ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics

http://www.econometricsociety.org/

Econometrica, Vol. 79, No. 6 (November, 2011), 1785–1821

SHARP IDENTIFICATION REGIONS IN MODELS WITH CONVEX MOMENT PREDICTIONS

ARIE BERESTEANU
University of Pittsburgh, Pittsburgh, PA 15260, U.S.A.

ILYA MOLCHANOV
Institute of Mathematical Statistics and Actuarial Science, University of Bern, CH-3012 Bern, Switzerland

FRANCESCA MOLINARI
Cornell University, Ithaca, NY 14853-7601, U.S.A.

The copyright to this Article is held by the Econometric Society. It may be downloaded, printed and reproduced only for educational or research purposes, including use in course packs. No downloading or copying may be done for any commercial purpose without the explicit permission of the Econometric Society. For such commercial purposes contact the Office of the Econometric Society (contact information may be found at the website http://www.econometricsociety.org or in the back cover of *Econometrica*). This statement must be included on all copies of this Article that are made available electronically or in any other format.

Econometrica, Vol. 79, No. 6 (November, 2011), 1785–1821

SHARP IDENTIFICATION REGIONS IN MODELS WITH CONVEX MOMENT PREDICTIONS

BY ARIE BERESTEANU, ILYA MOLCHANOV, AND FRANCESCA MOLINARI[1]

We provide a tractable characterization of the sharp identification region of the parameter vector 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 observable model variables. In short, we call these *models with convex moment predictions*. Examples include static, simultaneous-move finite games of complete and incomplete information in the presence of multiple equilibria; best linear predictors with interval outcome and covariate data; and random utility models of multinomial choice in the presence of interval regressors 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. 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, finite static games, multiple equilibria, random utility models, interval data, best linear prediction.

1. INTRODUCTION

Overview

THIS PAPER PROVIDES a simple, novel, and computationally feasible procedure to determine the *sharp* identification region of the parameter vector θ that characterizes 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 observable variables. In short, throughout the paper, we call these *models with convex moment predictions*. Our use of the

[1]This paper supersedes two earlier working papers (Beresteanu, Molchanov, and Molinari (2008, 2009)). We thank a co-editor and three anonymous reviewers for very useful suggestions that substantially improved the paper. We are grateful to H. Bar, T. Bar, L. Barseghyan, I. Canay, G. Chamberlain, P. Ellickson, Y. Feinberg, K. Hirano, T. Magnac, T. O'Donoghue, M. Peski, A. Rosen, X. Shi, J. Stoye, A. Sweeting, and, especially, L. Blume, D. Easley, and C. Manski for comments. We also thank seminar participants at many institutions and conferences where this paper was presented. We are grateful to M. Grant for advice on the use of the CVX software, which we used to run simulations in MatLab. Financial support from the U.S. NSF Grants SES-0617559 and SES-0922373 (Beresteanu) and SES-0617482 and SES-0922330 (Molinari); from the Swiss National Science Foundation Grants 200021-117606 and 200021-126503 (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.

© 2011 The Econometric Society DOI: 10.3982/ECTA8680

term "model" encompasses econometric frameworks ranging from structural semiparametric models to nonparametric best predictors under square loss. In the interest of clarity of exposition, in this paper we focus on the semiparametric case. We exemplify our methodology by applying it to static, simultaneous-move finite games of complete and incomplete information in the presence of multiple equilibria; best linear predictors with interval outcome and covariate data; and random utility models of multinomial choice in the presence of interval regressors data.

Models with convex moment predictions can be described as follows. For a given value of the parameter vector θ and realization of (a subset of) model variables, the economic model predicts a set of values for a vector of variables of interest. These are the model set-valued predictions, which are *not* necessarily convex. No restriction is placed on the manner in which, in the data generating process, a specific model prediction is selected from this set. When the researcher takes conditional expectations of the resulting elements of this set, the unrestricted process of selection yields a convex set of moments for the model variables—this is the model's *convex set of moment predictions*. If this set were almost surely single valued, the researcher would be able to identify θ by matching the model-implied vector of moments to the one observed in the data. When the model's moment predictions are set-valued, one may find many values for the parameter vector θ which, when coupled with specific selection mechanisms picking one of the model set-valued predictions, generate the same conditional expectation as the one observed in the data. Each of these values of θ is observationally equivalent, and the question becomes how to characterize the collection of observationally equivalent θ's in a tractable manner.

Although previous literature has provided tractable characterizations of the sharp identification region for certain models with convex moment predictions (see, e.g., Manski (2003) for the analysis of nonparametric best predictors under square loss with interval outcome data), there exist many important problems, including the examples analyzed in this paper, in which such a characterization is difficult to obtain. The analyzes of Horowitz, Manski, Ponomareva, and Stoye (2003; HMPS henceforth), and Andrews, Berry, and Jia (2004; ABJ henceforth), and Ciliberto and Tamer (2009; CT henceforth) are examples of research studying, respectively, the identified features of best linear predictors with missing outcome and covariate data, and finite games with multiple pure strategy Nash equilibria. HMPS provided sharp identification regions, but these may have prohibitive computational cost. To make progress not only on identification analysis, but also on finite sample inference, ABJ and CT proposed regions of parameter values which are not sharp.

Establishing whether a conjectured region for the identified features of an incomplete econometric model is sharp is a key step 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 generating process consistent with the maintained assumptions. Examples of sharp identification regions for parameters of incomplete models are given 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 moment 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 moment predictions. That is, this Aumann expectation gives the convex set, implied by the candidate θ, of moments for the relevant variables which are consistent with *all* the model's implications.[2] 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 the 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. Assume that the model is correctly specified. Then θ is in Θ_I if and only if the conditional Aumann expectation (a convex set) of the properly defined random set associated with θ contains the conditional expectation of a properly defined vector of random variables observed in the data (a point). This is because when such a condition is satisfied, there exists a vector of conditional expectations associated with θ that is consistent with all the implications of the model and that coincides with the vector of conditional expectations observed in the data. The methodology that we propose allows us to verify this condition by checking whether the *support function* of such a point is dominated by the support function of the θ-dependent convex set.[3] The latter can be evaluated exactly or approximated by simulation, depending on the complexity of the model. Showing that this dominance holds amounts to checking whether the difference between the support function of a point (a linear function) and the support function of a convex set (a sublinear

[2]We formally define the notion of random closed set in Appendix A and the notion of conditional Aumann expectation in Section 2.

[3]"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, p. 37)). See Rockafellar (1970, Chapter 13) or Schneider (1993, Section 1.7) for a thorough discussion of the support function of a closed convex set and its properties.

function) in a direction given by a vector u attains a maximum of zero as u ranges in the unit ball of appropriate dimension. This amounts to maximizing a superlinear function over a convex set, a task which can be carried out efficiently using algorithms in convex programming (e.g., Boyd and Vandenberghe (2004), Grant and Boyd (2008)).

It is natural to wonder which model with set-valued predictions may *not* belong to the class of models to which our methodology applies. Our approach is specifically tailored toward frameworks where Θ_I can be characterized via conditional or unconditional expectations of observable random vectors and model predictions. Within these models, if restrictions are imposed on the selection process, nonconvex sets of moments may result. We are chiefly interested in the case that no untestable assumptions are imposed on the selection process; therefore, exploring identification in models with nonconvex moment predictions is beyond the scope of this paper.

There are no precedents for our general characterization of the sharp identification region of models with convex moment predictions. However, there is one precedent for the use of the Aumann expectation as a key tool to describe fundamental features of partially identified models. This is the work of Beresteanu and Molinari (2006, 2008), who were the first to illustrate the benefits of using elements of the theory of random sets 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.

In important complementary work, Galichon and Henry (2009a) studied finite games of complete information with multiple *pure* strategy Nash equilibria. For this class of models, they characterized the sharp identification region of θ through 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).[4] They also established that powerful tools of optimal transportation theory can be employed to obtain computational simplifications when the model satisfies certain monotonicity conditions. With pure strategies only, the characterization based on the capacity functional is "dual" to ours, as we formally establish in the Supplemental Material (Beresteanu, Molchanov, and Molinari (2011, Appendix D.2)). It cannot, however, be extended to the general case where mixed strategies are allowed for, as discussed by Galichon and Henry (2009a, Section 4), or to other solution concepts such as, for exam-

[4]Galichon and Henry (2006) used 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. They then defined the notion of "core determining" classes of sets to find a manageable class of sets for which to check that the dominance condition is satisfied. Beresteanu and Molinari (2006, 2008) used Artstein's (1983) result to establish sharpness of the identification region of the parameters of a best linear predictor with interval outcome data.

ple, correlated equilibrium. Our methodology can address these more general game theoretic models.

While our main contribution lies in the identification analysis that we carry out, our characterization 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 sample criterion function. In the Supplemental Material (Beresteanu, Molchanov, and Molinari (2011, Appendix B)), we establish that the methodology of Andrews and Shi (2009) can be applied in our context to obtain confidence sets that uniformly cover each element of the sharp identification region with a prespecified asymptotic probability. Related methods for statistical inference in partially identified models include, among others, Chernozhukov, Hong, and Tamer (2004, 2007), Pakes, Porter, Ho, and Ishii (2006), Beresteanu and Molinari (2008), Rosen (2008), Chernozhukov, Lee, and Rosen (2009), Galichon and Henry (2009b), Kim (2009), Andrews and Soares (2010), Bugni (2010), Canay (2010), Romano and Shaikh (2010), and Ponomareva (2010).

Structure of the Paper

In Section 2, we describe formally the class of econometric models to which our methodology applies and we provide our characterization of the sharp identification region. In Section 3, we analyze in detail the identification problem in static, simultaneous-move finite games of complete information in the presence of multiple mixed strategy Nash equilibria (MSNE), and show how the results of Section 2 can be applied. In Section 4, we show how our methodology can be applied to best linear prediction with interval outcome and covariate data. Section 5 concludes. Appendix A contains definitions taken from random set theory, proofs of the results appearing in the main text, and details concerning the computational issues associated with our methodology (for concreteness, we focus on the case of finite games of complete information).

Appendices B–F are given in the Supplemental Material (Beresteanu, Molchanov, and Molinari (2011)). Appendix B establishes applicability of the methodology of Andrews and Shi (2009) for statistical inference in our class of models. Appendix C shows that our approach easily applies also to finite games of incomplete information, and characterizes Θ_I through a finite number of moment inequalities. Appendix D specializes our results, in the context of complete information games, to the case that players are restricted to use pure strategies only and Nash equilibrium is the solution concept. Also in this case, Θ_I is characterized through a finite number of moment inequalities and further insights are provided on how to reduce the number of inequalities to be checked so as to compute it. Appendix E shows that our methodology is applicable to static simultaneous-move finite games regardless of the solution concept used. Specifically, 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. Appendix F applies the results of Section 2 to the analysis of individual decision making in random utility models of multinomial choice in the presence of interval regressors data.

2. SEMIPARAMETRIC MODELS WITH CONVEX MOMENT PREDICTIONS

Notation: Throughout the paper, we use capital Latin letters to denote sets and random sets. We use lowercase 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. We denote by Δ^{d-1} the unit simplex in $\Re^d$. Given two nonempty sets $A, B \subset \Re^d$, we denote the directed Hausdorff distance from A to B, the Hausdorff distance between A and B, and the Hausdorff norm of B, respectively, by

$$d_H(A, B) = \sup_{a\in A} \inf_{b\in B} \|a - b\|,$$

$$\rho_H(A, B) = \max\{d_H(A, B), d_H(B, A)\}, \quad \|B\|_H = \sup_{b\in B} \|b\|.$$

Outline: In this section, we describe formally the class of econometric models to which our methodology applies and we provide our characterization of the sharp identification region. In Sections 3 and 4, we illustrate how empirically relevant models fit into this general framework. In particular, we show how to verify, for these models, the assumptions listed below.

2.1. *Framework*

Consider an econometric model which specifies a vector z of random variables observable by the researcher, a vector ξ of random variables unobservable by the researcher, and an unknown parameter vector $\theta \in \Theta \subset \Re^p$, with Θ the parameter space. Maintain the following assumptions:

ASSUMPTION 2.1—Probability Space: *The random vectors (z, ξ) are defined on a probability space $(\Omega, \mathfrak{F}, \mathbf{P})$. The σ-algebra $\mathfrak{F}$ is generated by (z, ξ). The researcher conditions her analysis on a sub-σ-algebra of $\mathfrak{F}$, denoted $\mathfrak{G}$, which is generated by a subvector of z. The probability space contains no $\mathfrak{G}$ atoms. Specifically, for all $A \in \mathfrak{F}$ having positive measure, there is a $B \subseteq A$ such that $0 < \mathbf{P}(B|\mathfrak{G}) < \mathbf{P}(A|\mathfrak{G})$ with positive probability.*

ASSUMPTION 2.2—Set-Valued Predictions: *For a given value of θ, the model maps each realization of (z, ξ) to a nonempty closed set $Q_\theta(z, \xi)$ which is a subset of the finite dimensional Euclidean space $\Re^d$. The functional form of this map is known to the researcher.*

ASSUMPTION 2.3—Absolutely Integrable Random Closed Set: *For every compact set C in $\Re^d$ and all $\theta \in \Theta$,*

$$\left\{\omega \in \Omega : Q_\theta(z(\omega), \xi(\omega)) \cap C \neq \emptyset\right\} \in \mathfrak{F}.$$

Moreover, $\mathbf{E}(\|Q_\theta(z, \xi)\|_H) < \infty$.

Assumption 2.1 requires the probability space to be nonatomic with respect to the σ-algebra $\mathfrak{G}$ on which the researcher conditions her analysis. This technical assumption is not restrictive for most economic applications, as we show in Sections 3 and 4. For example, it is satisfied whenever the distribution of ξ conditional on $\mathfrak{G}$ is continuous.

Assumption 2.2 requires the model to have set-valued predictions (models with singleton predictions are a special case of the more general ones analyzed here). As we further explain below, the set $Q_\theta(z, \xi)$ is the fundamental object that we use to relate the convex set of model moment predictions to the observable moments of random vectors. In Sections 3 and 4, we provide examples of how $Q_\theta(z, \xi)$ needs to be constructed in specific applications to exploit all the model information.

Assumption 2.3 is a measurability condition, requiring $Q_\theta(z, \xi)$ to be an *integrably bounded random closed set*; see Definitions A.1 and A.2 in Appendix A. It guarantees that any ($\mathfrak{F}$-measurable) random vector q such that $q(\omega) \in Q_\theta(z(\omega), \xi(\omega))$ a.s. is absolutely integrable.

In what follows, for ease of notation, we write the set $Q_\theta(z, \xi)$ and its realizations, respectively, as Q_θ and $Q_\theta(\omega) \equiv Q_\theta(z(\omega), \xi(\omega))$, $\omega \in \Omega$, omitting the explicit reference to z and ξ. The researcher wishes to learn θ from the observed distribution of z. Because the model makes set-valued predictions, we maintain the following assumption:

ASSUMPTION 2.4—Selected Prediction: *The econometric model can be augmented with a selection mechanism which selects one of the model predictions, yielding a map ψ which depends on z and ξ, may depend on θ, and satisfies the following conditions*:

(i) $\psi(z(\omega), \xi(\omega), \theta) \in Q_\theta(\omega)$ *for almost all* $\omega \in \Omega$.

(ii) $\psi(z(\omega), \xi(\omega), \theta)$ *is* $\mathfrak{F}$*-measurable for all* $\theta \in \Theta$.

Assumption 2.4 requires that the econometric model can be "completed" with an unknown selection mechanism. Economic theory often provides no guidance on the form of the selection mechanism, which therefore we leave completely unspecified. For each $\omega \in \Omega$, the process of selection results in a random element ψ which takes values in Q_θ, that is, is a model's selected prediction.[5] The map ψ is unknown and constitutes a nonparametric component

[5]For expository clarity, we observe that even for $\omega_1 \neq \omega_2$ such that $z(\omega_1) = z(\omega_2)$ and $\xi(\omega_1) = \xi(\omega_2)$, $\psi(z(\omega_1), \xi(\omega_1), \theta)$ may differ from $\psi(z(\omega_2), \xi(\omega_2), \theta)$.

of the model; it may depend on unobservable variables even after conditioning on observable variables. We insert θ as an argument of ψ to reflect the fact that Assumption 2.4(i) requires ψ to belong to the θ-dependent set Q_θ.

In this paper, we restrict attention to models where the set of observationally equivalent parameter vectors θ, denoted Θ_I, can be characterized by a finite number of conditional expectations of observable random vectors and model predictions. One may find many values for the parameter vector θ which, when coupled with maps ψ satisfying Assumption 2.4, generate the same moments as those observed in the data. Hence, we assume that Θ_I can be characterized through selected predictions as follows.

ASSUMPTION 2.5—Sharp Identification Region: *Given the available data and Assumptions* 2.1–2.3, *the sharp identification region of* θ *is*

$$\Theta_I = \big\{\theta \in \Theta : \exists \psi(z, \xi, \theta) \textit{ satisfying Assumption } 2.4, \\ s.t.\ \mathbf{E}(w(z)|\mathfrak{G}) = \mathbf{E}(\psi(z, \xi, \theta)|\mathfrak{G})\ a.s.\big\}, \tag{2.1}$$

where $w(\cdot)$ *is a known function mapping* z *into vectors in* $\Re^d$ *and* $\mathbf{E}(w(z)|\mathfrak{G})$ *is identified by the data.*

The process of "unrestricted selection" yielding ψ's satisfying Assumption 2.4 builds all possible mixtures of elements of Q_θ. When one takes expectations of these mixtures, the resulting set of expectations is the *convex set of moment predictions*:

$$\big\{\mathbf{E}(\psi(z, \xi, \theta)|\mathfrak{G}) : \psi(z, \xi, \theta) \text{ satisfies Assumption } 2.4\big\}.$$

Convexity of this set is formally established in the next section.

Using the notion of selected prediction, Assumption 2.5 characterizes abstractly the sharp identification region of a large class of incomplete econometric models in a fairly intuitive manner. This characterization builds on previous ones given by Berry and Tamer (2007) and Tamer (2010, Section 3). However, because ψ is a rather general random function, it may constitute an infinite dimensional nuisance parameter, which creates great difficulties for the computation of Θ_I and for inference. In this paper, we provide a complementary approach based on tools of random set theory. We characterize Θ_I by avoiding altogether the need to deal with ψ, thereby contributing to a stream of previous literature which has provided tractable characterizations of sharp identification regions without making any reference to the selection mechanism or the selected prediction (see, e.g., Manski (2003) and Manski and Tamer (2002)).

2.2. *Representation Through Random Set Theory*

As suggested by Aumann (1965), one can think of a random closed set (or correspondence in Aumann's work) as a bundle of random variables—its *measurable selections* (see Definition A.3 in Appendix A). We follow this idea and denote by $\text{Sel}(Q_\theta)$ the collection of $\mathfrak{F}$-measurable random elements q with values in $\Re^d$ such that $q(\omega) \in Q_\theta(\omega)$ for almost all $\omega \in \Omega$. As it turns out, there is not just a simple assonance between "selected prediction" and "measurable selection." Our first result establishes a one-to-one correspondence between them.

LEMMA 2.1: *Let Assumptions 2.1–2.3 hold. For any given $\theta \in \Theta$, $q \in \text{Sel}(Q_\theta)$ if and only if there exists a selected prediction $\psi(z, \xi, \theta)$ satisfying Assumption 2.4, such that $q(\omega) = \psi(z(\omega), \xi(\omega), \theta)$ for almost all $\omega \in \Omega$.*

The definition of the sharp identification region in Assumption 2.5 indicates that one needs to take conditional expectations of the elements of $\text{Sel}(Q_\theta)$. Observe that by Assumption 2.3, Q_θ is an integrably bounded random closed set and, therefore, all its selections are integrable. Hence, we can define the conditional Aumann expectation (Aumann (1965)) of Q_θ as

$$\mathbb{E}(Q_\theta|\mathfrak{G}) = \{\mathbf{E}(q|\mathfrak{G}) : q \in \text{Sel}(Q_\theta)\},$$

where the notation $\mathbb{E}(\cdot|\mathfrak{G})$ denotes the conditional Aumann expectation of the random set in parentheses, while we reserve the notation $\mathbf{E}(\cdot|\mathfrak{G})$ 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 $\mathfrak{F}$ contains no $\mathfrak{G}$ atoms, and because the random set Q_θ takes its realizations in a subset of the finite dimensional space $\Re^d$, it follows from Theorem 1.2 of Dynkin and Evstigneev (1976) and from Theorem 2.1.24 of Molchanov (2005) that $\mathbb{E}(Q_\theta|\mathfrak{G})$ is a closed convex set a.s., such that $\mathbb{E}(Q_\theta|\mathfrak{G}) = \mathbb{E}(\text{co}[Q_\theta]|\mathfrak{G})$, with $\text{co}[\cdot]$ the convex hull of the set in square brackets.

Our second result establishes that $\mathbb{E}(Q_\theta|\mathfrak{G})$ coincides with the convex set of the model's moment predictions:

LEMMA 2.2: *Let Assumptions 2.1–2.3 hold. For any given $\theta \in \Theta$,*

$$\mathbb{E}(Q_\theta|\mathfrak{G}) = \big\{\mathbf{E}(\psi(z, \xi, \theta)|\mathfrak{G}) : \psi(z, \xi, \theta) \textit{ satisfies Assumption 2.4}\big\},$$

and therefore the latter set is convex.

Hence, the set of observationally equivalent parameter values in Assumption 2.5 can be written as

$$\Theta_I = \big\{\theta \in \Theta : \mathbf{E}(w(z)|\mathfrak{G}) \in \mathbb{E}(Q_\theta|\mathfrak{G}) \text{ a.s.}\big\}.$$

The fundamental result of this paper provides two tractable characterizations of the sharp identification region Θ_I.

THEOREM 2.1: *Let Assumptions 2.1–2.5 be satisfied. Let* $h(Q_\theta, u) \equiv \sup_{q \in Q_\theta} u'q$ *denote the support function of* Q_θ *in direction* $u \in \Re^d$. *Then*

$$\Theta_I = \left\{\theta \in \Theta : \max_{u \in B}\big(u'\mathbf{E}(w(z)|\mathfrak{G}) - \mathbf{E}[h(Q_\theta, u)|\mathfrak{G}]\big) = 0 \; a.s.\right\} \tag{2.2}$$

$$= \left\{\theta \in \Theta : \int_B \big(u'\mathbf{E}(w(z)|\mathfrak{G}) - \mathbf{E}[h(Q_\theta, u)|\mathfrak{G}]\big)_+ \, d\mathcal{U} = 0 \; a.s.\right\}, \tag{2.3}$$

where $B = \{u \in \Re^d : \|u\| \leq 1\}$, $\mathcal{U}$ *is any probability measure on* B *with support equal to* B, *and, for any* $a \in \Re$, $(a)_+ = \max\{0, a\}$.

PROOF: The equivalence between equations (2.2) and (2.3) follows immediately, observing that the integrand in equation (2.3) is continuous in u and both conditions inside the curly brackets are satisfied if and only if

$$u'\mathbf{E}(w(z)|\mathfrak{G}) - \mathbf{E}[h(Q_\theta, u)|\mathfrak{G}] \leq 0 \quad \forall u \in B \text{ a.s.} \tag{2.4}$$

To establish sharpness, it suffices to show that for a given $\theta \in \Theta$, expression (2.4) holds if and only if $\theta \in \Theta_I$ as defined in equation (2.1). Take $\theta \in \Theta$ such that expression (2.4) holds. Theorem 2.1.47(iv) in Molchanov (2005) assures that

$$\mathbf{E}[h(Q_\theta, u)|\mathfrak{G}] = h(\mathbb{E}(Q_\theta|\mathfrak{G}), u) \quad \forall u \in \Re^d \text{ a.s.} \tag{2.5}$$

Recalling that the support function is positive homogeneous, equation (2.4) holds if and only if

$$u'\mathbf{E}(w(z)|\mathfrak{G}) \leq h(\mathbb{E}(Q_\theta|\mathfrak{G}), u) \quad \forall u \in \Re^d \text{ a.s.} \tag{2.6}$$

Standard arguments in convex analysis (see, e.g., Rockafellar (1970, Theorem 13.1)) assure that equation (2.6) holds if and only if $\mathbf{E}(w(z)|\mathfrak{G}) \in \mathbb{E}(Q_\theta|\mathfrak{G})$ a.s., and, therefore, by Lemma 2.2, $\theta \in \Theta_I$. Conversely, take $\theta \in \Theta_I$ as defined in equation (2.1). Then there exists a selected prediction ψ satisfying Assumption 2.4, such that $\mathbf{E}(w(z)|\mathfrak{G}) = \mathbf{E}(\psi(z, \xi, \theta)|\mathfrak{G})$. By Lemma 2.2 and the above argument, it follows that expression (2.4) holds. *Q.E.D.*

It is well known (e.g., Rockafellar (1970, Chapter 13), Schneider (1993, Section 1.7)) that the support function of a nonempty closed convex set is a continuous convex sublinear function.[6] This holds also for the support function of

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

the convex set of moment predictions. However, calculating this set is computationally prohibitive in many cases. The fundamental simplification comes from equation (2.5), which assures that one can work directly with the conditional expectation of $h(Q_\theta, u)$. This expectation is quite straightforward to compute. Hence, the characterization in equation (2.2) is computationally very attractive, because for each candidate $\theta \in \Theta$, it requires maximizing an easy-to-compute superlinear, hence concave, function over a convex set and checking whether 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, the book by Boyd and Vandenberghe (2004) and the MatLab software for disciplined convex programming CVX by Grant and Boyd (2010). Similarly, the characterization in equation (2.3) can be implemented by calculating integrals of concave functions over a convex set, a task which can be carried out in random polynomial time (see, e.g., Dyer, Frieze, and Kannan (1991) and Lovász and Vempala (2006)).

REMARK 2.1: Using the method proposed by Andrews and Shi (2009), expression (2.4) can be transformed, using appropriate instruments, into a set of unconditional moment inequalities indexed by the instruments and by $u \in B$, even when the conditioning variables have a continuous distribution. Equations (2.2) and (2.3) can be modified accordingly to yield straightforward criterion functions which are minimized by every parameter in the sharp identification region. When faced with a finite sample of data, one can obtain a sample analog of these criterion functions by replacing the unconditional counterpart of the moment $u'\mathbf{E}(w(z)|\mathfrak{G}) - \mathbf{E}[h(Q_\theta, u)|\mathfrak{G}]$ with its sample analog. The resulting statistics can be shown to correspond, respectively, to the Kolmogorov–Smirnov (KS) and the Cramér–von Mises (CvM) statistics introduced by Andrews and Shi (2009; see their equations (3.6), (3.7), and (3.8), and their Section 9). When the assumptions imposed by Andrews and Shi are satisfied, one can obtain confidence sets that have correct uniform asymptotic coverage probability for the true parameter vector by inverting the KS or the CvM tests. Under mild regularity conditions, these assumptions are satisfied using our characterization, because our moment function in expression (2.4) is Lipschitz in u. In Appendix B of the Supplemental Material, we formally establish this for the models in Sections 3 and 4.

3. APPLICATION I: FINITE GAMES OF COMPLETE INFORMATION

3.1. *Model Setup*

We consider simultaneous-move games of complete information (normal form games) in which each player has a finite set of actions (pure strategies) $\mathcal{Y}_j$, $j = 1, \ldots, J$, with J the number of players. Let $t = (t_1, \ldots, t_J) \in \mathcal{Y}$ denote a generic vector specifying an action for each player, with $\mathcal{Y} = \times_{j=1}^{J} \mathcal{Y}_j$ and

$\mathcal{Y}_{-j} = \times_{i \neq j} \mathcal{Y}_i$. Let $y = (y_1, \ldots, y_J)$ denote a (random) vector specifying the action chosen by each player; observe that the realizations of y are in $\mathcal{Y}$. Let $\pi_j(t_j, t_{-j}, x_j, \varepsilon_j, \theta)$ denote the payoff function for player j, where t_{-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. Let $\sigma_j : \mathcal{Y}_j \to [0, 1]$ denote the mixed strategy for player j that assigns to each action $t_j \in \mathcal{Y}_j$ a probability $\sigma_j(t_j) \geq 0$ that it is played, with $\sum_{t_j \in \mathcal{Y}_j} \sigma_j(t_j) = 1$ for each $j = 1, \ldots, J$. Let $\Sigma(\mathcal{Y}_j)$ denote the mixed extension of $\mathcal{Y}_j$ and let $\Sigma(\mathcal{Y}) = \times_{j=1}^{J} \Sigma(\mathcal{Y}_j)$. With the usual slight abuse of notation, denote by $\pi_j(\sigma_j, \sigma_{-j}, x_j, \varepsilon_j, \theta)$ the expected payoff associated with the mixed strategy profile $\sigma = (\sigma_1, \ldots, \sigma_J)$. With respect to the general notation used in Section 2, $z = (y, \underline{x})$, $\xi = \varepsilon$, $\mathfrak{F}$ is the σ-algebra generated by $(y, \underline{x}, \varepsilon)$, and $\mathfrak{G}$ is the σ-algebra generated by $\underline{x}$. We formalize our assumptions on the games and sampling processes as follows. These assumptions are fairly standard in the literature.[7]

ASSUMPTION 3.1: (i) *The set of outcomes of the game $\mathcal{Y}$ is finite. Each player j has $\kappa_{\mathcal{Y}_j} \geq 2$ pure strategies to choose from. The number of players is $J \geq 2$.*

(ii) *The observed outcome of the game results from static, simultaneous-move, Nash play.*

(iii) *The parametric form of the payoff functions $\pi_j(t_j, t_{-j}, x_j, \varepsilon_j, \theta)$, $j = 1, \ldots, J$, is known and for a known action $\bar{t}$, it is normalized to $\pi_j(\bar{t}_j, \bar{t}_{-j}, x_j, \varepsilon_j, \theta) = 0$ for each j. The payoff functions are continuous in x_j and ε_j. The parameter space Θ is compact.*

In the above assumptions, continuity is needed to establish measurability and closedness of certain sets. A location normalization is needed because if we add a constant to the payoff of each action, the set of equilibria does not change.

ASSUMPTION 3.2: *The econometrician observes data that identify $\mathbf{P}(y|\underline{x})$. The observed matrix of payoff shifters $\underline{x}$ comprises the nonredundant elements of x_j, $j = 1, \ldots, J$. The unobserved random vector $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_J)$ has a continuous conditional distribution function $F_\theta(\varepsilon|\underline{x})$ that is known up to a finite dimensional parameter vector that is part of θ.*

[7]We assume that players' actions and the outcomes observable by the econometrician coincide. This is a standard assumption in the literature; see, for example, ABJ, CT, Berry and Tamer (2007), and Bajari, Hong, and Ryan (2010). Our results, however, apply to the more general case that the strategy profiles determine the outcomes observable by the econometrician through an outcome rule known by the econometrician, as we illustrate with a simple example in Appendix D.1 in the Supplemental Material. Of course, the outcome rule needs to satisfy assumptions guaranteeing that it conveys some information about players' actions.

REMARK 3.1: Under Assumption 3.2, Assumption 2.1 is satisfied.

It is well known that the games and sampling processes satisfying Assumptions 3.1 and 3.2 may lead to multiple MSNE and partial identification of the model parameters; see, for example, Berry and Tamer (2007) for a thorough discussion of this problem. To achieve point identification, Bjorn and Vuong (1985), Bresnahan and Reiss (1987, 1990, 1991), Berry (1992), Mazzeo (2002), Tamer (2003), and Bajari, Hong, and Ryan (2010), for example, add assumptions concerning the nature of competition, heterogeneity of firms, availability of covariates with sufficiently large support and/or instrumental variables, and restrictions on the selection mechanism which, in the data generating process, determines the equilibrium played in the regions of multiplicity.[8]

We show that the models considered in this section satisfy Assumptions 2.1–2.5 and, therefore, our methodology gives a computationally feasible characterization of Θ_I. Our approach does not impose any assumption on the nature of competition, on the form of heterogeneity across players, or on the selection mechanism. It does not require availability of covariates with large support or instruments, but fully exploits their identifying power if they are present.

3.2. *The Sharp Identification Region*

For a given realization of $(\underline{x}, \varepsilon)$, the mixed strategy profile $\sigma = (\sigma_1, \dots, \sigma_J)$ constitutes a Nash equilibrium if $\pi_j(\sigma_j, \sigma_{-j}, x_j, \varepsilon_j, \theta) \geq \pi_j(\tilde{\sigma}_j, \sigma_{-j}, x_j, \varepsilon_j, \theta)$ for all $\tilde{\sigma}_j \in \Sigma(\mathcal{Y}_j)$ and $j = 1, \dots, J$. Hence, for a given realization of $(\underline{x}, \varepsilon)$, we define the θ-dependent set of MSNE as

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

EXAMPLE 3.1: Consider a simple two player entry game similar to the one in Tamer (2003), omit the covariates, and assume that players' payoffs are given by $\pi_j = t_j(t_{-j}\theta_j + \varepsilon_j)$, where $t_j \in \{0, 1\}$ 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 $\varepsilon_1, \varepsilon_2$.

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. Proposition 3.1 establishes that the set S_θ is a random closed set in $\Sigma(\mathcal{Y})$.

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

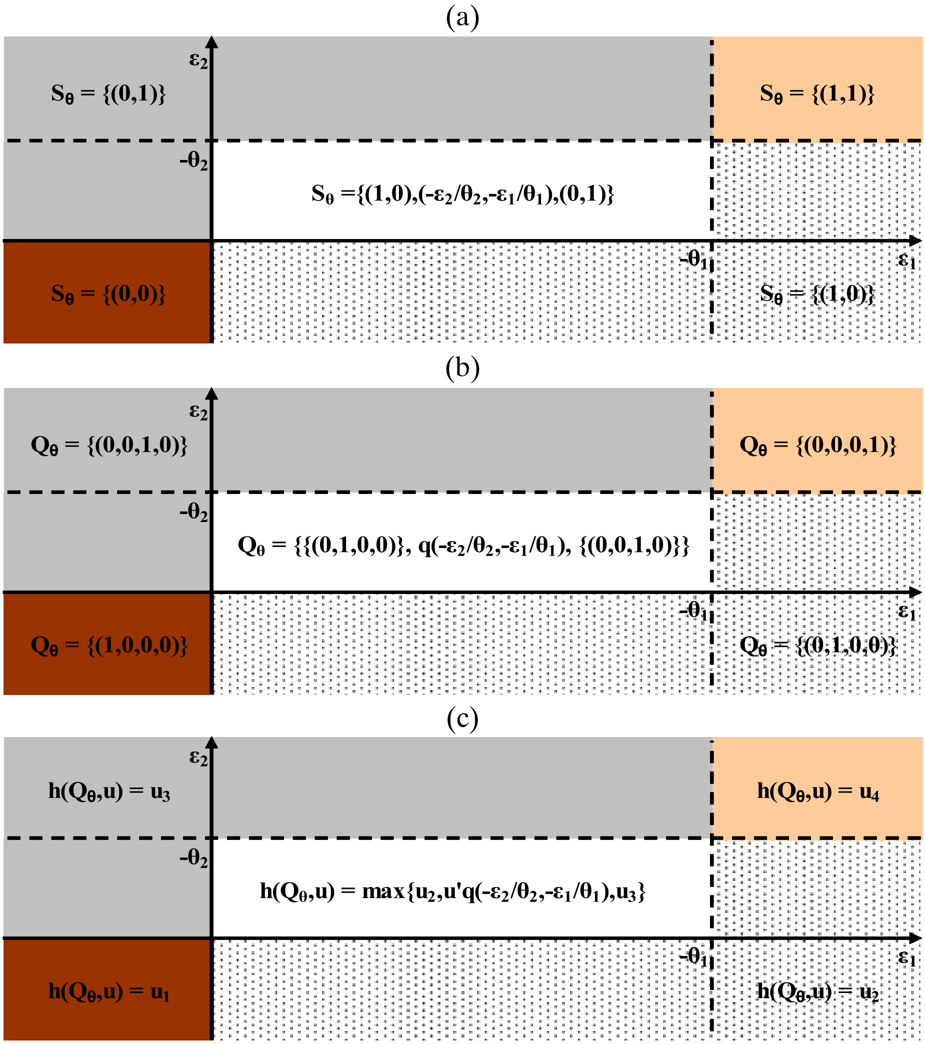

FIGURE 1.—Two player entry game. (a) The random set of mixed strategy NE profiles, S_θ, as a function of ε_1, ε_2. (b) The random set of probability distributions over outcome profiles implied by mixed strategy NE, Q_θ, as a function of ε_1, ε_2. (c) The support function in direction u of the random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

PROPOSITION 3.1: *Let Assumption 3.1 hold. Then the set S_θ is a random closed set in $\Sigma(\mathcal{Y})$ as per Definition A.1 in Appendix A.*

For a given $\theta \in \Theta$ and $\omega \in \Omega$, with some abuse of notation, we denote by $\sigma_j(\omega) : \mathcal{Y}_j \to [0, 1]$ the mixed strategy that assigns to each action $t_j \in \mathcal{Y}_j$ a probability $\sigma_j(\omega, t_j) \geq 0$ that it is played, with $\sum_{t_j \in \mathcal{Y}_j} \sigma_j(\omega, t_j) = 1,\ j = 1, \ldots, J$. We let $\sigma(\omega) \equiv (\sigma_1(\omega), \ldots, \sigma_J(\omega)) \in S_\theta(\omega)$ denote one of the admissible mixed strategy Nash equilibrium profiles (taking values in $\Sigma(\mathcal{Y})$) associated with the realizations $\underline{x}(\omega)$ and $\varepsilon(\omega)$. The resulting random elements σ are the selections of S_θ. We denote the collection of these selections by $\mathrm{Sel}(S_\theta)$; see Definition A.3 in Appendix A.

EXAMPLE 3.1—Continued: Consider the set S_θ plotted in Figure 1(a). Let $\Omega^M = \{\omega \in \Omega : \varepsilon(\omega) \in [0, -\theta_1] \times [0, -\theta_2]\}$. 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

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

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

Index the set $\mathcal{Y} = \bigtimes_{j=1}^J \mathcal{Y}_j$ in some (arbitrary) way such that $\mathcal{Y} = \{t^1, \ldots, t^{\kappa_{\mathcal{Y}}}\}$ and $t^k \equiv (t_1^k, \ldots, t_J^k)$, $k = 1, \ldots, \kappa_{\mathcal{Y}}$. Then for a given parameter value $\theta \in \Theta$ and realization $\sigma(\omega)$, $\omega \in \Omega$, of a selection $\sigma \in \mathrm{Sel}(S_\theta)$, the implied probability that y is equal to t^k is given by $\prod_{j=1}^J \sigma_j(\omega, t_j^k)$. Hence, we can use a selection $\sigma \in \mathrm{Sel}(S_\theta)$ to define a random vector $q(\sigma)$ whose realizations have coordinates

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

By construction, the random point $q(\sigma)$ is an element of $\Delta^{\kappa_{\mathcal{Y}}-1}$. For given $\omega \in \Omega$, each vector $([q(\sigma(\omega))]_k, k = 1, \ldots, \kappa_{\mathcal{Y}})$ is the multinomial distribution over outcomes of the game (a J-tuple of actions) determined by the mixed strategy equilibrium $\sigma(\omega)$. Repeating the above construction for each $\sigma \in \mathrm{Sel}(S_\theta)$, we obtain

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

REMARK 3.2: The set $Q_\theta \equiv Q_\theta(\underline{x}, \varepsilon)$ satisfies Assumption 2.2 by construction. By Proposition 3.1, Q_θ is a random closed set in $\Delta^{\kappa_{\mathcal{Y}}-1}$, because it is given by a continuous map applied to the random closed set S_θ. Because every realization of $q \in \mathrm{Sel}(Q_\theta)$ is contained in $\Delta^{\kappa_{\mathcal{Y}}-1}$, Q_θ is integrably bounded. Hence, Assumption 2.3 is satisfied.

EXAMPLE 3.1—Continued: Consider 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

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

Figure 1(b) plots the set Q_θ resulting from the possible realizations of ε_1 and ε_2.

Because Q_θ is an integrably bounded random closed set, all its selections are integrable and its conditional Aumann expectation is

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

EXAMPLE 3.1—Continued: Consider the set Q_θ plotted in Figure 1(b). Let $\Omega^M = \{\omega \in \Omega : \varepsilon(\omega) \in [0, -\theta_1] \times [0, -\theta_2]\}$. Then for $\omega \notin \Omega^M$, the set Q_θ has only one selection, since the equilibrium is unique. For $\omega \in \Omega^M$, the selections of Q_θ are

$$q(\sigma(\omega)) = \begin{cases} [0\ \ 1\ \ 0\ \ 0]', & \text{if } \omega \in \Omega_1^M, \\ q\left(\dfrac{\varepsilon_2(\omega)}{-\theta_2}, \dfrac{\varepsilon_1(\omega)}{-\theta_1}\right), & \text{if } \omega \in \Omega_2^M, \\ [0\ \ 0\ \ 1\ \ 0]', & \text{if } \omega \in \Omega_3^M, \end{cases}$$

for all measurable partitions $\{\Omega_i^M\}_{i=1}^3$ of Ω^M. In the above expression,

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

The expectations of the selections of Q_θ build the set $\mathbb{E}(Q_\theta)$, which is a convex subset of Δ^3 with infinitely many extreme points.

The set $\mathbb{E}(Q_\theta|\underline{x})$ collects vectors of probabilities with which each outcome of the game can be observed. It is obtained by integrating the probability distribution over outcomes of the game implied by each mixed strategy equilibrium

σ given $\underset{\sim}{x}$ and ε (that is, by integrating each element of $\mathrm{Sel}(Q_\theta)$) against the probability measure of $\varepsilon|\underset{\sim}{x}$. 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, $\mathbb{E}(Q_\theta|\underset{\sim}{x})$ is the set of probability distributions over action profiles conditional on $\underset{\sim}{x}$ which are consistent with the maintained modeling assumptions, that is, with *all* the model's implications. In other words, it is the convex set of moment predictions.

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 $\underset{\sim}{x}$, $\mathbf{P}(y|\underset{\sim}{x})$, is a point in the set $\mathbb{E}(Q_\theta|\underset{\sim}{x})$ for $\underset{\sim}{x}$-a.s., where $\mathbf{P}(y|\underset{\sim}{x}) \equiv [\mathbf{P}(y = t^k|\underset{\sim}{x}), k = 1, \dots, \kappa_{\mathcal{Y}}]$. This is because by the definition of $\mathbb{E}(Q_\theta|\underset{\sim}{x})$, $\mathbf{P}(y|\underset{\sim}{x}) \in \mathbb{E}(Q_\theta|\underset{\sim}{x})$, $\underset{\sim}{x}$-a.s., if and only if there exists $q \in \mathrm{Sel}(Q_\theta)$ such that $\mathbf{E}(q|\underset{\sim}{x}) = \mathbf{P}(y|\underset{\sim}{x})$, $\underset{\sim}{x}$-a.s. Hence, the set of observationally equivalent parameter values that form the sharp identification region is given by

$$\Theta_I = \{\theta \in \Theta : \mathbf{P}(y|\underset{\sim}{x}) \in \mathbb{E}(Q_\theta|\underset{\sim}{x}), \underset{\sim}{x}\text{-a.s.}\}. \tag{3.2}$$

THEOREM 3.2: *Let Assumptions 3.1 and 3.2 hold. Then*

$$\Theta_I = \Big\{\theta \in \Theta : \max_{u \in B}\big(u'\mathbf{P}(y|\underset{\sim}{x}) - \mathbf{E}[h(Q_\theta, u)|\underset{\sim}{x}]\big) = 0,\ \underset{\sim}{x}\text{-}a.s.\Big\} \tag{3.3}$$

$$= \Big\{\theta \in \Theta : \int_B \big(u'\mathbf{P}(y|\underset{\sim}{x}) - \mathbf{E}[h(Q_\theta, u)|\underset{\sim}{x}]\big)_+ \, d\mathcal{U} = 0,\ \underset{\sim}{x}\text{-}a.s.\Big\}, \tag{3.4}$$

where $h(Q_\theta, u) = \max_{q \in Q_\theta} u'q = \max_{\sigma \in S_\theta} \sum_{k=1}^{\kappa_{\mathcal{Y}}} u_k \prod_{j=1}^{J} \sigma_j(t_j^k)$ *and* $u' = [u_1\ u_2\ \cdots\ u_{\kappa_{\mathcal{Y}}}]$.[9]

Theorem 3.2 follows immediately from Theorem 2.1, because Assumptions 2.1–2.5 are satisfied for this application, as summarized in Remarks 3.1, 3.2, and 3.3 (the latter given below).

By Wilson's (1971) result, the realizations of the set of MSNE, S_θ, are almost surely finite sets. Therefore, the same holds for Q_θ. Hence, for given $\omega \in \Omega$, $h(Q_\theta(\omega), u)$ is given by the maximum among the inner product of u with a finite number of vectors, the elements of $Q_\theta(\omega)$. These elements are known functions of $(\underset{\sim}{x}(\omega), \varepsilon(\omega))$. Hence, given Q_θ, the expectation of $h(Q_\theta, u)$ is easy to compute.

EXAMPLE 3.1—Continued: Consider the set Q_θ plotted in Figure 1(b). Pick a direction $u \equiv [u_1\ u_2\ u_3\ u_4]' \in B$. Then for $\omega \in \Omega$ such that $\varepsilon(\omega) \in (-\infty, 0] \times (-\infty, 0]$, we have $Q_\theta(\omega) = \{[1\ 0\ 0\ 0]'\}$ and $h(Q_\theta(\omega), u) = u_1$. For

[9]Recall that B is the unit ball in $\Re^{\kappa_{\mathcal{Y}}}$ and $\mathcal{U}$ is any probability measure on B with support equal to B. Recall also that $\mathcal{Y} = \{t^1, t^2, \dots, t^{\kappa_{\mathcal{Y}}}\}$ is the set of possible outcomes of the game, and $t^k \equiv (t_1^k, \dots, t_J^k)$ is a J-tuple specifying one action in $\mathcal{Y}_j$ for each player $j = 1, \dots, J$.

$\omega \in \Omega$ such that $\varepsilon(\omega) \in [0, -\theta_1] \times [0, -\theta_2]$, we have $Q_\theta(\omega) = \{[0\ 1\ 0\ 0]',$ $q(\frac{\varepsilon_2(\omega)}{-\theta_2}, \frac{\varepsilon_1(\omega)}{-\theta_1}),\ [0\ 0\ 1\ 0]'\}$, and $h(Q_\theta(\omega), u) = \max(u_2, u'q(\frac{\varepsilon_2(\omega)}{-\theta_2}, \frac{\varepsilon_1(\omega)}{-\theta_1}), u_3)$. Figure 1(c) plots $h(Q_\theta(\omega), u)$ against the possible realizations of $\varepsilon_1, \varepsilon_2$.

By a way of comparison with the previous literature, and to show how Assumptions 2.4 and 2.5 can be verified, we provide the abstract definition of Θ_I given by Berry and Tamer (2007, equation (2.21), p. 67) for the case of a two player entry game, extending it to finite games with potentially more than two players and two actions. A finite game with multiple equilibria can be completed by a random vector which has almost surely nonnegative entries that sum to 1 and which gives the probability with which each equilibrium in the regions of multiplicity is played when the game is defined by $(\underline{x}, \varepsilon, \theta)$. Denote such (random) discrete distribution by $\lambda(\cdot; \underline{x}, \varepsilon, \theta) : S_\theta \to \Delta^{\kappa_{S_\theta}-1}$. Notice that $\lambda(\cdot; \underline{x}, \varepsilon, \theta)$ is left unspecified and can depend on market unobservables even after conditioning on market observables. By definition, the sharp identification region includes all the parameter values for which one can find a random vector $\lambda(\cdot; \underline{x}, \varepsilon, \theta)$ satisfying the above conditions, such that the model augmented with this selection mechanism generates the joint distribution of the observed variables. Hence,

$$
\begin{aligned}
(3.5)\qquad \Theta_I = \Bigg\{ & \theta \in \Theta : \exists \lambda(\cdot; \underline{x}, \varepsilon, \theta) : S_\theta \to \Delta^{\kappa_{S_\theta}-1} \text{ for } (\underline{x}, \varepsilon)\text{-a.s.,} \\
& \text{such that } \forall k = 1, \dots, \kappa_{\mathcal{Y}}, \\
& \mathbf{P}(y = t^k | \underline{x}) = \int \Bigg(\sum_{\sigma \in S_\theta(\underline{x}, \varepsilon)} \lambda(\sigma; \underline{x}, \varepsilon, \theta) \\
& \times \prod_{j=1}^{J} \sigma_j(t_j^k) \Bigg) \, dF(\varepsilon | \underline{x}),\ \underline{x}\text{-a.s.} \Bigg\}.
\end{aligned}
$$

Compared with this definition, our characterization in Theorem 3.2 has the advantage of avoiding altogether the need to deal with the specification of a selection mechanism. The latter may constitute an infinite dimensional nuisance parameter and may, therefore, create difficulties for the computation of Θ_I and for inference.

Notice that with respect to the general notation used in Section 2, $w(z) = [1(y = t^k), k = 1, \dots, \kappa_{\mathcal{Y}}]$. Finally, observe that using $\lambda(\cdot; \underline{x}, \varepsilon, \theta)$ one can construct a selected prediction $\psi(\underline{x}, \varepsilon, \theta)$ as a random vector whose realizations given $\underline{x}$ and ε are equal to

$$
\left[\prod_{j=1}^{J} \sigma_j(t_j^k), k = 1, \dots, \kappa_{\mathcal{Y}} \right]
$$

$$
\text{with probability } \lambda(\sigma; \underline{x}, \varepsilon, \theta),\ \ \sigma \in \operatorname{Sel}(S_\theta).
$$

REMARK 3.3: The random vector $\psi(\underline{x}, \varepsilon, \theta)$ is a selected prediction satisfying Assumption 2.4. Observing that

$$\mathbf{E}(\psi(\underline{x}, \varepsilon, \theta)|\underline{x}) = \int \Bigg[\sum_{\sigma \in S_\theta(\underline{x}, \varepsilon)} \lambda(\sigma; \underline{x}, \varepsilon, \theta) \times \prod_{j=1}^{J} \sigma_j(t_j^k),\ k = 1, \dots, \kappa_{\mathcal{Y}} \Bigg] dF(\varepsilon|\underline{x}),$$

where the integral is taken coordinatewise, Assumption 2.5 is verified.

REMARK 3.4: Appendix B in the Supplemental Material verifies Andrews and Shi's (2009) regularity conditions for models satisfying Assumptions 3.1 and 3.2 under the additional assumption that the researcher observes an i.i.d. sequence of equilibrium outcomes and observable payoff shifters $\{y_i, \underline{x}_i\}_{i=1}^n$. Andrews and Shi's (2009) generalized moment selection procedure with infinitely many conditional moment inequalities can, therefore, be applied to obtain confidence sets that have correct uniform asymptotic coverage.

REMARK 3.5: We conclude this section by observing that static finite games of incomplete information with multiple equilibria can be analyzed using our methodology in a manner which is completely analogous to how we have addressed the case of complete information. Moreover, our methodology characterizes the sharp identification region for this class of models through a finite number of conditional moment inequalities. We establish this formally in Appendix C in the Supplemental Material. Grieco (2009) introduced an important model, where each player has a vector of payoff shifters that are unobservable by the researcher. Some of the elements of this vector are private information to the player, while the others are known to all players. Our results in Section 2 apply to this setup as well, by the same arguments as in Section 3 and in Appendix C in the Supplemental Material.

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. Beresteanu, Molchanov, and Molinari (2009, Section 3.3) revisited Example 3.1 in light of this comparison.

In the presence of multiple equilibria, ABJ observed that an implication of the model is that for a given $t^k \in \mathcal{Y}$, $\mathbf{P}(y = t^k|\underline{x})$ 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 $t^l \in \mathcal{Y}$ which is also a possible equilibrium outcome of the game, and when both are possible, t^k is selected only part of the time. CT pointed out that additional information can be learned from the model. In particular, $\mathbf{P}(y = t^k|\underline{x})$ 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 following proposition rewrites the outer regions originally proposed by ABJ and CT, denoted Θ_O^{ABJ} and Θ_O^{CT}, using our notation. It then establishes their connection with Θ_I.

PROPOSITION 3.3: *Let Assumptions 3.1 and 3.2 hold. Then the outer regions proposed by ABJ and CT are, respectively,*

$$\begin{aligned}(3.6)\qquad \Theta_O^{\mathrm{ABJ}} = \Bigg\{\theta \in \Theta:& \\ &\mathbf{P}(y = t^k|\underline{x}) \le \max\bigg(\int [q(\sigma)]_k \, dF_\theta(\varepsilon|\underline{x}) : \sigma \in \mathrm{Sel}(S_\theta)\bigg), \\ &\text{for } k = 1, \ldots, \kappa_{\mathcal{Y}}, \underline{x}\text{-a.s.}\Bigg\}\end{aligned}$$

and

$$\begin{aligned}(3.7)\qquad \Theta_O^{\mathrm{CT}} = \Bigg\{\theta \in \Theta: &\min\bigg(\int [q(\sigma)]_k \, dF_\theta(\varepsilon|\underline{x}) : \sigma \in \mathrm{Sel}(S_\theta)\bigg) \\ &\le \mathbf{P}(y = t^k|\underline{x}) \le \max\bigg(\int [q(\sigma)]_k \, dF_\theta(\varepsilon|\underline{x}) : \sigma \in \mathrm{Sel}(S_\theta)\bigg), \\ &\text{for } k = 1, \ldots, \kappa_{\mathcal{Y}}, \underline{x}\text{-a.s.}\Bigg\}.\end{aligned}$$

Θ_O^{ABJ} can be obtained by solving the maximization problem in equation (3.3) over the restricted set of u's equal to the canonical basis vectors in $\Re^{\kappa_{\mathcal{Y}}}$. Θ_O^{CT} can be obtained by solving the maximization problem in equation (3.3) over the restricted set of u's equal to the canonical basis vectors in $\Re^{\kappa_{\mathcal{Y}}}$ and each of these vectors multiplied by -1.

Hence, the approaches of ABJ and CT can be interpreted on the basis of our analysis as follows. For each $\theta \in \Theta$, ABJ's inequalities give the closed half-spaces delimited by hyperplanes that are parallel to the axis and that support $\mathbb{E}(Q_\theta|\underline{x})$. Θ_O^{ABJ} is the collection of θ's such that $\mathbf{P}(y|\underline{x})$ is contained in the nonnegative part of such closed half-spaces $\underline{x}$-a.s. CT used a more refined approach and for each $\theta \in \Theta$, their inequalities give the smallest hypercube containing $\mathbb{E}(Q_\theta|\underline{x})$. Θ_O^{CT} is the collection of θ's such that $\mathbf{P}(y|\underline{x})$ is contained in such a hypercube $\underline{x}$-a.s. The more $\mathbb{E}(Q_\theta|\underline{x})$ 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. A graphical intuition for this relationship is given in Figure 2.

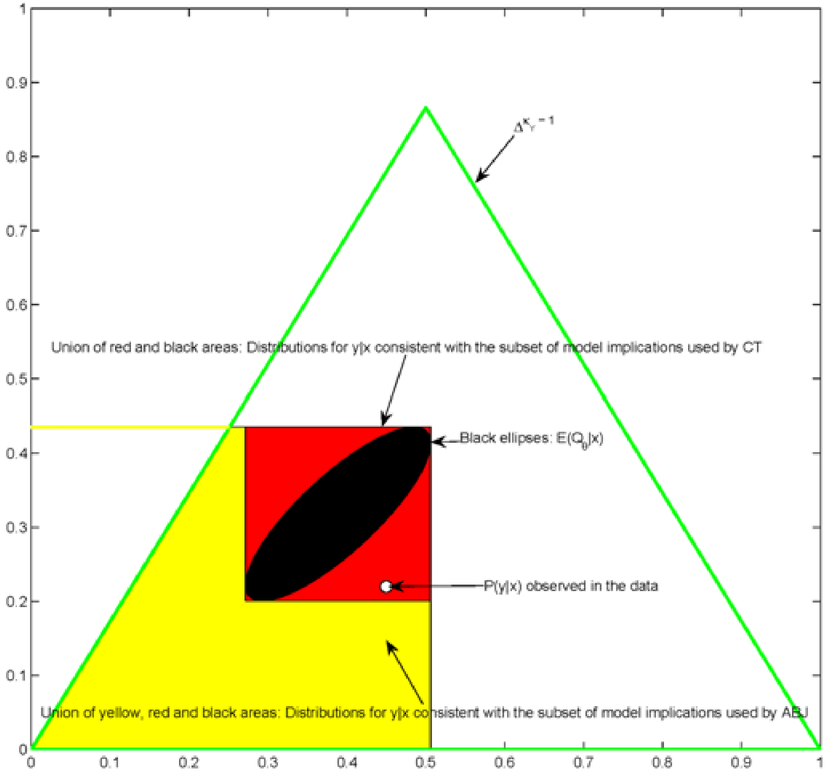

FIGURE 2.—A comparison between the logic behind the approaches of ABJ, CT, and ours obtained by projecting in $\Re^2 : \Delta^{\kappa_{\mathcal{Y}}-1}$, $\mathbb{E}(Q_\theta|\underline{x})$, and the hypercubes used by ABJ and CT. A candidate $\theta \in \Theta$ is in Θ_I if $\mathbf{P}(y|\underline{x})$,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

3.4. *Two Player Entry Game—An Implementation*

This section presents an implementation of our method and a series of numerical illustrations of the identification gains that it affords in the two player entry game in Example 3.1, both with and without covariates in the payoff functions. The set S_θ for this example (omitting $\underline{x}$) is plotted in Figure 1. Appendix A.3 provides details on the method used to compute Θ_O^{ABJ}, Θ_O^{CT}, and Θ_I.

For all the data generating processes (DGPs), we let $(\varepsilon_1, \varepsilon_2) \overset{\text{i.i.d.}}{\sim} N(0,1)$. The DGPs without covariates are designed as follows. We build a grid of 36 equally spaced values for $\theta_1^\star, \theta_2^\star$ on $[-1.8, -0.8] \times [-1.7, -0.7]$, yielding multiple equilibria with a probability that ranges from substantial (0.21) to small (0.07). We match each point on the $\theta_1^\star, \theta_2^\star$ grid, with each point on a grid of 10 values for

$$\lambda^\star = \Bigg[\mathbf{P}((0,1) \text{ is chosen} | \varepsilon \in \mathcal{E}_{\theta^\star}^M) \;\; \mathbf{P}((1,0) \text{ is chosen} | \varepsilon \in \mathcal{E}_{\theta^\star}^M) \;\; \mathbf{P}\bigg(\bigg(\frac{\varepsilon_2}{-\theta_2}, \frac{\varepsilon_1}{-\theta_1}\bigg) \text{ is chosen} \bigg| \varepsilon \in \mathcal{E}_{\theta^\star}^M\bigg)\Bigg]',$$

where $\mathcal{E}_{\theta^\star}^M \equiv [0, -\theta_1^\star] \times [0, -\theta_2^\star]$. The grid of values for $\lambda^\star$ is

$$\lambda^\star \in \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \frac{1}{4} \\ 0 \\ \frac{3}{4} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{3}{4} \\ 0 \\ \frac{1}{4} \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{4} \\ \frac{3}{4} \end{bmatrix}, \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ \frac{1}{4} \\ \frac{1}{4} \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{3}{4} \\ \frac{1}{4} \end{bmatrix} \right\}.$$

This results in 360 distinct DGPs, each with a corresponding vector $[\mathbf{P}(y = t),\ t \in \{(0,0), (1,0), (0,1), (1,1)\}]$. We compute Θ_I, Θ_O^{CT}, and Θ_O^{ABJ} for each DGP, letting the parameter space be $\Theta = [-4.995, -0.005]^2$. We then rank the results according to

$$\frac{\operatorname{length}(\operatorname{Proj}(\Theta_I|1)) + \operatorname{length}(\operatorname{Proj}(\Theta_I|2))}{\operatorname{length}(\operatorname{Proj}(\Theta_O^{\mathrm{CT}}|1)) + \operatorname{length}(\operatorname{Proj}(\Theta_O^{\mathrm{CT}}|2))},$$

where $\operatorname{Proj}(\cdot|i)$ is the projection of the set in parentheses on dimension i, and $\operatorname{length}(\operatorname{Proj}(\cdot|i))$ is the length of such projection. To conserve space, in Table I, we report only the results of our "top 15% reduction," "median reduction,"

TABLE I

THE PARAMETERS USED TO GENERATE THE DISTRIBUTION $\mathbf{P}(y|\underline{x})$ AND THE RESULTS[a]

DGP True Values[c]	Projections[b] Θ_O^{ABJ}	Θ_O^{CT}	Θ_I	% Width Reduction[d]	% Region Reduction[e]
$\theta_1^\star = -1.0$	$[-3.22, -0.22]$	$[-3.22, -0.28]$	$[-2.21, -0.30]$	35.03%	63.81%
$\theta_2^\star = -1.3$	$[-3.22, -1.05]$	$[-3.22, -1.15]$	$[-2.32, -1.16]$	43.96%	
$\lambda^\star = [0\ \ \frac{3}{4}\ \ \frac{1}{4}]'$					
$\theta_1^\star = -0.8$	$[-1.82, -0.53]$	$[-1.82, -0.57]$	$[-1.51, -0.58]$	25.60%	56.87%
$\theta_2^\star = -1.1$	$[-1.82, -0.59]$	$[-1.82, -0.64]$	$[-1.55, -0.64]$	22.88%	
$\lambda^\star = [\frac{1}{2}\ \ \frac{1}{4}\ \ \frac{1}{4}]'$					
$\theta_1^\star = -1.2$	$[-2.19, -0.75]$	$[-2.19, -0.90]$	$[-2.11, -1.05]$	17.83%	26.02%
$\theta_2^\star = -1.5$	$[-2.19, -0.75]$	$[-2.19, -0.79]$	$[-2.13, -0.90]$	12.14%	
$\lambda^\star = [\frac{1}{4}\ \ 0\ \ \frac{3}{4}]'$					

[a]The results reported are selected from a set of 360 DGPs as described in Section 3.4 for a two player entry game with mixed strategy Nash equilibrium as solution concept.

[b]Projections of Θ_O^{ABJ}, Θ_O^{CT}, and Θ_I are on each dimension (hence yielding bounds for θ_1 and θ_2).

[c]$\lambda^\star = [\mathbf{P}((0,1) \text{ is chosen}|\varepsilon \in \mathcal{E}_{\theta^\star}^M)\ \mathbf{P}((1,0) \text{ is chosen}|\varepsilon \in \mathcal{E}_{\theta^\star}^M)\ \mathbf{P}((\frac{\varepsilon_2}{-\theta_2}, \frac{\varepsilon_1}{-\theta_1}) \text{ is chosen}|\varepsilon \in \mathcal{E}_{\theta^\star}^M)]'$, $\mathcal{E}_{\theta^\star}^M = [0, -\theta_1^\star] \times [0, -\theta_2^\star]$.

[d]The reduction in width of the sharp bounds compared to CT's outer bounds. Calculated as $\frac{\mathrm{Proj}(\Theta_O^{\mathrm{CT}}|j) - \mathrm{Proj}(\Theta_I|j)}{\mathrm{Proj}(\Theta_O^{\mathrm{CT}}|j)}$, where $\mathrm{Proj}(\cdot|j)$ is the projection of the set in parentheses on dimension j.

[e]The reduction in area of Θ_I compared to Θ_O^{CT}. Calculated as $\frac{\mathrm{Area}(\Theta_O^{\mathrm{CT}}) - \mathrm{Area}(\Theta_I)}{\mathrm{Area}(\Theta_O^{\mathrm{CT}})}$, where $\mathrm{Area}(\cdot)$ is the area of the set in parenthesis).

and "bottom 15% reduction."[10] Figure 3 plots Θ_I, Θ_O^{CT}, and Θ_O^{ABJ} for each of these DGPs.

To further illustrate the computational feasibility of our methodology, we allow for covariates in the payoff functions. Specifically, we let $\pi_j = t_j(t_{-j}\theta_j + \beta_{0j} + x_{1j}\beta_{1j} + x_{2j}\beta_{2j} + \varepsilon_j)$, $j = 1, 2$, where $[x_{11}\ \ x_{21}]$, the covariates for player 1, take four different values, $\{[-2\ \ 1], [1\ \ -1.5], [0\ \ 0.75], [-1.5\ \ -1]\}$, and $[x_{12}\ \ x_{22}]$, the covariates for player 2, take five different values, $\{[1\ \ -1.75], [-1.25\ \ 1], [0\ \ 0], [0.6\ \ 0.5], [0.5\ \ -0.5]\}$. The parameter vector of interest is $\theta = [(\theta_j\ \ \beta_{0j}\ \ \beta_{1j}\ \ \beta_{2j})_{j=1,2}]$. In generating $\mathbf{P}(y|\underline{x})$, we use the values of $\lambda^\star$ and $\theta_1^\star, \theta_2^\star$ which yield the top 15% reduction, median reduction, and bottom 15% reduction in the DGPs with no x variables, and pair them with $[\beta_{01}^\star\ \ \beta_{11}^\star\ \ \beta_{21}^\star] = [0\ \ 1/2\ \ 1/3]$ and $[\beta_{02}^\star\ \ \beta_{12}^\star\ \ \beta_{22}^\star] = [0\ \ -1/3\ \ -1/2]$. This results in three different DGPs. Compared to the case with no covariates, for each of these DGPs, the computational time required to verify whether a can-

[10]The full set of results is available from the authors on request. Our best result has a 97% reduction in size of Θ_I compared to Θ_O^{CT}. Our worst result has a 20% reduction in size of Θ_I compared to Θ_O^{CT}. Only 6% of the DGPs yield a reduction in size of Θ_I compared to Θ_O^{CT} of less than 25%.

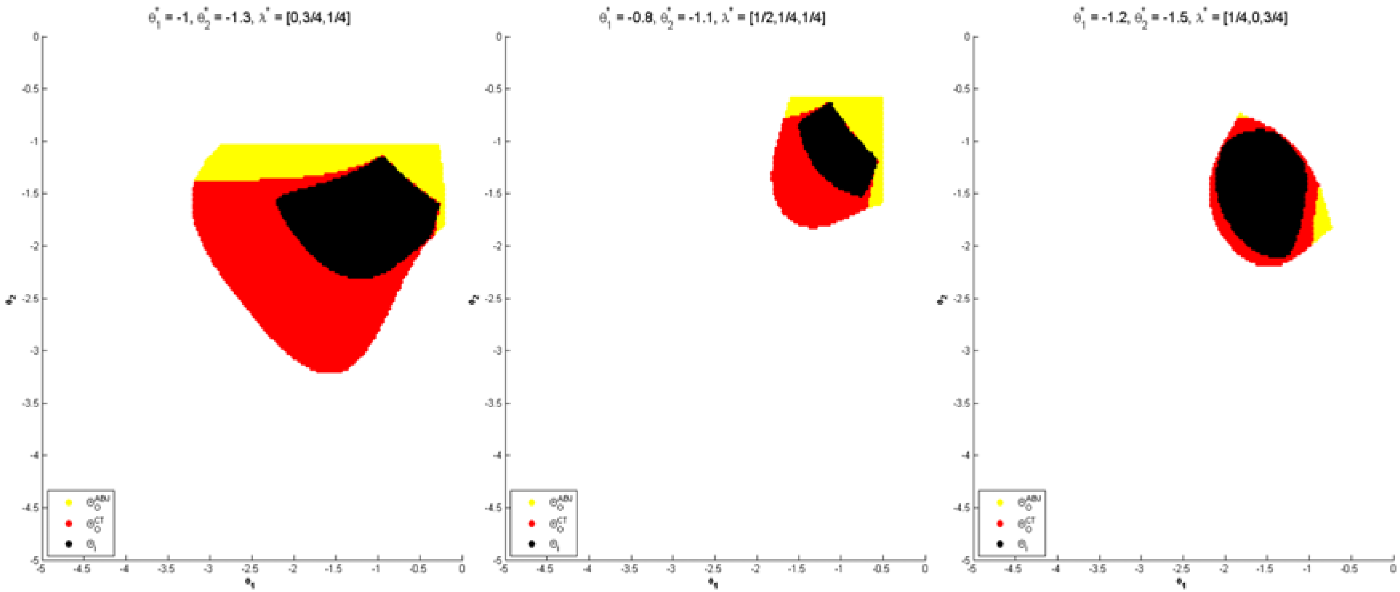

FIGURE 3.—Identification regions in a two player entry game with mixed strategy Nash equilibrium as the solution concept, for three different DGPs; see Table I.

TABLE II

PROJECTIONS OF Θ_O^{ABJ}, Θ_O^{CT}, AND Θ_I AND REDUCTION IN VOLUME OF Θ_I COMPARED TO Θ_O^{CT}. TWO PLAYER ENTRY GAME WITH MIXED STRATEGY NASH EQUILIBRIUM AS SOLUTION CONCEPT

	Projections of DGP 1[a]			Projections of DGP 2[b]			Projections of DGP 3[c]		
	Θ_O^{ABJ}	Θ_O^{CT}	Θ_I	Θ_O^{ABJ}	Θ_O^{CT}	Θ_I	Θ_O^{ABJ}	Θ_O^{CT}	Θ_I
Player 1									
θ_1	$[-3.81, -0.50]$	$[-3.78, -0.60]$	$[-3.18, -0.63]$	$[-1.51, -0.45]$	$[-1.47, -0.61]$	$[-1.41, -0.62]$	$[-2.37, -0.46]$	$[-1.99, -0.69]$	$[-1.77, -0.80]$
β_{01}	$[-0.09, 0.43]$	$[-0.02, 0.41]$	$[-0.01, 0.35]$	$[-0.16, 0.17]$	$[-0.05, 0.15]$	$[-0.04, 0.12]$	$[-0.24, 0.19]$	$[-0.13, 0.12]$	$[-0.12, 0.08]$
β_{11}	$[0.37, 0.66]$	$[0.42, 0.64]$	$[0.44, 0.62]$	$[0.41, 0.61]$	$[0.46, 0.58]$	$[0.47, 0.57]$	$[0.37, 0.57]$	$[0.44, 0.56]$	$[0.45, 0.55]$
β_{21}	$[0.20, 0.49]$	$[0.25, 0.47]$	$[0.27, 0.44]$	$[0.24, 0.43]$	$[0.29, 0.40]$	$[0.30, 0.39]$	$[0.21, 0.42]$	$[0.27, 0.40]$	$[0.29, 0.38]$
% Region Reduction[d]			64%			38%			52%
Player 2									
θ_2	$[-3.87, -0.95]$	$[-3.86, -1.23]$	$[-3.24, -1.25]$	$[-1.65, -0.60]$	$[-1.58, -0.77]$	$[-1.47, -0.79]$	$[-2.34, -0.65]$	$[-2.25, -0.78]$	$[-2.14, -0.92]$
β_{02}	$[-0.13, 0.11]$	$[-0.06, 0.09]$	$[-0.06, 0.06]$	$[-0.06, 0.06]$	$[-0.03, 0.05]$	$[-0.03, 0.05]$	$[-0.09, 0.04]$	$[-0.05, 0.04]$	$[-0.04, 0.03]$
β_{12}	$[-0.48, -0.20]$	$[-0.45, -0.28]$	$[-0.43, -0.29]$	$[-0.41, -0.26]$	$[-0.38, -0.31]$	$[-0.37, -0.31]$	$[-0.39, -0.25]$	$[-0.38, -0.29]$	$[-0.36, -0.30]$
β_{22}	$[-0.65, -0.38]$	$[-0.64, -0.46]$	$[-0.61, -0.48]$	$[-0.57, -0.43]$	$[-0.54, -0.48]$	$[-0.54, -0.48]$	$[-0.57, -0.41]$	$[-0.55, -0.46]$	$[-0.53, -0.47]$
% Region Reduction[d]			63%			33%			46%

[a] $\lambda^\star = [0\ \ \frac{3}{4}\ \ \frac{1}{4}]$, $\theta_1^\star = -1.0$, $\theta_2^\star = -1.3$, $\beta_1^\star = [0\ \ 1/2\ \ 1/3]$, $\beta_2^\star = [0\ \ -1/3\ \ -1/2]$.

[b] $\lambda^\star = [\frac{1}{2}\ \ \frac{1}{4}\ \ \frac{1}{4}]$, $\theta_1^\star = -0.8$, $\theta_2^\star = -1.1$, $\beta_1^\star = [0\ \ 1/2\ \ 1/3]$, $\beta_2^\star = [0\ \ -1/3\ \ -1/2]$.

[c] $\lambda^\star = [\frac{1}{4}\ \ 0\ \ \frac{3}{4}]$, $\theta_1^\star = -1.2$, $\theta_2^\star = -1.5$, $\beta_1^\star = [0\ \ 1/2\ \ 1/3]$, $\beta_2^\star = [0\ \ -1/3\ \ -1/2]$.

[d] Calculated as $\frac{\mathrm{Vol}(\Theta_O^{\mathrm{CT}}|j) - \mathrm{Vol}(\Theta_I|j)}{\mathrm{Vol}(\Theta_O^{\mathrm{CT}}|j)}$, where $\mathrm{Vol}(\cdot|j)$ is the volume of the set in parentheses projected on the parameters for player j (approximated by the box-grid).

didate θ is in Θ_I is linear in the number of values that $\underline{x}$ can take. The reductions in size of Θ_I compared to the outer regions of ABJ and CT is of similar magnitude to the case with no covariates. Table II reports the results.

4. APPLICATION II: BEST LINEAR PREDICTION WITH INTERVAL OUTCOME AND COVARIATE DATA

Here we consider the problem of best linear prediction under square loss, when both outcome and covariate data are interval valued. When thinking about best linear prediction (BLP), 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." HMPS studied the related problem of identification of the BLP parameters with missing data on outcome and covariates, and provided a characterization of the identification region of each component of the vector θ. While their characterization is sharp, we emphasize that the computational complexity of the problem in the HMPS formulation grows with the number of points in the support of the outcome and covariate variables, 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 Θ_I which remains computationally feasible regardless of the support of outcome and covariate variables.[11]

We let $y^\star$ and $x^\star$ denote the unobservable outcome and covariate variables. To simplify the exposition, we let $x^\star$ be scalar, although this assumption can be relaxed and is not essential for our methodology. We maintain the following assumption:

ASSUMPTION 4.1: *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, *and* x_U *such that* $\mathbf{P}(y_L \leq y^\star \leq y_U) = 1$ *and* $\mathbf{P}(x_L \leq x^\star \leq x_U) = 1$. $\mathbf{E}(|y_i|)$, $\mathbf{E}(|x_j|)$, $\mathbf{E}(|y_i x_j|)$, *and* $\mathbf{E}(x_j^2)$ *are all finite for each* $i, j = L, U$. *One of the following statements holds*: (i) *at least one of* $y_L, y_U, x_L, x_U, y^\star, x^\star$ *has a continuous distribution or* (ii) $(\Omega, \mathfrak{F}, \mathbf{P})$ *is a nonatomic probability space*.

With respect to the general notation used in Section 2, $z = (y_L, y_U, x_L, x_U)$, $\xi = (y^\star, x^\star)$, and $\mathfrak{F}$ is the σ-algebra generated by $(y_L, y_U, x_L, x_U, y^\star, x^\star)$. The researcher works with unconditional moments.

[11] Beresteanu and Molinari (2008) studied identification and statistical inference for the BLP parameters $\theta \in \Theta$ when only the outcome variable is interval valued. See also Bontemps, Magnac, and Maurin (2011) for related results. Here we significantly generalize their identification results by allowing also for interval-valued covariates. This greatly complicates computation of Θ_I and inference, because Θ_I is no longer a linear transformation of an Aumann expectation.

REMARK 4.1: Under Assumption 4.1, Assumption 2.1 is satisfied.

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 $\varepsilon^\star = y^\star - \theta_1^\star - \theta_2^\star x^\star$ satisfies $\mathbf{E}(\varepsilon^\star) = 0$ and $\mathbf{E}(\varepsilon^\star x^\star) = 0$. For any candidate $\theta \in \Theta$, we extend the construction of the prediction error to the case of interval valued data. We let $Y = [y_L, y_U]$ and $X = [x_L, x_U]$. It is easy to show that these are random closed sets in $\Re$ as per Definition A.1 (see Beresteanu and Molinari (2008, Lemma A.3)). We build the set

$$Q_\theta = \left\{ q = \begin{bmatrix} y - \theta_1 - \theta_2 x \\ (y - \theta_1 - \theta_2 x)x \end{bmatrix} : (y, x) \in \operatorname{Sel}(Y \times X) \right\}. \tag{4.1}$$

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.

REMARK 4.2: The set Q_θ satisfies Assumption 2.2 by construction. Because it is given by a continuous map applied to the random closed sets Y and X, Q_θ is a random closed set in $\Re^2$. By Assumption 4.1, the set Q_θ is integrably bounded; see Beresteanu and Molinari (2008, Proof of Theorem 4.2). By the fundamental selection theorem (Molchanov (2005, Theorem 1.2.13)) and by Lemma 2.1, there exist selected predictions $\psi(y_L, y_U, x_L, x_U, y^\star, x^\star, \theta)$ that satisfy Assumption 2.4. The last step in the proof of Theorem 4.1, given in Appendix A, establishes that Assumption 2.5 holds.

Given the set Q_θ, one can relate conceptually our approach in Section 2 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 $(y, x) \in \operatorname{Sel}(Y \times X)$ is equal to the (unconditional) Aumann expectation $\mathbb{E}(Q_\theta) = \{\mathbf{E}(q) : q \in \operatorname{Sel}(Q_\theta)\}$. Because the probability space is nonatomic and Q_θ belongs to a finite dimensional space, $\mathbb{E}(Q_\theta)$ is a closed convex set. If $\mathbb{E}(Q_\theta)$ 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 $x^\star$ (hence, with respect to the general notation used in Section 2, $w(z) = [0\ \ 0]'$). This is because if the condition just mentioned is satisfied, then for the candidate $\theta \in \Theta$, there exists a selection in $\operatorname{Sel}(Y \times X)$, 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 4.1.

THEOREM 4.1: *Let Assumption 4.1 hold. Then*

$$\Theta_I = \left\{\theta \in \Theta : \max_{u \in B}\big(-\mathbf{E}[h(Q_\theta, u)]\big) = 0\right\}$$
$$= \left\{\theta \in \Theta : \int_B \big(\mathbf{E}[h(Q_\theta, u)]\big)_- \, d\mathcal{U} = 0\right\}.$$

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

$$h(Q_\theta, u) = \max_{q \in Q_\theta} u'q = \max_{y \in Y, x \in X} [u_1(y - \theta_1 - \theta_2 x) + u_2(yx - \theta_1 x - \theta_2 x^2)]. \tag{4.2}$$

For given $\theta \in \Theta$ and $u \in B$, this maximization problem can be efficiently solved using the gradient method, regardless of whether $(y_i, x_i)_{i=L,U}$, $(y^\star, x^\star)$ are continuous or discrete random variables. Hence, $h(Q_\theta, u)$ is an easy-to-calculate continuous-valued convex sublinear function of u. Membership of a candidate θ to the set Θ_I can be verified by using efficient algorithms in convex programming or by taking integrals of concave functions.

REMARK 4.3: Appendix B in the Supplemental Material verifies Andrews and Shi's (2009) regularity conditions for models that satisfy Assumption 4.1, under the additional assumption that the researcher observes an i.i.d. sequence $\{y_{iL}, y_{iU}, x_{iL}, x_{iU}\}_{i=1}^n$ and that these have finite fourth moments.

5. CONCLUSIONS

This paper introduces a computationally feasible characterization for the sharp identification region Θ_I of the parameters of incomplete econometric models with convex moment predictions. Our approach is based on characterizing, for each $\theta \in \Theta$, the set of moments which are consistent with *all* the model's implications, as the (conditional) Aumann expectation of a properly defined random set. If the model is correctly specified, one can then build Θ_I as follows. A candidate θ is in Θ_I if and only if it yields a conditional Aumann expectation which, for $\underline{x}$-a.s., contains the relevant expectations of random variables observed in the data. Because, in general, for each $\theta \in \Theta$, the conditional Aumann expectation may have infinitely many extreme points, characterizing the set Θ_I entails checking that an infinite number of moment inequalities are satisfied. 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 maximum of an easy-to-compute superlinear (hence concave) function over the unit ball is equal to zero. We exemplify our methodology by applying it to empirically relevant models for which a feasible characterization of Θ_I was absent in the literature.

We acknowledge that the method proposed in this paper may, for some models, be computationally more intensive than existing methods (e.g., ABJ and CT in the analysis of finite games of complete information 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.

APPENDIX A: PROOFS AND AUXILIARY RESULTS FOR SECTIONS 3 AND 4

A.1. *Definitions*

The theory of random closed sets generally applies to the space of closed subsets of a locally compact Hausdorff second countable topological space $\mathbb{F}$ (e.g., Molchanov (2005)). For the purposes of this paper, it suffices to consider $\mathbb{F} = \Re^d$, which simplifies the exposition. Denote by $\mathcal{F}$ the family of closed subsets of $\Re^d$.

DEFINITION A.1: 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 $\Re^d$, $\{\omega \in \Omega : Z(\omega) \cap K \neq \emptyset\} \in \mathfrak{F}$.

DEFINITION A.2: A random closed set $Z : \Omega \to \mathcal{F}$ is called *integrably bounded* if $\|Z\|_H$ has a finite expectation.

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

A.2. *Proofs*

PROOF OF LEMMA 2.1: For any given $\theta \in \Theta$, if $\psi(z, \xi, \theta)$ is a selected prediction, then $\psi(z, \xi, \theta)$ is a random element as a composition of measurable functions and it belongs to Q_θ for almost all $\omega \in \Omega$ by Assumption 2.4(i). Conversely, for any given $\theta \in \Theta$, let $q \in \mathrm{Sel}(Q_\theta)$. Because q is $\mathfrak{F}$-measurable, by the Doob–Dynkin lemma (see, e.g., Rao and Swift (2006, Proposition 3, Chapter 1)), q can be represented as a measurable function of z and ξ, which is then the selected prediction, and satisfies conditions (i) and (ii) in Assumption 2.4. This selected prediction can also be obtained using a selection mechanism which picks a prediction equal to $q(\omega)$ for each $\omega \in \Omega$. *Q.E.D.*

PROOF OF LEMMA 2.2: For any given $\theta \in \Theta$, let $\mu \in \mathbb{E}(Q_\theta|\mathfrak{G})$. Then by the definition of the conditional Aumann expectation, there exists a $q \in \mathrm{Sel}(Q_\theta)$

such that $\mathbf{E}(q|\mathfrak{G}) = \mu$. By Lemma 2.1, there exists a $\psi(z, \xi, \theta)$ satisfying Assumption 2.4 such that $q(\omega) = \psi(z(\omega), \xi(\omega), \theta)$ for almost all $\omega \in \Omega$, and, therefore, $\mu \in \{\mathbf{E}(\psi(z, \xi, \theta)|\mathfrak{G}) : \psi(z, \xi, \theta)$ satisfies Assumption 2.4$\}$. A similar argument yields the reverse conclusion. *Q.E.D.*

PROOF OF PROPOSITION 3.1: Write the set

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

where $\tilde{\pi}_j(\sigma_{-j}, x_j, \varepsilon_j, \theta) = \sup_{\tilde{\sigma}_j \in \Sigma(\mathcal{Y}_j)} \pi_j(\tilde{\sigma}_j, \sigma_{-j}, x_j, \varepsilon_j, \theta)$. Since $\pi_j(\sigma_j, \sigma_{-j}, x_j, \varepsilon_j, \theta)$ is a continuous function of σ, x_j, and ε_j, its supremum $\tilde{\pi}_j(\sigma_{-j}, x_j, \varepsilon_j, \theta)$ is a continuous function. Continuity in x_j and ε_j follows from Assumption 3.1(iii). Continuity in σ follows because, by definition,

$$\pi_j(\sigma, x_j, \varepsilon_j, \theta) \equiv \sum_{t^k \in \mathcal{Y}} \left[\prod_{j=1}^{J} \sigma_j(t_j^k) \right] \pi_j(t^k, x_j, \varepsilon_j, \theta),$$

where $t^k \equiv (t_1^k, \ldots, t_J^k)$, $k = 1, \ldots, \kappa_{\mathcal{Y}}$ and $\mathcal{Y}$ can be ordered arbitrarily so that $\mathcal{Y} = \{t^1, \ldots, t^{\kappa_{\mathcal{Y}}}\}$. 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). *Q.E.D.*

PROOF OF PROPOSITION 3.3: To see that the expression in equation (3.6) is the outer region proposed by ABJ, observe that $\max(\int [q(\sigma)]_k \, dF(\varepsilon|\underline{x}) : \sigma \in \operatorname{Sel}(S_\theta))$ 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 1, in each region of multiplicity, the mixed strategy profile which yields the highest probability that t^k is the outcome of the game. To see that the expression in equation (3.7) is the outer region proposed by CT, observe that $\min(\int [q(\sigma)]_k \, dF(\varepsilon|\underline{x}) : \sigma \in \operatorname{Sel}(S_\theta))$ 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 1, in each region of multiplicity, the mixed strategy profile which yields the lowest probability that t^k is the outcome of the game.

To obtain Θ_O^{ABJ} by solving the maximization problem in equation (3.3) over the restricted set of u's equal to the canonical basis vectors in $\Re^{\kappa_{\mathcal{Y}}}$, take the vector $u^k \in \Re^{\kappa_{\mathcal{Y}}}$ to have all entries equal to zero except entry k, which is equal to 1. Then

$$\begin{aligned} \mathbf{P}(y = t^k|\underline{x}) &= u^{k\prime}\mathbf{P}(y|\underline{x}) \leq h(\mathbb{E}(Q_\theta|\underline{x}), u^k) \\ &= \max\big(\mathbf{E}\big([q(\sigma)]_k|\underline{x}\big) : \sigma \in \operatorname{Sel}(S_\theta)\big). \end{aligned}$$

To obtain Θ_O^{CT} by solving the maximization problem in equation (3.3) over the restricted set of u's equal to the canonical basis vectors in $\Re^{\kappa_{\mathcal{Y}}}$ and each of these vectors multiplied by -1, observe that 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 -1. Then

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

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

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

Q.E.D.

PROOF OF THEOREM 4.1: It follows from our discussion in Section 2 that $\min_{u\in B} \mathbf{E}[h(Q_\theta, u)] = 0$ if and only if $0 \leq h(\mathbb{E}(Q_\theta), u)$ for all $u \in B$, which in turn holds if and only if $[0\ \ 0]' \in \mathbb{E}(Q_\theta)$. By the definition of the Aumann expectation, this holds if and only if $\mathbf{E}(q) = [0\ \ 0]'$ for some $q \in \mathrm{Sel}(Q_\theta)$. This is equivalent to saying that a candidate θ belongs to Θ_I if and only if a selection (y, x) of $(Y \times X)$ yields, together with θ, a prediction error $\varepsilon = y - \theta_1 - \theta_2 x$ such that $\mathbf{E}(\varepsilon) = 0$ and $\mathbf{E}(\varepsilon x) = 0$. 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 previous step and Lemma 2.1 also verify Assumption 2.5. *Q.E.D.*

A.3. *Computational Aspects of the Problem*

In this section, we focus on games of complete information. The case of games of incomplete information can be treated analogously, and we refer to Grieco (2009) for a thorough discussion of how to compute the set of Bayesian Nash equilibria. The case of BLP with interval data is straightforward.

When computing Θ_I (or Θ_O^{ABJ} and Θ_O^{CT}), one faces three challenging tasks. We describe them here, and note how each task is affected by the number of players, the number of strategy profiles, and the presence of covariates in the payoff functions. For comparison purposes, we also discuss the differences in computational costs associated with our methodology versus those associated with ABJ's and CT's methodology.

The first step in the procedure requires one to compute the set of all MSNE for given realizations of the payoff shifters, $S_\theta(\underline{x}, \varepsilon)$. This is a computationally challenging problem, although a well studied one which can be performed using the Gambit software described by McKelvey and McLennan (1996).[12] The complexity of this task grows quickly with the number of players and the number of actions that each player can choose from. Notice, however, 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, as in Bajari, Hong, and Ryan (2010), 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. For example, Bajari, Hong, and Ryan (2010) worked with an empirical application which has a very large number of players, but they grouped the smaller ones together to reduce the number of players to 4. Similarly, application of our method to games with multiple *mixed* strategies Nash equilibria requires a limited number of players.[13]

The second task involves verifying whether a candidate $\theta \in \Theta$ is in the region of interest. The difficulty of this task varies depending on whether one wants to check that $\theta \in \Theta_I$, or that $\theta \in \Theta_O^{\mathrm{ABJ}}$ or $\theta \in \Theta_O^{\mathrm{CT}}$. As established in Proposition 3.3, in all cases one needs to work with $\mathbf{E}[h(Q_\theta, u)|\underline{x}]$, so we first describe, for a given $u \in \Re^{\kappa_{\mathcal{Y}}}$, how to obtain this quantity by simulation (see, e.g., McFadden (1989) and Pakes and Pollard (1989)). Recall that for given

[12]The Gambit software can be downloaded for free at http://www.gambit-project.org/. Bajari, Hong, and Ryan (2010) recommend the use of this software to compute the set of mixed strategy Nash equilibria in finite normal form games.

[13]On the other hand, our method is applicable to models with a larger set of players, when players are restricted to playing pure strategies, or when the game is one of incomplete information.

$\theta \in \Theta$ and realization of $\underline{x}$,

$$\mathbf{E}[h(Q_\theta, u)|\underline{x}] = \mathbf{E}\Big[\max_{\sigma \in S_\theta(\underline{x},\varepsilon)} u' q(\sigma)\big|\underline{x}\Big]$$

$$= \int \max_{\sigma \in S_\theta(\underline{x},\varepsilon)} \sum_{k=1}^{\kappa_{\mathcal{Y}}} u_k \prod_{j=1}^{J} \sigma_j(t_j^k)\, dF_\theta(\varepsilon|\underline{x}),$$

where $u' = [u_1\ u_2\ \cdots\ u_{\kappa_{\mathcal{Y}}}]$ and $\mathcal{Y} = \{t^1, \dots, t^{\kappa_{\mathcal{Y}}}\}$ is the set of possible outcomes of the game. One can simulate this multidimensional integral using the following procedure.[14] Let $\mathcal{X}$ denote the support of $\underline{x}$. For any $\underline{x} \in \mathcal{X}$, draw realizations of ε, denoted ε^r, $r = 1, \dots, R$, according to the distribution $F(\varepsilon|\underline{x})$ with identity covariance matrix. These draws stay fixed throughout the remaining steps. Transform the realizations ε^r, $r = 1, \dots, R$, into draws with covariance matrix specified by θ. For each ε^r, compute the payoffs $\pi_j(\cdot, x_j, \varepsilon_j^r, \theta)$, $j = 1, \dots, J$, and obtain the set $S_\theta(\underline{x}, \varepsilon^r)$. Then compute the set $Q_\theta(\underline{x}, \varepsilon^r)$ as the set of multinomial distributions over outcome profiles implied by each element of $S_\theta(\underline{x}, \varepsilon^r)$. Pick a $u \in \Re^{\kappa_{\mathcal{Y}}}$, compute the support function $h(Q_\theta(\underline{x}, \varepsilon^r), u)$, and average it over a large number of draws of ε^r. Call the resulting average $\hat{\mathbf{E}}_R[h(Q_\theta, u)|\underline{x}]$. Note that $\mathbf{E}_{F_\theta(\varepsilon|\underline{x})}(\hat{\mathbf{E}}_R[h(Q_\theta, u)|\underline{x}]) = \mathbf{E}[h(Q_\theta, u)|\underline{x}]$ because each summand is a function of ε^r and these are i.i.d. draws from the distribution $F_\theta(\varepsilon|\underline{x})$.

Having obtained $\hat{\mathbf{E}}_R[h(Q_\theta, u)|\underline{x}]$, to verify whether $\theta \in \Theta_O^{\mathrm{ABJ}}$ and $\theta \in \Theta_O^{\mathrm{CT}}$, it suffices to verify conditional moment inequalities involving, respectively, $\kappa_{\mathcal{Y}}$ and $2\kappa_{\mathcal{Y}}$ evaluations of $\hat{\mathbf{E}}_R[h(Q_\theta, u)|\underline{x}]$, which correspond to the choices of u detailed in Proposition 3.3. As illustrated in our examples, however, using only these inequalities may lead to outer regions which are much larger than Θ_I. Verifying whether $\theta \in \Theta_I$ using the method described in this paper involves solving $\max_{u \in B}(u'F(y|\underline{x}) - \hat{\mathbf{E}}_R[h(Q_\theta, u)|\underline{x}])$ and checking whether the resulting value function is equal to zero for each value of $\underline{x}$ (see Theorem B.1 in the Supplemental Material for a further reduction in the dimensionality of this maximization problem). We emphasize that the dimensionality of u does *not* depend in any way on the number of equilibria of the game (just on the number of players and strategies) or on the number R of draws of ε taken to simulate $\mathbf{E}[h(Q_\theta, u)|\underline{x}]$. As stated before, for given $\underline{x} \in \mathcal{X}$, the criterion function to be maximized is concave and the maximization occurs over a convex subset of $\Re^{\kappa_{\mathcal{Y}}-1}$. In a two player entry game with payoffs linear in $\underline{x}$, we have experienced that efficient algorithms in convex programming, such as the CVX software for MatLab (Grant and Boyd (2010)), can solve this maximization problem with

[14] The procedure described here is very similar to the one proposed by Ciliberto and Tamer (2009). When the assumptions maintained by Bajari, Hong, and Ryan (2010, Section 3) are satisfied, their algorithm can be used to significantly reduce the computational burden associated with simulating the integral.

a handful of iterations, on the order of 10–25, depending on the candidate θ. We have also experienced that a simple Nelder–Mead algorithm programmed in Fortran 90 works very well, yielding the usual speed advantages of Fortran over MatLab. When $\underline{x}$ is discrete, for each parameter candidate, the above maximization problem needs to be solved for all possible value of $\underline{x} \in \mathcal{X}$, and one needs to check whether all required conditions are satisfied. Therefore, it is reasonable to say that the computational burden of this stage is linear in the number of values that $\underline{x}$ takes. When $\underline{x}$ is continuous, one can apply the methodology of Andrews and Shi (2009) as detailed in the Supplemental Material.

Finally, the region of interest needs to be computed. This means that the researcher should search over the parameter space Θ and collect all the points in Θ_I or Θ_O^{CT} or Θ_O^{ABJ}. This is of course a theoretical set and, in practice, the researcher seeks to collect enough points that belong to the region of interest, such that it can be covered reasonably well. While easy to program, a grid search over Θ is highly inefficient, especially when Θ belongs to a high-dimensional space. CT proposed to search over Θ using a method based on simulated annealing. In this paper, we use an alternative algorithm called differential evolution. We give here a short description of this method, focusing mainly on its complexity. We refer to Price, Storn, and Lampinen (2004) for further details. Differential evolution (DE) is a type of genetic algorithm that is often used to solve optimization problems. The algorithm starts from a population of N_p points picked randomly from the set Θ. It then updates this list of points at each stage, creating a new generation of N_p points to replace the previous one. A candidate to replace a current member of the population (child) is created by combining information from members of the current population (parents). This new candidate is accepted into the population as a replacement for a current member if it satisfies a certain criterion. In our application, the criterion for being admitted into the new generation is to be a member of Θ_I (or Θ_O^{CT} or Θ_O^{ABJ}, when computing each of these regions). The process of finding a replacement for each of the current N_p points is repeated N times, yielding $N \cdot N_p$ maximizations of the criterion function (respectively, evaluation of the conditional inequalities for CT and ABJ). During this process, we record the points which were found to belong to the regions of interest. In our simulations, we experienced that this method explores Θ in a very efficient way. Price, Storn, and Lampinen (2004) recommended for N_p to grow linearly with the dimensionality of Θ. The number of iterations (generations) N depends on how well one wants to cover the region of interest. For example, in a two player entry game with $\Theta \subset \Re^4$, we found that setting $N_p = 40$ and $N = 1000$ gave satisfactory results, and when N was increased to 5000, the regions of interest seemed to be very well covered, while the projections on each component of θ remained very similar to what we obtained with $N = 1000$. Creating candidates to replace members of the population involves trivial algebraic operations whose number grows linearly with the dimensionality of Θ.

These operations involve picking two tuning parameters, but satisfactory rules of thumb exist in the literature; see Price, Storn, and Lampinen (2004).

REFERENCES

ANDREWS, D. W. K., AND X. SHI (2009): "Inference Based on Conditional Moment Inequalities," Mimeo, Yale University. [1789,1795,1803,1812,1818]

ANDREWS, D. W. K., AND G. SOARES (2010): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 78, 119–157. [1789]

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, Yale University. [1786]

ARADILLAS-LOPEZ, A., AND E. TAMER (2008): "The Identification Power of Equilibrium in Simple Games," *Journal of Business & Economic Statistics*, 26, 261–310. [1789]

ARTSTEIN, Z. (1983): "Distributions of Random Sets and Random Selections," *Israel Journal of Mathematics*, 46, 313–324. [1788]

AUMANN, R. J. (1965): "Integrals of Set Valued Functions," *Journal of Mathematical Analysis and Applications*, 12, 1–12. [1792,1793]

BAJARI, P., H. HONG, AND S. RYAN (2010): "Identification and Estimation of a Discrete Game of Complete Information," *Econometrica*, 78, 1529–1568. [1796,1797,1816,1817]

BERESTEANU, A., AND F. MOLINARI (2006): "Asymptotic Properties for a Class of Partially Identified Models," Working Paper CWP10/06, CeMMAP. [1788]

——— (2008): "Asymptotic Properties for a Class of Partially Identified Models," *Econometrica*, 76, 763–814. [1788,1789,1810,1811]

BERESTEANU, A., I. MOLCHANOV, AND F. MOLINARI (2008): "Sharp Identification Regions in Games," Working Paper CWP15/08, CeMMAP. [1785]

——— (2009): "Sharp Identification Regions in Models With Convex Predictions: Games, Individual Choice, and Incomplete Data," Working Paper CWP27/09, CeMMAP. [1785, 1803]

——— (2011): "Supplement to 'Sharp Identification Regions in Models With Convex Moment Predictions'," *Econometrica Supplemental Material*, 79, http://www.econometricsociety.org/ecta/Supmat/8680_extensions.pdf; http://www.econometricsociety.org/ecta/Supmat/8680_data_and_programs.zip. [1788,1789]

BERRY, S. T. (1992): "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60, 889–917. [1797]

BERRY, S. T., AND E. TAMER (2007): "Identification in Models of Oligopoly Entry," in *Advances in Economics and Econometrics: Theory and Application, Ninth World Congress*, Vol. II. Cambridge: Cambridge University Press, Chapter 2, 46–85. [1792,1796,1797,1802,1803]

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," Working Paper 557, CalTech DHSS. [1797]

BONTEMPS, C., T. MAGNAC, AND E. MAURIN (2011): "Set Identified Linear Models," *Econometrica* (forthcoming). [1810]

BOYD, S., AND L. VANDENBERGHE (2004): *Convex Optimization*. Cambridge: Cambridge University Press. [1788,1795]

BRESNAHAN, T. F., AND P. C. REISS (1987): "Do Entry Conditions Vary Across Markets?" *Brookings Papers on Economic Activity*, 18, 833–882. [1797]

——— (1990): "Entry in Monopoly Markets," *Review of Economic Studies*, 57, 531–553. [1797]

——— (1991): "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99, 977–1009. [1797]

BUGNI, F. A. (2010): "Bootstrap Inference in Partially Identified Models Defined by Moment Inequalities: Coverage of the Identified Set," *Econometrica*, 78, 735–753. [1789]

CANAY, I. A. (2010): "EL Inference for Partially Identified Models: Large Deviations Optimality and Bootstrap Validity," *Journal of Econometrics*, 156, 408–425. [1789]

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. [1789]

——— (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75, 1243–1284. [1789]

CHERNOZHUKOV, V., S. LEE, AND A. ROSEN (2009): "Intersection Bounds: Estimation and Inference," Working Paper CWP19/09, CeMMAP. [1789]

CILIBERTO, F., AND E. TAMER (2009): "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 77, 1791–1828. [1786,1817]

DYER, M., A. FRIEZE, AND R. KANNAN (1991): "A Random Polynomial-Time Algorithm for Approximating the Volume of Convex Bodies," *Journal of the Association for Computing Machinery*, 38, 1–17. [1795]

DYNKIN, E. B., AND I. V. EVSTIGNEEV (1976): "Regular Conditional Expectations of Correspondences," *Theory of Probability and Its Applications*, 21, 325–338. [1793]

GALICHON, A., AND M. HENRY (2006): "Inference in Incomplete Models," Mimeo, École Polytechnique and Université de Montréal. [1788]

——— (2009a): "Set Identification in Models With Multiple Equilibria," Mimeo, École Polytechnique and Université de Montréal. [1788]

——— (2009b): "A Test of Non-Identifying Restrictions and Confidence Regions for Partially Identified Parameters," *Journal of Econometrics*, 152, 186–196. [1789]

GRANT, M., AND S. BOYD (2008): "Graph Implementations for Nonsmooth Convex Programs," in *Recent Advances in Learning and Control*, ed. by V. Blondel, S. Boyd, and H. Kimura. London: Springer-Verlag, 95–110, available at http://www.stanford.edu/~boyd/papers/graph_dcp.html. [1788]

——— (2010): "CVX: Matlab Software for Disciplined Convex Programming, Version 1.21," available at http://cvxr.com/cvx. [1795,1817]

GRIECO, P. L. E. (2009): "Discrete Choice Games With Flexible Information Structure: An Application to Local Grocery Markets," Mimeo, Pennsylvania State University. [1803,1816]

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. [1786]

KIM, K.-I. (2009): "Set Estimation and Inference With Models Characterized by Conditional Moment Inequalities," Mimeo, University of Minnesota. [1789]

LOVÁSZ, L., AND S. VEMPALA (2006): "Simulated Annealing in Convex Bodies and an $O^*(n^4)$ Volume Algorithm," *Journal of Computer and System Sciences*, 72, 392–417. [1795]

MANSKI, C. F. (1989): "Anatomy of the Selection Problem," *Journal of Human Resources*, 24, 343–360. [1787]

——— (2003): *Partial Identification of Probability Distributions*. New York: Springer-Verlag. [1786,1787,1792,1815]

MANSKI, C. F., AND E. TAMER (2002): "Inference on Regressions With Interval Data on a Regressor or Outcome," *Econometrica*, 70, 519–546. [1787,1792]

MAZZEO, M. (2002): "Product Choice and Oligopoly Market Structure," *RAND Journal of Economics*, 33, 221–242. [1797]

MCFADDEN, D. (1989): "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 57, 995–1026. [1816]

MCKELVEY, R. D., AND A. MCLENNAN (1996): "Computation of Equilibria in Finite Games," in *Handbook of Computational Economics*, Vol. 1. Amsterdam: North-Holland, Chapter 2, 87–142. [1816]

MOLCHANOV, I. (2005): *Theory of Random Sets*. London: Springer-Verlag. [1787,1793,1794,1811, 1813,1814]

MOLINARI, F. (2008): "Partial Identification of Probability Distributions With Misclassified Data," *Journal of Econometrics*, 144, 81–117. [1787]

PAKES, A., AND D. POLLARD (1989): "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57, 1027–1057. [1816]

PAKES, A., J. PORTER, K. HO, AND J. ISHII (2006): "Moment Inequalities and Their Application," Mimeo, Harvard University. [1789]

PONOMAREVA, M. (2010): "Inference in Models Defined by Conditional Moment Inequalities With Continuous Covariates," Mimeo, University of Western Ontario. [1789]

PRICE, K. V., R. M. STORN, AND J. A. LAMPINEN (2004): *Differential Evolution: A Practical Approach to Global Optimization*. Natural Computing Series. Berlin: Springer-Verlag, with 1 CD-ROM (Windows, Macintosh, and UNIX). [1818,1819]

RAO, M. M., AND R. J. SWIFT (2006), *Probability Theory With Applications*. Mathematics and Its Applications, Vol. 582. New York: Springer-Verlag. [1813]

ROCKAFELLAR, R. T. (1970): *Convex Analysis*. Princeton, NJ: Princeton University Press. [1787, 1794]

ROMANO, J. P., AND A. M. SHAIKH (2010): "Inference for the Identified Set in Partially Identified Econometric Models," *Econometrica*, 78, 169–211. [1789]

ROSEN, A. M. (2008): "Confidence Sets for Partially Identified Parameters That Satisfy a Finite Number of Moment Inequalities," *Journal of Econometrics*, 146, 107–117. [1789]

SCHNEIDER, R. (1993): *Convex Bodies: The Brunn–Minkowski Theory*. Cambridge: Cambridge University Press. [1787,1794]

TAMER, E. (2003): "Incomplete Simultaneous Discrete Response Model With Multiple Equilibria," *Review of Economic Studies*, 70, 147–165. [1797]

——— (2010): "Partial Identification in Econometrics," *Annual Reviews of Economics*, 2, 167–195. [1792]

WILSON, R. (1971): "Computing Equilibria of N-Person Games," *SIAM Journal on Applied Mathematics*, 21, 80–87. [1801]

Dept. of Economics, University of Pittsburgh, 4510 Posvar Hall, Pittsburgh, PA 15260, U.S.A.; arie@pitt.edu,

Institute of Mathematical Statistics and Actuarial Science, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland; ilya@stat.unibe.ch,

and

Dept. of Economics, Cornell University, 492 Uris Hall, Ithaca, NY 14853-7601, U.S.A.; fm72@cornell.edu.

Manuscript received July, 2009; final revision received March, 2011.

Econometrica Supplementary Material

SUPPLEMENTAL TO "SHARP IDENTIFICATION REGIONS IN MODELS WITH CONVEX MOMENT PREDICTIONS"

(*Econometrica*, Vol. 79, No. 6, November 2011, 1785–1821)

BY ARIE BERESTEANU, ILYA MOLCHANOV, AND FRANCESCA MOLINARI

OUTLINE

THIS SUPPLEMENT INCLUDES four appendices. Appendix B establishes that the methodology of Andrews and Shi (2009) can be applied in our context to obtain confidence sets that uniformly cover each element of the sharp identification region with a prespecified asymptotic probability. Appendix C shows that our approach easily applies also to finite games of incomplete information and characterizes Θ_I through a finite number of moment inequalities. Appendix D specializes our results in the context of complete information games, to the case that players are restricted to use pure strategies only and Nash equilibrium is the solution concept. In this case, Θ_I is characterized through a finite number of moment inequalities, and further insights are provided on how to reduce the number of inequalities to be checked so as to compute it. Appendix E shows that our methodology is applicable to static simultaneous-move finite games regardless of the solution concept used.[1] Appendix F applies the results in Section 2 of the main paper to the analysis of individual decision making, looking at random utility models of multinomial choice in the presence of interval regressors data.

APPENDIX B: APPLICABILITY OF ANDREWS AND SHI'S GENERALIZED MOMENT SELECTION PROCEDURE[2]

B.1. *Finite Games of Complete and Incomplete Information*

Andrews and Shi (2009, Section 9; AS henceforth) considered conditional moment inequality problems of the form $\mathbf{E}(m_d(y, \underline{x}, \theta, u)|\underline{x}) \geq 0$ for all $u \in B$, $\underline{x}$-a.s., $d = 1, \dots, D$. They showed that the conditional moment inequalities can be transformed into equivalent unconditional moment inequalities, by choosing appropriate weighting functions (instruments) $g \in \mathcal{G}$, with $\mathcal{G}$ a collection of instruments and g that depend on $\underline{x}$. This yields $\mathbf{E}(m_d(y, \underline{x}, \theta, g, u)) \geq 0$ for all $u \in B$, $g = [g_1, \dots, g_D]' \in \mathcal{G}$, and $d = 1, \dots, D$, where $m_d(y, \underline{x}, \theta, g, u) = m_d(y, \underline{x}, \theta, u)g(\underline{x})$. In the models that we analyzed in Section 3 and in Appendix C below, the conditional moment inequalities are of the $\leq$ type, and

$$m(y, \underline{x}, \theta, u) = u'[1(y = t^k), k = 1, \dots, \kappa_y] - \mathbf{E}[h(Q_\theta, u)|\underline{x}],$$

[1]Specifically, we illustrate this by looking at games where rationality of level 1 is the solution concept (a problem first studied by Aradillas-Lopez and Tamer (2008)) and by looking at games where correlated equilibrium is the solution concept.

[2]We are grateful to Xiaoxia Shi for several discussions that helped us develop this section.

© 2011 The Econometric Society DOI: 10.3982/ECTA8680

$$m(y, \underline{x}, \theta, g, u) = \big(u'[1(y = t^k), k = 1, \dots, \kappa_{\mathcal{Y}}] \\ - \mathbf{E}[h(Q_\theta, u)|\underline{x}]\big)g(\underline{x}).$$

Notice that $\mathbf{E}[h(Q_\theta, u)|\underline{x}]$ is a known (or simulated) function of θ, u, and $\underline{x}$, and that for each $u \in B$, we have only one inequality. Notice also that by the positive homogeneity of the support function, our moment inequalities can be written equivalently as $\mathbf{E}(m(y, \underline{x}, \theta, g, u)) \leq 0$ for all $g \in \mathcal{G}$ and $u \in S \equiv \{u \in \Re^{\kappa_{\mathcal{Y}}} : \|u\| = 1\}$. Hence, they are invariant to rescaling of the moment function, which is important for finite sample inference (see, e.g., Andrews and Soares (2010)).

In all that follows, to simplify the exposition, we abstract from the choice of $\mathcal{G}$. Once we establish that our problem fits into the general framework of AS, we can choose instruments g as detailed in Section 3 of AS. To avoid ambiguity, in this section we denote $F(y|\underline{x}) \equiv [\mathbf{P}(y = t^k|\underline{x}), k = 1, \dots, \kappa_{\mathcal{Y}}]$. We first establish that Θ_I can be equivalently defined using only the first $\kappa_{\mathcal{Y}} - 1$ entries of $\mathcal{Y}$, thereby avoiding the problems for inference associated with linear dependence among the entries of $F(y|\underline{x})$ and also lowering the dimension over which the maximization is performed. Let $\widetilde{F}(y|\underline{x})$ denote the first $\kappa_{\mathcal{Y}} - 1$ rows of $F(y|\underline{x})$, $B^{\kappa_{\mathcal{Y}}-1} = \{u \in \Re^{\kappa_{\mathcal{Y}}-1} : \|u\| \leq 1\}$, $S^{\kappa_{\mathcal{Y}}-1} = \{u \in \Re^{\kappa_{\mathcal{Y}}-1} : \|u\| = 1\}$, and

$$\tilde{Q}_\theta = \big\{\tilde{q} = [[q(\sigma)]_k, k = 1, \dots, \kappa_{\mathcal{Y}} - 1], \sigma \in \mathrm{Sel}(S_\theta)\big\}.$$

THEOREM B.1: *Let Assumptions* 3.1 (*or* C.1 *below*) *and* 3.2 *hold*. *Then*

$$\begin{aligned}\tilde{\Theta}_I &\equiv \Big\{\theta \in \Theta : \max_{u \in B^{\kappa_{\mathcal{Y}}-1}} \big(u'\widetilde{F}(y|\underline{x}) - \mathbf{E}[h(\tilde{Q}_\theta, u)|\underline{x}]\big) = 0, \underline{x}\text{-}a.s.\Big\} \\ &= \Big\{\theta \in \Theta : \Big[\max_{u \in S^{\kappa_{\mathcal{Y}}-1}} \big(u'\widetilde{F}(y|\underline{x}) - \mathbf{E}[h(\tilde{Q}_\theta, u)|\underline{x}]\big)\Big]_+ = 0, \underline{x}\text{-}a.s.\Big\} \\ &= \Theta_I.\end{aligned}$$

PROOF: The equality between the two representations above follows by standard arguments; see, for example, Beresteanu and Molinari (2008, Lemma A.1). To establish that $\tilde{\Theta}_I = \Theta_I$, observe that $\theta \in \tilde{\Theta}_I$ if and only if $\widetilde{F}(y|\underline{x}) \in \mathbb{E}(\tilde{Q}_\theta|\underline{x})$. Pick $\theta \in \Theta_I$. Then $F(y|\underline{x}) = \mathbf{E}(q|\underline{x})$ for some $q \in \mathrm{Sel}(Q_\theta)$. Notice that this implies $\widetilde{F}(y|\underline{x}) = \mathbf{E}(\tilde{q}|\underline{x})$ for $\tilde{q} \in (\tilde{Q}_\theta)$; hence, $\theta \in \tilde{\Theta}_I$. Conversely, pick $\theta \in \tilde{\Theta}_I$. Then $\widetilde{F}(y|\underline{x}) = \mathbf{E}(\tilde{q}|\underline{x})$ for some $\tilde{q} \in \mathrm{Sel}(\tilde{Q}_\theta)$, which in turn implies that $q = [\tilde{q}; 1 - \sum_{k=1}^{\kappa_{\mathcal{Y}}-1} \tilde{q}] \in \mathrm{Sel}(Q_\theta)$ and $F(y|\underline{x}) = \mathbf{E}(q|\underline{x})$; hence, $\theta \in \Theta_I$. *Q.E.D.*

AS proposed a confidence set with nominal value $1 - \alpha$ for the true parameter vector as

$$\mathrm{CS}_n = \{\theta \in \Theta : T_n(\theta) \leq c_{n,1-\alpha}(\theta)\},$$

where $T_n(\theta)$ is a test statistic and $c_{n,1-\alpha}(\theta)$ is a corresponding critical value for a test with nominal significance level α. AS established that, under certain assumptions, this confidence set has correct uniform asymptotic size.[3] To apply the construction in AS, we maintain the following assumption:

ASSUMPTION B.1: *The researcher observes an i.i.d. sequence of equilibrium outcomes and observable payoff shifters* $\{y_i, \underline{x}_i\}_{i=1}^n$. *Define* $\tilde{\Sigma}_{\underline{x}} = \operatorname{diag}(\widetilde{F}(y|\underline{x})) - \widetilde{F}(y|\underline{x})\widetilde{F}(y|\underline{x})'$ *and let* $\tilde{\Sigma}_{\underline{x}}$ *be nonsingular with* $a < \|\tilde{\Sigma}_{\underline{x}}\| < b$, $\underline{x}$*-a.s. for some constants* $0 < a < b < \infty$, *where* $\|\tilde{\Sigma}_{\underline{x}}\|$ *is a matrix norm for* $\tilde{\Sigma}_{\underline{x}}$ *compatible with the Euclidean norm*.

AS proposed various criterion functions T_n: some of the Cramér–von Mises type, some of the Kolmogorov–Smirnov type. Here, we work with a mix of Cramér–von Mises and Kolmogorov–Smirnov statistic using a modification of the function S_1 on page 10 of AS. Specifically, we use

$$
\begin{aligned}
\text{(B.1)}\qquad T_n(\theta) &= \int \Big(\max_{u \in B^{\kappa_{\mathcal{Y}}-1}} \sqrt{n}\bar{m}_n(\theta, g, u) \Big)^2 \, d\Gamma \\
&= \int \Big(\max_{u \in S^{\kappa_{\mathcal{Y}}-1}} \sqrt{n}\bar{m}_n(\theta, g, u) \Big)_+^2 \, d\Gamma \\
&= \int \max_{u \in S^{\kappa_{\mathcal{Y}}-1}} (\sqrt{n}\bar{m}_n(\theta, g, u))_+^2 \, d\Gamma,
\end{aligned}
$$

where Γ denotes a probability measure on $\mathcal{G}$ whose support is $\mathcal{G}$ as detailed in Section 3 of AS, the second equality follows from the proof of Theorem B.1, and

$$
\begin{aligned}
&\bar{m}_n(\theta, g, u) = \frac{1}{n}\sum_{i=1}^{n} (u'w(y_i) - f(\underline{x}_i, \theta, u)) g(\underline{x}_i), \\
&f(\underline{x}_i, \theta, u) = \mathbf{E}[h(\tilde{Q}_\theta, u)|\underline{x}_i], \\
&w(y_i) = [1(y_i = t^k), k = 1, \ldots, \kappa_{\mathcal{Y}} - 1],
\end{aligned}
$$

[3] Imbens and Manski (2004) discussed the difference between confidence sets that uniformly cover the true parameter vector with a prespecified asymptotic probability, and confidence sets that uniformly cover Θ_I (see also Stoye (2009)). Providing methodologies to obtain asymptotically valid confidence sets of either type when the conditioning variables have a continuous distribution is a developing area of research, to which the method of AS belongs. In certain empirically relevant models (see, for example, Appendix C and Appendix D), the characterization in Theorem 2.1 yields a finite number of (conditional) moment inequalities. In such cases, the methods of Chernozhukov, Hong, and Tamer (2007) and Romano and Shaikh (2010) can be applied after discretizing the conditioning variables to obtain confidence sets which cover Θ_I with a prespecified asymptotic probability, uniformly in the case of Romano and Shaikh (2010). Ciliberto and Tamer (2009) verified the required regularity conditions for finite games of complete information.

so that $\bar{m}_n(\theta, g, u)$ is the sample analog of a version of $\mathbf{E}(m(y, \underline{x}, \theta, g, u))$, which is based on the first $\kappa_{\mathcal{Y}} - 1$ entries of $\mathcal{Y}$ and on $\tilde{Q}_\theta$. Note that by the same argument which follows, our problem specified as in equation (3.6) corresponds to the Cramér–von Mises test statistic of AS, with modified function S_1.

Below we show that our modified function S_1 satisfies Assumptions S1–S4 of AS and that Assumption M2 of AS is also satisfied. This establishes that their generalized moment selection procedure with infinitely many conditional moment inequalities is applicable. We note that one can take the confidence set CS_n applied with confidence level 1/2 to obtain half-median-unbiased estimated sets; see AS and Chernozhukov, Lee, and Rosen (2009). Finally, one can also take the criterion function in Theorem B.1, replace there $\widetilde{F}(y|\underline{x})$ with its sample analog, and construct a Hausdorff-consistent estimator of Θ_I using the methodology proposed by Chernozhukov, Hong, and Tamer (2007, equation (3.2) and Theorem 3.1). To see that their results are applicable, recall that the payoff functions are assumed to be continuous in (x_j, ε_j). Hence, the Nash equilibrium correspondence has a closed graph; see Fudenberg and Tirole (1991, Section 1.3.2). This implies that Q_θ has a closed graph and, therefore, the same is true for $\mathbb{E}(Q_\theta|\underline{x})$; see Aumann (1965, Corollary 5.2). In turn, this yields $\limsup_{\theta_n \to \theta} \mathbb{E}(Q_{\theta_n}|\underline{x}) \subseteq \mathbb{E}(Q_\theta|\underline{x})$. Observe that

$$\max_{u \in B^{\kappa_{\mathcal{Y}}-1}} \big(u'\widetilde{F}(y|\underline{x}) - \mathbf{E}[h(\tilde{Q}_\theta, u)|\underline{x}]\big) = d_H(\widetilde{F}(y|\underline{x}), \mathbb{E}(\tilde{Q}_\theta|\underline{x})).$$

The criterion function $s(\theta) \equiv \int d_H(\widetilde{F}(y|\underline{x}), \mathbb{E}(\tilde{Q}_\theta|\underline{x}))\, dF_{\underline{x}}$, with $F_{\underline{x}}$ the probability distribution of $\underline{x}$ (or a probability measure which dominates it), is therefore lower semicontinuous in θ, because

$$\begin{aligned} \liminf_{\theta_n \to \theta} s(\theta_n) &\geq \int \liminf_{\theta_n \to \theta} d_H(\widetilde{F}(y|\underline{x}), \mathbb{E}(\tilde{Q}_{\theta_n}|\underline{x}))\, dF_{\underline{x}} \\ &\geq \int d_H(\widetilde{F}(y|\underline{x}), \limsup \mathbb{E}(\tilde{Q}_{\theta_n}|\underline{x}))\, dF_{\underline{x}} \\ &\geq \int d_H(\widetilde{F}(y|\underline{x}), \mathbb{E}(\tilde{Q}_\theta|\underline{x}))\, dF_{\underline{x}} = s(\theta). \end{aligned}$$

Conditions (c)–(e) in Assumption C1 of Chernozhukov, Hong, and Tamer (2007) are verified by standard arguments.

We now verify AS's assumptions.

THEOREM B.2: *Let Assumption* B.1 *hold*. *Then Assumptions* S1–S4 *and* M2 *of AS are satisfied*.

PROOF: Assumption S1(a) follows because the moment inequalities are defined for $u \in S^{\kappa_{\mathcal{Y}}-1}$; hence any rescaling of the moment function is absorbed by a corresponding rescaling in u. The rest of Assumption S1 and Assumptions S2–S4 are verified by AS. To verify Assumption M2, observe that

$$\widetilde{m}(y, \underset{\sim}{x}, \theta, u) \equiv u'w(y) - f(\underset{\sim}{x}, \theta, u)$$

is given by the sum of a linear function of u and a Lipschitz function of u, with Lipschitz constant equal to 1. It is immediate that the processes $\{u'w(y_{in}),\ u \in S^{\kappa_{\mathcal{Y}}-1},\ i \leq n,\ n \geq 1\}$ satisfy Assumption M2. We now show that the same holds for the processes $\{f(\underset{\sim}{x}_{in}, \theta_n, u),\ u \in S^{\kappa_{\mathcal{Y}}-1},\ i \leq n,\ n \geq 1\}$. Assumption M2(a) holds because for all $u \in S^{\kappa_{\mathcal{Y}}-1}$,

$$\begin{aligned}\left|\frac{f(\underset{\sim}{x}, \theta, u)}{\operatorname{Var}(\widetilde{m}(y, \underset{\sim}{x}, \theta, u))}\right| &\leq \left|\frac{f(\underset{\sim}{x}, \theta, u)}{\mathbf{E}(u'\tilde{\Sigma}_{\underset{\sim}{x}}u)}\right| \leq c\left|\mathbf{E}[h(\tilde{Q}_\theta, u)|\underset{\sim}{x}]\right| \\ &\leq c\mathbf{E}(\|\tilde{Q}_\theta\|_H|\underset{\sim}{x}) \leq c, \quad \underset{\sim}{x}\text{-a.s.},\end{aligned}$$

where the first inequality follows from the variance decomposition formula, c is a constant that depends on a and b from Assumption B.1, and the last inequality follows by recalling that $\tilde{Q}_\theta$ takes its realizations in the unit simplex which is a subset of the unit ball. Assumption M2(b) follows immediately because the envelope function is a constant. Assumption M2(c) is verified by observing that $f(\underset{\sim}{x}, \theta, u)$ is Lipschitz in u, with Lipschitz constant equal to 1. By Lemma 2.13 in Pakes and Pollard (1989), the class of functions $\{f(\cdot, u),\ u \in S^{\kappa_{\mathcal{Y}}-1}\}$ is Euclidean with envelope equal to a constant and, therefore, is manageable. Assumption M2 for the processes $\{(u'w(y_{in}) - f(\underset{\sim}{x}_{in}, \theta_n, u)),\ u \in S^{\kappa_{\mathcal{Y}}-1},\ i \leq n,\ n \geq 1\}$ then follows by Lemma E1 of AS. *Q.E.D.*

B.2. *BLP With Interval Outcome and Covariate Data*

We maintain the following assumption:

ASSUMPTION B.2: *The researcher observes an i.i.d. sequence of tuples* $\{y_{iL}, y_{iU}, x_{iL}, x_{iU}\}_{i=1}^n$. $\mathbf{E}(|y_i|^2)$, $\mathbf{E}(|x_j|^2)$, $\mathbf{E}(|y_i x_j|^2)$, *and* $\mathbf{E}(x_j^4)$ *are all finite for each* $i, j = L, U$.

Let $Q_{\theta i}$ be the mapping defined as in equation (5.1) using $(y_{iL}, y_{iU}, x_{iL}, x_{iU})$. Beresteanu and Molinari (2008, Lemmas A.4 and A.5, and proof of Theorem 4.2) established that $\{Q_{\theta i}\}_{i=1}^n$ is a sequence of i.i.d. random closed sets, such that $\mathbf{E}(\|Q_{\theta i}\|_H^2) < \infty$. Define $T_n(\theta)$ similarly to the previous section,

$$\begin{aligned}T_n(\theta) &= \left(\max_{u\in B}(-\sqrt{n}\bar{m}_n(\theta, u))\right)^2 = \left(\max_{u\in S} -\sqrt{n}\bar{m}_n(\theta, u)\right)_+^2 \\ &= \max_{u\in S}(-\sqrt{n}\bar{m}_n(\theta, u))_+^2,\end{aligned}$$

$$\bar{m}_n(\theta, u) = \frac{1}{n}\sum_{i=1}^{n} h(Q_{\theta i}, u),$$

where, again, the fact that $u \in S$ guarantees that the above test statistic is invariant to rescaling of the moment function. This preserves concavity of the objective function. We then have the following result:

THEOREM B.3: *Let Assumptions* 5.1 *and* B.2 *hold*. *Then Assumption* EP *of AS* (*p*. 37) *is satisfied*.

PROOF: Let $m(y_{iL}, y_{iU}, x_{iL}, x_{iU}, \theta, u) = h(Q_{\theta i}, u)$. Following AS notation, define

$$\sqrt{n}\bar{m}_n(\theta, u) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} h(Q_{\theta i}, u),$$

$$\gamma_{1,n}(\theta, u) = \sqrt{n}\mathbf{E}[h(Q_{\theta i}, u)],$$

$$\gamma_2(\theta, u, u^*) = \mathbf{E}[h(Q_{\theta i}, u)h(Q_{\theta i}, u^*)] - \mathbf{E}[h(Q_{\theta i}, u)]\mathbf{E}[h(Q_{\theta i}, u^*)],$$

$$\nu_n(\theta, u) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big[h(Q_{\theta i}, u) - \mathbf{E}(h(Q_{\theta i}, u))\big].$$

Given the above definitions, we have

$$\sqrt{n}\bar{m}_n(\theta, u) = \nu_n(\theta, u) + \gamma_{1,n}(\theta, u).$$

By the central limit theorem for i.i.d. sequences of random sets (Molchanov (2005, Theorem 2.2.1)),

$$\nu_n(\theta, \cdot) \Longrightarrow \nu_{\gamma_2(\theta)}(\cdot),$$

a Gaussian process with mean zero, covariance kernel $\gamma_2(\theta, u, u^*)$, and continuous sample paths. It follows from the strong law of large numbers in Banach spaces of Mourier (1955) that the sample analog estimator $\hat{\gamma}_{2,n}(\theta, u, u^*)$ which replaces population moments with sample averages, satisfies $\hat{\gamma}_{2,n}(\theta, \cdot, \cdot) \stackrel{\text{a.s.}}{\to} \gamma_2(\theta, \cdot, \cdot)$ uniformly in u, u^*. *Q.E.D.*

APPENDIX C: ENTRY GAMES OF INCOMPLETE INFORMATION

We now consider the case that players have incomplete information (see, e.g. Aradillas-López (2010), Brock and Durlauf (2001, 2007), Seim (2006), Sweeting (2009)). We retain the notation introduced in the main paper, but we substitute for Assumption 3.1 the following assumption, which is fairly standard in the literature. We continue to maintain Assumption 3.2.

ASSUMPTION C.1: (i) *The set of outcomes of the game* $\mathcal{Y}$ *is finite. The observed outcome of the game results from simultaneous-move, pure strategy Bayesian Nash play.*

(ii) *All players and the researcher observe payoff shifters* x_j, $j = 1, \dots, J$. *The payoff shifter* ε_j *is private information to player* $j = 1, \dots, J$, *and unobservable to the researcher. Conditional on* $\{x_j, j = 1, \dots, J\}$, ε_j *is independent of* $\{\varepsilon_i\}_{i \neq j}$. *Players have correct common prior* $F_\theta(\varepsilon | \underline{x})$.

(iii) *The payoffs are additively separable in* $\varepsilon : \pi_j(y_j, y_{-j}, x_j, \varepsilon_j; \theta) = \tilde{\pi}_j(y_j, y_{-j}, x_j; \theta) + \varepsilon_j$. *Assumption* 3.1(iii) *holds.*

The independence condition in Assumption C.1(iii) substantially simplifies the task of calculating the set of Bayesian Nash equilibria (BNE). Conceptually, however, our methodology applies also when players' types are correlated. The resulting difficulties associated with calculating the set of BNE are to be faced with any methodology for inference in this class of games. The correct-common-prior condition in Assumption C.1(iii) can be relaxed, but we maintain it here for simplicity.

For the sake of brevity, we restrict attention to two player entry games. However, this restriction is not necessary. Our results easily extend, with appropriate modifications to the notation and the definition of the set of pure strategy Bayesian Nash equilibria, to the case of $J \geq 2$ players, each with $2 \leq \kappa_{\mathcal{Y}_j} < \infty$ strategies. In what follows, we characterize the set of BNE of the game, borrowing from the treatment in Grieco (2009, Section 4), and then apply our methodology to this set.[4] To conserve space, we do not explicitly verify Assumptions 2.1–2.5. Assumptions 2.1–2.3 follow by similar arguments as in Section 3. Assumptions 2.4 and 2.5 follow by the same construction that we provide at the end of Section 3, replacing equation (3.7) with equation (8) in Grieco (2009, Theorem 4).

With incomplete information, players' strategies are decision rules $y_j : \mathcal{E} \to \{0, 1\}$, with $\mathcal{E}$ the support of ε. The set of outcomes of the game is $\mathcal{Y} = \{(0, 0), (1, 0), (0, 1), (1, 1)\}$. Given $\theta \in \Theta$ and a realization of $\underline{x}$ and ε_j, player j enters the market if and only if his expected payoff is nonnegative. Therefore, equilibrium mappings (decision rules) are step functions determined by a threshold: $y_j(\varepsilon_j) = 1(\varepsilon_j \geq t_j)$, $j = 1, 2$. As a result, player j's beliefs about player $-j$'s probability of entry under the common prior assumption is $\int y_{-j}(\varepsilon_{-j})\, dF_\theta(\varepsilon_{-j} | \underline{x}) = 1 - F_\theta(t_{-j} | \underline{x})$ and, therefore, player j's best response

[4] We refer to Grieco (2009) for a thorough discussion of the related literature and of identification problems in games of incomplete information with multiple BNE. See also Berry and Tamer (2007, Section 3).

cutoff is[5]

$$t_j^b(t_{-j}, \underline{x}; \theta) = -\tilde{\pi}_j(1, 0, x_j; \theta)F_\theta(t_{-j}|\underline{x}) \\ - \tilde{\pi}_j(1, 1, x_j; \theta)(1 - F_\theta(t_{-j}|\underline{x})).$$

Hence, the set of equilibria can be defined as the set of cutoff rules

$$T_\theta(\underline{x}) = \{(t_1, t_2) : t_j = t_j^b(t_{-j}, \underline{x}; \theta)\ \forall j = 1, 2\}.$$

Note that the equilibrium thresholds are functions of $\underline{x}$ only. The set $T_\theta(\underline{x})$ might contain a finite number of equilibria (e.g., if the common prior is the Normal distribution) or a continuum of equilibria. For ease of notation we write the set $T_\theta(\underline{x})$ and its realizations, respectively, as T_θ and $T_\theta(\omega) \equiv T_\theta(\underline{x}(\omega))$, $\omega \in \Omega$.

For a given realization of the random variables that characterize the model, that is, for given $\omega \in \Omega$, we need to map the set of equilibrium decision rules of each player into outcomes of the game. Consider the realization $t(\omega)$ of $t \in \mathrm{Sel}(T_\theta)$. Through the threshold decision rule, such a realization implies the action profile

$$q(t(\omega)) = \begin{bmatrix} 1(\varepsilon_1(\omega) \le t_1(\omega), \varepsilon_2(\omega) \le t_2(\omega)) \\ 1(\varepsilon_1(\omega) \ge t_1(\omega), \varepsilon_2(\omega) \le t_2(\omega)) \\ 1(\varepsilon_1(\omega) \le t_1(\omega), \varepsilon_2(\omega) \ge t_2(\omega)) \\ 1(\varepsilon_1(\omega) \ge t_1(\omega), \varepsilon_2(\omega) \ge t_2(\omega)) \end{bmatrix} \in \Delta^3, \tag{C.1}$$

with Δ^3 the simplex in $\Re^4$. The vector $q(t(\omega))$ indicates which of the four possible pairs of actions is played with probability 1, when the realization of $(\underline{x}, \varepsilon)$ is $(\underline{x}(\omega), \varepsilon(\omega))$ and the equilibrium threshold is $t(\omega) \in T_\theta(\underline{x}(\omega))$. Applying this construction to all measurable selections of T_θ, we construct a random closed set in Δ^3:

$$Q_\theta = \{q(t) : t \in \mathrm{Sel}(T_\theta)\}.$$

For given $\underline{x}$ and $\theta \in \Theta$, define the conditional Aumann expectation

$$\mathbb{E}(Q_\theta|\underline{x}) = \{\mathbf{E}(q(t)|\underline{x}) : t \in \mathrm{Sel}(T_\theta)\}.$$

Notice that for a specific selection $t \in \mathrm{Sel}(T_\theta)$, given the independence assumption on $\varepsilon_1, \varepsilon_2$, the first entry of the vector $\mathbf{E}(q(t)|\underline{x})$ is

$$\mathbf{E}(1(\varepsilon_1 \le t_1, \varepsilon_2 \le t_2)|\underline{x}) = (1 - F_\theta(t_1|\underline{x}))(1 - F_\theta(t_2|\underline{x})),$$

[5]For example, with payoffs linear in $\underline{x}$ and given by $\pi(y_j, y_{-j}, \underline{x}, \varepsilon_j; \theta) = y_j(y_{-j}\theta_{1j} + x_j\theta_{2j} + \varepsilon_j)$, we have that player 1 enters if and only if $(\varepsilon_1 + x_1\theta_{21})F_\theta(t_2|\underline{x}) + (\varepsilon_1 + x_1\theta_{21} + \theta_{11})(1 - F_\theta(t_2|\underline{x})) \ge 0$. Therefore, the cutoff is $t_j^b(t_{-j}, \underline{x}; \theta) = -x_1\theta_{21}F_\theta(t_2|\underline{x}) - (x_1\theta_{21} + \theta_{11})(1 - F_\theta(t_2|\underline{x})) = -x_1\theta_{21} - \theta_{11}(1 - F_\theta(t_2|\underline{x}))$.

and similarly for other entries of $\mathbf{E}(q(t)|\underline{x})$. This yields the multinomial distribution over outcome profiles determined by equilibrium threshold $t \in \mathrm{Sel}(T_\theta)$. By the same logic as in Section 3, $\mathbb{E}(Q_\theta|\underline{x})$ is the set of probability distributions over action profiles conditional on $\underline{x}$ which are consistent with the maintained modeling assumptions, that is, with *all* the model's implications. By the same results that we applied in the main papers, the set $\mathbb{E}(Q_\theta|\underline{x})$ is closed and convex.

Observe that regardless of whether T_θ contains a finite number of equilibria or a continuum, Q_θ can take on only a finite number of realizations that correspond to each of the vertices of Δ^3, because the vectors $q(t)$ in equation (C.1) collect threshold decision rules.[6] As we show in the proof of Theorem C.1, this implies that $\mathbb{E}(Q_\theta|\underline{x})$ is a closed convex polytope $\underline{x}$-a.s., fully characterized by a finite number of supporting hyperplanes. In turn, this allows us to characterize Θ_I through a finite number of moment inequalities and to compute it using efficient algorithms in linear programming.

THEOREM C.1: *Let Assumptions* C.1 *and* 3.2 *hold*. *Then*

$$\begin{aligned}\Theta_I &= \Big\{\theta \in \Theta : \max_{u \in B}\big(u'\mathbf{P}(y|\underline{x}) - \mathbf{E}[h(Q_\theta, u)|\underline{x}]\big) = 0, \underline{x}\text{-}a.s.\Big\}\\ &= \big\{\theta \in \Theta : u'\mathbf{P}(y|\underline{x}) \le \mathbf{E}[h(Q_\theta, u)|\underline{x}]\ \forall u \in D, \underline{x}\text{-}a.s.\big\},\end{aligned}$$

where $D = \{u = [u_1 \ \cdots \ u_{\kappa_{\mathcal{Y}}}]' : u_i \in \{0, 1\}, i = 1, \dots, \kappa_{\mathcal{Y}}\}$.

PROOF: By the same argument as in the proof of Theorem 2.1,

$$\begin{aligned}\Theta_I &= \{\theta \in \Theta : \mathbf{P}(y|\underline{x}) \in \mathbb{E}(Q_\theta|\underline{x}), \underline{x}\text{-a.s.}\}\\ &= \Big\{\theta \in \Theta : \max_{u \in B}\big(u'\mathbf{P}(y|\underline{x}) - \mathbf{E}[h(Q_\theta, u)|\underline{x}]\big) = 0, \underline{x}\text{-a.s.}\Big\}\\ &= \big\{\theta \in \Theta : u'\mathbf{P}(y|\underline{x}) \le \mathbf{E}[h(Q_\theta, u)|\underline{x}]\ \forall u \in B, \underline{x}\text{-a.s.}\big\}.\end{aligned}$$

It remains to show equivalence of the conditions

$$\text{(i)}\quad u'\mathbf{P}(y|\underline{x}) \le \mathbf{E}[h(Q_\theta, u)|\underline{x}] \quad \forall u \in B,$$

$$\text{(ii)}\quad u'\mathbf{P}(y|\underline{x}) \le \mathbf{E}[h(Q_\theta, u)|\underline{x}] \quad \forall u \in D.$$

By the positive homogeneity of the support function, condition (i) is equivalent to $u'\mathbf{P}(y|\underline{x}) \le \mathbf{E}[h(Q_\theta, u)|\underline{x}]\ \forall u \in \Re^{\kappa_{\mathcal{Y}}}$. It is obvious that this condition implies condition (ii). To see why condition (ii) implies condition (i), observe that because the set Q_θ and the set $\mathrm{co}[Q_\theta]$ are simple, one can find a finite measurable

[6]Hence, the set Q_θ is a "simple" random closed set in Δ^3, in the sense that there exists a finite measurable partition $\Omega_1, \dots, \Omega_m$ of Ω and sets $K_1, \dots, K_m \in \mathcal{F}$ such that $Q_\theta(\omega) = K_i$ for all $\omega \in \Omega_i$, $1 \le i \le m$.

partition $\Omega_1, \dots, \Omega_m$ of Ω and convex sets $K_1, \dots, K_m \in \Delta^{\kappa_\mathcal{Y}-1}$, such that by Theorem 2.1.21 in Molchanov (2005),

$$\mathbb{E}(Q_\theta|\underline{x}) = K_1\mathbf{P}(\Omega_1|\underline{x}) \oplus K_2\mathbf{P}(\Omega_2|\underline{x}) \oplus \cdots \oplus K_m\mathbf{P}(\Omega_m|\underline{x}),$$

with K_i the value that $\operatorname{co}[Q_\theta(\omega)]$ takes for $\omega \in \Omega_i$, $i = 1, \dots, m$ (see Molchanov (2005, Definition 1.2.8)). By the properties of the support function (see Schneider (1993, Theorem 1.7.5)),

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

Finally, for each $i = 1, \dots, m$, the vertices of K_i are a subset of the vertices of $\Delta^{\kappa_\mathcal{Y}-1}$. Hence the supporting hyperplanes of K_i, $i = 1, \dots, m$, 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 $u \in D$. Therefore, the supporting hyperplanes of $\mathbb{E}(Q_\theta|\underline{x})$ are a subset of the supporting hyperplanes of $\Delta^{\kappa_\mathcal{Y}-1}$. *Q.E.D.*

REMARK 1: Grieco (2009) introduced an important model, where each player has a vector of payoff shifters that are unobservable by the researcher. Some of the elements of this vector are private information to the player, while the others are known to all players. Our results in Section 2 apply to this setup as well, by the same arguments as in Section 3 and in this appendix.

REMARK 2: Appendix B verifies the regularity conditions required by AS for models that satisfy Assumptions C.1 and 3.2 under the additional assumption that the researcher observes an i.i.d. sequence of equilibrium outcomes and observable payoff shifters $\{y_i, \underline{x}_i\}_{i=1}^n$.

APPENDIX D: PURE STRATEGIES ONLY: FURTHER SIMPLIFICATIONS

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 the possible nonexistence 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 D.1 below:

ASSUMPTION D.1: *One of the following statements holds*:

(i) *For a subset of values of $\theta \in \Theta$ which includes 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, $S_\theta(\underline{x}, \varepsilon) = \operatorname{vert}(\Sigma(\mathcal{Y}))$, with $\operatorname{vert}(\cdot)$ the vertices of the set in parentheses.*

Assumption D.1(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 $\underline{x}$, ε. On the other hand, Assumption D.1(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 D.1(ii) is more conservative than Assumption D.1(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 D.1(i) or D.1(ii) is imposed. Moreover, one may choose not to impose Assumption D.1 at all and to use a different solution concept. In that case as well, as we illustrate in Appendix E, our approach can be applied to characterize the sharp identification region.

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

EXAMPLE 1: Consider a simple two player entry game similar to the one in Tamer (2003), omit the covariates, and assume that players' payoffs are given by $\pi_j = y_j(y_{-j}\theta_j + \varepsilon_j)$, where $y_j \in \{0, 1\}$ 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 S.1 plots the set S_θ resulting from the possible realizations of ε_1, ε_2. In this case, S_θ assumes only five values:

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

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. Consequently, also the set Q_θ assumes only five values, equal, respectively, to $\{[1\ 0\ 0\ 0]'\}$, $\{[0\ 1\ 0\ 0]'\}$, $\{[0\ 0\ 1\ 0]'\}$, $\{[0\ 0\ 0\ 1]'\}$, and $\{[0\ 1\ 0\ 0]', [0\ 0\ 1\ 0]'\}$.

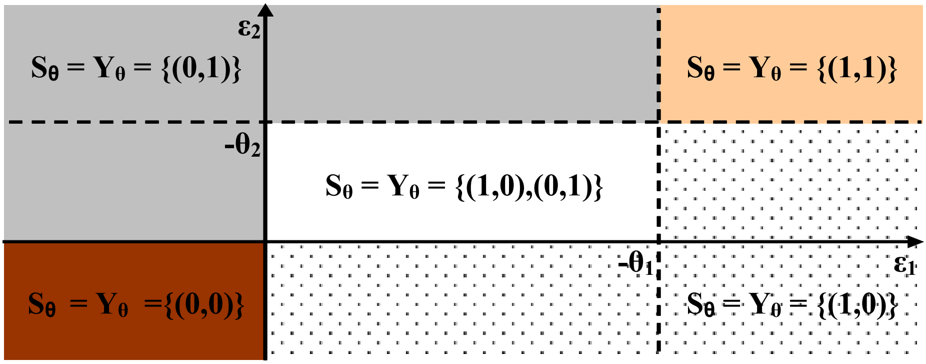

FIGURE S.1.—The random set of pure strategy Nash equilibrium profiles S_θ and the random set of pure strategy Nash equilibrium outcomes Y_θ as a function of $\varepsilon_1, \varepsilon_2$ in a two player entry game. In this simple example, the two sets coincide.

Hence, the sets S_θ and Q_θ are "simple" random closed sets in $\Sigma(\mathcal{Y})$ and $\Delta^{\kappa_{\mathcal{Y}}-1}$, respectively. Because the probability space is nonatomic and Q_θ is simple, $\mathbb{E}(Q_\theta|\underline{x})$ is a closed convex polytope, fully characterized by a finite number of supporting hyperplanes.

EXAMPLE 1—Continued: Consider again the simple two player entry game with pure strategies only in Example 1. Then for $\varepsilon \in \mathcal{E}_\theta^M$, the set Q_θ contains only two points, $[0\ \ 1\ \ 0\ \ 0]'$ and $[0\ \ 0\ \ 1\ \ 0]'$, and for $\varepsilon \notin \mathcal{E}_\theta^M$ it is a singleton. Therefore, the expectations of the selections of Q_θ are given by

$$\begin{aligned}\mathbf{E}(q) = {}& \Big[\mathbf{P}\big(\varepsilon \in \mathcal{E}_\theta^{(0,0)}\big)\ \ \mathbf{P}\big(\varepsilon \in \mathcal{E}_\theta^{(1,0)}\big)\ \ \mathbf{P}\big(\varepsilon \in \mathcal{E}_\theta^{(0,1)}\big)\ \ \mathbf{P}\big(\varepsilon \in \mathcal{E}_\theta^{(1,1)}\big)\Big]' \\ & + [0\ \ p_1\ \ 1-p_1\ \ 0]'\mathbf{P}(\varepsilon \in \mathcal{E}_\theta^M),\end{aligned}$$

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

Hence, checking whether $\mathbf{P}(y|\underline{x}) \in \mathbb{E}(Q_\theta|\underline{x})$ amounts to checking whether a point belongs to a polytope, that is, whether a finite number of moment inequalities hold $\underline{x}$-a.s. In Theorem D.1, we show that these inequalities are obtained by checking inequality $u'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}[h(Q_\theta, u)|\underline{x}]$ for the $2^{\kappa_{\mathcal{Y}}}$ possible u vectors whose entries are either equal to 0 or to 1.

THEOREM D.1: *Assume that players use only pure strategies, that Assumptions* 3.1 *and* 3.2 *in BMM and Assumption* D.1 *are satisfied. Then for $\underline{x}$-a.s. these two conditions are equivalent*:

(i) $u'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}[h(Q_\theta, u)|\underline{x}]\ \forall u \in \Re^{\kappa_{\mathcal{Y}}}$.

(ii) $u'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}[h(Q_\theta, u)|\underline{x}]$ $\forall u \in D = \{u = [u_1 \ \cdots \ u_{\kappa_{\mathcal{Y}}}]' : u_i \in \{0, 1\},\ i = 1, \ldots, \kappa_{\mathcal{Y}}\}$.

The proof follows using the same argument as in the proof of Theorem C.1.

In Appendix D.2, 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 Appendix D.2, 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 D.1 are redundant.

To conclude this appendix, 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.[7] When the distribution of ε is continuous, Q_θ may take a continuum of values as a function of ε, and $\mathbb{E}(Q_\theta|\underline{x})$ may have infinitely many extreme points. Therefore, one needs an infinite number of moment inequalities to determine whether $\mathbf{P}(y|\underline{x})$ belongs to it. In this case, the most practical approach to obtain the sharp identification region is by solving the maximization problem in Theorem 3.2.

D.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 Ciliberto and Tamer (2009, CT henceforth). We adopt the version of this model described in Berry and Tamer (2007, pp. 84 and 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

[7]For example, in the two player entry game in Example 1, for $\varepsilon \in \mathcal{E}_M^\theta$, $S_\theta = \{(0, 1), (\frac{\varepsilon_2}{-\theta_2}, \frac{\varepsilon_1}{-\theta_1}), (1, 0)\}$. 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.

and type 2 have, respectively, the payoff functions

$$\pi_{j1}(a_{j1}, a_{-j1}, a_{12}, a_{22}, \varepsilon_1) = y_{j1}(\theta_{11}(a_{-j1} + a_{12} + a_{22}) - \varepsilon_1), \tag{D.1}$$

$$\pi_{j2}(a_{j2}, a_{-j2}, a_{11}, a_{21}, \varepsilon_2) = a_{j2}(\theta_{21}(a_{11} + a_{21}) + \theta_{22}a_{-j2} - \varepsilon_2). \tag{D.2}$$

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 let $y_2 = a_{12} + a_{22}$ denote 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 nine 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 players' actions and observable outcomes of the game differ. Figure S.2 plots the outcomes of the game against the realizations of

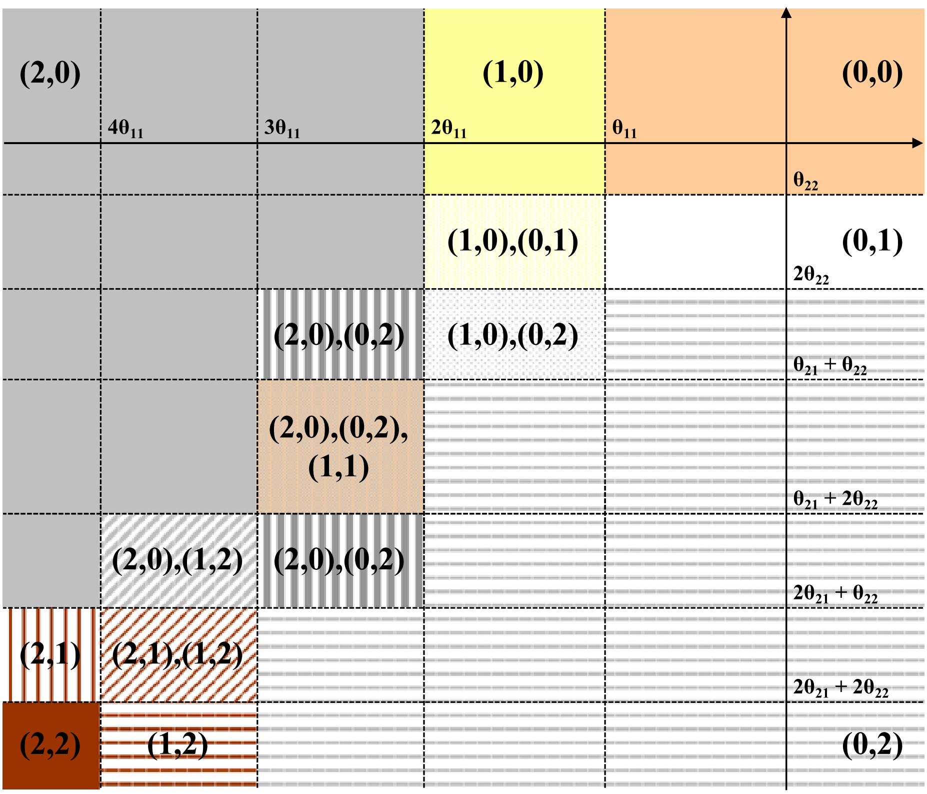

FIGURE S.2.—The random set of pure strategy Nash equilibrium outcomes as a function of $\varepsilon_1, \varepsilon_2$ in a four player, two type entry game.

$\varepsilon_1, \varepsilon_2$. In this case, Q_θ 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 $Q_\theta(\omega) = \{[1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]'\}$; 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 $Q_\theta(\omega) = \{[0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0]', [0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0]'\}$; and so forth.

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. Because we are allowing only pure strategy equilibria, the realizations of any $\sigma \in S_\theta$ are vectors of 0's and 1's. Hence, for all $\omega \in \Omega$, $[q(\sigma(\omega))]_k = 1$ if $\prod_{j=1}^{J} \sigma_j(\omega, t_j^k) = 1$ and equals 0 otherwise. Consider two equilibria $t^k, t^l \in \mathcal{Y}$, $1 \leq k \neq l \leq \kappa_{\mathcal{Y}}$, such that

$$\left\{\omega : \prod_{j=1}^{J} \sigma_j(\omega, t_j^k) = 1 \Big| \underline{x}\right\} \cap \left\{\omega : \prod_{j=1}^{J} \sigma_j(\omega, t_j^l) = 1 \Big| \underline{x}\right\} = \emptyset, \tag{D.3}$$

that is, the set of ω for which S_θ admits both t^k and t^l as equilibria has probability 0. Let u^k be a vector with each entry equal to 0 and entry k equal to 1, and similarly for u^l. Then the inequality $(u^k + u^l)'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}[h(Q(S_\theta), u^k + u^l)|\underline{x}]$ does not add any information beyond that provided by the inequalities $u'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}[h(Q(S_\theta), u)|\underline{x}]$ 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 $\kappa_{\mathcal{Y}}$. Applying this simple reasoning, the sharp identification region that we give in this example 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 S.3 and Table S.I report Θ_I, Θ_O^{CT} (the outer region proposed by CT), and Θ_O^{ABJ} (the outer region proposed by Andrews, Berry, and Jia (2004, ABJ henceforth)), in a simple example with $(\varepsilon_1, \varepsilon_2) \overset{\text{i.i.d.}}{\sim} N(0, 1)$ 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} is given 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 also use the equalities on $\mathbf{P}(y = (0, 0))$ and $\mathbf{P}(y = (2, 2))$ 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 $\mathbf{P}(y) =$

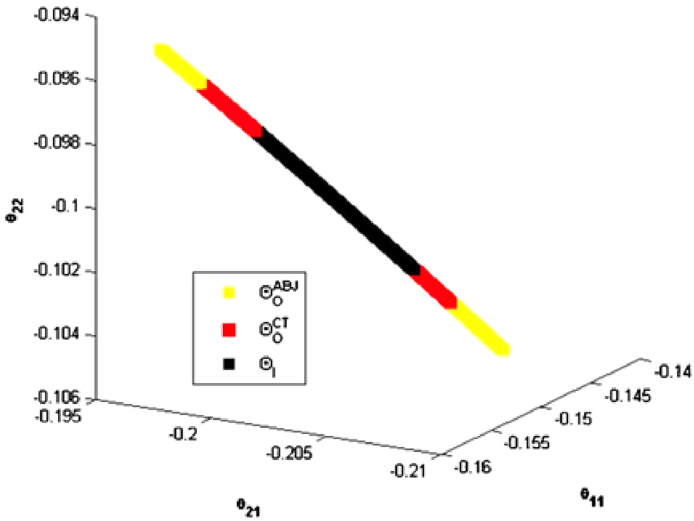

FIGURE S.3.—Identification regions in a four player, two type entry game with pure strategy Nash equilibrium as the solution concept.

[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.

To further illustrate the computational advantages of our characterization of Θ_I in Theorem 3.2, we also recalculated the sharp identification region for this example solving for each candidate $\theta \in \Theta$ the problem $\max_{u\in B}(u'\mathbf{P}(y|\underline{x}) - \mathbf{E}[h(Q_\theta, u)|\underline{x}])$, *without* taking advantage of our knowledge of the structure

TABLE S.I

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 THE SOLUTION CONCEPT

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

of the game that reduces the number of inequalities to be checked to 26. We modified the simple Nelder–Mead algorithm described in Section 3.4 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 1 minute.

D.2. *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[8] is

$$Y_\theta(\underline{x}, \varepsilon) = \{y \in \mathcal{Y} : \pi_j(y_j, y_{-j}, x_j, \varepsilon_j, \theta) \geq \pi_j(\tilde{y}_j, y_{-j}, x_j, \varepsilon_j, \theta) \\ \forall \tilde{y}_j \in \mathcal{Y}_j \ \forall j\}. \tag{D.4}$$

As we did for S_θ, we omit the explicit reference to this set's dependence on $\underline{x}$ and ε. Given Assumption 3.1, one can easily show that Y_θ is a random closed set in $\mathcal{Y}$ (see Definition A.1). Because the realizations of Y_θ are subsets of the finite set $\mathcal{Y}$, it immediately follows that Y_θ is a random closed set in $\mathcal{Y}$ without any requirement on the payoff functions.

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 3.1, 3.2, and D.1, 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_θ, $\underline{x}$-a.s.; see Definition A.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 Section 1.4.8), and amounts[9] to

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

[8]Restrict the set S_θ to be a set of pure strategy Nash equilibria. Then when players' actions and outcomes of the game coincide, Y_θ coincides with S_θ. However, under the more general assumption that $y = g(a)$, where $a \in \mathcal{A}$ is a strategy profile and g is an outcome rule, these two sets differ and

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

[9]Beresteanu and Molinari (2006, 2008, Proposition 4.1) used 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) used it to define a correctly specified partially identified structural model, and derived a Kolmogorov–Smirnov test for Choquet capacities.

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

$$\mathbf{P}\{y \in K|\underline{x}\} \leq \mathbf{P}\{Y_\theta \cap K \neq \emptyset|\underline{x}\} \quad \forall K \subset \mathcal{Y}, \underline{x}\text{-a.s.} \tag{D.5}$$

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 (D.5) is called the capacity functional of Y_θ given $\underline{x}$. The following definitions formally introduce the unconditional version of this functional and a few related ones:

DEFINITION D.1: Let Z be a random closed set in $\Re^d$ and denote by $\mathcal{K}$ the family of compact subsets of $\Re^d$. 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

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

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 set K, the following relationship holds:

$$\mathbf{C}_Z(K) = 1 - \mathbf{T}_Z(K^c). \tag{D.6}$$

EXAMPLE 2: Consider again the simple two player entry game in Example 1. Figure S.1 plots the set Y_θ against the realizations of $\varepsilon_1, \varepsilon_2$. In this case, $\mathbf{T}_{Y_\theta}(\{(0,0)\}) = \mathbf{P}(\varepsilon_1 \leq 0, \varepsilon_2 \leq 0)$, $\mathbf{T}_{Y_\theta}(\{(1,0)\}) = \mathbf{P}(\varepsilon_1 \geq 0, \varepsilon_2 \leq -\theta_2)$, $\mathbf{T}_{Y_\theta}(\{(0,1)\}) = \mathbf{P}(\varepsilon_1 \leq -\theta_1, \varepsilon_2 \geq 0)$, $\mathbf{T}_{Y_\theta}(\{(1,1)\}) = \mathbf{P}(\varepsilon_1 \geq -\theta_1, \varepsilon_2 \geq -\theta_2)$, and $\mathbf{T}_{Y_\theta}(\{(1,0),(0,1)\}) = \mathbf{T}_{Y_\theta}(\{(1,0)\}) + \mathbf{T}_{Y_\theta}(\{(0,1)\}) - \mathbf{P}(0 \leq \varepsilon_1 \leq -\theta_1, 0 \leq \varepsilon_2 \leq -\theta_2)$. The capacity functional of the remaining subsets of $\mathcal{Y}$ can be calculated similarly.

Notice that given equation (D.6), inequalities (D.5) can be equivalently written as

$$\mathbf{C}_{Y_\theta|\underline{x}}(K) \leq \mathbf{P}\{y \in K|\underline{x}\} \leq \mathbf{T}_{Y_\theta|\underline{x}}(K) \quad \forall K \subset \mathcal{Y}, \underline{x}\text{-a.s.}, \tag{D.7}$$

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 (D.5) when discussing the relationship between our analysis and that of CT. Clearly, if one considers all $K \subset \mathcal{Y}$, the left-hand side inequality in (D.7) is superfluous: when the inequalities in (D.7) are used, only subsets $K \subset \mathcal{Y}$ of cardinality up to half of the cardinality of $\mathcal{Y}$ are needed.

We can redefine the identified set of parameters θ as

$$\Theta_I = \{\theta \in \Theta : \mathbf{P}\{y \in K|\underline{x}\} \leq \mathbf{T}_{Y_\theta|\underline{x}}(K)\ \forall K \subset \mathcal{Y}, \underline{x}\text{-a.s.}\}. \tag{D.8}$$

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:

$$\Theta_O^{\mathrm{ABJ}} = \{\theta \in \Theta : \mathbf{P}\{y = t|\underline{x}\} \leq \mathbf{T}_{Y_\theta|\underline{x}}(t)\ \forall t \in \mathcal{Y}, \underline{x}\text{-a.s.}\}, \tag{D.9}$$

$$\Theta_O^{\mathrm{CT}} = \{\theta \in \Theta : \mathbf{C}_{Y_\theta|\underline{x}}(t) \leq \mathbf{P}\{y = t|\underline{x}\} \leq \mathbf{T}_{Y_\theta|\underline{x}}(t)\ \forall t \in \mathcal{Y}, \underline{x}\text{-a.s.}\}. \tag{D.10}$$

Both ABJ and CT acknowledged that the parameter regions they gave are not sharp. Comparing the sets in equations (D.9) and (D.10) with the set in equation (D.8), one observes that Θ_O^{ABJ} is obtained by applying inequality (D.5) only for $K = \{t\}$ for all $t \in \mathcal{Y}$. Similarly, Θ_O^{CT} is obtained by applying inequality (D.7) only for $K = \{t\}$ (or, equivalently, applying inequality (D.5) for $K = \{t\}$ and $K = \mathcal{Y} \setminus \{t\}$ for all $t \in \mathcal{Y}$). 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: (i) whether Θ_I as defined in equation (D.8) yields the sharp identification region of θ and (ii) if and by how much Θ_I differs from Θ_O^{ABJ} and Θ_O^{CT}. We answer here the first question. Appendix D.1 answers the second question by looking at a simple example.

THEOREM D.2: *Assume that players use only pure strategies, and that Assumptions* 3.1, 3.2, *and* D.1 *are satisfied. Then for* $\underline{x}$*-a.s., the following two conditions are equivalent*:

(i) $u'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}[h(Q_\theta, u)|\underline{x}]\ \forall u \in \Re^{\kappa_{\mathcal{Y}}}$.
(ii) $\mathbf{P}\{y \in K|\underline{x}\} \leq \mathbf{T}_{Y_\theta|\underline{x}}(K)\ \forall K \subset \mathcal{Y}$.

For the proof, see Beresteanu, Molchanov, and Molinari (2008, Theorem 4.1).

D.3. *On the Number of Inequalities to Be Checked in the Pure Strategies Case*

As discussed in Appendix D.1, 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 D.2, one can show that the result in equation (D.3) 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

$$\{\omega : Y_\theta(\omega) \cap K_1 \neq \emptyset|\underline{x}\} \cap \{\omega : Y_\theta(\omega) \cap K_2 \neq \emptyset|\underline{x}\} = \emptyset, \tag{D.11}$$

that is, the set of ω for which Y_θ intersects both K_1 and K_2 has probability 0, 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 four players, if the number of entrants is identified, the number of inequalities to be verified reduces from 65,536 to at most 100. Theorem D.3 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 D.3: *Take $\theta \in \Theta$, and let Assumptions* 3.1, 3.2, *and* D.1 *hold. Consider a partition of Ω into sets $\Omega^1, \dots, \Omega^M$ of positive probability. Let*

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

denote the range of $Y_\theta(\omega)$ for $\omega \in \Omega^i$. If $\mathcal{Y}_1, \dots, \mathcal{Y}_M$ are disjoint, then it suffices to check (D.5) *only for all subsets K such that there is $i = 1, \dots, M$ for which $K \subseteq \mathcal{Y}_i$.*

For the proof, see Beresteanu, Molchanov, and Molinari (2008, Theorem 5.1).

A simple corollary to Theorem D.3, the proof of which is omitted, follows:

COROLLARY D.4: *Take $\theta \in \Theta$, and let Assumptions* 3.1, 3.2, *and* D.1 *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 $\mathcal{Y}_i = \bigcup_{\omega \in \Omega^i} Y_\theta(\omega)$, $i = 1, 2$, and assume that $\mathcal{Y}_1 \cap \mathcal{Y}_2 = \emptyset$ and that $\kappa_{\mathcal{Y}_2} \leq 2$. Then inequalities* (D.5) *hold if*

$$\mathbf{P}\{Y_\theta = \{t\}|\underline{x}\} \leq \mathbf{P}\{y = t|\underline{x}\} \leq \mathbf{P}\{t \in Y_\theta|\underline{x}\}, \tag{D.12}$$

$\underline{x}$-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.[10] 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 (D.3) shows that actually the sharp identification region can be obtained by checking only five inequalities which have to hold for $\underline{x}$-a.s. and are given by inequalities (D.5) for $K = \{(0, 0)\}$, $\{(1, 0)\}$, $\{(0, 1)\}$, $\{(1, 1)\}$, $\{(1, 0), (0, 1)\}$. On the other hand, the example in Section 3.4 shows that CT's approach does not yield the sharp identification region when mixed strategies are allowed for.

[10] 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 discussed (see p. 32).

When no prior knowledge of the game such as, for example, that required in Theorem D.3 is available, it is still possible to use the insight in equation (D.11) 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."

THEOREM D.5: *Let Assumptions* 3.1, 3.2, *and* D.1 *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* $\mathbf{P}(\tilde{\mathcal{X}}) > 0$, *there exist* $t^1, t^2 \in \mathcal{Y}$ *with*

$$\mathbf{I}_{Y_\theta|\underline{x}}(t^1, t^2) > 0. \tag{D.13}$$

(a) *If* $\mathbf{P}\{\{t^1, t^2\} \cap Y_\theta \neq \emptyset|\underline{x}\} < 1$ *for all* $t^1, t^2 \in \mathcal{Y}$, *then there exists a random vector* z *which satisfies inequalities* (D.5) *for* $K = \{t\}$ *for all* $t \in \mathcal{Y}$ *but is not a selection of* Y_θ.

(b) *If*

$$\mathbf{P}\big\{\kappa_{Y_\theta} > 1|\underline{x}\big\} > \mathbf{I}_{Y_\theta|\underline{x}}(t^1) + \mathbf{I}_{Y_\theta|\underline{x}}(t^2) - \mathbf{C}_{Y_\theta|\underline{x}}(t^1) - \mathbf{C}_{Y_\theta|\underline{x}}(t^2), \tag{D.14}$$

then there exists a random vector z *which satisfies inequalities* (D.5) *for* $K = \{t\}$ *and* $K = \mathcal{Y} \setminus \{t\}$ *for all* $t \in \mathcal{Y}$ *but is not a selection of* Y_θ.

See Beresteanu, Molchanov, and Molinari (2008, Theorems 5.2 and 5.3) for a proof.

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 D.5(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 1, and it has at least three different equilibria. The assumptions of Theorem D.5(b) are satisfied whenever (a) there are regions of the unobservables of positive probability where two different outcomes can result from equilibrium strategy profiles and (b) the probability that the cardinality of Y_θ is greater than 1 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 D.1.

APPENDIX E: EXTENSIONS TO OTHER SOLUTION CONCEPTS

While in Section 3 and Appendix D, we focus on economic models of games in which Nash equilibrium is the solution concept employed, our approach easily applies 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.

E.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. 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

$$R_\theta(\underset{\sim}{x}, \varepsilon) = \big\{\sigma \in \Sigma(\mathcal{Y}) : \forall j\ \exists \tilde{\sigma}_{-j} \in \Sigma(\mathcal{Y}_{-j}) \text{ s.t.}$$
$$\pi_j(\sigma_j, \tilde{\sigma}_{-j}, x_j, \varepsilon_j, \theta) \geq \pi_j(\sigma'_j, \tilde{\sigma}_{-j}, x_j, \varepsilon_j, \theta)\ \forall \sigma'_j \in \Sigma(\mathcal{Y}_j)\big\}.$$

Omitting the explicit reference to its dependence on $\underset{\sim}{x}$ and ε, R_θ is the set of level-1 rational strategy profiles of the game. By arguments similar to those we used above, this is a random closed set in $\Sigma(\mathcal{Y})$. Figure S.4 plots this set against the possible realizations of $\varepsilon_1, \varepsilon_2$, in a simple two player simultaneous-move, complete information, static game of entry in which players' payoffs are given by $\pi_j = y_j(y_{-j}\theta_j + \varepsilon_j)$, $y_j \in \{0, 1\}$, and θ_1 and θ_2 are assumed to be negative.

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

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

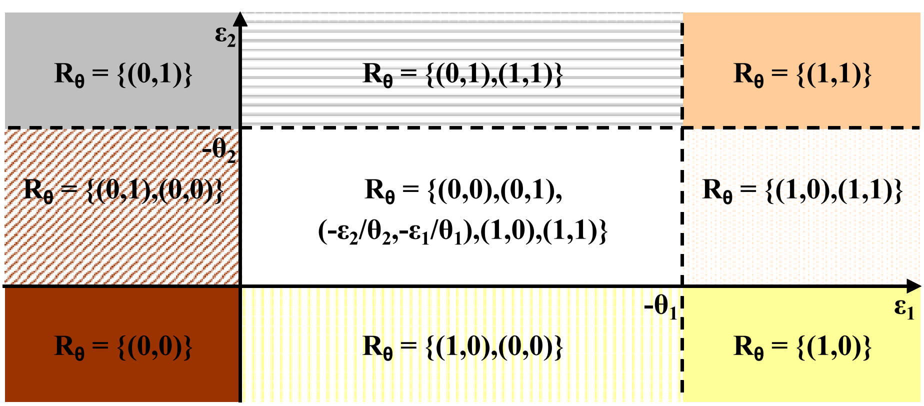

FIGURE S.4.—The random set of level-1 rational profiles as a function of $\varepsilon_1, \varepsilon_2$ in a two player entry game.

with

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

where $[q(\sigma)]_k, k = 1, \dots, \kappa_{\mathcal{Y}}$, is defined in Section 3.

In our simple example in Figure S.4, with omitted covariates, for any $\omega \in \Omega$ such that $\varepsilon(\omega) \in [0, -\theta_1] \times [0, -\theta_2]$,

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

and, therefore, it follows that $\mathbb{E}(Q(R_\theta))$ is equal to $\mathbb{E}(Q(\tilde{R}_\theta))$, with $\tilde{R}_\theta$ restricted to be the set of level-1 rational pure strategies. Hence, by Theorem D.1 below, 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) suggested to obtain the sharp identification region as the set of parameter values that return value 0 for the objective function of a linear programming problem. For the general case in which ε may have a continuous distribution, AT applied the same insight of CT and characterized an outer identification region through eight moment inequalities similar to those in equation (D.10). 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 (D.9). 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 and θ_2. Figure S.5 plots the upper contours of Θ_I, Θ_O^{CT}, and Θ_O^{ABJ} in a simple example with $(\varepsilon_1, \varepsilon_2) \overset{\text{i.i.d.}}{\sim} N(0,1)$ and $\Theta = [-5, 0]^2$. The data are generated with $\theta_1^\star = -1.15$ and $\theta_2^\star = -1.4$, 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 outcomes $(1,0)$ 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 and 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 $(-0.54, -0.61)$.

E.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

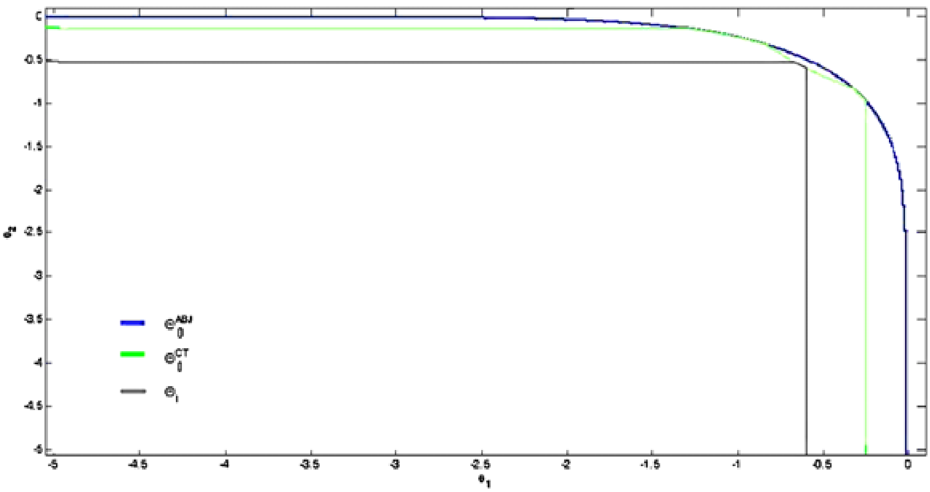

FIGURE S.5.—Upper contours of the identification regions in a two player entry game with level-1 rationality as the solution concept.

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 $\gamma \in \Delta^{\kappa_{\mathcal{Y}}-1}$, where $\Delta^{\kappa_{\mathcal{Y}}-1}$ denotes the set of probability distributions on $\mathcal{Y}$, is an equilibrium if, conditional on knowing that her own instruction is to play y_j, each player j has no incentive to deviate to any other strategy y_j', assuming that the other players follow their own instructions. In this case, one can define the θ-dependent set

$$\begin{aligned} C_\theta(\underline{x}, \varepsilon) = \Bigg\{ \gamma \in \Delta^{\kappa_{\mathcal{Y}}-1} : & \sum_{y_{-j} \in \mathcal{Y}_{-j}} \gamma(y_j, y_{-j}) \pi_j(y_j, y_{-j}, x_j, \varepsilon_j, \theta) \\ & \geq \sum_{y_{-j} \in \mathcal{Y}_{-j}} \gamma(y_j, y_{-j}) \pi_j(y_j', y_{-j}, x_j, \varepsilon_j, \theta) \\ & \forall y_j \in \mathcal{Y}_j,\ \forall y_j' \in \mathcal{Y}_j,\ \forall j \Bigg\}. \end{aligned}$$

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 $\Delta^{\kappa_{\mathcal{Y}}-1}$. Notice that C_θ is defined by a finite number of linear inequalities on the set $\Delta^{\kappa_{\mathcal{Y}}-1}$ of correlated strategies and, therefore, it is a nonempty polytope. Yang (2008) was the first to use this fact, along with the fact that $\mathrm{co}[Q(S_\theta)] \subset C_\theta$, to develop a computationally easy-to-implement estimator for an outer identification region of θ when the solution concept employed is Nash equilibrium. Here we provide a simple characteriza-

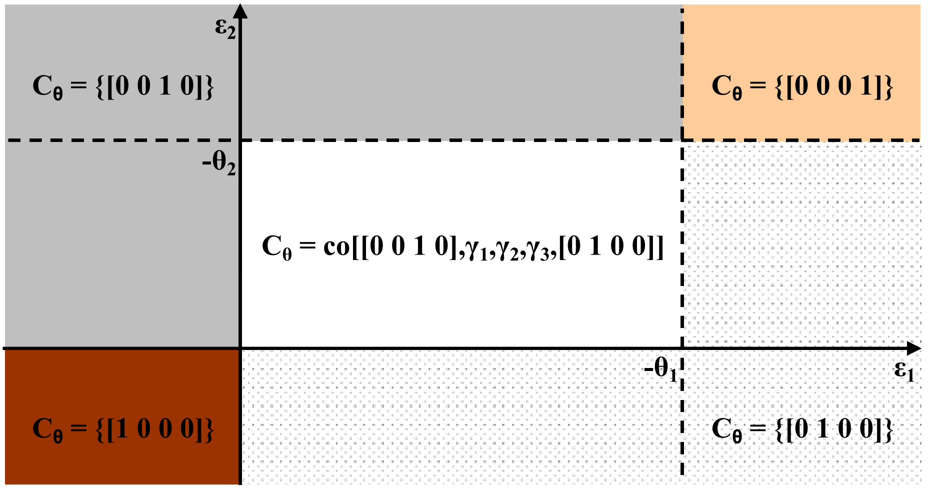

FIGURE S.6.—The random set of correlated equilibria as a function of $\varepsilon_1, \varepsilon_2$ in a two player entry game. The correlated equilibria γ_1, γ_2, and γ_3 are defined in Section E.2.

tion of the sharp identification region Θ_I when the solution concept employed is objective correlated equilibrium. In particular, the same approach of Section 3 allows us to obtain the sharp identification region for θ as

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

In our simple two player simultaneous-move, complete information, static game of entry, $\mathcal{Y}_j = \{0, 1\}$, $j = 1, 2$, $\mathcal{Y} = \{(0, 0), (1, 0), (0, 1), (1, 1)\}$. Again omitting the covariates, we assume that players' payoffs are given by $\pi_j = y_j(y_{-j}\theta_j + \varepsilon_j)$, where $y_j \in \{0, 1\}$ and θ_j is assumed to be negative (monopoly payoffs are higher than duopoly payoffs), $j = 1, 2$. Figure S.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 $Q_\theta(\omega)$ that results 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))—and is given by

$$\begin{aligned}
\gamma_0(\omega) &= [0 \ \ 0 \ \ 1 \ \ 0]', \\
\gamma_1(\omega) &= \left[1 \ \ -\frac{\varepsilon_2(\omega)}{\theta_2 + \varepsilon_2(\omega)} \ \ -\frac{\varepsilon_1(\omega)}{\theta_1 + \varepsilon_1(\omega)} \ \ 0\right]' \\
&\qquad \times \left(1 - \frac{\varepsilon_1(\omega)}{\theta_1 + \varepsilon_1(\omega)} - \frac{\varepsilon_2(\omega)}{\theta_2 + \varepsilon_2(\omega)}\right)^{-1},
\end{aligned}$$

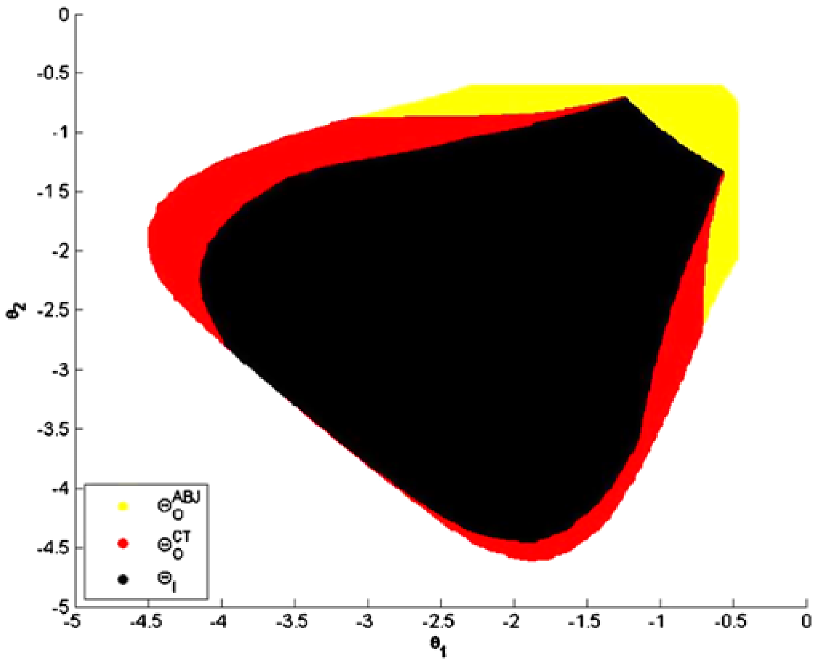

FIGURE S.7.—Identification regions in a two player entry game with correlated equilibrium as the solution concept.

$$\gamma_2(\omega) = \Bigg[\left(1+\frac{\varepsilon_2(\omega)}{\theta_2}\right)\left(1+\frac{\varepsilon_1(\omega)}{\theta_1}\right) \quad -\frac{\varepsilon_2(\omega)}{\theta_2}\left(1+\frac{\varepsilon_1(\omega)}{\theta_1}\right)$$
$$-\left(1+\frac{\varepsilon_2(\omega)}{\theta_2}\right)\frac{\varepsilon_1(\omega)}{\theta_1} \quad \frac{\varepsilon_2(\omega)}{\theta_2}\frac{\varepsilon_1(\omega)}{\theta_1}\Bigg]',$$

$$\gamma_3(\omega) = \left[0 \quad -\frac{\varepsilon_2(\omega)}{\theta_2+\varepsilon_2(\omega)} \quad -\frac{\varepsilon_1(\omega)}{\theta_1+\varepsilon_1(\omega)} \quad \frac{\varepsilon_1(\omega)}{\theta_1+\varepsilon_1(\omega)}\frac{\varepsilon_2(\omega)}{\theta_2+\varepsilon_2(\omega)}\right]'$$
$$\times\left(\frac{\varepsilon_1(\omega)}{\theta_1+\varepsilon_1(\omega)}\frac{\varepsilon_2(\omega)}{\theta_2+\varepsilon_2(\omega)} - \frac{\varepsilon_1(\omega)}{\theta_1+\varepsilon_1(\omega)} - \frac{\varepsilon_2(\omega)}{\theta_2+\varepsilon_2(\omega)}\right)^{-1},$$

$$\gamma_4(\omega) = [0 \;\; 1 \;\; 0 \;\; 0]'.$$

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.

Figure S.7 and Table S.II report Θ_I, Θ_O^{CT}, and Θ_O^{ABJ} in a simple example with $(\varepsilon_1, \varepsilon_2) \overset{\text{i.i.d.}}{\sim} N(0,1)$ 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} is given by the union of the red and black areas; Θ_I is the black region. The data are generated with $\theta_1^\star =$

TABLE S.II

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		
		Θ_O^{ABJ}	Θ_O^{CT}	Θ_I
$\theta_1^\star$	-1.15	$[-4.475, -0.485]$	$[-4.475, -0.585]$ (2.5%)	$[-4.125, -0.595]$ (11.5%)
$\theta_2^\star$	-1.40	$[-4.585, -0.625]$	$[-4.585, -0.725]$ (2.4%)	$[-4.425, -0.735]$ (6.8%)
Approximate reduction in total area compared to Θ_O^{ABJ}			(7.9%)	(23.1%)

-1.15 and $\theta_2^\star = -1.4$, and using a selection mechanism which picks each of outcomes $(0, 0)$ and $(1, 1)$ for 10% of $\omega : \varepsilon(\omega) \in [0, -\theta_1^\star] \times [0, -\theta_2^\star]$, and each of outcomes $(1, 0)$ 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.

APPENDIX F: MULTINOMIAL CHOICE MODELS WITH INTERVAL REGRESSORS DATA

This section of the supplement applies the methodology introduced in Section 2 to provide a tractable characterization of the sharp identification region of the parameters θ that characterize 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 $\mathcal{C} = \{0, \dots, \kappa_\mathcal{C} - 1\}$ 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 $(y, w, \{z_k, \varepsilon_k\}_{k=0}^{\kappa_\mathcal{C}-1})$ is defined on a nonatomic 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, z_k, and ε_k, $k = 0, \dots, \kappa_\mathcal{C} - 1$, are all scalars. 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 $F(\varepsilon)$ that is known. We let $\theta = [\{\alpha_k\}_{k=1}^{\kappa_\mathcal{C}-1} \;\; \delta \;\; \{\beta_k\}_{k=1}^{\kappa_\mathcal{C}-1}]' \in \Theta$ 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 $\varepsilon = [\{\varepsilon^k\}_{k=1}^{\kappa_\mathcal{C}-1}]$. 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 that arises when the econometrician observes only realizations of $\{y, z_{kL}, z_{kU}, w\}$, but not realizations of z_k, $k = 1, \dots, \kappa_\mathcal{C} - 1$. Following Manski and Tamer (2002), we assume that for each $k = 1, \dots, \kappa_\mathcal{C} - 1$, $\mathbf{P}(z_{kL} \le z_k \le z_{kU}) = 1$ 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 $\underline{x} = [1 \;\; \{z_{kL}\}_{k=1}^{\kappa_\mathcal{C}-1} \;\; \{z_{kU}\}_{k=1}^{\kappa_\mathcal{C}-1} \;\; w]$. Incompleteness of the data on z_k, $k = 1, \dots, \kappa_\mathcal{C} - 1$, 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) established that the sharp identification region for θ is given by

$$\begin{aligned}\Theta_I = \big\{&\theta \in \Theta : \mathbf{P}(x_{1L}\theta + \varepsilon^1 > 0 | \underline{x}) \le \mathbf{P}(y = 1 | \underline{x}) \\ &\le \mathbf{P}(x_{1U}\theta + \varepsilon^1 > 0 | \underline{x}), \underline{x}\text{-a.s.}\big\}.\end{aligned}$$

This construction is based on the observation that if the agent chooses alternative 1, this implies that $\varepsilon^1 > -x_1\theta \ge -x_{1U}\theta$. On the other hand, $\varepsilon^1 > -x_{1L}\theta \ge -x_1\theta$ implies that the agent chooses alternative 1.[11] In the case of more than two choices, one may wish to apply a similar insight as in the work of CT and construct the region

$$\begin{aligned}\text{(F.1)} \qquad \Theta_O = \big\{&\theta \in \Theta : \forall m \in \mathcal{C}, \underline{x}\text{-a.s.}, \\ &\mathbf{P}\big(x_m\theta_m + \varepsilon^m \ge x_k\theta_k + \varepsilon^k \; \forall (x_m, x_k) \in [x_{mL}, x_{mU}] \times [x_{kL}, x_{kU}], \\ &\forall k \in \mathcal{C},\; k \ne m | \underline{x}\big) \\ &\le \mathbf{P}(y = m | \underline{x}) \\ &\le \mathbf{P}\big(\exists x_m \in [x_{mL}, x_{mU}] \text{ s.t. } \forall k \in \mathcal{C},\; k \ne m, \exists x_k \in [x_{kL}, x_{kU}] \\ &\text{with } x_m\theta_m + \varepsilon^m \ge x_k\theta_k + \varepsilon^k | \underline{x}\big)\big\}.\end{aligned}$$

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

The lower bound on $\mathbf{P}(y = m|\underline{x})$ in equation (F.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 D.2 provides further insights to explain the lack of sharpness of Θ_O.[12]

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

$$\begin{aligned}\pi(m; x_m, \varepsilon_m, \theta_m) &= x_m\theta_m + \varepsilon_m \geq x_k\theta_k + \varepsilon_k \\ &= \pi(k; x_k, \varepsilon_k, \theta_k) \quad \forall k \in \mathcal{C}, k \neq m.\end{aligned}$$

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

$$\begin{aligned}M_\theta(\underline{x}, \varepsilon) = \big\{&m \in \mathcal{C} : \exists x_m \in [x_{mL}, x_{mU}] \text{ s.t. } \forall k \in \mathcal{C}, k \neq m, \\ &\exists x_k \in [x_{kL}, x_{kU}] \text{ with } x_m\theta_m + \varepsilon^m \geq x_k\theta_k + \varepsilon^k\big\}.\end{aligned} \tag{F.2}$$

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 $x_k \in [x_{kL}, x_{kU}]$, $k \in \mathcal{C}$, and, therefore, form the set of the model's predictions. As we did in Section 3, we write the set $M_\theta(\underline{x}, \varepsilon)$ and its realizations, respectively, as M_θ and $M_\theta(\omega) \equiv M_\theta(\underline{x}(\omega), \varepsilon(\omega))$, 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 A.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 $m \in \mathrm{Sel}(M_\theta)$, the individual chooses alternative $k = 0, \dots, \kappa_{\mathcal{C}} - 1$ if and only if $m(\omega) = k$. Hence, we can use a selection $m \in \mathrm{Sel}(M_\theta)$ to define a random point $q(m)$ whose realizations have coordinates $[q(m(\omega))]_k = 1(m(\omega) = k)$, $k = 0, \dots, \kappa_{\mathcal{C}} - 1$, with $1(\cdot)$ the indicator function of the event in parentheses. 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 $\mathcal{C}$,

[12] Appendix D.2 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 (F.1).

the set resulting from repeating the above construction for each $m \in \operatorname{Sel}(M_\theta)$ and given by

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

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

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

Because the probability space is nonatomic and the random set $Q(M_\theta)$ takes its realizations in a subset of the finite dimensional space $\Re^{\kappa_C}$, 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, that is, 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 $\mathbf{P}(y|\underline{x}) \equiv [\mathbf{P}(y = k|\underline{x}), k = 0, \ldots, \kappa_C - 1]$.

Using the same mathematical tools that lead to Theorem 3.2, we obtain that the set of observationally equivalent parameter values which form the sharp identification region is given by

$$\Theta_I = \left\{\theta \in \Theta : \max_{u \in B}(u'\mathbf{P}(y|\underline{x}) - \mathbf{E}[h(Q(M_\theta), u)|\underline{x}]) = 0, \underline{x}\text{-a.s.}\right\}, \tag{F.3}$$

with B the unit ball in $\Re^{\kappa_C}$.

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 $\Delta^{\kappa_C - 1}$. 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 D.1, a candidate θ belongs to Θ_I as defined in equation (F.3) if and only if $u'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}[h(Q(M_\theta), u)|\underline{x}]$ for each of the 2^{κ_C} possible u vectors whose entries are equal to either 0 or 1. Hence, Θ_I can be obtained through a finite set of moment inequalities which have to hold for $\underline{x}$-a.s.

REFERENCES

ANDREWS, D. W. K., AND X. SHI (2009): "Inference Based on Conditional Moment Inequalities," Mimeo, Yale University. [1]

ANDREWS, D. W. K., AND G. SOARES (2010): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 78, 119–157. [2]

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, Yale University. [15]

ARADILLAS-LÓPEZ, A. (2010): "Semiparametric Estimation of a Simultaneous Game With Incomplete Information," *Journal of Econometrics*, 157, 409–431. [6]
ARADILLAS-LOPEZ, A., AND E. TAMER (2008): "The Identification Power of Equilibrium in Simple Games," *Journal of Business & Economic Statistics*, 26, 261–310. [1,22]
ARTSTEIN, Z. (1983): "Distributions of Random Sets and Random Selections," *Israel Journal of Mathematics*, 46, 313–324. [17]
AUMANN, R. J. (1965): "Integrals of Set Valued Functions," *Journal of Mathematical Analysis and Applications*, 12, 1–12. [4]
——— (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics*, 1, 67–96. [23]
BERESTEANU, A., AND F. MOLINARI (2006): "Asymptotic Properties for a Class of Partially Identified Models," Working Paper CWP10/06, CeMMAP. [17]
——— (2008): "Asymptotic Properties for a Class of Partially Identified Models," *Econometrica*, 76, 763–814. [2,5,17]
BERESTEANU, A., I. MOLCHANOV, AND F. MOLINARI (2008): "Sharp Identification Regions in Games," Working Paper CWP15/08, CeMMAP. [19-21]
BERRY, S. T., AND E. TAMER (2007): "Identification in Models of Oligopoly Entry," in *Advances in Economics and Econometrics: Theory and Application*, *Ninth World Congress*, Vol. II. Cambridge: Cambridge University Press, Chapter 2, 46–85. [7,13]
BRESNAHAN, T. F., AND P. C. REISS (1990): "Entry in Monopoly Markets," *Review of Economic Studies*, 57, 531–553. [13]
——— (1991): "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99, 977–1009. [13]
BROCK, W. A., AND S. N. DURLAUF (2001): "Interactions-Based Models," in *Handbook of Econometrics*. Amsterdam: North-Holland, Chapter 54, 3297–3380. [6]
——— (2007): "Identification of Binary Choise Models With Social Interactions," *Journal of Econometrics*, 140, 52–75. [6]
CALVÓ-ARMENGOL, A. (2006): "The Set of Correlated Equilibria of 2 by 2 Games," Mimeo, Universitat Autònoma de Barcelona. [25]
CHERNOZHUKOV, V., H. HONG, AND E. TAMER (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75, 1243–1284. [3,4]
CHERNOZHUKOV, V., S. LEE, AND A. ROSEN (2009): "Intersection Bounds: Estimation and Inference," Working Paper CWP19/09, CeMMAP. [4]
CILIBERTO, F., AND E. TAMER (2009): "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 77, 1791–1828. [3,13]
FUDENBERG, D., AND J. TIROLE (1991): *Game Theory*. Cambridge, MA: MIT Press. [4]
GALICHON, A., AND M. HENRY (2006): "Inference in Incomplete Models," Mimeo, École Polytechnique and Université de Montréal. [17]
GRIECO, P. L. E. (2009): "Discrete Choice Games With Flexible Information Structure: An Application to Local Grocery Markets," Mimeo, Pennsylvania State University. [7,10]
IMBENS, G. W., AND C. F. MANSKI (2004): "Confidence Intervals for Partially Identified Parameters," *Econometrica*, 72, 1845–1857. [3]
MANSKI, C. F., AND E. TAMER (2002): "Inference on Regressions With Interval Data on a Regressor or Outcome," *Econometrica*, 70, 519–546. [27,28]
MOLCHANOV, I. (2005): *Theory of Random Sets*. London: Springer-Verlag. [6,10,17]
MOURIER, E. (1955): "L-Random Elements and L^*-Random Elements in Banach Spaces," in *Proceedings of the Third Berkeley Symposium in Mathematical Statistics and Probability*, Vol. 2. Berkeley: University of California Press, 231–242. [6]
NORBERG, T. (1992): "On the Existence of Ordered Couplings of Random Sets—With Applications," *Israel Journal of Mathematics*, 77, 241–264. [17]
PAKES, A., AND D. POLLARD (1989): "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57, 1027–1057. [5]

ROMANO, J. P., AND A. M. SHAIKH (2010): "Inference for the Identified Set in Partially Identified Econometric Models," *Econometrica*, 78, 169–211. [3]
SCHNEIDER, R. (1993): *Convex Bodies: The Brunn–Minkowski Theory*. Cambridge: Cambridge University Press. [10]
SEIM, K. (2006): "An Empirical Model of Firm Entry With Endogenous Product-Type Choices," *RAND Journal of Economics*, 37, 619–640. [6]
STOYE, J. (2009): "More on Confidence Intervals for Partially Identified Parameters," *Econometrica*, 77, 1299–1315. [3]
SWEETING, A. (2009): "The Strategic Timing of Radio Commercials: An Empirical Analysis Using Multiple Equilibria," *RAND Journal of Economics*, 40, 710–742. [6]
TAMER, E. (2003): "Incomplete Simultaneous Discrete Response Model With Multiple Equilibria," *Review of Economic Studies*, 70, 147–165. [11]
YANG, Z. (2008): "Correlated Equilibrium and the Estimation of Discrete Games of Complete Information," Mimeo, Virginia Tech. [24]

Dept. of Economics, University of Pittsburgh, 4510 Posvar Hall, Pittsburgh, PA 15260, U.S.A.; arie@pitt.edu,

Institute of Mathematical Statistics and Actuarial Science, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland; ilya@stat.unibe.ch,

and

Dept. of Economics, Cornell University, 492 Uris Hall, Ithaca, NY 14853-7601, U.S.A.; fm72@cornell.edu.

Manuscript received July, 2009; final revision received March, 2010.