cemmap
centre for microdata methods and practice

Sharp Identification Regions in Games

Arie Beresteanu
Ilya Molchanov
Francesca Molinari

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP15/08

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

An ESRC Research Centre

Sharp Identification Regions in Games*

Arie Beresteanu[†] Ilya Molchanov[‡] Francesca Molinari[§]

June 15, 2008

Abstract

We study identification in static, simultaneous move finite games of complete information, where the presence of multiple Nash equilibria may lead to partial identification of the model parameters. The identification regions for these parameters proposed in the related literature are known not to be sharp. Using the theory of random sets, we show that the sharp identification region can be obtained as the set of minimizers of the distance from the conditional distribution of game's outcomes given covariates, to the conditional Aumann expectation given covariates of a properly defined random set. This is the random set of probability distributions over action profiles given profit shifters implied by mixed strategy Nash equilibria. The sharp identification region can be approximated arbitrarily accurately through a finite number of moment inequalities based on the support function of the conditional Aumann expectation. When only pure strategy Nash equilibria are played, the sharp identification region is exactly determined by a finite number of moment inequalities. We discuss how our results can be extended to other solution concepts, such as for example correlated equilibrium or rationality and rationalizability.

We show that calculating the sharp identification region using our characterization is computationally feasible. We also provide a simple algorithm which finds the set of inequalities that need to be checked in order to insure sharpness. We use examples analyzed in the literature to illustrate the gains in identification afforded by our method.

Keywords: Identification, Random Sets, Aumann Expectation, Support Function, Capacity Functional, Normal Form Games, Inequality Constraints.

*We are grateful to Talia Bar, David Easley, Paul Ellickson, Yossi Feinberg, Keisuke Hirano, Adam Rosen, Andrew Sweeting, and especially Larry Blume for comments. The results in this paper on sharp identification when the solution concept adopted is pure strategy Nash equilibrium begun circulating in February 2007 under the title "Asymptotic Behavior of Inequality Constrained Models - with Applications to Office Supply Superstores." We benefitted from the opportunity to present our research, under that title, at the 2007 Summer Meetings of the Econometric Society and at the 2008 Winter Meetings of the Econometric Society, and to present our research, under the current title, at the Hebrew University of Jerusalem, Olin Business School, Tel Aviv University, University of Arizona, UCL, the 2007 Triangle Econometrics Conference, and the 2008 Cemmap-Northwestern University Conference on "Inference in Partially Identified Models with Applications." Financial support from the USA NSF Grants SES-0617559 (Beresteanu) and SES-0617482 (Molinari), from the Swiss National Science Foundation Grant Nr. 200021-117606 (Molchanov), and from the Center for Analytic Economics at Cornell University is gratefully acknowledged. Ilya Molchanov gratefully acknowledges the hospitality of the Department of Economics at Cornell University

[†]Department of Economics, Duke University, arie@econ.duke.edu.

[‡]Department of Mathematical Statistics and Actuarial Science, University of Bern, ilya.molchanov@stat.unibe.ch.

[§]Department of Economics, Cornell University, fm72@cornell.edu.

1 Introduction

This paper belongs to the literature on identification in incomplete econometric models. Examples of such models may arise when the data are incomplete (sample realizations are not fully observable) or when the model asserts that the relationship between the outcome variable and the exogenous variables is a correspondence rather than a function. When the econometric model is incomplete, the sampling process and the maintained assumptions may be consistent with a set of parameter vectors or functionals, rather than a single one. In this case, the model is partially identified. The analyses in Manski (1989, 2003), Manski and Tamer (2002), Haile and Tamer (2003), Ciliberto and Tamer (2004) and Andrews, Berry, and Jia (2004) are examples of research studying the identified features of incomplete econometric models.

Our main contribution is to provide a simple and novel procedure to determine the *sharp identification region* of the parameters characterizing static, simultaneous move finite games of complete information in the presence of multiple Nash equilibria. By contrast, the identification region for this class of models provided in the related literature is known not to be sharp. Establishing whether a conjectured region for the identified features of an incomplete model is *sharp* is a key question in identification analysis. For simplicity, we focus on the parametric case. An econometric model then consists of a sampling process and a set of maintained modeling assumptions summarized by an unknown finite dimensional parameter vector θ. This vector, or one of its subvectors, is the focus of empirical research. Given the joint distribution of the observed variables, a researcher asks herself what parameters θ are consistent with this distribution. The region in the parameter space which includes all possible parameter values that could generate the same distribution of observables for some data generation process consistent with the maintained modeling assumptions, and no other parameter value, is called the *sharp identification region*. Examples of sharp identification regions for parameters of incomplete models are given in Manski (2003) and Manski and Tamer (2002), among others. In some cases, researchers are only able to characterize a region in the parameter space that includes all the parameter values that may have generated the observables, but may include other (infeasible) parameter values as well. These larger regions are called *outer regions*. Examples of outer regions for parameters of incomplete models are given in Ciliberto and Tamer (2004) and Andrews, Berry, and Jia (2004). The inclusion in the outer regions of parameter values which are infeasible may weaken the model's ability to make useful predictions, and the researcher's ability to test for model misspecification.

Point identification of the class of models treated in this paper has been previously studied by

Bjorn and Vuong (1985), Bresnahan and Reiss (1988, 1990, 1991), Berry (1992), Mazzeo (2002), Tamer (2003), and Bajari, Hong, and Ryan (2007) among others. These authors achieve point identification of the payoff parameters by adding assumptions to the model.[1] Examples of such restrictions include assumptions on the nature of competition, heterogeneity of firms, availability of covariates with sufficiently large support and/or instrumental variables, and restrictions on the selection mechanism which, in the data generating process, picks an equilibrium in the regions of multiplicity. By contrast, we do not impose any assumption on the selection mechanism, on the nature of competition, or on the form of heterogeneity across players. Our approach does not require availability of covariates with large support or instruments, but fully exploits their identifying power if they are present. Andrews, Berry, and Jia (2004), Ciliberto and Tamer (2004), and Berry and Tamer (2007) study partial identification of the same class of models as we do. Their work is the closest in spirit to ours. However, they only provide outer regions for the model parameters.

While throughout most of the paper we assume that players follow Nash behavior, we show that our methodology easily extends to other solution concepts for the game. We illustrate this by looking at games where rationality of level-1 is the solution concept (a problem first studied by Aradillas-Lopez and Tamer (2008)), and by looking at games where correlated equilibrium is the solution concept.[2]

Our paper is exclusively about identification. However, our characterization of the sharp identification region leads to an obvious sample analog counterpart which can be used when the researcher is confronted with a finite sample of observations. This sample analog is given by the set of minimizers of a criterion function obtained from a finite number of sample moment equalities and inequalities, so that the recent contributions of Chernozhukov, Hong, and Tamer (2007), Andrews and Guggenberger (2007), Andrews and Soares (2007), Galichon and Henry (2006), Romano and Shaikh (2006), Rosen (2006), and Pakes, Porter, Ho, and Ishii (2006), among others, can be applied for estimation and statistical inference.

1.1 Overview

The literature on identification in games with multiple equilibria often describes the sharp identification region of the model parameters using the concept of "selection mechanism." An admissible

[1] Tamer (2003) also suggests an approach to partially identify the model's parameters when no additional assumptions are imposed.

[2] Yang (2008) exploits the fact that all Nash equilibria are correlated equilibria to provide simple-to-compute outer regions for the model parameters when Nash equilibrium is the solution concept.

selection mechanism is the probability distribution of a random variable which chooses the equilibrium played in the regions of the sample space where the model admits multiple equilibria. By definition, the sharp identification region includes all the parameter values for which one can find an admissible selection mechanism, such that the model augmented with this selection mechanism generates the joint distribution of the observed variables. If, as it is the case in this paper, no assumptions are placed on it, the selection mechanism may represent an infinite dimensional nuisance parameter. Hence, "standard" approaches to characterizing the sharp identification region, which are based on the selection mechanism, entail dealing with an infinite dimensional nuisance parameter. This task is sufficiently difficult that Berry and Tamer (2007, page 68) have suggested to give up on obtaining sharp identification regions. Rather, they suggest focusing on outer regions for the model parameters that do not exploit all the information contained in the model, but are practically appealing because they are defined by a finite number of moment inequalities. These moment inequalities have to hold for $\underline{x}-a.s.$, with $\underline{x}$ the observable payoff shifters (in what follows, for simplicity we do not keep repeating that all statements concerning moment inequalities and set membership have to hold for $\underline{x}-a.s.$).

The methodology that we propose allows us to bypass the need to directly deal with infinite dimensional nuisance parameters. Our treatment of the problem distinguishes between two cases. When only pure strategies are considered, the main benefit of our approach is that the sharp identification region is obtained through a finite number of moment inequalities. When mixed strategies are also allowed, one loses the ability to characterize the sharp identification region through a finite number of moment inequalities. Intuitively, this is because there is additional information provided by the fact that players must be indifferent among the actions that they play with positive probability according to a given equilibrium strategy.[3] However, we show that even in this case one can approximate the sharp identification region arbitrarily accurately through a finite number of moment inequalities.

We achieve our results by using the theory of random sets (Molchanov (2005)). The key insight that leads to our characterization of the sharp identification region is the following. Suppose that the researcher observes game's outcomes y and payoff shifters $\underline{x}$ from a cross section of markets in which players play the same game and follow Nash behavior. Profits also depend on payoff shifters ε that are unobservable by the econometrician, but observable by the players (the game is one of complete information). Let the cross section be large enough that the distribution of y conditional on $\underline{x}$, denoted $\mathbf{P}(y|\underline{x})$, can be learned exactly. Parametrize the payoff functions of

[3] Section 4 further discusses this fact.

the game, and fix a given value of the parameter vector θ. Each realization of $\underline{x}$ and ε implies a (necessarily non-empty) set of mixed strategy Nash equilibria, which we denote by $S_\theta(\underline{x}, \varepsilon)$. Each of the equilibria in this set determines a probability distribution over the game's outcomes, conditional on the realization of $\underline{x}$ and ε. Let the *random closed set* of probability distributions over the game's outcomes implied by $S_\theta(\underline{x}, \varepsilon)$ be denoted $Q(S_\theta(\underline{x}, \varepsilon))$. In Section 3 we establish that the entire collection of model's predicted probability distributions of the game's outcomes conditional on $\underline{x}$ is given by the *Aumann expectation* of $Q(S_\theta(\underline{x}, \varepsilon))$ conditional on $\underline{x}$, denoted $\mathbb{E}(Q(S_\theta(\underline{x}, \varepsilon))|\,\underline{x})$.[4]

Framing the set of the model's predicted probability distributions in terms of an Aumann expectation is practically very advantageous. A candidate value for the parameter vector may have generated the observed conditional distribution $\mathbf{P}(y|\,\underline{x})$ if and only if $\mathbf{P}(y|\,\underline{x})$ belongs to the conditional Aumann expectation associated with that parameter vector. Hence, the sharp identification region of the model parameters is given by the collection of θ's that determine a conditional Aumann expectation $\mathbb{E}(Q(S_\theta(\underline{x}, \varepsilon))|\,\underline{x})$ that contains $\mathbf{P}(y|\,\underline{x})$ for $\underline{x} - a.s.$

Given a candidate value for θ, one can verify whether it belongs to the sharp identification region in the following way. The set $\mathbb{E}(Q(S_\theta(\underline{x}, \varepsilon))|\,\underline{x})$ can be evaluated exactly or approximated by simulation, depending on the complexity of the game. The candidate parameter vector θ is in the sharp identification region if and only if $\mathbf{P}(y|\,\underline{x})$ is an element of this set. This corresponds to checking whether the *support function* of $\mathbf{P}(y|\,\underline{x})$ is dominated by the support function of $\mathbb{E}(Q(S_\theta(\underline{x}, \varepsilon))|\,\underline{x})$.[5] Showing that one support function dominates another amounts to checking an infinite number of inequalities, each associated with a point on a unit sphere of an appropriate dimension. A finite set of moment inequalities can be obtained by discretizing this unit sphere. The properties of this approximation are discussed in Section 5.2.

A substantial simplification is possible in the special case where one assumes that players do not randomize across their actions, and pure strategy Nash equilibria exist (see Assumption 4). We show that the number of inequalities to be checked in order to assure that the support function of $\mathbf{P}(y|\,\underline{x})$ is dominated by the support function of $\mathbb{E}(Q(S_\theta(\underline{x}, \varepsilon))|\,\underline{x})$ is finite, without the need to discretize the unit sphere. This is because when players are only allowed to play pure strategies, $\mathbb{E}(Q(S_\theta(\underline{x}, \varepsilon))|\,\underline{x})$ is a closed convex polytope, fully characterized by a finite number of supporting hyperplanes, i.e., by its support function evaluated at a finite number of directions in the unit sphere. These directions are trivial to determine. We connect this result to a related notion in

[4] We formally define the notions of both random closed set and Aumann expectation in Section 3.

[5] See Schneider (1993, Section 1.7) for a thorough discussion of the support function of a closed convex set, and its properties.

the theory of random sets, that of a *capacity functional*, which is the probability distribution of a random closed set. We show that our characterization of the sharp identification region based on the support function of $\mathbb{E}\left(Q\left(S_\theta\left(\underline{x},\varepsilon\right)\right)\middle|\,\underline{x}\right)$ is dual to a characterization based on the capacity functional of the random set of pure strategy equilibrium outcomes, by exploiting a result due to Artstein (1983). While the number of inequalities to be checked in order to obtain the sharp identification region is finite, in some applications it may be quite large. However, we show that in some cases this number can be substantially reduced by exploiting basic notions of set algebra.

There are no precedents to our characterization of the sharp identification region of the payoff function parameters characterizing static, simultaneous move finite games of complete information in the presence of multiple Nash equilibria. However, there are two precedents, with respect to this paper, to the use of the theory of random sets within the econometrics literature. Both of them are mainly focused on statistical inference. One is given by the work of Beresteanu and Molinari (2006, 2008). They study a class of partially identified models in which the sharp identification region of the parameter vector of interest can be written as a transformation of the Aumann expectation of a properly defined random set.[6] For this class of models, they propose to use a sample analog estimator given by a transformation of a Minkowski average of properly defined random sets. They use limit theorems for independent and identically distributed sequences of random sets, to establish consistency of this estimator with respect to the Hausdorff metric. They propose two Wald-type test statistics, based on the Hausdorff metric and on the lower Hausdorff hemimetric, to test hypothesis and make confidence statements about the entire sharp identification region and its subsets.

The other is given by the work of Galichon and Henry (2006). The goal of their paper is to provide a specification test for partially identified structural models. They introduce various formulations of the notion of a correctly specified structural model which is partially identified, and establish their equivalence. One of the notions that they use is based on the (Choquet) capacity functional of a random set. In particular, they show that the model is correctly specified if the distribution of the observed outcome is dominated by the capacity functional of the random correspondence between the latent variables and the outcome variables characterizing the model. This allows them to extend the Kolmogorov-Smirnov test of correct model specification to partially identified models. They then define the notion of "core determining" classes of sets, to find a manageable class of sets for which to check that the dominance condition is satisfied.

[6] In order to establish sharpness of the identification region of the parameters of a best linear predictor with interval outcome data, Beresteanu and Molinari (2006, 2008) use the same result involving the capacity functional of a random set due to Artstein (1983) that we use in this paper.

1.2 Structure of the Paper

In Section 2 we introduce notation and assumptions, and present the identification problem for the class of models treated in this paper. In order to clearly connect our work to the related literature, we discuss the definition of the sharp identification region provided by Berry and Tamer (2007). In Section 3 we give a computationally feasible characterization of the sharp identification region when players may randomize across their actions. We illustrate the gains in identification afforded by our methodology through the simple example of a two player entry game. In Section 4 we show how our approach further simplifies when only pure strategy Nash equilibria are played. We illustrate the gains in identification afforded by our methodology through the simple example of a four player, two type entry game. In Section 5 we address the computational issues associated with our characterization of the sharp identification region. We first discuss how $\mathbb{E}\left(Q\left(S_\theta\left(\underline{x},\varepsilon\right)\right)|\,\underline{x}\right)$ and its support function can be approximated by simulation. We then show how one can use the result of this simulation to compute the sharp identification region in the case that only pure strategies are considered. We also provide a very simple algorithm that may allow one to significantly reduce the number of inequalities to check, in order to obtain the sharp identification region. Finally, we discuss how to approximate the sharp identification region when mixed strategies are played. In Section 6 we show how our methodology can be extended to the case that players are only assumed to be level-1 rational, or are assumed to play correlated strategies. Section 7 concludes. Appendix A collects all the proofs. Appendix B gives additional details about one of our examples, a two player entry game with mixed strategies.

2 Notation, Assumptions, and Identification Problem

We focus on simultaneous-move games of complete information (normal form games) in which each player has a finite set of pure strategies. A key example is the static game of entry, in which each player can choose between two actions: "enter" or "not enter." Partial identification of this model is studied by Andrews, Berry, and Jia (2004, ABJ henceforth), Ciliberto and Tamer (2004, CT henceforth) and Berry and Tamer (2007).

2.1 Notation and Assumptions

Throughout the paper, we use capital Latin letters to denote sets and random sets. We use lower case Latin letters for random vectors. We denote parameter vectors and sets of parameter vectors, respectively by θ and Θ. For a given finite set W, we denote by κ_W its cardinality. Given two nonempty sets $B, C \subset \Re^d$, we denote by

$$\begin{aligned} d_H(C,B) &= \sup_{c \in C} \inf_{b \in B} \|c - b\|, \\ \rho_H(C,B) &= \max\{d_H(C,B), d_H(B,C)\}, \end{aligned}$$

respectively, the directed Hausdorff distance from C to B and the Hausdorff distance between C and B. The Hausdorff norm of B is denoted $\|B\|_H = \rho_H(B, \{0\}) = \sup\{\|b\| : b \in B\}$.

We let J denote the number of players. Each player has a finite set of actions (pure strategies) $\mathcal{A}_j$, $j = 1, \ldots, J$. We denote by $a = (a_1, \ldots, a_J) \in \mathcal{A}$ a generic vector specifying an action for each player (a pure strategy profile), with $\mathcal{A} = \times_{j=1}^{J} \mathcal{A}_j$. We denote by $\pi_j(a_j, a_{-j}, x_j, \varepsilon_j, \theta)$ the payoff function for player j, where a_{-j} is the vector of player j's opponents' actions, $x_j \in \mathcal{X}$ is a vector of observable payoff shifters, ε_j is a payoff shifter observed by the players but unobserved by the econometrician, and $\theta \in \Theta \subset \Re^p$ is a vector of parameters to be estimated, with Θ the parameter space. We denote by $\sigma_j : \mathcal{A}_j \to [0,1]$ the mixed strategy for player j that assigns to each action $a_j \in \mathcal{A}_j$ a probability $\sigma_j(a_j) \geq 0$ that it is played, with $\sum_{a_j \in \mathcal{A}_j} \sigma_j(a_j) = 1$ for each $j = 1, \ldots, J$. We let $\Delta(\mathcal{A}_j)$ denote the mixed extension of $\mathcal{A}_j$, and $\Delta(\mathcal{A}) = \times_{j=1}^{J} \Delta(\mathcal{A}_j)$. With the usual slight abuse of notation, we denote by $\pi_j(\sigma_j, \sigma_{-j}, x_j, \varepsilon_j, \theta)$ the expected payoff associated with the mixed strategy profile $\sigma = (\sigma_1, \ldots, \sigma_J)$. We denote by $y \in \mathcal{Y}$ the vector of potentially observable outcomes of the game. In the remainder of this section, we formalize our assumptions on the games and sampling processes.

Assumption 1 *(i) The set $\mathcal{A}$ of pure strategy profiles and the set $\mathcal{Y}$ of potentially observable outcomes are finite. Each player has a finite number $\kappa_{\mathcal{A}_j} \geq 2$ of pure strategies to choose from. The number of players is $J \geq 2$.*
(ii) Players follow Nash behavior. They move simultaneously and only once.
(iii) The strategy profiles determine the outcomes observable by the econometrician through a continuous mapping $g : \mathcal{A} \to \mathcal{Y}$, the "outcome rule". This outcome rule is known by the econometrician.
(iv) The parametric form of the payoff functions $\pi_j(a_j, a_{-j}, x_j, \varepsilon_j, \theta)$, $j = 1, \ldots, J$, is known, and for a known action $\bar{a}$ it is normalized to $\pi_j(\bar{a}_j, \bar{a}_{-j}, x_j, \varepsilon_j, \theta) = 0$ for each j. The payoff functions are continuous in the observable and unobservable payoff shifters. The parameter space Θ is compact.

Assumption 1-(i) assures that there is a finite set of strategies for each player, and a finite set of possible outcomes observable by the econometrician. It restricts attention to normal form games. Part (ii) of the assumption requires that players follow Nash behavior, so that for given payoff shifters x_j and ε_j, the mixed strategy profile σ constitutes a Nash equilibrium if each player's mixed strategy is a best response. Part (iii) of the assumption requires continuity of the outcome rule. Part (iv) of the assumption requires continuity of the payoff functions in x_j and ε_j. These conditions are needed to establish measurability and closedness of certain sets. Assumption 1-(iv) also provides a location normalization. Such normalization is implicit in entry models, where players are commonly assumed to earn zero payoffs if they do not enter the market (regardless of the action chosen by their opponents).

In many normal form games, such as the static simultaneous move entry games analyzed by ABJ, CT, and Berry and Tamer (2007), players' actions and the outcomes observable by the econometrician coincide. We simplify the exposition in all that follows, by restricting attention to games satisfying this condition:

Assumption 2 *The outcome rule $g(\cdot)$ is the identity mapping, so that $y = a$.*

Our results, however, apply to the more general case stated in Assumption 1-(iii), as we illustrate in Section 4.2 with a simple example.

Assumption 3 *The econometrician observes a random sample $(y_i, \underline{x}_i)_{i=1}^n$ of equilibrium outcomes and observable payoff shifters from a large cross section of n markets drawn from a population distribution that satisfies Assumption 1. The outcome vector for market i is $y_i = (y_{i1}, \ldots, y_{iJ})$. The observed matrix of payoff shifters for market i, $\underline{x}_i$, is comprised of the non-redundant elements of x_{ji}, $j = 1, \ldots, J$ (the observable payoff shifters of firm j in market i). The unobserved random vector $\varepsilon_i = (\varepsilon_{1i}, \ldots, \varepsilon_{Ji})$ is independently and identically distributed across i with a distribution function F that is known up to a finite dimensional parameter that is part of θ. The random vectors $(y, \underline{x}, \varepsilon)$ are defined on a non-atomic product probability space $(\Omega, \mathfrak{F}, \mathbf{P}) = (\Omega_1, \mathfrak{F}_1, \mathbf{P}_1) \times (\Omega_2, \mathfrak{F}_2, \mathbf{P}_2)$.*

Assumption 3 allows us to identify $\mathbf{P}(y|\underline{x})$, the population distribution of observed equilibrium outcomes given covariates. Since our focus in this paper is identification, we treat identified distributions as population distributions.

The requirement that the probability space is non-atomic is fairly weak and facilitates some of the technical details below. We use a product probability space to clearly differentiate between the randomness in the payoff functions due to the payoff shifters, represented by Ω_1, and the randomness in the actions taken by the players for a given mixed strategy profile, represented by Ω_2. Hence, for any $\omega_1 \in \Omega_1$, $\omega_2 \in \Omega_2$, $(\underline{x}(\omega_1, \omega_2), \varepsilon(\omega_1, \omega_2)) = (\underline{x}(\omega_1), \varepsilon(\omega_1))$. On the other hand, $y(\omega_1, \omega_2)$ depends both on ω_1, which determines the equilibrium mixed strategy profiles $\sigma(\omega_1)$, and ω_2, which determines the specific action taken by the players when they randomize according to the mixed strategy profile $\sigma(\omega_1)$.

2.2 The Identification Problem

It is well known that the games and sampling processes satisfying Assumptions 1-3 may lead to multiple Nash equilibria. Multiplicity implies that there are regions of values of the exogenous variables where the econometric model predicts more than one outcome. Therefore, the relationship between the outcome variable of interest and the exogenous variables is a correspondence rather than a function. Hence, the parameters of the payoff functions may not be point identified, see for example Berry and Tamer (2007) for a thorough discussion of this problem.

Nevertheless, these parameters can be partially identified given knowledge of $\mathbf{P}(y|\underline{x})$ for all $\underline{x}$. In particular, their identification region is given by the set of parameter vectors which are consistent with the sampling process and the maintained modeling assumptions, and therefore may have generated the distribution of observables. If the conjectured region for the parameters of interest contains all its observationally equivalent feasible values and no other, the region is sharp. Berry and Tamer (2007, equation (2.21), page 67) provide an abstract formulation for the sharp identification region in a two player entry model. Here we report their formulation, modified to allow for games with more than two players and two actions. This formulation, however, requires the introduction of an infinite dimensional nuisance parameter, and thus is considered to be impractical. In the next section we show that an alternative and practical formulation, which avoids the need for an infinite dimensional nuisance parameter, delivers the same sharp identification region.

We start with some additional notation. Let $S_\theta(\underline{x}, \varepsilon)$ denote the set of mixed strategy Nash equilibria associated with a specific realization of the payoff shifters $\underline{x}$ and ε (this set is defined formally in equation (3.2) below). Let $\psi(\sigma; S_\theta(\underline{x}, \varepsilon))$ denote a selection mechanism giving the probability that an equilibrium $\sigma \in S_\theta(\underline{x}, \varepsilon)$ is selected. Observe that for this selection mechanism to be admissible it is required that $\psi(\sigma; S_\theta(\underline{x}, \varepsilon)) \geq 0$ for all $\sigma \in S_\theta(\underline{x}, \varepsilon)$, and that $\sum_{\sigma \in S_\theta(\underline{x}, \varepsilon)} \psi(\sigma; S_\theta(\underline{x}, \varepsilon)) = 1$. This summation is well defined because except on a set of $\underline{x}, \varepsilon$ realizations of measure zero, the set $S_\theta(\underline{x}, \varepsilon)$ contains a finite number of equilibria (Wilson (1971)). Notice that the equilibrium selection mechanism ψ is left unspecified and can depend on market unobservables. Then we have the following definition.

Definition 1 *In a game which satisfies Assumptions 1-3, the sharp identification region for the* *parameter vector $\theta \in \Theta$ is given by:*

$$\text{(2.1)}\quad \Theta_I^\star = \left\{ \theta \in \Theta : \begin{array}{l} \exists\ \psi \text{ such that } \forall\ t \in \mathcal{Y}, \\ \mathbf{P}\left(y = t \middle| \underline{x}\right) = \int \left(\sum_{\sigma \in S_\theta(\underline{x},\varepsilon)} \psi\left(\sigma; S_\theta\left(\underline{x}, \varepsilon\right)\right) \prod_{j=1}^{J} \sigma_j\left(t_j\right) \right) dF\left(\varepsilon \middle| \underline{x}\right)\ \ \underline{x} - a.s. \end{array} \right\}$$

where ψ is an admissible equilibrium selection mechanism as described above.

Let $\mathbf{P}\left(y \middle| \underline{x}; \theta, \psi\right)$ denote the integral on the right hand side of the second line of equation (2.1) above. Berry and Tamer explain this formulation and the practical difficulties involved in computing the set $\Theta_I^\star$ as follows (page 68):

> "The set $\Theta_I^\star$ is the sharp identified set, i.e., the set of parameters θ that are consistent with the data and the model. Heuristically, a $\theta \in \Theta_I^\star$ if and only if there exists a (proper) selection mechanism ψ (...) such that the induced probability distribution $\mathbf{P}\left\{y \middle| \underline{x}; \theta, \psi\right\}$ matches the choice probabilities $\mathbf{P}\left(y \middle| \underline{x}\right)$ for all $\underline{x}$ almost everywhere. So, the presence of multiple equilibria introduces nuisance parameters that are not specified and hence makes it harder to identify the parameter θ. (...) Inference on the set $\Theta_I^\star$ based on definition (2.1) [(2.21) in the original] though theoretically attractive is not practically feasible since one needs to deal with infinite dimensional nuisance parameters (the ψ's). A practical approach to inference in this class of models follows the approach in Ciliberto and Tamer (2004) by exploiting the fact that the selection mechanism ψ is a probability and hence bounded between zero and one. Although this approach does not provide a sharp set, it is practically attractive."

In the following sections we provide a tractable characterization of the sharp identification region. In the special case that players play only pure strategy Nash equilibria (Section 4 below), we show that the sharp identification region is given by a finite number of moment inequalities which have to hold for $\underline{x} - a.s.$

3 The Sharp Identification Region

3.1 The Random Set of Mixed Strategy Equilibrium Profiles

We assume that players in each market follow Nash behavior. For a given realization of $\underline{x}$ and ε, the mixed strategy profile $\sigma = (\sigma_1, \ldots, \sigma_J)$ constitutes a Nash equilibrium if

$$\text{(3.1)}\qquad \pi_j\left(\sigma_j, \sigma_{-j}, x_j, \varepsilon_j, \theta\right) \geq \pi_j\left(\tilde{\sigma}_j, \sigma_{-j}, x_j, \varepsilon_j, \theta\right)\ \forall \tilde{\sigma}_j \in \Delta\left(\mathcal{A}_j\right)\ \forall j.$$

Hence, we define the following θ-dependent set:

$$S_\theta(\underline{x}, \varepsilon) = \{\sigma \in \Delta(\mathcal{A}) : \pi_j(\sigma_j, \sigma_{-j}, x_j, \varepsilon_j, \theta) \geq \pi_j(\tilde{\sigma}_j, \sigma_{-j}, x_j, \varepsilon_j, \theta)\ \forall \tilde{\sigma}_j \in \Delta(\mathcal{A}_j)\ \forall j\}. \tag{3.2}$$

For a given value of θ and realization of $\underline{x}$ and ε, this is the set of mixed strategy Nash equilibrium profiles. For ease of notation we write $S_\theta \equiv S_\theta(\underline{x}, \varepsilon)$ and omit the explicit reference to $\underline{x}$ and ε. Given Assumption 1, S_θ is a random closed set in $\Delta(\mathcal{A})$.

Definition 2 *Denoting by $\mathcal{F}$ the family of closed subsets of a topological space $\mathbb{F}$, a map $Z : \Omega \to \mathcal{F}$ is called a* ***random closed set****, also known as a closed set valued random variable, if for every compact set K in $\mathbb{F}$, $Z^{-1}(K) = \{\omega \in \Omega : Z(\omega) \cap K \neq \emptyset\} \in \mathfrak{F}$.*

The fact that the set S_θ satisfies the conditions in Definition 2 can be shown by writing the set S_θ as follows:

$$S_\theta = \bigcap_{j=1}^{J} \{\sigma \in \Delta(\mathcal{A}) : \pi_j(\sigma_j, \sigma_{-j}, x_j, \varepsilon_j, \theta) \geq \tilde{\pi}_j(\sigma_{-j}, x_j, \varepsilon_j, \theta)\},$$

where

$$\tilde{\pi}_j(\sigma_{-j}, x_j, \varepsilon_j, \theta) = \sup_{\tilde{\sigma}_j \in \Delta(\mathcal{A}_j)} \pi_j(\tilde{\sigma}_j, \sigma_{-j}, x_j, \varepsilon_j, \theta).$$

Since $\pi_j(\sigma_j, \sigma_{-j}, x_j, \varepsilon_j, \theta)$ is a continuous function of $\sigma, x_j, \varepsilon_j$, its supremum $\tilde{\pi}_j(\sigma_{-j}, x_j, \varepsilon_j, \theta)$ is a continuous function. Therefore S_θ is the finite intersection of sets defined as solutions of inequalities for continuous (random) functions. Thus, S_θ is a random closed set, see Molchanov (2005, Section 1.1).

For a given parameter $\theta \in \Theta$, each element $\sigma(\omega_1) \in S_\theta(\omega_1)\ \mathbf{P} - a.s.$ determines a distribution of players' actions in $\mathcal{A}_1 \times \cdots \times \mathcal{A}_J$. Such random elements σ are the selections of S_θ:

Definition 3 *Let Z be a random closed set in a topological space $\mathbb{F}$. A random element z with values in $\mathbb{F}$ is called a (measurable) selection of Z if $z(\omega) \in Z(\omega)$ for almost all $\omega \in \Omega$. The family of all selections of Z is denoted by* $\mathrm{Sel}(Z)$.

Each realization of the selection $\sigma(\omega_1) \in S_\theta(\omega_1)$ takes values in $\Delta(\mathcal{A}_1) \times \cdots \times \Delta(\mathcal{A}_J)$ and is one of the admissible mixed strategy Nash equilibrium profiles associated with the realizations $\underline{x}(\omega_1)$ and $\varepsilon(\omega_1)$ determined by $\omega_1 \in \Omega_1$. As briefly discussed after Assumption 3, the vector of players actions $y(\omega_1, \omega_2)$ depends not only on ω_1 through the selection $\sigma(\omega_1) \in S_\theta(\omega_1)$, but also on $\omega_2 \in \Omega_2$, which determines players' choice of which action to take in accordance with the mixed strategy profile $\sigma(\omega_1)$. Let $\sigma(\omega_1) = (\sigma_1(\omega_1), \ldots, \sigma_J(\omega_1))$. By definition of a mixed strategy profile, for each $j = 1, \ldots, J$, $\sigma_j(\omega_1) : \mathcal{A}_j \to [0,1]$ assigns to each action $a_j \in \mathcal{A}_j$ a probability $\sigma_j(\omega_1, a_j) \geq 0$ that it is played, with $\sum_{a_j \in \mathcal{A}_j} \sigma_j(\omega_1, a_j) = 1$.

Recall that by Assumption 2, the realizations of y coincide with the actions a taken with positive probability and $\mathcal{Y} = \mathcal{A}$. Index the set $\mathcal{Y}$ in some (arbitrary) way, such that $\mathcal{Y} = \{t^1, \ldots, t^{\kappa_\mathcal{Y}}\}$. Then by the law of iterated expectations, for a given parameter value $\theta \in \Theta$ and selection $\sigma \in \mathrm{Sel}(S_\theta)$, the implied probability that $y = t^k$, $k = 1, \ldots, \kappa_\mathcal{Y}$, is given by[7]

$$\mathbf{E}\left(\mathbf{P}\left(y(\omega_1, \omega_2) = t^k \middle| \omega_1; \theta\right)\right) = \mathbf{E} \prod_{j=1}^{J} \sigma_j\left(\omega_1, t_j^k\right).$$

Hence, we can use $\sigma = (\sigma_1, \ldots, \sigma_J)$ to define a random point $q(\sigma)$ whose realizations have coordinates

$$[q(\sigma(\omega_1))]_k = \prod_{j=1}^{J} \sigma_j\left(\omega_1, t_j^k\right), \quad k = 1, \ldots, \kappa_\mathcal{Y}, \tag{3.3}$$

This random point lies in a space of dimension equals to $\kappa_\mathcal{Y}$ and is such that $[q(\sigma(\omega_1))]_k \geq 0$ for each $k = 1, \ldots, \kappa_\mathcal{Y}$ and $\sum_{k=1}^{\kappa_\mathcal{Y}} [q(\sigma(\omega_1))]_k = 1$. Hence, it is an element of the $\kappa_\mathcal{Y} - 1$ dimensional simplex, denoted $\Delta^{\kappa_\mathcal{Y}-1}$. The resulting set

$$Q(S_\theta) = \{([q(\sigma)]_k, k = 1, \ldots, \kappa_\mathcal{Y}) : \sigma \in \mathrm{Sel}(S_\theta)\}, \tag{3.4}$$

is a closed random set in $\Delta^{\kappa_\mathcal{Y}-1}$. Each vector $([q(\sigma(\omega_1))]_k, k = 1, \ldots, \kappa_\mathcal{Y}) \in Q(S_\theta(\omega_1))$ gives the probability with which each outcome (a J-tuple of actions under Assumption 2) of the game is observed under the mixed strategy equilibrium $\sigma(\omega_1)$ when the realization of $\underline{x}$ is $\underline{x}(\omega_1)$ and the realization of ε is $\varepsilon(\omega_1)$. Section 3.3 below illustrates these ideas through the simple example of a two player complete information static game of entry.

3.2 The Sharpness Result

Every realization of $q \in \mathrm{Sel}(Q(S_\theta))$ is contained in $\Delta^{\kappa_\mathcal{Y}-1}$, and therefore $Q(S_\theta)$ is an integrably bounded random closed set, see Molchanov (2005, Definition 2.1.11), and all its selections are integrable. Hence we can define the set

$$\begin{aligned} \mathbb{E}(Q(S_\theta)|\underline{x}) &= \{\mathbf{E}(q|\underline{x}) : q \in \mathrm{Sel}(Q(S_\theta))\} \\ &= \{(\mathbf{E}([q(\sigma)]_k|\underline{x}), k = 1, \ldots, \kappa_\mathcal{Y}) : \sigma \in \mathrm{Sel}(S_\theta)\}. \end{aligned}$$

[7] Under the more general Assumption 1-(iii), this probability is equal to

$$\mathbf{E}\left(\mathbf{P}\left(y(\omega_1, \omega_2) = t^k \middle| \omega_1; \theta\right)\right) = \mathbf{E} \prod_{j=1}^{J} \prod_{a_j \in \mathcal{A}_j} (\sigma_j(\omega_1, a_j))^{1\left(t_j^k = g(a_j)\right)}.$$

The set $\mathbb{E}\left(Q\left(S_\theta\right)\middle|\, \underline{x}\right)$ is the conditional Aumann expectation[8] of $Q\left(S_\theta\right)$. The following example illustrates the notion of Aumann expectation. To keep it as simple as possible, we omit covariates and do not base it on a game theoretic model.

Example 1 *Let Z be a random closed set in $\Re^4$ defined as follows:*

$$Z(\omega) = \begin{cases} \left\{\left[\frac{1}{3}\ \frac{2}{3}\ 0\ 0\right], \left[\frac{2}{3}\ \frac{1}{3}\ 0\ 0\right]\right\} & \text{for } \omega \in \Omega_1', \\ \left\{\left[\frac{2}{3}\ \frac{1}{3}\ 0\ 0\right], \left[0\ 0\ \frac{1}{3}\ \frac{2}{3}\right]\right\} & \text{for } \omega \in \Omega_1 \backslash \Omega_1', \end{cases}$$

where $\mathbf{P}\left(\Omega_1'\right) = \frac{1}{2}$. Consider the set Ω_1'. Then all selections of $Z(\omega)$ for $\omega \in \Omega_1'$ can be obtained as

$$z(\omega) = \begin{cases} \left[\frac{1}{3}\ \frac{2}{3}\ 0\ 0\right] & \text{for } \omega \in \Omega_1'', \\ \left[\frac{2}{3}\ \frac{1}{3}\ 0\ 0\right] & \text{for } \omega \in \Omega_1' \backslash \Omega_1'', \end{cases}$$

for all measurable $\Omega_1'' \subset \Omega_1'$. Since $\mathbf{E}\left(z \middle|\, \omega \in \Omega_1'\right) = \left[\frac{1}{3}\ \frac{2}{3}\ 0\ 0\right]\mathbf{P}\left(\Omega_1''\middle|\,\Omega_1'\right) + \left[\frac{2}{3}\ \frac{1}{3}\ 0\ 0\right]\left(1 - \mathbf{P}\left(\Omega_1''\middle|\,\Omega_1'\right)\right)$, the range of expectations of selections depends on the atomic structure of the underlying probability space. If the probability space has no atoms, then the possible values for $\mathbf{P}\left(\Omega_1''\middle|\,\Omega_1'\right)$ fill in the whole segment $[0,1]$. Hence, $\mathbb{E}\left(Z \middle|\, \omega \in \Omega_1'\right)$ is the interval with extreme points given by $\left[\frac{1}{3}\ \frac{2}{3}\ 0\ 0\right]$, $\left[\frac{2}{3}\ \frac{1}{3}\ 0\ 0\right]$. A similar reasoning allows one to conclude that

$$\mathbb{E}(Z) = \mathrm{co}\left[\left\{\left[\frac{2}{3}\ \frac{1}{3}\ 0\ 0\right]\right\}, \left\{\left[\frac{1}{2}\ \frac{1}{2}\ 0\ 0\right]\right\}, \left\{\left[\frac{1}{6}\ \frac{1}{3}\ \frac{1}{6}\ \frac{1}{3}\right]\right\}, \left\{\left[\frac{1}{3}\ \frac{1}{6}\ \frac{1}{6}\ \frac{1}{3}\right]\right\}\right],$$

where $\mathrm{co}[\cdot]$ denotes the convex hull of the set in square brackets. Hence, in this simple example $\mathbb{E}(Z)$ is a parallelogram. □

By Theorem 2.1.46 in Molchanov (2005) the conditional Aumann expectation exists and is unique. Because by Assumption 3 the probability space is non-atomic,[9] and because the random set $Q\left(S_\theta\right)$ takes its realizations in a subset of the finite dimensional space $\Re^{\kappa_\mathcal{Y}}$, it follows by Theorem 2.1.15 and Theorem 2.1.24 of Molchanov (2005) that $\mathbb{E}\left(Q\left(S_\theta\right)\middle|\,\underline{x}\right)$ is a closed convex set for $\underline{x} - a.s.$, and $\mathbb{E}\left(Q\left(S_\theta\right)\middle|\,\underline{x}\right) = \mathbb{E}\left(\mathrm{co}\left[Q\left(S_\theta\right)\right]\middle|\,\underline{x}\right)$.

The set $\mathbb{E}\left(Q\left(S_\theta\right)\middle|\,\underline{x}\right)$ collects vectors of probabilities with which each outcome of the game can be observed, by averaging over $\omega_1 \in \Omega_1$ the corresponding probability with which each outcome of the game is observed under the mixed strategy equilibrium $\sigma\left(\omega_1\right)$. We emphasize that in case of multiplicity, a different mixed strategy equilibrium $\sigma\left(\omega_1\right) \in S_\theta\left(\omega_1\right)$ may be selected for each ω_1. If the model is correctly specified, there exists at least one value of $\theta \in \Theta$ such that the observed

[8] Aumann (1965) introduces the notion of integrals for set valued functions that we use here.

[9] When the probability space contains atoms, all the discussion that follows applies, with $\mathrm{co}\left[Q\left(S_\theta\right)\right]$ replacing $Q\left(S_\theta\right)$.

conditional distribution of y given $\underline{x}$, $\mathbf{P}(y|\underline{x})$, is a point in the set $\mathbb{E}(Q(S_\theta)|\underline{x})$ for $\underline{x}-a.s.$[10] Let the support function of a nonempty closed convex set $B \in \Re^{\kappa_\mathcal{Y}}$ be denoted $h(B,\cdot)$, with

$$h(B,u) = \sup_{b\in B} u'b, \ \ u \in \Re^{\kappa_\mathcal{Y}},$$

Then $\mathbf{P}(y|\underline{x}) \in \mathbb{E}(Q(S_\theta)|\underline{x})$ if and only if

$$u'\mathbf{P}(y|\underline{x}) \leq h(\mathbb{E}(Q(S_\theta)|\underline{x}),u) = \mathbf{E}[h(Q(S_\theta),u)|\underline{x}] \ \ \forall\, u \in \Re^{\kappa_\mathcal{Y}}, \tag{3.5}$$

where $\mathbf{P}(y|\underline{x}) \equiv \left[\mathbf{P}\left(y=t^k\middle|\underline{x}\right), k=1,\ldots,\kappa_\mathcal{Y}\right]$, and where the last equality follows by Theorem 2.1.47-(iv) in Molchanov (2005). Because the support function is positively homogeneous, i.e., $h(Q(S_\theta),cu) = ch(Q(S_\theta),u)$ for all $c>0$ and for all $u \in \Re^{\kappa_\mathcal{Y}}$, condition (3.5) is equivalent to

$$u'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}[h(Q(S_\theta),u)|\underline{x}] \ \ \forall\, u \in \mathfrak{S}, \tag{3.6}$$

where $\mathfrak{S} = \{u \in \Re^{\kappa_\mathcal{Y}} : \|u\| = 1\}$ denotes the unit sphere in $\Re^{\kappa_\mathcal{Y}}$.

The key result of this paper is the following:

Theorem 3.1 *Let Assumptions 1-3 be satisfied, and no other information be available. Then*

$$\Theta_I = \left\{\theta \in \Theta : u'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}[h(Q(S_\theta),u)|\underline{x}] \ \ \forall\, u \in \mathfrak{S}\ \underline{x}-a.s.\right\} \tag{3.7}$$

is the sharp identification region for the parameter vector $\theta \in \Theta$.

By standard arguments, condition (3.6) is equivalent to

$$d_H(\mathbf{P}(y|\underline{x}), \mathbb{E}(Q(S_\theta)|\underline{x})) = 0.$$

Hence, Θ_I can be defined in three equivalent ways:

$$\begin{aligned}
\Theta_I &= \{\theta \in \Theta : d_H(\mathbf{P}(y|\underline{x}), \mathbb{E}(Q(S_\theta)|\underline{x})) = 0\ \underline{x}-a.s.\} \\
&= \left\{\theta \in \Theta : u'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}[h(Q(S_\theta),u)|\underline{x}] \ \ \forall\, u \in \mathfrak{S}\ \underline{x}-a.s.\right\} \\
&= \left\{\theta \in \Theta : u'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}[h(Q(S_\theta),u)|\underline{x}] \ \ \forall\, u \in \mathfrak{S}^\Delta\ \underline{x}-a.s.\right\},
\end{aligned} \tag{3.8}$$

where

$$\mathfrak{S}^\Delta = \left\{u \in \Re^{\kappa_\mathcal{Y}} : u_1 + \ldots + u_{\kappa_\mathcal{Y}} = 1 \text{ and } \|u\| = 1\right\}.$$

The last equality follows because $\mathbb{E}(Q(S_\theta)|\underline{x})$ is a subset of the $\kappa_\mathcal{Y}-1$ dimensional simplex $\Delta^{\kappa_\mathcal{Y}-1}$. Therefore, it suffices to restrict the directions used in condition (3.6) to those parallel to the $(\kappa_\mathcal{Y}-1)$-dimensional hyperplane that defines the unit simplex.

[10] By the definition of $\mathbb{E}(Q(S_\theta)|\underline{x})$, $\mathbf{P}(y|\underline{x}) \in \mathbb{E}(Q(S_\theta)|\underline{x})$ if and only if $\exists\, q \in \mathrm{Sel}(Q(S_\theta)) : \mathbf{E}(q|\underline{x}) = \mathbf{P}(y|\underline{x})$.

The three definitions of Θ_I given above are equivalent. The definition based on the support function can be straightforwardly applied, by discretizing the unit sphere in the relevant space and checking a finite number of moment inequalities which have to hold for $\underline{x}-a.s.$ We discuss in detail how to do this in Section 5.2 below. In practice, the definition based on $u \in \mathfrak{S}^{\Delta}$ should be preferred to the one based on $\mathfrak{S}$, because $\mathfrak{S}^{\Delta}$ is of a lower dimension than $\mathfrak{S}$.

The definition based on the distance from $\mathbf{P}(y|\underline{x})$ to $\mathbb{E}(Q(S_\theta)|\underline{x})$ may be used depending on whether it is computationally more convenient in the specific application, as it yields a straightforward criterion function[11] which is minimized by every parameter in the identification region:

$$\mathbf{W}(\theta) = \int d_H\left(\mathbf{P}(y|\underline{x}), \mathbb{E}(Q(S_\theta)|\underline{x})\right) dF_{\underline{x}}, \tag{3.9}$$

where $F_{\underline{x}}$ denotes the joint distribution of $\underline{x}$. Clearly, $\mathbf{W}(\theta) \geq 0$ for all $\theta \in \Theta$, and $\mathbf{W}(\theta) = 0$ if and only if $\theta \in \Theta_I$.

3.3 Example: Two Players Entry Game

In this case $\mathcal{A}_j = \{0,1\}$ for $j = 1,2$, $\mathcal{A} = \mathcal{Y} = \{0,1\} \times \{0,1\}$, and $\kappa_{\mathcal{Y}} = 4$. Let $\mathbf{P}(y|\underline{x}) = [\mathbf{P}(y = t|\underline{x}), t \in \{(0,0),(1,0),(0,1),(1,1)\}]$. Let $\sigma_j \in [0,1]$ denote the probability that player j enters the market, with $1-\sigma_j$ the probability that he does not. Omit the regressors $\underline{x}$ in all that follows, and let players' payoffs be $\pi_j = a_j(a_{-j}\theta_j + \varepsilon_j)$, $j = 1,2$. Figure 1 plots the set of mixed strategy equilibrium profiles S_θ resulting from the possible realizations of $\varepsilon_1, \varepsilon_2$.

With only two players and two actions, we have that

$$Q(S_\theta) = \left\{ \begin{bmatrix} (1-\sigma_1)(1-\sigma_2) \\ \sigma_1(1-\sigma_2) \\ (1-\sigma_1)\sigma_2 \\ \sigma_1\sigma_2 \end{bmatrix} : \sigma \in \mathrm{Sel}(S_\theta) \right\}.$$

Appendix B provides additional details giving the values of the coordinates of $\mathbf{E}(q)$, $q \in \mathrm{Sel}(Q(S_\theta))$, and relating our approach to the discussion in Berry and Tamer (2007, pages 65-70). In order not to deal with the infinite dimensional nuisance parameter (the selection mechanism) discussed in section 2.2, Berry and Tamer suggest to estimate an outer region for the parameter vector of interest, based on the insight in CT. Using our notation, their outer region is given by

[11] This criterion function was used, in the case of pure strategies only, by Ciliberto and Tamer (2004). However, in their case the distance is from $\mathbf{P}(y|\underline{x})$ to a superset of $\mathbb{E}(Q(S_\theta)|\underline{x})$.

$$\Theta_I^{CT} = \left\{ \theta \in \Theta : \begin{array}{c} \mathbf{C}_{S_\theta}((0,0)) \leq \mathbf{P}(y=(0,0)) \leq \mathbf{C}_{S_\theta}((0,0)) + \int\limits_{M^\theta} \left(1+\frac{\varepsilon_2}{\theta_2}\right)\left(1+\frac{\varepsilon_1}{\theta_1}\right) dF(\varepsilon) \\ \mathbf{C}_{S_\theta}((1,0)) \leq \mathbf{P}(y=(1,0)) \leq \mathbf{C}_{S_\theta}((1,0)) + \mathbf{P}\left(\varepsilon \in M^\theta\right) \\ \mathbf{C}_{S_\theta}((0,1)) \leq \mathbf{P}(y=(0,1)) \leq \mathbf{C}_{S_\theta}((0,1)) + \mathbf{P}\left(\varepsilon \in M^\theta\right) \\ \mathbf{C}_{S_\theta}((1,1)) \leq \mathbf{P}(y=(1,1)) \leq \mathbf{C}_{S_\theta}((1,1)) + \int\limits_{M^\theta} \frac{\varepsilon_2}{\theta_2}\frac{\varepsilon_1}{\theta_1} dF(\varepsilon) \end{array} \right\},$$

where $M^\theta = \{\varepsilon : \varepsilon \in [0, -\theta_1] \times [0, -\theta_2]\}$ and $\mathbf{C}_{S_\theta}(\sigma) = \mathbf{P}(S_\theta = \{\sigma\})$, see Definition 4 below.

Alternatively, one may adopt the insight of ABJ and define another outer region (which contains Θ_I^{CT}) which, using our notation, is given by

$$\Theta_I^{ABJ} = \left\{ \theta \in \Theta : \begin{array}{c} \mathbf{P}(y=(0,0)) \leq \mathbf{C}_{S_\theta}((0,0)) + \int\limits_{M^\theta} \left(1+\frac{\varepsilon_2}{\theta_2}\right)\left(1+\frac{\varepsilon_1}{\theta_1}\right) dF(\varepsilon) \\ \mathbf{P}(y=(1,0)) \leq \mathbf{C}_{S_\theta}((1,0)) + \mathbf{P}\left(\varepsilon \in M^\theta\right) \\ \mathbf{P}(y=(0,1)) \leq \mathbf{C}_{S_\theta}((0,1)) + \mathbf{P}\left(\varepsilon \in M^\theta\right) \\ \mathbf{P}(y=(1,1)) \leq \mathbf{C}_{S_\theta}((1,1)) + \int\limits_{M^\theta} \frac{\varepsilon_2}{\theta_2}\frac{\varepsilon_1}{\theta_1} dF(\varepsilon) \end{array} \right\}$$

Appendix B shows that, compared with Θ_I as defined in equation (3.8), Θ_I^{CT} and Θ_I^{ABJ} are obtained by checking the support function dominance only for u equal to the canonical basis vectors in $\Re^4$ (Θ_I^{ABJ}), and by taking the canonical basis vectors in $\Re^4$ and each of these vectors multiplied by -1 (Θ_I^{CT}). Clearly, these inequalities are a small subset of the ones required to obtain the sharp identification region, and therefore give an outer region for θ.

Figure 6 and Table 1 report Θ_I, Θ_I^{CT}, and Θ_I^{ABJ} in a simple example with $(\varepsilon_1, \varepsilon_2) \overset{iid}{\sim} N(0,1)$ and $\Theta = [-5, 0]^2$. In the figure, Θ_I^{ABJ} is given by the union of the yellow, red, and black areas, and Θ_I^{CT} by the union of the red and black areas. Θ_I is the black region. In Section 5 we explain how this region is calculated. The data is generated with $\theta_1^\star = -1.15$, $\theta_2^\star = -1.4$, and using a selection mechanism which picks each of outcome $(0,0)$ and $(1,1)$ for 10% of $\omega : \varepsilon(\omega) \in [0, -\theta_1^\star] \times [0, -\theta_2^\star]$, and each of outcome $(1,0)$ and $(0,1)$ for 40% of $\omega : \varepsilon(\omega) \in [0, -\theta_1^\star] \times [0, -\theta_2^\star]$. Hence, the observed distribution is $\mathbf{P}(y) = [0.26572\ 0.34315\ 0.36531\ 0.02582]$. Our results clearly show that Θ_I is substantially smaller than Θ_I^{CT} and Θ_I^{ABJ} : the sharp identification region has an area which is 43.5% of Θ_I^{ABJ}, and 52% of Θ_I^{CT}.

4 Pure Strategies Only: A Further Simplification

4.1 The Random Set of Equilibrium Outcomes Generated by Pure Strategies

We now assume that players in each market do not randomize across their actions. In this case, the set S_θ takes its realizations in the vertices of $\Delta(\mathcal{A})$. For a given realization of $\underline{x}$ and ε, a strategy profile $a \in \mathcal{A}$ is a pure strategy Nash Equilibrium if

$$\pi_j(a_j, a_{-j}, x_j, \varepsilon_j, \theta) \geq \pi_j(\tilde{a}_j, a_{-j}, x_j, \varepsilon_j, \theta) \ \forall \tilde{a}_j \in \mathcal{A}_j \ \forall j. \tag{4.1}$$

Using this inequality, we define the following θ-dependent set:

$$Y_\theta(\underline{x}, \varepsilon) = \{y \in \mathcal{Y} : \exists \ a \in \mathcal{A} \text{ s.t. } y = a, \ \pi_j(a_j, a_{-j}, x_j, \varepsilon_j, \theta) \geq \pi_j(\tilde{a}_j, a_{-j}, x_j, \varepsilon_j, \theta) \ \forall \tilde{a}_j \in \mathcal{A}_j \ \forall j\}. \tag{4.2}$$

For a given value of θ, this is the set of outcomes generated by pure strategies Nash equilibria.[12] As we did for S_θ, we omit the explicit reference to this set's dependence on $\underline{x}$ and ε. Given Assumption 1, one can easily show that Y_θ is a random closed set in $\mathcal{Y}$ (see Definition 2). Because the realizations of Y_θ are subsets of the finite set $\mathcal{Y}$, it suffices that $\pi(\cdot)$ is a measurable (rather than continuous) function of $\underline{x}$ and ε in order to establish that Y_θ is a random closed set in $\mathcal{Y}$.

For the model to be correctly specified, it is necessary that at least for some parameter values a pure strategy Nash equilibrium exists $\mathbf{P} - a.s.$ Hence, we impose the following assumption:

Assumption 4 *For a subset of values of $\theta \in \Theta$ which include the values of θ that have generated the observed outcomes y, a pure strategy Nash equilibrium exists $\mathbf{P} - a.s.$*

Under Assumptions 1-4, the observed outcomes y are consistent with Nash behavior if and only if there exists at least one $\theta \in \Theta$ such that $y(\omega) \in Y_\theta(\omega)$ $\mathbf{P} - a.s.$ (i.e., y is a selection of Y_θ, see Definition 3). In what follows, we exploit this insight to provide an equivalent characterization of the identification region in equation (3.8) based on a finite number of moment inequalities which have to hold for $\underline{x} - a.s.$ Mathematically, this simplification is due to the fact that when only pure strategies are played, $\mathbb{E}(Q(S_\theta)|\underline{x})$ is a closed convex polytope, fully characterized by a finite number of supporting hyperplanes. The vertices of $Q(S_\theta)$ are determined by the selections of S_θ, which in turn are degenerate mixed strategy profiles placing probability one on a specific action for each player. Hence, the supporting hyperplanes determining $\mathbb{E}(Q(S_\theta)|\underline{x})$ can be easily obtained (see Theorem 4.1 below). On the other hand, when players randomize across their actions, in equilibrium they must be indifferent among the actions over which they place positive probability. This implies that the equilibrium mixed strategy profiles are a function of both θ and ε. If ε has a discrete

[12] Restrict the set S_θ to be a set of pure strategy Nash equilibria. Then under Assumption 2, Y_θ coincides with S_θ. However, under the more general Assumption 1-(iii), these two sets differ, and

$$Y_\theta(\underline{x}, \varepsilon) = \{y \in \mathcal{Y} : \exists \ a \in \mathcal{A} \text{ s.t. } y = g(a), \ \pi_j(a_j, a_{-j}, x_j, \varepsilon_j, \theta) \geq \pi_j(\tilde{a}_j, a_{-j}, x_j, \varepsilon_j, \theta) \ \forall \tilde{a}_j \in \mathcal{A}_j \ \forall j\}.$$

distribution, $\mathbb{E}\left(Q\left(S_\theta\right)|\, \underline{x}\right)$ remains a convex polytope, and one can exactly calculate its supporting hyperplanes. But when the distribution of ε is continuous, $\mathbb{E}\left(Q\left(S_\theta\right)|\, \underline{x}\right)$ may have infinitely many extreme points, and therefore one needs an infinite number of inequalities to determine whether $\mathbf{P}\left(y|\, \underline{x}\right)$ belongs to it.

For the case that there are no covariates $\underline{x}$, one can determine whether a random vector y is a selection of an almost surely non-empty random closed set Y_θ by using the results of Artstein (1983), Norberg (1992) and Molchanov (2005, Theorem 1.2.20 and Section 1.4.8). These results establish that $y \in \mathrm{Sel}\left(Y_\theta\right)$ if and only if[13]

$$\mathbf{P}\{y \in K\} \leq \mathbf{P}\{Y_\theta \cap K \neq \emptyset\} \text{ for all compact sets } K \subset \mathcal{Y}. \tag{4.3}$$

Because $\mathcal{Y}$ is finite, all subsets of $\mathcal{Y}$ are compact. The functional $\mathbf{P}\{Y_\theta \cap K \neq \emptyset\}$ on the right-hand side of (4.3) is called the capacity functional of Y_θ. The following definitions formally introduce this functional and a few related ones:

Definition 4 *Let Z be a random closed set in the topological space $\mathbb{F}$, and denote by $\mathcal{K}$ the family of compact subsets of $\mathbb{F}$.*
1. *A functional $\mathbf{T}_Z : \mathcal{K} \to [0,1]$ given by*

$$\mathbf{T}_Z\left(K\right) = \mathbf{P}\{Z \cap K \neq \emptyset\},\ K \in \mathcal{K}$$

*is said to be the **capacity functional** of Z.*
2. *A functional $\mathbf{C}_Z : \mathcal{K} \to [0,1]$ given by*

$$\mathbf{C}_Z\left(K\right) = \mathbf{P}\{Z \subset K\},\ K \in \mathcal{K}$$

*is said to be the **containment functional** of Z.*
3. *A functional $\mathbf{I}_Z : \mathcal{K} \to [0,1]$ given by*

$$\mathbf{I}_Z\left(K\right) = \mathbf{P}\{K \subset Z\},\ K \in \mathcal{K}$$

*is said to be the **inclusion functional** of Z.*

The following relationships hold:

$$\begin{aligned} \mathbf{C}_Z\left(K\right) &= 1 - \mathbf{T}_Z\left(K^c\right), \\ \mathbf{I}_Z\left(K\right) &= 1 - \mathbf{T}_{Z^c}\left(K\right). \end{aligned} \tag{4.4}$$

[13] Beresteanu and Molinari (2008, Proposition 4.1) use this result to establish sharpness of the identification region of the parameters of a best linear predictor with interval outcome data. Galichon and Henry (2006) use it to define a correctly specified partially identified structural model, and derive a Kolmogorov-Smirnov test for Choquet capacities.

where K^c and Z^c denote, respectively, the complement of the sets K and Z.

Example 2 *Consider a simple two player entry game similar to the one in Tamer (2003), omit the covariates, assume that players' payoffs are given by* $\pi_j = a_j (a_{-j}\theta_j + \varepsilon_j)$, *where* $a_j \in \{0,1\}$ *and* $\theta_j < 0$, $j = 1, 2$, *and assume that players only play pure strategy equilibria. Figure 2 plots the set* Y_θ *against the realizations of* $\varepsilon_1, \varepsilon_2$. *In this case,* $\mathbf{T}_{Y_\theta}(\{(0,0)\}) = \mathbf{P}(\varepsilon_1 \leq 0, \varepsilon_2 \leq 0)$, $\mathbf{T}_{Y_\theta}(\{(1,0)\}) = \mathbf{P}(\varepsilon_1 \geq 0, \varepsilon_2 \leq -\theta_2)$, $\mathbf{T}_{Y_\theta}(\{(0,1)\}) = \mathbf{P}(\varepsilon_1 \leq -\theta_1, \varepsilon_2 \geq 0)$, $\mathbf{T}_{Y_\theta}(\{(1,1)\}) = \mathbf{P}(\varepsilon_1 \geq -\theta_1, \varepsilon_2 \geq -\theta_2)$, $\mathbf{T}_{Y_\theta}(\{(1,0),(0,1)\}) = \mathbf{T}_{Y_\theta}(\{(1,0)\}) + \mathbf{T}_{Y_\theta}(\{(0,1)\}) - \mathbf{P}(0 \leq \varepsilon_1 \leq -\theta_1, 0 \leq \varepsilon_2 \leq -\theta_2)$. *The capacity functional of the remaining subsets of* $\mathcal{Y}$ *can be calculated similarly. Corollary 5.1 and Algorithm 5.1 show that the capacity functional of those remaining subsets can be obtained as sums of the capacity functional of the subsets reported here.* $\square$

For the case that there are covariates $\underline{x}$, the researcher observes the tuple $(y, \underline{x})$, and the random set Y_θ is a function of $\underline{x}$ (and of course ε). Hence, one needs to work with the pair $(Y_\theta, \underline{x})$. In this case, the results of Artstein (1983), Norberg (1992) and Molchanov (2005, Theorem 1.2.20 and Section 1.4.8) imply that $(y, \underline{x})$ is a selection of $(Y_\theta, \underline{x})$ if and only if

$$\mathbf{P}\{(y, \underline{x}) \in K \times L\} \leq \mathbf{P}\{(Y_\theta, \underline{x}) \cap K \times L \neq \emptyset\} \ \forall K \subset \mathcal{Y}, \forall \text{ compact sets } L \subset \mathcal{X}.$$

This inequality can be written as

$$\mathbf{P}(y \in K | \underline{x} \in L)\, \mathbf{P}(\underline{x} \in L) \leq \mathbf{P}\{Y_\theta \cap K \neq \emptyset | \underline{x} \in L\}\, \mathbf{P}(\underline{x} \in L) \ \ \forall K \subset \mathcal{Y}, \forall \text{ compact sets } L \subset \mathcal{X},$$

and it is satisfied if and only if

$$\mathbf{P}(y \in K | \underline{x}) \leq \mathbf{P}\{Y_\theta \cap K \neq \emptyset | \underline{x}\} \ \ \forall K \subset \mathcal{Y} \ \underline{x} - a.s. \tag{4.5}$$

Notice that given equation (4.4), inequalities (4.5) can be equivalently written as

$$\mathbf{C}_{Y_\theta|\underline{x}}(K) \leq \mathbf{P}(y \in K | \underline{x}) \leq \mathbf{T}_{Y_\theta|\underline{x}}(K) \ \ \forall K \subset \mathcal{Y} \ \underline{x} - a.s., \tag{4.6}$$

where the subscript $Y_\theta|\,\underline{x}$ denotes that the functional is for the random set Y_θ conditional on $\underline{x}$. We return to this representation of inequalities (4.5) when discussing the relationship between our analysis and that of CT. Clearly, if one considers all $K \subset \mathcal{Y}$, the left-hand side inequality in (4.6) is superfluous: when the inequalities in (4.6) are used, only subsets $K \subset \mathcal{Y}$ of cardinality up to half of the cardinality of $\mathcal{Y}$ are needed.

We define the identified set of parameters θ as

$$\Theta_I = \left\{\theta \in \Theta : \mathbf{P}(y \in K | \underline{x}) \leq \mathbf{T}_{Y_\theta|\underline{x}}(K) \ \ \forall K \subset \mathcal{Y} \ \underline{x} - a.s.\right\}. \tag{4.7}$$

For comparison purposes, we reformulate the definition of the identified sets given by ABJ and CT respectively through the capacity functional and the containment functional:

$$\Theta_I^{ABJ} = \left\{\theta \in \Theta : \mathbf{P}\left\{y = t | \underline{x}\right\} \leq \mathbf{T}_{Y_\theta|\underline{x}}(t) \;\; \forall t \in \mathcal{Y} \;\; \underline{x} - a.s.\right\}, \tag{4.8}$$

$$\Theta_I^{CT} = \left\{\theta \in \Theta : \mathbf{C}_{Y_\theta|\underline{x}}(t) \leq \mathbf{P}\left\{y = t | \underline{x}\right\} \leq \mathbf{T}_{Y_\theta|\underline{x}}(t) \;\; \forall t \in \mathcal{Y} \;\; \underline{x} - a.s.\right\}. \tag{4.9}$$

Both ABJ and CT acknowledge that the identification regions they give are not sharp. Comparing the sets in equations (4.8)-(4.9) with the set in equation (4.7), one observes that Θ_I^{ABJ} is obtained applying inequality (4.5) only for $K = \{y\}$ $\forall y \in \mathcal{Y}$. Similarly, Θ_I^{CT} is obtained applying inequality (4.6) only for $K = \{y\}$ (or, equivalently, applying inequality (4.5) for $K = \{y\}$ and $K = \mathcal{Y} \backslash \{y\}$ $\forall y \in \mathcal{Y}$). Clearly both ABJ and CT do not use the information contained in the remaining subsets of $\mathcal{Y}$, while this information is used to obtain Θ_I. Two questions arise: (1) whether Θ_I as defined in equation (4.7) coincides with Θ_I as defined in equation (3.8), hence yielding the sharp identification region of θ through a finite number of moment inequalities which need to hold for $\underline{x} - a.s.$; and (2) if and by how much Θ_I differs from Θ_I^{ABJ} and Θ_I^{CT}.

We answer here the first question. Section 4.2 answers the second question by looking at a simple example.

Theorem 4.1 *Assume that players use only pure strategies, that Assumptions 1-3 are satisfied, that θ is such that Assumption 4 is satisfied, and that no other information is available. Then for $\underline{x} - a.s.$ these two conditions are equivalent:*

1. $u'\mathbf{P}(y | \underline{x}) \leq \mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right) | \underline{x}\right] \; \forall \; u \in \mathfrak{S}$,
2. $\mathbf{P}(y \in K | \underline{x}) \leq \mathbf{T}_{Y_\theta|\underline{x}}(K) \; \forall K \subset \mathcal{Y}$.

Notice that a candidate value of $\theta \in \Theta$ such that with positive probability a pure strategy Nash equilibrium does not exist for a set of values of $\underline{x}$ of positive probability is trivially rejected as a member of Θ_I using either definitions of the identification region in equations (3.8) and (4.7).

4.2 Example: Entry Game With 2 Types of Players and Pure Strategies Only

Consider a game where in each market there are four potential entrants, two of each type. The two types differ from each other by their payoff function. This model is an extension of the seminal papers by Bresnahan and Reiss (1990, 1991). An empirical application of a version of this model appears in Ciliberto and Tamer (2004). We adopt the version of this model described in Berry and

Tamer (2007, pages 84-85), and for illustration purposes we simplify it by omitting the observable payoff shifters $\underline{x}$ and by setting to zero the constant in the payoff function.

Let $a_{jk} \in \{0,1\}$ be the strategy of firm $j = 1,2$ of type $k = 1,2$. Entry is denoted by $a_{jk} = 1$, with $a_{jk} = 0$ denoting staying out. Let $y_1 = a_{11} + a_{21}$ denote the number of rivals of type 1 and $y_2 = a_{12} + a_{22}$ the number of rivals of type 2 that a firm faces so that $y_k \in \{0,1,2\}$. Players $j = 1,2$ of type 1 and type 2 have respectively the following payoff functions:

$$\pi_{j1}(a_{j1}, a_{-j1}, a_{12}, a_{22}, \varepsilon_1) = a_{j1}(\theta_{11}(a_{-j1} + a_{12} + a_{22}) - \varepsilon_1), \tag{4.10}$$

$$\pi_{j2}(a_{j2}, a_{-j2}, a_{11}, a_{21}, \varepsilon_2) = a_{j2}(\theta_{21}(a_{11} + a_{21}) + \theta_{22}a_{-j2} - \varepsilon_2). \tag{4.11}$$

We assume that θ_{11}, θ_{21} and θ_{22} are strictly negative and that $\theta_{22} > \theta_{21}$. This means that a type 2 firm is worried more about rivals of type 1 than of rivals of its own type. Since firms of a given type are indistinguishable to the econometrician, the observable outcome is the number of firms of each type which enter the market. Hence there are 9 possible outcomes to this game: $\mathcal{Y} = \{(0,0), (0,1), (1,0), (1,1), (2,0), (0,2), (1,2), (2,1), (2,2)\}$. Figure 3 plots the set Y_θ against the realizations of $\varepsilon_1, \varepsilon_2$.

We use this example to illustrate our methodology. We first define the specific form taken by the set Y_θ given equations (4.10)-(4.11):

$$Y_\theta = \left\{ y \in \mathcal{Y} : \exists\, a \in \mathcal{A} \text{ s.t. } \begin{array}{l} y_1 = a_{11} + a_{21}, \\ y_2 = a_{12} + a_{22}, \\ a_{j1}(\theta_{11}(a_{-j1} + y_2) - \varepsilon_1) \geq (1 - a_{j1})(\theta_{11}(a_{-j1} + y_2) - \varepsilon_1), \ j = 1,2, \\ a_{j2}(\theta_{21}y_1 + \theta_{22}a_{-j2} - \varepsilon_2) \geq (1 - a_{j2})(\theta_{21}y_1 + \theta_{22}a_{-j2} - \varepsilon_2), \ j = 1,2. \end{array} \right\}$$

Because the set $\mathcal{Y}$ has cardinality 9, in principle there are $2^9 = 512$ inequality restrictions to consider, corresponding to each compact subset $K \subset \mathcal{Y}$. However, the number of inequalities to be checked is significantly smaller. In particular, by a simple application of Algorithm 5.1 below, the sharp identification region that we give is based on 26 inequalities, whereas the identification region obtained following CT's insight is based on 18 inequalities. Section 5 below addresses formally the issue of how to reduce the number of inequalities to be checked.

Figure 7 and Table 2 report Θ_I, Θ_I^{CT}, and Θ_I^{ABJ} in a simple example with $(\varepsilon_1, \varepsilon_2) \overset{iid}{\sim} N(0,1)$ and $\Theta = [-5,0]^3$. In the figure, Θ_I^{ABJ} is given by the union of the yellow, red and black segments, and Θ_I^{CT} by the union of the red and black segments. Θ_I is the black segment. Notice that the identification regions are segments because the outcomes $(0,0)$ and $(2,2)$ can only occur as unique equilibrium outcomes, and therefore imply two moment equalities which make θ_{21} and θ_{22} a function of θ_{11}. While, strictly speaking, the approach in ABJ does not take into account this fact, as it uses only upper bounds on the probabilities that each outcome occurs, it is clear (and indicated in their paper) that one can incorporate equalities into their method. Hence, we use the equalities on $\mathbf{P}(y=(0,0))$ and $\mathbf{P}(y=(2,2))$ also when calculating Θ_I^{ABJ}. We generate the data with $\theta_{11}^{\star}=-0.15$, $\theta_{21}^{\star}=-0.20$, and $\theta_{22}^{\star}=-0.10$ and use a selection mechanism to choose the equilibrium played in the many regions of multiplicity. The resulting observed distribution is $\mathbf{P}(y)=[0.3021\ 0.0335\ 0.0231\ 0.0019\ 0.2601\ 0.2779\ 0.0104\ 0.0158\ 0.0752]$. Our results clearly show that Θ_I is substantially smaller than Θ_I^{CT} and Θ_I^{ABJ}. The width of the bounds on each parameter vector obtained using our method is about 46% of the width obtained using ABJ's method, and about 63% of the width obtained using CT's method.

5 Computational Aspects of the Problem

In order to compute the sharp identification region, we need to calculate the support function of the random set $Q(S_\theta)$. This is achieved by applying the Method of Simulated Moments, see McFadden (1989) and Pakes and Pollard (1989). The first step in the procedure requires one to compute the set of all mixed strategy Nash equilibria for given realizations of the payoff shifters, $S_\theta(\underline{x},\varepsilon)$. This is a computationally challenging problem, though a well studied one which can be performed using the Gambit software described by McKelvey and McLennan (1996).[14] Notice that this step has to be performed regardless of which features of normal form games are identified: whether sufficient conditions are imposed for point identification of the parameter vector of interest, or this vector is restricted to lie in an outer region, or its sharp identification region is characterized through the methodology proposed in this paper.

In our case, for given realizations of $\underline{x}$ and ε, computation of the set $S_\theta(\underline{x},\varepsilon)$ is needed in order to obtain by simulation, for each $u \in \mathfrak{S}^{\Delta}$,

$$\mathbf{E}\left[h\left(Q\left(S_\theta\right),u\right)\middle|\,\underline{x}\right]=\mathbf{E}\left[\sup_{\sigma\in S_\theta(\underline{x},\varepsilon)} u'q(\sigma)\middle|\,\underline{x}\right]=\int \sup_{\sigma\in S_\theta(\underline{x},\varepsilon)} u'q(\sigma)\,dF(\varepsilon|\,\underline{x}).$$

One can simulate this integral using the following procedure.[15] For any $\underline{x}\in\mathcal{X}$, draw realizations of ε, denoted ε^b, $b=1,\ldots,B$, according to the distribution $F(\cdot|\,\underline{x})$ with identity covariance matrix. These draws stay fixed throughout the remaining steps. Transform the realizations ε^b, $b=1,\ldots,B$,

[14] The Gambit software can be freely downloaded at http://gambit.sourceforge.net/. Bajari, Hong, and Ryan (2007) use this software to compute the set of mixed strategy Nash equilibria in finite normal form games whose parameters are point identified.

[15] The procedure described here is very similar to the one proposed by Ciliberto and Tamer (2004). When the assumptions maintained by Bajari, Hong, and Ryan (2007, Section 3) are satisfied, their algorithm can be used to significantly reduce the computational burden associated with simulating the integral.

into draws with covariance matrix specified by θ. For each ε^b, compute the payoffs $\pi_j\left(\cdot, x_j, \varepsilon_j^b, \theta\right)$ for $j = 1, \ldots, J$ and obtain the set $S_\theta\left(\underline{x}, \varepsilon^b\right)$. Then compute the set $Q\left(S_\theta\left(\underline{x}, \varepsilon^b\right)\right)$. Pick a $u \in \mathfrak{S}^\Delta$, compute the support function $h\left(\operatorname{co}\left[Q\left(S_\theta\left(\underline{x}, \varepsilon^b\right)\right)\right], u\right)$, and average it over a large number of draws of ε^b.

The same reasoning gives that the conditional Aumann expectation $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$ can be approximated by a simulated Minkowski average,

$$\widehat{\mathbb{E}}_B\left(Q\left(S_\theta\right)|\,\underline{x}\right) = \frac{1}{B} \bigoplus_{b=1}^{B} \operatorname{co}\left[Q\left(S_\theta\left(\underline{x}, \varepsilon^b\right)\right)\right].$$

The strong law of large numbers in Molchanov (2005, Theorem 3.1.6) guarantees that as $B \to \infty$, i.e., the number of simulations increases, $\rho_H\left(\widehat{\mathbb{E}}_B\left(Q\left(S_\theta\right)|\,\underline{x}\right), \mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)\right) \to 0$ $\mathbf{P} - a.s.$ This in turn implies almost sure convergence of $\hat{\mathbf{E}}_B\left[h\left(\operatorname{co}\left[Q\left(S_\theta\left(\underline{x}, \varepsilon^b\right)\right)\right], u\right)\middle|\,\underline{x}\right]$ to $\mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right)|\,\underline{x}\right]$, uniformly in $u \in \mathfrak{S}^\Delta$, see Schneider (1993, Theorem 1.8.12).

Denoting by $\mathbf{W}_B(\theta) = \int d_H\left(\mathbf{P}\left(y|\,\underline{x}\right), \widehat{\mathbb{E}}_B\left(Q\left(S_\theta\right)|\,\underline{x}\right)\right) dF_{\underline{x}}$ the analog of $\mathbf{W}(\theta)$ from equation (3.9), with $\widehat{\mathbb{E}}_B\left(Q\left(S_\theta\right)|\,\underline{x}\right)$ replacing $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$, we have that by triangle inequality

$$\begin{aligned}
\sup_{\theta\in\Theta}\left|\mathbf{W}_B(\theta) - \mathbf{W}(\theta)\right| &\leq \sup_{\theta\in\Theta} \int \rho_H\left(\widehat{\mathbb{E}}_B\left(Q\left(S_\theta\right)|\,\underline{x}\right), \mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)\right) dF_{\underline{x}} \\
&\leq \sup_{\theta\in\Theta} \rho_H\left(\widehat{\mathbb{E}}_B\left(Q\left(S_\theta\right)\right), \mathbb{E}\left(Q\left(S_\theta\right)\right)\right),
\end{aligned}$$

where the last inequality follows by the properties of the conditional Aumann expectation, Molchanov (2005, Theorem 2.1.47-(v)). For each $\theta \in \Theta$, $\sqrt{B}\rho_H\left(\widehat{\mathbb{E}}_B\left(Q\left(S_\theta\right)\right), \mathbb{E}\left(Q\left(S_\theta\right)\right)\right)$ converges in distribution to the supremum of a Gaussian process (Molchanov (2005, Theorem 2.2.1)). Hence, the arguments in Manski and Tamer (2002, Proposition 5), Ciliberto and Tamer (2004), and Chernozhukov, Hong, and Tamer (2007) assure that an identification region based on the simulated conditional Aumann expectation and its support function delivers an approximation of Θ_I which converges to Θ_I with respect to the Hausdorff metric as $B \to \infty$. In what follows we do not differentiate between the set Θ_I, and its counterpart resulting from numerical simulations.

5.1 Computing the Identification Region in the Pure Strategies Case

When it is assumed that players play only pure strategies, one needs to calculate the capacity functional of the random set Y_θ. As established in Theorem 4.1, $\mathbf{T}_{Y_\theta|\underline{x}}(K)$, $K \subset \mathcal{Y}$, is equal to the expectation of the support function of the set $Q\left(S_\theta\right)$ evaluated at u equal to each of the $2^{\kappa_\mathcal{Y}}$ vectors with each entry equal to either 1 or 0. Hence $\mathbf{T}_{Y_\theta|\underline{x}}(K)$, $K \subset \mathcal{Y}$, can be approximated through the procedure described above.

The set Θ_I is defined by $2^{\kappa_{\mathcal{Y}}}$ inequalities which have to hold for $\underline{x} - a.s.$ This number can be, in practice, very large. However, we emphasize that once the set Y_θ has been computed, evaluating whether all $2^{\kappa_{\mathcal{Y}}}$ inequalities are satisfied is a matter of "bookkeeping." Nevertheless, it can be a demanding task when the number of players or the number of actions each player can take is large. Fortunately, in many cases there is no need to verify the complete set of $2^{\kappa_{\mathcal{Y}}}$ inequalities, because many are redundant.[16] In particular, if K_1 and K_2 are two disjoint subsets of $\mathcal{Y}$ such that

$$\{\omega : Y_\theta(\omega) \cap K_1 \neq \emptyset | \underline{x}\} \cap \{\omega : Y_\theta(\omega) \cap K_2 \neq \emptyset | \underline{x}\} = \emptyset, \tag{5.1}$$

that is, the set of ω for which Y_θ intersects both K_1 and K_2 has probability zero, then the inequality $\mathbf{P}\{y \in K_1 \cup K_2 | \underline{x}\} \leq \mathbf{P}\{Y_\theta \cap (K_1 \cup K_2) \neq \emptyset | \underline{x}\}$ does not add any information beyond that provided by the inequalities $\mathbf{P}\{y \in K_1 | \underline{x}\} \leq \mathbf{P}\{Y_\theta \cap K_1 \neq \emptyset | \underline{x}\}$ and $\mathbf{P}\{y \in K_2 | \underline{x}\} \leq \mathbf{P}\{Y_\theta \cap K_2 \neq \emptyset | \underline{x}\}$. Therefore, prior knowledge of some properties of the game can be very helpful in eliminating unnecessary inequalities. For example, in a Bresnahan and Reiss entry model with 4 players, if the number of entrants is identified, the number of inequalities to be verified reduces from 65,536 to at most 100. Theorem 5.1 below gives the general result. While its proof is simple, this result is conceptually and practically important.

Theorem 5.1 *Take $\theta \in \Theta$ and let Assumptions 1-4 hold. Consider a partition of Ω into sets $\Omega^1, \ldots, \Omega^M$ of positive probability. Let $\mathcal{Y}_i$*

$$\mathcal{Y}_i = \cup\{Y_\theta(\omega) : \omega \in \Omega^i\}.$$

denote the range of $Y_\theta(\omega)$ for $\omega \in \Omega^i$. Assume that $\mathcal{Y}_1, \ldots, \mathcal{Y}_M$ are disjoint. Then it suffices to check (4.5) only for all subsets K such that there is $i = 1, \ldots, M$ for which $K \subseteq \mathcal{Y}_i$.

A simple Corollary of Theorem 5.1, the proof of which is omitted, is the following:

Corollary 5.1 *Take $\theta \in \Theta$ and let Assumptions 1-4 hold. Assume that $\Omega = \Omega^1 \cup \Omega^2$ with $\Omega^1 \cap \Omega^2 = \emptyset$, such that $Y_\theta(\omega)$ is a singleton almost surely for $\omega \in \Omega^1$. Let $\mathcal{Y}_i = \cup_{\omega \in \Omega^i} Y_\theta(\omega)$, $i = 1, 2$, and assume that $\mathcal{Y}_1 \cap \mathcal{Y}_2 = \emptyset$ and that $\kappa_{\mathcal{Y}_2} \leq 2$. Then inequalities (4.5) hold if*

$$\mathbf{P}\{Y_\theta = \{t\} | \underline{x}\} \leq \mathbf{P}\{y = t | \underline{x}\} \leq \mathbf{P}\{t \in Y_\theta | \underline{x}\} \tag{5.2}$$

$\underline{x} - a.s.$ for all $t \in \mathcal{Y}$.

[16] The game we described in Section 4.2 above is an example for the possible elimination of redundant inequalities.

An implication of this Corollary is that in a static entry game with two players in which only pure strategies are played, the identification region proposed by CT coincides with ours, and is sharp.[17] In this example, $\mathcal{Y}_1 = \{(0,0), (1,1)\}$, $\mathcal{Y}_2 = \{(0,1), (1,0)\}$, and $\Omega^2 = \{\omega : Y_\theta \cap \mathcal{Y}_2 \neq \emptyset\}$. An application of Algorithm 5.1 below shows that actually the sharp identification region can be obtained by checking only five inequalities which have to hold for $\underline{x} - a.s.$, given by inequalities (4.5) for $K = \{(0,0)\}, \{(1,0)\}, \{(0,1)\}, \{(1,1)\}, \{(1,0), (0,1)\}$. On the other hand, the example in Section 3.3 above shows that CT's approach does not yield the sharp identification region when mixed strategies are allowed for.

When no prior knowledge of the game such as, for example, that required in Theorem 5.1 is available, it is still possible to use the insight in equation (5.1) within an algorithm that determines which inequalities yield the sharp identification region. In particular, one can use the following procedure to build a collection of sets $\mathcal{C}$ such that checking inequalities (4.5) for each element of $\mathcal{C}$ suffices for sharpness. That is, the algorithm decomposes $\mathcal{Y}$ into subsets such that Y_θ does not jointly hit any two of them with positive probability. Observe that

$$\mathbf{P}\{Y_\theta \cap K_1 \neq \emptyset, Y_\theta \cap K_2 \neq \emptyset | \underline{x}\} = \int 1(Y_\theta(\underline{x}, \varepsilon) \cap K_1 \neq \emptyset) 1(Y_\theta(\underline{x}, \varepsilon) \cap K_2 \neq \emptyset) dF(\varepsilon | \underline{x}),$$

so that this probability can be easily approximated by simulation for any $K_1, K_2 \subset \mathcal{Y}$ as described above. Hence, one can use the following algorithm to determine which inequalities to check.

Algorithm 5.1

0) Set $\mathcal{C}^1 = \mathcal{Y}$ and $\mathcal{C}^2 = \emptyset$.

1) For each $t^i, t^j \in \mathcal{C}^1$ $i \neq j$, if there exists a set $\tilde{\mathcal{X}}_{ij} \subset \mathcal{X}$ such that $\mathbf{P}\left(\tilde{\mathcal{X}}_{ij}\right) > 0$ and

$$\mathbf{P}\left\{Y_\theta \cap t^i \neq \emptyset, Y_\theta \cap t^j \neq \emptyset \middle| \underline{x}\right\} > 0 \ \forall \underline{x} \in \tilde{\mathcal{X}}_{ij}$$

let $\mathcal{C}^2 = \mathcal{C}^2 \cup \left\{t^i, t^j\right\}$. If $\mathcal{C}^2 \neq \emptyset$, set $\mathcal{C}^3 = \emptyset$.

2) For each $\left\{t^i, t^j\right\} \in \mathcal{C}^2$, for each $t^k \in \mathcal{C}^1, k \neq i, j$, if there exists a set $\tilde{\mathcal{X}}_{ijk} \subset \mathcal{X}$ such that $\mathbf{P}\left(\tilde{\mathcal{X}}_{ijk}\right) > 0$ and

$$\mathbf{P}\left\{Y_\theta \cap \left\{t^i, t^j\right\} \neq \emptyset, Y_\theta \cap t^k \neq \emptyset \middle| \underline{x}\right\} > 0 \ \forall \underline{x} \in \tilde{\mathcal{X}}_{ijk},$$

let $\mathcal{C}^3 = \mathcal{C}^3 \cup \left\{t^i, t^j, t^k\right\}$. If for all $\left\{t^i, t^j\right\} \in \mathcal{C}^2$, for each $t^k \in \mathcal{C}^1, k \neq i, j$, the above condition is not satisfied, let $\bar{m} = 2$ and stop. Else, let $\bar{m} = 3$, set $\mathcal{C}^{\bar{m}} = \emptyset$, and go to the next step.

[17] A literal application of ABJ's approach does not take into account the fact that in this game $(0,0)$ and $(1,1)$ only occur as unique equilibria of the game, and therefore does not yield the sharp identification region as ABJ discuss (see page 32)

3) For $\kappa_{\mathcal{Y}} \geq \bar{m} \geq 3$, repeat the same operation as follows. For each $\left\{t^{i_1}, \ldots, t^{i_{m-1}}\right\} \in \mathcal{C}^{m-1}$, for each $t^k \in \mathcal{C}^1, k \neq i_1, \ldots, i_{m-1}$, if there exists a set $\tilde{\mathcal{X}}_{i_1,\ldots,i_{m-1}} \subset \mathcal{X}$ such that $\mathbf{P}\left(\tilde{\mathcal{X}}_{i_1,\ldots,i_{m-1}}\right) > 0$ and

$$\mathbf{P}\left\{Y_\theta \cap \left\{t^{i_1}, \ldots, t^{i_{m-1}}\right\} \neq \emptyset, Y_\theta \cap t^k \neq \emptyset \middle| \underline{x}\right\} > 0 \ \forall \underline{x} \in \tilde{\mathcal{X}}_{i_1,\ldots,i_{m-1}},$$

let $\mathcal{C}^m = \mathcal{C}^m \cup \left\{t^{i_1}, \ldots, t^{i_{m-1}}, t^k\right\}$. If for all $\left\{t^{i_1}, \ldots, t^{i_{m-1}}\right\} \in \mathcal{C}^{m-1}$, for each $t^k \in \mathcal{C}^1, k \neq i_1, \ldots, i_{m-1}$, the above condition is not satisfied, let $\bar{m} = m - 1$ and stop. Else, let $\bar{m} = m$, set $\mathcal{C}^{\bar{m}} = \emptyset$, and continue.

The set Θ_I is then given by equation (4.7) for $K \in \bigcup_{m=1}^{\bar{m}} \mathcal{C}^m$.

One may wonder whether in general the above algorithm will yield a different set of inequalities compared to those used by ABJ or CT. The following result shows that in general the system of constraints (4.5) obtained by restricting attention to K being singletons, as it is done by ABJ, does not yield a full characterization of the random set Y_θ, and therefore is not suited to yield the sharp identification region. Hence, Algorithm 5.1 returns a different set of inequalities to be checked compared to ABJ when the assumptions of Theorem 5.2 are satisfied.

Theorem 5.2 *Let Assumptions 1-4 hold. Assume that there exists $\theta \in \Theta$, with $Y_\theta \neq \emptyset$ $\mathbf{P} - a.s.$, such that for all $\underline{x} \in \tilde{\mathcal{X}} \subset \mathcal{X}$, with $\mathbf{P}\left(\tilde{\mathcal{X}}\right) > 0$, the expected cardinality of Y_θ given $\underline{x}$ is strictly greater than one, and such that $\mathbf{P}\left\{\left\{t^1, t^2\right\} \cap Y_\theta \neq \emptyset \middle| \underline{x}\right\} < 1$ for all $t^1, t^2 \in \mathcal{Y}$. Then there exists a random vector z which satisfies inequalities (4.5) for $K = \{t\}$ $\forall t \in \mathcal{Y}$ but is not a selection of Y_θ.*

This result shows that the extra inequalities matter in general, compared to those used by ABJ, to fully characterize Y_θ and determine if $y \in \mathrm{Sel}(Y_\theta)$. In fact, the assumptions of the theorem are satisfied whenever the model has multiple equilibria with positive probability, which implies that the expected cardinality of Y_θ given $\underline{x}$ is strictly greater than one, and it has at least three different equilibria.

On the other hand, CT strengthen the use of the singleton-based inequalities through an insight that corresponds to the observation that

$$\text{(5.3)} \quad \mathbf{P}\{Y_\theta = \{t\}| \underline{x}\} = \mathbf{C}_{Y_\theta|\underline{x}}(\{t\}) \leq \mathbf{P}\{\mathbf{y} = t| \underline{x}\} \leq \mathbf{P}\{t \in Y_\theta | \underline{x}\} = \mathbf{T}_{Y_\theta|\underline{x}}(\{t\}) \ \forall t \in \mathcal{Y} \ \underline{x} - a.s.$$

For simplicity write $\mathbf{I}(t)$ instead of $\mathbf{I}(\{t\})$, and $\mathbf{I}(t^1, t^2)$ instead of $\mathbf{I}(\{t^1, t^2\})$ and the same for the capacity functional $\mathbf{T}$ and the containment functional $\mathbf{C}$. The following result shows that under certain assumptions the system of constraints (4.6) obtained by restricting attention to K being singletons, as it is done by CT, does not yield a full characterization of the random set Y_θ, and therefore is not suited to yield the sharp identification region. Hence when the assumptions of Theorem 5.3 are satisfied, Algorithm 5.1 returns a different set of inequalities to be checked than those used by CT.

Theorem 5.3 *Let Assumptions 1-4 hold. Assume that there exists* $\theta \in \Theta$, *with* $Y_\theta \neq \emptyset$ $\mathbf{P} - a.s.$, *such that for all* $\underline{x} \in \tilde{\mathcal{X}} \subset \mathcal{X}$, *with* $\mathbf{P}\left(\tilde{\mathcal{X}}\right) > 0$, *there exist* $t^1, t^2 \in \mathcal{Y}$ *with*

$$\mathbf{I}_{Y_\theta|\underline{x}}(t^1, t^2) > 0 \tag{5.4}$$

and

$$\mathbf{P}\{\kappa_{Y_\theta} > 1|\,\underline{x}\} > \mathbf{I}_{Y_\theta|\underline{x}}(t^1) + \mathbf{I}_{Y_\theta|\underline{x}}(t^2) - \mathbf{C}_{Y_\theta|\underline{x}}(t^1) - \mathbf{C}_{Y_\theta|\underline{x}}(t^2). \tag{5.5}$$

Then there exists a random vector z *which satisfies inequalities (5.3) but is not a selection of* Y_θ.

This result shows that the extra inequalities matter in general, compared to those used by CT, to fully characterize Y_θ and determine if $y \in \operatorname{Sel}(Y_\theta)$. In fact, the assumptions of the theorem are satisfied whenever (1) there are regions of the unobservables of positive probability where two different outcomes can result from equilibrium strategy profiles; and (2) the probability that the cardinality of Y_θ is greater than one exceeds the probability that each of these two outcomes is not a unique equilibrium. It is easy to see that these assumptions are not satisfied in a two player entry game where players are allowed only to play pure strategies, but they are satisfied in the four player, two type game described in Section 4.2.

5.2 Computing the Identification Region in the Mixed Strategies Case

Consider now the case where players randomize across their actions. The support function of the conditional Aumann expectation $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$ can be approximated by simulation as described at the beginning of Section 5. We now discuss how to discretize the unit sphere $\mathfrak{S}^\Delta$ in order to transform the definition of Θ_I based on the support function in equation (3.8) into a definition involving a finite number of moment inequalities which have to hold for $\underline{x} - a.s.$

Our proposal is to use a ν_k-net on the sphere. For a given $0 < \nu_k < 1$, a ν_k-net of $\mathfrak{S}^\Delta$ is given by a finite subset of vectors $U_k = \{u_1, \ldots, u_k\} \subset \mathfrak{S}^\Delta$ such that for every vector $u \in \mathfrak{S}^\Delta$ there is an $i \in \{1, \ldots, k\}$ with the property that u is within distance ν_k from u_i. By construction,

$$\nu_k = \max_{u \in \mathfrak{S}^\Delta} \min_{1 \le i \le k} \|u - u_i\|.$$

One can build this ν_k-net such that it contains a number of points that is $O\left(\nu_k^{-(\kappa_{\mathcal{Y}}-1)}\right)$, see Gardner and Milanfar (2003, Lemma 7.1). An easy to implement algorithm which allows one to build such a ν_k-net is provided in Lovisolo and DaSilva (2001, Section 2). This is the algorithm used for the construction of the ν_k-net used in the examples in Sections 3.3 and 6.2. In those examples, we approximate Θ_I using 1,160 points uniformly distributed over the unit sphere.[18] This corresponds to a value of ν_k equal to 0.222.

Once the ν_k-net is constructed, one can define the set

$$\Theta_{IU_k} = \left\{\theta \in \Theta : u'\mathbf{P}\left(y|\,\underline{x}\right) \leq \mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right)|\,\underline{x}\right] \ \forall\ u \in U_k \ \underline{x} - a.s.\right\}. \tag{5.6}$$

Then by construction $\Theta_I \subset \Theta_{IU_k}$. As $\nu_k \to 0$, Θ_{IU_k} decreases to Θ_I, a result formally established in the following Theorem.

Theorem 5.4 *Let Θ_I and Θ_{IU_k} be defined by equations (3.8) and (5.6) respectively. Let U_K be a ν_k-net of $\mathfrak{S}^\Delta$, with $\nu_k \to 0$ as $k \to \infty$. Then $\rho_H\left(\Theta_I, \Theta_{IU_k}\right) \to 0$ as $k \to \infty$.*

6 Extensions to Other Solution Concepts

While in the main body of this paper we focus on economic models of games in which Nash Equilibrium is the solution concept employed, our approach easily extends to other solution concepts. Here we consider the case that players are assumed to be only level-1 rational, and the case that they are assumed to play correlated strategies. For simplicity, we exemplify these extensions using a two player simultaneous move static game of entry with complete information.

6.1 Level-1 Rationality

Suppose that players are only assumed to be level-1 rational. The identification problem under this weaker solution concept was first studied by Aradillas-Lopez and Tamer (2008, AT henceforth). Let the econometrician observe players' actions, so that Assumption 2 is satisfied. A level-1 rational profile is given by a mixed strategy for each player that is a best response to one of the possible mixed strategies of her opponent. In this case one can define the θ-dependent set

$$R_\theta\left(\underline{x}, \varepsilon\right) = \left\{\sigma \in \Delta\left(\mathcal{A}\right) : \begin{array}{l} \forall j\ \exists\ \tilde{\sigma}_{-j} \in \Delta\left(\mathcal{A}_{-j}\right) \text{ s.t.} \\ \pi_j\left(\sigma_j, \tilde{\sigma}_{-j}, x_j, \varepsilon_j, \theta\right) \geq \pi_j\left(\sigma'_j, \tilde{\sigma}_{-j}, x_j, \varepsilon_j, \theta\right)\ \forall \sigma'_j \in \Delta\left(\mathcal{A}_j\right) \end{array}\right\}.$$

[18] We experimented with a significantly larger number of points, but this did not appreciably change the approximation of Θ_I.

Omitting the explicit reference to its dependence on $\underline{x}$ and ε, R_θ is the set of level-1 rational strategy profiles of the game. By similar arguments to what we used above, this is a random closed set in $\Delta(\mathcal{A})$. Figure 4 plots this set against the possible realizations of $\varepsilon_1, \varepsilon_2$, in a simple two player simultaneous move, complete information, static game of entry. We ignore covariates for ease of exposition, and assume that players' payoffs are given by $\pi_j = a_j(a_{-j}\theta_j + \varepsilon_j)$, where $a_j \in \{0,1\}$ and θ_j is assumed to be negative (monopoly payoffs are higher than duopoly payoffs), $j = 1, 2$.

The same approach of Section 3 allows us to obtain the sharp identification region for θ as

$$\Theta_I = \left\{\theta \in \Theta : u'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}\left[h(Q(R_\theta), u)|\underline{x}\right] \ \forall\, u \in \mathfrak{S}^{\Delta}\ \underline{x} - a.s.\right\},$$

with

$$q(R_\theta) = \left\{\left([q(\sigma)]_k, k = 1, \ldots, \kappa_{\mathcal{Y}}\right) : \sigma \in \mathrm{Sel}(R_\theta)\right\},$$

where $[q(\sigma)]_k, k = 1, \ldots, \kappa_{\mathcal{Y}}$, is defined in equation (3.3).

Observing, however, that in our simple example for $\omega \in \Omega_1$ such that $\varepsilon(\omega) \in [0, -\theta_1] \times [0, -\theta_2]$,

$$\left[q\left(\left(\frac{\varepsilon_2(\omega)}{-\theta_2}, \frac{\varepsilon_1(\omega)}{-\theta_1}\right)\right)\right]_k \in \mathrm{co}\left[\left\{[q(0,0)]_k, [q(1,0)]_k, [q(0,1)]_k, [q(1,1)]_k\right\}\right],$$

$k = 1, \ldots, 4$, it follows that $\mathbb{E}(Q(R_\theta)|\underline{x})$ is equal to $\mathbb{E}\left(Q\left(\tilde{R}_\theta\right)\Big|\underline{x}\right)$, with $\tilde{R}_\theta$ restricted to be the set of level-1 rational pure strategies. Hence, by Theorem 4.1, Θ_I can be obtained by checking a finite number of moment inequalities.

For the case that ε has a discrete distribution, AT (Section 3.1) suggest to obtain the sharp identification region as the set of parameter values that return value zero for the objective function of a linear programming problem. For the general case in which ε may have a continuous distribution, AT apply the same insight of CT and characterize an outer identification region through eight moment inequalities as in equation (4.9). One may also extend ABJ's approach to this problem, and obtain a larger outer region through four moment inequalities as in equation (4.8). Our approach, which yields the sharp identification region, in this simple example requires one to check 14 inequalities.

As shown in AT (Figure 3), the model with level-1 rationality only places upper bounds on θ_1, θ_2. Figure 8 plots the upper contours of Θ_I, Θ_I^{CT}, and Θ_I^{ABJ} in a simple example with $(\varepsilon_1, \varepsilon_2) \overset{iid}{\sim} N(0,1)$ and $\Theta = [-5, 0]^2$. The data is generated with $\theta_1^\star = -1.15$, $\theta_2^\star = -1.4$, and using a selection mechanism which picks outcome $(0,0)$ for 40% of $\omega : \varepsilon(\omega) \in [0, -\theta_1^\star] \times [0, -\theta_2^\star]$, outcome $(1,1)$ for 10% of $\omega : \varepsilon(\omega) \in [0, -\theta_1^\star] \times [0, -\theta_2^\star]$, and each of outcome $(1,0)$ and $(0,1)$ for 25% of $\omega : \varepsilon(\omega) \in [0, -\theta_1^\star] \times [0, -\theta_2^\star]$. Hence, the observed distribution is $\mathbf{P}(y) = [0.5048\ 0.2218\ 0.1996\ 0.0738]$. Our methodology allows us to obtain significantly lower upper contours compared to AT (and CT) and ABJ. The upper bounds on θ_1, θ_2 resulting from the projections of Θ_I^{ABJ}, Θ_I^{CT} and Θ_I are, respectively, $(-0.02, -0.02,)$, $(-0.15, -0.26)$, and $(-0.54, -0.61)$.

6.2 Objective Correlated Equilibria

Suppose that players play correlated equilibria, a notion introduced by Aumann (1974). A correlated equilibrium can be interpreted as the distribution of play instructions given by some "trusted authority" to the players. Each player is given her instruction privately but does not know the instruction received by others. The distribution of instructions is common knowledge across all players. Then a correlated joint strategy $\gamma \in \Delta^{\kappa_{\mathcal{A}}-1}$, where $\Delta^{\kappa_{\mathcal{A}}-1}$ denotes the set of probability distribution on $\mathcal{A}$, is an equilibrium if, conditional on knowing that her own instruction is to play a_j, each player j has no incentive to deviate to any other strategy a_j', assuming that the other players follow their own instructions. In this case one can define the θ-dependent set

$$C_\theta(\underline{x}, \varepsilon) = \left\{ \gamma \in \Delta^{\kappa_{\mathcal{A}}-1} : \begin{array}{l} \sum\limits_{a_{-j} \in \mathcal{A}_{-j}} \gamma(a_j, a_{-j}) \pi_j(a_j, a_{-j}, x_j, \varepsilon_j, \theta) \geq \\ \sum\limits_{a_{-j} \in \mathcal{A}_{-j}} \gamma(a_j, a_{-j}) \pi_j\left(a_j', a_{-j}, x_j, \varepsilon_j, \theta\right), \ \forall a_j \in \mathcal{A}_j, \ \forall a_j' \in \mathcal{A}_j, \ \forall j \end{array} \right\}.$$

Omitting the explicit reference to its dependence on $\underline{x}$ and ε, C_θ is the set of correlated equilibrium strategies of the game. By similar arguments as those used before, it is a random closed set in $\Delta^{\kappa_{\mathcal{A}}-1}$. Notice that C_θ is defined by a finite number of linear inequalities on the set $\Delta^{\kappa_{\mathcal{A}}-1}$ of correlated strategies, and therefore it is a non-empty polytope. Yang (2008) is the first to use this fact, along with the fact that $\mathrm{co}[Q(S_\theta)] \subset C_\theta$, to develop a computationally easy-to-implement estimator for an outer identification region of θ, when the solution concept employed is Nash equilibrium. Here we provide a simple characterization of the sharp identification region Θ_I, when the solution concept employed is objective correlated equilibrium. In particular, the same approach of Section 3 allows us to obtain the sharp identification region for θ as

$$\Theta_I = \left\{ \theta \in \Theta : u'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}[h(C_\theta, u)|\underline{x}] \ \forall\, u \in \mathfrak{S}^\Delta \ \underline{x} - a.s. \right\}.$$

In our simple two player simultaneous move, complete information, static game of entry, $\mathcal{A}_j = \{0, 1\}$, $j = 1, 2$, $\mathcal{A} = \{(0,0), (1,0), (0,1), (1,1)\}$. We ignore covariates for ease of exposition, and assume that players' payoffs are given by $\pi_j = a_j(a_{-j}\theta_j + \varepsilon_j)$, where $a_j \in \{0, 1\}$ and θ_j is assumed to be negative (monopoly payoffs are higher than duopoly payoffs), $j = 1, 2$. Figure 5 plots the set C_θ against the possible realizations of $\varepsilon_1, \varepsilon_2$, for this example. Notice that for $\omega \in \Omega_1$ such that $\varepsilon(\omega) \notin [0, -\theta_1] \times [0, -\theta_2]$, the game is dominance solvable and therefore $C_\theta(\omega)$ is given by the singleton $Q(S_\theta(\omega))$ resulting from the unique Nash equilibrium in those regions. For $\omega \in \Omega_1$ such that $\varepsilon(\omega) \in [0, -\theta_1] \times [0, -\theta_2]$, $C_\theta(\omega)$ is given by a polytope with five vertices, three of which are implied by Nash equilibria, see Calvó-Armengol (2006).[19] Also in this case one can extend the approaches of ABJ and CT to obtain outer identification regions defined, respectively, by four and eight moment inequalities as in equations (4.8)-(4.9).

Figure 9 and Table 3 report Θ_I, Θ_I^{CT}, and Θ_I^{ABJ} in a simple example with $(\varepsilon_1, \varepsilon_2) \overset{iid}{\sim} N(0,1)$ and $\Theta = [-5, 0]^2$. In the figure, Θ_I^{ABJ} is given by the union of the yellow, red and black areas, and Θ_I^{CT} by the union of the red and black areas. Θ_I is the black region. The data is generated with $\theta_1^\star = -1.15$, $\theta_2^\star = -1.4$, and using a selection mechanism which picks each of outcome $(0,0)$ and $(1,1)$ for 10% of $\omega : \varepsilon(\omega) \in [0, -\theta_1^\star] \times [0, -\theta_2^\star]$, and each of outcome $(1,0)$ and $(0,1)$ for 40% of $\omega : \varepsilon(\omega) \in [0, -\theta_1^\star] \times [0, -\theta_2^\star]$. Hence, the observed distribution is $\mathbf{P}(y) = [0.26572\ 0.34315\ 0.36531\ 0.02582]$. Also in this case Θ_I is smaller than Θ_I^{CT} and Θ_I^{ABJ}, although the reduction in the size of the identification region is less pronounced than in the case where mixed strategy Nash equilibrium is the solution concept.

7 Conclusions

This paper introduces a computationally feasible characterization of the sharp identification region of the model parameters in static, simultaneous move finite games of complete information in the presence of multiple equilibria. The methodology that we propose allows us to bypass the need to directly deal with infinite dimensional nuisance parameters, the selection mechanisms, a simplification that was considered unattainable in the related literature (see, e.g., Berry and Tamer (2007)).

For the case that players are assumed to play only pure strategies, we show that the sharp identification region is given by a finite number of moment inequalities which have to hold for $\underline{x}-a.s.$ While finite, this number of moment inequalities can be very large in certain games. However, we

[19] These vertices are

$$\begin{aligned}
\gamma_0(\omega) &= [0,0,1,0] \\
\gamma_1(\omega) &= \left[1, -\tfrac{\varepsilon_2(\omega)}{\theta_2+\varepsilon_2(\omega)}, -\tfrac{\varepsilon_1(\omega)}{\theta_1+\varepsilon_1(\omega)}, 0\right]\left(1 - \tfrac{\varepsilon_1(\omega)}{\theta_1+\varepsilon_1(\omega)} - \tfrac{\varepsilon_2(\omega)}{\theta_2+\varepsilon_2(\omega)}\right)^{-1} \\
\gamma_2(\omega) &= \left[\left(1+\tfrac{\varepsilon_2(\omega)}{\theta_2}\right)\left(1+\tfrac{\varepsilon_1(\omega)}{\theta_1}\right), -\tfrac{\varepsilon_2(\omega)}{\theta_2}\left(1+\tfrac{\varepsilon_1(\omega)}{\theta_1}\right), -\left(1+\tfrac{\varepsilon_2(\omega)}{\theta_2}\right)\tfrac{\varepsilon_1(\omega)}{\theta_1}, \tfrac{\varepsilon_2(\omega)}{\theta_2}\tfrac{\varepsilon_1(\omega)}{\theta_1}\right] \\
\gamma_3(\omega) &= \left[0, -\tfrac{\varepsilon_2(\omega)}{\theta_2+\varepsilon_2(\omega)}, -\tfrac{\varepsilon_1(\omega)}{\theta_1+\varepsilon_1(\omega)}, \tfrac{\varepsilon_1(\omega)}{\theta_1+\varepsilon_1(\omega)}\tfrac{\varepsilon_2(\omega)}{\theta_2+\varepsilon_2(\omega)}\right]\left(\tfrac{\varepsilon_1(\omega)}{\theta_1+\varepsilon_1(\omega)}\tfrac{\varepsilon_2(\omega)}{\theta_2+\varepsilon_2(\omega)} - \tfrac{\varepsilon_1(\omega)}{\theta_1+\varepsilon_1(\omega)} - \tfrac{\varepsilon_2(\omega)}{\theta_2+\varepsilon_2(\omega)}\right)^{-1} \\
\gamma_4(\omega) &= [0,1,0,0]
\end{aligned}$$

show that many such inequalities may be redundant, and we provide a simple algorithm that allows the researcher to determine a (often significantly) smaller set of moment inequalities that are sufficient to preserve sharpness.

When players may also randomize across their actions, the sharp identification region cannot in general be characterized through a finite number of moment inequalities. Intuitively, this is because there is additional information provided by the fact that players must be indifferent among the actions that they play with positive probability according to a given equilibrium strategy. While there is an infinite number of inequalities characterizing the sharp identification region, we show that this region can be approximated arbitrarily accurately through a finite number of moment inequalities, which again have to hold for $\underline{x} - a.s.$ As this number of moment inequalities increases, the approximated identification region converges to the sharp identification region with respect to the Hausdorff metric.

We acknowledge that the method proposed in this paper may be, for some models, computationally more intensive than existing methods (e.g., Andrews, Berry, and Jia (2004), Ciliberto and Tamer (2004)). However, the benefits in terms of identification coming from considering these additional inequalities may be substantial, as illustrated by our examples.

A Proofs

Theorem 3.1.

Proof. In order to establish sharpness, it suffices to show that $\Theta_I = \Theta_I^{\star}$. Take $\theta \in \Theta_I$. Then $\exists\ q \in \text{Sel}\left(Q\left(S_\theta\right)\right) : \mathbf{E}\left(q|\,\underline{x}\right) = \mathbf{P}\left(y|\,\underline{x}\right)$. Hence a selection mechanism that selects with probability 1 a $\sigma \in \text{Sel}\left(S_\theta\right) : q{=}\left(\left[q\left(\sigma\right)\right]_k, k = 1, \ldots, \kappa_{\mathcal{Y}}\right)$ is admissible and assures that $\theta \in \Theta_I^{\star}$ (notice that by the definition of $Q\left(S_\theta\right)$, such a $\sigma \in \text{Sel}\left(S_\theta\right)$ exists). Conversely, take $\theta \in \Theta_I^{\star}$. Then there exists an admissible selection mechanism ψ which picks a selection $\sigma \in \text{Sel}\left(S_\theta\right)$, such that $\mathbf{P}\left(y|\,\underline{x}\right) = \mathbf{P}\left(y|\,\underline{x};\theta,\psi\right) = \mathbf{E}\left(q|\,\underline{x}\right)$ for $\underline{x} - a.s.$ for $q{=}\left(\left[q\left(\sigma\right)\right]_k, k = 1, \ldots, \kappa_{\mathcal{Y}}\right)$. Hence $\theta \in \Theta_I$. ∎

Theorem 4.1.

Proof. Let condition (1) hold. Notice that because the support function is positively homogeneous, this condition is equivalent to $u'\mathbf{P}\left(y|\,\underline{x}\right) \leq \mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right)|\,\underline{x}\right]\ \forall\ u \in \Re^{\kappa_{\mathcal{Y}}}$. Take any $\bar{u} \in \Re^{\kappa_{\mathcal{Y}}}$ such that its entries are zeros and ones. The coordinates of $\bar{u}$ correspond to the vertices of $\Delta\left(\mathcal{A}_1\right) \times \ldots \times \Delta\left(\mathcal{A}_J\right)$, and determine a subset $K_{\bar{u}}$ of $\mathcal{Y}$. By definition, the scalar product $\bar{u}'\mathbf{P}\left(y|\,\underline{x}\right)$ equals $\mathbf{P}\left(y \in K_{\bar{u}}|\,\underline{x}\right)$. Moreover,

$$h\left(\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right), \bar{u}\right) = \mathbf{E}\left[h\left(Q\left(S_\theta\right), \bar{u}\right)|\,\underline{x}\right] = \mathbf{E}\left[\left.\sup_{\sigma \in S_\theta} \bar{u}'q\left(\sigma\right)\right|\underline{x}\right].$$

Because we are allowing only pure strategy equilibria, the realizations of any $\sigma \in S_\theta$ are vectors of zeros and ones. Hence, $\forall \omega_1 \in \Omega_1$, $\left[q\left(\sigma\left(\omega_1\right)\right)\right]_k = 1$ if $\prod_{j=1}^{J} \sigma_j\left(\omega_1, t_j^k\right) = 1$, and zero otherwise. Thus, given the choice of $\bar{u}$, the value of $\bar{u}'q\left(\sigma\left(\omega_1\right)\right)$ equals one if $y\left(\omega_1\right) \in K_{\bar{u}}$ and zero otherwise. Hence, condition (1) reduces to

$$\begin{aligned}\mathbf{P}\left(y \in K_{\bar{u}}|\,\underline{x}\right) &= \bar{u}'\mathbf{P}\left(y|\,\underline{x}\right) \leq \mathbf{E}\left[h\left(Q\left(S_\theta\right), \bar{u}\right)|\,\underline{x}\right] = \mathbf{E}\left[\left.\sup_{\sigma \in S_\theta} \bar{u}'q\left(\sigma\right)\right|\underline{x}\right] \\ &= \mathbf{E}\left[1\left(Y_\theta \cap K_{\bar{u}} \neq \emptyset\right)|\,\underline{x}\right] = \mathbf{P}\left\{Y_\theta \cap K_{\bar{u}} \neq \emptyset|\,\underline{x}\right\}.\end{aligned}$$

Choosing $\bar{u}$ equal to each of the $2^{\kappa_{\mathcal{Y}}}$ vectors with entries equal to either 1 or 0, yields condition (2).

Suppose now that condition (2) holds. Then $y \in \text{Sel}\left(Y_\theta\right)$. This means that $\exists\ \sigma \in \text{Sel}\left(S_\theta\right)$: $\sigma_j\left(\omega_1, a_j\right) = 1$ if $y_j\left(\omega_1\right) = a_j$ and zero otherwise. Hence, $\mathbf{P}\left(y|\,\underline{x}\right) \in \mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$ for $\underline{x} - a.s.$, and therefore condition (1) is satisfied. ∎

Theorem 5.1.

Proof. If $K_i = K \cap \mathcal{Y}_i$ for $i = 1, \ldots, M$, then

$$\mathbf{P}\{y \in K|\,\underline{x}\} = \sum_{i=1}^{M} \mathbf{P}\{y \in K_i|\,\underline{x}\} \leq \sum_{i=1}^{M} \mathbf{P}\{Y_\theta \cap K_i \neq \emptyset|\,\underline{x}\} = \mathbf{P}\{Y_\theta \cap K \neq \emptyset|\,\underline{x}\},\ \underline{x} - a.s.,$$

since Y_θ cannot hit both K_i and K_j simultaneously in view of the disjointedness assumption. ■

Theorem 5.2.

Proof. To simplify the notation, we omit the conditioning on $\underline{x} \in \tilde{\mathcal{X}}$ and the subscript $Y_\theta|\underline{x}$ indexing the inclusion functional in all that follows. For each finite set $K = \{t^1, \ldots, t^k\} \subset \mathcal{Y}$ write $\mathbf{I}(t^1, \ldots, t^k) = \mathbf{P}\{K \subset Y_\theta\}$ for the inclusion functional of K. Since the expected cardinality of Y_θ is more than one, there exist $t^1, t^2 \in \mathcal{Y}$ such that $\mathbf{I}(t^1, t^2) > 0$.

Assume that $\mathbf{I}(t^1) + \mathbf{I}(t^2) \geq 1$. Then choose a random element z which takes values t^1 and t^2 with probabilities that sum to one and are dominated by $\mathbf{I}(t^1)$ and $\mathbf{I}(t^2)$ respectively. Then (4.5) holds for all singletons K, while if $K = \{t^1, t^2\}$, then

$$\mathbf{P}(z \in K) = \mathbf{P}\left(z = t^1\right) + \mathbf{P}\left(z = t^2\right) = 1$$

cannot be smaller than $\mathbf{P}\left\{Y_\theta \cap \left\{t^1, t^2\right\} \neq \emptyset\right\}$, since the latter is less than one by assumption.

Assume that $\mathbf{I}(t^1) + \mathbf{I}(t^2) < 1$. Then construct a random element z that takes values t^1 and t^2 with probabilities $\mathbf{I}(t^1)$ and $\mathbf{I}(t^2)$ and some values t outside of $\left\{t^1, t^2\right\}$ with probabilities dominated by $\mathbf{I}(t)$ for the chosen t, such that $\sum_{t \in \mathcal{Y}} \mathbf{P}\{z = t\} = 1$. This is possible, since the total sum of $\mathbf{I}(t)$ over $t \in \mathcal{Y}$ equals the expected cardinality of Y_θ and so is at least one. Then

$$\mathbf{P}\left(z = t^1\right) + \mathbf{P}\left(z = t^2\right) = \mathbf{I}(t^1) + \mathbf{I}(t^2)$$

whereas

$$\mathbf{P}\left\{Y_\theta \cap \left\{t^1, t^2\right\} \neq \emptyset\right\} = \mathbf{I}(t^1) + \mathbf{I}(t^2) - \mathbf{I}(t^1, t^2) < \mathbf{I}(t^1) + \mathbf{I}(t^2) = \mathbf{P}\{z \in \left\{t^1, t^2\right\}\}.$$

■

Theorem 5.3.

Proof. To simplify the notation, we omit the conditioning on $\underline{x} \in \tilde{\mathcal{X}}$ and the subscript $Y_\theta|\underline{x}$ indexing the containment, inclusion, and capacity functionals in all that follows. We use notation from the proof of Theorem 5.2. Note that

$$\sum_{t \in \mathcal{Y}} \mathbf{C}(t) = \mathbf{P}\{\kappa_{Y_\theta} = 1\} \tag{A.1}$$

and

$$\sum_{t \in \mathcal{Y}} \mathbf{I}(t) = \mathbf{E}\left(\kappa_{Y_\theta}\right).$$

Note that the expected cardinality $\mathbf{E}\left(\kappa_{Y_\theta}\right)$ is greater than one, since the random set Y_θ is almost surely non-empty and has cardinality at least two with positive probability.

Take t^1, t^2 to satisfy the assumptions of Theorem 5.3. If $\mathbf{I}(t^1) + \mathbf{I}(t^2) \geq 1$, then define z which takes values $t \neq t^1, t^2$ with probabilities $\mathbf{C}(t)$. Furthermore, assume that z takes values t^1 and t^2 with the probabilities $p_{t^1} = \mathbf{C}(t^1) + \delta_{t^1}$ and $p_{t^2} = \mathbf{C}(t^2) + \delta_{t^2}$ dominated by $\mathbf{I}(t^1)$ and $\mathbf{I}(t^2)$ respectively, with $\delta_{t^1}, \delta_{t^2}$ two non-negative constants for which

$$\sum_{t \in \mathcal{Y}} \mathbf{C}(t) + \delta_{t^1} + \delta_{t^2} = 1.$$

Such constants $\delta_{t^1}, \delta_{t^2}$ exist, because the left-hand side in the above expression is less than 1 for $\delta_{t^1} = \delta_{t^2} = 0$ since $\kappa_{Y_\theta} > 1$ with positive probability, and the left-hand side is greater than or equal to 1 for $\delta_{t^1} = \mathbf{I}(t^1) - \mathbf{C}(t^1)$ and $\delta_{t^2} = \mathbf{I}(t^2) - \mathbf{C}(t^2)$. This together with equation (A.1) implies that $\mathbf{P}\{\kappa_{Y_\theta} = 1\} + \delta_{t^1} + \delta_{t^2} = 1$, whence

$$p_{t^1} + p_{t^2} = \mathbf{C}(t^1) + \mathbf{C}(t^2) + \delta_{t^1} + \delta_{t^2} = \mathbf{C}(t^1) + \mathbf{C}(t^2) + \mathbf{P}\{\kappa_{Y_\theta} > 1\}.$$

By (5.5), the right-hand side is strictly greater than $\mathbf{I}(t^1) + \mathbf{I}(t^2) > \mathbf{I}(t^1) + \mathbf{I}(t^2) - \mathbf{I}(t^1, t^2) = \mathbf{T}(t^1, t^2) = \mathbf{P}\left\{Y_\theta \cap \left\{t^1, t^2\right\} \neq \emptyset\right\}$, where the inequality follows by equation (5.4). Thus, $\mathbf{P}\{z \in K\}$ is not dominated by $\mathbf{T}(K)$ for $K = \{t^1, t^2\}$, i.e. z is not a selection of Y_θ.

Consider now the case that $\mathbf{I}(t^1) + \mathbf{I}(t^2) < 1$. Then construct a random element z that takes values t^1 and t^2 with probabilities $p_{t^1} = \mathbf{I}(t^1)$ and $p_{t^2} = \mathbf{I}(t^2)$ and some values t outside of $\{t^1, t^2\}$ with probabilities $p_t = \lambda I(t) + (1-\lambda)\mathbf{C}(t)$ for some $\lambda \in [0,1]$. One can find values of λ such that these assignments give a probability distribution, since $\sum_t p_t$ equals $\sum_t \mathbf{I}(t) > 1$ if $\lambda = 1$, while $\lambda = 0$ yields that

$$\sum_t p_t = \sum_t \mathbf{C}(t) + \mathbf{I}(t^1) - \mathbf{C}(t^1) + \mathbf{I}(t^2) - \mathbf{C}(t^2) = 1 - (\mathbf{P}\{\kappa_{Y_\theta} > 1\} - \mathbf{I}(t_1) - \mathbf{I}(t_2) + \mathbf{C}(t_1) + \mathbf{C}(t_2)) < 1$$

by (5.5). Finally,

$$\mathbf{P}\{z \in \{t^1, t^2\}\} = p_{t^1} + p_{t^2} = \mathbf{I}(t^1) + \mathbf{I}(t^2)$$

while (5.4) yields that

$$\mathbf{T}(t^1, t^2) = \mathbf{P}\{\{t^1, t^2\} \cap Y_\theta \neq \emptyset\} = \mathbf{I}(t^1) + \mathbf{I}(t^2) - \mathbf{I}(t^1, t^2) < \mathbf{I}(t^1) + \mathbf{I}(t^2).$$

Thus, $\mathbf{P}\{z \in K\}$ is not dominated by $\mathbf{T}(K)$ for $K = \{t^1, t^2\}$, i.e. in this case it is also possible to construct a random element that satisfies (5.3), but fails to satisfy (4.5). ■

Theorem 5.4.

Proof. Observe that $\{\Theta_{IU_k}\}_{k \in \mathbb{N}}$, is a decreasing sequence in the set of non-empty compact subsets of $\Re^p$. Moreover,

$$\bigcap_{k=1}^{\infty} \Theta_{IU_k} = \bigcap_{k=1}^{\infty} \left\{\theta \in \Theta : u'\mathbf{P}\left(y | \underline{x}\right) \leq \mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right) | \underline{x}\right] \ \forall\ u \in U_k \ \underline{x} - a.s.\right\} = \Theta_I.$$

Hence, $\rho_H(\Theta_I, \Theta_{IU_k}) \to 0$, see, e.g., Schneider (1993, Lemma 1.8.1). ■

B Details for the Two Players Entry Game

Elements of the Aumann Expectation

For a given $q \in \text{Sel}(Q(S_\theta))$, the values of the coordinates of $\mathbf{E}(q)$ are:

$$\begin{aligned}
\mathbf{E}[(1-\sigma_1)(1-\sigma_2)] &= \mathbf{P}(\varepsilon_1 \le 0, \varepsilon_2 \le 0) \\
&\quad +\mathbf{E}\left[(1-\sigma_1)(1-\sigma_2)1\left(0 \le \varepsilon_1 \le -\theta_1, 0 \le \varepsilon_2 \le -\theta_2, \sigma_1 = \frac{\varepsilon_2}{-\theta_2}, \sigma_2 = \frac{\varepsilon_1}{-\theta_1}\right)\right], \\
\mathbf{E}[\sigma_1(1-\sigma_2)] &= \mathbf{P}((\varepsilon_1,\varepsilon_2) \in [-\theta_1, +\infty) \times (-\infty, -\theta_2] \cup [0, -\theta_1] \times (-\infty, 0]) \\
&\quad +\mathbf{P}(0 \le \varepsilon_1 \le -\theta_1, 0 \le \varepsilon_2 \le -\theta_2, \sigma_1 = 1, \sigma_2 = 0) \\
&\quad +\mathbf{E}\left[\sigma_1(1-\sigma_2)1\left(0 \le \varepsilon_1 \le -\theta_1, 0 \le \varepsilon_2 \le -\theta_2, \sigma_1 = \frac{\varepsilon_2}{-\theta_2}, \sigma_2 = \frac{\varepsilon_1}{-\theta_1}\right)\right], \\
\mathbf{E}[(1-\sigma_1)\sigma_2] &= \mathbf{P}((\varepsilon_1,\varepsilon_2) \in (-\infty, 0] \times [0, +\infty) \cup [0, -\theta_1] \times [-\theta_2, +\infty)) \\
&\quad +\mathbf{P}(0 \le \varepsilon_1 \le -\theta_1, 0 \le \varepsilon_2 \le -\theta_2, \sigma_1 = 0, \sigma_2 = 1) \\
&\quad +\mathbf{E}\left[(1-\sigma_1)\sigma_2 1\left(0 \le \varepsilon_1 \le -\theta_1, 0 \le \varepsilon_2 \le -\theta_2, \sigma_1 = \frac{\varepsilon_2}{-\theta_2}, \sigma_2 = \frac{\varepsilon_1}{-\theta_1}\right)\right] \\
\mathbf{E}[\sigma_1\sigma_2] &= \mathbf{P}(\varepsilon_1 \ge -\theta_1, \varepsilon_2 \ge -\theta_2) \\
&\quad +\mathbf{E}\left[\sigma_1\sigma_2 1\left(0 \le \varepsilon_1 \le -\theta_1, 0 \le \varepsilon_2 \le -\theta_2, \sigma_1 = \frac{\varepsilon_2}{-\theta_2}, \sigma_2 = \frac{\varepsilon_1}{-\theta_1}\right)\right]
\end{aligned}$$

where $1(\cdot)$ denotes the indicator function of the event in brackets.

Further Exemplification of the Sharpness Result

In order to further illustrate the sharpness result, we apply Berry and Tamer (2007) formulation to this game. Let

$$\begin{aligned}
U_t^\theta &= \{\varepsilon : t \text{ is the unique equilibrium outcome given } \theta\}, \\
M_D^\theta &= \{\varepsilon : D \text{ is the set of multiple equilibrium outcomes given } \theta\},
\end{aligned}$$

where $t \in \mathcal{Y}$ is an equilibrium outcome, and D is the set $\{(0,1),(1,0),(\sigma_1,\sigma_2)\}$, with $(0,1),(1,0)$ being pure strategy equilibria, and (σ_1,σ_2) being a mixed strategy equilibrium. Notice that in this case, $M_D^\theta = \{\varepsilon : \varepsilon \in [0,-\theta_1] \times [0,-\theta_2]\}$. If $\omega_1 : \varepsilon(\omega_1) \in [0,-\theta_1] \times [0,-\theta_2]$, let d be a random variable denoting which of the possible equilibria is selected in the region of multiplicity, with $d = 1$ if $(1,0)$ is selected, $d = 2$ if $(0,1)$ is selected, and $d = 3$ if the mixed strategy equilibrium is selected. Let $\psi(\varepsilon) = [\mathbf{P}(d = i|\,\varepsilon),\ i = 1,2,3]$ denote an admissible equilibrium selection mechanism in the region of multiplicity. As in Berry and Tamer, this equilibrium selection mechanism is left unspecified and can depend on market unobservables. Then for a given equilibrium selection mechanism

$\psi(\varepsilon)$, one has:

$$\begin{aligned}
\mathbf{P}[(0,0)|\,\theta,\psi] &= \mathbf{P}\left(\varepsilon \in U^{\theta}_{(0,0)}\right) + \int\limits_{M^{\theta}_{D}} \left(1-\frac{\varepsilon_2}{-\theta_2}\right)\left(1-\frac{\varepsilon_1}{-\theta_1}\right)\psi_3(\varepsilon)\,dF(\varepsilon),\\
\mathbf{P}[(1,0)|\,\theta,\psi] &= \mathbf{P}\left(\varepsilon \in U^{\theta}_{(1,0)}\right) + \int\limits_{M^{\theta}_{D}} \left[\psi_1(\varepsilon)+\frac{\varepsilon_2}{-\theta_2}\left(1-\frac{\varepsilon_1}{-\theta_1}\right)\psi_3(\varepsilon)\right]dF(\varepsilon),\\
\mathbf{P}[(0,1)|\,\theta,\psi] &= \mathbf{P}\left(\varepsilon \in U^{\theta}_{(0,1)}\right) + \int\limits_{M^{\theta}_{D}} \left[\psi_2(\varepsilon)+\left(1-\frac{\varepsilon_2}{-\theta_2}\right)\frac{\varepsilon_1}{-\theta_1}\psi_3(\varepsilon)\right]dF(\varepsilon),\\
\mathbf{P}[(1,1)|\,\theta,\psi] &= \mathbf{P}\left(\varepsilon \in U^{\theta}_{(1,1)}\right) + \int\limits_{M^{\theta}_{D}} \frac{\varepsilon_2}{-\theta_2}\frac{\varepsilon_1}{-\theta_1}\psi_3(\varepsilon)\,dF(\varepsilon),
\end{aligned}$$

where $\mathbf{P}\left(U^{\theta}_{(0,0)}\right) = \mathbf{P}(\varepsilon_1 \le 0, \varepsilon_2 \le 0)$ etc. Comparing these equations with the ones above defining $\mathbf{E}(q)$, $q \in \mathrm{Sel}(Q(S_\theta))$, one observes that the expressions are identical. In fact, each selection $q \in \mathrm{Sel}(Q(S_\theta))$ determines an admissible selection mechanism (observing that $q \in \mathrm{Sel}(Q(S_\theta))$ if and only if $\exists\, \sigma \in \mathrm{Sel}(S_\theta)$ such that $q = q(\sigma)$), and each admissible selection mechanism determines a selection $\sigma \in \mathrm{Sel}(S_\theta)$.

ABJ and CT Inequalities as a Special Case of the Support Function Inequalities

Finally, we show that the inequalities defining the set Θ_I^{CT} in Section 3.3 (and therefore those defining the set Θ_I^{ABJ}) are a subset of the inequalities defining Θ_I in equation (3.8). Using the information in Figure 1, we have that for $\omega \in \Omega_1$,

$$Q(S_{\boldsymbol{\theta}}(\omega)) = \begin{cases} \left\{[1\ 0\ 0\ 0]'\right\} & \text{if } \varepsilon(\omega) \in (-\infty,0]\times(-\infty,0],\\ \left\{[0\ 1\ 0\ 0]'\right\} & \text{if } \varepsilon(\omega) \in [-\theta_1,+\infty)\times(-\infty,-\theta_2]\cup[0,-\theta_1]\times(-\infty,0],\\ \left\{[0\ 0\ 1\ 0]'\right\} & \text{if } \varepsilon(\omega) \in (-\infty,0]\times[0,+\infty)\cup[0,-\theta_1]\times[-\theta_2,+\infty),\\ \left\{[0\ 0\ 0\ 1]'\right\} & \text{if } \varepsilon(\omega) \in [-\theta_1,+\infty)\times[-\theta_2,+\infty),\\ \left\{\left\{[0\ 1\ 0\ 0]'\right\},\{\delta(\omega)\}\left\{[0\ 0\ 1\ 0]'\right\}\right\} & \text{if } \varepsilon(\omega) \in [0,-\theta_1]\times[0,-\theta_2], \end{cases}$$

where

$$\delta(\omega) = \left[\left(1+\tfrac{\varepsilon_2(\omega)}{\theta_2}\right)\left(1+\tfrac{\varepsilon_1(\omega)}{\theta_1}\right), -\tfrac{\varepsilon_2(\omega)}{\theta_2}\left(1+\tfrac{\varepsilon_1(\omega)}{\theta_1}\right), -\left(1+\tfrac{\varepsilon_2(\omega)}{\theta_2}\right)\tfrac{\varepsilon_1(\omega)}{\theta_1}, \tfrac{\varepsilon_2(\omega)}{\theta_2}\tfrac{\varepsilon_1(\omega)}{\theta_1}\right]$$

Hence,

$$h(Q(S_\theta(\omega)),u) = \begin{cases} u_1 & \text{if } \varepsilon(\omega) \in (-\infty,0]\times(-\infty,0],\\ u_2 & \text{if } \varepsilon(\omega) \in [-\theta_1,+\infty)\times(-\infty,-\theta_2]\cup[0,-\theta_1]\times(-\infty,0],\\ u_3 & \text{if } \varepsilon(\omega) \in (-\infty,0]\times[0,+\infty)\cup[0,-\theta_1]\times[-\theta_2,+\infty),\\ u_4 & \text{if } \varepsilon(\omega) \in [-\theta_1,+\infty)\times[-\theta_2,+\infty),\\ \max\left(u_2,\delta(\omega)'u,u_3\right) & \text{if } \varepsilon(\omega) \in [0,-\theta_1]\times[0,-\theta_2]. \end{cases}$$

Take $u = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}'$. Then $\max\left(u_2,\delta(\omega)'u,u_3\right) = \delta(\omega)'u = \left(1-\frac{\varepsilon_2(\omega)}{-\theta_2}\right)\left(1-\frac{\varepsilon_1(\omega)}{-\theta_1}\right)$, and

$$\begin{aligned}
\mathbf{P}(y=(0,0)) &\le \mathbf{E}(h(Q(S_\theta),u))\\
&= \mathbf{P}(S_\theta = \{(0,0)\}) + \mathbf{E}\left(\left(1-\frac{\varepsilon_2}{-\theta_2}\right)\left(1-\frac{\varepsilon_1}{-\theta_1}\right)1(\varepsilon \in [0,-\theta_1]\times[0,-\theta_2])\right).
\end{aligned}$$

Take $u = \begin{bmatrix} -1 & 0 & 0 & 0 \end{bmatrix}'$. Then $\max\left(u_2, \delta(\omega)' u, u_3\right) = 0$, and

$$-\mathbf{P}(y = (0,0)) \leq \mathbf{E}(h(Q(S_\theta), u)) = -\mathbf{P}(S_\theta = \{(0,0)\}).$$

A similar argument applies to the remaining canonical basis vectors in $\Re^4$, and to these vectors multiplied by -1. This yields the desired result.

References

ANDREWS, D. W. K., S. T. BERRY, AND P. JIA (2004): "Confidence Regions for Parameters in Discrete Games with Multiple Equilibria, with an Application to Discount Chain Store Location," mimeo.

ANDREWS, D. W. K., AND P. GUGGENBERGER (2007): "Validity of Subsampling and Plug-in Asymptotic Inference for Parameters Defined by Moment Inequalities," working paper, Cowles Foundation, Yale University.

ANDREWS, D. W. K., AND G. SOARES (2007): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," Working Paper, Cowles Foundation, Yale University.

ARADILLAS-LOPEZ, A., AND E. TAMER (2008): "The Identification Power of Equilibrium in Simple Games," *Journal of Business and Economic Statistics*, forthcoming.

ARTSTEIN, Z. (1983): "Distributions of Random Sets and Random Selections," *Israel Journal of Mathematics*, 46(4), 313–324.

AUMANN, R. J. (1965): "Integrals of Set Valued Functions," *Journal of Mathematical Analysis and Applications*, 12, 1–12.

——— (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics*, 1, 67–96.

BAJARI, P., H. HONG, AND S. RYAN (2007): "Identification and Estimation of a Discrete Game of Complete Information," mimeo.

BERESTEANU, A., AND F. MOLINARI (2006): "Asymptotic Properties for a Class of Partially Identified Models," *CeMMAP working papers*, CWP10/06.

——— (2008): "Asymptotic Properties for a Class of Partially Identified Models," *Econometrica*, forthcoming.

BERRY, S. T. (1992): "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60(4), 889–917.

Berry, S. T., and E. Tamer (2007): "Identification in Models of Oligopoly Entry," in *Advances in Economics and Econometrics: Theory and Application*, vol. II, chap. 2, pp. 46–85. Cambridge University Press, Ninth World Congress.

Bjorn, P. A., and Q. H. Vuong (1985): "Simultaneous Equations Models for Dummy Endogenous Variables: A Game Theoretic Formulation with an Application to Labor Force Participation," CalTech DHSS Working Paper Number 557.

Bresnahan, T. F., and P. C. Reiss (1988): "Do Entry Conditions Vary Across Markets?," *Brookings Papers on Economic Activity*, pp. 833–871.

——— (1990): "Entry in Monopoly Markets," *Review of Economic Studies*, 57, 531–553.

——— (1991): "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99(5), 977–1009.

Calvó-Armengol, A. (2006): "The Set of Correlated Equilibria of 2 by 2 Games," mimeo.

Chernozhukov, V., H. Hong, and E. Tamer (2007): "Estimation and Confidence Regions for Parameter Sets In Econometric Models," *Econometrica*, 75, 1243–1284.

Ciliberto, F., and E. Tamer (2004): "Market Structure and Multiple Equilibria in Airline Markets," mimeo.

Galichon, A., and M. Henry (2006): "Inference in Incomplete Models," Working Paper, Columbia University.

Gardner, R. J., and P. Milanfar (2003): "Reconstruction of Convex Bodies from Brightness Functions," *Discrete and Computational Geometry*, 29, 279–303.

Haile, P., and E. Tamer (2003): "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, 111, 1–51.

Lovisolo, L., and E. A. B. DaSilva (2001): "Uniform Distribution of Points on a Hyper-Sphere with Applications to Vector Bit-Plane Encoding," *IEE Proceedings. Vision, Image, and Signal Processing*, 148, 187–193.

Manski, C. F. (1989): "Anatomy of the Selection Problem," *Journal of Human Resources*, 24, 343–360.

——— (2003): *Partial Identification of Probability Distributions*. Springer Verlag, New York.

Manski, C. F., and E. Tamer (2002): "Inference on Regressions with Interval Data on a Regressor or Outcome," *Econometrica*, 70, 519–546.

Mazzeo, M. (2002): "Product Choice and Oligopoly Market Structure," *RAND Journal of Economics*, 33(2), 221–242.

McFadden, D. (1989): "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 57, 995–1026.

McKelvey, R. D., and A. McLennan (1996): "Computation of Equilibria in Finite Games," in *Handbook of Computational Economics*, vol. 1, chap. 2, pp. 87–142. Elsevier Science.

Molchanov, I. S. (2005): *Theory of Random Sets*. Springer Verlag, London.

Norberg, T. (1992): "On the Existence of Ordered Couplings of Random Sets – with Applications," *Israel Journal of Mathematics*, 77, 241–264.

Pakes, A., and D. Pollard (1989): "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57, 1027–1057.

Pakes, A., J. Porter, K. Ho, and J. Ishii (2006): "Moment Inequalities and Their Application," mimeo.

Romano, J. P., and A. M. Shaikh (2006): "Inference for the Identified Set in Partially Identified Econometric Models," mimeo.

Rosen, A. (2006): "Confidence Sets for Partially Identified Parameters That Satisfy a Finite Number of Moment Inequalities," CeMMAP working papers, CWP25/06.

Schneider, R. (1993): *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge Univ. Press.

Tamer, E. (2003): "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria," *Review of Economic Studies*, 70, 147–165.

Wilson, R. (1971): "Computing Equilibria of N-Person Games," *SIAM Journal on Applied Mathematics*, 21, 80–87.

Yang, Z. (2008): "Correlated Equilibrium and the Estimation of Discrete Games of Complete Information," mimeo.

Table 1: Projections of Θ_I^{ABJ}, Θ_I^{CT} and Θ_I, reduction in bounds width (in parentheses), and area of the identification regions compared to ABJ. Two player entry game with mixed strategy Nash equilibrium as solution concept.

	True Values	Projections of:		
		Θ_I^{ABJ}	Θ_I^{CT}	Θ_I
θ_1	-1.15	$[-2.715, -0.485]$	$[-2.715, -0.585]$	$[-2.205, -0.605]$
			(4.5%)	(28.3%)
θ_2	-1.40	$[-2.785, -0.625]$	$[-2.785, -0.725]$	$[-2.245, -0.745]$
			(4.6%)	(30.6%)
Approximate Reduction in Total Area Compared to Θ_I^{ABJ}			(16.4%)	(56.5%)

Table 2: Projections of Θ_I^{ABJ}, Θ_I^{CT} and Θ_I, and reduction in bounds width compared to ABJ. Four player, two type entry game with pure strategy Nash equilibrium as solution concept.

	True Values	Projections of:		
		Θ_I^{ABJ}	Θ_I^{CT}	Θ_I
θ_{11}	-0.15	$[-0.154, -0.144]$	$[-0.153, -0.146]$	$[-0.152, -0.147]$
			(27%)	(54%)
θ_{21}	-0.20	$[-0.206, -0.195]$	$[-0.204, -0.197]$	$[-0.203, -0.198]$
			(27%)	(54%)
θ_{22}	-0.10	$[-0.106, -0.096]$	$[-0.104, -0.097]$	$[-0.103, -0.098]$
			(27%)	(54%)

Table 3: Projections of Θ_I^{ABJ}, Θ_I^{CT} and Θ_I, reduction in bounds width (in parentheses), and area of the identification regions compared to ABJ. Two player entry game with correlated equilibrium as solution concept.

	True Values	Projections of:		
		Θ_I^{ABJ}	Θ_I^{CT}	Θ_I
θ_1	-1.15	$[-4.475, -0.485]$	$[-4.475, -0.585]$	$[-4.125, -0.595]$
			(2.5%)	(11.5%)
θ_2	-1.40	$[-4.585, -0.625]$	$[-4.585, -0.725]$	$[-4.425, -0.735]$
			(2.4%)	(6.8%)
Approximate Reduction in Total Area Compared to Θ_I^{ABJ}			(7.9%)	(23.1%)

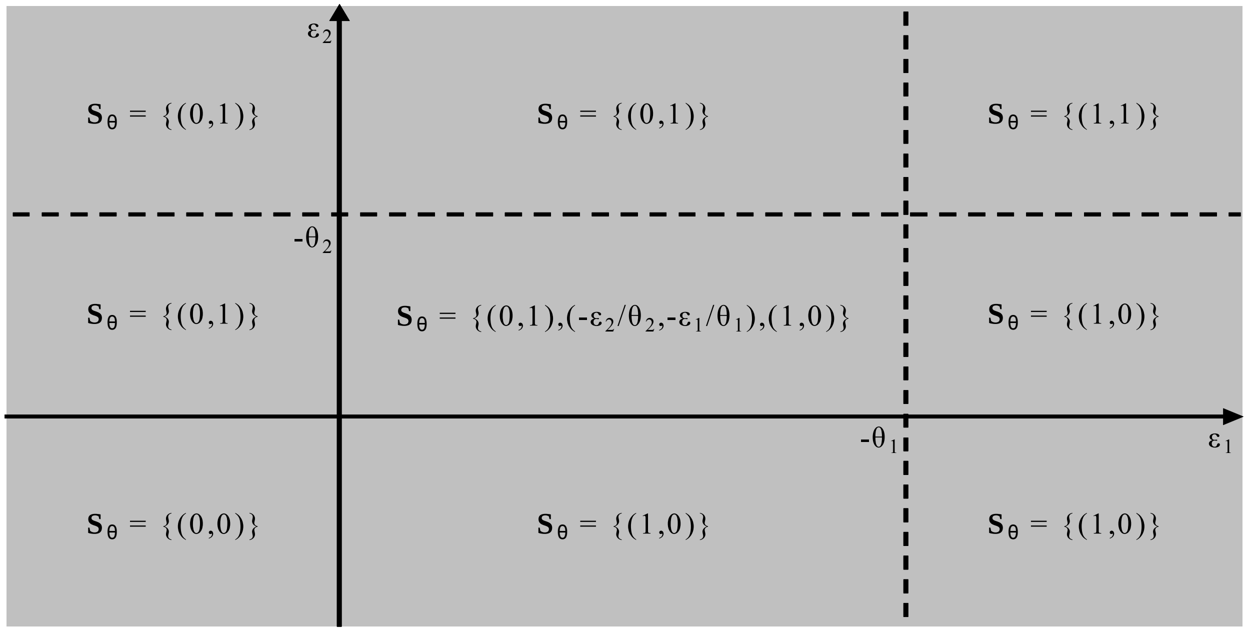

Figure 1: The random set of mixed strategy NE profiles as a function of $\varepsilon_1, \varepsilon_2$ in a two player entry game.

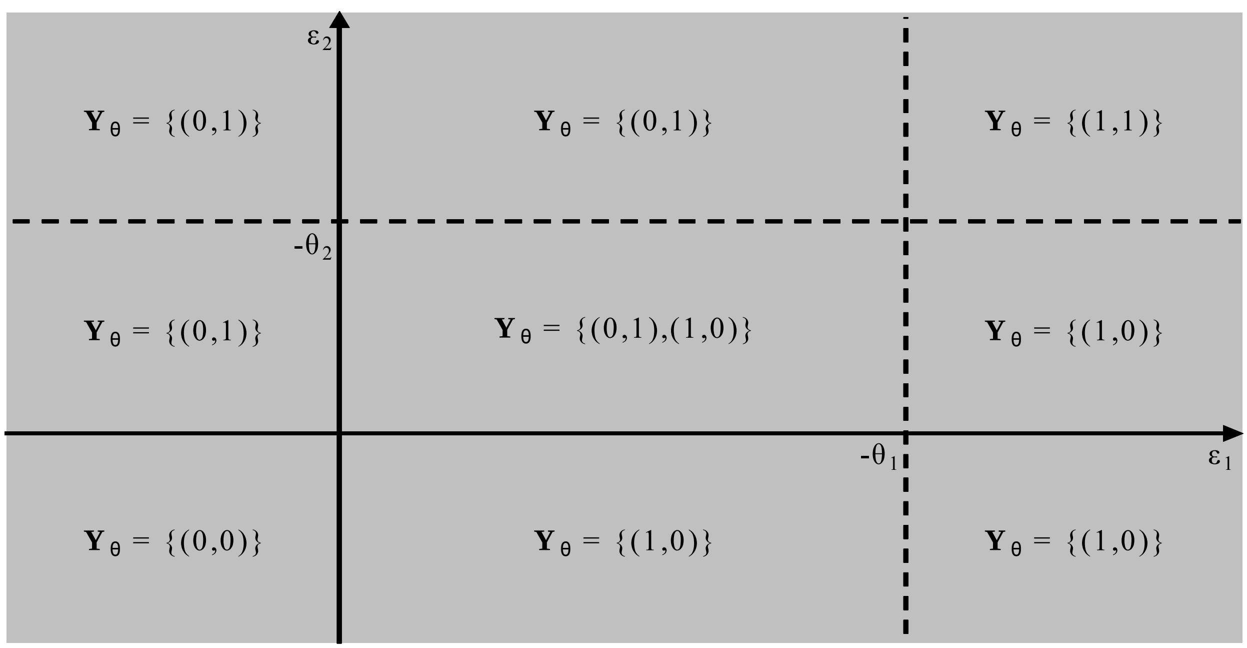

Figure 2: The random set of pure strategy NE outcomes as a function of $\varepsilon_1, \varepsilon_2$ in a two player entry game.

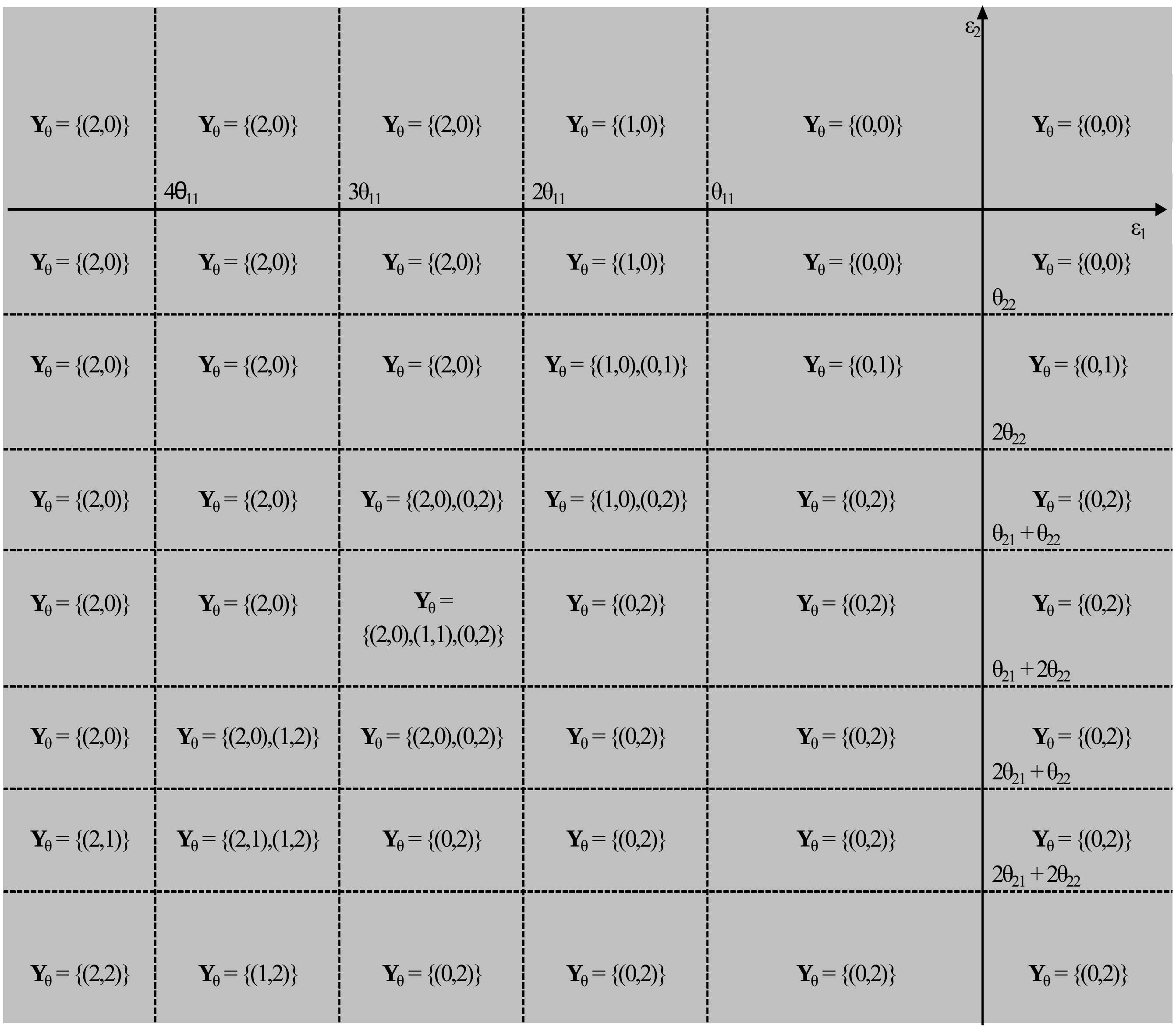

Figure 3: The random set of pure strategy NE outcomes as a function of $\varepsilon_1, \varepsilon_2$ in a four player, two type entry game.

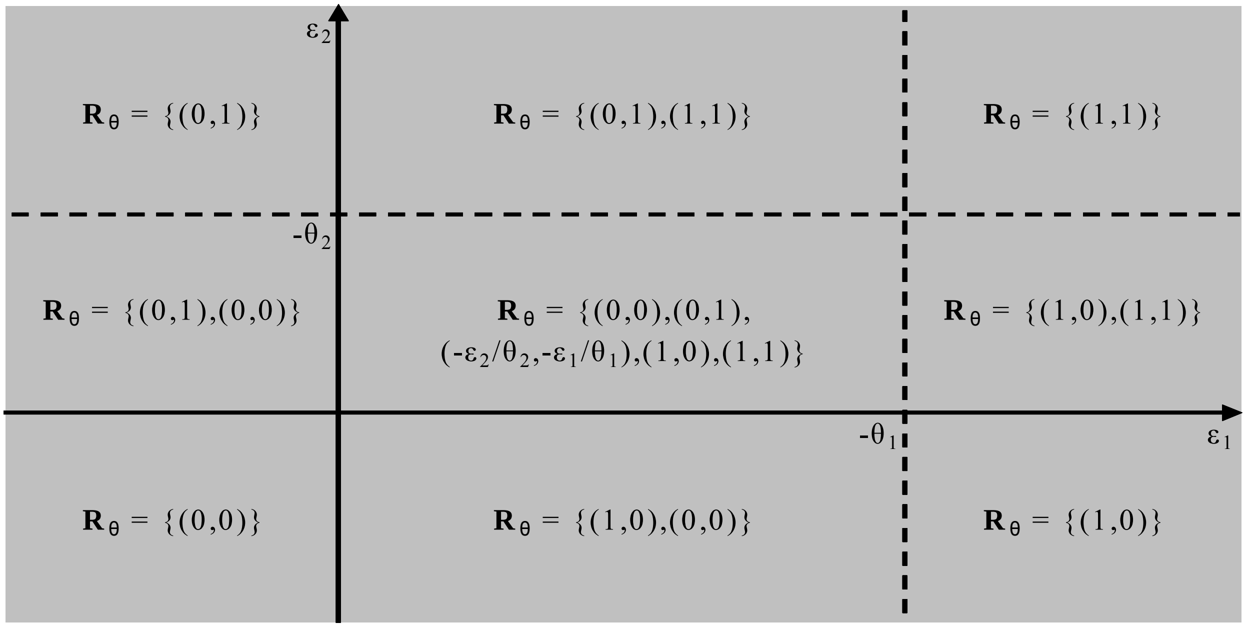

Figure 4: The random set of level-1 rational profiles as a function of $\varepsilon_1, \varepsilon_2$ in a two player entry game.

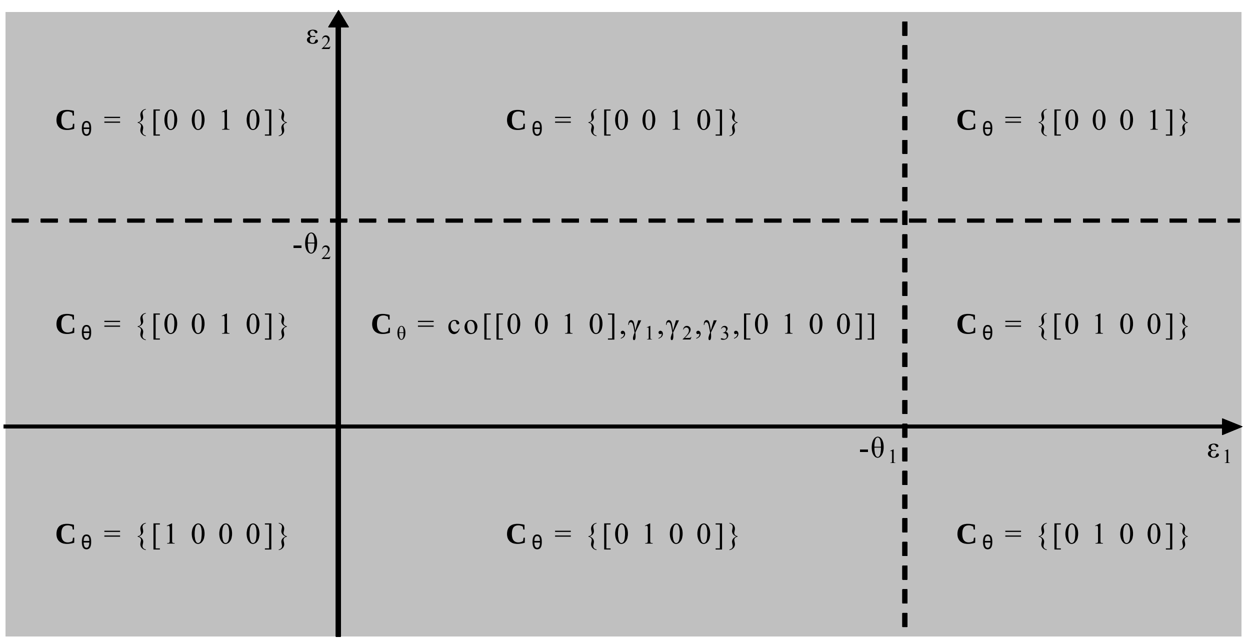

Figure 5: The random set of correlated equilibria for different values of ε_1 and ε_2 in a two player entry game. The correlated equilibria $\gamma_1, \gamma_2, \gamma_3$ are defined in Section 6.2.

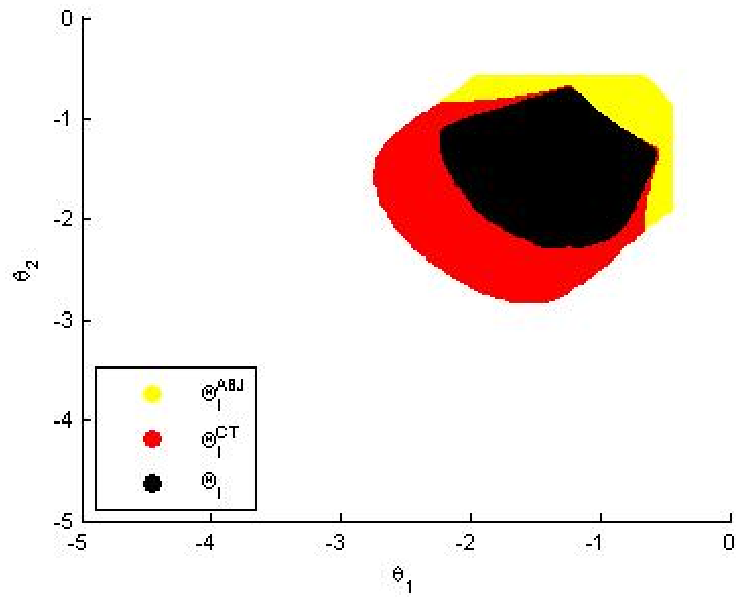

Figure 6: Identification regions in a two player entry game with mixed strategy Nash equilibrium as solution concept.

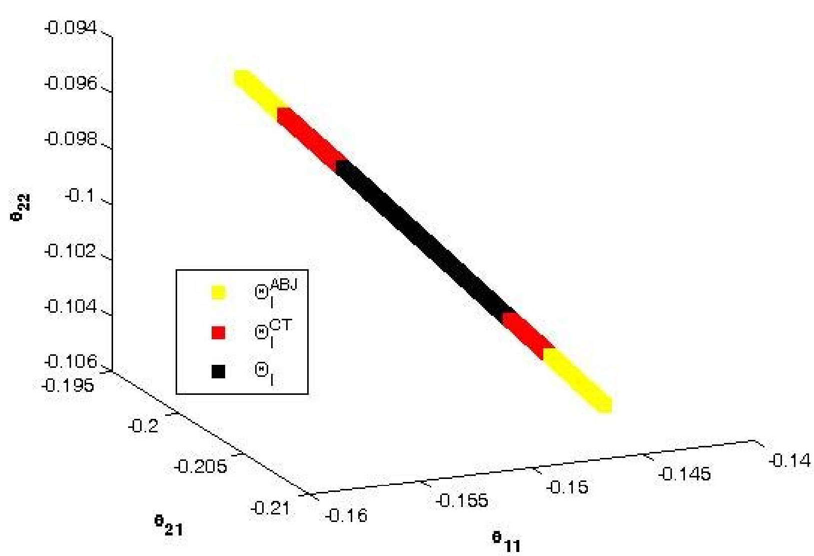

Figure 7: Identification regions in a four player, two type entry game with pure strategy Nash equilibrium as solution concept.

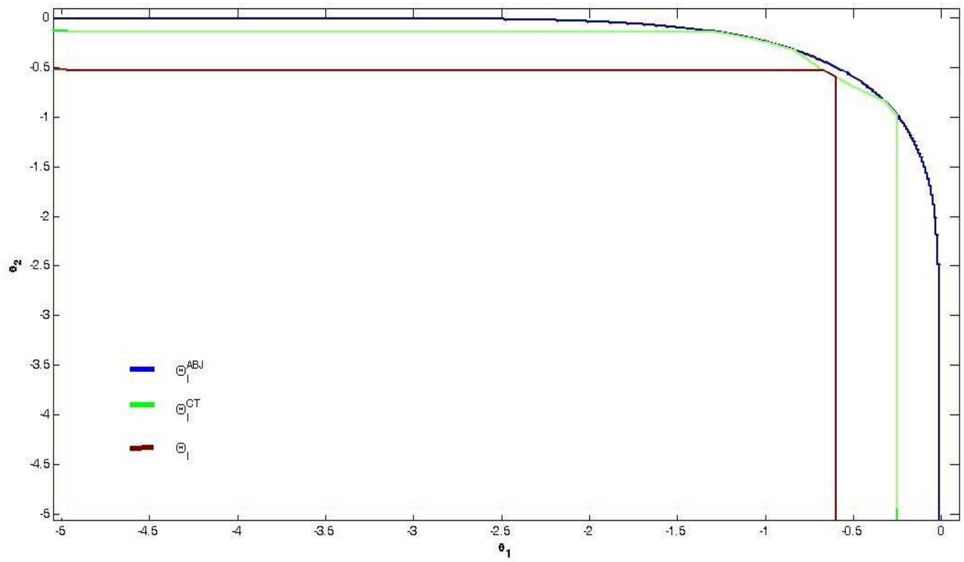

Figure 8: Upper contours of the identification regions in a two player entry game with level-1 rationality as solution concept.

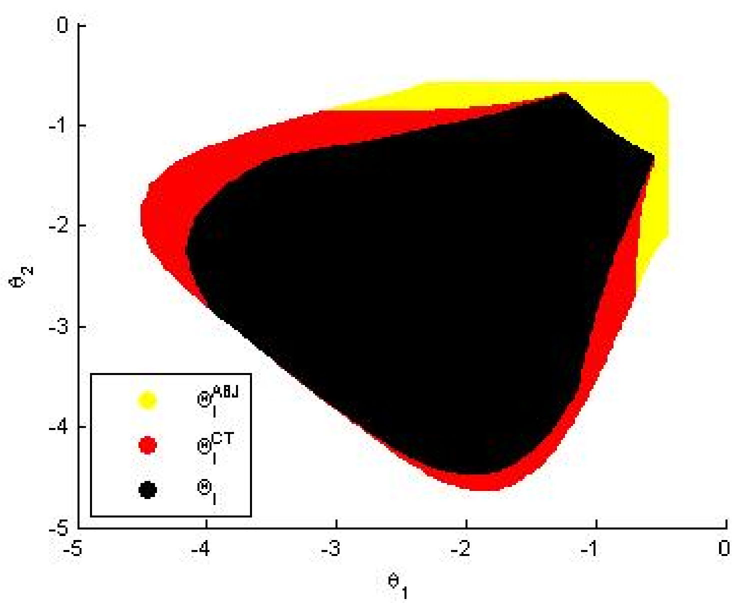

Figure 9: Identification regions in a two player entry game with correlated equilibrium as solution concept.