cemmap
centre for microdata methods and practice

Sharp identification regions in models with convex predictions: games, individual choice, and incomplete data

Arie Beresteanu
Ilya Molchanov
Francesca Molinari

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP27/09

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL
An ESRC Research Centre

Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE **Tel:** 44 (0)20 7291 4800 **Email:** info@cemmap.ac.uk **www.cemmap.ac.uk**

Sharp Identification Regions in Models with Convex Predictions: Games, Individual Choice, and Incomplete Data*

Arie Beresteanu† Ilya Molchanov‡ Francesca Molinari§

This draft: February 25, 2009

Abstract

We provide a tractable characterization of the sharp identification region of the parameters θ in a broad class of incomplete econometric models. Models in this class have set-valued predictions that yield a convex set of conditional or unconditional moments for the model variables. In short, we call these *models with convex predictions*. Examples include static, simultaneous move finite games of complete information in the presence of multiple mixed strategy Nash equilibria; random utility models of multinomial choice in the presence of interval regressors data; and best linear predictors with interval outcome and covariate data. Given a candidate value for θ, we establish that the convex set of moments yielded by the model predictions can be represented as the Aumann expectation of a properly defined random set. The sharp identification region of θ, denoted Θ_I, can then be obtained as the set of minimizers of the distance from a properly specified vector of moments of random variables to this Aumann expectation. We show that algorithms in convex programming can be exploited to efficiently verify whether a candidate θ is in Θ_I. We use examples analyzed in the literature to illustrate the gains in identification and computational tractability afforded by our method.

Keywords: Partial Identification, Random Sets, Aumann Expectation, Support Function, Normal Form Games, Multiple Equilibria, Random Utility Models, Interval Data, Best Linear Prediction.

---

*This paper supersedes an earlier working paper by the same authors, "Sharp Identification Regions in Games," Cemmap Working Paper CWP 15/08. We are grateful to Haim Bar, Talia Bar, Ivan Canay, Gary Chamberlain, David Easley, Paul Ellickson, Yossi Feinberg, Keisuke Hirano, Thierry Magnac, Ted O'Donoghue, Marcin Peski, Adam Rosen, Andrew Sweeting, and especially Larry Blume and Chuck Manski for comments. The results in this paper on finite games with multiple pure strategy Nash equilibria begun circulating on February 15, 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 at Arizona, BU, Cornell, Georgetown, Hebrew University of Jerusalem, Iowa State, Maryland, Michigan State, MIT-Harvard, Northwestern, NYU, Ohio State, Olin Business School, Stanford, Tel Aviv, UBC, UCL, UCLA, UT Austin, Wisconsin, the 2007 Triangle Econometrics Conference, the 2008 Cemmap-Northwestern University Conference "Inference in Partially Identified Models with Applications," the IIOC 6th annual meetings, the 2007 and 2008 NASM and the 2008 NAWM of the Econometric Society, the 2008 ERID conference at Duke, the 2008 CIREQ Conference "Inference with Incomplete Models," and the 2009 ICEEE Conference. 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 in October 2006.

†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 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 (Manski (2003)).

Our main contribution is to provide a simple, novel, and computationally feasible procedure to determine the *sharp* identification region of the parameters θ characterizing a broad class of incomplete econometric models. Models in this class have set-valued predictions which yield a convex set of conditional or unconditional moments for the model variables. In short, throughout the entirety of the paper, we call these *models with convex predictions*. Our use of the term "model" encompasses econometric frameworks ranging from structural parametric models, to nonparametric best predictors under square loss. In the interest of clarity of exposition, in this paper we focus on the parametric case. We consider incomplete structural parametric models, as well as parametric approximations to best predictors in the presence of incomplete data.

Structural parametric models with convex predictions can be described as follows. For a given value of the parameter vector θ and realization of the exogenous variables, the economic model predicts a set of values for the outcome variable of interest; these are the model (not necessarily convex) set-valued predictions. No restriction is placed on the manner in which a specific model predicted outcome is selected from this set. Hence, once the unobservable exogenous variables are "integrated out," the researcher obtains a convex set of conditional probability distributions for the outcome variable given regressors, rather than a single one. This is the convex set of conditional moments for the model variables, which we refer to as the *convex set of model predictions*. The identification problem arises because, apart from the regressors, the researcher only observes a single realized outcome, rather than the entire set of model predicted outcome values. When the economic model is correctly specified, the observed outcome is a realization of a random variable which follows a conditional distribution given regressors selected from the convex set of model predictions. Hence, one may find many values for the parameter vector θ which, when coupled with specific distributions selected from its associated convex set of model predictions, generate the same distribution of outcome and regressors as the one observed in the data. Specific examples of structural parametric models with convex predictions include: static, simultaneous move finite games of complete information in the presence of multiple mixed strategy Nash equilibria; and random utility models of multinomial choice in the presence of interval regressors data.

Nonparametric best predictors under square loss amount to conditional expectations of outcome variables given covariates. We consider the case that the researcher wishes to obtain a best linear approximation to the conditional expectation, that is, wishes to obtain the best linear predictor (BLP) of the outcome variable given the covariates, but these variables are only known to lie in observable random intervals.[1] When thinking about best linear prediction, no "model" is assumed in any substantive sense. However, with some abuse of terminology, for a given value of the BLP parameter vector θ, we refer to the set of prediction errors associated with each logically possible outcome and covariate variables in the observable random intervals, as the "model set-valued predictions." The relevant (unconditional) moments for each prediction error in this set are given by the prediction error's expectation, and its covariance with each of the covariates associated with that specific prediction error. The collection of such moments is convex, and we refer to this set as the "convex set of model predictions." If the data were complete this set would be a singleton, and under standard regularity conditions, there would be only one parameter vector θ determining a prediction error with expectation and covariance with each covariate equal to zero. However, due to the incompleteness of the data, one may find many values for the parameter vector θ which, when coupled with random variables in the observable intervals, yield prediction errors with these moments equal to zero.

Although previous literature has provided tractable characterizations of the sharp identification region for certain models with convex predictions (see, e.g. the analysis of nonparametric best predictors under square loss with interval outcome data), there exist many important problems, including the examples listed above, in which such a characterization is difficult to obtain. The analyses of Horowitz, Manski, Ponomareva, and Stoye (2003), Ciliberto and Tamer (2004), and Andrews, Berry, and Jia (2004) are examples of research studying the identified features of best linear predictors with missing outcome and covariates data, and finite games with multiple pure strategy equilibria, in which the regions of parameter values proposed are either infeasible to compute, or not sharp.

Establishing whether a conjectured region for the identified features of an incomplete econometric model is sharp is a key question in identification analysis. Given the joint distribution of the observed variables, a researcher asks herself what parameters θ are consistent with this distribution. The sharp identification region is the collection of parameter values that could generate the same distribution of observables as the one in the data, for some data generation process consistent with the maintained assumptions. Examples of sharp identification regions for parameters of incomplete models are given

[1] Beresteanu and Molinari (2006, 2008) provide a computationally tractable characterization of the sharp identification region for the BLP parameters in the case that only the outcome variable is interval measured. Here we significantly extend their identification results by allowing also for interval valued covariates.

in Manski (1989, 2003), Manski and Tamer (2002), and Molinari (2008), 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*. The inclusion in the outer regions of parameter values which are infeasible may weaken the researcher's ability to make useful predictions, and to test for model misspecification.

Using the theory of random sets (Molchanov (2005)), we provide a general methodology that allows us to characterize the sharp identification region for the parameters of models with convex predictions in a computationally tractable manner. Our main insight is that for a given candidate value of θ, the (conditional or unconditional) *Aumann expectation* of a properly defined θ-dependent *random closed set* coincides with the convex set of model predictions.[2] That is, this Aumann expectation gives the set, implied by the candidate θ, of moments for the relevant variables which are consistent with *all* the model's implications. This is a crucial advancement compared to the related literature, where researchers are often unable to fully exploit the information provided by the model that they are studying, and work with just a subset of model's implications. In turn, this advancement allows us to characterize the sharp identification region of θ, denoted Θ_I, through a simple necessary and sufficient condition. We explain this condition focusing on the case of structural models; a similar condition applies in the case of best linear prediction, see Section 5. Assume the model is correctly specified. Then θ is in Θ_I if and only if for all values of the covariates (except possibly a set of measure zero), the conditional Aumann expectation of the properly defined random set associated with θ contains the probability distribution of outcomes given covariates observed in the data. This is because when such condition is satisfied, there exists a probability distribution of outcomes given covariates associated with θ that is consistent with all the implications of the model, and coincides with the distribution observed in the data. The methodology that we propose allows us to verify this condition by determining whether a point (the distribution observed in the data) belongs to a θ-dependent convex set (the Aumann expectation). We show that this can be accomplished by minimizing a sublinear, hence convex, function over a convex set, and checking whether the resulting objective value is equal to zero. Computationally this is a very simple task, which can be carried out efficiently using algorithms in convex programming (e.g., Boyd and Vandenberghe (2004)).

It is natural to wonder which model with set-valued predictions may *not* belong to the class of models to which our methodology applies. If the model is not augmented with restrictions on the manner in which a specific outcome[3] is selected from the set-valued predictions, then these set-valued

[2] We formally define the notions of both random closed set and Aumann expectation in Section 3.
[3] Prediction error in the case of BLP with incomplete data.

predictions yield a convex set of moments for the model variables. This is because the unrestricted process of outcome selection convexifies the set-valued predictions of the model, yielding a convex set of moments. Therefore, our methodology applies. However, if restrictions are imposed on the selection process, non-convex sets of moments may result. We are chiefly interested in the case that no untestable assumptions are imposed on the selection process, and therefore exploring identification in models with non-convex predictions is beyond the scope of this paper.

There are no precedents to our general characterization of the sharp identification region of models with convex predictions. However, there is one precedent to the use of the Aumann expectation as a key tool to describe fundamental features of partially identified models. This is given by the work of Beresteanu and Molinari (2006, 2008), who were the first to illustrate the benefits of using elements of random sets theory to conduct identification analysis and statistical inference for incomplete econometric models in the space of sets, in a manner which is the exact analog of how these tasks are commonly performed for point identified models in the space of vectors.[4]

While our contribution lies in the identification analysis that we carry out, 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, so that the recent contributions of Chernozhukov, Hong, and Tamer (2004, 2007), Andrews and Guggenberger (2007), Andrews and Soares (2007), Canay (2008), Galichon and Henry (2006), Pakes, Porter, Ho, and Ishii (2006), Romano and Shaikh (2006), and Rosen (2008), among others, can be applied for estimation and statistical inference.

1.1 Overview for the Case of Finite Games with Multiple Equilibria

While our approach is general and applies to the entire class of models with convex predictions, in the interest of clarity of exposition we focus this section and the first part of the paper on identification analysis in static, simultaneous move finite games of complete information in the presence of multiple mixed strategy Nash equilibria. Our choice of finite games with multiple equilibria as the main example of our methodology is based on two considerations. First, the problem of identification for parameters characterizing these models has received a large amount of attention in the literature. See Section 2.2 for many references. Yet, unless untestable and often not credible assumptions are imposed to obtain point identification of the model's parameters, none of the methodologies developed in this vast literature provide a tractable characterization of Θ_I. Second, when players may randomize across their actions, a number of technical challenges arise for identification analysis. Certain features of our

[4] Beresteanu and Molinari (2006, 2008) study a class of partially identified models in which the sharp identification region of θ can itself be written as a transformation of the Aumann expectation of a properly defined random set.

approach allow us to overcome these challenges. Once these features are clearly explained, it becomes fairly simple to apply our methodology to other models with convex predictions. We illustrate this by analyzing random utility models of multinomial choice in the presence of interval regressors data, and best linear predictors with interval outcome and covariate data, arguably two of the most widely used tools in applied microeconomics.

We first explain why finite games with multiple equilibria are a special case of the broad class of models with convex predictions. Suppose that players follow Nash behavior and that their profits depend on their own actions, the actions of their opponents, payoff shifters $\underline{x}$ which are observable by the players and the econometrician, and payoff shifters ε that are unobservable by the econometrician, but observable by the players (the game is one of complete information). Parametrize the payoff functions of 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 $S_\theta(\underline{x}, \varepsilon)$. These equilibria are the model set-valued predictions. Each of the equilibria in $S_\theta(\underline{x}, \varepsilon)$ 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 collection of probability distributions over outcomes of the game conditional on $\underline{x}$ which are consistent with the model (i.e., with *all* its implications) is given by the Aumann expectation of $Q(S_\theta(\underline{x}, \varepsilon))$ conditional on $\underline{x}$, denoted $\mathbb{E}(Q(S_\theta(\underline{x}, \varepsilon))|\underline{x})$, which is a closed convex set. This is the convex set of model predictions. The researcher only observes an outcome of the game, denoted y, and payoff shifters $\underline{x}$. While data on y and $\underline{x}$ identify the distribution of y conditional on $\underline{x}$, denoted $\mathbf{P}(y|\underline{x})$, the observed outcome is determined by a mixed strategy equilibrium selected from the set of model predicted mixed strategy equilibria. Hence, it is just the realization of a random mixing draw from that equilibrium mixed strategy profile.

Framing the set of the model's predicted probability distributions in terms of an Aumann expectation is extremely useful for determining Θ_I. 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, Θ_I is given by the collection of θ's yielding a conditional Aumann expectation $\mathbb{E}(Q(S_\theta(\underline{x}, \varepsilon))|\underline{x})$ that contains $\mathbf{P}(y|\underline{x})$ for $\underline{x} - a.s.$ This is our fundamental identification result.

Given a candidate value for θ, one can verify whether it belongs to Θ_I by 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] The latter

[5] "The support function [of a nonempty closed convex set B in direction u] $h(B, u)$ is the signed distance of the support plane to B with exterior normal vector u from the origin; the distance is negative if and only if u points into the open half space containing the origin," Schneider (1993, page 37). See Rockafellar (1970, Chapter 13) or Schneider

can be evaluated exactly or approximated by simulation, depending on the complexity of the game. Showing that this dominance holds amounts to checking whether the difference between the support function of $\mathbb{E}\left(Q\left(S_\theta\left(\underline{x},\varepsilon\right)\right)\middle|\,\underline{x}\right)$ and the support function of $\mathbf{P}\left(y\middle|\,\underline{x}\right)$ in a direction given by a vector u attains a minimum of zero as u ranges in the unit ball of appropriate dimension. Because the support function of $\mathbb{E}\left(Q\left(S_\theta\left(\underline{x},\varepsilon\right)\right)\middle|\,\underline{x}\right)$ is a sublinear (hence convex) function and the support function of the vector $\mathbf{P}\left(y\middle|\,\underline{x}\right)$ is a linear function, this amounts to minimizing a sublinear function over a convex set, a task which can be carried out efficiently using algorithms in convex programming.

A further simplification is possible in the special case where one assumes that players do not randomize across their actions. We show that the support function of $\mathbf{P}\left(y\middle|\,\underline{x}\right)$ is dominated by the support function of $\mathbb{E}\left(Q\left(S_\theta\left(\underline{x},\varepsilon\right)\right)\middle|\,\underline{x}\right)$ as long as this dominance condition holds for a finite number of directions u. This is because when players are only allowed to play pure strategies, $\mathbb{E}\left(Q\left(S_\theta\left(\underline{x},\varepsilon\right)\right)\middle|\,\underline{x}\right)$ 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 ball. These directions are trivial to determine. While the number of inequalities to be checked in order to obtain Θ_I is finite, in some applications it may be quite large. However, we show that in many cases this number can be substantially reduced by exploiting basic notions of set algebra. Moreover, obtaining Θ_I by solving the minimization problem described above remains feasible.

1.2 Structure of the Paper

In Sections 2 and 3 we analyze in great detail the identification problem in static, simultaneous move finite games of complete information in the presence of multiple mixed strategy Nash equilibria. In Section 2 we introduce notation and assumptions, and present the identification problem. In order to clearly connect our work to the related literature, we present the definition of the sharp identification region provided by Berry and Tamer (2007), which however is infeasible to compute. In Section 3 we give our computationally feasible characterization of Θ_I. We first construct the convex set of model predictions (Section 3.1), and then use this set to obtain Θ_I in the case that players are allowed to use mixed strategies (Section 3.2). In Section 3.3 we compare our methodology to the earlier contributions of Andrews, Berry, and Jia (2004) and Ciliberto and Tamer (2004), and clarify why our approach leads to the sharp identification region, while theirs to outer regions. In Section 3.4 we specialize our results for the case that players are restricted to use pure strategies only. We illustrate the gains in identification afforded by our methodology through the simple examples of a two player entry game with mixed strategies (Section 3.3.1), and a four player, two type entry game with pure

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

strategies only (Section 3.4.1). In Section 3.5 we address the computational issues associated with our methodology.

In Section 4 we provide a computationally tractable characterization of the sharp identification region of the parameters of random utility models of multinomial choice in the presence of interval regressors data, by building on our treatment of finite games with multiple pure strategy Nash equilibria. In Section 5 we provide a computationally tractable characterization of the sharp identification region of the parameters of a best linear predictor with interval outcome and covariate data, by building on our treatment of finite games with multiple mixed strategy Nash equilibria. Section 6 concludes. While throughout Sections 2-3 we assume that players follow Nash behavior, our analysis of finite games with multiple equilibria does not depend on this assumption, but easily extends to other solution concepts for the game. In Appendix A 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.[6] Appendix B gives a dual representation of the sharp identification region in the special case that only pure strategies are played, and provides further insights on how to reduce the number of inequalities to be checked in order to compute it.[7]

2 Set-up of the Problem in Finite Games with Multiple Equilibria

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 non-empty sets $B, C \subset \Re^d$, we denote the directed Hausdorff distance from C to B, the Hausdorff distance

[6] 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.

[7] In particular, in Appendix B we show that in the very special case that players are restricted to playing pure strategies (and *only* in this case) our characterization of the sharp identification region based on the support function of $\mathbb{E}\left(Q\left(S_\theta\left(\underline{x},\varepsilon\right)\right)|\,\underline{x}\right)$ is dual to a characterization based on the *capacity functional* (i.e., the "probability distribution") of the random set of pure strategy equilibrium outcomes, by exploiting a result due to Artstein (1983). Galichon and Henry (2006) also use the notion of capacity functional of a properly defined random set and the results of Artstein (1983), to provide a specification test for partially identified structural models, thereby extending the Kolmogorov-Smirnov test of correct model specification to partially identified models. The analysis in Galichon and Henry (2006) does not treat the broad class of problems considered in this paper, nor does it relate the treatment of the identification problem in finite games with multiple pure strategy Nash equilibria based on the tools of random sets theory to the analysis in Berry and Tamer (2007), Ciliberto and Tamer (2004), and Andrews, Berry, and Jia (2004). For the very special case of finite games with multiple pure strategy Nash equilibria, Galichon and Henry (2008) address these questions. Their paper is subsequent to ours.

between C and B, and the Hausdorff norm of B, respectively, by

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

We focus on simultaneous-move games of complete information (normal form games) in which each player has a finite set of actions (pure strategies) $\mathcal{A}_j$, $j = 1, \ldots, J$, with J the number of players. 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 of interest, 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 outcomes of the game; this vector is observable by the econometrician. 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 potential outcomes of the game are finite. Each player has $\kappa_{\mathcal{A}_j} \geq 2$ 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 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 of the game. 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. These assumptions restrict attention to normal (strategic) form games. Part (iii) of the assumption requires that the outcome rule is known to the econometrician. Part (iv) of the assumption requires continuity of the payoff functions in x_j and ε_j.

This condition is 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 ones analyzed by Andrews, Berry, and Jia (2004, ABJ henceforth), Ciliberto and Tamer (2004, CT henceforth), Berry and Tamer (2007), and Bajari, Hong, and Ryan (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 3.4.1 with a simple example.

Assumption 3 *The econometrician observes data that identify $\mathbf{P}(y|\underline{x})$. The observed matrix of payoff shifters $\underline{x}$ is comprised of the non-redundant elements of x_j, $j = 1, \ldots, J$. The unobserved random vector $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_J)$ has 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 probability space $(\Omega, \mathfrak{F}, \mathbf{P})$.*

Assumption 3 requires that the researcher can 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.

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 unless more assumptions are added to the model, see for example Berry and Tamer (2007) for a thorough discussion of this problem. Bjorn and Vuong (1985), Bresnahan and Reiss (1988, 1990, 1991), Berry (1992), Mazzeo (2002), Tamer (2003), and Bajari, Hong, and Ryan (2007), among others, add restrictions to some of the models treated here to guarantee point identification of the payoff parameters.[8] Examples of

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

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 mechanism which, in the data generating process, determines the equilibrium played in the regions of multiplicity.

In the absence of additional assumptions of this kind, the payoff 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, is based on the concept of "selection mechanism." An admissible selection mechanism, in its most general definition, is a random vector whose entries almost surely have nonnegative realizations that sum up to one. The selection mechanism determines the probability with which each equilibrium, in the regions of the sample space where the model admits multiple equilibria, is played. 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 no assumptions are placed on it, the selection mechanism may represent an infinite dimensional nuisance parameter. One may consider the use of sieves to approximate such a nuisance parameter. While the sieve approach might be theoretically possible (though no results are available to establish its validity in this context), in practice it will result in extremely demanding (if not unfeasible) computational challenges. Given these difficulties, 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, see, e.g., Andrews, Berry, and Jia (2004) and Ciliberto and Tamer (2004). These moment inequalities have to hold for $\underline{x} - a.s.$

In Section 3 we show that contrary to what was indicated in the previous literature, the sharp identification region can in fact be characterized in a computationally feasible manner. We provide such a characterization, avoiding the need to deal with infinite dimensional nuisance parameters. Our approach does not impose any assumption on the selection mechanism, on the nature of competition, or on the form of heterogeneity across players. It does not require availability of covariates with large support or instruments, but fully exploits their identifying power if they are present.

Before moving on to the presentation of our results, for comparison purposes we give the (extension) of the abstract formulation of the sharp identification region provided by Berry and Tamer (2007, equation (2.21), page 67). We start with some additional notation. As in Section 1.1, 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 (3.1) below). Let $\psi(\cdot;\underline{x},\varepsilon) : S_\theta(\underline{x},\varepsilon) \to \Delta^{\kappa_{S_\theta(\underline{x},\varepsilon)}-1}$ denote a selection mechanism giving the probability that an equilibrium $\sigma \in S_\theta(\underline{x},\varepsilon)$ is selected, with $\Delta^{\kappa_{S_\theta(\underline{x},\varepsilon)}-1}$ denoting the unit simplex in the space of dimension $\kappa_{S_\theta(\underline{x},\varepsilon)}$. Observe that for this selection mechanism to be admissible it is required that $\psi(\sigma;\underline{x},\varepsilon) \geq 0$ for all $\sigma \in S_\theta(\underline{x},\varepsilon)$, and that $\sum_{\sigma\in S_\theta(\underline{x},\varepsilon)} \psi(\sigma;\underline{x},\varepsilon) = 1$. Under simple regularity conditions (e.g., sufficient assumptions on $\underline{x},\varepsilon$ to guarantee continuity of the distribution of the payoffs), 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 even after conditioning on market observables. 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:*

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

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

Let $\mathbf{P}(y|\,\underline{x};\theta,\psi)$ 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 $\mathbf{P}(y|\,\underline{x};\theta,\psi)$ matches the choice probabilities $\mathbf{P}(y|\,\underline{x})$ 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."

3 The Sharp Identification Region in Finite Games with Multiple Equilibria

This section develops the characterization of the sharp identification region which constitutes the main contribution of our paper in the analysis of finite games with multiple equilibria. For given $\theta \in \Theta$ and realization of $(\underline{x}, \varepsilon)$, in Section 3.1 we derive the set of (mixing) probability distributions over outcome profiles which are consistent with the modeling assumptions. Then we show how one can properly "integrate out" ε conditional on $\underline{x}$, hence obtaining the set of predicted choice probabilities consistent with the modeling assumptions, i.e. with *all* the implications of the model. This step is crucial, because it gives the set of probability distributions to which, if the model is correctly specified, $\mathbf{P}(y|\underline{x})$ belongs $\underline{x} - a.s.$ for each observationally equivalent $\theta \in \Theta$. In Section 3.2 we formalize this intuition, and give the characterization of the set of parameters $\theta \in \Theta$ which are consistent both with the modeling assumptions and the data – this is the sharp identification region of θ. In Section 3.3 we relate our methodology to the earlier contributions of ABJ and CT, and clarify the profound difference in our approaches. This difference leads to the sharp region in our case, and to outer regions in the cases of ABJ and CT. In Section 3.3.1 we illustrate the gains in identification afforded by our methodology through a simple two player entry game.

3.1 Construction of the Convex Set of Model Predictions

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

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

Hence, we define the following θ-dependent set:

$$(3.1) \qquad S_\theta(\underline{x}, \varepsilon) = \{\sigma \in \Delta(\mathcal{A}) : \pi_j(\sigma_j, \sigma_{-j}, x_j, \varepsilon_j, \theta) \geq \pi_j(\tilde{\sigma}_j, \sigma_{-j}, x_j, \varepsilon_j, \theta) \ \forall \tilde{\sigma}_j \in \Delta(\mathcal{A}_j) \ \forall j\}.$$

For a given value of θ and realization of $(\underline{x}, \varepsilon)$, this is the set of mixed strategy Nash equilibrium profiles.

Example 1 *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* $\theta_j < 0$, $j = 1, 2$. *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. Figure 1-(a) plots the set of mixed strategy equilibrium profiles* $S_\theta(\varepsilon)$

resulting from the possible realizations of $\varepsilon_1, \varepsilon_2$. Formally,

$$S_\theta(\varepsilon) = \begin{cases} \{(0,0)\} & \text{if } \varepsilon \in \mathcal{E}_\theta^{(0,0)} \equiv (-\infty, 0] \times (-\infty, 0], \\ \{(1,0)\} & \text{if } \varepsilon \in \mathcal{E}_\theta^{(1,0)} \equiv [-\theta_1, +\infty) \times (-\infty, -\theta_2] \cup [0, -\theta_1] \times (-\infty, 0], \\ \{(0,1)\} & \text{if } \varepsilon \in \mathcal{E}_\theta^{(0,1)} \equiv (-\infty, 0] \times [0, +\infty) \cup [0, -\theta_1] \times [-\theta_2, +\infty), \\ \{(1,1)\} & \text{if } \varepsilon \in \mathcal{E}_\theta^{(1,1)} \equiv [-\theta_1, +\infty) \times [-\theta_2, +\infty), \\ \left\{(0,1), \left(\frac{\varepsilon_2}{-\theta_2}, \frac{\varepsilon_1}{-\theta_1}\right), (1,0)\right\} & \text{if } \varepsilon \in \mathcal{E}_\theta^{M} \equiv [0, -\theta_1] \times [0, -\theta_2], \end{cases}$$

where in the above expressions $\mathcal{E}_\theta^{(\cdot,\cdot)}$ denotes a region of values for ε such that the game admits the pair in the superscript as a unique equilibrium, and $\mathcal{E}_\theta^M$ denotes the region of values for ε such that the game has multiple equilibria. In all of Section 3, whenever revisiting this example, we use the notations $\mathcal{E}_\theta^{(\cdot,\cdot)}$ and $\mathcal{E}_\theta^M$ without repeating their definition. $\square$

For ease of notation we write the set $S_\theta(\underline{x}, \varepsilon)$ and its realizations, respectively, as S_θ and $S_\theta(\omega) \equiv S_\theta(\underline{x}(\omega), \varepsilon(\omega))$, $\omega \in \Omega$, omitting 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.[9] 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 $\theta \in \Theta$ and $\omega \in \Omega$, each element $\sigma(\omega) \equiv (\sigma_1(\omega), \ldots, \sigma_J(\omega)) \in S_\theta(\omega)$ is one of the admissible mixed strategy Nash equilibrium profiles associated with the realizations $\underline{x}(\omega)$ and $\varepsilon(\omega)$, and it takes values in $\Delta(\mathcal{A}_1) \times \cdots \times \Delta(\mathcal{A}_J)$. The resulting random elements $\sigma = \{\sigma(\omega), \omega \in \Omega\}$ are the selections of S_θ:

[9] Continuity in x_j, ε_j follows from Assumption 1-(iv). Continuity in σ follows because by definition

$$\pi_j(\sigma, x_j, \varepsilon_j, \theta) \equiv \sum_{a \in \mathcal{A}} \left[\prod_{k=1}^{J} \sigma_k(a_k)\right] \pi_j(a, x_j, \varepsilon_j, \theta).$$

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* $\operatorname{Sel}(Z)$.

Example 1 (Cont.) *Consider again the simple two player entry game in Example 1, with the set S_θ plotted in Figure 1-(a). Let $\Omega^M = \{\omega \in \Omega : \varepsilon(\omega) \in \mathcal{E}_\theta^M\}$. Then for $\omega \notin \Omega^M$ the set S_θ has only one selection, since the equilibrium is unique. For $\omega \in \Omega^M$, S_θ contains a rich set of selections, which can be obtained as*

$$\sigma(\omega) = (\sigma_1(\omega), \sigma_2(\omega)) = \begin{cases} (1,0) & \text{if } \omega \in \Omega_1^M, \\ \left(\frac{\varepsilon_2(\omega)}{-\theta_2}, \frac{\varepsilon_1(\omega)}{-\theta_1}\right) & \text{if } \omega \in \Omega_2^M, \\ (0,1) & \text{if } \omega \in \Omega_3^M, \end{cases}$$

for all measurable $\Omega_i^M \subset \Omega^M$, $i = 1, 2, 3$, such that $\Omega_1^M \cup \Omega_2^M \cup \Omega_3^M = \Omega^M$. $\square$

By definition of a mixed strategy profile, for each $j = 1, \ldots, J$, $\sigma_j(\omega) : \mathcal{A}_j \to [0,1]$ assigns to each action $a_j \in \mathcal{A}_j$ a probability $\sigma_j(\omega, a_j) \geq 0$ that it is played, with $\sum_{a_j \in \mathcal{A}_j} \sigma_j(\omega, a_j) = 1$. 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 for a given parameter value $\theta \in \Theta$ and realization $\sigma(\omega)$, $\omega \in \Omega$, of a selection $\sigma \in \operatorname{Sel}(S_\theta)$, the implied probability that y is equal to $t^k \equiv (t_1^k, \ldots, t_J^k)$, $k = 1, \ldots, \kappa_\mathcal{Y}$, is given by $\prod_{j=1}^J \sigma_j\left(\omega, t_j^k\right)$. Hence, we can use a selection $\sigma \in \operatorname{Sel}(S_\theta)$ to define a random point $q(\sigma)$ whose realizations have coordinates

$$[q(\sigma(\omega))]_k = \prod_{j=1}^J \sigma_j\left(\omega, t_j^k\right), \quad k = 1, \ldots, \kappa_\mathcal{Y}. \tag{3.2}$$

The random point $q(\sigma)$ lies in a space of dimension equals to $\kappa_\mathcal{Y}$ and is such that for $\omega \in \Omega$, $[q(\sigma(\omega))]_k \geq 0$ for each $k = 1, \ldots, \kappa_\mathcal{Y}$ and $\sum_{k=1}^{\kappa_\mathcal{Y}} [q(\sigma(\omega))]_k = 1$. Hence, it is an element of the unit simplex in $\Re^{\kappa_\mathcal{Y}}$, denoted $\Delta^{\kappa_\mathcal{Y}-1}$. Because S_θ is a random closed set in $\Delta(\mathcal{A})$, the set resulting from repeating the above construction for each $\sigma \in \operatorname{Sel}(S_\theta)$ and given by

$$Q(S_\theta) = \{([q(\sigma)]_k, k = 1, \ldots, \kappa_\mathcal{Y}) : \sigma \in \operatorname{Sel}(S_\theta)\}, \tag{3.3}$$

is a closed random set in $\Delta^{\kappa_\mathcal{Y}-1}$

Example 1 (Cont.) *Consider again the simple two player entry game in Example 1, with the set S_θ plotted in Figure 1-(a). Index the set $\mathcal{Y}$ so that $\mathcal{Y} = \{(0,0), (1,0), (0,1), (1,1)\}$. Then*

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

Figure 1-(b) plots the set $Q(S_\theta)$ resulting from the possible realizations of $\varepsilon_1, \varepsilon_2$. $\square$

For given $\omega \in \Omega$, each vector $\left([q(\sigma(\omega))]_k, k = 1, \ldots, \kappa_{\mathcal{Y}}\right) \in Q(S_\theta(\omega))$ 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 $\sigma(\omega)$ when the realization of $(\underline{x}, \varepsilon)$ is $(\underline{x}(\omega), \varepsilon(\omega))$.

Notice that each $q(\sigma), \sigma \in \operatorname{Sel}(S_\theta)$, can be obtained by choosing an admissible selection mechanism $\psi(\cdot; \underline{x}, \varepsilon)$ which, for each $(\underline{x}, \varepsilon)$, picks the specific $\sigma \in S_\theta(\underline{x}, \varepsilon)$ with probability one, and constructing the implied probability distribution over action profiles. However, it is clear from the discussion in Section 2.2 that an admissible selection mechanism also gives *mixtures* of the probability distributions over action profiles associated with the mixed strategy Nash equilibria. All such mixtures can be obtained by considering the convex hull of the set $Q(S_\theta)$, denoted $\operatorname{co}[Q(S_\theta)]$. In fact, each realization of $\operatorname{co}[Q(S_\theta)]$ gives the entire set of probability distributions over action profiles associated with the corresponding realization of $(\underline{x}, \varepsilon)$ which are consistent with the maintained modeling assumptions. The task left is then to integrate such probability distributions against the distribution of ε given $\underline{x}$, that is, to "integrate out" ε for each selection of $\operatorname{co}[Q(S_\theta)]$.

Observe that every realization of $q \in \operatorname{Sel}(Q(S_\theta))$ is contained in $\Delta^{\kappa_{\mathcal{Y}}-1}$, and therefore $Q(S_\theta)$ is an integrably bounded random closed set, i.e., $\mathbf{E}(\|Q(S_\theta)\|_H) < \infty$, see Molchanov (2005, Definition 2.1.11). This implies that all its selections are integrable. Hence we can define the conditional Aumann expectation[10] of $Q(S_\theta)$ as

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

where the notation $\mathbb{E}(\cdot|\,\underline{x})$ denotes the conditional Aumann expectation of the random set in parentheses, while we reserve the notation $\mathbf{E}(\cdot|\,\underline{x})$ for the conditional expectation of a random vector. 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, and because the random set $Q(S_\theta)$ takes its realizations in a subset of the finite dimensional space $\Re^{\kappa_{\mathcal{Y}}}$, it follows from Theorem 2.1.15 and Theorem 2.1.24 of Molchanov (2005) that $\mathbb{E}(Q(S_\theta)|\,\underline{x})$ is a closed convex set for $\underline{x} - a.s.$, and $\mathbb{E}(Q(S_\theta)|\,\underline{x}) = \mathbb{E}(\operatorname{co}[Q(S_\theta)]|\,\underline{x})$.

Example 1 (Cont.) *Consider again the simple two player entry game in Example 1, with the set $Q(S_\theta)$ plotted in Figure 1-(b). Let $\Omega^M = \left\{\omega \in \Omega : \varepsilon(\omega) \in \mathcal{E}_\theta^M\right\}$. Suppose for simplicity that ε has a discrete distribution on $\mathcal{E}_\theta^M$, with $\varepsilon = e_1^\theta \equiv \left(-\frac{\theta_1}{3}, -\frac{\theta_2}{3}\right)$ with probability $\frac{1}{2}$, and $\varepsilon = e_2^\theta \equiv \left(-\frac{2\theta_1}{3}, -\frac{2\theta_2}{3}\right)$ with probability $\frac{1}{2}$. Then for $\omega \in \Omega^M$ such that $\varepsilon(\omega) = e_1^\theta$ the set $Q(S_\theta)$ contains three points: $[0 \;\; 1 \;\; 0 \;\; 0]'$, $\left[\frac{4}{9} \;\; \frac{2}{9} \;\; \frac{2}{9} \;\; \frac{1}{9}\right]'$, and $[0 \;\; 0 \;\; 1 \;\; 0]'$. Hence, conditional on $\varepsilon = e_1^\theta$, the expectations of the*

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

selections of $Q(S_\theta)$ are given by

$$\mathbf{E}\left(q|\,\Omega^M, \varepsilon = e_1^\theta\right) = \left[\tfrac{4}{9}p_2 \quad p_1 + \tfrac{2}{9}p_2 \quad \tfrac{2}{9}p_2 + p_3 \quad \tfrac{1}{9}p_2\right]',$$

where $p_i = \mathbf{P}\left(\Omega_i^M\middle|\,\Omega^M, \varepsilon = e_1^\theta\right)$, $i = 1,2,3$, for all measurable $\Omega_i^M \subset \{\omega \in \Omega^M : \varepsilon(\omega) = e_1^\theta\}$, $i = 1,2,3$, such that $\Omega_1^M \cup \Omega_2^M \cup \Omega_3^M = \{\omega \in \Omega^M : \varepsilon(\omega) = e_1^\theta\}$. Notice that 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 p_i, $i = 1,2,3$, fill in the whole two dimensional unit simplex. Hence, $\mathbb{E}\left(Q(S_\theta)|\,\Omega^M, \varepsilon = e_1^\theta\right)$ is a triangle in Δ^3 with extreme points $[0 \;\; 1 \;\; 0 \;\; 0]'$, $\left[\frac{4}{9} \;\; \frac{2}{9} \;\; \frac{2}{9} \;\; \frac{1}{9}\right]'$, and $[0 \;\; 0 \;\; 1 \;\; 0]'$. A similar result holds for $\mathbb{E}\left(Q(S_\theta)|\,\Omega^M, \varepsilon = e_2^\theta\right)$, so that $\mathbb{E}\left(Q(S_\theta)|\,\omega \in \Omega^M\right)$ is given by a Minkowski average of two triangles with weights $\frac{1}{2}$ each.[11] Hence,

$$\begin{aligned}\mathbb{E}\left(Q(S_\theta)|\,\omega \in \Omega^M\right) \;=\; co\Big\{&[0 \;\; 1 \;\; 0 \;\; 0]', \left[0 \;\; \tfrac{1}{2} \;\; \tfrac{1}{2} \;\; 0\right]', \left[\tfrac{4}{18} \;\; \tfrac{11}{18} \;\; \tfrac{2}{18} \;\; \tfrac{1}{18}\right]', \left[\tfrac{4}{18} \;\; \tfrac{2}{18} \;\; \tfrac{11}{18} \;\; \tfrac{1}{18}\right]', \\ &\left[\tfrac{1}{18} \;\; \tfrac{11}{18} \;\; \tfrac{2}{18} \;\; \tfrac{4}{18}\right]', \left[\tfrac{1}{18} \;\; \tfrac{2}{18} \;\; \tfrac{11}{18} \;\; \tfrac{4}{18}\right]', \left[\tfrac{5}{18} \;\; \tfrac{4}{18} \;\; \tfrac{4}{18} \;\; \tfrac{5}{18}\right]', [0 \;\; 0 \;\; 1 \;\; 0]'\Big\}\end{aligned}$$

is a three dimensional polytope in Δ^3. Given that for $\omega \notin \Omega^M$ the set $Q(S_\theta)$ is a singleton, a similar conclusion holds for $\mathbb{E}(Q(S_\theta))$. If the distribution of ε on $\mathcal{E}_\theta^M$ is continuous, one can show that $\mathbb{E}(Q(S_\theta))$ is a convex body in Δ^3 with infinitely many extreme points. □

The set $\mathbb{E}(Q(S_\theta)|\underline{x})$ collects vectors of probabilities with which each outcome of the game can be observed. It is a conditional Aumann expectation obtained by integrating the probability distribution over outcomes of the game implied by each mixed strategy equilibrium σ given $\underline{x}$ and ε, that is, by integrating each element of $\text{Sel}(Q(S_\theta))$, against the probability measure of $\varepsilon|\underline{x}$. We emphasize that in case of multiplicity, a different mixed strategy equilibrium $\sigma(\omega) \in S_\theta(\omega)$ may be selected (with different probability) for each ω. By construction, $\mathbb{E}(Q(S_\theta)|\underline{x})$ is the set of probability distributions over action profiles conditional on $\underline{x}$ which are consistent with the maintained modeling assumptions, i.e., with *all* the model's implications. Hence, it is the convex set of model predictions.

3.2 Characterization of the Sharp Identification Region

If the model is correctly specified, there exists at least one value of $\theta \in \Theta$ such that the observed 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 $\underline{x} - a.s.$[12] Hence, the set of observationally equivalent parameter values which form the sharp identification region is given by

$$\Theta_I = \{\theta \in \Theta : \mathbf{P}(y|\,\underline{x}) \in \mathbb{E}(Q(S_\theta)|\,\underline{x}) \;\; \underline{x} - a.s.\}, \tag{3.4}$$

[11] The *Minkowski sum* of two sets B and C in $\Re^d$ is given by $B \oplus C = \left\{r \in \Re^d : r = b + c,\ b \in B,\ c \in C\right\}$.
[12] 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 $\exists\, q \in \text{Sel}(Q(S_\theta)) : \mathbf{E}(q|\,\underline{x}) = \mathbf{P}(y|\,\underline{x})$.

where $\mathbf{P}(y|\underline{x}) \equiv \left[\mathbf{P}\left(y = t^k \middle| \underline{x}\right), k = 1, \ldots, \kappa_{\mathcal{Y}}\right]$. While this is the fundamental identification result in this paper, as written in equation (3.4) the set Θ_I may remain computationally challenging to obtain in certain cases. However, a dramatic simplification is possible if one uses a representation of the set Θ_I obtained through the notion of *support function* of a set.

Let the support function of a non-empty compact convex set $B \in \Re^{\kappa_{\mathcal{Y}}}$ be denoted $h(B, \cdot)$, with

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

It is well known (e.g., Rockafellar (1970, Chapter 13), Schneider (1993, Section 1.7)) that the support function of a non-empty compact convex set $B \in \Re^{\kappa_{\mathcal{Y}}}$ is a continuous convex sublinear function.[13] Standard arguments in convex analysis (e.g., Rockafellar (1970, Theorem 13.1)) give that $\mathbf{P}(y|\underline{x}) \in \mathbb{E}(Q(S_\theta)|\underline{x})$ if and only if

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

Theorem 2.1.47-(iv) in Molchanov (2005), a fundamental result in random sets theory, assures that $h(\mathbb{E}(Q(S_\theta)|\underline{x}), u) = \mathbf{E}[h(Q(S_\theta), u)|\underline{x}] \; \forall \, u \in \Re^{\kappa_{\mathcal{Y}}}$, and therefore

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

For our purposes, this further simplifies the problem, because $h(Q(S_\theta), u)$ is a continuous valued random variable whose expectation is simple to compute.

Example 1 (Cont.) *Consider again the simple two player entry game in Example 1, with the set S_θ plotted in Figure 1-(a) and the set $Q(S_\theta)$ plotted in Figure 1-(b). Pick a direction $u \equiv [u_1 \; u_2 \; u_3 \; u_4]' \in \Re^4$. Then, for $\omega \in \Omega$ such that $\varepsilon(\omega) \in \mathcal{E}_\theta^{(0,0)}$, we have $Q(S_\theta(\omega)) = \left\{[1 \;\; 0 \;\; 0 \;\; 0]'\right\}$, and $h(Q(S_\theta(\omega)), u) = u_1$. For $\omega \in \Omega$ such that $\varepsilon(\omega) \in \mathcal{E}_\theta^M$, we have $Q(S_\theta(\omega)) = \left\{[0 \;\; 1 \;\; 0 \;\; 0]', \; q\left(\frac{\varepsilon_2(\omega)}{-\theta_2}, \frac{\varepsilon_1(\omega)}{-\theta_1}\right), \; [0 \;\; 0 \;\; 1 \;\; 0]'\right\}$, and therefore $h(Q(S_\theta(\omega)), u) = \max\left(u_2, u'q\left(\frac{\varepsilon_2(\omega)}{-\theta_2}, \frac{\varepsilon_1(\omega)}{-\theta_1}\right), u_3\right)$, where*

$$q\left(\frac{\varepsilon_2(\omega)}{-\theta_2}, \frac{\varepsilon_1(\omega)}{-\theta_1}\right) = \left[\left(1 + \frac{\varepsilon_2(\omega)}{\theta_2}\right)\left(1 + \frac{\varepsilon_1(\omega)}{\theta_1}\right) \quad -\frac{\varepsilon_2(\omega)}{\theta_2}\left(1 + \frac{\varepsilon_1(\omega)}{\theta_1}\right) \quad -\left(1 + \frac{\varepsilon_2(\omega)}{\theta_2}\right)\frac{\varepsilon_1(\omega)}{\theta_1} \quad \frac{\varepsilon_2(\omega)}{\theta_2}\frac{\varepsilon_1(\omega)}{\theta_1}\right]'.$$

Figure 1-(c) plots the support function $h(Q(S_\theta(\omega)), u)$ resulting from the possible realizations of $\varepsilon_1, \varepsilon_2$. □

Because the support function is positively homogeneous, condition (3.5) is equivalent to

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

[13] In particular, $h(B, u+v) \leq h(B, u) + h(B, v)$ for all $u, v \in \Re^{\kappa_{\mathcal{Y}}}$ and $h(B, cu) = ch(B, u)$ for all $c > 0$ and for all $u \in \Re^{\kappa_{\mathcal{Y}}}$. Additionally, one can show that the support function of a bounded set $B \in \Re^d$ is Lipschitz with Lipschitz constant $\|B\|_H$, see Molchanov (2005, Theorem F.1).

where $\mathfrak{B} = \{u \in \Re^{\kappa_\mathcal{Y}} : \|u\| \leq 1\}$ denotes the unit ball in $\Re^{\kappa_\mathcal{Y}}$. By standard arguments, the inequality in condition (3.6) can be rewritten as

$$d_H\left(\mathbf{P}\left(y|\,\underline{x}\right), \mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)\right) = \max_{u \in \mathfrak{B}} \left(u'\mathbf{P}\left(y|\,\underline{x}\right) - \mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right)|\,\underline{x}\right]\right) = 0. \tag{3.7}$$

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*

$$\Theta_I = \left\{\theta \in \Theta : \min_{u \in \mathfrak{B}} \left(\mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right)|\,\underline{x}\right] - u'\mathbf{P}\left(y|\,\underline{x}\right)\right) = 0 \;\; \underline{x} - a.s.\right\} \tag{3.8}$$

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

Proof. In order to establish sharpness, it suffices to show that $\Theta_I = \Theta_I^\star$, with $\Theta_I^\star$ defined in equation (2.1). Take $\theta \in \Theta_I$. Then $\exists\ \tilde{\sigma} \in \mathrm{Sel}\left(S_\theta\right) : \mathbf{E}\left(q\left(\tilde{\sigma}\right)|\,\underline{x}\right) = \mathbf{P}\left(y|\,\underline{x}\right)$. Note that $\tilde{\sigma}$ specifies which equilibrium is chosen for each $\omega \in \Omega$, i.e. for each realization of $\left(\underline{x}, \varepsilon\right)$. Hence one can build a selection mechanism such that $\mathbf{P}\left(\psi\left(\tilde{\sigma}; \underline{x}, \varepsilon\right) = 1\right) = 1$, and this selection mechanism is admissible because for each realization of $\left(\underline{x}, \varepsilon\right)$ it selects the corresponding realization of $\tilde{\sigma}$ with probability 1. Hence $\theta \in \Theta_I^\star$. Conversely, take $\theta \in \Theta_I^\star$. Then for each realization of $\left(\underline{x}, \varepsilon\right)$ there exists an admissible selection mechanism $\psi\left(\cdot; \underline{x}, \varepsilon\right)$, which gives the probability of picking an equilibrium in each region of multiplicity associated with the specific realization of $\left(\underline{x}, \varepsilon\right)$, such that $\mathbf{P}\left(y|\,\underline{x}\right) = \mathbf{P}\left(y|\,\underline{x}; \theta, \psi\right)\ \underline{x} - a.s.$ Let $\tilde{q}$ denote the random variable which for each realization of $\left(\underline{x}, \varepsilon\right)$ takes on a realization equal to a mixture of the points in $Q\left(S_\theta\left(\underline{x}, \varepsilon\right)\right)$, with mixing coefficients ψ. Then $\tilde{q} \in \mathrm{Sel}\left(\mathrm{co}\left[Q\left(S_\theta\right)\right]\right)$ and by construction $\mathbf{P}\left(y|\,\underline{x}\right) = \mathbf{P}\left(y|\,\underline{x}; \theta, \psi\right) = \mathbf{E}\left(\tilde{q}|\,\underline{x}\right)\ \underline{x} - a.s.$ Hence $\theta \in \Theta_I$. The representation of Θ_I in equation (3.8) follows from equations (3.5)-(3.7). ■

The result in equation (3.8) gives a computationally very attractive characterization of the sharp identification region, because for each candidate $\theta \in \Theta$ it requires to minimize a sublinear, hence convex, function over a convex set, and check if the resulting objective value is equal to zero. This problem is computationally tractable and several efficient algorithms in convex programming are available to solve it, see for example Boyd and Vandenberghe (2004).[14]

Equation (3.8) yields a straightforward criterion function which is maximized by every parameter in the identification region:

$$w\left(\theta\right) = \int \min_{u \in \mathfrak{B}} \left(\mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right)|\,\underline{x}\right] - u'\mathbf{P}\left(y|\,\underline{x}\right)\right) dF_{\underline{x}} = \int -d_H\left(\mathbf{P}\left(y|\,\underline{x}\right), \mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)\right) dF_{\underline{x}}, \tag{3.9}$$

[14]While the support function of a convex set may not be differentiable at a countable number of points $u \in \mathfrak{B}$ (see, e.g., Schneider (1993, Theorem 1.5.2)), any support function can be approximated arbitrarily accurately by a support function of class C^∞. In particular, Schneider (1993, Theorem 3.3.1) gives a detailed convolution-based regularization process to implement this approximation. Hence, in principle one may use the gradient method to approximately solve the minimization problem in equation (3.8). However, given the recent developments in convex programming, this is not necessary, as convex programming algorithms are extremely efficient.

where $F_{\underline{x}}$ denotes the joint distribution of $\underline{x}$. Clearly, because the expression in parentheses vanishes for $u = 0 \in \mathfrak{B}$, $w(\theta) \leq 0$ for all $\theta \in \Theta$, and $w(\theta) = 0$ if and only if $\theta \in \Theta_I$.

3.3 Comparison with the Outer Regions of ABJ and CT

While ABJ and CT discuss only the case that players are restricted to use pure strategies, it is clear and explained in Berry and Tamer (2007, pp. 65-70) that their insights can be extended to the case that players are allowed to randomize over their strategies. Here we discuss the relationship between such extensions, and the methodology that we propose.

In the presence of multiple equilibria, ABJ observe that an implication of the model is that for a given $k = 1, ..., \kappa_{\mathcal{Y}}$, $\mathbf{P}\left(y = t^k \middle| \underline{x}\right)$ cannot be larger than the probability that t^k is *a possible* equilibrium outcome of the game. This is because for given $\theta \in \Theta$ and realization of $(\underline{x}, \varepsilon)$ such that t^k is a possible equilibrium outcome of the game, there can be another outcome $t^{k'} \in \mathcal{Y}$ which is also a possible equilibrium outcome of the game, and when both are possible t^k is selected only part of the time. While $\mathbf{P}\left(y = t^k \middle| \underline{x}\right)$ can be learned from the data, the probability that t^k is a possible equilibrium depends on the parameters of the model, and therefore these inequalities place restrictions on the values that the parameters can take. Using our notation, one can write the outer region proposed by ABJ as

(3.10)
$$\Theta_O^{ABJ} = \left\{\theta \in \Theta : \mathbf{P}\left(y = t^k \middle| \underline{x}\right) \leq \max\left(\int [q(\sigma)]_k \, dF(\varepsilon|\underline{x}) : \sigma \in \operatorname{Sel}(S_\theta)\right), \text{ for } k = 1, ..., \kappa_{\mathcal{Y}}, \ \underline{x} - a.s.\right\}.$$

In the above expression, $\mathbf{P}\left(y = t^k \middle| \underline{x}\right)$ is the probability that t^k is the equilibrium outcome of the game in the data. The expression $\max\left(\int [q(\sigma)]_k \, dF(\varepsilon|\underline{x}) : \sigma \in \operatorname{Sel}(S_\theta)\right)$ gives the probability that t^k is a possible equilibrium outcome of the game according to the model. It is obtained by selecting with probability one, in each region of multiplicity, the mixed strategy profile which yields the highest probability that t^k is the outcome of the game.

Example 1 (Cont.) *Consider again the simple two player entry game in Example 1, with the set S_θ plotted in Figure 1-(a). In this case, the expression for Θ_O^{ABJ} in equation (3.10) simplifies to*

$$\Theta_O^{ABJ} = \left\{\theta \in \Theta : \begin{array}{c} \mathbf{P}(y = (0,0)) \leq \mathbf{P}\left(\varepsilon \in \mathcal{E}_\theta^{(0,0)}\right) + \int_{\mathcal{E}_\theta^M} \left(1 + \frac{\varepsilon_2}{\theta_2}\right)\left(1 + \frac{\varepsilon_1}{\theta_1}\right) dF(\varepsilon) \\ \mathbf{P}(y = (1,0)) \leq \mathbf{P}\left(\varepsilon \in \mathcal{E}_\theta^{(1,0)}\right) + \mathbf{P}\left(\varepsilon \in \mathcal{E}_\theta^M\right) \\ \mathbf{P}(y = (0,1)) \leq \mathbf{P}\left(\varepsilon \in \mathcal{E}_\theta^{(0,1)}\right) + \mathbf{P}\left(\varepsilon \in \mathcal{E}_\theta^M\right) \\ \mathbf{P}(y = (1,1)) \leq \mathbf{P}\left(\varepsilon \in \mathcal{E}_\theta^{(1,1)}\right) + \int_{\mathcal{E}_\theta^M} \frac{\varepsilon_2}{\theta_2}\frac{\varepsilon_1}{\theta_1} dF(\varepsilon) \end{array}\right\}. \quad (3.11)$$

To see how this simplification occurs, take for example $t^2 = (1,0)$. Then we need to maximize the expectation of $[q(\sigma)]_2 = \sigma_1(1-\sigma_2)$ over all selections $\sigma \in \mathrm{Sel}(S_\theta)$. For $\varepsilon \in \mathcal{E}_\theta^{(1,0)}$, the only selection is $(1,0)$, and correspondingly $[q(\sigma)]_2 = 1$. For $\varepsilon \in \mathcal{E}_\theta^M$, the set S_θ consists of three points and $(1,0)$ is among them, so that the expectation of $[q(\sigma)]_2$ is maximized by the selection $\sigma = (1,0) \in \mathrm{Sel}(S_\theta)$ which yields $[q(\sigma)]_2 = 1$. For ε in any of the other regions, $[q(\sigma)]_2 = 0$. Thus, the maximum of the expectation of $[q(\sigma)]_2$ is the probability that $(1,0)$ is a selection of S_θ, that is

$$\begin{aligned}\max\left(\int [q(\sigma)]_2\, dF(\varepsilon) : \sigma \in \mathrm{Sel}(S_\theta)\right) &= \int_{\mathcal{E}_\theta^{(1,0)}} dF(\varepsilon|\underline{x}) + \max\left(\int_{\mathcal{E}_\theta^M} dF(\varepsilon), \int_{\mathcal{E}_\theta^M} -\frac{\varepsilon_2}{\theta_2}\left(1+\frac{\varepsilon_1}{\theta_1}\right) dF(\varepsilon), 0\right) \\ &= \mathbf{P}\left(\varepsilon \in \mathcal{E}_\theta^{(1,0)}\right) + \mathbf{P}\left(\varepsilon \in \mathcal{E}_\theta^M\right).\end{aligned}$$

This corresponds to the expression on the right hand side of the second inequality in (3.11). □

Comparing the expression for Θ_O^{ABJ} with that for Θ_I, one can see that Θ_O^{ABJ} can be obtained by applying inequality (3.5) for u equal to the canonical basis vectors in $\Re^{\kappa_\mathcal{Y}}$. In fact, take the vector $u^k \in \Re^{\kappa_\mathcal{Y}}$ to have all entries equal to zero except entry k which is equal to one, $k = 1, ..., \kappa_\mathcal{Y}$. Then

$$\mathbf{P}\left(y = t^k \middle| \underline{x}\right) = u^{k\prime}\mathbf{P}(y|\underline{x}) \le h\left(\mathbb{E}(Q(S_\theta)|\underline{x}), u^k\right) = \max\left(\mathbf{E}\left([q(\sigma)]_k|\underline{x}\right) : \sigma \in \mathrm{Sel}(S_\theta)\right).$$

Ciliberto and Tamer (2006) point out that additional information can be learned from the model. In particular, $\mathbf{P}\left(y = t^k \middle| \underline{x}\right)$ cannot be smaller than the probability that t^k is the *unique* equilibrium outcome of the game. This is because t^k is certainly realized whenever it is the only possible equilibrium outcome, but it can additionally be realized when it belongs to a set of multiple equilibrium outcomes. The probability that t^k is the unique equilibrium of the game also depends on the parameters of the model, and so the additional inequality further restricts the values that these parameters can take. Using our notation and applying the same logic as above, one can write the outer region proposed by CT as

$$\Theta_O^{CT} = \left\{\theta \in \Theta : \begin{array}{l} \min\left(\int [q(\sigma)]_k\, dF(\varepsilon|\underline{x}) : \sigma \in \mathrm{Sel}(S_\theta)\right) \le \mathbf{P}\left(y = t^k \middle| \underline{x}\right) \le \\ \max\left(\int [q(\sigma)]_k\, dF(\varepsilon|\underline{x}) : \sigma \in \mathrm{Sel}(S_\theta)\right), \text{ for } k = 1, ..., \kappa_\mathcal{Y},\ \underline{x} - a.s. \end{array}\right\}. \tag{3.12}$$

The expression $\min\left(\int [q(\sigma)]_k\, dF(\varepsilon|\underline{x}) : \sigma \in \mathrm{Sel}(S_\theta)\right)$ gives the probability that t^k is the unique equilibrium outcome of the game according to the model. It is obtained by selecting with probability one, in each region of multiplicity, the mixed strategy profile which yields the lowest probability that t^k is the outcome of the game.

Example 1 (Cont.) *Consider again the simple two player entry game in Example 1, with the set S_θ plotted in Figure 1-(a). In this case, the expression for Θ_O^{CT} in equation (3.12) simplifies to*

(3.13)

$$\Theta_O^{CT}=\left\{\theta\in\Theta:\begin{array}{c}\mathbf{P}\left(\varepsilon\in\mathcal{E}_\theta^{(0,0)}\right)\le\mathbf{P}\left(y=(0,0)\right)\le\mathbf{P}\left(\varepsilon\in\mathcal{E}_\theta^{(0,0)}\right)+\int_{\mathcal{E}_\theta^M}\left(1+\frac{\varepsilon_2}{\theta_2}\right)\left(1+\frac{\varepsilon_1}{\theta_1}\right)dF\left(\varepsilon\right)\\ \mathbf{P}\left(\varepsilon\in\mathcal{E}_\theta^{(1,0)}\right)\le\mathbf{P}\left(y=(1,0)\right)\le\mathbf{P}\left(\varepsilon\in\mathcal{E}_\theta^{(1,0)}\right)+\mathbf{P}\left(\varepsilon\in\mathcal{E}_\theta^{M}\right)\\ \mathbf{P}\left(\varepsilon\in\mathcal{E}_\theta^{(0,1)}\right)\le\mathbf{P}\left(y=(0,1)\right)\le\mathbf{P}\left(\varepsilon\in\mathcal{E}_\theta^{(0,1)}\right)+\mathbf{P}\left(\varepsilon\in\mathcal{E}_\theta^{M}\right)\\ \mathbf{P}\left(\varepsilon\in\mathcal{E}_\theta^{(1,1)}\right)\le\mathbf{P}\left(y=(1,1)\right)\le\mathbf{P}\left(\varepsilon\in\mathcal{E}_\theta^{(1,1)}\right)+\int_{\mathcal{E}_\theta^M}\frac{\varepsilon_2}{\theta_2}\frac{\varepsilon_1}{\theta_1}dF\left(\varepsilon\right)\end{array}\right\}.$$

To see how this simplification occurs, take for example $t^2=(1,0)$. Then the upper bound follows from the same reasoning as above. For the lower bound, we need to minimize the expectation of $[q(\sigma)]_2=\sigma_1(1-\sigma_2)$ over all selections $\sigma\in\mathrm{Sel}(S_\theta)$. For $\varepsilon\in\mathcal{E}_\theta^{(1,0)}$, the only selection is $(1,0)$, and correspondingly $[q(\sigma)]_2=1$. For $\varepsilon\in\mathcal{E}_\theta^M$, the set S_θ consists of three points and $(0,1)$ is among them, so that the expectation of $[q(\sigma)]_2$ is minimized by the selection $\sigma=(0,1)\in\mathrm{Sel}(S_\theta)$ which yields $[q(\sigma)]_2=0$. For ε in any of the other regions, $[q(\sigma)]_2=0$. Thus, the minimum of the expectation of $[q(\sigma)]_2$ is the probability that $(1,0)$ is the unique selection of S_θ, that is

$$\begin{aligned}\min\left(\int[q(\sigma)]_2\,dF(\varepsilon):\sigma\in\mathrm{Sel}(S_\theta)\right)&=\int_{\mathcal{E}_\theta^{(1,0)}}dF\left(\varepsilon|\underline{x}\right)+\min\left(\int_{\mathcal{E}_\theta^M}dF(\varepsilon),\int_{\mathcal{E}_\theta^M}-\frac{\varepsilon_2}{\theta_2}\left(1+\frac{\varepsilon_1}{\theta_1}\right)dF(\varepsilon),0\right)\\&=\mathbf{P}\left(\varepsilon\in\mathcal{E}_\theta^{(1,0)}\right).\end{aligned}$$

This corresponds to the expression on the left hand side of the second inequality in (3.13). □

Comparing the expression for Θ_O^{CT} with that for Θ_I, one can see that Θ_O^{CT} can be obtained by applying inequality (3.5) for u equal to the canonical basis vectors in $\Re^{\kappa_\mathcal{Y}}$ and each of these vectors multiplied by -1. The statement for the upper bound follows by the argument given above when considering Θ_O^{ABJ}. To verify the statement for the lower bound, take the vector $-u^k\in\Re^{\kappa_\mathcal{Y}}$ to have all entries equal to zero except entry k which is equal to minus one, $k=1,...,\kappa_\mathcal{Y}$. Then[15]

$$\begin{aligned}-\mathbf{P}\left(y=t^k\middle|\underline{x}\right)&=-u^{k\prime}\mathbf{P}\left(y|\underline{x}\right)\\&\le h\left(\mathbb{E}\left(Q(S_\theta)|\underline{x}\right),\left(-u^k\right)\right)=h\left(-\mathbb{E}\left(Q(S_\theta)|\underline{x}\right),u^k\right)\\&=-\min\left(\int[q(\sigma)]_k\,dF\left(\varepsilon|\underline{x}\right):\sigma\in\mathrm{Sel}(S_\theta)\right).\end{aligned}$$

[15] Equivalently, taking u to be a vector with each entry equal to 1, except entry k which is set to 0, one has that

$$\begin{aligned}1-\mathbf{P}\left(y=t^k\middle|\underline{x}\right)&=u'\mathbf{P}\left(y|\underline{x}\right)\le h\left(\mathbb{E}\left(Q(S_\theta)|\underline{x}\right),u\right)=\max\left(\sum_{i\neq k}\int[q(\sigma)]_i\,dF\left(\varepsilon|\underline{x}\right):\sigma\in\mathrm{Sel}(S_\theta)\right)\\&=\max\left(1-\int[q(\sigma)]_k\,dF\left(\varepsilon|\underline{x}\right):\sigma\in\mathrm{Sel}(S_\theta)\right)=1-\min\left(\int[q(\sigma)]_k\,dF\left(\varepsilon|\underline{x}\right):\sigma\in\mathrm{Sel}(S_\theta)\right).\end{aligned}$$

Hence, the approaches of ABJ and CT can be interpreted on the base of our analysis as follows. For each $\theta \in \Theta$, ABJ's inequalities give the smallest hypercube with sides that include the positive part of the axes, which contains $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$. The outer region Θ_O^{ABJ} is the collection of θ's such that $\mathbf{P}\left(y|\,\underline{x}\right)$ is contained in such hypercube $\underline{x}-a.s.$ CT use a more refined approach, and for each $\theta \in \Theta$ their inequalities give the smallest hypercube containing $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$. The outer region Θ_O^{CT} is the collection of θ's such that $\mathbf{P}\left(y|\,\underline{x}\right)$ is contained in such hypercube $\underline{x}-a.s.$ The more $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$ differs from the hypercubes used by ABJ and CT, the more likely it is that a candidate value θ belongs to Θ_O^{ABJ} and Θ_O^{CT}, but not to Θ_I. In order to provide a graphical intuition for this relationship, Figure 2 plots a projection into $\Re^2$ of: $\Delta^{\kappa_\mathcal{Y}-1}$; $\mathbf{P}\left(y|\,\underline{x}\right)$ – given by the white dot; $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$ – given by the black ellipsoid; the hypercube used by CT – given by the red square, on which is superimposed $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$; and the hypercube used by ABJ – given by the yellow square, on which are superimposed the hypercube used by CT and $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$. The intersection of ABJ's hypercube with $\Delta^{\kappa_\mathcal{Y}-1}$ gives the collection of probability distributions over outcome profiles (choice probabilities) consistent with the subset of model implications used by ABJ, namely the necessary conditions for each single outcome of the game to be a possible Nash equilibrium outcome. The intersection of CT's hypercube with $\Delta^{\kappa_\mathcal{Y}-1}$ gives the collection of probability distributions over outcome profiles consistent with the subset of model implications used by CT, namely the necessary conditions for each single outcome of the game to be a possible Nash equilibrium outcome, and the sufficient conditions for each single outcome of the game to be the unique Nash equilibrium outcome. The closed convex set $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$ gives the collection of choice probabilities consistent with *all* implications of the model. Hence, Θ_O^{ABJ} and Θ_O^{CT} are not sharp because they are based on checking whether the observed distribution of outcome profiles belongs to a set of choice probabilities which depends on θ, but is not consistent with all the model's implications.

3.3.1 A Simple Implementation in the Two Player Entry Game

This section provides a simple implementation of our method, and a numerical illustration of the identification gains that it affords, in the two player entry game in Example 1. The set S_θ for this example is plotted in Figure 1-(a). We generate data with $(\varepsilon_1, \varepsilon_2) \overset{iid}{\sim} N(0,1)$, $\theta_1^\star = -1.15$, $\theta_2^\star = -1.4$, and using a selection mechanism which, for each $\omega : \varepsilon(\omega) \in \mathcal{E}_{\theta^\star}^M$, picks each of outcome $(0,0)$ and $(1,1)$ for 10% of the cases and each of outcome $(1,0)$ and $(0,1)$ for 40% of the cases. Hence, the observed distribution is $\mathbf{P}(y) = [0.26572\ 0.34315\ 0.36531\ 0.02582]'$. The parameter space is assumed to be $\Theta = [-5, 0]^2$. Figure 7 and Table 1 report Θ_I, Θ_O^{CT}, and Θ_O^{ABJ}. In the figure, Θ_O^{ABJ} is given by the union of the yellow, red, and black areas, and Θ_O^{CT} by the union of the red and black areas. Θ_I is the black region. Our results clearly show that Θ_I is substantially smaller than Θ_O^{CT} and Θ_O^{ABJ}: Θ_I

has an area which is 43.5% of the area of Θ_O^{ABJ}, and 52% of that of Θ_O^{CT}.

The set Θ_I is calculated using an extremely simple application of the Nelder-Mead (NM) algorithm. To initialize the algorithm, we chose the four canonical basis vectors and the zero vector in $\Re^4$ as the five starting points (since for points $\theta \in \Theta_I$ the minimum of the objective function is achieved at the origin, choosing the vector zero as one of the starting points may reduce the number of iterations of the NM algorithm). For each candidate $\theta \in \Theta$, the value function was computed for the five starting points of the NM algorithm, and in each iteration of the algorithm the point yielding the worst (i.e., the biggest) value function was replaced with a new point following the NM procedure. The search for the minimum was stopped either if the algorithm encountered a negative value for the objective function, which guarantees that $\theta \notin \Theta_I$ without the need to reach full convergence, or if the difference between the value function for the worse point and for the best point fell below a certain tolerance value. We computed the set Θ_I both using this minimization procedure, and checking inequality $u'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}[h(Q(S_\theta), u)|\underline{x}]$ for each u in a fine grid of the unit sphere, as suggested in Beresteanu, Molchanov, and Molinari (2008, Section 5.2). The minimization procedure was an order of magnitude faster, while producing the exact same set Θ_I. The computation of the set Θ_I based on equation (3.8) took 15 minutes when the program was written in Fortran 90 and was compiled and ran on a Unix machine with a single processor of 3.2 GHz. The run time for each candidate θ varied widely depending on whether such point was in Θ_I, in which case on average it took about 0.0038 seconds to verify the condition in equation (3.8), or outside of Θ_I, in which case it took about 0.00047 seconds. Superior results are expected if parallel processing and more advanced platforms are used, and if efficient convex programming algorithms are employed.

3.4 Pure Strategies Only: A Further Simplification

We now assume that players in each market do not randomize across their actions. In a finite game, when restricting attention to pure strategies, one necessarily contends with the issue of possible non-existence of an equilibrium for certain parameter values $\theta \in \Theta$ and realizations of $(\underline{x}, \varepsilon)$. To deal with this problem, one can impose Assumption 4 below:

Assumption 4 *One of the following holds:*
(i) 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 $(\underline{x}, \varepsilon) - a.s.$
(ii) For each $\theta \in \Theta$ and realizations of $\underline{x}, \varepsilon$ such that a pure strategy Nash equilibrium does not exist, $S_\theta(\underline{x}, \varepsilon) = \mathrm{vert}(\Delta(\mathcal{A}))$, with $\mathrm{vert}(\cdot)$ the vertices of the set in parenthesis.

Assumption 4-(i) requires an equilibrium always to exist for the values of θ that have generated the observed outcomes y. If the model is correctly specified and players in fact follow pure strategy Nash behavior, then this assumption is satisfied. However, the assumption implicitly imposes strong restrictions on the parameter vector θ, the payoff functions, and the payoff shifters $\underline{x}, \varepsilon$. On the other hand, Assumption 4-(ii) posits that if the model does not have an equilibrium for a given $\theta \in \Theta$ and realization of $(\underline{x}, \varepsilon)$, then the model has no prediction on what should be the action taken by the players, and "anything can happen." In this respect, one may argue that Assumption 4-(ii) is more conservative than Assumption 4-(i). We do not take a stand here on which solution to the existence problem the applied researcher should follow. Either way, the approach that we propose delivers the sharp identification region Θ_I, although the set Θ_I will differ depending on whether Assumption 4-(i) or 4-(ii) is imposed. Moreover, one may choose not to impose Assumption 4 at all, and use a different solution concept. In that case as well, as we illustrate in Appendix A, our approach can be extended to deliver the sharp identification region.

When players play only pure strategies, the set S_θ takes its realizations as subsets of the vertices of $\Delta(\mathcal{A})$, because each pure strategy Nash equilibrium is equivalent to a degenerate mixed strategy Nash equilibrium placing probability one on a specific pure strategy profile. Hence, the realizations of the set $Q(S_\theta)$ lie in the subsets of the vertices of $\Delta^{\kappa_\mathcal{Y}-1}$.

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* $\theta_j < 0$, $j = 1, 2$. *Assume that players do not randomize across their actions, so that each* σ_j, $j = 1, 2$, *can take only values* 0 *and* 1. *Figure 3 plots the set* S_θ *resulting from the possible realizations of* $\varepsilon_1, \varepsilon_2$. *In this case,* S_θ *assumes only five values:* $\{(0,0)\}$ *for* $\varepsilon \in \mathcal{E}_\theta^{(0,0)}$, $\{(1,0)\}$ *for* $\varepsilon \in \mathcal{E}_\theta^{(1,0)}$, $\{(0,1)\}$ *for* $\varepsilon \in \mathcal{E}_\theta^{(0,1)}$, $\{(1,1)\}$ *for* $\varepsilon \in \mathcal{E}_\theta^{(1,1)}$, *and* $\{(0,1),(1,0)\}$ *for* $\varepsilon \in \mathcal{E}_\theta^{M}$. *Consequently, also the set* $Q(S_\theta)$ *assumes only five values, equal respectively to* $\{[1\ \ 0\ \ 0\ \ 0]'\}$, $\{[0\ \ 1\ \ 0\ \ 0]'\}$, $\{[0\ \ 0\ \ 1\ \ 0]'\}$, $\{[0\ \ 0\ \ 0\ \ 1]'\}$, *and* $\{[0\ \ 1\ \ 0\ \ 0]', [0\ \ 0\ \ 1\ \ 0]'\}$. $\square$

Hence, the sets S_θ and $Q(S_\theta)$ are "simple" random closed sets in $\Delta(\mathcal{A})$ and $\Delta^{\kappa_\mathcal{Y}-1}$, respectively.

Definition 4 *Denoting by* $\mathcal{F}$ *the family of closed subsets of a topological space* $\mathbb{F}$, *a random closed set* Z *in* $\mathcal{F}$ *is called **simple** if it assumes at most a finite number of values, so that there exists a finite measurable partition* $\Omega_1, \ldots, \Omega_m$ *of* Ω *and sets* $K_1, ..., K_m \in \mathcal{F}$ *such that* $Z(\omega) = K_i$ *for all* $\omega \in \Omega_i$, $1 \leq i \leq m$.

Because the probability space is non-atomic, the conditional Aumann expectation of $Q(S_\theta)$ is a convex set and is given by the weighted Minkowski sum of the possible realizations of $\text{co}[Q(S_\theta)]$, see Molchanov

(2005, Theorem 2.1.21). Each of these realizations is a polytope, and the Minkowski sum of polytopes is a polytope. Hence, $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$ is a closed convex polytope, fully characterized by a finite number of supporting hyperplanes.

Example 2 (Cont.) *Consider again the simple two player entry game with pure strategies only in Example 2. Then for $\varepsilon \in \mathcal{E}_\theta^M$ the set $Q\left(S_\theta\right)$ contains only two points, $[0\quad 1\quad 0\quad 0]'$ and $[0\quad 0\quad 1\quad 0]'$, and for $\varepsilon \notin \mathcal{E}_\theta^M$ it is a singleton. Therefore, the expectations of the selections of $Q\left(S_\theta\right)$ are given by*

$$\mathbf{E}\left(q\right) = \left[\mathbf{P}\left(\varepsilon \in \mathcal{E}_\theta^{(0,0)}\right) \quad \mathbf{P}\left(\varepsilon \in \mathcal{E}_\theta^{(1,0)}\right) \quad \mathbf{P}\left(\varepsilon \in \mathcal{E}_\theta^{(0,1)}\right) \quad \mathbf{P}\left(\varepsilon \in \mathcal{E}_\theta^{(1,1)}\right)\right]' + [0 \quad p_1 \quad 1-p_1 \quad 0]'\,\mathbf{P}\left(\varepsilon \in \mathcal{E}_\theta^M\right)$$

where $p_1 = \mathbf{P}\left(\Omega_1^M \middle| \omega : \varepsilon\left(\omega\right) \in \mathcal{E}_\theta^M\right)$, for all measurable $\Omega_1^M \subset \left\{\omega : \varepsilon\left(\omega\right) \in \mathcal{E}_\theta^M\right\}$, $i = 1, 2$. If the probability space has no atoms, then the possible values for p_1 fill in the whole $[0, 1]$ segment. Hence, $\mathbb{E}\left(Q\left(S_\theta\right)\right)$ is a segment in Δ^3. $\square$

The supporting hyperplanes determining $\operatorname{co}\left[Q\left(S_\theta\right)\right]$ can be easily obtained, and similarly for the supporting hyperplanes determining $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$. Hence, checking whether $\mathbf{P}\left(y|\,\underline{x}\right) \in \mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$ amounts to checking whether a point belongs to a polytope, i.e. whether a finite number of moment inequalities hold $\underline{x} - a.s.$ In Theorem 3.2 we show that these inequalities are obtained by checking inequality $u'\mathbf{P}\left(y|\,\underline{x}\right) \leq \mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right)|\,\underline{x}\right]$ for the $2^{\kappa_\mathcal{Y}}$ possible u vectors whose entries are either equal to zero or to one.

Theorem 3.2 *Assume that players use only pure strategies, that Assumptions 1-4 are satisfied, and that no other information is available. Then for $\underline{x} - a.s.$ these two conditions are equivalent:*
1. $u'\mathbf{P}\left(y|\,\underline{x}\right) \leq \mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right)|\,\underline{x}\right]\ \forall\ u \in \Re^{\kappa_\mathcal{Y}}$,
2. $u'\mathbf{P}\left(y|\,\underline{x}\right) \leq \mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right)|\,\underline{x}\right]\ \forall\ u \in \left\{u = \left[u_1 \quad ... \quad u_{\kappa_\mathcal{Y}}\right]' : u_i \in \{0, 1\},\ i = 1, ..., \kappa_\mathcal{Y}\right\}$.

Proof. It is obvious that condition (1) implies condition (2). To see why condition (2) implies condition (1), observe that because the set $Q\left(S_\theta\right)$ and the set $\operatorname{co}\left[Q\left(S_\theta\right)\right]$ are simple, one can find a finite measurable partition $\Omega_1, \ldots, \Omega_m$ of Ω and sets $K_1, ..., K_m \in \Delta^{\kappa_\mathcal{Y}-1}$, such that by Theorem 2.1.21 in Molchanov (2005)

$$\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right) = K_1\mathbf{P}\left(\Omega_1|\,\underline{x}\right) \oplus K_2\mathbf{P}\left(\Omega_2|\,\underline{x}\right) \oplus ... \oplus K_m\mathbf{P}\left(\Omega_m|\,\underline{x}\right),$$

with K_i the value that $\operatorname{co}\left[Q\left(S_\theta\left(\omega\right)\right)\right]$ takes for $\omega \in \Omega_i$, $i = 1, ..., m$ (see Definition 4). By the properties of the support function, see Schneider (1993, Theorem 1.7.5),

$$h\left(\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right), u\right) = \sum_{i=1}^{m} \mathbf{P}\left(\Omega_i|\,\underline{x}\right) h\left(K_i, u\right).$$

Finally, for each $i = 1, ..., m$, $\text{vert}(K_i) \subset \text{vert}\left(\Delta^{\kappa_{\mathcal{Y}}-1}\right)$, with $\text{vert}(\cdot)$ the vertices of the set in parenthesis. Hence the supporting hyperplanes of K_i, $i = 1, ..., m$, are a subset of the supporting hyperplanes of the simplex $\Delta^{\kappa_{\mathcal{Y}}-1}$, which in turn are obtained through its support function evaluated in directions $u \in \left\{u = \begin{bmatrix} u_1 & ... & u_{\kappa_{\mathcal{Y}}} \end{bmatrix}' : u_i \in \{0,1\},\ i = 1, ..., \kappa_{\mathcal{Y}}\right\}$. Therefore the supporting hyperplanes of $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$ are a subset of the supporting hyperplanes of $\Delta^{\kappa_{\mathcal{Y}}-1}$. ∎

In Appendix B we connect this result to a related notion in the theory of random sets, that of a *capacity functional* (the "probability distribution" of a random closed set), and we provide an equivalent characterization of the sharpness result which gives further insights into our approach. In Section 3.5 and Appendix B we provide results that significantly reduce the number of inequalities to be checked, by showing that depending on the model under consideration, many of the $2^{\kappa_{\mathcal{Y}}}$ inequalities in Theorem 3.2 are redundant.

To conclude this section, it is important to discuss why the sharp identification region cannot in general be obtained through a finite number of moment inequalities. When players are not allowed to randomize over their actions, the family of possible equilibria is finite. Hence, the range of values that ε takes can be partitioned into areas in which the set of equilibria remains constant, that is, does not depend on ε any longer. However, when players randomize across their actions, in equilibrium they must be indifferent among the actions over which they place positive probability. This implies that there exist regions in the sample space where the equilibrium mixed strategy profiles are a function of ε directly.[16] If ε has a discrete distribution, $Q\left(S_\theta\right)$ continues to be a simple random set, i.e. a random set which takes only a finite number of values, and $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$ remains a convex polytope whose supporting hyperplanes can be characterized exactly, so that the sharp identification region can be obtained through a finite number of moment inequalities. However, when the distribution of ε is continuous, $Q\left(S_\theta\right)$ may take a continuum of values as a function of ε, and $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$ may have infinitely many extreme points. Therefore, one needs an infinite number of moment inequalities to determine whether $\mathbf{P}\left(y|\,\underline{x}\right)$ belongs to it. In this case, the most practical approach to obtain the sharp identification region is by solving the minimization problem that we lay out in equation (3.8).

3.4.1 Example: Two Type, Four Player Entry Game with 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

[16] For example, in the two player entry game in Example 1, for $\varepsilon \in \mathcal{E}_M^\theta$, $S_\theta = \left\{(0,1), \left(\frac{\varepsilon_2}{-\theta_2}, \frac{\varepsilon_1}{-\theta_1}\right), (1,0)\right\}$. However, if one restricts players to use pure strategies, then for $\varepsilon \in \mathcal{E}_M^\theta$, $S_\theta = \{(0,1), (1,0)\}$, with no additional dependence of the equilibria on ε.

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 $a_{jm} \in \{0,1\}$ be the strategy of firm $j = 1,2$ of type $m = 1,2$. Entry is denoted by $a_{jm} = 1$, with $a_{jm} = 0$ denoting staying out. Players $j = 1,2$ of type 1 and type 2 have respectively the following payoff functions:

$$\pi_{j1}(a_{j1}, a_{-j1}, a_{12}, a_{22}, \varepsilon_1) = a_{j1}(\theta_{11}(a_{-j1} + a_{12} + a_{22}) - \varepsilon_1), \tag{3.14}$$

$$\pi_{j2}(a_{j2}, a_{-j2}, a_{11}, a_{21}, \varepsilon_2) = a_{j2}(\theta_{21}(a_{11} + a_{21}) + \theta_{22}a_{-j2} - \varepsilon_2). \tag{3.15}$$

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. Let $y_1 = a_{11} + a_{21}$ denote the number of entrants of type 1 and $y_2 = a_{12} + a_{22}$ the number of entrants of type 2 that a firm faces, so that $y_m \in \{0,1,2\}$, $m = 1,2$. Then there are 9 possible outcomes to this game, ordered as follows: $\mathcal{Y} = \{(0,0), (0,1), (1,0), (1,1), (2,0), (0,2), (1,2), (2,1), (2,2)\}$. Notice that here Assumption 1-(iii) rather than Assumption 2 holds. Figure 4 plots the outcomes of the game against the realizations of $\varepsilon_1, \varepsilon_2$. In this case, $Q(S_\theta)$ takes its realizations in the vertices of Δ^8. For example, for $\omega : \varepsilon_1(\omega) \geq \theta_{11}$, $\varepsilon_2(\omega) \geq \theta_{22}$, the game has a unique equilibrium outcome, $y = (0,0)$, and $Q(S_\theta(\omega)) = \{[1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]'\}$; for $\omega : 2\theta_{11} \leq \varepsilon_1(\omega) \leq \theta_{11}$, $2\theta_{22} \leq \varepsilon_2(\omega) \leq \theta_{22}$, the game has two equilibrium outcomes, $y = (0,1)$ and $y = (1,0)$, and $Q(S_\theta(\omega)) = \{[0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0]', [0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0]'\}$; etc.

Because the set $\mathcal{Y}$ has cardinality 9, in principle there are $2^9 = 512$ inequality restrictions to consider, corresponding to each binary vector of length 9. However, the number of inequalities to be checked is significantly smaller. In particular, by a simple application of the insight in equation (3.17) below, the sharp identification region that we give is based on 26 inequalities, whereas Θ_O^{ABJ} and Θ_O^{CT} are based, respectively, on 9 and 18 inequalities. Hence, the computational burden is essentially equivalent.

Figure 8 and Table 2 report Θ_I, Θ_O^{CT}, and Θ_O^{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_O^{ABJ} is given by the union of the yellow, red and black segments, and Θ_O^{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_O^{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_O^{CT} and Θ_O^{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.

In order to further illustrate the computational advantages of our characterization of Θ_I in equation (3.8), we also recalculated the sharp identification region for this example *without* taking advantage of our knowledge of the structure of the game which reduces the number of inequalities to be checked to 26, but by simply solving for each candidate $\theta \in \Theta$ the problem $\min_{u \in \mathfrak{B}} \left(\mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right) | \underline{x}\right] - u'\mathbf{P}\left(y | \underline{x}\right)\right)$. We modified the simple Nelder-Mead algorithm described in Section 3.3.1 to apply to a minimization in $\Re^9$, wrote it as a program in Fortran 90, and compiled and ran it on a Unix machine with a single processor of 3.2 GHz. Our recalculation of Θ_I yielded exactly the same result as described above, and checking 10^6 candidate values for $\theta \in \Theta$ took less than one minute.

3.5 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).[17] 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 $u \in \mathfrak{B}$,

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

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

One can simulate this integral using the following procedure.[18] 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, $b = 1, \ldots, B$, 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 $u \in \mathfrak{B}$, compute the support function $h\left(Q\left(S_\theta\left(\underline{x}, \varepsilon^b\right)\right), u\right)$, and average it over a large number of draws of ε^b. Call the resulting average $\hat{\mathbf{E}}_B\left[h\left(Q\left(S_\theta\right), u\right)|\underline{x}\right]$. The strong law of large numbers for closed random sets in Molchanov (2005, Theorem 3.1.6) guarantees that as $B \to \infty$, i.e., as the number of simulations increases, $\hat{\mathbf{E}}_B\left[h\left(Q\left(S_\theta\right), u\right)|\underline{x}\right]$ converges to $\mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right)|\underline{x}\right]$ almost surely, uniformly in u.

Based on the above simulation procedure, one can rewrite the analog of the criterion function $w(\theta)$ from equation (3.9) as

$$w_B(\theta) = \int \min_{u \in \mathfrak{B}} \left(\hat{\mathbf{E}}_B\left[h\left(Q\left(S_\theta\right), u\right)|\underline{x}\right] - u'\mathbf{P}(y|\underline{x})\right) dF_{\underline{x}} = \int -d_H\left(\mathbf{P}(y|\underline{x}), \widehat{\mathbb{E}}_B\left(Q\left(S_\theta\right)|\underline{x}\right)\right) dF_{\underline{x}},$$

with $\widehat{\mathbb{E}}_B\left(Q\left(S_\theta\right)|\underline{x}\right)$ denoting the convex body with support function equal to $\hat{\mathbf{E}}_B\left[h\left(Q\left(S_\theta\right), u\right)|\underline{x}\right]$ $\forall$ $u \in \Re^{\kappa_{\mathcal{Y}}}$. By triangle inequality[19]

$$\begin{aligned} \sup_{\theta\in\Theta}\left|w_B(\theta) - w(\theta)\right| &\leq \sup_{\theta\in\Theta}\int \max_{u\in\mathfrak{B}}\left|\hat{\mathbf{E}}_B\left[h\left(Q\left(S_\theta\right), u\right)|\underline{x}\right] - \mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right)|\underline{x}\right]\right| dF_{\underline{x}} \\ &\leq \sup_{\theta\in\Theta}\left(\max_{u\in\mathfrak{B}}\left|\hat{\mathbf{E}}_B\left[h\left(Q\left(S_\theta\right), u\right)\right] - \mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right)\right]\right|\right), \end{aligned} \tag{3.16}$$

where the last inequality in the above expression follows by the properties of the conditional Aumann expectation, see Molchanov (2005, Theorem 2.1.47-(v)). For each $\theta \in \Theta$, the process in parentheses in expression (3.16), when multiplied by $\sqrt{B}$, 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 expectation of the support function of $Q(S_\theta)$ delivers an approximation of Θ_I which converges to Θ_I with respect to the Hausdorff metric as $B \to \infty$.

[18] 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.

[19] Here we are using the fact that

$$\begin{aligned} \left|d_H\left(\mathbf{P}(y|\underline{x}), \mathbb{E}\left(Q\left(S_\theta\right)|\underline{x}\right)\right) - d_H\left(\mathbf{P}(y|\underline{x}), \widehat{\mathbb{E}}_B\left(Q\left(S_\theta\right)|\underline{x}\right)\right)\right| &\leq \rho_H\left(\mathbb{E}\left(Q\left(S_\theta\right)|\underline{x}\right), \widehat{\mathbb{E}}_B\left(Q\left(S_\theta\right)|\underline{x}\right)\right) \\ &= \max_{u\in\mathfrak{B}}\left|\hat{\mathbf{E}}_B\left[h\left(Q\left(S_\theta\right), u\right)|\underline{x}\right] - \mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right)|\underline{x}\right]\right|. \end{aligned}$$

Consider now the general task of computing Θ_I when mixed strategy equilibria are allowed for. Given the simulated values for $\mathbf{E}\left[h\left(Q\left(S_\theta\right),u\right)\right]$, the minimization problems defining Θ_I in equation (3.8) can be solved for each $\theta \in \Theta$ and for each $\underline{x} \in \mathcal{X}$. In small scale problems, such as the two player entry game in Example 1, even the simple use of the Nelder-Mead algorithm delivers the sharp identification region very quickly, as discussed in Sections 3.3.1 and 3.4.1. In larger problems, the use of efficient convex programming algorithms preserves computational feasibility.

Consider now the special case that players play only pure strategies. Then Theorem 3.2 guarantees that it suffices to evaluate $\hat{\mathbf{E}}_B\left[h\left(Q\left(S_\theta\right),u\right)\middle|\underline{x}\right] - u'\mathbf{P}\left(y\middle|\underline{x}\right)$ for u equal to each of the $2^{\kappa_\mathcal{Y}}$ binary vectors with each entry equal to either 1 or 0. Hence, one can calculate the set Θ_I either by finding the parameter values that satisfy these $2^{\kappa_\mathcal{Y}}$ inequalities which have to hold for $\underline{x} - a.s.$, or by solving the minimization problem in equation (3.8). In the former case, the number of inequalities to check can be, in practice, very large. However, often many such inequalities are redundant. In particular, because we are allowing only pure strategy equilibria, the realizations of any $\sigma \in S_\theta$ are vectors of zeros and ones. Hence, $\forall \omega \in \Omega$, $\left[q\left(\sigma\left(\omega\right)\right)\right]_k = 1$ if $\prod_{j=1}^{J} \sigma_j\left(\omega, t_j^k\right) = 1$, and zero otherwise. Consider two equilibria $t^k, t^l \in \mathcal{Y}$, $1 \leq k \neq l \leq \kappa_\mathcal{Y}$, such that

$$\left\{\omega : \prod_{j=1}^{J} \sigma_j\left(\omega, t_j^k\right) = 1 \middle| \underline{x}\right\} \cap \left\{\omega : \prod_{j=1}^{J} \sigma_j\left(\omega, t_j^l\right) = 1 \middle| \underline{x}\right\} = \emptyset, \tag{3.17}$$

that is, the set of ω for which S_θ admits both t^k and t^l as equilibria has probability zero. Let u^k be a vector with each entry equal to zero, and entry k equal to 1, and similarly for u^l. Then the inequality $\left(u^k + u^l\right)' \mathbf{P}\left(y\middle|\underline{x}\right) \leq \mathbf{E}\left[h\left(Q\left(S_\theta\right), u^k + u^l\right)\middle|\underline{x}\right]$ does not add any information beyond that provided by the inequalities $u'\mathbf{P}\left(y\middle|\underline{x}\right) \leq \mathbf{E}\left[h\left(Q\left(S_\theta\right),u\right)\middle|\underline{x}\right]$ for $u = u^k$ and for $u = u^l$. The same reasoning can be extended to tuples of pure strategy equilibria of size up to $\kappa_\mathcal{Y}$. Therefore, prior knowledge of some properties of the game can be very helpful in eliminating unnecessary inequalities. The game we described in Section 3.4.1 is an example for the possible elimination of redundant inequalities: $\kappa_\mathcal{Y} = 512$, but it suffices to consider 26 inequalities to obtain the sharp identification region. By comparison, the calculation of Θ_O^{ABJ} and Θ_O^{CT} requires checking 9 and 18 inequalities respectively, a task which is computationally equivalent to that required to calculate Θ_I, while the gain in size of the identification region is of the order of $27\% - 54\%$.

Theorems B.2-B.3 in Appendix B provide general results on how to reduce the number of inequalities to be checked. We emphasize, however, that as the example in Section 3.4.1 illustrates, even when such results do not apply, the minimization problem in equation (3.8) can be solved extremely efficiently.

4 Multinomial Choice Models with Interval Regressors Data

This Section of the paper extends the methodology introduced in Section 3 to provide a tractable characterization of the sharp identification region of the parameters θ characterizing random utility models of multinomial choice, when only interval information is available on regressors. In doing so, we extend the seminal contribution of Manski and Tamer (2002), who considered the same inferential problem in the case of binary choice models. For these models, Manski and Tamer (2002) provided a tractable characterization of the sharp identification region, and proposed set estimators which are consistent with respect to the Hausdorff distance. However, their characterization of the sharp identification region does not easily extend to models in which the agents face more than two choices, as we illustrate below.

We assume that an agent chooses an alternative y from a finite choice set $\mathcal{C} = \{0, \ldots, \kappa_{\mathcal{C}} - 1\}$ to maximize her utility. The agent possesses a vector of socioeconomic characteristics w. Each alternative $k \in \mathcal{C}$ is characterized by an observable vector of attributes z_k and an attribute ε_k which is observable by the agent but not by the econometrician. The vector $\left(y, w, \{z_k, \varepsilon_k\}_{k=0}^{\kappa_{\mathcal{C}}-1}\right)$ is defined on a non-atomic probability space $(\Omega, \mathfrak{F}, \mathbf{P})$. The agent is assumed to possess a random utility function of known parametric form.

To simplify the exposition, we assume that the random utility is linear, and that w and z_k are scalars for $k = 0, \ldots, \kappa_{\mathcal{C}} - 1$; however, *all* these assumptions can be relaxed and are in no way essential for our methodology. We let the random utility be $\pi(k; x_k, \varepsilon_k, \theta_k) = \alpha_k + z_k\delta + w\beta_k + \varepsilon_k \equiv x_k\theta_k + \varepsilon_k$, $k \in \mathcal{C}$, with $x_k = [1 \;\; z_k \;\; w]$, and $\theta_k = [\alpha_k \;\; \delta \;\; \beta_k]'$. We normalize $\pi(0; x_0, \varepsilon_0, \theta_0) = \varepsilon_0$. For simplicity, we assume that ε_k is independently and identically distributed across choices with a continuous distribution function F that is known. We let $\theta = \left[\{\alpha_k\}_{k=1}^{\kappa_{\mathcal{C}}-1} \quad \delta \quad \{\beta_k\}_{k=1}^{\kappa_{\mathcal{C}}-1}\right]' \in \Theta$ be the vector of parameters of interest, with Θ the parameter space. We denote $\varepsilon^k = \varepsilon_k - \varepsilon_0$, $k \in \mathcal{C}$, and $\varepsilon = \left[\left\{\varepsilon^k\right\}_{k=1}^{\kappa_{\mathcal{C}}-1}\right]$. Under these assumptions, if the econometrician observes a random sample of choices, socioeconomic characteristics, and alternatives' attributes, the parameter vector θ is point identified.

Here we consider the identification problem arising when the econometrician observes only realizations of $\{y, z_{kL}, z_{kU}, w\}$, but not realizations of z_k, $k = 1, \ldots, \kappa_{\mathcal{C}} - 1$. Following Manski and Tamer (2002), we assume that for each $k = 1, \ldots, \kappa_{\mathcal{C}} - 1$, $\mathbf{P}(z_{kL} \leq z_k \leq z_{kU}) = 1$, and that $\delta > 0$. We let $x_{kL} = [1 \;\; z_{kL} \;\; w]$, $x_{kU} = [1 \;\; z_{kU} \;\; w]$, $\underline{x}_k = [1 \;\; z_{kL} \;\; z_{kU} \;\; w]$, and $\underline{x} = [1 \quad \{z_{kL}\}_{k=1}^{\kappa_{\mathcal{C}}-1} \quad \{z_{kU}\}_{k=1}^{\kappa_{\mathcal{C}}-1} \quad w]$. Incompleteness of the data on z_k, $k = 1, \ldots, \kappa_{\mathcal{C}} - 1$, implies that there are regions of values of the exogenous variables where the econometric model predicts that more than one choice may maximize utility. 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 utility functions may not be point identified.

In the case of binary choice, Manski and Tamer (2002) establish that the sharp identification region for θ is given by

$$\Theta_I = \left\{\theta \in \Theta : \mathbf{P}\left(x_{1L}\theta + \varepsilon^1 > 0 \middle| \underline{x}\right) \leq \mathbf{P}\left(y = 1 | \underline{x}\right) \leq \mathbf{P}\left(x_{1U}\theta + \varepsilon^1 > 0 \middle| \underline{x}\right),\ \underline{x} - a.s.\right\}.$$

This construction is based on the observation that if the agent chooses alternative 1, this implies that $\varepsilon^1 > -x_1\theta \geq -x_{1U}\theta$. On the other hand, $\varepsilon^1 > -x_{1L}\theta \geq -x_1\theta$ implies that the agent chooses alternative 1.[20] In the case of more than two choices, one may wish to apply a similar insight as in the work of Ciliberto and Tamer (2006), and construct the region

(4.1)

$$\Theta_O = \left\{ \begin{array}{c} \theta \in \Theta : \forall m \in \mathcal{C},\ \underline{x} - a.s., \\ \mathbf{P}\left(x_m\theta_m + \varepsilon^m \geq x_k\theta_k + \varepsilon^k\ \forall \left(x_m, x_k\right) \in \left[x_{mL}, x_{mU}\right] \times \left[x_{kL}, x_{kU}\right],\ \forall k \in \mathcal{C},\ k \neq m \middle| \underline{x}\right) \\ \leq \mathbf{P}\left(y = m | \underline{x}\right) \leq \\ \mathbf{P}\left(\exists\ x_m \in \left[x_{mL}, x_{mU}\right] \text{ s.t. } \forall k \in \mathcal{C},\ k \neq m, \exists\ x_k \in \left[x_{kL}, x_{kU}\right] \text{ with } x_m\theta_m + \varepsilon^m \geq x_k\theta_k + \varepsilon^k \middle| \underline{x}\right) \end{array} \right\}.$$

The lower bound on $\mathbf{P}\left(y = m | \underline{x}\right)$ in equation (4.1) is given by the probability that ε falls in the regions where choice $m \in \mathcal{C}$ is the only optimal alternative. The upper bound is given by the probability that ε falls in the regions where choice $m \in \mathcal{C}$ is one of the possible optimal alternatives. Similarly to the case of Θ_O^{CT} in the finite games analyzed in Section 3, Θ_O is just an outer region for θ, and is not sharp in general. Appendix B provides further insights to explain the lack of sharpness of Θ_O.[21]

We begin our treatment of the identification problem by noticing that, if x_k were observed for each $k \in \mathcal{C}$, one would conclude that a choice $m \in \mathcal{C}$ maximizes utility if

$$\pi\left(m; x_m, \varepsilon_m, \theta_m\right) = x_m\theta_m + \varepsilon_m \geq x_k\theta_k + \varepsilon_k = \pi\left(k; x_k, \varepsilon_k, \theta_k\right)\ \forall\ k \in \mathcal{C},\ k \neq m.$$

Hence, for a given $\theta \in \Theta$ and realization of $\underline{x}$ and ε, we can define the following θ-dependent set:

(4.2)

$$M_\theta\left(\underline{x}, \varepsilon\right) = \left\{m \in \mathcal{C} : \exists\ x_m \in \left[x_{mL}, x_{mU}\right] \text{ s.t. } \forall k \in \mathcal{C}, k \neq m, \exists\ x_k \in \left[x_{kL}, x_{kU}\right] \text{ with } x_m\theta_m + \varepsilon^m \geq x_k\theta_k + \varepsilon^k\right\}.$$

This is the set of choices associated with a specific value of θ and realization of $\underline{x}$ and ε, which are optimal for some combination of $x_k \in \left[x_{kL}, x_{kU}\right]$, $k \in \mathcal{C}$, and therefore form the set of model's predictions. As we did in Section 3, we write the set $M_\theta\left(\underline{x}, \varepsilon\right)$ and its realizations, respectively, as M_θ and $M_\theta(\omega) \equiv M_\theta\left(\underline{x}(\omega), \varepsilon(\omega)\right)$, omitting the explicit reference to $\underline{x}$ and ε. Because M_θ is a subset of a discrete space, and any event of the type $\{m \in M_\theta\}$ can be represented as a combination of measurable events determined by ε_k, $k \in \mathcal{C}$, M_θ is a random closed set in $\mathcal{C}$, see Definition 2.

[20] For $-x_{1U}\theta \leq \varepsilon^1 \leq -x_{1L}\theta$, the model predicts that either alternative 0 or 1 may maximize the agent's utility.

[21] Appendix B focuses on the lack of sharpness of Θ_O^{CT} in finite games with multiple pure strategy Nash equilibria. The same reasoning applies to the set Θ_O in equation (4.1).

We now apply to the random closed set M_θ the same logic that we applied to the random closed set S_θ in Section 3. The treatment which follows is akin to the treatment of static, simultaneous move finite games of complete information, when players use only pure strategies.

For a given parameter value $\theta \in \Theta$ and realization $m(\omega)$, $\omega \in \Omega$, of a selection $m \in \operatorname{Sel}(M_\theta)$, the individual chooses alternative $k = 0, \ldots, \kappa_{\mathcal{C}} - 1$ if and only if $m(\omega) = k$. Hence, we can use a selection $m \in \operatorname{Sel}(M_\theta)$ to define a random point $q(m)$ whose realizations have coordinates $[q(m(\omega))]_k = 1(m(\omega) = k)$, $k = 0, \ldots, \kappa_{\mathcal{C}} - 1$, with $1(\cdot)$ the indicator function of the event in parenthesis. Clearly, the random point $q(m)$ is an element of the unit simplex in the space of dimension $\kappa_{\mathcal{C}}$, denoted $\Delta^{\kappa_{\mathcal{C}}-1}$. Because M_θ is a random closed set in $\mathcal{C}$, the set resulting from repeating the above construction for each $m \in \operatorname{Sel}(M_\theta)$ and given by

$$Q(M_\theta) = \{([q(m)]_k, \ k = 0, \ldots, \kappa_{\mathcal{C}} - 1) : m \in \operatorname{Sel}(M_\theta)\},$$

is a closed random set in $\Delta^{\kappa_{\mathcal{C}}-1}$. Hence we can define the set

$$\mathbb{E}(Q(M_\theta)|\,\underline{x}) = \{\mathbf{E}(q|\,\underline{x}) : q \in \operatorname{Sel}(Q(M_\theta))\} = \{(\mathbf{E}([q(m)]_k|\,\underline{x}), k = 0, \ldots, \kappa_{\mathcal{C}} - 1) : m \in \operatorname{Sel}(M_\theta)\}.$$

Because the probability space is non-atomic and the random set $Q(M_\theta)$ takes its realizations in a subset of the finite dimensional space $\Re^{\kappa_{\mathcal{C}}}$, the set $\mathbb{E}(Q(M_\theta)|\,\underline{x})$ is a closed convex set for $\underline{x} - a.s.$ By construction, it is the set of probability distributions over alternatives conditional on $\underline{x}$ which are consistent with the maintained modeling assumptions, i.e., with *all* the model implications. If the model is correctly specified, there exists at least one value of $\theta \in \Theta$ such that the observed conditional distribution of y given $\underline{x}$, $\mathbf{P}(y|\,\underline{x})$, is a point in the set $\mathbb{E}(Q(M_\theta)|\,\underline{x})$ for $\underline{x} - a.s.$, where $\mathbf{P}(y|\,\underline{x}) \equiv [\mathbf{P}(y = k|\,\underline{x}), k = 0, \ldots, \kappa_{\mathcal{C}} - 1]$.

Using the same mathematical tools leading to Theorem 3.1, we obtain that the set of observationally equivalent parameter values which form the sharp identification region is given by

$$\Theta_I = \left\{\theta \in \Theta : \min_{u \in \mathfrak{B}} \left(\mathbf{E}[h(Q(M_\theta), u)|\,\underline{x}] - u'\mathbf{P}(y|\,\underline{x})\right) = 0 \ \underline{x} - a.s.\right\}, \tag{4.3}$$

with $\mathfrak{B}$ the unit ball in $\Re^{\kappa_{\mathcal{C}}}$.

Notice that the set $Q(M_\theta)$ assumes at most a finite number of values, and its realizations lie in the subsets of the vertices of $\Delta^{\kappa_{\mathcal{C}}-1}$. The conditional Aumann expectation of $Q(M_\theta)$ is given by the weighted Minkowski sum of the possible realizations of $\operatorname{co}[Q(M_\theta)]$. Each of these realizations is a polytope, and therefore $\mathbb{E}(Q(M_\theta)|\,\underline{x})$ is a closed convex polytope. By Theorem 3.2, a candidate θ belongs to Θ_I as defined in equation (4.3) if and only if $u'\mathbf{P}(y|\,\underline{x}) \le \mathbf{E}[h(Q(M_\theta), u)|\,\underline{x}]$ for each of the $2^{\kappa_{\mathcal{C}}}$ possible u vectors whose entries are either equal to zero or to one. Hence, Θ_I can be obtained through a finite set of moment inequalities which have to hold for $\underline{x} - a.s.$

5 Best Linear Prediction with Interval Outcome and Covariate Data

Beresteanu and Molinari (2008) study identification and statistical inference for the parameters $\theta \in \Theta$ of the Best Linear Predictor (BLP) under square loss of a random variable $y^\star$ conditional on a random vector $x^\star$, when $y^\star$ is only observed to lie in a random interval. Here we significantly generalize their identification results, by considering the case that both outcome and covariate data are interval valued. Earlier on, Horowitz, Manski, Ponomareva, and Stoye (2003) studied the related problem of identification of the BLP parameters with missing data on both $y^\star$ and $x^\star$, and provided a characterization of the identification region of each component of the vector θ. While their characterization is sharp, the bounds that they provide are obtained as solutions to non-convex mathematical programming problems for which global optimization techniques are needed. The computational complexity of the problem in the Horowitz et al.'s (2003) formulation grows with the number of points in the support of the random variables $y^\star$ and $x^\star$, and becomes essentially unfeasible if these variables are continuous, unless one discretizes their support quite coarsely. Using the same approach as in the previous part of the paper, we provide a characterization of the sharp identification region of θ which remains computationally feasible regardless of the support of $y^\star$ and $x^\star$.

To simplify the exposition, we let $x^\star$ be scalar, though this assumption can be relaxed and is in no way essential for our methodology. We assume that the researcher does not observe the realizations of $(y^\star, x^\star)$, but rather the realizations of real valued random variables y_L, y_U, x_L, x_U such that $\mathbf{P}(y_L \leq y^\star \leq y_U) = 1$ and $\mathbf{P}(x_L \leq x^\star \leq x_U) = 1$. We assume that $\mathbf{E}(|y_i|)$, $\mathbf{E}(|x_j|)$, $\mathbf{E}(|y_i x_j|)$, and $\mathbf{E}\left(x_j^2\right)$ are all finite, for each $i = L, U$ and $j = L, U$, that the vector (y_L, y_U, x_L, x_U) is defined on a non-atomic probability space $(\Omega, \mathfrak{F}, \mathbf{P})$, and that Θ is a compact set. We let $Y = [y_L, y_U]$, $X = [x_L, x_U]$; one can easily show that these sets are random closed sets in $\Re$ (see Definition 2).

When $y^\star$ and $x^\star$ are perfectly observed, it is well known that the BLP problem can be expressed through a linear projection model, where the prediction error associated with the BLP parameters $\theta^\star$ and given by $\varepsilon^\star = y^\star - \theta_1^\star - \theta_2^\star x^\star$ by construction satisfies $\mathbf{E}(\varepsilon^\star) = 0$ and $\mathbf{E}(\varepsilon^\star x^\star) = 0$. For any candidate $\theta \in \Theta$, we extend this construction of the prediction error to the case of interval valued data, and build the set

$$G_\theta = \left\{ g = \begin{bmatrix} y - \theta_1 - \theta_2 x \\ (y - \theta_1 - \theta_2 x)\, x \end{bmatrix} : \ (y, x) \in \operatorname{Sel}(Y \times X) \right\}.$$

This is the (not necessarily convex) θ-dependent set of prediction errors and prediction errors multiplied by covariate which are implied by the intervals Y and X. Because Y and X are random closed sets in $\Re$, G_θ is a random closed set in $\Re^2$. It follows from the restrictions on the moments of y_L, y_U, x_L, x_U, that the set G_θ is integrably bounded. Because the probability space is non-atomic and G_θ belongs to a finite dimensional space, its Aumann expectation $\mathbb{E}(G_\theta) = \{\mathbf{E}(g) : g \in \mathrm{Sel}(G_\theta)\}$ is a closed convex set.

When first thinking about the problem that we study in this section, it might not be obvious why its treatment fits in the general methodology that we develop in the previous part of the paper. Here, the researcher observes data on y_L, y_U, x_L, x_U, which identify the joint distribution of the random sets Y and X, but are silent about $y^\star$, $x^\star$ beyond providing random intervals to which these variables belong. However, given the set G_θ, one can relate our approach in the first part of the paper to the problem that we study here, as follows. For a candidate $\theta \in \Theta$, each selection (y, x) from the random intervals Y and X yields a moment for the prediction error $\varepsilon = y - \theta_1 - \theta_2 x$ and its product with the covariate x. The collection of such moments for all $(y, x) \in \mathrm{Sel}(Y \times X)$ is a convex set equal to $\mathbb{E}(G_\theta)$. If this (unconditional) Aumann expectation contains the vector $[0 \;\; 0]'$ as one of its elements, then the candidate value of θ is one of the observationally equivalent parameters of the BLP of $y^\star$ given $x^\star$. This is because if the condition just mentioned is satisfied, then for the candidate $\theta \in \Theta$ there exists a selection in $\mathrm{Sel}(Y \times X)$, that is, a pair of admissible random variables y and x, which implies a prediction error that has mean zero and is uncorrelated with x, hence satisfying the requirements for the BLP prediction error. This intuition is formalized in Theorem 5.1.

Theorem 5.1 *Let $o = [0 \;\; 0]'$. The sharp identification region for $\theta \in \Theta$ is given by*

$$\begin{aligned}
\Theta_I &= \{\theta : o \in \mathbb{E}(G_\theta)\} \\
&= \{\theta : 0 \leq h(\mathbb{E}(G_\theta), u) \;\; \forall \, u \in \mathfrak{B}\} \\
&= \left\{\theta : \min_{u \in \mathfrak{B}} \mathbf{E}\left[h(G_\theta, u)\right] = 0\right\}.
\end{aligned}$$

Proof. By definition of the Aumann expectation, $o \in \mathbb{E}(G_\theta)$ if and only if $\exists \; g \in \mathrm{Sel}(G_\theta) : \mathbf{E}(g) = o$. In words, this is equivalent to saying that a candidate θ belongs to Θ_I if and only if one can find a selection $(y, x) \in \mathrm{Sel}(Y \times X)$ which yields, together with θ, a prediction error $\varepsilon = y - \theta_1 - \theta_2 x$ such that $\mathbf{E}(\varepsilon) = 0$ and $\mathbf{E}(\varepsilon x) = 0$. Observe that by Theorem 2.1 in Artstein (1983) (see also the discussion in Molchanov (2005, pp. 34-35)), $(y, x) \in \mathrm{Sel}(Y \times X)$ if and only if $\mathbf{P}((y, x) \in K \times L) \leq \mathbf{P}((Y \times X) \cap K \times L \neq \emptyset) = \mathbf{P}(y_U > \inf K, y_L < \sup K, x_U > \inf L, x_L < \sup L)$ for all compact intervals $K, L \subset \Re$. Hence, the above condition is equivalent to being able to find a pair of random variables (y, x) with a joint distribution $\mathbf{P}(y, x)$ that belongs to the (sharp) identification region of $\mathbf{P}(y^\star, x^\star)$ as defined by Manski (2003, Chapter 3), such that $\theta = \arg\min_{\vartheta \in \Theta} \int (y - \vartheta_1 - \vartheta_2 x)^2 d\mathbf{P}(y, x)$. It then follows that the set Θ_I is equivalent to the sharp identification region characterized by Manski (2003, Complement 3B, pp. 56-58). The definition of Θ_I in terms of support function follows from our discussion in Section 3.2. ■

The support function of G_θ can be easily calculated. In particular, for any $u = [u_1 \ \ u_2]' \in \mathfrak{B}$,

$$h(G_\theta, u) = \max_{g \in G_\theta} u'g = \max_{y \in Y, x \in X} \left[u_1 (y - \theta_1 - \theta_2 x) + u_2 \left(yx - \theta_1 x - \theta_2 x^2 \right) \right]. \tag{5.1}$$

For given $\theta \in \Theta$ and $u \in \mathfrak{B}$, this maximization problem can be solved extremely quickly using the gradient method, regardless of whether y_L, y_U, x_L, x_U, $x^\star$, $y^\star$, are continuous or discrete random variables. The support function $h(G_\theta, u)$ that results in (5.1) is a continuous-valued convex sublinear function of u (see, e.g., Molchanov (2005, p. 421)). Hence, membership of a candidate θ to the set Θ_I can be verified extremely easily using convex programming techniques, as discussed in Section 3.

6 Conclusions

This paper introduces a computationally feasible characterization of the sharp identification region of the parameters of incomplete econometric models with set-valued predictions which yield a convex set of (conditional or unconditional) moments for the variables characterizing them. Examples of models in this class include static, simultaneous move finite games of complete information in the presence of multiple mixed strategy Nash equilibria, multinomial choice models with interval regressors data, and best linear predictors with interval outcome and covariate data. We summarize our results focusing on the case of finite games with multiple equilibria. In this context, 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 (e.g., Berry and Tamer (2007)).

Our approach is based on characterizing, for each $\theta \in \Theta$, the set of probability distributions of outcomes given covariates which are consistent with *all* the model's implications. If the model is correctly specified, one can then characterize the set of observationally equivalent parameter values $\theta \in \Theta$ which are consistent with the data and the modeling assumptions (which we call Θ_I), as follows. A candidate θ is in Θ_I if and only if it gives a set of model's predicted probability distributions of outcomes given covariates which, for $\underline{x} - a.s.$, contains the probability distribution of outcomes given covariates which is observed in the data.

Because in general, for each $\theta \in \Theta$, the set of probability distributions of outcomes given covariates may have infinitely many extreme points, characterizing the set Θ_I in principle entails checking that an infinite number of moment inequalities is satisfied for $\underline{x} - a.s.$ However, we show that this computational hardship can be avoided, and the sharp identification region can be characterized as the set of parameter values for which the minimum of a sublinear (hence convex) function over the unit ball is equal to zero.

For the case that players are assumed to play only pure strategies (and for the case of multinomial choice models with interval regressors data) we show that the sharp identification region is given by a finite number of moment inequalities which have to hold for $\underline{x} - a.s.$ This is because in this very special case, for each $\theta \in \Theta$, the set of probability distributions of outcomes given covariates which are consistent with *all* the model's implications, simplifies to being a polytope whose supporting hyperplanes are easy to determine. While finite, the number of moment inequalities to be checked can be very large in certain cases. However, we show that many such inequalities may be redundant, and we provide a simple condition that allows the researcher to determine a (often significantly) smaller set of moment inequalities that are sufficient to preserve sharpness. Moreover, the minimization problem used in the general case of finite games with mixed strategy Nash equilibria remains feasible in this simpler case, and can be applied when the number of inequalities cannot be reduced a-priori and is very large

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) in the analysis of finite games with multiple equilibria). However, advanced computational methods in convex programming made available in recent years, along with the use of parallel processing, can substantially alleviate this computational burden. On the other hand, the benefits in terms of identification yielded by our methodology may be substantial, as illustrated in our examples.

A Extensions to Other Solution Concepts

While in Sections 2-3 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.

A.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

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

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 5 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 in which players' payoffs are given by $\pi_j = a_j(a_{-j}\theta_j + \varepsilon_j)$, $a_j \in \{0,1\}$, and θ_1, θ_2 are assumed to be negative.

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

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

with

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

where $[q(\sigma)]_k$, $k = 1, \ldots, \kappa_{\mathcal{Y}}$, is defined in equation (3.2).

In our simple example in Figure 5, with omitted covariates, for any $\omega \in \Omega$ such that $\varepsilon(\omega) \in [0, -\theta_1] \times [0, -\theta_2]$,

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

$k = 1, \ldots, 4$, and therefore it follows that $\mathbb{E}(Q(R_\theta))$ is equal to $\mathbb{E}\left(Q\left(\tilde{R}_\theta\right)\right)$, with $\tilde{R}_\theta$ restricted to be the set of level-1 rational pure strategies. Hence, by Theorem 3.2, 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 similar to those in equation (B.7). One may also extend ABJ's approach to this problem, and obtain a larger outer region through four moment inequalities similar to those in equation (B.6). Our approach, which yields the sharp identification region, in this simple example requires one to check just 14 inequalities.

As shown in AT (Figure 3), the model with level-1 rationality only places upper bounds on θ_1, θ_2. Figure 9 plots the upper contours of Θ_I, Θ_O^{CT}, and Θ_O^{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_O^{ABJ}, Θ_O^{CT} and Θ_I are, respectively, $(-0.02, -0.02)$, $(-0.15, -0.26)$, and $(-0.54, -0.61)$.

A.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 distributions 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

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

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 $\operatorname{co}[Q(S_\theta)] \subset C_\theta$, 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

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

In our simple two player simultaneous move, complete information, static game of entry, $\mathcal{A}_j = \{0,1\}$, $j = 1,2$, $\mathcal{A} = \{(0,0), (1,0), (0,1), (1,1)\}$. Omitting again the covariates, we 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 6 plots the set C_θ against the possible realizations of $\varepsilon_1, \varepsilon_2$, for this example. Notice that for $\omega \in \Omega$ 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 these regions. For $\omega \in \Omega$ 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).[22] Also in this case one can extend the approaches of ABJ and CT to obtain outer regions defined, respectively, by four and eight moment inequalities similar to those in equations (3.11)-(3.13).

Figure 10 and Table 3 report Θ_I, Θ_O^{CT}, and Θ_O^{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_O^{ABJ} is given by the union of the yellow, red and black areas, and Θ_O^{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_O^{CT} and Θ_O^{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.

B Dual Characterization of the Sharpness Result in the Pure Strategies Case

For a given realization of $(\underline{x}, \varepsilon)$ and value of $\theta \in \Theta$, the set of outcomes generated by pure strategy Nash equilibria[23] is

(B.1) $$Y_\theta(\underline{x}, \varepsilon) = \{y \in \mathcal{Y} : y = a \in \mathcal{A} \text{ and } \pi_j(a_j, a_{-j}, x_j, \varepsilon_j, \theta) \geq \pi_j(\tilde{a}_j, a_{-j}, x_j, \varepsilon_j, \theta)\ \forall \tilde{a}_j \in \mathcal{A}_j\ \forall j\}.$$

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 $\mathcal{Y}$.

The researcher observes the tuple $(y, \underline{x})$, and the random set Y_θ is a function of $\underline{x}$ (and of course ε). Under Assumptions 1-4, and given the covariates $\underline{x}$, the observed outcomes y are consistent with the model if and only if there exists at least one $\theta \in \Theta$ such that $y(\omega) \in Y_\theta(\omega)\ \underline{x} - a.s.$ (i.e., y is a selection of $Y_\theta\ \underline{x} - a.s.$, see Definition 3). A necessary and sufficient condition which guarantees that a random vector $(y, \underline{x})$ is a selection of $(Y_\theta, \underline{x})$ is given by the results of Artstein (1983), Norberg (1992) and Molchanov (2005, Theorem 1.2.20 and

[22] These vertices are

$$\begin{aligned}
\gamma_0(\omega) &= [0\ \ 0\ \ 1\ \ 0]' \\
\gamma_1(\omega) &= \left[1 \quad -\tfrac{\boldsymbol{\varepsilon}_2(\omega)}{\theta_2+\boldsymbol{\varepsilon}_2(\omega)} \quad -\tfrac{\boldsymbol{\varepsilon}_1(\omega)}{\theta_1+\boldsymbol{\varepsilon}_1(\omega)} \quad 0\right]' \left(1 - \tfrac{\boldsymbol{\varepsilon}_1(\omega)}{\theta_1+\boldsymbol{\varepsilon}_1(\omega)} - \tfrac{\boldsymbol{\varepsilon}_2(\omega)}{\theta_2+\boldsymbol{\varepsilon}_2(\omega)}\right)^{-1} \\
\gamma_2(\omega) &= \left[\left(1 + \tfrac{\boldsymbol{\varepsilon}_2(\omega)}{\theta_2}\right)\left(1 + \tfrac{\boldsymbol{\varepsilon}_1(\omega)}{\theta_1}\right) \quad -\tfrac{\boldsymbol{\varepsilon}_2(\omega)}{\theta_2}\left(1 + \tfrac{\boldsymbol{\varepsilon}_1(\omega)}{\theta_1}\right) \quad -\left(1 + \tfrac{\boldsymbol{\varepsilon}_2(\omega)}{\theta_2}\right)\tfrac{\boldsymbol{\varepsilon}_1(\omega)}{\theta_1} \quad \tfrac{\boldsymbol{\varepsilon}_2(\omega)}{\theta_2}\tfrac{\boldsymbol{\varepsilon}_1(\omega)}{\theta_1}\right]' \\
\gamma_3(\omega) &= \left[0 \quad -\tfrac{\boldsymbol{\varepsilon}_2(\omega)}{\theta_2+\boldsymbol{\varepsilon}_2(\omega)} \quad -\tfrac{\boldsymbol{\varepsilon}_1(\omega)}{\theta_1+\boldsymbol{\varepsilon}_1(\omega)} \quad \tfrac{\boldsymbol{\varepsilon}_1(\omega)}{\theta_1+\boldsymbol{\varepsilon}_1(\omega)}\tfrac{\boldsymbol{\varepsilon}_2(\omega)}{\theta_2+\boldsymbol{\varepsilon}_2(\omega)}\right]' \left(\tfrac{\boldsymbol{\varepsilon}_1(\omega)}{\theta_1+\boldsymbol{\varepsilon}_1(\omega)}\tfrac{\boldsymbol{\varepsilon}_2(\omega)}{\theta_2+\boldsymbol{\varepsilon}_2(\omega)} - \tfrac{\boldsymbol{\varepsilon}_1(\omega)}{\theta_1+\boldsymbol{\varepsilon}_1(\omega)} - \tfrac{\boldsymbol{\varepsilon}_2(\omega)}{\theta_2+\boldsymbol{\varepsilon}_2(\omega)}\right)^{-1} \\
\gamma_4(\omega) &= [0\ \ 1\ \ 0\ \ 0]'
\end{aligned}$$

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

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

Section 1.4.8), and amounts to the following:[24]

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

This inequality can be written as $\mathbf{P}(y \in K|\,\underline{x} \in L)\,\mathbf{P}(\underline{x} \in L) \leq \mathbf{P}\{Y_\theta \cap K \neq \emptyset|\,\underline{x} \in L\}\,\mathbf{P}(\underline{x} \in L)$ $\forall K \subset \mathcal{Y}$, $\forall$ compact sets $L \subset \mathcal{X}$ such that $\mathbf{P}(\underline{x} \in L) > 0$, and it is satisfied if and only if

$$\text{(B.2)} \qquad \mathbf{P}(y \in K|\,\underline{x}) \leq \mathbf{P}\{Y_\theta \cap K \neq \emptyset|\,\underline{x}\} \ \ \forall K \subset \mathcal{Y} \ \underline{x} - a.s.$$

Because $\mathcal{Y}$ is finite, all its subsets are compact. The functional $\mathbf{P}\{Y_\theta \cap K \neq \emptyset|\,\underline{x}\}$ on the right-hand side of (B.2) is called the capacity functional of Y_θ given $\underline{x}$. The following definitions formally introduce this functional and a few related ones:

Definition 5 *Let Z be a random closed set in the topological space $\mathbb{F}$, and denote by $\mathcal{K}$ the family of compact subsets of $\mathbb{F}$. The functionals $\mathbf{T}_Z : \mathcal{K} \to [0,1]$, $\mathbf{C}_Z : \mathcal{K} \to [0,1]$, and $\mathbf{I}_Z : \mathcal{K} \to [0,1]$, given by*

$$\mathbf{T}_Z(K) = \mathbf{P}\{Z \cap K \neq \emptyset\}, \quad \mathbf{C}_Z(K) = \mathbf{P}\{Z \subset K\}, \quad \mathbf{I}_Z(K) = \mathbf{P}\{K \subset Z\}\ , \ K \in \mathcal{K},$$

*are said to be, respectively, the **capacity functional** of Z, the **containment functional** of Z, and the **inclusion functional** of Z.*

Denoting by K^c the complement of the sets K, the following relationship holds:

$$\text{(B.3)} \qquad \mathbf{C}_Z(K) = 1 - \mathbf{T}_Z(K^c).$$

Example 3 *Consider again the simple two player entry game in Example 2. Figure 3 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.* □

Notice that given equation (B.3), inequalities (B.2) can be equivalently written as

$$\text{(B.4)} \qquad \mathbf{C}_{Y_\theta|\underline{x}}(K) \leq \mathbf{P}(y \in K|\,\underline{x}) \leq \mathbf{T}_{Y_\theta|\underline{x}}(K) \ \ \forall K \subset \mathcal{Y} \ \underline{x} - a.s.,$$

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 (B.2) 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 (B.4) is superfluous: when the inequalities in (B.4) are used, only subsets $K \subset \mathcal{Y}$ of cardinality up to half of the cardinality of $\mathcal{Y}$ are needed.

We can re-define the identified set of parameters θ as

$$\text{(B.5)} \qquad \Theta_I = \left\{\theta \in \Theta : \mathbf{P}(y \in K|\,\underline{x}) \leq \mathbf{T}_{Y_\theta|\underline{x}}(K) \ \ \forall K \subset \mathcal{Y} \ \underline{x} - a.s.\right\}.$$

[24] Beresteanu and Molinari (2006, 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.

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

(B.6) $$\Theta_O^{ABJ} = \left\{\theta \in \Theta : \mathbf{P}\left\{y = t|\, \underline{x}\right\} \leq \mathbf{T}_{Y_\theta|\underline{x}}(t) \;\; \forall t \in \mathcal{Y} \;\, \underline{x} - a.s.\right\},$$

(B.7) $$\Theta_O^{CT} = \left\{\theta \in \Theta : \mathbf{C}_{Y_\theta|\underline{x}}(t) \leq \mathbf{P}\left\{y = t|\, \underline{x}\right\} \leq \mathbf{T}_{Y_\theta|\underline{x}}(t) \;\; \forall t \in \mathcal{Y} \;\, \underline{x} - a.s.\right\}.$$

Both ABJ and CT acknowledge that the parameter regions they give are not sharp. Comparing the sets in equations (B.6)-(B.7) with the set in equation (B.5), one observes that Θ_O^{ABJ} is obtained applying inequality (B.2) only for $K = \{t\}$ $\forall t \in \mathcal{Y}$. Similarly, Θ_O^{CT} is obtained applying inequality (B.4) only for $K = \{t\}$ (or, equivalently, applying inequality (B.2) for $K = \{t\}$ and $K = \mathcal{Y}\setminus\{t\}$ $\forall t \in \mathcal{Y}$). 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 (B.5) yields the sharp identification region of θ; and (2) if and by how much Θ_I differs from Θ_O^{ABJ} and Θ_O^{CT}. We answer here the first question. Section 3.4.1 answers the second question by looking at a simple example.

Theorem B.1 *Assume that players use only pure strategies, that Assumptions 1-4 are satisfied, and that no other information is available. Then for $\underline{x} - a.s.$ these two conditions are equivalent:*
1. $u'\mathbf{P}(y|\,\underline{x}) \leq \mathbf{E}[h(Q(S_\theta), u)|\,\underline{x}]\ \forall\ u \in \Re^{\kappa_\mathcal{Y}}$,
2. $\mathbf{P}(y \in K|\,\underline{x}) \leq \mathbf{T}_{Y_\theta|\underline{x}}(K)\ \forall K \subset \mathcal{Y}$.

Proof. Beresteanu, Molchanov, and Molinari (2008, Theorem 4.1). ■

As per inequalities (B.2), condition 2 in the above theorem means that y is a selection of Y_θ conditionally on x. As discussed by Galichon and Henry (2008), the distributions of all selections y of a random set Y correspond to the *core* of Y, the latter being the family of probability measures that are dominated by the capacity functional of Y. Galichon and Henry (2008) devote particular attention to the concept of *core determining classes*, namely those sub-families of compact sets K such that the dominance condition on the members of this sub-family guarantees that a candidate probability measure belongs to the core. The core of a random closed set in turn determines the distribution of the set, and therefore the core determining class is exactly the class such that the values of the capacity functional on it determine uniquely the distribution of the random closed set. It is well known in random sets theory that by imposing some restrictions on the family of realizations of a random set it is possible to arrive at some more economical distribution determining classes, see Molchanov (1984), later summarized (and corrected) as Molchanov (2005, Theorem 1.7.7), and Norberg (1989). In particular, if $Y(\omega_1)$ and $Y(\omega_2)$ are ordered by inclusion for almost all $\omega_1, \omega_2 \in \Omega$, then one can always choose intervals as the sets that determine the distribution of Y, and so the core of it. This rather strong assumption, in the form of an ordering relation, is imposed in Galichon and Henry (2008) and brings to a reduction in the size of the core determining class, which exactly corresponds to the setting in the earlier papers by Molchanov (1984) and Norberg (1989).

B.1 On the Number of Inequalities to Be Checked in the Pure Strategies Case

As discussed in Section 3.5, when it is assumed that players play only pure strategies, often there is no need to verify the complete set of $2^{\kappa_\mathcal{Y}}$ inequalities, because many are redundant. Using the insight in Theorem B.1, one can show that the result in equation (3.17) can be restated using the set Y_θ and its capacity functional. In particular, if K_1 and K_2 are two disjoint subsets of $\mathcal{Y}$ such that

$$\{\omega : Y_\theta(\omega) \cap K_1 \neq \emptyset | \underline{x}\} \cap \{\omega : Y_\theta(\omega) \cap K_2 \neq \emptyset | \underline{x}\} = \emptyset, \tag{B.8}$$

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 B.2 below gives a general result which may lead to a dramatic reduction in the number of inequalities to be checked. While its proof is simple, this result is conceptually and practically important.

Theorem B.2 *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 $\omega \in \Omega^i$. If $\mathcal{Y}_1, \ldots, \mathcal{Y}_M$ are disjoint, then it suffices to check (B.2) only for all subsets K such that there is $i = 1, \ldots, M$ for which $K \subseteq \mathcal{Y}_i$.

Proof. Beresteanu, Molchanov, and Molinari (2008, Theorem 5.1). ∎
A simple Corollary of Theorem B.2, the proof of which is omitted, is the following:

Corollary B.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 (B.2) hold if*

$$\mathbf{P}\{Y_\theta = \{t\} | \underline{x}\} \leq \mathbf{P}\{y = t | \underline{x}\} \leq \mathbf{P}\{t \in Y_\theta | \underline{x}\} \tag{B.9}$$

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

An implication of this Corollary is that in a static entry game with two players in which only pure strategies are played, the outer region proposed by CT coincides with ours, and is sharp.[25] 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 equation (3.17) 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 (B.2) for $K = \{(0,0)\}, \{(1,0)\}, \{(0,1)\}, \{(1,1)\}, \{(1,0), (0,1)\}$. On the other hand, the example in Section 3.3.1 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 B.2 is available, it is still possible to use the insight in equation (B.8) to determine which inequalities yield the sharp identification region, by decomposing $\mathcal{Y}$ into subsets such that Y_θ does not jointly hit any two of them with positive probability. One may wonder whether in general the set of inequalities yielding the sharp identification region is different from the set of inequalities used by ABJ or CT. The following result shows that in general the answer to this question is "yes".

[25] 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).

Theorem B.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*

$$\mathbf{I}_{Y_\theta|\underline{x}}(t^1, t^2) > 0. \tag{B.10}$$

(a) If $\mathbf{P}\left\{\left\{t^1, t^2\right\} \cap Y_\theta \neq \emptyset \middle| \underline{x}\right\} < 1$ *for all* $t^1, t^2 \in \mathcal{Y}$*, then there exists a random vector* z *which satisfies inequalities (B.2) for* $K = \{t\}$ $\forall t \in \mathcal{Y}$ *but is not a selection of* Y_θ*.*
(b) If

$$\mathbf{P}\{\kappa_{Y_\theta} > 1|\underline{x}\} > \mathbf{I}_{Y_\theta|\underline{x}}(t^1) + \mathbf{I}_{Y_\theta|\underline{x}}(t^2) - \mathbf{C}_{Y_\theta|\underline{x}}(t^1) - \mathbf{C}_{Y_\theta|\underline{x}}(t^2), \tag{B.11}$$

then there exists a random vector z *which satisfies inequalities (B.2) for* $K = \{t\}$ *and* $K = \mathcal{Y} \setminus \{t\}$ $\forall t \in \mathcal{Y}$ *but is not a selection of* Y_θ*.*

Proof. Beresteanu, Molchanov, and Molinari (2008, Theorems 5.2 and 5.3) ■

These results show that the extra inequalities matter in general, compared to those used by ABJ, and CT, to fully characterize Y_θ and determine if $y \in \mathrm{Sel}(Y_\theta)$. In fact, the assumptions of Theorem B.3-(a) 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. The assumptions of Theorem B.3-(b) 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 3.4.1.

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*, 26(3), 261–310.

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., I. S. Molchanov, and F. Molinari (2008): "Sharp Identification Regions in Games," CeMMAP Working Paper CWP15/08.

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*, 76, 763–814.

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.

Boyd, S., and L. Vandenberghe (2004): *Convex Optimization*. Cambridge University Press, New York.

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.

Canay, I. A. (2008): "EL Inference for Partially Identified Models: Large Deviations Optimality and Bootstrap Validity," mimeo.

Chernozhukov, V., H. Hong, and E. Tamer (2004): "Parameter Set Inference in a Class of Econometric Models," Discussion paper, MIT, Duke University and Northwestern University.

——— (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.

——— (2006): "Market Structure and Multiple Equilibrium in Airline Markets," *Univeristy of Virginia Working Papers*.

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

——— (2008): "Inference in Models with Multiple Equilibria," mimeo.

Horowitz, J. L., C. F. Manski, M. Ponomareva, and J. Stoye (2003): "Computation of Bounds on Population Parameters When the Data Are Incomplete," *Reliable Computing*, 9, 419–440.

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. (1984): "A Generalization of the Choquet Theorem for Random Sets with a Given Class of Realizations," *Theory of Probability and Mathematical Statistics*, 28, 99–106.

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

Molinari, F. (2008): "Partial Identification of Probability Distributions with Misclassified Data," *Journal of Econometrics*, 144(1), 81–117.

Norberg, T. (1989): "Existence Theorems for Measures on Continuous Posets, with Applications to Random Sets Theory," *Mathematica Scandinavica*, 64, 15–51.

——— (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.

Rockafellar, R. (1970): *Convex Analysis*. Princeton University Press.

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

Rosen, A. (2008): "Confidence Sets for Partially Identified Parameters That Satisfy a Finite Number of Moment Inequalities," *Journal of Econometrics*, 146(1), 107–117.

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_O^{ABJ}, Θ_O^{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_O^{ABJ}	Θ_O^{CT}	Θ_I
$\theta_1^\star$	-1.15	$[-2.715, -0.485]$	$[-2.715, -0.585]$	$[-2.205, -0.605]$
			(4.5%)	(28.3%)
$\theta_2^\star$	-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_O^{ABJ}			(16.4%)	(56.5%)

Table 2: Projections of Θ_O^{ABJ}, Θ_O^{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_O^{ABJ}	Θ_O^{CT}	Θ_I
$\theta_{11}^\star$	-0.15	$[-0.154, -0.144]$	$[-0.153, -0.146]$	$[-0.152, -0.147]$
			(27%)	(54%)
$\theta_{21}^\star$	-0.20	$[-0.206, -0.195]$	$[-0.204, -0.197]$	$[-0.203, -0.198]$
			(27%)	(54%)
$\theta_{22}^\star$	-0.10	$[-0.106, -0.096]$	$[-0.104, -0.097]$	$[-0.103, -0.098]$
			(27%)	(54%)

Table 3: Projections of Θ_O^{ABJ}, Θ_O^{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_O^{ABJ}	Θ_O^{CT}	Θ_I
$\theta_1^\star$	-1.15	$[-4.475, -0.485]$	$[-4.475, -0.585]$	$[-4.125, -0.595]$
			(2.5%)	(11.5%)
$\theta_2^\star$	-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_O^{ABJ}			(7.9%)	(23.1%)

Figure 1: Two player entry game. Panel (a): The random set of mixed strategy NE profiles, S_θ, as a function of ε_1,ε_2. Panel (b): The random set of probability distributions over outcome profiles implied by mixed strategy NE, $Q(S_\theta)$, as a function of ε_1,ε_2. Panel (c): The support function in direction u of the random set of probability distributions over outcome profiles implied by mixed strategy NE, $h(Q(S_\theta), u)$, as a function of ε_1,ε_2.

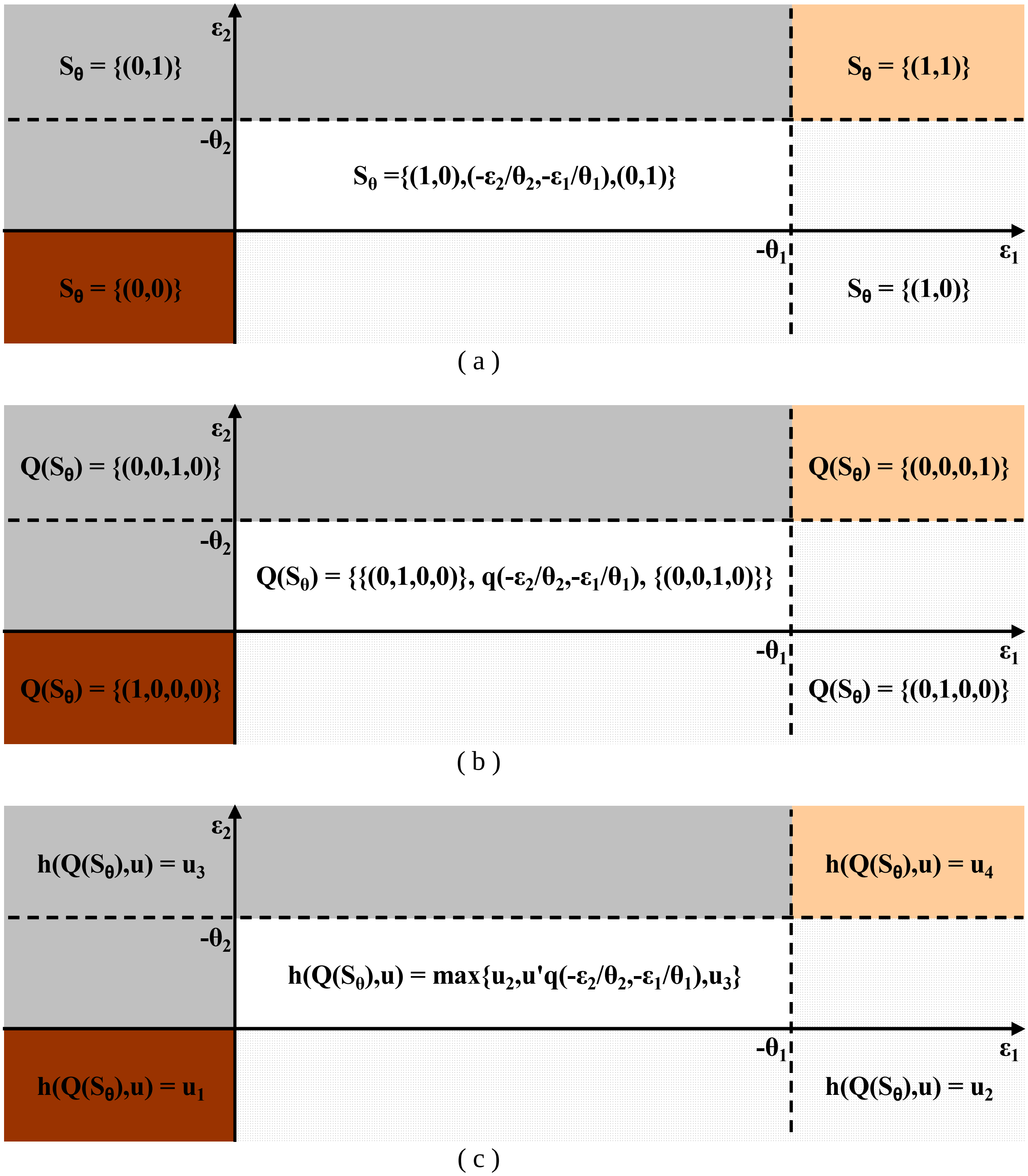

Figure 2: A comparison between the logic behind the approaches of ABJ, CT, and this paper, obtained by projecting in $\Re^2 : \Delta^{\kappa_{\mathcal{Y}}-1}$, $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$, and the hypercubes used by ABJ and CT. A candidate $\theta \in \Theta$ is in Θ_I if $\mathbf{P}\left(y|\,\underline{x}\right)$, the white dot in the picture, belongs to the black ellipses $\mathbb{E}\left(Q\left(S_\theta\right)|\,\underline{x}\right)$, which gives the set of probability distributions consistent with *all* the model's implications. The same θ is in Θ_O^{CT} if $\mathbf{P}\left(y|\,\underline{x}\right)$ belongs to the red region or to the black ellipses, which give the set of probability distributions consistent with the subset of model's implications used by CT. The same θ is in Θ_O^{ABJ} if $\mathbf{P}\left(y|\,\underline{x}\right)$ belongs to the yellow region or to the red region or to the black ellipses, which give the set of probability distributions consistent with the subset of model's implications used by ABJ.

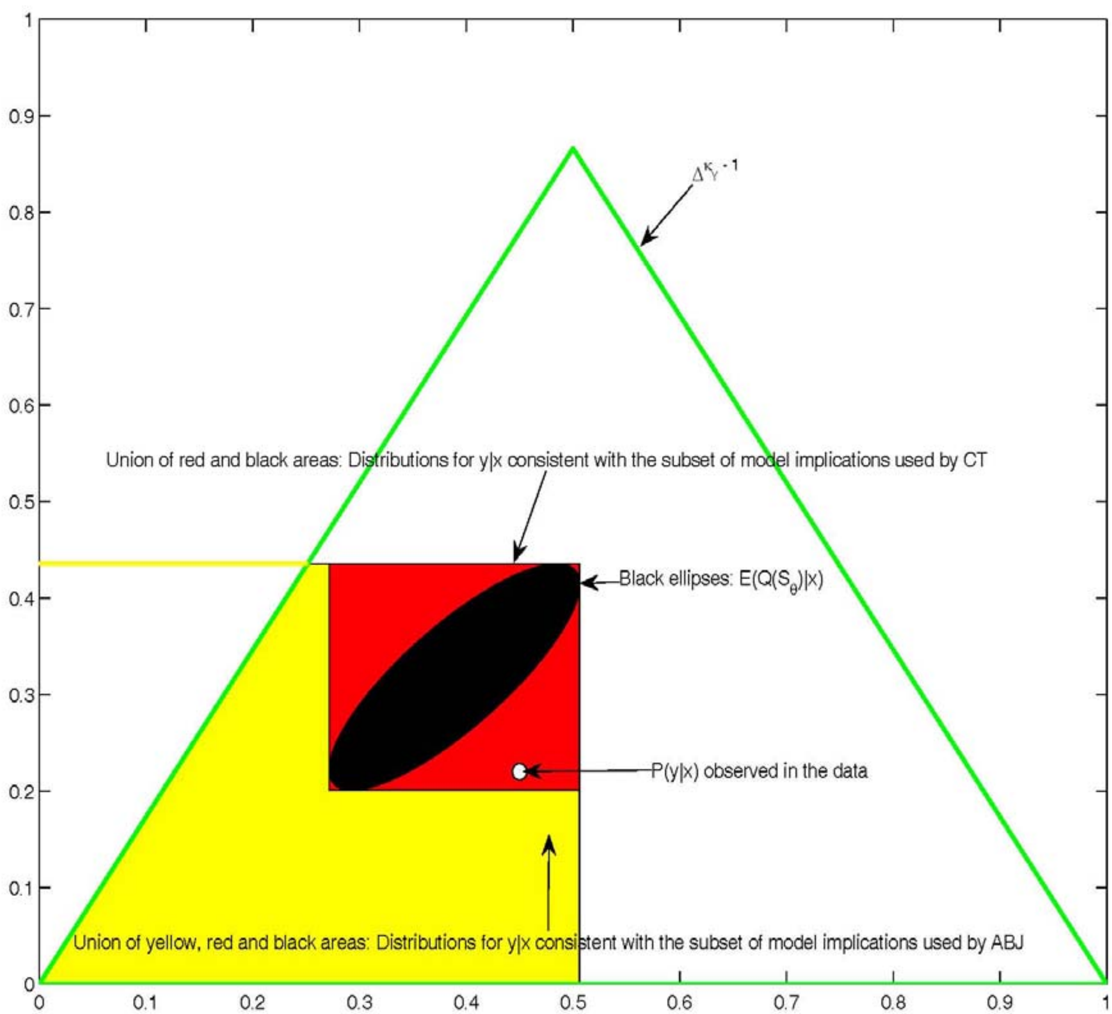

Figure 3: The random set of pure strategy NE profiles, S_θ, and the random set of pure strategy NE outcomes, Y_θ, as a function of $\varepsilon_1, \varepsilon_2$ in a two player entry game. In this simple example the two sets coincide.

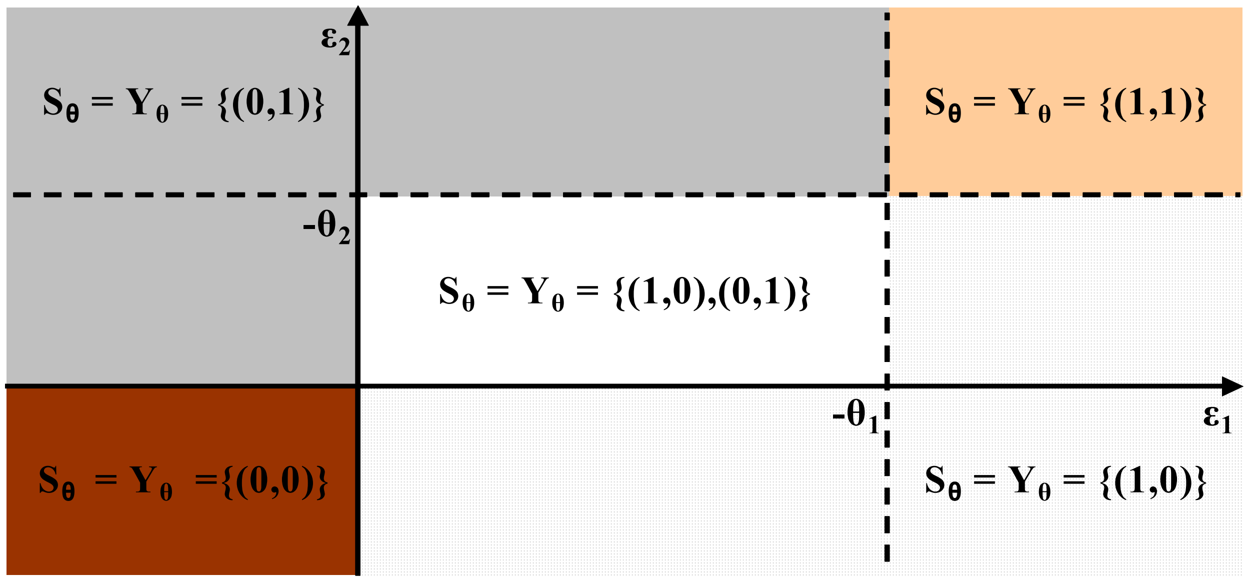

Figure 4: The random set of pure strategy NE outcomes as a function of $\varepsilon_1, \varepsilon_2$ in a four player, two type entry game.

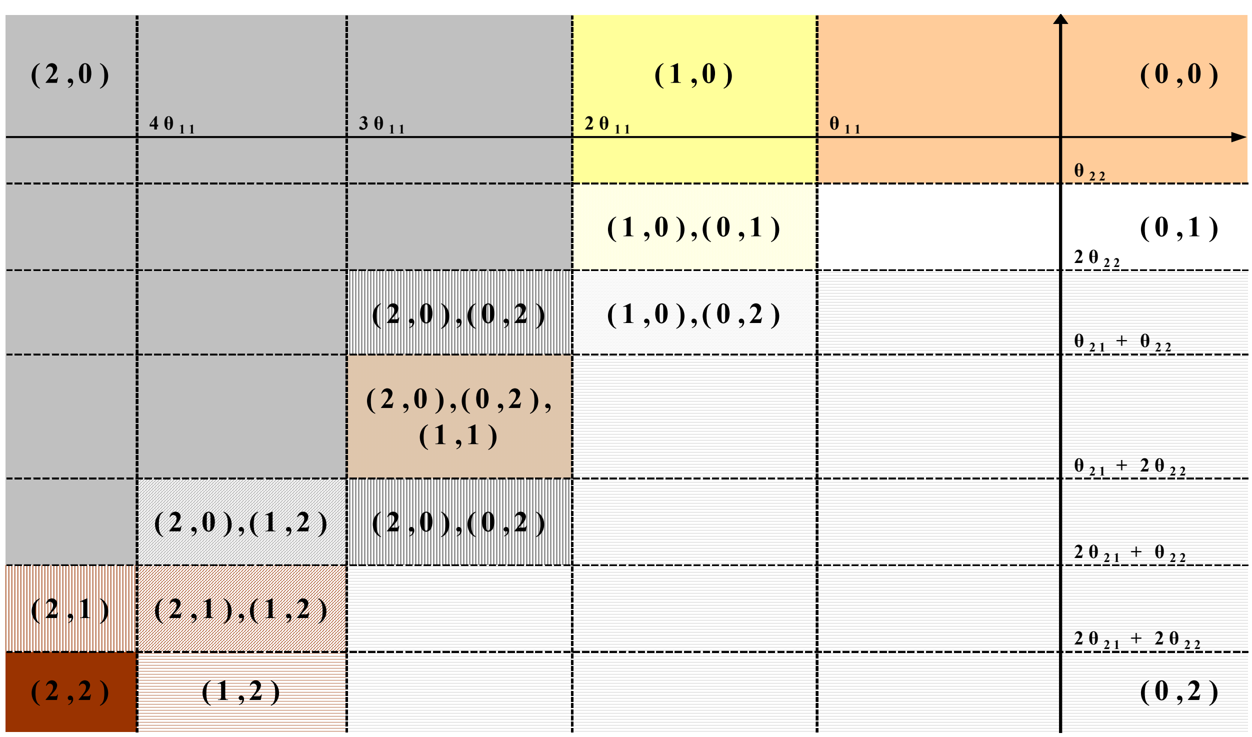

Figure 5: The random set of level-1 rational profiles as a function of $\varepsilon_1, \varepsilon_2$ in a two player entry game.

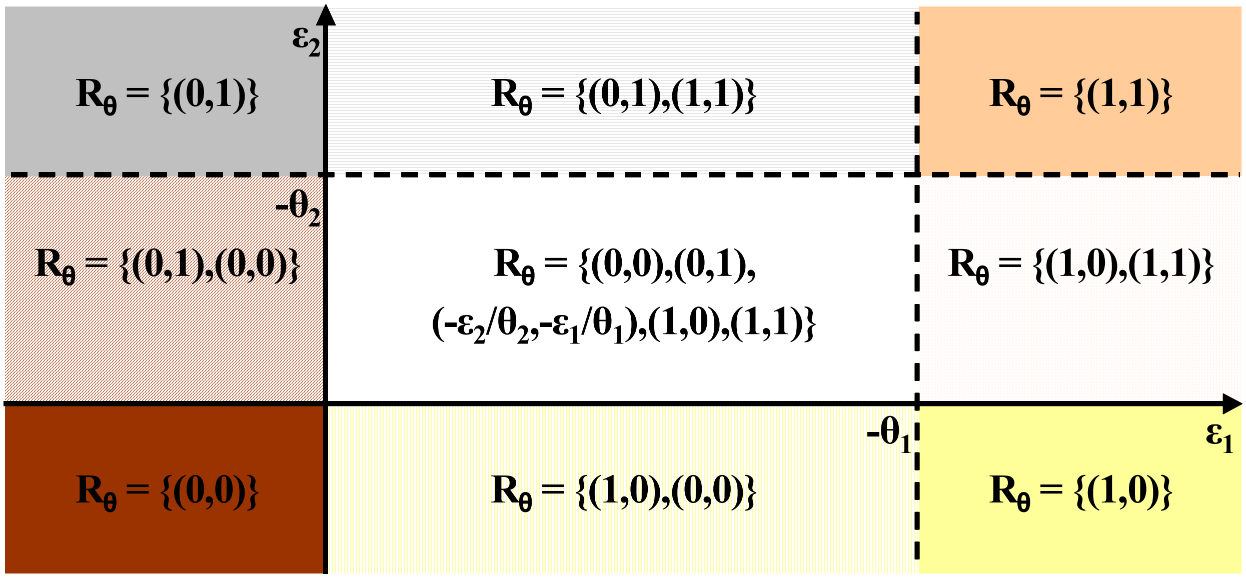

Figure 6: The random set of correlated equilibria as a function of ε_1,ε_2 in a two player entry game. The correlated equilibria $\gamma_1, \gamma_2, \gamma_3$ are defined in Section A.2.

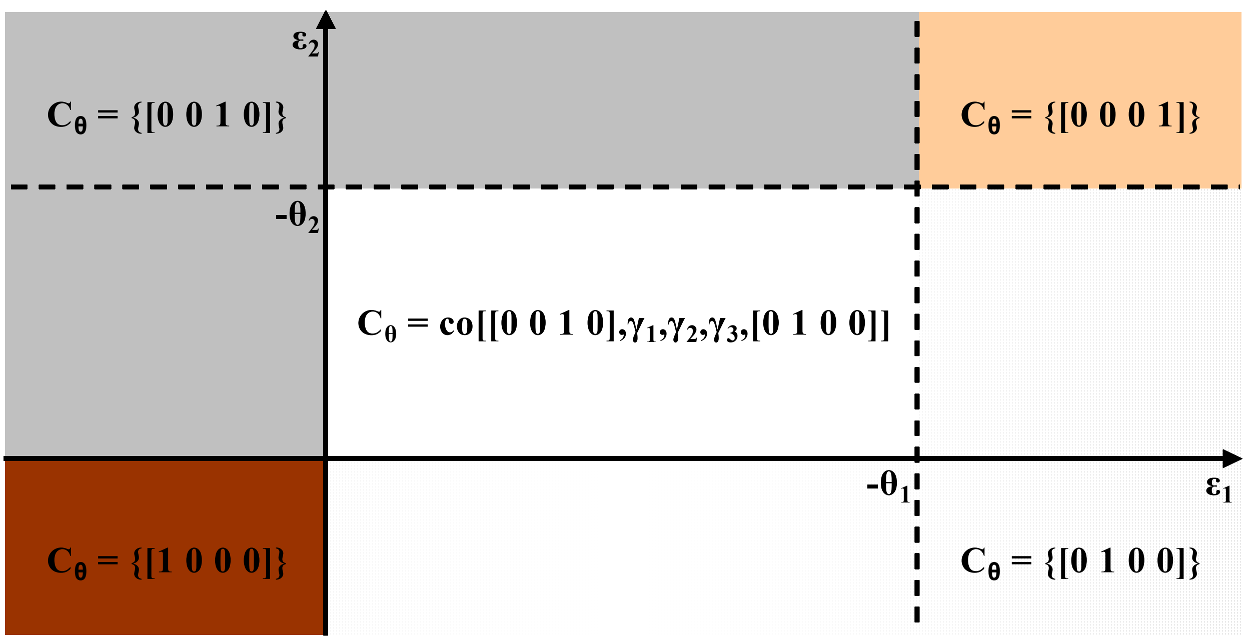

Figure 7: Identification regions in a two player entry game with mixed strategy Nash equilibrium as solution concept.

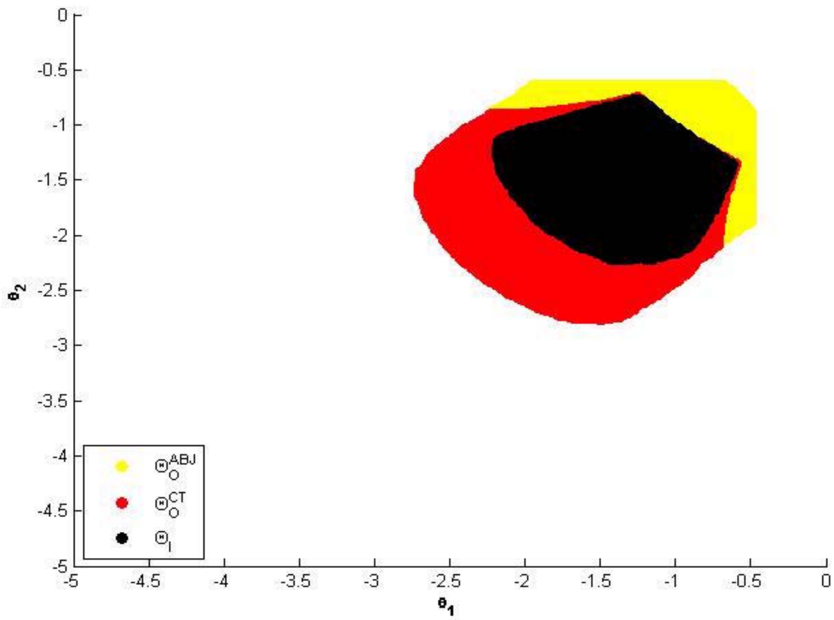

Figure 8: Identification regions in a four player, two type entry game with pure strategy Nash equilibrium as solution concept.

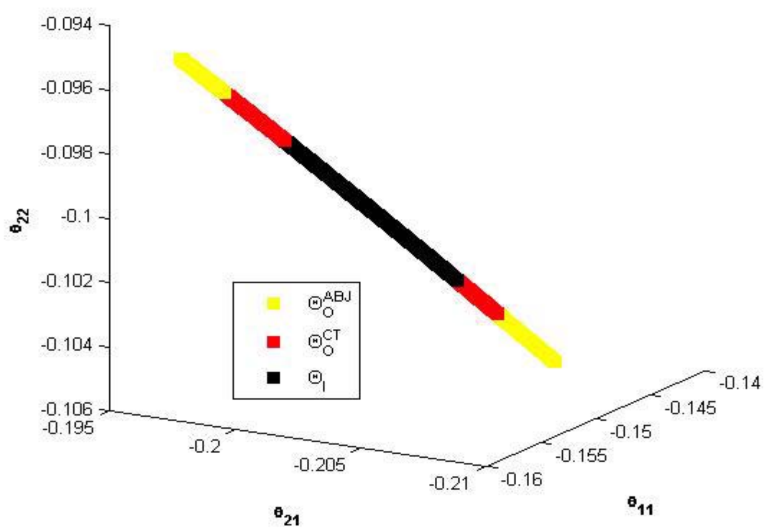

Figure 9: Upper contours of the identification regions in a two player entry game with level-1 rationality as solution concept.

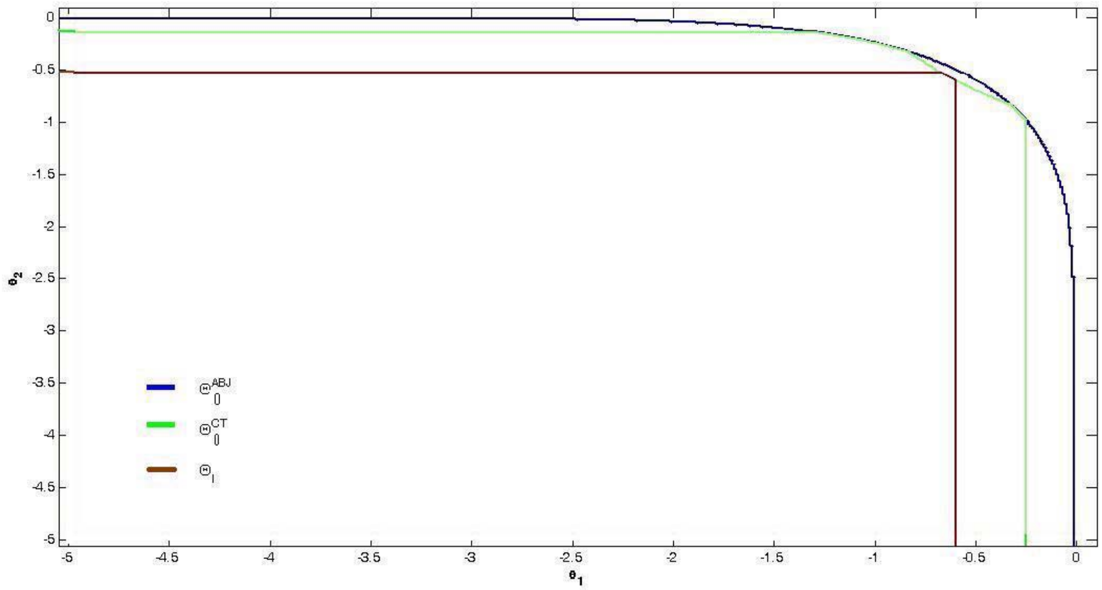

Figure 10: Identification regions in a two player entry game with correlated equilibrium as solution concept.