Simultaneous Choice Models: The Sandwich Approach to Nonparametric Analysis (Preliminary Draft)

Natalia Lazzati*†

March 09, 2014

Abstract

We study simultaneous choice models from a revealed preference approach given limited data. By limited data we mean that we observe a single equilibrium from the equilibrium set for a collection of related models, or games. The objective of our analysis is twofold: We first use exclusion and monotone shape restrictions to provide out-of-sample predictions of equilibrium points. We then propose conditions on the data so that the empirical evidence can be rationalized as the Nash equilibria of a supermodular game. The approach we propose is nonparametric, allows for unobserved heterogeneity, and does not rely on any equilibrium selection mechanism.

JEL codes: C31, D01

Keywords: Simultaneous choice models, supermodular games, partial identification, out-of-sample predictions, monotone comparative statics, revealed preferences.

*Natalia Lazzati, Department of Economics, University of Michigan, Ann Arbor (e-mail: nlazzati@umich.edu).

†I thank Dan Ackerberg, Andres Aradillas-Lopez, Panle Jia, Amilcar Menichini, and John Quah for very useful comments. I also thank participants at the economics seminars of Cornell U, Penn State U, U of Pittsburgh, Carnegie Mellon U, and Binghamton U.

1 Introduction

Much of the empirical content of economics lies in the comparative statics predictions it generates [Kreps (2013)].

The celebrated monotone comparative statics (MCS) for games is a powerful yet simple method to contrast equilibrium points for an indexed family of models as we vary the parameter of interest.[1] From a theoretical perspective, MCS has proved very useful. There are at least two plausible explanations for its success: First, it works under conditions that are often easy to justify on economic grounds. Second, it allows us to compare equilibrium points in games with multiple solutions. From an empirical perspective, on the other hand, it has not received much attention. One of the reasons is that, with partial observability of equilibrium points, the standard predictions of MCS are neither testable nor useful for counterfactual analysis unless we impose specific equilibrium selection rules on the data-generating process.[2] One may conclude too readily that, in contrast to our initial claim, MCS has little empirical content and is thereby of limited use in the applied studies of games.

We give a new look at the importance of MCS for counterfactual predictions in the context of simultaneous choice models. The results we provide can be applied to a wide variety of models, such as those of consumer behavior in network markets, strategic R&D, crime decisions, peer effects in schooling, among many others. The idea is simple: There is a set of agents that, under different circumstances, make interdependent decisions. We capture each of these circumstances by a set of covariates. For instance, in a model of demand with network effects these covariates might be the price of the good and the income levels. They may also refer to individual characteristics such as sex or age. We have limited data on past equilibrium choices for various realized covariates. The data are limited in that, for each realized set of covariate values, we observe only one element from the equilibrium set: the equilibrium selected by the group. It is also limited because we only have access to a few covariate values. The objective of our

[1] See, e.g., Milgrom and Roberts (1990, 1994), Milgrom and Shannon (1994), Vives (1990), and Topkis (1979).

[2] Echenique and Komunjer (2009) provide a test for MCS that is based on a stochastic equilibrium selection devise that places strict positive probability on the extremal elements of the equilibrium set. Lazzati (2014) takes advantage of specific equilibrium selection rules to provide counterfactual predictions in models of social interaction via MCS. Uetake and Watanabe (2013) exploit MCS to provide parametric estimation results for supermodular games.

analysis is twofold: We first want to combine the data with credible restrictions to predict the set of actions that could be taken by the group members if they were to face a set of covariates that differ from the ones we observe in the data. We refer to this set of choices as either the counterfactual or the out-of-sample predictions. Our second objective is to state conditions on the data under which they can be rationalized as the Nash equilibria of a supermodular game.

The approach we propose to achieve our goals is nonparametric and does not impose any equilibrium selection mechanism on the data-generating process.[3] It relies on exclusion and monotone shape restrictions on the choice function of each agent. An advantage of our approach is that the assumptions we impose are indeed testable. That is, they can be proved false given the data. By exclusion restrictions we mean that individual covariates or characteristics may affect the choice of one agent without entering the choice functions of the other agents in the group. For instance, in the model of demand with network effects, income levels are natural candidates for the exclusion restrictions. In the same model, prices may or may not play the role of exclusion restrictions depending on whether we allow for price discrimination. These exclusion restrictions allow us to track the choice function of each individual for alternative action profiles of the others. We elaborate next on the role of the monotone shape restrictions we sustain.

The basics of our model are a system of interdependent behavioral functions (or best-replies in a game) that relate the choice of each agent to the decisions of the other group members, its characteristics or covariates, and a term of unobserved heterogeneity. The two monotone conditions we impose are as follows: the choice function of each individual increases with the action profile of the others; and it varies monotonically with the covariate levels. These conditions imply clear restrictions on the equilibrium set: The system of structural equations has a smallest and a largest solution. In addition, these extremal solutions vary monotonically with the covariates. This is the celebrated MCS of equilibrium points for games. Since this comparative statics result only refers to the extremal equilibria, it is neither testable nor useful for counterfactual predictions unless we impose specific equilibrium selection rules on the data-generating process. However, we argue that the opposite is true regarding its two underlying assumptions. Specifically, when combined with the exclusion restrictions, they allow us to

[3] The econometric literature has recently emphasized on the relevance of monotone restrictions to provide identification results. See, e.g., Kline and Tamer (2012), Lazzati (2014), and Manski (1997, 2011).

construct sharp bounds for individual choice functions given the empirical evidence. By sharp we mean that the true set of choice functions must lie between the bounds we construct, and that we cannot reject the hypothesis that the true set coincides with one or the other. We then use these extremal choice functions to bound equilibrium behavior by an approach that relies on a nonparametric version of the MCS method. This result addresses our initial objective of analysis, namely, deriving out-of-sample predictions of equilibrium points. **We refer to our method as the sandwich approach to the out-of-sample predictions**.

We present two sets of results for the out-of-sample predictions. We first follow the standard revealed preference approach pioneered by Samuelson, and assume that choices are non-random. In other words, we assume that agents with the same covariates or characteristics behave identically. This initial setting facilitates a transparent exposition of our main ideas.[4] We then extend our results to the case with unobserved heterogeneity by taking advantage of the method proposed by Galichon and Henry (2011) to identify models with multiple equilibria (see also Beresteanu, Molinari, and Molchanov (2011) and Cheser and Rosen (2012)). The incorporation of unobserved heterogeneity into the analysis facilitates the empirical implementation of the results we offer.

The second objective of our analysis is to show that, if the data satisfy the monotone conditions we propose, then the empirical evidence can always be rationalized as the Nash equilibria of a supermodular game.[5] This result provides a theoretical justification of the assumptions we impose as well as an assessment of the empirical content of this class of games. Given the previous results, we achieve our second goal by showing that any set of monotone choice functions can be rationalized as the set of best-reply functions of a game among the agents. Moreover, we show that we can always build the payoffs of the players in such a way that they satisfy the conditions of the supermodular games. **We refer to this result as the supermodular rationalization of the data**.

The rest of the paper is organized as follows. Section 2 presents the model and the objective of our analysis. Section 3 provides out-of-sample predictions and rationalization results in the

[4] As Echenique et al. (2013) explain, this restriction can also be justified as a worst case scenario for testability and identification.

[5] The literature on testability of simultaneous choice models includes, among other papers, Browning and Chiappori (1998), Cherchye, Demuynck, and De Rock (2013), Carvajal (2002), Carbajal et al. (2013), Deb (2009), Echenique et al. (2013), and Sprumont (2000).

context of no unobservables. Section 4 extends this analysis to the case with unobserved heterogeneity and non-monotone covariates. Section 5 includes some extensions and remarks, and Section 6 concludes. All proofs are collected in Section 7.

2 The Model

This section presents the model and elaborates on the objective of our analysis.

Consider a set of agents $\mathrm{N} = \{1, 2, ..., n\}$. Agent $i \in \mathrm{N}$ selects an element a_i from a set of available options $\mathrm{A}_i \subseteq \mathbb{R}^d$. We assume A_i is a complete lattice.[6] His decision depends on the options selected by the others, $a_{-i} = (a_j : j \in \mathrm{N}, j \neq i) \in \mathrm{A}_{-i} = \times_{j \in \mathrm{N}, j \neq i} \mathrm{A}_j$, a set of covariates or individual characteristics $\tau_i \in \mathrm{Z}_i$, and a term $\varepsilon_i \in \mathcal{E}_i$ that reflects unobserved individual heterogeneity. We assume the spaces of actions and covariates are finite. The behavior of agent i is described by the choice function

$$\mathrm{C}_i (a_{-i}, \tau_i, \varepsilon_i) : \mathrm{A}_{-i} \times \mathrm{Z}_i \times \mathcal{E}_i \rightarrow \mathrm{A}_i.$$

If the choice functions derive from an underlying game, then

$$\mathrm{C}_i (a_{-i}, \tau_i, \varepsilon_i) = \arg\max \{\mathrm{U}_i (a_i, a_{-i}, \tau_i, \varepsilon_i) : a_i \in \mathrm{A}_i\} \tag{1}$$

where U_i is the payoff function of agent i and C_i can be interpreted as his best-reply function.

Let $\mathrm{C} = (\mathrm{C}_i : i \in \mathrm{N})$ be the collection of choice functions of all agents in N. We refer to C as the microeconomic specification of our model. Putting all the ingredients together, our model can be described by the tuple $\Gamma_{\tau,\varepsilon} = (\mathrm{N}, \mathrm{C}, \mathrm{A})$ with $\mathrm{A} = (\mathrm{A}_i : i \in \mathrm{N})$, $\tau = (\tau_i : i \in \mathrm{N}) \in \mathrm{Z} = \times_{i \in \mathrm{N}} \mathrm{Z}_i$, and $\varepsilon = (\varepsilon_i : i \in \mathrm{N}) \in \mathcal{E} = \times_{i \in \mathrm{N}} \mathcal{E}_i$. We next introduce our solution concept.

Definition (Solution Concept): We say $\mu(\tau, \varepsilon) = (\mu_i(\tau, \varepsilon) : i \in \mathrm{N})$ is consistent with equilibrium behavior generated by $\Gamma_{\tau,\varepsilon}$ if, for each $i \in \mathrm{N}$,

$$\mu_i(\tau, \varepsilon) = \mathrm{C}_i (a_{-i}, \tau_i, \varepsilon_i) \quad \text{with} \quad a_{-i} = \left(\mu_j(\tau, \varepsilon) : j \in \mathrm{N}, j \neq i\right).$$

[6] A partially ordered set $(\mathrm{L}, \leq)$ is a complete lattice if every subset A of L has both a greatest lower bound or infimum (i.e., the meet) and a least upper bound or supremum (i.e., the join) in $(\mathrm{L}, \leq)$.

We indicate by $\Delta_{\tau,\varepsilon} : \mathrm{Z} \times \mathcal{E} \rightrightarrows \mathrm{A}$, with $\Delta_{\tau,\varepsilon} \subseteq \mathrm{A}$, the equilibrium correspondence.[7]

That is, we say $\mu(\tau, \varepsilon)$ is consistent with equilibrium behavior if each agent selects an option according to both his choice function and the equilibrium behavior of the other agents. In a game theoretic context, our solution concept is a Nash equilibrium in pure strategies.

We are now ready to describe the objective of our analysis.

Objective of Analysis: We have limited data on realized choices for a set of covariates

$$a^m = (a_i^m : i \in \mathrm{N}) \text{ and } \tau^m = (\tau_i^m : i \in \mathrm{N}) \text{ for } m = 1, ..., M.$$

We assume these observations are consistent with equilibrium behavior. Our data is limited because, for each realized covariates, we only observe an element of the equilibrium set: the equilibrium selected by the group. It is also limited in that we have access to a restricted set of covariate values. **Our first objective is to learn the actions that could be taken by this set of agents if they were to face a set of covariates and individual heterogeneity** τ, ε. In other words, we observe an equilibrium outcome $\mu(\tau^m, \varepsilon^m) \in \Delta_{\tau^m,\varepsilon^m}$ for a set of realized covariates $(\tau^m : m \leq M)$ and want to use this information to learn about the equilibrium set $\Delta_{\tau,\varepsilon}$ for an alternative τ, ε. **Our second objective is to state conditions on the data under which the empirical evidence can be rationalized as the Nash equilibria of a supermodular game among the agents.** This is the same as to provide conditions on the data so that we can construct a set of payoff functions $\mathrm{U} = (\mathrm{U}_i : i \in \mathrm{N})$, with each U_i given as in (1) and satisfying the conditions of the supermodular games, that are able to generate the empirical observations as the Nash equilibria of the induced game.

The next example illustrates the concepts and ideas we just described. We will invoke this example again in the next section to motivate the restrictions we impose to the model.

Example 1: Let us consider a demand model with network effects. The choice function of each agent indicates whether the person acquires a network good. The covariates consist on

[7] Nash equilibrium in pure strategies requires two conditions: (i) individual optimization performed by each player; and (ii) consistency of choices across players. The distinction between these two conditions has been recently emphasized by De Clippel (2014) who provides implementation results in mechanism design by only relying on condition (ii). In line with his results, the out-of-sample predictions we provide only rely on the second requirement.

the price the individual has to pay for the good, p, and his income level, m_i. In this model,

$$\mathrm{C}_i : \{0,1\}^{n-1} \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{E}_i \to \{0,1\} \quad \text{with} \quad \mathrm{C}_i \left(a_{-i}, p, m_i, \varepsilon_i\right) \in \{0,1\}$$

represents the demand function of consumer i.[8] An equilibrium $\mu\left(p, m, \varepsilon\right)$, with $m = (m_i : i \in \mathrm{N})$, is a vector in $\{0,1\}^n$ that indicates the set of people that acquires the good.

We have limited data on past consumption decisions for a set of different prices and income levels. We want to use the empirical evidence to predict market demand for a vector p, m that differs from the ones we observe in the data. The method we use allows for price discrimination. That is, we can let prices depend on the names of the agents by substituting p by p_i. In this context, the counterfactual predictions could be used to evaluate target marketing strategies that aim to identify the influential consumers, that is, the set of consumers that have a large impact on the choices of other consumers in the economy (see, e.g., Candogan et al. (2012)). ■

3 Another Look at Monotone Comparative Statics

This section provides out-of sample predictions and rationalization of equilibrium points assuming unobserved heterogeneity does not vary across observations. The next section allows the unobservables to be randomly distributed.

3.1 Out-of-Sample Predictions of Equilibrium Points

This sub-section addresses our first goal. That is, we provide out-of-sample predictions of equilibrium points in simultaneous choice models, or games, given limited data. We do so without imposing either functional form restrictions on the choice functions or equilibrium selection rules on the data-generating process. Instead, we take advantage of monotone shape restrictions that, as we show in the next sub-section, can be motivated by economic theory. We also explain the role of the exclusion restrictions in our analysis.

Simultaneous choice models may have no equilibrium. Lack of equilibrium existence means we cannot use our solution concept to predict consistent behavior. From an econometrics perspective, this possibility introduces additional difficulties into the analysis. For instance, data

[8]Though in this application p, m belong *a priori* to an infinite set $\mathbb{R}_+ \times \mathbb{R}_+$, we can always restrict the latter to a large, but finite, sub-space. In fact, this is the case when prices and incomes are measured in dollars and cents and we rule out the possibility of either infinite prices or income levels.

on past choices may not be informative about the beliefs of each agent with respect to the behavior of others and it is therefore hard to recover their choice functions. The set of restrictions we introduce rule out this possibility and, at the same time, facilitate the comparative statics analysis we exploit for counterfactual predictions.[9]

Our first condition requires the choice function of each agent to be increasing in the action profile of the other agents.

(A1) For each $i \in \mathrm{N}$, $\mathrm{C}_i\left(a_{-i}, \tau_i, \varepsilon_i\right)$ increases in a_{-i} for all τ_i, ε_i.

Under condition A1, our model displays positive interaction effects. In Example 1, cellphones and software are goods for which researchers and practitioners often believe this restriction is naturally fulfilled. The next result states that, if A1 holds, then the solution set is non-empty and has extremal elements for each set of covariates.

Proposition 1 *Under A1, there exist extremal equilibria, $\underline{\mu}$ and $\overline{\mu}$, such that any μ of $\Delta_{\tau,\varepsilon}$ satisfies $\overline{\mu}\left(\tau, \varepsilon\right) \geq \mu \geq \underline{\mu}\left(\tau, \varepsilon\right)$.*

Proposition 1 follows directly from Tarski's fixed point theorem. Though important, the existence result we just described does not impose enough structure on the model in order to make the empirical evidence really informative. To do so, we will connect the alternative models in such a way that we can use the equilibrium behavior revealed by the data for a set of covariates to learn about the equilibrium set corresponding to a different set of covariate values. We next add a second restriction on C that establishes this connection.

(A2) For each $i \in \mathrm{N}$, $\mathrm{C}_i\left(a_{-i}, \tau_i, \varepsilon_i\right)$ increases in τ_i for all a_{-i}, ε_i.

Remark The monotonicity of choices with respect to covariates makes the out-of-sample predictions more informative. Nevertheless, as we explain in Sub-section 4.2, all our results can still be applied if some (or all) of the covariates have non-monotone effects on choices. On the other hand, the monotonicity behind A1 is critical to our analysis.

[9] Chesher and Rosen (2012) provide identification results for models that have no solution for some covariate values. Many other papers guarantee non-emptiness solution sets by allowing for mixed strategy equilibria (see, e.g., Beresteanu, Molinari, and Molchanov (2011)). Echenique and Edlin (2004) show that, under the assumptions we impose, mixed strategy equilibria are often unstable. Their result offers a plausible justification for the solution concept we use in this work.

In Example 1, this condition requires the demand function of agent i to be increasing in $\tau_i = -p, m_i$. (Note that we can always reverse the order of one covariate to make this condition work.) The next proposition states that, under A1 and A2, we can compare the extremal elements of the equilibrium sets for every pair of ordered covariate values.

Proposition 2 *Under A1 and A2,* $\underline{\mu}(\tau, \varepsilon)$ *and* $\overline{\mu}(\tau, \varepsilon)$ *increase in* τ.

Proposition 2 is the standard monotone comparative statics (MCS) result of equilibrium points. We next explain, via example, the difficulties of using this result directly for out-of-sample predictions. To this end, we show that the MCS result, in and of itself, is not useful for counterfactual analysis with partial observability of equilibrium points unless we impose specific equilibrium selection rules in the data-generating process. We then show that its underlying assumptions—namely, A1 and A2—are nevertheless extremely powerful to this end. We finally use the example to highlight the role of the exclusion restrictions in our work.

Example 2: Let $\mathrm{N} = \{1, 2\}$, $\mathrm{Z}_1 = \mathrm{Z}_2 = \{\underline{\tau}, \overline{\tau}\}$ (with $\overline{\tau} > \underline{\tau}$), and $\mathrm{A}_i = \{\underline{a}_i, \overline{a}_i\}$ (with $\overline{a}_i > \underline{a}_i$) for $i = 1, 2$. (To simplify the exposition, we avoid unobserved heterogeneity in this example.) The data we have are described by the next tables.

Data

		$\underline{a}_2$	$\overline{a}_2$
$\tau_1 = \overline{\tau}, \tau_2 = \overline{\tau}$	$\underline{a}_1$		
	$\overline{a}_1$		$\bullet$

		$\underline{a}_2$	$\overline{a}_2$
$\tau_1 = \underline{\tau}, \tau_2 = \overline{\tau}$	$\underline{a}_1$		$\bullet$
	$\overline{a}_1$		

		$\underline{a}_2$	$\overline{a}_2$
$\tau_1 = \overline{\tau}, \tau_2 = \underline{\tau}$	$\underline{a}_1$	$\bullet$	
	$\overline{a}_1$		

That is, we know the actions selected by these two agents under three different sets of covariates. The selected profiles are indicated by dots in the tables. We also know these data were generated by equilibrium behavior. Our goal is to describe all possible equilibrium profiles for the low covariate levels that are consistent with the empirical evidence and satisfy our initial assumptions. Specifically, we want to characterize $\Delta(\underline{\tau}, \underline{\tau})$. We write ✓ for consistent profiles and use x for the inconsistent ones.

It is readily verified that the MCS result, in and of itself, does not allow us to rule out any action profile. That is, given the data we have, by only relying on this prediction, every single pair of choices is consistent with equilibrium behavior. Nevertheless, if we know that A1 and A2 are satisfied, then we can make a unique prediction. The table below captures this observation.

Out-of-Sample Prediction Under A1 and A2

		$\underline{a}_2$	$\overline{a}_2$
$\tau_1 = \underline{\tau}, \tau_2 = \underline{\tau}$	$\underline{a}_1$	✓	x
	$\overline{a}_1$	x	x

We now elaborate on how we find $\mu(\underline{\tau}, \underline{\tau})$ under A1 and A2. We know the data corresponds to equilibrium behavior. Thus, each data point allows us to recover an element of the choice functions, e.g., $\mathrm{C}_1(\overline{a}_2, \underline{\tau}) = \underline{a}_1$. Once we recover all these elements we can invoke A1 and A2 to infer the other ones. For instance, since $\mathrm{C}_1(\overline{a}_2, \underline{\tau}) = \underline{a}_1$, then, by A1, it must be that $\mathrm{C}_1(\underline{a}_2, \underline{\tau}) = \underline{a}_1$. Proceeding in this way, there is only one element for which we do not have any information, $\mathrm{C}_2(\overline{a}_1, \underline{\tau})$. For this element, all we can say is that $\mathrm{C}_2(\overline{a}_1, \underline{\tau}) \in \{\underline{a}_2, \overline{a}_2\}$. Using this procedure, we find two sets of choice functions that are consistent with the data and satisfy A1 and A2, namely, B and D below.

$$\mathrm{B} = \begin{pmatrix} \mathrm{B}_1(a_2, \tau_1) = \underline{a}_1 1(a_2, \tau_1 < \overline{a}_2, \overline{\tau}) + \overline{a}_1 [1 - 1(a_2, \tau_1 < \overline{a}_2, \overline{\tau})] \\ \mathrm{B}_2(a_1, \tau_2) = \underline{a}_2 1(a_1, \tau_2 \leq \overline{a}_1, \underline{\tau}) + \overline{a}_2 [1 - 1(a_1, \tau_2 \leq \overline{a}_1, \underline{\tau})] \end{pmatrix}$$

$$\mathrm{D} = \begin{pmatrix} \mathrm{D}_1(a_2, \tau_1) = \underline{a}_1 1(a_2, \tau_1 < \overline{a}_2, \overline{\tau}) + \overline{a}_1 [1 - 1(a_2, \tau_1 < \overline{a}_2, \overline{\tau})] \\ \mathrm{D}_2(a_1, \tau_2) = \underline{a}_2 1(a_1, \tau_2 \leq \underline{a}_1, \underline{\tau}) + \overline{a}_2 [1 - 1(a_1, \tau_2 \leq \underline{a}_1, \underline{\tau})] \end{pmatrix}$$

We finally compute all possible equilibria for B and D. These action profiles constitute our out-of-sample predictions. Doing so, we find only one equilibrium, $(\underline{a}_1, \underline{a}_2)$. This result illustrates the power of A1 and A2 to obtain out-of-sample prediction of equilibrium points and it is the leading force of the analysis that follows.

In this example, by exclusion restrictions we mean that the covariates can take on different values for the two individuals. This allows us to track the choice function of each agent for a given covariate value and alternative choices of the other one. While the method we propose for out-of-sample predictions can still be applied without these exclusion restrictions, if they do not hold, then our bounds for equilibrium behavior would almost always be just the trivial ones, i.e., the extremal elements of the set of available options. To illustrate this claim, note that, in absence of the exclusion restrictions, the available data and the corresponding predictions would be as follows.

Data

$\tau_1 = \overline{\tau}, \tau_2 = \overline{\tau}$		$\underline{a}_2$	$\overline{a}_2$
	$\underline{a}_1$		
	$\overline{a}_1$		•

Out-of-Sample Prediction Under A1 and A2

$\tau_1 = \underline{\tau}, \tau_2 = \underline{\tau}$		$\underline{a}_2$	$\overline{a}_2$
	$\underline{a}_1$	✓	✓
	$\overline{a}_1$	✓	✓

Thus, even assuming that A1 and A2 hold, we cannot rule out any action profile. ■

Example 2 shows the power of A1 and A2 and the exclusion restrictions for making counterfactual predictions. The problem with the approach we just described is that it becomes less tractable as we add more feasible options. We next show that, in our model, the out-of-sample predictions can always be obtained by an easy-to-implement procedure.

Let us indicate by

$$\mathrm{B} = (\mathrm{B}_i : i \in \mathrm{N}) \quad \text{and} \quad \mathrm{D} = (\mathrm{D}_i : i \in \mathrm{N})$$

two microeconomic specifications for the model. Let $\underline{\nu}, \overline{\nu}$ and $\underline{\lambda}, \overline{\lambda}$ be the extremal equilibria corresponding to B and D, respectively, when such equilibria exist. We say $\mathrm{D} \geq \mathrm{C} \geq \mathrm{B}$ if, for each $i \in \mathrm{N}$, we have that

$$\mathrm{D}_i (a_{-i}, \tau_i, \varepsilon_i) \geq \mathrm{C}_i (a_{-i}, \tau_i, \varepsilon_i) \geq \mathrm{B}_i (a_{-i}, \tau_i, \varepsilon_i) \text{ for all } a_{-i}, \tau_i, \varepsilon_i.$$

The next proposition states that if B, C, and D satisfy A1 and C is *sandwiched* by B and D as above, then the extremal equilibria associated to B and D bound every element of the equilibrium set corresponding to C.

Proposition 3 *Assume B, C, and D satisfy A1 and $D \geq C \geq B$. Then, for each τ, ε,*

$$\underline{\lambda}(\tau, \varepsilon) \geq \underline{\mu}(\tau, \varepsilon) \geq \underline{\nu}(\tau, \varepsilon) \quad \textit{and} \quad \overline{\lambda}(\tau, \varepsilon) \geq \overline{\mu}(\tau, \varepsilon) \geq \overline{\nu}(\tau, \varepsilon).$$

From a theoretical perspective, Proposition 3 is just a nonparametric variant of Proposition 2.[10] We next show that if the data is consistent with our modelling restrictions, then the set of all choice functions that could have generated the empirical evidence has extremal elements such as B and D above. After doing so, we take advantage of our last result to bound out-of-sample predictions of equilibrium points.

Recall that we have limited data on realized equilibrium choices for a set of covariates

$$a^m = (a_i^m : i \in \mathrm{N}) \text{ and } \tau^m = (\tau_i^m : i \in \mathrm{N}) \text{ for } m = 1, ..., M.$$

We indicate by $\mathcal{D}_i$ the set of all a_{-i}^m, τ_i^m with $m \leq M$, and by $\mathcal{A}_i$ the set of all a^m with $m \leq M$. For each $i \in \mathrm{N}$, the observed choice function

$$\mathrm{C}_i^{\mathrm{data}} : \mathcal{D}_i \to \mathcal{A}_i \text{ with } \mathrm{C}_i^{\mathrm{data}}\left(a_{-i}^m, \tau_i^m, \varepsilon_i\right) = a_i^m$$

associates the action that has been chosen by the agent to each observed action profile of the other agents and covariates. We say C is consistent with our model if it satisfies A1 and A2 and, for each $i \in \mathrm{N}$, we have that

$$\mathrm{C}_i(a_{-i}, \tau_i, \varepsilon_i) = \mathrm{C}_i^{\mathrm{data}}(a_{-i}, \tau_i, \varepsilon_i) \text{ for all } a_{-i}, \tau_i \in \mathcal{D}_i.$$

That is, if C is monotone and extends $\mathrm{C}^{\mathrm{data}} = \left(\mathrm{C}_i^{\mathrm{data}} : i \in \mathrm{N}\right)$ to the full domain of the choice functions. We write $\mathrm{H}\{\mathrm{C}\}$ for the set of all consistent sets of choice functions.

If follows directly from our definition, that a consistent C exists only if, for each $i \in \mathrm{N}$,

$$\mathrm{C}_i^{\mathrm{data}}(a_{-i}, \tau_i, \varepsilon_i) \text{ increases in } a_{-i}, \tau_i \text{ on } \mathcal{D}_i.$$

That is, if $\mathrm{C}^{\mathrm{data}}$ satisfies A1 and A2. We show below that this condition is also sufficient. In addition, we show that $\mathrm{H}\{\mathrm{C}\}$ has a nice structure. It is a complete lattice in the sense that, for each subset $\mathrm{h} \subseteq \mathrm{H}\{\mathrm{C}\}$, we have that

$$\inf(\mathrm{h}) \in \mathrm{H}\{\mathrm{C}\} \quad \text{and} \quad \sup(\mathrm{h}) \in \mathrm{H}\{\mathrm{C}\}$$

[10] Amir (2008) shows that a similar result can be obtained even if the intermediate mapping is not increasing.

where inf and sup of h are calculated pointwise for the choice function of each agent in N. That is, for each collection of consistent choice functions $\mathrm{h} = \left(\mathrm{C}^k : k \in \mathrm{K}\right)$, we let $\inf(\mathrm{h}) = \left(\inf\left(\mathrm{C}_i^k : k \in \mathrm{K}\right) : i \in \mathrm{N}\right)$ where $\inf\left(\mathrm{C}_i^k : k \in \mathrm{K}\right)$ indicates the pointwise infimum of the set of functions corresponding to agent $i \in \mathrm{N}$. We similarly define $\sup(\mathrm{h})$. Finally, the extremal elements of H{C}, $\mathrm{B} = (\mathrm{B}_i : i \in \mathrm{N})$ and $\mathrm{D} = (\mathrm{D}_i : i \in \mathrm{N})$, can be easily described: for each $i \in \mathrm{N}$ and all $a_{-i}, \tau_i \in \mathrm{A}_{-i} \times \mathrm{Z}_i$,

$$\begin{aligned} \mathrm{B}_i\left(a_{-i}, \tau_i, \varepsilon_i\right) &= \inf\left\{a_i \in \mathrm{A}_i : a_i \geq \sup\left\{a_i^m : \tau_i^m, a_{-i}^m \leq \tau_i, a_{-i}, m \leq M\right\}\right\} \\ \mathrm{D}_i\left(a_{-i}, \tau_i, \varepsilon_i\right) &= \sup\left\{a_i \in \mathrm{A}_i : a_i \leq \inf\left\{a_i^m : \tau_i^m, a_{-i}^m \geq \tau_i, a_{-i}, m \leq M\right\}\right\} \end{aligned}.$$

The next proposition formalizes all these claims.

Proposition 4 *H{C} is non-empty if and only if C^{data} satisfies A1 and A2. Moreover, H{C} is a complete lattice with extremal elements given by B and D.*

From a technical standpoint, Proposition 4 relates to the literature on interpolation of monotone functions that map partially ordered sets into lattices. Our results substantially simplify the burden of recovering choice functions from the available data. As we explain next, they also facilitate the computation of out-of-sample predictions of equilibrium points.

We finally combine Propositions 3 and 4 to provide bounds for $\Delta_{\tau,\varepsilon}$. (Recall that $\underline{\nu}$ and $\overline{\lambda}$ indicate the smallest and the largest equilibrium corresponding to B and D, respectively.)

Corollary 5 *If C^{data} satisfies A1 and A2, then, for each τ, ε,*

$$\overline{\lambda}(\tau, \varepsilon) \geq \mu \geq \underline{\nu}(\tau, \varepsilon) \text{ for all } \mu \in \Delta_{\tau,\varepsilon}.$$

Moreover, the bounds are sharp.

Corollary 5 provides bounds for out-of-sample predictions of equilibrium points given the empirical evidence. By sharp we mean that the equilibrium behavior of the group under τ, ε can actually coincide with either $\overline{\lambda}(\tau, \varepsilon)$ or $\underline{\nu}(\tau, \varepsilon)$.

We next summarize the main findings in this sub-section.

The Sandwich Approach to the Out-of-Sample Predictions: Observed choices are consistent with our model if and only if they are monotone. When this is the case, for any value of the covariates or individual characteristics τ, we can construct two extremal sets of monotone choice functions (B and D) such that the true one (C) must lie between them and can actually coincide with one or the other. The extremal equilibria of these two sets of functions provide sharp bounds for the equilibrium set corresponding to τ.

3.2 Rationalization by a Supermodular Game

This sub-section addresses our second goal. That is, we show that if the data satisfies monotonicity, then the empirical evidence can always be rationalized as the Nash equilibria of a supermodular game. We describe this class of games next.[11]

Definition (Supermodular Games): Let $\mathrm{U} = (\mathrm{U}_i : i \in \mathrm{N})$ be the set of payoff functions of all players in N. When the choice functions result from an underlying game, then

$$\mathrm{C}_i(\tau_i, a_{-i}, \varepsilon_i) = \arg\max \{\mathrm{U}_i(a_i, a_{-i}, \tau_i, \varepsilon_i) : a_i \in \mathrm{A}_i\}$$

where U_i is the payoff function of agent i and C_i is his best-reply function.

We say U_i is supermodular in a_i if, for all $a_i, a_i' \in \mathrm{A}_i$,

$$\mathrm{U}_i\left(\sup\left(a_i, a_i'\right), a_{-i}, \tau_i, \varepsilon_i\right) + \mathrm{U}_i\left(\inf\left(a_i, a_i'\right), a_{-i}, \tau_i, \varepsilon_i\right) \geq \mathrm{U}_i(a_i, a_{-i}, \tau_i, \varepsilon_i) + \mathrm{U}_i\left(a_i', a_{-i}, \tau_i, \varepsilon_i\right).$$

In addition, U_i has increasing differences in $(a_i; a_{-i})$ if, for all $a_i \geq a_i'$ and $a_{-i} \geq a_{-i}'$,

$$\mathrm{U}_i(a_i, a_{-i}, \tau_i, \varepsilon_i) - \mathrm{U}_i\left(a_i', a_{-i}, \tau_i, \varepsilon_i\right) \geq \mathrm{U}_i\left(a_i, a_{-i}', \tau_i, \varepsilon_i\right) - \mathrm{U}_i\left(a_i', a_{-i}', \tau_i, \varepsilon_i\right).$$

That is, if the extra payoff of increasing own action increases with the choices of the others. If U_i is supermodular in a_i and has increasing differences in $(a_i; a_{-i})$, then $\mathrm{C}_i(a_{-i}, \tau_i, \varepsilon_i)$ increases in a_{-i} (i.e., A1 holds). These two conditions are the distinctive feature of any supermodular game.

We say U_i has increasing differences in $(a; \tau_i)$ if, for all $a_i \geq a_i'$ and $\tau_i \geq \tau_i'$,

$$\mathrm{U}_i(a_i, a_{-i}, \tau_i, \varepsilon_i) - \mathrm{U}_i\left(a_i', a_{-i}, \tau_i, \varepsilon_i\right) \geq \mathrm{U}_i\left(a_i, a_{-i}, \tau_i', \varepsilon_i\right) - \mathrm{U}_i\left(a_i', a_{-i}, \tau_i', \varepsilon_i\right).$$

[11] Our definition of supermodular games differs from the standard one in two regards. First, we assume best-replies are functions as opposed to just correspondences—we Sub-section 5.1 for a natural extension of our results to the case of multi-valued best-replies. Second, we require the payoff functions to have increasing differences between actions and covariates.

That is, if the extra payoff of increasing own action increases with τ_i. If U_i is supermodular in a_i and has increasing differences in $(a_i; \tau_i)$, then $\mathrm{C}_i(a_{-i}, \tau_i, \varepsilon_i)$ increases in τ_i (i.e., A2 holds).

Thus, the monotone assumptions A1 and A2 are satisfied by the class of supermodular games we just described and our results can be therefore applied to each them. We next show that the opposite is also true. That is, any set of monotone choice functions C can be rationalized as the set of best-replies of a supermodular game. Since we showed in the last sub-section that monotonicity of $\mathrm{C}^{\mathrm{data}}$ is necessary and sufficient for the existence of a consistent C, then this result completes the proof of our initial claim.

Before presenting the main idea, we define the concept of supermodular rationalization.

Definition (Supermodular Rationalization): We say a set of payoff functions $\mathrm{U} = (\mathrm{U}_i : i \in N)$, with $\mathrm{U}_i : \mathrm{A}_i \times \mathrm{A}_{-i} \times \mathrm{Z}_i \times \mathcal{E}_i \to \mathrm{A}_i$, rationalizes C if, for each $i \in \mathrm{N}$,

$$\mathrm{C}_i(\tau_i, a_{-i}, \varepsilon_i) = \arg\max \{\mathrm{U}_i(a_i, a_{-i}, \tau_i, \varepsilon_i) : a_i \in \mathrm{A}_i\}.$$

If, in addition, for each $i \in \mathrm{N}$, U_i is supermodular in a_i and has increasing differences in both $(a_i; a_{-i})$ and $(a_i; \tau_i)$, then we say U is a supermodular rationalization of C.

The next proposition formalizes our previous claim.

Proposition 6 *There is a supermodular rationalization U for any C that satisfies A1 and A2.*

Topkis (1998), Theorem 2.8.9, shows that any increasing collection of multi-valued functions defined on a complete ordered set (where increasing is defined with respect to the strong set order) can be represented as the set of unconstrained maxima of a supermodular function. Proposition 6 shows that, by restricting attention to functions as opposed to correspondences, we can extend this result to domains that are just partially ordered.

We next summarize the main findings in this sub-section.

Supermodular Rationalization: Observed choices are consistent with our model if and only if they can be rationalized as the Nash equilibrium of a supermodular game.

4 Unobservables and Non-Monotone Covariates

4.1 Unobserved Heterogeneity

This section incorporates unobserved heterogeneity into the analysis. Adding unobservables into the analysis is of critical importance for the empirical implementation of our ideas. We keep the role of covariates as exclusion restrictions by sustaining that unobservables do not depend on covariates. We next discuss two extra conditions.

The first additional restriction refers to the distribution of unobservables.

(A3) ε is distributed according to the probability function $\upsilon(\mathrm{U})$ for each $\mathrm{U} \subseteq \times_{i \in \mathrm{N}} [1, 2, ..., u_i]$.

Condition A3 requires the unobservables to be discretely distributed. Given that all observables in our model are discrete, this assumption does not impose any actual constraint. Saying differently, for any continuous distribution, there exists a discrete one that is able to generate the same observations. Without assuming any functional form restriction on the choice functions, the distribution of unobservables is generally unidentified. Thus, we assume it is known by the analyst (see, e.g., Blundell, Kristensen, and Matzkin (2012)).

Regarding the choice functions, we will add a monotone restriction that simply makes the analysis more tractable.

(A4) For each $i \in \mathrm{N}$, $\mathrm{C}_i(a_{-i}, \tau_i, \varepsilon_i)$ increases in ε_i for all a_{-i}, τ_i.

Condition A4 guarantees that the set of unobservables that can produce each action is an interval. That is,

$$\mathrm{C}_i^{-1}(a, \tau_i) = \{\varepsilon_i : \mathrm{C}_i(a_{-i}, \tau_i, \varepsilon_i) = a_i\}$$

is an interval in $[1, 2, ..., u_i]$. Finally, we let

$$\mathrm{C}^{-1}(a, \tau) = \times_{i \in \mathrm{N}} \mathrm{C}_i^{-1}(a, \tau_i).$$

Given our solution concept, this is the set of all unobservables that can generate the action profile a as an equilibrium outcome for a given τ, under the set of choice functions C. We next describe the set of all C that are consistent with our model and the data, i.e., H{C}.

Recall we have limited data on realized choices for a set of covariates. We assume the data correspond to independent realizations of the random vector of unobservables, ε. We indicate by $\Pr^{\text{data}}(a \mid \tau)$ the probability of observing action profile a conditional on covariate values τ in the empirical evidence. In addition, we let $\mathcal{Z} \subseteq \mathrm{Z}$ be the support of τ in the data. We finally let $\mathcal{D}_i$ and $\mathcal{A}_i$ be the corresponding supports of a_{-i}, τ_i and a_i in the data.

Let C^{data}, with $\mathrm{C}_i^{\text{data}}$: $\mathcal{D}_i \times [1, 2, ..., u_i] \rightarrow \mathcal{A}_i$, be any set of functions that satisfy (A1, A2, and A4). Let us assume that there is an action profile a' that, under C^{data}, is a unique equilibrium for some $\tau \in \mathcal{Z}$ whenever it is part of the equilibrium set. Then, it must be that

$$\upsilon\left(\mathrm{C}^{\text{data}-1}\left(a', \tau\right)\right) = \Pr^{\text{data}}\left(a' \mid \tau\right) \text{ a.s. for that } \tau \in \mathcal{Z}.$$

Under C^{data}, $\mathrm{C}^{\text{data}-1}(a', \tau)$ is the set of unobservables that can generate a' as an equilibrium outcome. The above condition simply states that the probability of those unobservables has to be equal to the probability of observing action profile a' in the data. Suppose next that a' is just *an* equilibrium for some $\tau \in \mathcal{Z}$. In this case, we must have that

$$\upsilon\left(\mathrm{C}^{\text{data}-1}\left(a', \tau\right)\right) \geq \Pr^{\text{data}}\left(a' \mid \tau\right) \text{ a.s. for that } \tau \in \mathcal{Z}.$$

Under multiple equilibria, it is possible that a' is not selected even if it is part of the equilibrium set. Thus, the last condition simply states that the probability of all the unobservables that can generate the equilibrium profile a' has to be at least as large as the probability of observing that action profile in the data. Let us define

$$\mathcal{A}(\tau) = \mathcal{A}(\tau)^{\mathrm{U}} \cup \mathcal{A}(\tau)^{\mathrm{M}} \text{ for each } \tau \in \mathcal{Z}$$

where $\mathcal{A}^{\mathrm{U}}(\tau)$ is the set of all action profiles that are unique equilibria under C^{data} for a given τ, and $\mathcal{A}^{\mathrm{M}}(\tau)$ is the set of all action profiles for which the equilibrium set is not a singleton. Let us finally consider the next set of restrictions. For each $\tau \in \mathcal{Z}$,

$$\begin{array}{rclll} \upsilon\left(\mathrm{C}^{\text{data}-1}(a, \tau)\right) & = & \Pr^{\text{data}}(a \mid \tau) & \text{a.s. for all} & a \in \mathcal{A}^{\mathrm{U}}(\tau) \\ \upsilon\left(\cup_{a \in \mathrm{B}} \mathrm{C}^{\text{data}-1}(a, \tau)\right) & \geq & \sum_{a \in \mathrm{B}} \Pr^{\text{data}}(a \mid \tau) & \text{a.s. for all} & \mathrm{B} \subseteq \mathcal{A}^{\mathrm{M}}(\tau) \end{array}. \tag{2}$$

The first equation is as before. The inequality requires the probability of all unobservables that can generate certain action profiles as an equilibrium in areas of multiplicity to be larger than the probability of observing those action profiles in the data. It follows immediately from Galichon and Henry (2011) that there exists a consistent C only if there is a C^{data} that satisfies the above conditions. (Though the partial identification results of Galichon and Henry (2011) are parametric, their method can be directly applied to nonparametric models with discrete observables. See, e.g., Cheser and Rosen (2012).) We next describe H{C}.

The characterization of H{C} combines our last results with the ones in Proposition 4. To make the idea clear we let $\mathrm{M}(\mathrm{C}^{\mathrm{data}})$ be the set of all monotone extensions of $\mathrm{C}^{\mathrm{data}}$. Specifically, $\mathrm{C} \in \mathrm{H}(\mathrm{C}^{\mathrm{data}})$ if C satisfies (A1, A2, and A4) and, for each $i \in \mathrm{N}$, $\mathrm{C}_i\left(a_{-i}, \tau_i, \varepsilon_i\right) = \mathrm{C}_i^{\mathrm{data}}\left(a_{-i}, \tau_i, \varepsilon_i\right)$ for all $a_{-i}, \tau_i, \varepsilon_i \in \mathcal{D}_i \times [1, 2, ..., u_i]$. The main result is as follows.

Proposition 7 *H{C} is non-empty if and only if there exists a C^{data} that satisfies (A1, A2, and A4) and (2) for all $\tau \in \mathcal{Z}$. For each C^{data}, $M(C^{data})$ is a complete lattice. Moreover,*

$$H\{C\} = \left\{ C \in M\left(C^{data}\right) : C^{data} \text{ satisfies (A1, A2, and A4) and (2) for all } \tau \in \mathcal{Z} \right\}.$$

Proposition 7 is a simple extension of Proposition 4. Since $\mathrm{M}(\mathrm{C}^{\mathrm{data}})$ is a complete lattice, it has extremal elements. These extremal elements can be computed analogously to B and D before—we just need to treat the unobservables as additional covariates. Thus, for each compatible $\mathrm{C}^{\mathrm{data}}$, we can use Proposition 5 to obtain sharp bounds for out-of-sample predictions of equilibrium points.

4.2 Non-Monotonicity of Choices Regarding Covariates

The monotonicity of individual choice functions regarding the choices of others (A1) is of particular importance in our study as it guarantees the model has at least one solution for each set of covariates and unobserved heterogeneity. The methodology we use crucially depends on this assumption. On the other hand, the monotonicity of choices regarding covariates (A2) simply helps us to improve our bounds. Though it contributes to make more informative predictions, all our ideas carry over to the cases in which some (or all) of the covariates have non-monotone effect on choices. The previous results should be simply understood conditional on the set of covariates that do not satisfy condition A2.

5 Extensions of the Model and Remarks

5.1 Choice Correspondences

Along the study we described the behavior of each agent by a choice function. We next show that we can still follow the sandwich approach to the out-of-sample predictions if we allow for choice correspondences, under slightly stronger restrictions.

Let us suppose the behavior of agent $i \in \mathrm{N}$ is described by the correspondence

$$\mathrm{C}_i\left(a_{-i}, \tau_i, \varepsilon_i\right) : \mathrm{A}_{-i} \times \mathrm{Z}_i \times \mathcal{E}_i \rightrightarrows \mathrm{A}_i \quad \text{with} \quad \mathrm{C}_i\left(a_{-i}, \tau_i, \varepsilon_i\right) \subseteq \mathrm{A}_i.$$

If we allow for choice correspondences, then we need to adapt our solution concept.

Definition (Solution Concept): We say $\mu(\tau, \varepsilon) = (\mu_i(\tau, \varepsilon) : i \in \mathrm{N})$ is consistent with equilibrium behavior generated by $\Gamma_{\tau,\varepsilon}$ if, for all $i \in \mathrm{N}$,

$$\mu_i(\tau, \varepsilon) \in \mathrm{C}_i\left(a_{-i}, \tau_i, \varepsilon_i\right) \quad \text{with} \quad a_{-i} = \left(\mu_j(\tau, \varepsilon) : j \in \mathrm{N}, j \neq i\right).$$

When we allow for correspondences, then observed choices are not enough to describe C_i even for those arguments that satisfy $\tau_i, a_{-i} = \tau_i^m, a_{-i}^m$ for some $m \leq M$. We next strength our monotone restrictions to address this difficulty.

To compare correspondences we will use the induced set order we specify next. We say A is higher than B if $a \geq b$ for any $a \in A$ and $b \in B$. That is, if each element in A is larger than every element in B. The next two restrictions play the role of conditions A1 and A2.

(A1′) For each $i \in \mathrm{N}$, $\mathrm{C}_i\left(a_{-i}, \tau_i, \varepsilon_i\right)$ increases in a_{-i} in the induced set order for all τ_i, ε_i.

(A2′) For each $i \in \mathrm{N}$, $\mathrm{C}_i\left(a_{-i}, \tau_i, \varepsilon_i\right)$ increases in τ_i in the induced set order for all a_{-i}, ε_i.

These two conditions guarantee that any selection from C_i is increasing in a_{-i}, τ_i for each ε_i. When the choice correspondences derive from an underlying game, then A1′ holds if U_i is supermodular in a_i and satisfies the next condition. For all $a_i > a_i'$ and $a_{-i} > a_{-i}'$,

$$\mathrm{U}_i\left(a_i, a_{-i}, \tau_i, \varepsilon_i\right) - \mathrm{U}_i\left(a_i', a_{-i}, \tau_i, \varepsilon_i\right) > \mathrm{U}_i\left(a_i, a_{-i}', \tau_i, \varepsilon_i\right) - \mathrm{U}_i\left(a_i', a_{-i}', \tau_i, \varepsilon_i\right).$$

That is, if it has strict increasing differences in $(a_i; a_{-i})$. Similarly, A2′ is satisfied if, for all $a_i > a_i'$ and $\tau_i > \tau_i'$,

$$\mathrm{U}_i\left(a_i, a_{-i}, \tau_i, \varepsilon_i\right) - \mathrm{U}_i\left(a_i', a_{-i}, \tau_i, \varepsilon_i\right) > \mathrm{U}_i\left(a_i, a_{-i}, \tau_i', \varepsilon_i\right) - \mathrm{U}_i\left(a_i', a_{-i}, \tau_i', \varepsilon_i\right).$$

Thus, conditions A1′ and A2′ arise in supermodular games with strict increasing differences.

Given A1′ and A2′, any $a_i \in \mathrm{C}_i\left(a_{-i}, \tau_i, \varepsilon_i\right)$ must satisfy

$$a_i^m \geq a_i \geq a_i^n \text{ whenever } \tau_i^m, a_{-i}^m > \tau_i, a_{-i} > \tau_i^n, a_{-i}^n \text{ for some } m, n \in M.$$

That is, we can still use the data to bound the choice functions. The difference with respect to our previous analysis is that in order to provide bounds for the choice correspondence evaluated at $a_{-i}, \tau_i, \varepsilon_i$ we need information regarding two pairs of covariates and choices of other agents that are strictly smaller and larger than the argument we are interested in.

It can be shown that Propositions 4 and 5 remain valid if we slightly modify the sharp bounds for C_i, B_i, and D_i, as follows.

$$\mathrm{B}_i'\left(a_{-i}, \tau_i, \varepsilon_i\right) = \inf\left\{a_i \in \mathrm{A}_i : a_i \geq \sup\left\{a_i^m : \tau_i^m, a_{-i}^m < \tau_i, a_{-i}, m \leq M\right\}\right\}$$

$$\mathrm{D}_i'\left(a_{-i}, \tau_i, \varepsilon_i\right) = \sup\left\{a_i \in \mathrm{A}_i : a_i \leq \inf\left\{a_i^m : \tau_i^m, a_{-i}^m > \tau_i, a_{-i}, m \leq M\right\}\right\}$$

That is, if we change two of the weak inequalities by strict ones.

5.2 Identification of Choice Functions vs Equilibrium Points

Manski (2010) clearly explains that identification of the choice functions by shape restrictions is quite different from direct identification of equilibrium points. By using examples of models where the solution set is a singleton, he shows that the assumptions we need to impose on the primitives of the models to identify one or the other object can be quite different. We next explain why the issue is more delicate when the solution set has multiple elements.

Our approach to identify solution sets relies on previous identification of the choice functions. Thus, in terms of Manski's discussion, the question is whether we can provide bounds for $\Delta_{\tau,\varepsilon}$ by imposing monotone conditions on C without previous identification of the latter. The answer to this question would be positive if, for instance, we were able to provide conditions on C so that

$$\Delta_{\tau,\varepsilon} \text{ increases in } \tau \text{ in the induced set order.} \tag{3}$$

(We defined this order in the previous subsection.) If we were able to do so, then, for any $\mu \in \Delta_{\tau,\varepsilon}$,

$$\begin{gathered}\sup\left\{a \in \mathrm{A} : a \leq \inf\left\{a^m : \tau^m \geq \tau, m \leq M\right\}\right\} \\ \geq \mu \geq \\ \inf\left\{a \in \mathrm{A} : a \geq \sup\left\{a^m : \tau^m \leq \tau, m \leq M\right\}\right\}.\end{gathered}$$

The problem with this approach is that, so far, there is no fixed point theorem that, allowing for multiple solutions, provides conditions on the primitives that guarantee (3). This negative statement highlights a key difference between the MCS method for equilibrium points as compared to maximizers of objective functions. (Recall that in the last sub-section we provided precise conditions on the objective functions to be able to compare the set of maximizers by using the induced set order.) This issue can be addressed by imposing specific equilibrium selection rules in the data-generating process (see, e.g., Lazzati (2013)).

5.3 Equilibrium Selection and Out-of-Sample Predictions

Along the study, we obtained out-of-sample predictions without assuming any equilibrium selection rule. It is nevertheless interesting to emphasize that the equilibrium selection rule in the data-generating process will affect the informativeness of our predictions. We illustrate this idea via a simple example.

Example 3: Let $\mathrm{N} = \{1, 2\}$, $\mathrm{Z}_1 = \mathrm{Z}_2 = \{\underline{\tau}, \overline{\tau}\}$ (with $\overline{\tau} > \underline{\tau}$), and $\mathrm{A}_i = \{\underline{a}_i, \overline{a}_i\}$ (with $\overline{a}_i > \underline{a}_i$) for $i = 1, 2$. (To simplify the exposition, we avoid unobserved heterogeneity.) The next table describes the equilibrium set for the high treatment level.

Equilibrium Set

		$\underline{a}_2$	$\overline{a}_2$
$\tau_1 = \overline{\tau}, \tau_2 = \overline{\tau}$	$\underline{a}_1$	$*$	
	$\overline{a}_1$		$*$

Let us suppose the data was generated by equilibrium behavior and A1 and A2 hold. Moreover, let us assume first that these agents always coordinate in the largest equilibrium (though the researcher does not know it). The problem we face is captured by the next tables.

Data

		$\underline{a}_2$	$\overline{a}_2$
$\tau_1 = \overline{\tau}, \tau_2 = \overline{\tau}$	$\underline{a}_1$		
	$\overline{a}_1$		$\bullet$

Out-of-Sample Predictions

		$\underline{a}_2$	$\overline{a}_2$
$\tau_1 = \underline{\tau}, \tau_2 = \underline{\tau}$	$\underline{a}_1$	✓	✓
	$\overline{a}_1$	✓	✓

That is, A1 and A2 are not informative in this case.

Let us next suppose that these agents always coordinate in the smallest equilibrium (though the researcher does not know it). Now the situation is

Data

		$\underline{a}_2$	$\overline{a}_2$
$\tau_1 = \overline{\tau}, \tau_2 = \overline{\tau}$	$\underline{a}_1$	•	
	$\overline{a}_1$		

Out-of-Sample Predictions

		$\underline{a}_2$	$\overline{a}_2$
$\tau_1 = \underline{\tau}, \tau_2 = \underline{\tau}$	$\underline{a}_1$	✓	x
	$\overline{a}_1$	x	✓

Under this alternative equilibrium selection rule our prediction is more informative. ■

We hypothesize that, with more data points, stochastic equilibrium selection rules would lead, on average, to smaller bounds as compared to extremal selections. The reason is that stochastic equilibrium selection rules would generate more variation in the selected action profiles and this would allow us to track individual choice functions more precisely.

6 Conclusion

This paper shows the power of monotone restrictions for obtaining out-of-sample predictions of equilibrium points in simultaneous choice models, or games. The primitives of our model are a system of interdependent choice functions. When they arise from an underlying game, they constitute the best-reply functions of the various players. We show that, while the standard implications of the MCS method for games may be neither testable nor useful for counterfactual predictions, its underlying assumptions are yet quite informative. Specifically, they allow us to partially identify the set of choice functions from observed behavior. We then propose an alternative use of the MCS method that translates the last result into bounds for counterfactual prediction of equilibrium points. We refer to this method as the sandwich approach to the out-of-sample predictions. We finally show that, if the data satisfy certain monotone restrictions, then the empirical evidence can always be rationalized as the Nash equilibria of a supermodular game. This last result offers a theoretical justification of our modeling restrictions and explains the empirical content of this class of games.

7 Proofs

Proof of Proposition 1: Let us consider a mapping $\mathrm{M}_{\tau,\varepsilon} : \times_{i\in\mathrm{N}}\mathrm{A}_i \to \times_{i\in\mathrm{N}}\mathrm{A}_i$, defined by

$$\mathrm{M}_{\tau,\varepsilon}(a_1, a_2, ..., a_n) = (\mathrm{C}_1(a_{-1}, \tau_1, \varepsilon_1), \mathrm{C}_2(a_{-2}, \tau_2, \varepsilon_2), ..., \mathrm{C}_n(a_{-n}, \tau_n, \varepsilon_n)).$$

Function $\mathrm{M}_{\tau,\varepsilon}$ maps n-dimensional actions into itself. Given our solution concept, the set of fixed point of $\mathrm{M}_{\tau,\varepsilon}$ coincides with $\Delta_{\tau,\varepsilon}$. Thus, the proof of Proposition 1 reduces to show that the set of fixed points of $\mathrm{M}_{\tau,\varepsilon}$ has extremal elements. It is well-known that if a product lattice is a product of all complete lattices, then it is itself complete. Thus, $\times_{i\in\mathrm{N}}\mathrm{A}_i$ is a complete lattice. Under A1, $\mathrm{M}_{\tau,\varepsilon}$ is increasing. The result follows by Tarskis$'$ fixed point theorem (see Tarski (1955)). ■

Proof of Proposition 2: Under A2, we have that, for all $\tau \geq \tau'$,

$$\mathrm{M}_{\tau,\varepsilon}(a_1, a_2, ..., a_n) \geq \mathrm{M}_{\tau',\varepsilon}(a_1, a_2, ..., a_n)$$

where $\mathrm{M}_{\tau,\varepsilon}$ and $\mathrm{M}_{\tau',\varepsilon}$ are defined as in the proof of Proposition 1. By the proof of Proposition 1, $\mathrm{M}_{\tau,\varepsilon}$ and $\mathrm{M}_{\tau',\varepsilon}$ map complete lattices into themselves. Under A1, $\mathrm{M}_{\tau,\varepsilon}$ and $\mathrm{M}_{\tau',\varepsilon}$ are increasing. Thus, the result follows by Topkis (1998, Corollary 2.5.2). ■

Proof of Proposition 3: Let us consider two mapping $\mathrm{L}_{\tau,\varepsilon}$ and $\mathrm{N}_{\tau,\varepsilon}$ from $\times_{i\in\mathrm{N}}\mathrm{A}_i$ into $\times_{i\in\mathrm{N}}\mathrm{A}_i$, defined as

$$\mathrm{L}_{\tau,\varepsilon}(a_1, a_2, ..., a_n) = (\mathrm{B}_i(\tau_i, a_{-i}, \varepsilon_i) : i \in \mathrm{N}) \text{ and } \mathrm{N}_{\tau,\varepsilon}(a_1, a_2, ..., a_n) = (\mathrm{D}_i(\tau_i, a_{-i}, \varepsilon_i) : i \in \mathrm{N}).$$

If $\mathrm{D} \geq \mathrm{C} \geq \mathrm{B}$, then

$$\mathrm{N}_{\tau,\varepsilon}(a_1, a_2, ..., a_n) \geq \mathrm{M}_{\tau,\varepsilon}(a_1, a_2, ..., a_n) \geq \mathrm{L}_{\tau,\varepsilon}(a_1, a_2, ..., a_n)$$

where $\mathrm{M}_{\tau,\varepsilon}$ is defined as in the proof of Proposition 1. By our argument in the proof of Proposition 2, they map complete lattices into themselves. Under A1, the three mappings are increasing. Thus, the result follows by Topkis (1998, Corollary 2.5.2). ■

Proof of Proposition 4: Let us first assume that $\mathrm{H}\{\mathrm{C}\}$ is non-empty. We next show that $\mathrm{H}\{\mathrm{C}\}$ is a complete lattice.

Let $\mathrm{h} \subseteq \mathrm{H}\{\mathrm{C}\}$ be given by $\mathrm{h} = \left(\mathrm{C}^k : k \in \mathrm{K}\right)$. By definition of consistency, for each $i \in \mathrm{N}$ and each $k \in \mathrm{K}$, we have that

$$\mathrm{C}_i^k\left(\tau_i, a_{-i}, \varepsilon_i\right) = \mathrm{C}_i^{\mathrm{data}}\left(\tau_i, a_{-i}, \varepsilon_i\right) \text{ for all } \tau_i, a_{-i} \in \mathcal{D}.$$

Thus, for all $\tau_i, a_{-i} \in \mathcal{D}$,

$$\mathrm{C}_i^{\mathrm{data}}\left(\tau_i, a_{-i}, \varepsilon_i\right) = \inf\left(\mathrm{C}_i^k\left(\tau_i, a_{-i}, \varepsilon_i\right) : k \in \mathrm{K}\right) = \sup\left(\mathrm{C}_i^k\left(\tau_i, a_{-i}, \varepsilon_i\right) : k \in \mathrm{K}\right).$$

We next show that $\inf\left(\mathrm{C}^k : k \in \mathrm{K}\right)$ and $\sup\left(\mathrm{C}^k : k \in \mathrm{K}\right)$ satisfy A1 and A2. By definition of consistency, for each $i \in \mathrm{N}$ and each $k \in \mathrm{K}$, we have that

$$\mathrm{C}_i^k\left(\tau_i, a_{-i}, \varepsilon_i\right) \geq \mathrm{C}_i^k\left(\tau_i', a_{-i}', \varepsilon_i\right) \text{ for all } \tau_i, a_{-i} \geq \tau_i', a_{-i}'.$$

Then, for all $\tau_i, a_{-i} \geq \tau_i', a_{-i}'$,

$$\begin{aligned} \inf\left(\mathrm{C}_i^k\left(\tau_i, a_{-i}, \varepsilon_i\right) : k \in \mathrm{K}\right) &\geq \inf\left(\mathrm{C}_i^k\left(\tau_i', a_{-i}', \varepsilon_i\right) : k \in \mathrm{K}\right) \\ \sup\left(\mathrm{C}_i^k\left(\tau_i, a_{-i}, \varepsilon_i\right) : k \in \mathrm{K}\right) &\geq \sup\left(\mathrm{C}_i^k\left(\tau_i', a_{-i}', \varepsilon_i\right) : k \in \mathrm{K}\right). \end{aligned}$$

In addition, since A_i is a complete lattice and $\mathrm{C}_i^k\left(\tau_i, a_{-i}, \varepsilon_i\right) \in \mathrm{A}_i$ for each $i \in \mathrm{N}$ and each $k \in \mathrm{K}$, we have that

$$\inf\left(\mathrm{C}_i^k\left(\tau_i, a_{-i}, \varepsilon_i\right) : k \in \mathrm{K}\right) \in \mathrm{A}_i \text{ and } \sup\left(\mathrm{C}_i^k\left(\tau_i, a_{-i}, \varepsilon_i\right) : k \in \mathrm{K}\right) \in \mathrm{A}_i.$$

It follows from all previous results that $\inf(\mathrm{h}) \in \mathrm{H}\{\mathrm{C}\}$ and $\sup(\mathrm{h}) \in \mathrm{H}\{\mathrm{C}\}$. We next show that $\mathrm{H}\{\mathrm{C}\}$ is non-empty if and only if $\mathrm{C}^{\mathrm{data}}$ satisfies A1 and A2.

The only if part follows directly by the definition of consistency. To show the if part we will prove that if $\mathrm{C}^{\mathrm{data}}$ satisfies A1 and A2, then B,D $\in \mathrm{H}\{\mathrm{C}\}$. Recall that

$$\begin{aligned} \mathrm{B}_i\left(a_{-i}, \tau_i, \varepsilon_i\right) &= \inf\left\{a_i \in \mathrm{A}_i : a_i \geq \sup\left\{a_i^m : \tau_i^m, a_{-i}^m \leq \tau_i, a_{-i}, m \leq M\right\}\right\} \\ \mathrm{D}_i\left(a_{-i}, \tau_i, \varepsilon_i\right) &= \sup\left\{a_i \in \mathrm{A}_i : a_i \leq \inf\left\{a_i^m : \tau_i^m, a_{-i}^m \geq \tau_i, a_{-i}, m \leq M\right\}\right\} \end{aligned}.$$

Note that if $\mathrm{C}_i^{\mathrm{data}}$ satisfies A1 and A2, then

$$\mathrm{C}_i^{\mathrm{data}}\left(\tau_i^m, a_{-i}^m, \varepsilon_i\right) = \mathrm{B}_i\left(\tau_i^m, a_{-i}^m, \varepsilon_i\right) = \mathrm{D}_i\left(\tau_i^m, a_{-i}^m, \varepsilon_i\right) = a^m \text{ for all } m \leq M.$$

We next show that $\mathrm{B}_i\left(a_{-i}, \tau_i, \varepsilon_i\right) \in \mathrm{A}_i$ and $\mathrm{D}_i\left(a_{-i}, \tau_i, \varepsilon_i\right) \in \mathrm{A}_i$.

Each of the sets

$$\left\{a_i \in \mathrm{A}_i : a_i \geq \sup\left\{a_i^m : \tau_i^m, a_{-i}^m \leq \tau_i, a_{-i}, m \leq M\right\}\right\}$$
$$\left\{a_i \in \mathrm{A}_i : a_i \leq \inf\left\{a_i^m : \tau_i^m, a_{-i}^m \geq \tau_i, a_{-i}, m \leq M\right\}\right\}$$

is the intersection of a complete lattice A_i and a complete sublattice of A_i. Thus, they are complete sublattices of A_i and have a minimal and a maximal element that belong to A_i. Thus, B_i and D_i select an element of A_i for all $a_{-i}, \tau_i, \varepsilon_i$ (i.e., they are non-empty). We next show that B_i and D_i satisfy A1 and A2.

Recall that

$$\mathrm{B}_i(\tau_i, a_{-i}, \varepsilon_i) = \inf\left\{a_i \in A_i : a_i \geq \sup\left\{a_i^m : \tau_i^m, a_{-i}^m \leq \tau_i, a_{-i}, m \leq M\right\}\right\}.$$

To show that A1 and A2 hold, let $\tau_i, a_{-i} \geq \tau_i', a_{-i}'$. It is readily verified that

$$\left\{a_i^m : \tau_i^m, a_{-i}^m \leq \tau_i, a_{-i},\ m \leq M\right\} \supseteq \left\{a_i^m : \tau_i^m, a_{-i}^m \leq \tau_i', a_{-i}',\ m \leq M\right\}.$$

Thus,

$$\sup\left\{a_i^m : \tau_i^m, a_{-i}^m \leq \tau_i, a_{-i},\ m \leq M\right\} \geq \sup\left\{a_i^m : \tau_i^m, a_{-i}^m \leq \tau_i', a_{-i}',\ m \leq M\right\}.$$

It follows that $\mathrm{B}_i(\tau_i, a_{-i}, \varepsilon_i) \geq \mathrm{B}_i\left(\tau_i', a_{-i}', \varepsilon_i\right)$ whenever $\tau_i, a_{-i} \geq \tau_i', a_{-i}'$ and A1 and A2 hold.

Recall that

$$\mathrm{D}_i(\tau_i, a_{-i}, \varepsilon_i) = \sup\left\{a_i \in A_i : a_i \leq \inf\left\{a_i^m : \tau_i^m, a_{-i}^m \geq \tau_i, a_{-i}, m \leq M\right\}\right\}.$$

To show that A1 and A2 hold, let $\tau_i, a_{-i} \geq \tau_i', a_{-i}'$. It follows immediately that

$$\left\{a_i^m : \tau_i^m, a_{-i}^m \geq \tau_i, a_{-i},\ m \leq M\right\} \subseteq \left\{a_i^m : \tau_i^m, a_{-i}^m \geq \tau_i', a_{-i}',\ m \leq M\right\}.$$

Thus,

$$\inf\left\{a_i^m : \tau_i^m, a_{-i}^m \leq \tau_i, a_{-i},\ m \leq M\right\} \geq \inf\left\{a_i^m : \tau_i^m, a_{-i}^m \leq \tau_i', a_{-i}',\ m \leq M\right\}.$$

It follows that $\mathrm{D}_i(\tau_i, a_{-i}, \varepsilon_i) \geq \mathrm{D}_i\left(\tau_i', a_{-i}', \varepsilon_i\right)$ whenever $\tau_i, a_{-i} \geq \tau_i', a_{-i}'$ and A1 and A2 are satisfied. This completes our claim. The fact that B and D are the minimal and maximal elements of H{C} follows by construction. ■

Proof of Corollary 5: By Proposition 4, B and D are the minimal and maximal microeconomic specifications for our model that are consistent with the data. Thus, the result follows by Proposition 3. ■

Proof of Proposition 6: Let C_i be the choice function of agent i for any C that satisfies A1 and A2. We need to show that can always construct a payoff function $\mathrm{U}_i(a_i, a_{-i}, \tau_i, \varepsilon_i)$ that is supermodular in a_i and has increasing differences in $(a_i; a_{-i}, \tau_i)$ such that

$$\mathrm{C}_i(a_{-i}, \tau_i, \varepsilon_i) = \arg\max\{\mathrm{U}_i(a_i, a_{-i}, \tau_i, \varepsilon_i) : a_i \in \mathrm{A}_i\}.$$

Let $\mathrm{F}_i(a_i, b_i) : \mathrm{A}_i \times \mathrm{A}_i \to \mathrm{A}_i$ be any function that satisfies the next three properties:

(i) $\mathrm{F}_i(a_i, a_i) > \mathrm{F}_i(a_i, b_i)$ for all $a_i \neq b_i$;

(ii) $\mathrm{F}_i(a_i, b_i)$ is supermodular in a_i for all b_i; and

(iii) $\mathrm{F}_i(a_i, b_i)$ has increasing differences in $(a_i; b_i)$.

Define $\mathrm{U}_i(a_i, a_{-i}, \tau_i) = \mathrm{F}_i(a_i, \mathrm{C}_i(a_{-i}, \tau_i, \varepsilon_i))$. We next show that, constructed in this way, U_i rationalizes C_i, is supermodular in a_i, and has increasing differences in $(a_i; a_{-i}, \tau_i)$.

By (i) we have that

$$\mathrm{C}_i(a_{-i}, \tau_i, \varepsilon_i) = \arg\max_{a_i}\{\mathrm{U}_i(a_i, a_{-i}, \tau_i) : a_i \in \mathrm{A}_i\}.$$

By (ii), U_i is supermodular in a_i. Let $a_i \geq a_i'$ and $a_{-i}, \tau_i \geq a_{-i}', \tau_i'$, then

$$\mathrm{F}_i(a_i, \mathrm{C}_i(a_{-i}, \tau_i, \varepsilon_i)) - \mathrm{F}_i\left(a_i', \mathrm{C}_i(a_{-i}, \tau_i, \varepsilon_i)\right) \geq \mathrm{F}_i\left(a_i, \mathrm{C}_i\left(a_{-i}', \tau_i', \varepsilon_i\right)\right) - \mathrm{F}_i\left(a_i', \mathrm{C}_i\left(a_{-i}', \tau_i', \varepsilon_i\right)\right)$$

by (iii) and the fact that $\mathrm{C}_i(a_{-i}, \tau_i, \varepsilon_i) \geq \mathrm{C}_i\left(a_{-i}', \tau_i', \varepsilon_i\right)$ by A1 and A2. Thus, U_i has increasing differences in $(a_i; a_{-i}, \tau_i)$. We next provide a function F that satisfies (i)-(iii).

The complete lattice A_i belongs to $\mathbb{R}^d$. Thus, $a_i = (a_{ij} : j \leq d)$. It can be easily checked that, for instance, $\mathrm{F}_i(a_i, b_i) = -\sum_{j=1}^{d}|a_{ij} - b_{ij}|$ satisfies conditions (i), (ii), and (iii). ■

Proof of Proposition 7: The fact that $\mathrm{M}(\mathrm{C}^{\mathrm{data}})$ is a complete lattice follows immediately from the proof of Proposition 4—we just need to treat each ε_i as an additional covariate.

Let C be any set of choice functions that satisfy (A1, A2, and A4). It follows from Galichon and Henry (2011) that C is consistent if and only if, for each $\tau \in \mathcal{Z}$,

$$\upsilon\left(\cup_{a \in \mathrm{B}}\mathrm{C}^{-1}(a, \tau)\right) \geq \sum_{a \in \mathrm{B}} \mathrm{Pr}^{\mathrm{data}}(a \mid \tau) \text{ a.s. for all } \mathrm{B} \subseteq \mathrm{A}. \quad (4)$$

Note that these set of inequalities only impose *actual* restrictions on the support of the observations. Thus, for each of these C, (4) is satisfied if and only if the restriction of C to the support of the data satisfies the set of inequalities. Call this restriction $\mathrm{C}^{\mathrm{data}}$. For each $\mathrm{C}^{\mathrm{data}}$, any monotone extension $\mathrm{C} \in \mathrm{M}(\mathrm{C}^{\mathrm{data}})$ can generate, for each ε and each $\tau \in \mathcal{Z}$, the same equilibria than $\mathrm{C}^{\mathrm{data}}$ and possibly more. By simply assuming that these extra equilibria are never selected, then the two functions can generate exactly the same empirical observations. Thus, each of these C satisfies (A1, A2, and A4) and is compatible with the data. This completes the proof. ■

References

[1] Amir, R. (2008). Comparative Statics of Nash Equilibria, mimeo.

[2] Aradillas-López, A. (2011). Nonparametric Probability Bounds for Nash Equilibrium Actions in a Simultaneous Discrete Game, *Quantitative Economics*, 2, 135-171.

[3] Becker, G. (1962). Irrational Behavior and Economic Theory, *Journal of Political Economy*, 70, 1-13.

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

[5] Blundell, R., M. Browing, and I. Crawford (2003). Nonparametric Engel Curves and Revealed Preferences, *Econometrica*, 71, 205-240.

[6] Blundell, R., D. Kristensen, R. Matzkin (2012). Bounding Quantile Demand Functions using Revealed Preference Inequalities, mimeo.

[7] Browning, M., and P. Chiappori (1998). Efficient Intra-Household Allocation: A General Characterization and Empirical Tests, *Econometrica*, 66 , 1241-78.

[8] Candogan, O., K. Bimpikis, and A. Ozdaglar (2012). Optimal Pricing in Networks with Externalities, *Operations Research*, 4, 883-905.

[9] Carvajal, A. (2002). Testable Restrictions of Nash Equilibrium in Games with Continuous Domains, mimeo.

[10] Carvajal, A., R. Deb, J. Fenske, and J. Quah (2013). Revealed Preference Tests of the Cournot Model, *Econometrica*, forthcoming.

[11] Chambers, C., F. Echenique, and E. Shmaye (2010). The Axiomatic Structure of Empirical Content, mimeo.

[12] Chesher A., and A. Rosen (2012). Simultaneous Equations Models for Discrete Outcomes: Coherence, Completeness, and Identification, cemmap working paper CWP21/12.

[13] Cherchye, L., T. Demuynck, and B. De Rock (2013). The Empirical Content of Cournot Competition, *Journal of Economic Theory*, 148, 1552-1581.

[14] Chesher, A., and A. Rosen (2012). Simultaneous Equations Models for Discrete Outcomes, Coherence, Completeness, and Identification, mimeo.

[15] De Clippel, G. (2014). Behavioral Implementation, *Americal Economic Review*, Forthcoming.

[16] De Clippel, G. and K. Rozen (2013). Bounded Rationality and Limited Datasets, mimeo.

[17] De Paula, Á. (2009). Inference in a Synchronization Game with Social Interactions, *Journal of Econometrics*, 148, 56-71.

[18] De Paula, Á., and X. Tang (2012). Inference of Signs of Interaction Effects in Simultaneous Games with Incomplete Information, *Econometrica*, 80, 143–172.

[19] Deb, R. (2009). A Testable Model of Consumption With Externalities, *Journal of Economic Theory*, 114, 1804-1816.

[20] Echenique, F., and A. Edlin (2004). Mixed Equilibria are Unstable in Games of Strategic Complements, *Journal of Economic Theory*, 118, 61-79.

[21] Echenique, F., and I. Komunjer (2009). Testing Models with Multiple Equilibria by Quantile Methods, *Econometrica*, 77, 1281–1297.

[22] Echenique, F., S. Lee, M. Shum, and M. Yenmez (2013). The Revealed Prefence Theory of Stable and Extremal Stable Matching, *Econometrica*, 81, 153–171.

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

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

[25] Kline, B., and E. Tamer (2012). Bounds for Best Response Functions in Binary Games, *Journal of Econometrics*, 166, 92–105.

[26] Kreps, D. (2013). Microeconomic Foundations I: Choice and Competitive Markets. Princeton University Press.

[27] Lazzati, N. (2014). Treatment Response with Social Interactions: Identification via Monotone Comparative Statics, *Quantitative Economics*, Accepted.

[28] Manski, C. (1997). Monotone Treatment Response, *Econometrica*, 65, 1311-1334.

[29] Manski, C. (2007). Partial Identification of Counterfactual Choice Probabilities, *International Economic Review*, 48, 1393-1410.

[30] Manski, C. (2011). Identification of Treatment Response with Social Interactions, *The Econometrics Journal*, 16, 1-23.

[31] Matzkin, R. (2008). Identification in Nonparametric Simultaneous Equations Models. Econometrica, 76, 5, 945-978.

[32] Milgrom, P., and C. Shannon (1994). Monotone Comparative Statics, *Econometrica*, 62, 157-180.

[33] Milgrom, P., and J. Roberts (1990). Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities, *Econometrica*, 58, 1255-1277.

[34] Milgrom, P., and J. Roberts (1994). Comparing Equilibria, *American Economic Review*, 84, 441-459.

[35] Molinari, F., and A. Rosen (2008). The Identification Power of Equilibrium in Games: The Supermodular Case, *Journal of Business & Economic Statistics*, 26, 297-302.

[36] Pollak, R. (1977). Price Dependent Preferences, *American Economic Review*, 67, 64-75.

[37] Quah, J. (2007). The Comparative Statics of Constrained Optimization Problems, *Econometrica*, 75, 401-431.

[38] Quah, J., and B. Strulovici (2009). Comparative Statics, Informativeness, and the Interval Dominance Order, *Econometrica*, 77, 1949–1992.

[39] Sprumont, Y. (2000). Testable Implications of Collective Choice Theories, *Journal of Economic Theory*, 93, 205-232.

[40] Tamer, E. (2003). Incomplete Simultaneous Discrete Response Model with Multiple Equilibria, *Review of Economic Studies*, Wiley Blackwell, 70, 147-165.

[41] Tarski, A. (1955). A Lattice Theoretical Fixpoint Theorem and its Applications, *Pacific Journal of Mathematics*, 5, 285-309.

[42] Topkis, D. (1979). Equilibrium Points in Nonzero-sum n-person Submodular Games, *Journal of Control and Optimization*, 17, 773-787.

[43] Topkis, D. (1998). Supermodularity and Complementarity, Frontiers of Economic Research, edited by D. Kreps and T. Sargent.

[44] Uetake, K., and Y. Watanabe (2013). A Note on Identification and Estimation of Supermodular Games, mimeo.

[45] Varian, H. (1982). The Nonparametric Approach to Demand Analysis, *Econometrica*, 50, 945-973.

[46] Vives, X. (1990). Nash Equilibrium with Strategic Complementarities, *Journal of Mathematical Economics*, 19, 305-321.