x}_1$. It follows that $\Pr(a_1, a_2 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ increases with respect to stochastic dominance in $\widetilde{x}_1$. In turn, this means that $\Pr(a_1 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ and $\Pr(a_2 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ increase in $\widetilde{x}_1$. The same holds with respect to $\widetilde{x}_2$.

Alternatively, if $\Phi(x') \leq 0$, then V is supermodular in $(a_1, -a_2)$ and (under Assumption 1) it has increasing differences in both $(a_1, \widetilde{x}_1)$ and $(-a_2, \widetilde{x}_1)$. Using the previous argument, we get that $\Pr(a_1 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ and $\Pr(a_2 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ increase and decrease in $\widetilde{x}_1$, respectively. In this case, the opposite holds regarding $\widetilde{x}_2$.

The last two results show that $\widetilde{x}_1$ affects $\Pr(a_2 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ in opposite directions depending on the sign of $\Phi(x')$. Thus, the sign of the interaction term can be identified from data. With this information, it is easy to check that Assumption 5 holds with $\widehat{a} = (1, 0)$ or $\widehat{a} = (0, 1)$ when $\Phi(x') \geq 0$, and $\widehat{a} = (0, 0)$ or $\widehat{a} = (1, 1)$ when $\Phi(x') \leq 0$.

4.2.2 Negative Interaction Effects

The second sufficient condition we propose relies on strategic substitutabilities among players. The next lemma shows that Assumption 5 holds if $V(a, x, \varepsilon)$ has the negative single-crossing property in $(a_i; a_{-i})$ for all $i \in \mathcal{N}$.

Lemma 4. *Assume* $V(a,x,\varepsilon)$ *has the negative single-crossing property in* $(a_i;a_{-i})$ *for all* $i\in\mathcal{N}$, *i.e., for all* $a_i'>a_i$ *and all* $a_{-i}'>a_{-i}$ *(in the coordinatewise order) we have*

$$V(a_i',a_{-i},x,\varepsilon)-V(a_i,a_{-i},x,\varepsilon)\leq(<)\,0\Longrightarrow V\left(a_i',a_{-i}',x,\varepsilon\right)-V\left(a_i,a_{-i}',x,\varepsilon\right)\leq(<)\,0.$$

Then Assumption 5 holds with $\widehat{a}=(0,0,...,0)$ *for all* $\widehat{x}'$.

In this case, identification of $F_{\varepsilon|x'}$ will use data on the fraction of markets or groups where all agents take action 0, i.e., $\Pr((0,0,...,0)\mid x)$. The single-crossing condition in Lemma 4 would be satisfied if, for instance, Γ were a game with negative interactions, i.e., a submodular game. This restriction is often assumed in the entry games often estimated in industrial organization.

4.2.3 Concavity for Discrete Domains

Potential games that admit a concave potential function have been widely studied in economics. The reason is that when the potential function has this property, then the set of maximizers coincides with the equilibrium set of the underlying game. Ui (2008) introduces the concept of discrete concavity for potential games with discrete strategy spaces. His condition guarantees that local optimality of a vector of choices implies global optimality. The next lemma is based on Ui (2008, Proposition 1).

Lemma 5. *Assume that, for all* $a,a'\in\{0,1\}^n$ *with* $\|a-a'\|=2$,

$$\max_{a'':\|a-a''\|=\|a'-a''\|=1}V(a'',x,\varepsilon)\begin{cases}>\min\{V(a,x,\varepsilon),V(a',x,\varepsilon)\} & \text{if } V(a,x,\varepsilon)\neq V(a',x,\varepsilon)\\ \geq V(a,x,\varepsilon)=V(a',x,\varepsilon) & \text{otherwise}\end{cases}.$$

Then Assumption 5 holds with any $\widehat{a}\in\{0,1\}^n$ *for all* $\widehat{x}'$.

Concavity imposes non-trivial restrictions on the cross effects of multivariate functions and, in our model, the support of unobservables and regressors. In particular, if the unobservables and regressors are allowed to take values on the entire real line, then the interaction effects need to be identically zero for concavity to hold globally. Thus, this approach is non-trivial only if the large support portions of Assumptions 2 and 4 are simultaneously relaxed.

For this reason, we cannot recommend basing identification off of concavity explicitly. However, one can view Assumption 5 as a weaker (local) version of concavity that is compatible with our other restrictions.

5 Two-Player Games Without Selection Rules

The potential equilibrium selection rule for point identification in binary action games is particularly relevant to the case of three or more players. Here we provide a proof of identification in two-player games that does not impose any equilibrium selection rule. As Bajari et al. (2010) point out, the data generating process is now a mixture over equilibria in the regions of ε under which there is more than one equilibrium. As the literature (Tamer 2003, etc.) has exploited for the parametric case, our identification result will rely on the fact that some outcomes only occur as unique equilibria in the two player case.

As we do not impose the potential selection rule, we do not require that interaction effects have the pairwise symmetry property described above. We allow the interaction effect for husbands' actions on wives' utilities to differ from that of wives' actions on husbands' utilities. Berry and Tamer (2006, Result 4), following Tamer (2003), study identification in two-player, submodular games where the key special regressors enter multiplicatively instead of additively. The advantage of our approach is that it does not rely on identification at infinity to recover the payoff terms. As it was earlier defined, we let $v_i(x')$ be the interaction effect of the other player's action on i's utility.

Theorem 2. *In the two-player game, assume that players follow a pure-strategy equilibrium (but not necessarily a potential function maximizer) and that the game is either supermodular or submodular. Under Assumptions 1–4,* $\Pi = \left((u_i, v_i)_{i\leq 2}, F_{\varepsilon|x'}\right)$ *is identified.*

The full proof is in the appendix. Here we provide a sketch of the proof for the supermodular case. When $v_i \geq 0$ for $i = 1, 2$, then $(1, 0)$ and $(0, 1)$ are unique equilibria. The probability that the first player chooses action 1 and the other one action 0 (given x) is given by

$$\Pr((1,0) \mid x) = \Pr\left(-\varepsilon_1 - u_1(x') \leq \widetilde{x}_1, \varepsilon_2 + u_2(x') + v_2(x') \leq -\widetilde{x}_2 \mid x\right).$$

Given that ε has a zero mean, we can use variation in $\widetilde{x}$ to identify $F_{-\varepsilon_1,\varepsilon_2|x'}$. The subutilities $u_1(x')$ and $u_2(x')+v_2(x')$ can then be identified via conditional expectations using variation in x'. Similarly, we can use $\Pr((0,1) \mid x)$ to recover $u_2(x')$ and $u_1(x')+v_1(x')$. Therefore $(u_i,v_i)_{i\leq 2}$ is identified. The distribution of unobservables can be recovered from $F_{-\varepsilon_1,\varepsilon_2|x'}$. A similar argument follows when the interactions are negative.

The nonparametric identification result in Theorem 2 relies on knowing the signs of the interaction terms, but is independent of the equilibrium selection rule that guides players choices. Thus, the two-player game offers a very natural environment to empirically test competing theories about equilibrium selection, e.g., risk dominance versus payoff dominance.

6 Demand for Bundles

Our model for potential games is formally equivalent to a multinomial choice demand model where consumers can pick bundles of individual alternatives. We draw out the equivalence here and therefore present new identification results for multinomial choice models where consumers can purchase bundles of goods.

Think of $\mathcal{N}$ as the set of available products for a consumer, and let us denote by $\widetilde{x} \equiv (\widetilde{x}_i)_{i\leq n}$ the vector of (negative) prices. Prices will play the role of our special regressors in Assumption 1. We need prices to vary across consumers in the sample. This could occur from observing consumers in different markets or at different points in time. We could use other special regressors than prices. For example, if a consumer is choosing between a set of medical offices (say a hospital and an out-patient clinic), each $\widetilde{x}_i$ could be the distance between the consumer's home and provider i. If the consumer's utility function is quasilinear in income, his or her maximization problem can be written as

$$\max\left\{V(a,x,\varepsilon)=\sum\nolimits_{i\in\mathcal{N}}\left(u_i(x')+\widetilde{x}_i+\varepsilon_i\right)+\sum_{S\in\mathcal{S}(a)}\Phi(S,x') \mid a\in\{0,1\}^n\right\}. \tag{3}$$

This is a special case of (2). Here $u_i(x')+\widetilde{x}_i+\varepsilon_i$ is the stand-alone utility the consumer gets from acquiring good i and $\Phi(S,x')$ is an interaction term that captures complementarities or substitutabilities among the subset of goods S. The model restricts unobserved heterogeneity to happen at the goods level; we do not allow unobservables in the $\Phi(S, x')$ terms. Because this is a special case of the potential game studied earlier, the following corollary to Theorem 1 holds.

Corollary 1. *Let there be two goods and let Assumptions 1–4 hold. Then* $\Pi = \left((u_i)_{i \leq 2}, \Phi, F_{\varepsilon|x'}\right)$ *is identified.*

Let there be three or more goods and let Assumptions 1–5 hold. Then $\Pi = \left((u_i)_{i \leq n}, \Phi, F_{\varepsilon|x'}\right)$ *is identified.*

The various sufficient conditions for Assumption 5 can be straightforwardly adapted to the demand problem; we emphasize the two goods case in the statement of the corollary. The other possible sufficient conditions are that all goods are substitutes or that utility is discrete concave. There is no need to refer to Theorem 2 here because equilibrium selection is not relevant in consumer choice. Note also that there is no need to generalize the model to allow, for say the two good case, $v_1(x') \neq v_2(x')$. There are no asymmetric complementarities in the demand model. We do not discuss price endogeneity here, although this is important in many empirical applications. Various techniques in the literature can be used to resolve price endogeneity in multinomial choice models (e.g., Berry and Haile 2010, Fox and Gandhi 2012).

Gentzkow (2007a) studies a demand model for print and online newspapers with bivariate normal unobservables. We offer a nonparametric version of his result. Importantly, like his paper we allow for correlation in the unobservable tastes for the two goods. We can therefore distinguish complementarities in utility from correlated preferences for products.

Similar conditions to our sufficient conditions for Assumption 5 have been used in tangentially related identification papers. For example, Berry, Gandhi and Haile (2011) provide a sufficient condition under which a system of demand functions (representing, say, aggregate demand) can be inverted in unobservables. Their condition applies to goods that are (connected) substitutes and is loosely related to our result. Matzkin (2007) studies the inversion of simultaneous equations (again representing, say, aggregate demand), imposing (quasi-)concavity as we do in Section 4.2.3.

For another example, we can think of $\mathcal{N}$ as the set of available practices a firm can use. The problem of the firm is to select the subset of activities that maximize its objective function (3), where $u_i(x') + \widetilde{x}_i + \varepsilon_i$ captures the marginal contribution of each activity and $\Phi(S, x')$ is the extra output the firm gets by using the activities in the subset S together. Athey and Stern (1998) provide a framework for testing theories about complementarity in organizational practices within a cross section of firms. They address identification for two specifications of the model that differ regarding the type of heterogeneity across firms: the Random Systems Model (RSM) and the Random Practice Model (RPM). The RSM allows unobserved heterogeneity across firms for both the marginal contribution of each activity and the interaction terms. The specification of the RPM is similar to (3). Our approach to identification and specification of the utility function differs from theirs in that we impose more conditions on payoffs but rely on fewer exclusion restrictions. One way to achieve identification in a model where a consumer can pick bundles of choices is to redefine the alternatives by incorporating in the choice set all combinations of individual options. Doing this, Athey and Stern provide an identification argument that relies on exclusion restrictions for each bundle of choices (i.e., the enlarged choice set). In the demand setting, this would require bundle-specific prices that flexibly vary across consumers or markets. Doing this, one can appeal to nonparametric identification results from the literature on multinomial choice by treating each bundle as a separate good. It is only because we do not rely on bundle-specific exclusion restrictions that our identification results are new.

7 Conclusion

We explore identification in binary choice games of complete information. We derive conditions under which a binary choice game is a potential game and impose as the equilibrium selection rule that the selected equilibrium maximizes the resulting potential function. This makes our game formally equivalent to a multinomial choice demand model where a consumer can elect to purchase any bundle of products.

We show that the model is identified. We recover from data the subutility function of each player, the interaction effects among each group of players, and the joint distribution of potentially correlated, player-specific unobservables. We state some alternative key assumptions: two players, a submodular game, and concavity of the game. These conditions also have analogs for demand for bundles. For the two-player case, we present a separate identification result that does not rely on equilibrium selection.

It is likely our results will lead to nonparametric identification outcomes in other classes of potential games. For example, Monderer and Shapley (1996) and Qin (1996) discuss the connection between noncooperative potential games and cooperative games. It follows from their results that many cooperative solution concepts can be expressed as in (2), and therefore our previous results can be applied to these types of interactions.

A Proofs

A.1 Proof of Theorem 1

We divide the proof in the next two Lemmas.

A.1.1 Proof of Lemma 1 (Identification of $F_{\varepsilon|x'}$)

By Assumption 5, for each x', there exists a known vector $\widehat{a} \in \{0,1\}^n$ for which for any $x = (x', \widetilde{x})$ and ε, $V(a, x, \varepsilon)$ is maximized at $\widehat{a}$ if $V(\widehat{a}_i, \widehat{a}_{-i}, x, \varepsilon) \geq V(a'_i, \widehat{a}_{-i}, x, \varepsilon)$ for all $i \in \mathcal{N}$ and $a'_i = 1 - \widehat{a}_i$. Under Assumption 1, this condition holds if, for all $i \in \mathcal{N}$,

$$\begin{aligned}&(1(\widehat{a}_i = 1) - 1(\widehat{a}_i = 0))\,\varepsilon_i \geq \\ &\sum\nolimits_{S \in \mathcal{S}(a')} \Phi(S, x') - \sum\nolimits_{S \in \mathcal{S}(\widehat{a})} \Phi(S, x') - (1(\widehat{a}_i = 1) - 1(\widehat{a}_i = 0))(u_i(x') + \widetilde{x}_i),\end{aligned} \tag{4}$$

where a' is obtained from $\widehat{a}$ by changing only $\widehat{a}_i$. By Assumptions 2 and 3, we will show that we can recover $F_{\varepsilon|x'}$ from variation in $\widetilde{x}$ using $\Pr(\widehat{a} \mid x', \widetilde{x})$. Applying Assumption 5 shows that

$$\Pr(\widehat{a} \mid x', \widetilde{x}) = \Pr((4) \text{ holds for all } i \in \mathcal{N} \mid x', \widetilde{x}).$$

Let the new random variable μ_i for each $i \in \mathcal{N}$ be

$$\begin{aligned}\mu_i = {} & (1(\widehat{a}_i = 1) - 1(\widehat{a}_i = 0))\,\varepsilon_i \\ & - \left(\sum\nolimits_{S \in \mathcal{S}(a')} \Phi(S, \widehat{x}') - \sum\nolimits_{S \in \mathcal{S}(\widehat{a})} \Phi(S, \widehat{x}')\right) + (1(\widehat{a}_i = 1) - 1(\widehat{a}_i = 0))(u_i(x')).\end{aligned}$$

Let $\mu = (\mu_1, \ldots, \mu_n)$, which is independent of $\widetilde{x}$ conditional on x'. Therefore,

$$\Pr\left(\widehat{a} \mid x', \widetilde{x}\right) = \Pr\left(\mu_i \geq \left(1\left(\widehat{a}_i = 1\right) - 1\left(\widehat{a}_i = 0\right)\right)\left(\widetilde{x}_i\right) \text{ for all } i \in \mathcal{N} \mid x', \widetilde{x}\right).$$

Therefore, we identify the upper probabilities of the vector μ, conditional on x', at all points

$$\left(\left(1\left(\widehat{a}_1 = 1\right) - 1\left(\widehat{a}_1 = 0\right)\right)\left(\widetilde{x}_1\right), \ldots, \left(1\left(\widehat{a}_n = 1\right) - 1\left(\widehat{a}_n = 0\right)\right)\left(\widetilde{x}_n\right)\right).$$

By Assumption 2 and the fact that $\left(1\left(\widehat{a}_n = 1\right) - 1\left(\widehat{a}_n = 0\right)\right)$ is at most a sign change,

$$\left(\left(1\left(\widehat{a}_1 = 1\right) - 1\left(\widehat{a}_1 = 0\right)\right)\left(\widetilde{X}_1\right), \ldots, \left(1\left(\widehat{a}_n = 1\right) - 1\left(\widehat{a}_n = 0\right)\right)\left(\widetilde{X}_n\right)\right)$$

has support on all of $\mathbb{R}^n$. Therefore, we learn the upper tail probabilities of μ conditional on x' for all points of evaluation μ^*. Upper tail probabilities completely determine a random vector's distribution, so we also identify the lower tail probabilities of μ conditional on x', also known as the joint CDF of μ conditional on x'. Note that ε_i is the only random variable in μ_i, conditional on x'. By Assumption 3(ii), $\mathrm{E}(\varepsilon \mid x') = 0$. Therefore, up to the possible sign change in $\left(1\left(\widehat{a}_i = 1\right) - 1\left(\widehat{a}_i = 0\right)\right)$, the distribution of ε conditional on x' is obtained from the distribution of $\mu - E\left(\mu \mid x'\right)$ conditional on x'. □

A.1.2 Proof of Lemma 2 (Identification of $\left((u_i)_{i \leq n}, \Phi\right)$

Under Assumption 1

$$V\left(a, x, \varepsilon\right) \equiv \sum\nolimits_{i \in \mathcal{N}} \left(u_i\left(x'\right) + \widetilde{x}_i + \varepsilon_i\right) 1\left(a_i = 1\right) + \sum\nolimits_{S \in \mathcal{S}(a)} \Phi\left(S, x'\right).$$

To facilitate the exposition, recall that

$$\triangle^j \varepsilon\left(a'\right) \equiv \sum\nolimits_{i \in \mathcal{N}} \varepsilon_i 1\left(a'_i = 1\right) - \sum\nolimits_{i \in \mathcal{N}} \varepsilon_i 1\left(a^j_i = 1\right) \text{ and } \triangle^j W\left(a', x\right) \equiv W\left(a^j, x\right) - W\left(a', x\right).$$

Thus $\triangle\varepsilon(a')$ is the vector $(\triangle^j\varepsilon(a'))_{j\leq 2^n, a^j\neq a'}$, which has a distribution $dF_{\triangle\varepsilon(a')|x'}$. Then,

$$\begin{aligned}&P\left(\triangle W(a',x)\mid x;F_{\varepsilon|x'}\right)\\&\equiv\int\ldots\int 1\left(\triangle^1\varepsilon(a')\leq\triangle^1 W(a',x)\right)\ldots 1\left(\triangle^{2^n}\varepsilon(a')\leq\triangle^{2^n}W(a',x)\right)dF_{\triangle\varepsilon(a')|x'}.\end{aligned}$$

Each $\left((u_i)_{i\leq n},\Phi\right)$ clearly leads to a different vector of functions $W(a,x)$. Therefore, consider $\widetilde{W}\neq W$. We first pick an x where the subvector $\widetilde{x}$ will satisfy certain properties that we delay discussion of until the end of the proof. Let

$$C\equiv\arg\max_a\left\{\left(W(a,x)-\widetilde{W}(a,x)\right)\mid a\in\{0,1\}^n\right\}$$

and suppose $\max_a\left(W(a,x)-\widetilde{W}(a,x)\right)>0$ (the other case follows by a similar argument). Define $D\equiv\{a\notin C\mid a\in\{0,1\}^n\}$. We know $C\neq\varnothing$. The fact that $D\neq\varnothing$ follows as

$$W(a=(0,0,...,0),x)=\widetilde{W}(a=(0,0,...,0),x)=0.$$

Fix some $a'\in C$. We know that $W(a',x)-\widetilde{W}(a',x)=W(a,x)-\widetilde{W}(a,x)$ for all $a\in C$, and $W(a',x)-\widetilde{W}(a',x)>W(a,x)-\widetilde{W}(a,x)$ for all $a\in D$. Rearranging terms,

$$\begin{aligned}&W(a',x)-W(a,x)=\widetilde{W}(a',x)-\widetilde{W}(a,x)\text{ for all }a\in C,\text{ and}\\&W(a',x)-W(a,x)>\widetilde{W}(a',x)-\widetilde{W}(a,x)\text{ for all }a\in D.\end{aligned}$$

Assume there exists some $a''\in D$ for which $W(a',x)\geq W(a'',x)$ does not necessarily hold whenever $W(a',x)\geq W(a,x)$ for all $a\neq a',a''$ (see Example 3). Saying differently, knowing that a' maximizes $W(a,x)$ over all $a\neq a',a''$ is not enough to conclude that $W(a',x)\geq W(a'',x)$. We show this always holds below. Then, by Assumption 4

$$\begin{aligned}&\int\ldots\int 1\left(\triangle^1\varepsilon(a')\leq\triangle^1 W(a',x)\right)\ldots 1\left(\triangle^{2^n}\varepsilon(a')\leq\triangle^{2^n}W(a',x)\right)dF_{\triangle\varepsilon(a')|x'}\\&>\int\ldots\int 1\left(\triangle^1\varepsilon(a')\leq\triangle^1\widetilde{W}(a',x)\right)\ldots 1\left(\triangle^{2^n}\varepsilon(a')\leq\triangle^{2^n}\widetilde{W}(a',x)\right)dF_{\triangle\varepsilon(a')|x'}\end{aligned}$$

i.e., $P\left(\triangle W\left(a',x\right) \mid x; F_{\varepsilon|x'}\right) > P\left(\triangle \widetilde{W}\left(a',x\right) \mid x; F_{\varepsilon|x'}\right)$. Thus W identified, and $\left(\left(u_i\right)_{i\leq n}, \Phi\right)$ can be recovered directly from these action probabilities.

We now prove that there exists some $a' \in C$ and $a'' \in D$ for which $W\left(a',x\right) \geq W\left(a'',x\right)$ is not implied by $W\left(a',x\right) \geq W\left(a,x\right)$ for all $a \neq a', a''$. By contradiction, we will show that if this were not true then $(0,0,...,0) \in C$, which is not possible as $\max_a \left(W\left(a,x\right) - \widetilde{W}\left(a,x\right)\right) > 0$.

For each $a' \in C$ there are (at least) n inequalities that are not implied by the others. These inequalities correspond to vectors of actions that differ from a' regarding the action of one single player. To shed light on this point, let a'' be equal to a' except for some player i who selects action 1 at a' and action 0 at a''. Then $W\left(a',x\right) \geq W\left(a'',x\right)$ if and only if

$$u_i\left(x'\right) + \widetilde{x}_i + \sum\nolimits_{S\in\mathcal{S}(a')} \Phi\left(S,x'\right) \geq \sum\nolimits_{S\in\mathcal{S}(a'')} \Phi\left(S,x'\right).$$

All other inequalities, $W\left(a',x\right) \geq W\left(a,x\right)$ with $a \neq a', a''$, will involve at least one other special regressor $\widetilde{x}_j$ with $(j \neq i)$. Thus, we can always find a vector $\left(\widetilde{x}_i\right)_{i\leq n}$ such that $W\left(a',x\right) \geq W\left(a,x\right)$ with $a \neq a', a''$ and yet $W\left(a',x\right) < W\left(a'',x\right)$. Thus assume $a'' \in C$. By repeating this process $\|a'\|$ times, we will need to assume $(0,0,...,0) \in C$. But this is not possible as we explained before. □

A.2 Proof of Lemma 3

The proof of Lemma 3 relies on Topkis' theorem (see, e.g., Topkis (1998)) and the concept of stochastic dominance. We include a simple version of these concepts next.

Proposition 2. ***Topkis' theorem:*** *Let $f\left(a_1,a_2,x\right) : \mathcal{A}_1 \times \mathcal{A}_2 \times \mathbb{R} \rightarrow \mathbb{R}$, where $\mathcal{A}_1$ and $\mathcal{A}_2$ are finite ordered sets. Assume that $f\left(a_1,a_2,x\right)$ (i) is supermodular in $\left(a_1,a_2\right)$; and (ii) has increasing differences in $\left(a_1,x\right)$ and $\left(a_2,x\right)$. Then,* $\arg\max\left\{f\left(a_1,a_2,x\right) \mid \left(a_1,a_2\right) \in \mathcal{A}_1 \times \mathcal{A}_2\right\}$ *increases in x with respect to the strong set order.*[12] *(According to this order, we write $\mathcal{A} \geq_S \mathcal{B}$ if for every $a \in \mathcal{A}$ and $b \in \mathcal{B}$, we have that $a \vee b \in \mathcal{A}$ and $a \wedge b \in \mathcal{B}$.)*

[12] For any two elements $a, a' \in \mathcal{A}_1 \times \mathcal{A}_2$ we write $a \vee a'$ $\left(a \wedge a'\right)$ for the least upper bound (greatest lower bound). We say $f\left(a_1,a_2,x\right)$ is supermodular in $\left(a_1,a_2\right)$ if, for all $a, a' \in \mathcal{A}_1 \times \mathcal{A}_2$,

$$f\left(a \vee a', x\right) + f\left(a \wedge a', x\right) \geq f\left(a,x\right) + f\left(a',x\right).$$

The concept of first order (or standard) stochastic dominance (FOSD), is based on upper sets. Let us consider $(\Omega, \geq)$, where Ω is a set and $\geq$ defines a partial order on it. A subset $U \subset \Omega$ is an upper set if and only if $x \in U$ and $x' \geq x$ imply $x' \in U$.

First Order Stochastic Dominance: Let $X', X \in \mathbb{R}^n$ be two random vectors. We say X' is higher than X with respect to first order stochastic dominance if

$$\Pr(X' \in U) \geq \Pr(X \in U)$$

for all upper set $U \subset \mathbb{R}^n$.

Proof of Lemma 3: **(i)** By Example 1 and Assumption 1, if Γ is a two-player game, then

$$V(a,x,\varepsilon) = (u_1(x') + \widetilde{x}_1 + \varepsilon_1)\,1(a_1 = 1) + (u_2(x') + \widetilde{x}_2 + \varepsilon_2)\,1(a_2 = 1) + \Phi(x')\,1(a_1 = 1, a_2 = 1).$$

If $\Phi(x') \geq 0$ then $V(a,x,\varepsilon)$ is supermodular in (a_1, a_2). In addition $V(a,x,\varepsilon)$ has increasing differences in $(a_1, \widetilde{x}_1)$ and $(a_2, \widetilde{x}_1)$ (here the cross partial derivative is 0). Then, by Topkis' theorem,

$$a^*(x', \widetilde{x}_1, \widetilde{x}_2, \varepsilon) \equiv \arg\max\left\{V(a,x,\varepsilon) \mid (a_1,a_2) \in \{0,1\}^2\right\}$$

increases in $\widetilde{x}_1$. By Assumption 4, the probability that the argmax is not unique is 0. By Assumption 3(i), the unobservables are independent of $\widetilde{x}_1$. Let $\widetilde{x}_1^* > \widetilde{x}_1$ and let U be an upper set in the space of pure-strategy equilibria, then

$$\begin{aligned}\Pr(a \in U \mid x', \widetilde{x}_1', \widetilde{x}_2) &= \Pr(a^*(x^*, \widetilde{x}_1', \widetilde{x}_2, \varepsilon) \in U \mid x', \widetilde{x}_1^*, \widetilde{x}_2) \\ &\geq \Pr(a^*(x', \widetilde{x}_1, \widetilde{x}_2, \varepsilon) \in U \mid x', \widetilde{x}_1, \widetilde{x}_2) = \Pr(a \in U \mid x', \widetilde{x}_1, \widetilde{x}_2),\end{aligned}$$

i.e., $\Pr(a_1, a_2 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ increases with respect to first order stochastic dominance in $\widetilde{x}_1$. Because first order stochastic dominance is preserved under marginalization, $\Pr(a_2 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ increases in $\widetilde{x}_1$ (see, e.g., Müller and Stoyan 2002, Theorem 3.3.10, p. 94). In a similar way,

We say $f(a_1, a_2, x)$ has increasing differences in (a_1, x) if, for all $a_1' > a_1$ and $x' > x$,

$$f(a_1', a_2, x') - f(a_1, a_2, x') \geq f(a_1', a_2, x) - f(a_1, a_2, x).$$

we can show that $\Pr(a_2 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ increases in $\widetilde{x}_2$.

On the other hand, $V(a, x, \varepsilon)$ is supermodular in $(a_1, -a_2)$ if $\Phi(x') \leq 0$. In addition, $V(a, x, \varepsilon)$ has increasing differences in $(a_1, \widetilde{x}_1)$ and $(-a_2, \widetilde{x}_1)$ (here the cross partial derivative is 0). By Assumption 4, the probability that the argmax is not unique is 0. By Assumption 3(i), the unobservables are independent of $\widetilde{x}_1$. Thus, $\Pr(a_1, -a_2 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ increases with respect to first order stochastic dominance in $\widetilde{x}_1$. Because first order stochastic dominance is preserved under marginalization, $\Pr(-a_2 = 0 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ increases in $\widetilde{x}_1$, and therefore $\Pr(a_2 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ decreases in $\widetilde{x}_1$. In a similar way, $\Pr(a_2 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ decreases in $\widetilde{x}_2$.

By Assumption 4, unless $\Phi(x') = 0$, neither $\Pr(a_2 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ nor $\Pr(a_1 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2)$ are constant as a function of $\widetilde{x}_1$ and $\widetilde{x}_2$, respectively. From the last two paragraphs, the sign of $\Phi(x')$ is identified from data.

(ii) For $\Phi(x') \geq 0$ let $\widehat{a}$, as defined in Assumption 5, be $(\widehat{a}_1, \widehat{a}_2) = (1, 0)$. We need to show that if $V((1,0), x, \varepsilon) \geq V((1,1), x, \varepsilon)$ and $V((1,0), x, \varepsilon) \geq V((0,0), x, \varepsilon)$ hold, then $V((1,0), x, \varepsilon) \geq V((0,1), x, \varepsilon)$. The first two conditions are satisfied if and only if

$$u_2(x') + \widetilde{x}_2 + \varepsilon_2 + \Phi(x') \leq 0 \tag{5}$$

$$u_1(x') + \widetilde{x}_1 + \varepsilon_1 \geq 0. \tag{6}$$

Since $\Phi(x') \geq 0$, then (5) implies $u_2(x') + \widetilde{x}_2 + \varepsilon_2 \leq 0$. By this observation and (6) we get that $u_1(x') + \widetilde{x}_1 + \varepsilon_1 \geq u_2(x') + \widetilde{x}_2 + \varepsilon_2$, i.e., $V((1,0), x, \varepsilon) \geq V((0,1), x, \varepsilon)$. A similar claim holds if we select $(\widehat{a}_1, \widehat{a}_2) = (0, 1)$.

For $\Phi(x') \leq 0$, Assumption 5 holds by selecting $(\widehat{a}_1, \widehat{a}_2) = (0, 0)$ or $(\widehat{a}_1, \widehat{a}_2) = (1, 1)$. The proof is almost identical so we omit it. □

A.3 Proof of Theorem 2

Assume $\Gamma(X, \varepsilon)$ is supermodular, i.e., $v_i(X) \geq 0$ for $i = 1, 2$. Under Assumption 4, when $(1, 0)$ is an equilibrium, it is the only equilibrium with positive probability. Thus, under Assumptions 1 and 4,

$$\Pr((1,0) \mid x) = \Pr(-\varepsilon_1 - u_1(x') \leq \widetilde{x}_1, \varepsilon_2 + u_2(x') + v_2(x') \leq -\widetilde{x}_2 \mid x).$$

By Assumption 3(i), ε is independent of $\widetilde{X}$. We can identify $F_{\alpha|x'}$ with

$$\alpha = (\alpha_1, \alpha_2) \equiv (-\varepsilon_1 - u_1(x'), \varepsilon_2 + u_2(x') + v_2(x'))$$

from variation in $\widetilde{x}$. Using the same logic, we can recover $F_{\beta|x'}$ with

$$\beta = (\beta_1, \beta_2) \equiv (\varepsilon_1 + u_1(x') + v_1(x'), -\varepsilon_2 - u_2(x')).$$

By Assumption 3(ii), ε is mean independent of X'. Thus,

$$\begin{aligned} \mathrm{E}[(\alpha_1, \alpha_2) \mid x'] &= (-u_1(x'), u_2(x') + v_2(x')) \\ \mathrm{E}[(\beta_1, \beta_2) \mid x'] &= (u_1(x') + v_1(x'), -u_2(x')) \\ \mathrm{E}[(\alpha_1, \alpha_2) \mid x'] + \mathrm{E}[(\beta_1, \beta_2) \mid x'] &= (v_1(x'), v_2(x')). \end{aligned}$$

Then $(u_i, v_i)_{i \leq 2}$ are identified from variation in x'. By Assumption 3(ii), ε has zero mean conditional on x'. Thus the previous information allows us to identify $F_{\varepsilon|x'}$.

We omit the proof for the submodular case, i.e., $v_i(X) \leq 0$ for $i = 1, 2$, as it is almost identical. □

A.4 Proof of Lemma 4

Assume $V(a, x, \varepsilon)$ has the negative single-crossing property on $(a_i; a_{-i})$ for all $i \in \mathcal{N}$, i.e., for all $a_i' > a_i$ and all $a_{-i}' > a_{-i}$ we have

$$V(a_i', a_{-i}, x, \varepsilon) - V(a_i, a_{-i}, x, \varepsilon) \leq (<) 0 \Longrightarrow V(a_i', a_{-i}', x, \varepsilon) - V(a_i, a_{-i}', x, \varepsilon) \leq (<) 0. \quad (7)$$

We next show that Assumption 5 holds with $\widehat{a} = (0, 0, ..., 0)$.

Assume $V(\widehat{a}_i = 0, \widehat{a}_{-i} = (0, 0, ..., 0), x, \varepsilon) \geq V(a_i = 1, \widehat{a}_{-i} = (0, 0, ..., 0), x, \varepsilon)$ for all $i \in \mathcal{N}$. Then, by (7), for all $a_{-i} \in \{0, 1\}^{n-1}$, and all $i \in \mathcal{N}$,

$$V(\widehat{a}_i = 0, a_{-i}, x, \varepsilon) \geq V(a_i = 1, a_{-i}, x, \varepsilon)$$

Thus, $V((0, .., 0), x, \varepsilon) \geq V(a, x, \varepsilon)$ for all $a \in \{0, 1\}^n$. □

A.5 Proof of Lemma 5

Assume the required conditions of the lemma hold. Then, by Ui (2008, Proposition 1), $V(\widehat{a}, X, \varepsilon) \geq V(a, X, \varepsilon)$ for all $a \in \{0,1\}^n$ with $\|\widehat{a} - a\| \leq 1$ is satisfied if and only if $V(\widehat{a}, X, \varepsilon) \geq V(a, X, \varepsilon)$ for all $a \in \{0,1\}^n$. Thus, Assumption 5 holds at any $\widehat{a} \in \{0,1\}^n$. $\square$

References

[1] Bajari, P., H. Hong, and S. Ryan (2010). Identification and estimation of a discrete game of complete information. Econometrica 78, 1529–1568.

[2] Berry, S. (1992). Estimation of a Model of Entry in the Airline Industry. Econometrica 60, 889-917.

[3] Berry, S., and P. Haile (2010). Nonparametric identification of multinomial choice demand models with heterogeneous consumers. Yale University working paper.

[4] Berry, S., A. Gandhi, and P. Haile (2011). Connected Substitutes and Invertibility of Demand. Yale University working paper.

[5] Berry, S., and E. Tamer (2007). Identification in Models of Oligopoly Entry. Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress, vol. 2.

[6] Bjorn, P., and Q. Vuong (1984). Simultaneous equations models for dummy endogenous variables: A game theoretic formulation with application to labor force participation. Mimeo.

[7] Blume, L. (1993). The statistical mechanics of strategic interaction. Games Econ. Behav. 5, 387–424.

[8] Bresnahan, T. and P. Reiss (1991). Entry and Competition in Concentrated Markets. Journal of Polical Economy, 99, 977-1009.

[9] Brock, W., and S. Durlauf (2001). Discrete choice with social interactions. Rev. Econ. Stud. 68, 235-260.

[10] Brock, W., and S. Durlauf (2007). Identification of binary choice models with social interactions. J. Econometrics 140, 52-75.

[11] Card, D. and L. Giuliano (2012), Peer Effects and Multiple Equilibria in the Risky Behavior of Friends. University of California, Berkeley working paper.

[12] Chen, R., and Y. Chen (2011). The potential of social identity for equilibrium selection. Amer. Econ. Rev. 101, 2562–2589.

[13] Chen, X. (2007). Large Sample Sieve Estimation of Semi-Nonparametric Models. Handbook of Econometrics.

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

[15] Cooper, R. , and A. John (1988). Coordinating coordination failures in Keynesian models. Quart. J. Econ. 103, 441-463.

[16] De Paula, A., and X. Tang. (2011). Inference of signs of interaction effects in simultaneous games with incomplete information. Econometrica 80, 143-172.

[17] Deb, R. (2009). A Testable Model of Consumption with Externalities. J. Econ. Theory. 144, 1804–1816

[18] Ellickson, P., S. Houghton and C. Timmins (2010). Estimating Network Economies in Retail Chains: A Revealed Preference Approach. Duke University working paper.

[19] Fox, J. and A. Gandhi (2012). Nonparametric Identification and Estimation of Random Coefficients in Multinomial Choice Models. University of Michigan working paper.

[20] Galichon, A. and M. Henry (2011). Set identification in models with multiple equilibria. Review of Economic Studies 78, 1264–1298.

[21] Gentzkow, M. (2007a). Valuing New Goods in a Model with Complementarity: Online Newspapers. The American Economic Review 97, 713-744.

[22] Gentzkow, M. (2007b). Valuing New Goods in a Model with Complementarity: Online Newspapers. University of Chicago working paper.

[23] Goeree, J., and C. Holt (2005). An experimental study of costly coordination. Games Econ. Behav. 51, 349–364.

[24] Gowrisankaran, G., and J. Stavins (2004). Network externalities and technology adoption: Lessons from electronic payments. RAND Journal of Economics 35, 260-276.

[25] Harsanyi, J., and S. Selten (1988). A General Theory of Equilibrium Selection in Games. MIT Press.

[26] Hartman, W. (2010). Demand estimation with social interactions and the implications for targeted marketing. Journal Marketing Science 29, 585-601.

[27] Heckman, J. (1978). Dummy Endogenous Variables in a Simultaneous Equation System. Econometrica 46, 931–959.

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

[29] Kajii, A., and S. Morris. (1997a). The robustness of equilibria to incomplete information. Econometrica, 65, 6, 1283-1309.

[30] Kajii, A., and S. Morris. (1997b). Refinements and Higher Order Beliefs: A Unified Survey. Northwestern University working paper.

[31] Lewbel, A. (1998). Semiparametric latent variable model estimation with endogenous or mismeasured regressors. Econometrica. 66, 105-121.

[32] Lewbel, A. (2000). Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables. 97, 145-177.

[33] Lewbel, A., and X. Tang (2011). Identification and estimation of games with incomplete Information using excluded regressors. Boston College working paper.

[34] Magnac, T. and E. Maurin (2007). Identification and information in monotone binary models. Journal of Econometrics. 139, 76–104.

[35] Manski, C (1988). Identification of Binary Response Models. Journal of the American Statistical Association. 83, 729-738.

[36] Matzkin, R. (1993). Nonparametric identification and estimation of polychotomous choice models. J. Econometrics 58, 137-168.

[37] Matzkin, R. (2007). Heterogeneous Choice. Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress of the Econometric Society.

[38] Mazzeo, M. (2002). Product Choice and Oligopoly Market Structure. RAND Journal of Economics 33, 221–242.

[39] Monderer, D., and L. Shapley (1996). Potential Games. Games Econ. Behav. 14, 124-143.

[40] Morris, S. and T. Ui (2005). Generalized potentials and robust sets of equilibria. Journal of Economic Theory 124, 45–78.

[41] Müller, A., and D. Stoyan (2002). Comparison Methods for Stochastic Models and Risks. John Wiley & Sons, Inc.

[42] Nelsen, R. (1971). Introduction to Copulas. Lecture Notes in Statistics: 139. Springer-Verlag New York, Inc.

[43] Qin, C. (1996). Endogenous formation of cooperation structures. J. Econ. Theory 69, 218-226

[44] Rosenthal, R. (1973). A class of games possessing pure-strategy Nash equilibria. Int. J. Game Theory 2, 65–67.

[45] Tamer, E. (2003). Incomplete Simultaneous Discrete Response Model with Multiple Equilibria. Review of Economic Studies 70, 147–165.

[46] Thompson, S. (1998). Identification of semiparametric discrete choice models. Mimeo.

[47] Topkis, D. M. (1998). Supermodularity and Complementarity. Princeton University Press.

[48] Ui, T. (2000). A Shapley value representation of potential games. Games Econ. Behav. 31, 121–135.

[49] Ui, T. (2001). Robust equilibria of potential games. Econometrica 69, 1373–1380.

[50] Ui, T. (2008). Discrete concavity for potential games. Int. J. Game Theory Review 10, 137-143.

[51] Van Huyck, J., R. Battalio and R. Beil. (1990). Tactic Coordination Games, Strategic Uncertainty, and Coordination Failure. American Economic Review 80, 234–248.