Identification of Discrete Choice Models for Bundles and Binary Games*

Jeremy T. Fox

University of Michigan and NBER[†]

Natalia Lazzati

University of Michigan[‡]

February 2013

Abstract

We study nonparametric identification of single-agent discrete choice models for bundles and binary games of complete information. We provide conditions under which we can recover both the interaction effects and the distributions of potentially correlated unobservables across goods in single-agent models and across players in games. We establish similarities in identification between these two models. Strengthening our assumptions for games, we provide an equivalence relation between discrete choice models for bundles and binary games that relies on the theory of potential games. Potential games are particularly useful for games of three or more players.

1 Introduction

1.1 Motivation

Single-agent discrete choice models for bundles and binary games play an important role in empirical work in industrial organization and a variety of other fields. Most studies on discrete choice models assume the available alternatives are mutually exclusive. That is, they model situations where an agent faces a set of alternatives and the acceptance of one alternative automatically excludes the other possibilities. We assume instead that the agent acquires the subset or bundle of the available alternatives that maximizes its expected payoff. Under the bundle specification, we show that single-agent discrete choice models and binary action games of complete information are quite similar from an identification perspective. Our identification proofs for the two models are nearly the same, for two goods and two players, respectively. In addition, we prove the two models are actually formally equivalent when attention is restricted to the class of potential games. Potential games are particularly useful for games of three or more players. Altogether, taking advantage of these connections this paper provides new identification results for discrete choice models and games.

As we just mentioned, potential games are at the heart of a formal equivalence between identification of discrete choice models for bundles and binary games. Monderer and Shapley (1996) define potential games as those games that admit an exact potential function. This potential function is defined over the set of action profiles of the players such that the function's maximizer is a pure strategy Nash equilibrium for the associated game. When a game admits a potential function, the function is uniquely defined up to an additive constant. Based on this observation, Monderer and Shapley (1996) argue that the maximizer of the potential offers, from a technical perspective, an equilibrium refinement. Recent theoretical and experimental work has provided economic justification to this equilibrium selection rule.[1] Contributing to this growing body of literature, we show that if the game admits a potential function and players coordinate on the potential maximizer, then the binary game is formally equivalent from an econometric standpoint to the discrete choice model for bundles. It follows that any identification result proved for one of these two models applies to the other one.

To the best of our knowledge, we are the first to provide nonparametric results for discrete choice models for bundles. In a recent study, Gentzkow (2007a) estimates a parametric probit model where a consumer can elect to consume no newspaper, a print newspaper, an online

---

[1]See Ui (2001) for a theoretical justification of the selection rule based on potential maximizers and Huyck et al. (1990), Goeree and Holt (2005) and Chen and Chen (2011) for experimental evidence.

newspaper, or both. In this setting, it is essential to distinguish true complementarities (or substitutabilities) in demand from correlated, unobserved heterogenous preferences for the goods. In the study, Gentzkow's data show that many consumers purchase both the print and online newspapers. This behavior may occur because the two goods are complements in consumption or because consumers with high unobserved stand-alone values for the print newspaper have high unobserved valuations for the online version as well. Our identification approach allows us to distinguish unobserved, correlated preferences from true interdependence of the alternatives in utility. Moreover, in cases where the alternatives are substitutes, our approach identifies a distribution of heterogeneous interaction effects, measuring differences across agents in the degree of substitutability.

In addition to providing insight about multinomial choice models, our findings can also inform studies of static, binary choice games of complete information.[2] This set of games includes games of product adoption with social interactions, entry games, and labor force participation games, among others. For the case of two players, our identification results do not rely on potential games and equilibrium selection rules. When the number of players is larger, we invoke the mentioned restrictions. In binary games, interaction effects capture the strategic interdependence of player decisions. As in the single-agent model, these effects are often hard to distinguish from unobserved, heterogeneous payoffs that are correlated across players, as these two sources of interdependence have similar observable implications. For example, similarities in smoking behavior among friends may be due to peer effects or due to correlated tastes for smoking. Likewise, firms may tend to enter geographic markets in clusters either because the competitive effects of entry are small or because of the unobservable profitability of certain markets. Though it is well known that not allowing for the correlation of unobservables can lead to biases in the estimates of interaction effects, much of the empirical literature on games assumes that any unobservables are independent of each other. By contrast, our identification approach can distinguish and separately identify these two sources of interdependence without relying on arguments based on identification at infinity.

The identification strategy we propose relies on exclusion restrictions at the level of each alternative (for the single-agent model) and for each player (for games). For the discrete choice model of bundles, these excluded regressors enter the standalone payoff of each alternative (additively separably) but not the interaction term determining whether a pair of alternatives are complements or substitutes. For games, these excluded regressors similarly affect the payoff

---

[2]See, e.g., Heckman (1978), Bjorn and Vuong (1984), Bresnahan and Reiss (1991a, 1991b), Berry (1992), Mazzeo (2002), Ciliberto and Tamer (2009), Hartmann (2009), Bajari et al. (2010), Beresteanu et al. (2011), Galichon and Henry (2011), and Card and Giuliano (2012).

of one player but not the others. These exclusion restrictions have been called special regressors in, for example, the literature on discrete choice without bundles (Lewbel 1998, 2000, Matzkin 2007, Berry and Haile 2010, Fox and Gandhi 2012). As we mentioned earlier, this approach allows us to nonparametrically identify the complete structure of both models: the utilities of all bundles as functions of covariates and the players' subutilities of choosing each action as functions of regressors and the actions of other players as well as, in both models, the distribution of potentially correlated unobservables. When there are two goods or players and the alternatives are substitutes or the game is submodular, we also identify a distribution of heterogeneous interaction terms.

Semiparametric methods that estimate interaction effects but not the distribution of unobservables do not allow for all the uses of structural models in empirical work. Consider the discrete choice model for bundles with prices of goods as regressors. A key use of the estimated model might be to predict demand under counterfactual prices. Without estimating the distribution of the unobservables, one cannot predict market shares, marginal effects of price changes and price elasticities. Our results identify the complete structure of each model and hence allow for such predictions.

1.2 Further Results and Literature Review

There is a small literature on the demand for bundles of discrete goods when those goods may have positive or negative interaction effects in utility (e.g., Gentzkow 2007a, Athey and Stern 1998).[3] Gentzkow uses a parametric, multivariate probit model. To our knowledge, we are the first to give conditions for nonparametric identification in models of demand for bundles, at least without relying on bundle-specific prices and treating each bundle as a separate good in the usual multinomial choice framework where a consumer picks only one product.[4] The sufficient conditions we propose to identify the distribution of the unobservables for models with three or more goods are economically similar to some of the requirements imposed in tangentially related identification papers. For example, Berry, Gandhi and Haile (2011) provide a sufficient

---

[3]Hendel (1999) and Dubé (2004) study models of demand for bundles where the bundles do not have interaction effects.

[4]Kim and Sher (2012) study identification where the researcher observes only the market share for each good, not market shares for each bundle. They assume that the true distribution of heterogeneity represents a finite number of consumers that is the same finite group across markets. They prove so-called generic identification, where counterexamples to identification are proved to have probability zero, if a distribution is placed over possible parameter values. This definition of identification is weak because, in other contexts, counterexamples to identification for parameters that researchers intuitively think are fundamentally unidentified may have probability zero.

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 demand when goods are substitutes. Matzkin (2008) studies the inversion of simultaneous equations in models with continuous outcomes, imposing quasi-concavity. We suggest a discrete notion of concavity.

We argue that identification in discrete choice models of bundles is quite similar to identification in binary games. For games of two players, our results extend those of Tamer (2003) and Berry and Tamer (2006, Result 4). Without assuming any equilibrium selection rule, both results, theirs and ours, require knowledge of the signs of the interactions terms. The advantage of our identification argument with respect to theirs is that we do not rely on identification at infinity to recover any of the unknown objects in the model. By identification at infinity, we mean recovering the interaction effects of one player by choosing values for special regressors of all other players in the range where all these other players have dominant strategies. Though conceptually simple, this identification technique is empirically unattractive because it relies mainly on extreme regressor values that are unlikely to appear in any finite data set. In other econometric contexts such as sample selection models, identification at infinity has been shown to result in slow rates of convergence (under certain tail thickness properties for the regressor and unobservables) for estimators based on the identification arguments (e.g., Andrews and Schafgans 1998 and Khan and Tamer 2010). Tamer (2003) and Berry and Tamer (2006, Result 4) do not consider heterogeneous interaction effects.

In games without an equilibrium selection rule, we do not show how to identify the signs of the interaction effects; they must be assumed. This is because identification relies on unique equilibria, equilibria that never occur as part of a set of multiple equilibria, and the identities of the unique equilibria depend on the signs of the interaction effects. In contrast, the sign of the interaction terms can be recovered from data on choices of bundles of at most two goods. The proof of this result invokes monotone comparative static techniques, and the result extends to potential games when players coordinate on potential maximizers.

Binary games with three or more players may typically have no unique equilibria. In addition, even for those that display such a feature (e.g., submodular games), we argue that the unique pure strategy equilibria are not enough to identify all the interaction effects, at least using the identification strategy we use for two players. So for the case of three or more players in games, we recommend imposing an equilibrium selection rule and exploiting the formal equivalence between potential games and models of demand for bundles.

We know of two simultaneous, independent papers on the identification of binary games of complete information. Neither paper discusses the link with single-agent discrete choice models

with bundles or potential games. For the two player case, Kline (2012) presents identification results that, like ours, do not rely on identification at infinity or on an equilibrium selection mechanism. In the case of three or more players, he uses identification at infinity to in effect turn the game into a two-player game. Identification at infinity does not recover the joint distribution of unobservables in games of three or more players unless extra assumptions are imposed on the copula of the multivariate distribution. Dunker, Hoderlein and Kaido (2012) study the two-player case and allow for random coefficients on all regressors, not just interaction effects, in a semiparametric specification. They also provide a computationally simple estimator for the density of random coefficients.

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 results for games depart from this literature in that our model is nonparametric and we show point rather than set identification. Given the natural trade-offs, our approach is more restrictive in other respects: we focus on the binary actions case (enter or not, smoke or not) and use the aforementioned special regressors. Binary and ordered games of incomplete information have received a lot of attention (e.g., Brock and Durlauf 2001, 2007, Aradillas-Lopez 2010, De Paula and Tang 2011, Lewbel and Tang 2011, Gandhi and Aradillas-Lopez 2012, Grieco 2012). As Bajari et al. (2010) explain, the challenges involved in complete information games are quite different. One advantage of the complete information setup is that players do not have ex-post regret: they do not prefer to take a different action after seeing the actions of their rivals. This is potentially unrealistic in models where the decision has longstanding consequences, such as entry. Another advantage of complete information with correlated unobservables over incomplete information games (without unobserved heterogeneity) is that the unobservables can reflect unmeasured characteristics of, say, the geographic market that are observed to the players.

Our discussion of how we avoid using identification at infinity is not meant to say that we do not rely on large support special regressors to identify the joint distribution of the unobservables. We do use those regressors just as they are used in the literature on discrete choice models without bundles to identify the joint distribution of the unobservables. We discuss how various results in that literature apply to the model for bundles and, hence, games, including some so-called mean independence assumptions that are particularly sensitive to values of the unobservables in the tails (e.g., Lewbel 1998, 2000; Magnac and Maurin 2007; Khan and Tamer 2010). It is important to differentiate our use of large support from identification at infinity. When applied to the model for bundles of goods, identification at infinity means that the agent would never consider certain goods. When applied to games, identification at infinity

identifies the payoff of a given player by making sure the other players are playing dominant strategies. That is, it identifies payoffs in the range of values of covariates where the game does not involve strategic interactions among players.

We do not formally explore estimation because establishing consistency for particular non-parametric estimators is somewhat straightforward conceptually once identification is established. Consider first models where the data generating process lacks multiple equilibria, such as discrete choice with bundles and potential games where players coordinate on the potential maximizer. Under the full independence of observable regressors and unobservables, 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). However, the results of Khan and Nekipelov (2012) on rates of convergence deserve some attention here. For semiparametric discrete games of complete information, Khan and Nekipelov establish that there is no estimator of the interaction effects that always converges at the parametric rate. They argue that the reason for the slow rates of convergence is the binary nature of the other player's action in a given player's best response, not the presence of multiple equilibria. Therefore, we also conjecture that there exists no estimator that converges at the parametric rate for semiparametric models of the demand for bundles and potential games. Khan and Nekipelov's impossibility result is not related to the issues with the use of identification at infinity in the previous literature on games. Our identification results for two-player binary games allow multiple equilibria; standard maximum likelihood is inconsistent even though we establish that the model is point identified. Modified likelihood estimators from the literature on multiple equilibria can be adapted instead (Bresnahan and Reiss 1991b, Berry 1992, Tamer 2003).

Section 2 describes our identification results for two-player games without imposing equilibrium selection rules. Section 3 presents identification results for discrete choice models for bundles with two goods. Section 4 introduces potential games and shows the equivalence of identification in such games to identification in discrete choice models for bundles. Section 5 extends our identification results to the case of three or more goods or players. Section 6 concludes, and all proofs and some basic results on comparative statics are collected in the Appendixes.

2 Two-Player Game

2.1 Game

We consider a simultaneous, discrete choice game with complete information. The set of players is $\{1, 2\}$. Each player i 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 (in the data) state variables (with x the realization of the random variable). We also let $\varepsilon \equiv (\varepsilon_1, \varepsilon_2) \in \mathbb{R}^2$ and $\eta = (\eta_1, \eta_2) \in I \subset \mathbb{R}^2$ indicate two vectors 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) $F_{\varepsilon|x}$; similarly η has the CDF $F_{\eta|x}$. Importantly, we do not assume that ε_1 and ε_2 are independent: we allow the unobservables to be correlated even after conditioning on the observable state variables. Likewise, we do not assume η_1 and η_2 are independent, but we will assume η and ε are independent.

The payoff of player i from choosing action 1 is

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

while the return from action 0, $U_{0,i}(a_{-i}, x, \varepsilon_i, \eta_i)$, is normalized to 0. The first element of (1), $u_i(x) + \varepsilon_i$, is the stand-alone value of action 1, and the interaction effect $v_i(x) + \eta_i$ captures the effect that the choice of the other player, a_{-i}, has on player i. The distinction between ε_i and η_i is that ε_i is an unobservable in standalone payoffs and η_i is an unobservable in interaction effects. We denote this game by $\Gamma(x, \varepsilon, \eta)$.[5]

A vector of decisions $a^* \equiv (a_1^*, a_2^*)$ is a pure strategy Nash equilibrium if, for $i = 1, 2$,

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

We write $\mathcal{D}(x, \varepsilon, \eta)$ for the equilibrium set (in pure strategies) of $\Gamma(x, \varepsilon, \eta)$.

The econometrician observes the vector of covariates and choices, $(x, (a_i)_{i \leq 2})$, for a cross section of independent two-player games that share the same structure $((u_i, v_i)_{i \leq 2}, F_{\varepsilon|x}, F_{\eta|x})$. These independent games may be different geographic markets in an entry application or different pairs of friends in a peer effects model. The objective of the analysis is to combine initial

---

[5]The additive separability of $u_i(x) + \varepsilon_i$ and $1(a_{-i} = 1)(v_i(x) + \eta_i)$ does not impose any restrictions on the model; it is just a convenient way to write the payoffs.

assumptions with data to learn about the structure of the game.

If the players are anonymous such that agent indices do not have common meaning across markets or plays of the game, then $u_1 = u_2$, $v_1 = v_2$, $F_{\varepsilon|x}$ is exchangeable in the arguments in ε, and likewise $F_{\eta|x}$ is exchangeable. Anonymous players are a special case of our analysis.

2.2 Identification of the Game

Our identification strategy for the two-player game relies on the following assumptions.

ASSUMPTION 1: Let $X \equiv \left(X', \widetilde{X}\right)$ with $\widetilde{X} \equiv \left(\tilde{X}_1, \tilde{X}_2\right)$, where each $\tilde{X}_i$ is a scalar. We assume $u_i(x) \equiv u_i(x') + \tilde{x}_i$ and $v_i(x) \equiv v_i(x')$ for $i = 1, 2$.

ASSUMPTION 2: The conditional distribution of $\widetilde{X}_i$ given the other covariates has support equal to $\mathbb{R}$ for $i = 1, 2$.

ASSUMPTION 3: (I) ε is independent of $\widetilde{X}$ (i.e., $F_{\varepsilon|x} = F_{\varepsilon|x'}$) and (II) $E(\varepsilon \mid x') = E(\varepsilon) = (0, 0)$.

ASSUMPTION 4: ε has an everywhere positive Lebesgue density on $\mathbb{R}$ conditional on x'.

ASSUMPTION 5: For each x', sign($v_1(x')$) = sign($v_2(x')$) and the econometrician knows the signs.

ASSUMPTION 6: For each x' with $v_1(x'), v_2(x') \leq 0$, (I) η is independent of $\widetilde{X}$ (i.e., $F_{\eta|x} = F_{\eta|x'}$), (II) $E(\eta \mid x') = E(\eta) = (0, 0)$, (III) for all η_i in its support I, $v_i(x') + \eta_i \leq 0$, and (IV) η and ε are independent. For each x' with $v_1(x'), v_2(x') \geq 0$, (V) $\eta \equiv (0, 0)$.

We discuss each assumption in turn. Assumption 1 requires exclusion restrictions at the level of each player. Specifically, it assumes $X \equiv \left(X', \widetilde{X}\right)$ includes a subvector $\widetilde{X}$ of player-specific factors that enter payoffs additively. These factors are the special regressors familiar from the literature on binary and multinomial choice models (Manksi 1988 and Lewbel 1998, 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. Consequently, interaction effects must be present if changes in the realization $\tilde{x}_i$ correspond to changes in the marginal probability of player j's action. In an entry game, $\tilde{x}_i$ might be the distance between chain i's headquarters and the geographic market in question.

Assumption 2 is a standard large support restriction on the special regressors. This restriction is necessary for identifying the tails of the distribution of $F_{\varepsilon|x'}$, as in the literature on binary and multinomial choice (Lewbel 2000). For example, if each consumer decides whether to purchase a bottle of soda for \$5, then the fraction of consumers with a willingness to pay of more than \$5 is identified by the fraction of consumers who purchase at that price. Varying

9

the price identifies the CDF of the consumers' willingness to pay. The failure of the special regressor to have large support naturally results in set identification of $F_{\varepsilon|x'}$ (and hence, under our approach, the other unknowns): we cannot learn this function evaluated at values in the tails of ε.

Assumption 3(i) is necessary to recover the distribution of unobservables ε from variation in $\widetilde{x}$. Assumption 3(ii) provides a location normalization. For any x', the mean of ε is otherwise not separately identified from $u_i(x')$, which subsumes the role of any intercept. This mean independence condition is economically weak. However, the mean of ε is naturally sensitive to the weight in the tails of ε. Consequently, under Assumption 3(ii) the rate of convergence of an assumed parametric portion of the model (say $u_i(x')$ is specified parametrically) may be slow and related to the thicknesses of the tails of ε and $\widetilde{X}$. This mean independence condition and its sensitivity to the support of $\tilde{X}$ (relative to the support of ε) has been examined in the case of semiparametric binary and multinomial choice models (e.g., Lewbel 1998, 2000; Khan and Tamer 2010; Dong and Lewbel 2012). Additional assumptions, such as tail symmetry in the distribution of ε, may remove some of these identification and estimation issues (Magnac and Maurin 2007). Other sets of identifying assumptions from the literature on binary and multinomial choice could be modified for our models.

Assumption 4 gives probability zero to tie events. In our definition of a pure strategy Nash equilibrium, it rules out with probability 1 that an agent is indifferent between actions 1 and 0.

Assumption 5 requires symmetry and knowledge of the signs of the interaction terms. It assumes the game is either a game of strategic complements or a game of strategic substitutes and the econometrician knows, for each $x' \in X'$, the type of the game she is studying. In many applications, such as entry games in industrial organization, economists are willing to assume that the signs of interaction effects are common and known, i.e. negative in entry games.

Assumption 6 is the analogous assumption to Assumption 3 for the unobservables in the interaction effects, η. Assumption 6 has separate implications for submodular games (such as entry games) and supermodular games (such as peer effect games). As we will explain, Assumption 6(v) implies that we only allow heterogeneous interaction effects in submodular games. For the restriction to submodular games to have content, Assumption 6(iii) imposes that each realization of η must be such that the game is still submodular.[6] We allow an otherwise unrestricted joint distribution for (η_1, η_2), but Assumption 6(iv) imposes that ε and η are independent.

Under the above set of assumptions, the conditional probability of an action profile a in the

---

[6]The allowed support for η thus depends on (v_1, v_2).

data is

$$\Pr\left(a \mid x\right) = \int 1\left[a \in \mathcal{D}\left(x, \varepsilon, \eta\right)\right] \Pr\left[a \text{ selected} \mid x, \varepsilon, \eta\right] dF_{\varepsilon|x'} dF_{\eta|x'}, \tag{2}$$

or the probability that a is an equilibrium and is the selected equilibrium for values of x, ε and η where multiple pure strategy equilibria exist. $\Pr\left(a \mid x\right)$ is a mixture over equilibria in the regions of non-uniqueness. The mixture distribution depends on all of x, ε and η.

Remark. Note that under Assumptions 5 and 6(iii), a pure strategy equilibrium to the game always exists. Indeed, Echenique (2004) shows that with probability 1 any two player, two action game either can be represented by a supermodular game or will have no equilibrium in pure strategies. The proof is simple: if the game is submodular we can reverse the strategy space of one of the players to convert the game to a supermodular one. The existence of equilibrium in pure strategies follows. If the interaction terms have different signs, then the game is isomorphic to matching pennies and has no equilibrium in pure strategies. Thus, two-player, two-action games with symmetric signs of the interaction terms always have an equilibrium in pure strategies. Moreover, in these cases some profiles of choices are unique pure strategy equilibria when they are pure strategy Nash equilibria. An equilibrium is said to be unique if $\mathcal{D}\left(x, \varepsilon, \eta\right) = \{a\}$ for all x, ε, η where $a \in \mathcal{D}\left(x, \varepsilon, \eta\right)$.

Following Tamer (2003) in a more parametric setup, we use the unique equilibria for identification. Our first result states the that the unknowns in the model are identified.

Theorem 1. *Under Assumptions 1–6,* $\left(\left(u_i, v_i\right)_{i \leq 2}, F_{\varepsilon|x'}, F_{\eta|x'}\right)$ *is identified.*

The proof relies on two unique equilibria. For submodular games (entry games), we use the probabilities of neither firm entering, $\Pr\left((0,0) \mid x\right)$, to identify u_1, u_2 and $F_{\varepsilon|x'}$ and then use the probabilities of both firms entering, $\Pr\left((1,1) \mid x\right)$, to identify v_1, v_2 and $F_{\eta|x'}$. Assumption 6(iv) imposes that ε and η are independent precisely so that $F_{\varepsilon|x'}$ and $F_{\eta|x'}$ are the structural objects of interest and so that they can be identified in separate steps. For supermodular games (peer effects), we use the probabilities of the unique equilibria, $(1,0)$ and $(0,1)$. In the first stage, we identify distributions of composite random variables and need to use the second stage to finally identify the primitives $\left(\left(u_i, v_i\right)_{i \leq 2}, F_{\varepsilon|x'}\right)$. This requirement to use both equilibria before any primitives are identified prevents us from including heterogeneous interaction effects, η.

Theorem 1 identifies the game under the key restrictions of symmetric signs of interaction effects and knowledge of them. 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 (without heterogeneous interaction effects). Their identification strategy and ours rely on the fact that, in both supermodular and submodular games,

two vectors of choices are unique equilibria when they are Nash equilibria in pure strategies. The advantage of our approach with respect to theirs is that ours does not rely on identification at infinity to recover the payoff terms. We discussed the simultaneous work of Kline (2012) and Dunker, Hoderlein and Kaido (2012) in the introduction.

3 Discrete Choice Over Bundles

3.1 Bundles Model

Consider an agent that faces two binary alternatives, $\{1, 2\}$. The only difference with standard models in discrete choice is that these alternatives are not mutually exclusive. The agent's choice set is therefore $\{(0,0), (1,0), (0,1), (1,1)\}$, where we use $a_i = 1$ to indicate that alternative i is selected and $a_i = 0$ to indicate that i is not selected. Let $a \equiv (a_1, a_2)$.

If the agent selects only alternative i then her payoff is $u_i(x) + \varepsilon_i$ for $i = 1, 2$. If she chooses both alternatives, her utility is

$$u_1(x) + \varepsilon_1 + u_2(x) + \varepsilon_2 + \Phi(x) + \eta.$$

That is, the value of the bundle is the sum of the utilities the agent gets from selecting each individual alternative plus an extra term $\Phi(x) + \eta$. We say that the alternatives are complements if $\Phi(x) + \eta \geq 0$; we say that they are substitutes if $\Phi(x) + \eta \leq 0$.

Again we denote by x the realization of the vector of observable state variables, $X \in \mathcal{X} \subseteq \mathbb{R}^k$, and let $\varepsilon \equiv (\varepsilon_1, \varepsilon_2) \in \mathbb{R}^2$ indicate a vector of random terms that are observed by the agent but not by the econometrician. The random vector ε is distributed according to the cumulative distribution function (CDF) $F_{\varepsilon|x}$; we allow correlation between the unobservables. In the discrete choice for bundles model, a scalar $\eta \in I \subset \mathbb{R}$ naturally captures heterogeneity in the interaction effect.

The agent chooses the set of goods a to maximize her utility

$$V(a, x, \varepsilon, \eta) \equiv \sum_{i=1,2} (u_i(x) + \varepsilon_i) \mathbf{1}(a_i = 1) + (\Phi(x) + \eta) \mathbf{1}(a_1 = 1, a_2 = 1).$$

The purpose of our analysis is to use data on chosen bundles and covariates, pairs (a, x), to recover the structure of the discrete choice for bundles model, $\left((u_i)_{i \leq 2}, \Phi, F_{\varepsilon|x}, F_{\eta|x}\right)$.

3.2 Identification of the Bundles Model

The model of discrete choice over bundles clearly differs from a binary game. Therefore, the restrictions we propose for identification have alternative interpretations in the two settings. However, from a technical standpoint many assumptions are identical. That is, our identification strategy for bundles relies on Assumptions 1–4 and 6 above with only two slight modifications.

ASSUMPTION 1*: Let $X \equiv \left(X', \widetilde{X}\right)$ with $\widetilde{X} \equiv \left(\widetilde{X}_1, \widetilde{X}_2\right)$. We assume $u_i(x) \equiv u_i(x') + \widetilde{x}_i$ for $i = 1, 2$ and $\Phi(x) \equiv \Phi(x')$.

ASSUMPTION 6*: For each x' for which $\Phi(x') \leq 0$, (I) η is independent of $\widetilde{X}$ (i.e., $F_{\eta|x} = F_{\eta|x'}$), (II) $\mathrm{E}(\eta \mid x') = \mathrm{E}(\eta) = 0$, (III) for all η in its support I, $\Phi(x') + \eta \leq 0$, and (IV) η and ε are independent. For each x' for which $\Phi(x') \geq 0$, (V) $\eta \equiv 0$.

For Assumption 1*, a common example of a special regressor is $\tilde{x}_i = -p_i$, or the price of good i. The price of a good is naturally excluded from the standalone payoff of the other good and the interaction value from consuming both of them. The intuition for identification is that the price of good 2 affects the purchase decision for good 1 only through the term $\Phi(x') + \eta$. We do not need regressors $\tilde{x}_i$ for each bundle, only each good.

Under Assumptions 1*, 2–4, and 6*, the conditional probability of observing the bundle a given the observables x is

$$\Pr(a \mid x) = \int 1 \left[a \in \arg\max_{\hat{a}} V(\hat{a}, x, \varepsilon, \eta)\right] dF_{\varepsilon|x'} dF_{\eta|x'}. \tag{3}$$

Compared to (2), there is no mixture over equilibria.

From an identification standpoint, there are two other key differences between the bundles model and the game. First, the common signs portion of Assumption 5 has no counterpart in the discrete choice model for bundles, as there is naturally only one agent's term $\Phi(x') + \eta$ capturing the interaction between the two goods. Second, the sign of $\Phi(x') + \eta$ can be recovered from available data on choices and covariates, under the portion of Assumption 6* that restricts the sign of $\Phi(x') + \eta$ to be the same for all η, for each x'.

Lemma 1. *Under Assumptions 1*, 2–4 and 6*, the sign of $\Phi(x') + \eta$ is identified for each x' and all η.*

The proof of Lemma 1 uses standard techniques from monotone comparative statics and relies on the concepts of supermodularity, increasing differences, and stochastic dominance. For

completeness, we describe Topkis' theorem and stochastic dominance in Appendix B. Using Lemma 1, Theorem 2 states identification of the model.

Theorem 2. *Under Assumptions 1*, 2–4 and 6*, $\left((u_i)_{i\leq 2}, \Phi, F_{\varepsilon|x'}, F_{\eta|x'}\right)$ is identified.*

For each x', identification proceeds by first learning the sign of $\Phi(x') + \eta$ using Lemma 1. The next steps of the proof are similar to those in the proof of Theorem 1. For the case of substitutes, we use the probabilities of no purchase, $\Pr((0,0) \mid x)$, to identify u_1, u_2 and $F_{\varepsilon|x'}$. We finally use the probabilities of buying the bundle, $\Pr((1,1) \mid x)$, to identify Φ and $F_{\eta|x'}$. Identification for the case of complementarities uses the probability of buying the first good and not the second one, $(1,0)$, followed by the probability of buying the second good but not the first, $(0,1)$. As with games, for complements we identify only a distribution of some composite primitives using the first bundle $(1,0)$, which explains why we cannot allow heterogeneous interaction effects η.

While we have mostly emphasized similarities in the identification arguments for games and models of discrete choice for bundles, there is another key difference in addition to Lemma 1. In a game, verifying that $(0,0)$ is a Nash equilibrium requires checking only unilateral deviations from $(0,0)$ to either $(1,0)$ or $(0,1)$. In the discrete choice model, the fact that an agent selects bundle $(0,0)$ is telling us she prefers this bundle over any other one, including $(1,1)$. Nevertheless, when goods are substitutes and $(0,0)$ is preferred over $(1,0)$ and $(0,1)$, then consuming none of the goods, $(0,0)$, is preferred to consuming both of them, $(1,1)$. The situation is similar for the case of complements. Therefore, the proof of Theorem 2 is nearly identical to the proof of Theorem 1.

Gentzkow (2007a, b) studies a discrete choice for bundles model for print and online newspapers where each u_i is specified parametrically and F_ε is bivariate normal. He does not include η.

4 Potential Games

A weakness of Theorem 1 that is not shared by Theorem 2 is that the result for games requires prior knowledge of the signs of the interaction effects, $v_1(x') + \eta_1$ and $v_2(x') + \eta_2$. This is because the identification argument behind Theorem 1 requires that certain equilibria, known to the econometrician, are unique whenever they occur. In this section, we present the argument that a stronger class of games, potential games, when combined with an equilibrium refinement, allows us to identify the signs of a common interaction effect.

14

In game theory, a potential function is a real-valued function defined on the space of pure strategy profiles of the players such that the change in any player's payoffs from a unilateral deviation is equal to the gain in the potential function. A game that admits such a function is called a potential game.[7] This concept was first used in economics as a way to prove the existence of Nash equilibria in pure strategies.[8] The reason is that the global maximum of the potential function corresponds to a pure strategy equilibrium of the related game. In a finite game, the potential function has a finite set of values; therefore it always has a global maximizer. It follows that the equilibrium set (in pure strategies) of any finite potential game is guaranteed to be non-empty.

Monderer and Shapley (1996) show that a potential function is uniquely defined up to an additive constant. As the set of maximizers does not depend on the specific potential function we use, the potential offers an equilibrium refinement. Important work has been done to address whether the selection rule based on potential maximizers 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 (1997b).

Remark. 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 (2001) 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."[9]

Lab experiments studying the so-called minimum effort game have showed that observed choice data are 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). These results are remarkable as this class of games often displays a very large number

---

[7]Monderer and Shapley (1996) also define ordinal and weighted potential games. We restrict attention to exact potential games.

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

[9]In a certain two-player, exact potential game, Blume (1993) argues that a log-linear strategy revision process selects potential maximizers. Morris and Ui (2005) discuss 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."

of equilibria in pure strategies.[10]

The equilibrium selection based on potential maximizers has received a lot of attention in both the theoretical and the experimental literatures on games. However, this selection rule has not yet been explored from an econometric perspective. We believe the exercise is interesting as the selection rule not only completes the model but also links observed choices to a single maximization problem identical to that for the discrete choice for bundles. Being identical in that respect, the identification result for the model of discrete choice for bundles can be directly extrapolated to the potential game.

We next specify the definition of a potential function for the two-player case and then provide a necessary and sufficient condition for $\Gamma\left(x, \varepsilon,\left(\eta_1, \eta_2\right)\right)$ to be a potential game.

Potential: A function $V: \{0,1\}^2 \times X \times \mathbb{R}^2 \times I \to \mathbb{R}$ for $I \subset \mathbb{R}^2$ is a **potential function** for $\Gamma\left(x, \varepsilon,\left(\eta_1, \eta_2\right)\right)$ if, for all $i \leq 2$, and all $a_{-i} \in \{0,1\}$,

$$V\left(a_i = 1, a_{-i}, x, \varepsilon,\left(\eta_1, \eta_2\right)\right) - V\left(a_i = 0, a_{-i}, x, \varepsilon,\left(\eta_1, \eta_2\right)\right) = U_{1,i}\left(a_{-i}, x, \varepsilon_i, \eta_i\right).$$

$\Gamma\left(x, \varepsilon,\left(\eta_1, \eta_2\right)\right)$ is called a **potential game** if it admits a potential function.

Monderer and Shapley (1996) argue that $\Gamma\left(x, \varepsilon,\left(\eta_1, \eta_2\right)\right)$ is a potential game if and only if the interaction effects on players' payoffs are symmetric. The next proposition reflects this idea.

Proposition 1. *$\Gamma\left(x, \varepsilon,\left(\eta_1, \eta_2\right)\right)$ is a potential game if and only if $v_1\left(x\right) + \eta_1 = v_2\left(x\right) + \eta_2 = \Phi\left(x\right) + \eta$, in which case we replace the argument $\left(\eta_1, \eta_2\right)$ in V with the scalar η and the potential function can be written as*

$$V\left(a, x, \varepsilon, \eta\right) = \sum_{i=1,2}\left(u_i\left(x\right) + \varepsilon_i\right) 1\left(a_i = 1\right) + \left(\Phi\left(x\right) + \eta\right) 1\left(a_1 = 1, a_2 = 1\right).$$

Borrowing notation from the model of discrete choice for bundles, we use η for a scalar interaction effect rather than the vector $\left(\eta_1, \eta_2\right)$. This proposition's proof follows directly from Ui (2000, Theorem 3) and is therefore omitted. Under Assumption 4, we can see that the maximizer of the potential function in Proposition 1 is unique with probability 1. By assuming that observed choices correspond to the potential maximizer, there is a unique equilibrium in pure strategies in the data generating process. When the equilibrium in the related game is

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

unique, it is always the potential maximizer. Under multiple equilibria, the potential function selects one of them depending on the sign of the interaction effect, as we explain next.

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

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

if not the maximizer is $(1,0)$. Thus, the equilibrium selection rule induced by the potential function predicts that the player choosing action 1 is the one with the highest standalone value.

Alternatively, when $\Phi(x) + \eta \geq 0$, as in a game of peer effects, and the game has multiple equilibria, then the equilibrium set is $\mathcal{D}(x, \varepsilon, \eta) = \{(0,0),(1,1)\}$. It can be easily checked that $(1,1)$ maximizes the potential if

$$(-u_1(x) - \varepsilon_1 - \Phi(x) - \eta)(-u_2(x) - \varepsilon_2 - \Phi(x) - \eta) > (u_1(x) + \varepsilon_1)(u_2(x) + \varepsilon_2);$$

if not the maximizer is $(0,0)$. 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).

Many empirical applications impose that a Pareto optimal equilibrium is 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 supermodular games. The criterion based on the potential maximizer makes predictions that are more in line with coordination failures (see, e.g., Cooper and John 1988). Which selection device is more appropriate depends on the specific application and more research needs to be done to address this fundamental question. Section 2 offers a natural environment for performing a nonparametric test of equilibrium selection. If the researcher is confident about the signs of the interaction terms, then the identification of the two player, two action game does not rely on an equilibrium selection rule. Being robust in that way, the econometrician can recover the structure of the game from observed choices and evaluate the criterion by which players select a particular equilibrium.

Remark. The equilibrium selection based on potential maximizers depends on both the observables x and the unobservables ε and η.

Proposition 1 implies that a binary game is a potential game whenever the interaction

effects are symmetric across players. In an entry game, the effect of player 1's entry on player 2's profit is the same as the effect of player 2's entry on player 1's profit. The standalone profits of both players can be flexibly asymmetric. We also allow the interaction effects to vary across games $\Gamma(x, \varepsilon, \eta)$, such as across markets in entry or across groups of friends in peer effects. We note that symmetry of interaction effects is assumed in much of the empirical literature estimating binary games. Symmetry of interaction effects is most attractive whenever players are anonymous, meaning in a potential game that $u_1 = u_2$ and $F_{\varepsilon|x}$ is exchangeable in the arguments in ε.

The following two restrictions allow us to use the previous identification strategy for bundles for potential games.

ASSUMPTION 7: $v_1(x') = v_2(x') = \Phi(x')$ for all $x' \in X'$ and $\eta_1 = \eta_2 = \eta$ for all realizations (η_1, η_2).

ASSUMPTION 8: Players select actions according to the potential maximizer.

Under Assumptions 7 and 8, identification in a potential game is formally equivalent to identification in the model where a consumer selects a bundle of choices. Note that this is not equivalent to stating that the equilibrium that maximizes the sum of player's utilities is always picked, as we argued previously.

Corollary 1. *Under Assumptions 1*, 2–4, 6*, 7 and 8 , $\left((u_i)_{i \leq 2}, \Phi, F_{\varepsilon|x'}, F_{\eta|x'}\right)$ is identified.*

Assumptions 7 and 8 allow us to use choice data to recover the sign of the interaction term, as in Lemma 1. The researcher does not need to know the sign ahead of time. A further advantage of any known equilibrium selection rule is that the model is coherent and maximum likelihood estimation can be used, for all equilibria. Also, any equilibrium selection rule can aid in counterfactuals.

5 Three or More Goods or Players

We extend our identification results to three or more goods and three or more players.

5.1 Discrete Choice Model for Bundles

5.1.1 Bundles Model with n Goods

Consider an agent that faces $\mathcal{N} = \{1, 2, ..., n\}$ binary alternatives that are not mutually exclusive. Thus, her choice set is $\{0, 1\}^n$, where as before we use 1 to indicate that the corresponding

alternative is selected.

If the agent selects only alternative i then her payoff is $u_i(x) + \varepsilon_i$ for $i = 1, 2, ..., n$. We write $\mathcal{S}(a) \subseteq \mathcal{N}$ for the set of alternatives that are selected in a, $\mathcal{S}(a) \equiv \{i \in \mathcal{N} : a_i = 1\}$. If the agent selects bundle $\mathcal{S}(a)$, her utility is

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

That is, the value of the bundle is the sum of the utilities of each individual alternative plus an extra term $\Phi(\mathcal{S}(a), x)$. We normalize $\Phi(\mathcal{S}(a), x)$ so that it is 0 when $|S(a)| < 2$. Importantly, we allow the interaction term to vary with the bundle of goods selected. This notation nests an example where joint consumption of goods 1 and 2 raises utility compared to the sum of standalone utilities and joint consumption of goods 1 and 3 lowers utility relative to the sum of standalone utilities.

Here again we denote by x the realization of the vector of observable state variables, $X \in \mathcal{X} \subseteq \mathbb{R}^k$. We also let $\varepsilon \equiv (\varepsilon_i)_{i \leq n} \in \mathbb{R}^n$ indicate a vector of random terms that are observed by the agent but not by the econometrician. The random vector ε is distributed according to the cumulative distribution function (CDF) $F_{\varepsilon|x}$; we allow an unrestricted joint distribution of the unobservables. We do not include heterogeneous interaction effects η, as we discuss below.

The agent chooses a vector of choices a to maximize $V(a, x, \varepsilon)$. The purpose of our analysis is to recover the structure of the discrete choice for bundles model, $\left((u_i)_{i \leq n}, \Phi, F_{\varepsilon|x}\right)$, from available data on choices and covariates (a, x).

5.1.2   Identification of the Model

The identification strategy relies on the following assumptions.

ASSUMPTION 1': Let $X \equiv \left(X', \widetilde{X}\right)$ with $\widetilde{X} \equiv \left(\widetilde{X}_i\right)_{i \leq n}$. We assume $u_i(x) \equiv u_i(x') + \widetilde{x}_i$ for all $i \in \mathcal{N}$ and $\Phi(., x) \equiv \Phi(., x')$.

ASSUMPTION 2': The conditional distribution of $\widetilde{X}_i$ given the other covariates has support on all $\mathbb{R}$ for all $i \in \mathcal{N}$.

ASSUMPTION 3': (I) ε is independent of $\widetilde{X}$ so that we write $F_{\varepsilon|x'}$ and (II) $\mathrm{E}(\varepsilon \mid x') = \mathrm{E}(\varepsilon) = (0, 0, ..., 0)$.

ASSUMPTION 4': ε has an everywhere positive Lebesgue density on $\mathbb{R}$ conditional on x'.

ASSUMPTION 5': For each x', there exists a known vector $\widehat{a} \in \{0, 1\}^n$ with the following property. For all $\widetilde{x}$, where $x = (x', \widetilde{x})$, and for all 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 $i \in \mathcal{N}$, where $a_i = 1 - \widehat{a}_i$.

Compared to the two-alternative case, the main extra restriction is Assumption 5'. It requires the existence of a vector of choices (known by the econometrician) such that if the vector is a local maximum, then the vector is also a global maximum. Assumption 5' can be thought of as a notion of local concavity for discrete domains. We use this assumption to trace out $F_{\varepsilon|x'}$ using variation in the special regressors. We state the identification result first and then provide three sufficient conditions for Assumption 5' to hold.

Bundle probabilities are of the form (3). Gentzkow (2007b, Proposition 2) explores comparative statics of choice probabilities in the special regressors $\tilde{x}_i$; they depend on the interaction effects in a somewhat complex manner.

The identification result is as follows.

Theorem 3. *Under Assumptions 1'–5', $\left((u_i)_{i\leq n}, \Phi, F_{\varepsilon|x'}\right)$ is identified.*

The proof works by first using Assumption 5' to trace $F_{\varepsilon|x'}$ using variation in the special regressors. We then use the known $F_{\varepsilon|x'}$ to show that different values for $\left((u_i)_{i\leq n}, \Phi\right)$ lead to different bundle probabilities. The last step of the proof involves a subtle detail not present in proofs of identification for multinomial choice models without bundles: sometimes one inequality will be implied by other inequalities. This was illustrated earlier for the two good case: in the case of substitutes, if it does not maximize utility to buy either good alone, then the bundle of both goods will also not maximize utility. In a discrete choice model without bundles, it is always the case that increasing the special regressor $\tilde{x}_i$ for good i will strictly increase the probability of purchasing good i. However, in the bundles case, it can be that changing regressor values locally does not affect choice probabilities, if those regressors enter only inequalities that are implied by other inequalities. This subtle detail makes the proof differ more from the literature on discrete choice without bundles than may otherwise be apparent.

The next three conditions are sufficient for Assumption 5' to hold.

Two Goods

Assumption 5' always holds if there are two goods, so that the assumption in this case is not stronger than the previous analysis. If Assumptions 1'–4' are satisfied and $\mathcal{N} = \{1, 2\}$, then by Lemma 1 the sign of $\Phi(x')$ is identified for all x'. It can be easily shown that Assumption 5' holds with $\hat{a} = (1, 0)$ or $\hat{a} = (0, 1)$ when $\Phi(x') \geq 0$ and $\hat{a} = (0, 0)$ or $\hat{a} = (1, 1)$ when $\Phi(x') \leq 0$.

Negative Interaction Effects

The second sufficient condition we propose relies on substitutabilities among alternatives. 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 2. *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 $a'_{-i} > a_{-i}$ (in the coordinatewise order) we have*

$$V(a_i = 1, a_{-i}, x, \varepsilon) - V(a_i = 0, a_{-i}, x, \varepsilon) \leq (<) \, 0 \implies V\left(a_i = 1, a'_{-i}, x, \varepsilon\right) - V\left(a_i = 0, a'_{-i}, x, \varepsilon\right) \leq (<) \, 0.$$

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

In the lemma's statement, $a'_{-i} > a_{-i}$ means that the consumer adds one or more goods other than i to her bundle. In this case, identification of $F_{\varepsilon|x'}$ uses data on the fraction of consumers who purchase none of the goods, i.e., $\Pr((0, 0, ..., 0) \mid x)$. The single-crossing condition in Lemma 2 is satisfied if, for instance, the alternatives are substitutes in some sense (we do not precisely define substitutes for this model).

Global Concavity for Discrete Domains

Assumption 5' is a local notion of concavity for discrete domains. The last sufficient condition for the assumption is a more global analog of concavity.

Lemma 3. *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 x'.

Global 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 can take values on the entire real line, then the interaction effects must be identically zero for concavity to hold globally. Thus, this requires the large support portions of Assumptions 2' and 4' to be simultaneously relaxed. For this reason, we cannot recommend basing identification on global concavity explicitly. However, Assumption 5' itself is a weaker (local) version of discrete concavity that is compatible with our other restrictions.

Remark. We do not include heterogeneous interaction terms η in our model. This is because the number of such interaction terms grows quickly with the number of goods. Indeed, there is

one interaction term for each bundle with more than two elements. Our strategy of identifying one distribution for each conditional probability $\Pr(a \mid x)$ will not allow the identification of the joint distribution of heterogeneous interaction terms.

5.2 n-Player Game

5.2.1 The Game

Consider an extension of the game to more than two players. The set of players is $\mathcal{N} = \{1, ..., n\}$. Each player $i \in \mathcal{N}$ chooses an action $a_i \in \{0, 1\}$. We denote by $X \in \mathcal{X} \subseteq \mathbb{R}^k$ a vector of observable state variables and by $\varepsilon \equiv (\varepsilon_i)_{i \leq n} \in \mathbb{R}^n$ a vector of random terms that are observed by the players but not by the econometrician. The random vector ε is distributed according to $F_{\varepsilon|x}$. Here again, 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

$$U_{1,i}(a_{-i}, x, \varepsilon_i) \equiv u_i(x) + v_i(a_{-i}, x) + \varepsilon_i, \tag{4}$$

while the return from action 0, $U_{0,i}(a_{-i}, x, \varepsilon_i)$, is normalized to 0. In addition, we normalize $v_i(a_{-i}, x)$ to 0 when the actions of all players but i are 0. We denote this game by $\Gamma(x, \varepsilon)$. The definition of a pure strategy Nash equilibrium $a^* \equiv (a_i^*)_{i \leq n}$ naturally extends from the two-player case. We write $\mathcal{D}(x, \varepsilon)$ for the equilibrium set (in pure strategies) of $\Gamma(x, \varepsilon)$. The same conditions that 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, $\left((u_i, v_i)_{i \leq n}, F_{\varepsilon|x'} \right)$, from available data on choices and covariates, (a, x), for each play of the game.

5.2.2 Previous Identification Strategy with No Selection Rule

We first explore our identification strategy without imposing an equilibrium selection rule, which for the two player case was done in Section 2. Recall that the identification strategy for two players used two unique equilibria. For a submodular (example, entry) game where $\eta \equiv (0, 0)$, we used the probabilities $\Pr(0, 0 \mid x)$ to identify $F_{\varepsilon|x'}$ and each $u_i(x')$. We then used the probabilities $\Pr(1, 1 \mid x)$ to identify the interaction effects $v_i(1, x')$.

In a submodular game of three players, the equilibria $(0, 0, 0)$ and $(1, 1, 1)$ will also be unique. However, the number of interaction effects is now quite large. For each of the three players, there are three interaction effect functions. For example, for player 1 the interaction effects are $v_1((1, 0), x')$, $v_1((0, 1), x')$, and $v_1((1, 1), x')$, yielding $3 \cdot 3 = 9$ interaction effects to identify

for each x'. While the probabilities $\Pr\left(0,0,0 \mid x\right)$ are enough to identify $F_{\varepsilon|x'}$ and each $u_i\left(x'\right)$, the probabilities $\Pr\left(1,1,1 \mid x\right)$ and the proof strategy used in Section 2 will not identify nine interaction effects.

The number of interaction effects increases quickly with the number of players. For sub-modular games of four players, there will be $6 \cdot 4 = 24$ interaction effects to identify from the probabilities $\Pr\left(1,1,1,1 \mid x\right)$. For these combinatorial reasons, the strategy of identifying all the model parameters by relying on the unique equilibria $(0,\ldots,0)$ and $(1,\ldots,1)$ for submodular games likely applies only to two-player games. In addition, supermodular games with three or more players may not even have unique equilibria.

5.2.3   Representation of the Game via Potential Functions

Because the identification strategy relying on unique equilibria does not appear to extend to the case of three or more players, we turn our attention to potential games.

We first specify the definition of a potential function for the n-player case and then provide a necessary and sufficient condition for $\Gamma\left(x, \varepsilon\right)$ to be a potential game.

Potential: A function $V : \{0,1\}^n \times X \times \mathbb{R}^n \to \mathbb{R}$ is a **potential function** for $\Gamma\left(x, \varepsilon\right)$ if, for all $i \leq n$, for all $a_{-i} \in \{0,1\}^{n-1}$,

$$V\left(a_i = 1, a_{-i}, x, \varepsilon\right) - V\left(a_i = 0, a_{-i}, x, \varepsilon\right) = U_{1,i}\left(a_{-i}, x, \varepsilon_i\right).$$

$\Gamma\left(x, \varepsilon\right)$ is a **potential game** if it admits a potential function.

We write

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

Each player's payoff from action 0 is normalized to 0; the notation $S \in \mathcal{S}(a, i)$ returns the empty set when $a_i = 0$. The next proposition derives from Ui (2000, Theorem 3).

Proposition 2. Γ *is a potential game if and only if there exists a function*

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

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

$$U_{1,i}\left(a, x, \varepsilon_i\right) = u_i\left(x\right) + \varepsilon_i + \sum\nolimits_{S \in \mathcal{S}(a,i)} \tilde{\Phi}\left(S, x\right).$$

23

A potential function is given by

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

An n-player game admits a potential representation if the interaction terms are group-wise symmetric. That is, if players i and j are in $S(a)$, then the corresponding interaction effects for players i and j are the same. The restriction does not imply that the effect of i's entry, say, on player j is the same as i's effect on player k. Rather, i's effect on j must be the same as j's effect on i.

The purpose of our analysis is to combine data with initial assumptions to learn about the structure of the game $\left((u_i)_{i \leq n}, \tilde{\Phi}, F_{\varepsilon | x'} \right)$.

5.2.4 Identification for Potential Games

For identification, we require that $\Gamma(x, \varepsilon)$ admits a potential function for all x and ε and that the observed action profiles correspond to potential maximizers. In addition to Assumptions 1'–4' above, the identification strategy relies on the following assumptions.

ASSUMPTION 6': The game admits a potential representation for all $x \in X$.
ASSUMPTION 7': Players select actions according to the potential maximizer.

Corollary 2. *Under Assumptions 1'–7',* $\left((u_i)_{i \leq n}, \tilde{\Phi}, F_{\varepsilon | x'} \right)$ *is identified.*

The corollary holds by appeal to Theorem 3 for the model of discrete choice for bundles. To formalize the idea, let us define $\Phi(\mathcal{S}(a), x) \equiv \sum_{S \in \mathcal{S}(a)} \tilde{\Phi}(S, x)$. By Theorem 3, $\left((u_i)_{i \leq n}, \Phi, F_{\varepsilon | x'} \right)$ is identified and it is readily verified that we can recover $\tilde{\Phi}$ from Φ. The three sufficient conditions provided above for the notion of local concavity in Assumption 5' have counterparts for the analysis of the game. That is, Assumption 5' holds if one or more of the following conditions hold: there are only two players; the game is of strategic substitutes – this restriction is often assumed in the entry games estimated in industrial organization; and the potential function is discrete globally concave.

It is likely our results will lead to nonparametric identification 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 a potential game, and therefore our previous results can be applied to these types of interactions.

6 Conclusion

We explore identification in discrete choice models with bundles and binary choice games of complete information. We argue that identification is quite similar in binary games (without an equilibrium selection rule) and discrete choice models for bundles. Further, there is an exact equivalence in potential games when the potential maximizer is always played.

We show how our models are identified. We recover from data the subutility function of each player or each good, the interaction effects among each set of players or each bundle of goods, and the joint distribution of potentially correlated, player- or good-specific unobservables. For some models, we recover the joint distribution of heterogeneous interaction effects. Identifying the distributions of the unobservables is necessary for computing marginal effects and conducting counterfactuals.

A Proofs

A.1 Proof of Theorem 1

By Assumptions 5 and 6, for each x', the game is either supermodular or submodular.

Submodular (entry game)

We first consider the case where $\Gamma(x, \varepsilon, \eta)$ is a submodular game, i.e., $v_i(x') + \eta_i \leq 0$ for $i = 1, 2$ and all η_i (Assumption 6(iii)). Under Assumption 4, when $(0, 0)$ is an equilibrium, it is a unique equilibrium (e.g., Tamer 2003). Under Assumptions 1 and 4 and using (2), the probability of both agents taking action 0 is

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

Define the random vector

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

By Assumption 3(i), ε is independent of $\widetilde{X}$. Therefore, for an arbitrary point of evaluation $\alpha^\star = (\alpha_1^\star, \alpha_2^\star)$ of the CDF $F_{\alpha|x'}$,

$$F_{\alpha|x'}(\alpha^\star) = \Pr(\alpha_1 \leq \alpha_1^\star, \alpha_2 \leq \alpha_2^\star \mid x') = \Pr(\alpha_1 \leq -\widetilde{x}_1, \alpha_2 \leq -\widetilde{x}_2 \mid x) = \Pr((0,0) \mid x),$$

for special regressor choices $\tilde{x}_1 = -\alpha_1^{\star}$ and $\tilde{x}_2 = -\alpha_2^{\star}$. Therefore, the variation in $\tilde{x}$ from Assumption 2 and $\Pr((0,0) \mid x)$ identifies the CDF of α, for each x'. Applying the same logic and Assumptions 3(i) and 6(i), we can use variation in $\tilde{x}$ and the probability $\Pr((1,1) \mid x)$ of the other unique equilibrium, where both agents take action 1, to recover $F_{\beta \mid x'}$, where the random vector β is

$$\beta = (\beta_1, \beta_2) \equiv (-\varepsilon_1 - u_1(x') - v_1(x') - \eta_1, -\varepsilon_2 - u_2(x') - v_2(x') - \eta_2).$$

By Assumptions 3(ii) and 6(ii), ε and η have zero means conditional on X'. Therefore,

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

One can see that $(u_i, v_i)_{i \leq 2}$ are identified for this x'. For example, $u_1(x')$ is identified as $E[\alpha_1 \mid x']$.

Once $(u_i, v_i)_{i \leq 2}$ are identified, it is a simple, known shift of the location of a random vector to go from the identified distribution of α to the distribution of $(\varepsilon_1, \varepsilon_2)$. Therefore, we identify $F_{\varepsilon \mid x'}$, which is a primitive of interest by Assumption 6(iv). Likewise, we can move from the distribution of β to the distribution of $(-\varepsilon_1 - \eta_1, -\varepsilon_2 - \eta_2)$ and, by a known multiplicative change of variable, the distribution of $(\varepsilon_1 + \eta_1, \varepsilon_2 + \eta_2)$. The distribution of $F_{\eta \mid x'}$ is identified from the distributions $F_{\varepsilon \mid x'}$ and $F_{\beta \mid x'}$ because the joint characteristic function of the sum of two independent (by Assumption 6(iv)) random vectors, here ε and η, is equal to the products of the joint characteristic functions of η and ε. Therefore, the characteristic function of η is the ratio of the characteristic functions of β and ε. One can find the distribution function of η once its characteristic function is identified, because there is a bijection between the spaces of distribution and characteristic functions.

Supermodular (peer effects game)

We next consider the case where $\Gamma(x, \varepsilon, \eta)$ is a supermodular game, i.e., $v_i(x') \geq 0$ for $i = 1, 2$. Recall that $\eta \equiv (0,0)$ by Assumption 6(v). Under Assumption 4, when $(1,0)$ is an equilibrium, it is a unique equilibrium (e.g., Tamer 2003). Under Assumptions 1 and 4 and using (2), the probability of the first agent taking action 1 and the second agent taking action 0 is

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

Define the random vector

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

By Assumption 3(i), ε is independent of $\widetilde{X}$. Therefore, for an arbitrary point of evaluation $\alpha^\star = (\alpha_1^\star, \alpha_2^\star)$ of the CDF $F_{\alpha|x'}$,

$$F_{\alpha|x'}(\alpha^\star) = \Pr(\alpha_1 \leq \alpha_1^\star, \alpha_2 \leq \alpha_2^\star \mid x') = \Pr(\alpha_1 \leq \widetilde{x}_1, \alpha_2 \leq -\widetilde{x}_2 \mid x) = \Pr((1,0) \mid x)$$

for choices $\tilde{x}_1 = \alpha_1^\star$ and $\tilde{x}_2 = -\alpha_2^\star$. Therefore, the variation in $\widetilde{x}$ from Assumption 2 and $\Pr((1,0) \mid x)$ identifies the CDF of α, for each x'. Applying the same logic, we can use variation in $\widetilde{x}$ and the probability $\Pr((0,1) \mid x)$ of the other unique equilibrium to recover $F_{\beta|x'}$, where the random vector β is

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

By Assumption 3(ii), ε has zero mean conditional on X'. Therefore,

$$E[(\alpha_1, \alpha_2) \mid x'] = (-u_1(x'), u_2(x') + v_2(x'))$$
$$E[(\beta_1, \beta_2) \mid x'] = (u_1(x') + v_1(x'), -u_2(x'))$$
$$E[(\alpha_1, \alpha_2) \mid x'] + E[(\beta_1, \beta_2) \mid x'] = (v_1(x'), v_2(x')).$$

One can see that $(u_i, v_i)_{i \leq 2}$ are identified for this x'. For example, $u_1(x')$ is identified as $-E[\alpha_1 \mid x']$.

Once $(u_i, v_i)_{i \leq 2}$ are identified, it is a known shift of the location of a random vector to go from the identified distribution of α to the distribution of $(-\varepsilon_1, \varepsilon_2)$. Then, a known multiplicative change of variables is used to learn the distribution of $(\varepsilon_1, \varepsilon_2)$ from the distribution of $(-\varepsilon_1, \varepsilon_2)$. Therefore, we identify $F_{\varepsilon|x'}$.

If we had included heterogeneous interaction effects η in the model, we could not have learned either $F_{\varepsilon|x'}$ or $F_{\eta|x'}$ from the distribution of $F_{\alpha|x'}$ alone.

A.2 Proof of Lemma 1

The proof of Lemma 1 involves three steps. The first two steps show that the sign of $\Phi(x') + \eta$ has different observable implications for the cases where the two goods are complements and substitutes. The last step exploits the observable implications. Appendix B reviews some of

the tools used in this proof.

Step 1 If $\Phi(x') \geq 0$, then $V(a, x, \varepsilon, \eta)$ is supermodular in (a_1, a_2). In addition, $V(a, x, \varepsilon, \eta)$ has increasing differences in both $(a_1, \widetilde{x}_1)$ and $(a_2, \widetilde{x}_1)$ (here the cross effect between a_2 and $\widetilde{x}_1$ is 0). By Assumption 4, the maximizer is unique with probability 1. Define the optimal vector of decisions as

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

Then, by Topkis's theorem, $a^*\left(x', \widetilde{x}_1, \widetilde{x}_2, \varepsilon, \eta\right)$ increases (in the coordinatewise order) in $\widetilde{x}_1$ with probability 1. By Assumption 3(i), the unobservables ε are independent of $\widetilde{x}_1$. Therefore, for all $\widetilde{y}_1 > \widetilde{x}_1$ and all upper sets U in $\{0, 1\}^2$, we have

$$\Pr\left(a^*\left(x', \widetilde{y}_1, \widetilde{x}_2, \varepsilon, \eta\right) \in U \mid x', \widetilde{y}_1, \widetilde{x}_2\right) \geq \Pr\left(a^*\left(x', \widetilde{x}_1, \widetilde{x}_2, \varepsilon, \eta\right) \in U \mid x', \widetilde{x}_1, \widetilde{x}_2\right).$$

That is, $\Pr\left(a^*\left(x', \widetilde{x}_1, \widetilde{x}_2, \varepsilon, \eta\right) \mid x', \widetilde{x}_1, \widetilde{x}_2\right)$ increases with respect to first order stochastic dominance in $\widetilde{x}_1$. Because stochastic dominance is preserved under marginalization, then $\Pr\left(a_2 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2\right)$ increases in $\widetilde{x}_1$ (see, e.g., Müller and Stoyan 2002, Theorem 3.3.10, p. 94). Similarly, we can show that $\Pr\left(a_1 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2\right)$ increases in $\widetilde{x}_2$.

Step 2 If $\Phi(x') + \eta \leq 0$ then $V(a, x, \varepsilon, \eta)$ is supermodular in $(a_1, -a_2)$. In addition, $V(a, x, \varepsilon, \eta)$ has increasing differences in both $(a_1, \widetilde{x}_1)$ and $(-a_2, \widetilde{x}_1)$ (here the cross effect between $-a_2$ and $\widetilde{x}_1$ is 0). By Assumption 4, the maximizer is unique with probability 1. Then, by Topkis's theorem, $\left(a_1^*\left(x', \widetilde{x}, \varepsilon, \eta\right), -a_2^*\left(x', \widetilde{x}, \varepsilon, \eta\right)\right)$ increases in $\widetilde{x}_1$ with probability 1. By Assumptions 3(i) and 6(i), the unobservables are independent of $\widetilde{x}_1$. Therefore, $\Pr\left(a_1, -a_2 \mid x', \widetilde{x}_1, \widetilde{x}_2\right)$ increases with respect to first order stochastic dominance in $\widetilde{x}_1$. Because stochastic dominance is preserved under marginalization, $\Pr\left(-a_2 = 0 \mid x', \widetilde{x}_1, \widetilde{x}_2\right)$ increases in $\widetilde{x}_1$ and thus $\Pr\left(a_2 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2\right)$ decreases in $\widetilde{x}_1$. Similarly, we can show that $\Pr\left(a_1 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2\right)$ decreases in $\widetilde{x}_2$.

Step 3 By Assumption 4, $\Pr\left(a_2 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2\right)$ and $\Pr\left(a_1 = 1 \mid x', \widetilde{x}_1, \widetilde{x}_2\right)$ are not constant as functions of $\widetilde{x}_1$ and $\widetilde{x}_2$ respectively if $\Phi(x') \neq 0$. From Steps 1 and 2, the sign of $\Phi(x')$ is identified from available data for each $x' \in X'$.

A.3 Proof of Theorem 2

As this proof is very similar to the proof of Theorem 1, we omit some of the more detailed arguments. The agent will maximize $V(a, x, \varepsilon, \eta)$. By Lemma 1, we can use data to learn the

sign of $\Phi\left(x'\right)+\eta$ for this x'.

Goods are Substitutes

Under Assumptions 1* and 4, the probability of buying neither good is

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

Because $\Phi\left(x'\right)+\eta<0$ by Assumption 6*(iii), the third inequality above is implied by the first two. Thus,

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

We omit the rest of the proof because it is similar to the proof of Theorem 1 for submodular games.

Goods are Complements

Consider the case where $\Phi\left(x'\right)\geq0$ and $\eta\equiv0$, by Assumption 6*(v). Under Assumption 4, when $\left(1,0\right)$ is selected, it is the only bundle that maximizes utility with positive probability. Thus, under Assumptions 1* and 4,

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

Because $\Phi\left(x'\right)\geq0$, the third inequality above is implied by the first two. Thus,

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

We omit the rest of the proof because it is similar to the proof of Theorem 1 for supermodular games.

A.4   Proof of Proposition 1

The proof of this result follows directly from Ui (2000, Theorem 3).   □

A.5   Proof of Corollary 1

The proof for this result follows directly from Assumptions 7 and 8, Proposition 1 and Theorem 2. The potential game is formally equivalent to a model of discrete choice for bundles.

A.6   Proof of Theorem 3

The proof of this result is divided into two steps. We first show that if Assumptions 1'–5' are satisfied, then $F_{\varepsilon|x'}$ is identified. We then show that if Assumptions 1'–4' are satisfied and $F_{\varepsilon|x'}$ is identified, then the stand-alone utility functions and interaction terms $\left((u_i)_{i \leq n}, \Phi\right)$ are also identified.

Step 1 (Identification of $F_{\varepsilon|x'}$) We start by showing that if Assumptions 1'–5' are satisfied, then $F_{\varepsilon|x'}$ is identified. By Assumption 5', for each x', there exists a known vector $\widehat{a} \in \{0, 1\}^n$ such that, 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 $i \in \mathcal{N}$,

$$(1(\widehat{a}_i = 1) - 1(\widehat{a}_i = 0))\varepsilon_i \geq \tag{5}$$
$$\Phi(\mathcal{S}(a'), x') - \Phi(\mathcal{S}(\widehat{a}), x') - (1(\widehat{a}_i = 1) - 1(\widehat{a}_i = 0))(u_i(x') + \widetilde{x}_i),$$

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

$$\Pr(\widehat{a} \mid x', \widetilde{x}) = \Pr((5) \text{ holds for all } i \in \mathcal{N} \mid x', \widetilde{x}).$$

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

$$\mu_i = (1(\widehat{a}_i = 1) - 1(\widehat{a}_i = 0))\varepsilon_i$$
$$- (\Phi(\mathcal{S}(a'), x') - \Phi(\mathcal{S}(\widehat{a}), x')) + (1(\widehat{a}_i = 1) - 1(\widehat{a}_i = 0))(u_i(x')).$$

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

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

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

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

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

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

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 $(1\,(\widehat{a}_i = 1) - 1\,(\widehat{a}_i = 0))$, the distribution of ε conditional on x' is obtained from the distribution of $\mu - E\,(\mu \mid x')$ conditional on x'.

Step 2 (Identification of $\left((u_i)_{i \leq n}, \Phi\right)$) This argument conditions on x'. Under Assumption 1',

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

To facilitate the exposition, let

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

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 $a^{2^n} = (1, 1, ..., 1)$. In addition, we define

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

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

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

respectively. Thus, for each a',

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

captures the probability of observing the choice vector a' conditional on $X = x$, or in simpler notation, $\Pr\,(a' \mid x)$.[11] The researcher can identify $\Pr\,(a' \mid x)$ directly from the data.

Let $\widetilde{W} \neq W$, meaning one or more of $(u_i\,(x'))_{i \in \mathcal{N}}$ and $\Phi\,(x')$ differ across $\tilde{W}$ and W for x'.

---

[11]More formally, let $F_{\triangle \varepsilon(a')|x'}$ be the distribution of $\triangle \varepsilon\,(a')$. Then,

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

We next show that $P\left(\triangle W\left(a',x\right)\mid x;F_{\varepsilon|x'}\right) > P\left(\triangle\widetilde{W}\left(a',x\right)\mid x;F_{\varepsilon|x'}\right)$, so that W is identified and we can then recover $\left((u_i)_{i\leq n},\Phi\right)$. Let

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

Note that by the formulas for $\left(W\left(a,x\right)-\widetilde{W}\left(a,x\right)\right)$, $C\left(x\right)$ is a constant set, as only $\tilde{x}$ varies. Also suppose $\max_a\left(W\left(a,x\right)-\widetilde{W}\left(a,x\right)\right) > 0$; the other case follows by a similar argument. Define $D\left(x\right)\equiv\{a\notin C\left(x\right)\mid a\in\{0,1\}^n\}$. We know $C\left(x\right)\neq\varnothing$. The fact that $D\left(x\right)\neq\varnothing$ follows as

$$W\left(a=(0,0,...,0),x\right)=\widetilde{W}\left(a=(0,0,...,0),x\right)=0 \tag{6}$$

and we supposed $\max_a\left(W\left(a,x\right)-\widetilde{W}\left(a,x\right)\right) > 0$. Fix some $a'\in C\left(x\right)$. We know that $W\left(a',x\right)-\widetilde{W}\left(a',x\right) = W\left(a,x\right)-\widetilde{W}\left(a,x\right)$ for all $a\in C\left(x\right)$, and $W\left(a',x\right)-\widetilde{W}\left(a',x\right) > W\left(a,x\right)-\widetilde{W}\left(a,x\right)$ for all $a\in D\left(x\right)$. Rearranging terms,

$$W\left(a',x\right)-W\left(a,x\right)=\widetilde{W}\left(a',x\right)-\widetilde{W}\left(a,x\right)\ \text{ for all }a\in C\left(x\right),\ \text{ and}$$
$$W\left(a',x\right)-W\left(a,x\right)>\widetilde{W}\left(a',x\right)-\widetilde{W}\left(a,x\right)\ \text{ for all }a\in D\left(x\right).$$

Then, by Assumption 4', the argument in the following paragraphs ensures that we can find $\tilde{x}$ such that, for $x=(\tilde{x},x')$,

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

i.e., $P\left(\triangle W\left(a',x\right)\mid x;F_{\varepsilon|x'}\right) > P\left(\triangle\widetilde{W}\left(a',x\right)\mid x;F_{\varepsilon|x'}\right)$. Therefore, W is non-constructively identified at x', and hence $\left((u_i\left(x'\right))_{i\leq n},\Phi\left(x'\right)\right)$ is identified as well.

As mentioned previously, we need to find an appropriate value for $\tilde{x}$. This choice of $\tilde{x}$ involves an additional detail that we address next. The inequalities that allow us to show that $P\left(\triangle W\left(a',x\right)\mid x;F_{\varepsilon|x'}\right) > P\left(\triangle\widetilde{W}\left(a',x\right)\mid x;F_{\varepsilon|x'}\right)$ are the ones that involve $a\in D\left(x\right)$. Some of the inequalities involving $a\in D\left(x\right)$ may be implied by other inequalities for some choices of $\tilde{x}$. To see this, suppose there are two, substitute goods and let $a'=(0,0)$, $C\left(x\right)=\{(0,0),(1,0),(0,1)\}$ and $D\left(x\right)=\{(1,1)\}$. Under substitutes, $W\left((0,0),x\right)\geq W\left((1,0),x\right)$ and $W\left((0,0),x\right)\geq W\left((0,1),x\right)$ together imply $W\left((0,0),x\right) > W\left((1,1),x\right)$. In this example, the fact that two inequalities imply a third means that marginal changes in the Φ term will not

affect the probability of the outcome $(0,0)$. Therefore, $a' = (0,0)$ does not allow us to effectively distinguish W from $\widetilde{W}$ as we just claimed. Notice that we assumed $(0,0) \in C(x)$, which is not possible as we are covering the case $\max_a \left(W(a,x) - \widetilde{W}(a,x) \right) > 0$ and (6) holds. The next argument extends this idea.

We now show that there always exists some $a'' \in D(x)$ for which the inequality $W(a',x) \geq W(a'',x)$ is not directly implied from $W(a',x) \geq W(a,x)$ for all $a \neq a', a''$. By contradiction, as we illustrated above, we will show that if this were not true, then $(0,0,...,0) \in C(x)$ which is not possible as $\max_a \left(W(a,x) - \widetilde{W}(a,x) \right) > 0$ and (6) holds.

For each $a' \in C(x)$, 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 one single good. To see this, let a'' be equal to a' except for some good i that is chosen at a' but not at a''. Then $W(a',x) \geq W(a'',x)$ if and only if

$$u_i(x') + \widetilde{x}_i + \Phi(\mathcal{S}(a'), x') \geq \Phi(\mathcal{S}(a''), x').$$

All other inequalities, $W(a',x) \geq W(a,x)$ 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 $(\widetilde{x}_i)_{i \leq n}$ such that $W(a',x) \geq W(a,x)$ with $a \neq a', a''$ and yet $W(a',x) < W(a'',x)$. Therefore assume $a'' \in C(x)$. By repeating this process $\|a'\|$ times (the number of goods being consumed in the bundle represented by a'), we need to assume $(0,0,...,0) \in C(x)$. But this is not possible, as we explained before.

A.7   Proof of Lemma 2

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

$$V(a'_i, a_{-i}, x, \varepsilon) - V(a_i, a_{-i}, x, \varepsilon) \leq (<) 0 \implies V\left(a'_i, a'_{-i}, x, \varepsilon\right) - V\left(a_i, a'_{-i}, x, \varepsilon\right) \leq (<) 0. \quad (7)$$

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.8   Proof of Lemma 3

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 $\widehat{a} \in \{0,1\}^n$.

A.9   Proof of Proposition 2

The proof of this result follows directly from Ui (2000, Theorem 3).

A.10   Proof of Corollary 2

By Assumptions 6' and 7', Proposition 2 and Theorem 3, we can identify $\left((u_i)_{i \leq n}, \Phi, F_{\varepsilon|x}\right)$. Because $\Phi(\mathcal{S}(a), x) \equiv \sum_{S \in \mathcal{S}(a)} \tilde{\Phi}(S, x)$, it is readily verified that we can recover $\tilde{\Phi}$ from Φ.

B   Monotone Comparative Statics and Stochastic Dominance

The proof of Lemma 1 relies on Topkis' theorem (Topkis (1998)) and the concept of stochastic dominance. We outline a simple version of these concepts.

Proposition 3. *Topkis' theorem:* *Let* $f(a_1, a_2, x) : \mathcal{A}_1 \times \mathcal{A}_2 \times \mathbb{R} \to \mathbb{R}$, *where* $\mathcal{A}_1$ *and* $\mathcal{A}_2$ *are finite ordered sets. Assume that* $f(a_1, a_2, x)$ *(i) is supermodular in* (a_1, a_2); *and (ii) has increasing differences in* (a_1, x) *and* (a_2, x). *Then,* $\arg\max\{f(a_1, a_2, x) \mid (a_1, a_2) \in \mathcal{A}_1 \times \mathcal{A}_2\}$ *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* $a \wedge b \in \mathcal{B}$.)

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 $x' \in U$.

---

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

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

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

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

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 $U \subset \mathbb{R}^n$.

References

[1] Andrews, D., and M. Schafgans (1998). Semiparametric Estimation of the Intercept of a Sample Selection Model. Review of Economic Studies 65, 497–517.

[2] Aradillas-Lopez, A. (2010). Semiparametric Estimation of a Simultaneous Game with Incomplete Information. Journal of Econometrics 157, 409–431.

[3] Aradillas-Lopez, A., and A. Gandhi (2012). Robust Inference in Ordered Response Games with Incomplete Information: Are Firms Strategic Substitutes? University of Wisconsin working paper.

[4] Athey, S., and S. Stern (1998). An Empirical Framework for Testing Theories About Complementarity in Organizational Design. Working paper.

[5] Bajari, P., H. Hong, and S. Ryan (2010). Identification and Estimation of a Discrete Game of Complete Information. Econometrica 78, 1529–1568.

[6] Beresteanu, A., I. Molchanov, and F. Molinari (2011). Identification Regions in Models with Convex Moment Predictions. Econometrica 79, 1785–1821.

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

[8] Berry, S., and P. Haile (2010). Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers. Yale University working paper.

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

[10] 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.

[11] Bjorn, P., and Q. Vuong (1984). Simultaneous Equations Models for Dummy Endogenous Variables: A game theoretic formulation with application to labor force participation. Mimeo.

[12] Blume, L. (1993). The Statistical Mechanics of Strategic Interaction. Games Economic Behavior 5, 387–424.

[13] Bresnahan, T., and P. Reiss (1991a). Entry and Competition in Concentrated Markets. Journal of Political Economy 99, 977–1009.

[14] Bresnahan, T., and P. Reiss (1991b). Empirical Models of Discrete Games. Journal of Econometrics 48, 57–81.

[15] Brock, W., and S. Durlauf (2001). Discrete Choice with Social Interactions. Review Economics Studies 68, 235–260.

[16] Brock, W., and S. Durlauf (2007). Identification of Binary Choice models with Social Interactions. Journal of Econometrics 140, 52–75.

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

[18] Chen, R., and Y. Chen (2011). The Potential of Social Identity for Equilibrium Selection. American Economic Review 101, 2562–2589.

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

[20] Ciliberto, F., and E. Tamer (2009). Market Structure and Multiple Equilibria in Airline Markets. Econometrica 77, 1791–1828.

[21] Cooper, R., and A. John (1988). Coordinating Coordination Failures in Keynesian Models. Quarterly Journal of Economics 103, 441–463.

[22] De Paula, A., and X. Tang (2011). Inference of Signs of Interaction Effects in Simultaneous Games with Incomplete Information. Econometrica 80, 143–172.

[23] Dong, Y., and A. Lewbel (2012). A Simple Estimator for Binary Choice Models With Endogenous Regressors. Boston College working paper.

[24] Dubé, J. (2004). Multiple Discreteness and Product Differentiation: Demand for Carbonated Soft Drinks. Marketing Science 23, 66–81.

[25] Dunker, F., S. Hoderlein, and H. Kaido (2012). Random Coefficients in Static Games of Complete Information. Boston College working paper.

[26] Echenique, F. (2004). A Characterization of Strategic Complementarities. Games and Economic Behavior 46, 325–347.

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

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

[29] Galichon, A., and M. Henry (2011). Set Identification in Models with Multiple Equilibria. Review of Economic Studies 78, 1264–1298.

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

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

[32] Goeree, J., and C. Holt (2005). An Experimental Study of Costly Coordination. Games Economics Behavior 51, 349–364.

[33] Gowrisankaran, G., and J. Stavins (2004). Network Externalities and Technology Adoption: Lessons from Electronic Payments. RAND Journal of Economics 35, 260–276.

[34] Grieco, P. (2012). Discrete Games with Flexible Information Structures: An Application to Local Grocery Markets. Penn State working paper.

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

[36] Hartman, W. (2010). Demand Estimation with Social Interactions and the Implications for Targeted Marketing. Journal Marketing Science 29, 585–601.

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

[38] Hendel, I. (1999). Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns. Review of Economic Studies 66, 423–446.

[39] Kajii, A., and S. Morris (1997a). The Robustness of Equilibria to Incomplete Information. Econometrica 65, 1283–1309.

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

[41] Khan, S., and D. Nekipelov (2012). Information Structure and Statistical Information in Discrete Response Models. Duke University working paper.

[42] Khan, S., and E. Tamer (2010). Irregular Identification, Support Conditions and Inverse Weight Estimation. Econometrica 78, 2021–2042.

[43] Kim, K., and I. Sher (2012). Identification of Demand Models of Multiple Purchases. University of Minnesota working paper.

[44] Lewbel, A. (1998). Semiparametric Latent Variable Model Estimation with Endogenous or Mismeasured Regressors. Econometrica 66, 105–121.

[45] Lewbel, A. (2000). Semiparametric Qualitative Response Model Estimation with Unknown Heteroskedasticity or Instrumental Variables. Journal of Econometrics 97, 145–177.

[46] Lewbel, A., and X. Tang (2011). Identification and Estimation of Games with Incomplete Information using Excluded Regressors. Boston College working paper.

[47] Magnac, T., and E. Maurin (2007). Identification and Information in Monotone Binary Models. Journal of Econometrics 139, 76–104.

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

[49] Matzkin, R. (2008). Identification in Nonparametric Simultaneous Equations. Econometrica 76, 945–978.

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

[51] Monderer, D., and L. Shapley (1996). Potential Games. Games Economic Behavior 14, 124-143.

[52] Morris, S., and T. Ui (2005). Generalized Potentials and Robust Sets of Equilibria. Journal of Economic Theory 124, 45–78.

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

[54] Qin, C. (1996). Endogenous Formation of Cooperation Structures. Journal Economic Theory 69, 218–226

[55] Rosenthal, R. (1973). A Class of Games Possessing Pure-strategy Nash Equilibria. International Journal of Game Theory 2, 65–67.

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

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

[58] Ui, T. (2000). A Shapley Value Representation of Potential Games. Games Economic Behavior 31, 121–135.

[59] Ui, T. (2001). Robust Equilibria of Potential Games. Econometrica 69, 1373–1380.

[60] Ui, T. (2008). Discrete Concavity for Potential Games. International Journal of Game Theory Review 10, 137–143.

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