Supplemental Material for:

Sharp Identification Regions in Models with Convex Moment Predictions

Arie Beresteanu* Ilya Molchanov† Francesca Molinari‡

First draft: February 15, 2007. Current draft: August 30, 2010

*Department of Economics, Duke University, arie@econ.duke.edu.
†Department of Mathematical Statistics and Actuarial Science, University of Bern, ilya@stat.unibe.ch.
‡Department of Economics, Cornell University, fm72@cornell.edu.

Outline

This Supplement includes four appendices. Appendix B specializes our results for 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 in order to compute it. Appendix C shows that our methodology is applicable to static simultaneous move finite games regardless of the solution concept used.[1] Appendix D applies the results in Section 2 of Beresteanu, Molchanov, and Molinari (2010, BMM henceforth) to the analysis of individual decision making, looking at random utility models of multinomial choice in the presence of interval regressors data.

B 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 possible non-existence of an equilibrium for certain parameter values $\theta \in \Theta$ and realizations of $(\underline{x}, \varepsilon)$. To deal with this problem, one can impose Assumption B.1 below:

Assumption B.1 *One of the following holds:*
(i) For a subset of values of $\theta \in \Theta$ which include the values of θ that have generated the observed outcomes y, a pure strategy Nash equilibrium exists $(\underline{x}, \varepsilon) - a.s.$
(ii) For each $\theta \in \Theta$ and realizations of $\underline{x}, \varepsilon$ such that a pure strategy Nash equilibrium does not exist, $S_\theta(\underline{x}, \varepsilon) = \text{vert}(\Sigma(\mathcal{Y}))$, with $\text{vert}(\cdot)$ the vertices of the set in parenthesis.

Assumption B.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 $\underline{x}, \varepsilon$. On the other hand, Assumption B.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 B.1-(ii) is more conservative than Assumption B.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

[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.

delivers the sharp identification region Θ_I, although the set Θ_I will differ depending on whether Assumption B.1-(i) or B.1-(ii) is imposed. Moreover, one may choose not to impose Assumption B.1 at all, and use a different solution concept. In that case as well, as we illustrate in Appendix C, 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 one 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, 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 1 plots the set* S_θ *resulting from the possible realizations of* $\varepsilon_1, \varepsilon_2$. *In this case,* S_θ *assumes only five values:*

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

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* $\{[0\ \ 1\ \ 0\ \ 0]', [0\ \ 0\ \ 1\ \ 0]'\}$. $\square$

Hence, the sets S_θ and Q_θ are "simple" random closed sets in $\Sigma(\mathcal{Y})$ and $\Delta^{\kappa_{\mathcal{Y}}-1}$, respectively (see Definition A.1 in BMM). Because the probability space is non-atomic 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 (Cont.) *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*

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

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

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, i.e. whether a finite number of moment inequalities hold $\underline{x} - a.s.$ In Theorem B.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 zero or to one.

Theorem B.1 *Assume that players use only pure strategies, that Assumptions 3.1 and 3.2 in BMM and Assumption B.1 are satisfied, and that no other information is available. Then for* $\underline{x} - a.s.$ *these two conditions are equivalent:*

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

2. $u'\mathbf{P}(y|\underline{x}) \leq \mathbf{E}[h(Q_\theta, u)|\underline{x}]\ \forall\ u \in D = \left\{u = \begin{bmatrix} u_1 & \dots & u_{\kappa_\mathcal{Y}} \end{bmatrix}' : u_i \in \{0,1\},\ i = 1, ..., \kappa_\mathcal{Y}\right\}$.

Proof. Same argument as in the proof of Theorem 4.1 in BMM. ■

In Appendix B.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 B.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 B.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.[2] 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 $\mathbf{P}(y|\underline{x})$

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

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 of BMM.

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

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

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

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

We assume that θ_{11}, θ_{21} and θ_{22} are strictly negative and that $\theta_{22} > \theta_{21}$. This means that a type 2 firm is worried more about rivals of type 1 than of rivals of its own type. Since firms of a given type are indistinguishable to the econometrician, the observable outcome is the number of firms of each type which enter the market. Let $y_1 = a_{11} + a_{21}$ denote the number of entrants of type 1 and $y_2 = a_{12} + a_{22}$ the number of entrants of type 2 that a firm faces, so that $y_m \in \{0,1,2\}$, $m = 1,2$. Then there are 9 possible outcomes to this game, ordered as follows: $\mathcal{Y} = \{(0,0), (0,1), (1,0), (1,1), (2,0), (0,2), (1,2), (2,1), (2,2)\}$. Notice that here players' actions and observable outcomes of the game differ. Figure 2 plots the outcomes of the game against the realizations of $\varepsilon_1, \varepsilon_2$. 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 $Q_\theta(\omega) = \{[0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0]', [0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0]'\}$; etc.

Because the set $\mathcal{Y}$ has cardinality 9, in principle there are $2^9 = 512$ inequality restrictions to consider, corresponding to each binary vector of length 9. However, the number of inequalities to be checked is significantly smaller. Because we are allowing only pure strategy equilibria, the realizations of any $\sigma \in S_\theta$ are vectors of zeros and ones. Hence, $\forall \omega \in \Omega$, $[q(\sigma(\omega))]_k = 1$ if $\prod_{j=1}^{J} \sigma_j\left(\omega, t_j^k\right) = 1$, and zero otherwise. Consider two equilibria $t^k, t^l \in \mathcal{Y}$, $1 \leq k \neq l \leq \kappa_{\mathcal{Y}}$, such that

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

that is, the set of ω for which S_θ admits both t^k and t^l as equilibria has probability zero. Let u^k be a vector with each entry equal to zero, and entry k equal to 1, and similarly for u^l. Then the inequality $\left(u^k + u^l\right)' \mathbf{P}\left(y | \underline{x}\right) \leq \mathbf{E}\left[h\left(Q\left(S_\theta\right), u^k + u^l\right) \middle| \underline{x}\right]$ does not add any information beyond that provided by the inequalities $u'\mathbf{P}\left(y | \underline{x}\right) \leq \mathbf{E}\left[h\left(Q\left(S_\theta\right), u\right) | \underline{x}\right]$ for $u = u^k$ and for $u = u^l$. The same reasoning can be extended to tuples of pure strategy equilibria of size up to $\kappa_{\mathcal{Y}}$. 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 3 and Table 1 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 $(\varepsilon_1, \varepsilon_2) \overset{iid}{\sim} N(0,1)$ and $\Theta = [-5, 0]^3$. In the figure, Θ_O^{ABJ} is given by the union of the yellow, red and black segments, and Θ_O^{CT} by the union of the red and black segments. Θ_I is the black segment. Notice that the identification regions are segments because the outcomes $(0,0)$ and $(2,2)$ can only occur as unique equilibrium outcomes, and therefore imply two moment equalities which make θ_{21} and θ_{22} a function of θ_{11}. While, strictly speaking, the approach in ABJ does not take into account this fact, as it uses only upper bounds on the probabilities that each outcome occurs, it is clear (and indicated in their paper) that one can incorporate equalities into their method. Hence, we use the equalities on $\mathbf{P}(y = (0,0))$ and $\mathbf{P}(y = (2,2))$ also when calculating Θ_O^{ABJ}. We generate the data with $\theta_{11}^\star = -0.15$, $\theta_{21}^\star = -0.20$, and $\theta_{22}^\star = -0.10$ and use a selection mechanism to choose the equilibrium played in the many regions of multiplicity. The resulting observed distribution is $\mathbf{P}(y) = [0.3021\ 0.0335\ 0.0231\ 0.0019\ 0.2601\ 0.2779\ 0.0104\ 0.0158\ 0.0752]'$. Our results clearly show that Θ_I is substantially smaller than Θ_O^{CT} and Θ_O^{ABJ}. The width of the bounds on each parameter vector obtained using our method is about 46% of the width obtained using ABJ's method, and about 63% of the width obtained using CT's method.

In order to further illustrate the computational advantages of our characterization of Θ_I in Theorem 3.2 of BMM, we also recalculated the sharp identification region for this example *without* taking advantage of our knowledge of the structure of the game which reduces the number of inequalities to be checked to 26, but by simply solving for each candidate $\theta \in \Theta$ the problem $\max_{u \in B}\left(u'\mathbf{P}\left(y | \underline{x}\right) - \mathbf{E}\left[h\left(Q_\theta, u\right) | \underline{x}\right]\right)$. We modified the simple Nelder-Mead algorithm described in Section 3.4 of BMM to apply to a minimization in $\Re^9$, wrote it as a program in Fortran 90, and compiled and ran it on a Unix machine with a single processor of 3.2 GHz. Our recalculation of Θ_I yielded exactly the same result as described above, and checking 10^6 candidate values for $\theta \in \Theta$ took less than one minute.

B.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[3] is

(B.4)

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

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 of BMM, one can easily show that Y_θ is a random closed set in $\mathcal{Y}$ (see Definition A.1 of BMM). Because the realizations of Y_θ are subsets of the finite set $\mathcal{Y}$, it suffices that $\pi(\cdot)$ is a measurable (rather than continuous) function of $\underline{x}$ and ε in order to establish that Y_θ is a random closed set in $\mathcal{Y}$.

The researcher observes the tuple $(y, \underline{x})$, and the random set Y_θ is a function of $\underline{x}$ (and of course ε). Under Assumptions 3.1 and 3.2 of BMM and Assumption B.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 in BMM). 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 to the following:[4]

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

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

(B.5) $$\mathbf{P}\{y \in K | \underline{x}\} \leq \mathbf{P}\{Y_\theta \cap K \neq \emptyset | \underline{x}\}\ \forall K \subset \mathcal{Y}\ \underline{x} - a.s.$$

[3] 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

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

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

Because $\mathcal{Y}$ is finite, all its subsets are compact. The functional $\mathbf{P}\{Y_\theta \cap K \neq \emptyset | \underline{x}\}$ on the right-hand side of (B.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 B.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*

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

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

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

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

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

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

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

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 (B.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 (B.7) is superfluous: when the inequalities in (B.7) are used, only subsets $K \subset \mathcal{Y}$ of cardinality up to half of the cardinality of $\mathcal{Y}$ are needed.

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

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

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:

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

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

Both ABJ and CT acknowledge that the parameter regions they give are not sharp. Comparing the sets in equations (B.9)-(B.10) with the set in equation (B.8), one observes that Θ_O^{ABJ} is obtained applying inequality (B.5) only for $K = \{t\}$ $\forall t \in \mathcal{Y}$. Similarly, Θ_O^{CT} is obtained applying inequality (B.7) only for $K = \{t\}$ (or, equivalently, applying inequality (B.5) for $K = \{t\}$ and $K = \mathcal{Y} \setminus \{t\}$ $\forall t \in \mathcal{Y}$). Clearly both ABJ and CT do not use the information contained in the remaining subsets of $\mathcal{Y}$, while this information is used to obtain Θ_I. Two questions arise: (1) whether Θ_I as defined in equation (B.8) yields the sharp identification region of θ; and (2) if and by how much Θ_I differs from Θ_O^{ABJ} and Θ_O^{CT}. We answer here the first question. Appendix B.1 answers the second question by looking at a simple example.

Theorem B.2 *Assume that players use only pure strategies, that Assumptions 3.1 and 3.2 in BMM and Assumption B.1 are satisfied, and that no other information is available. Then for $\underline{x}$ − a.s. these two conditions are equivalent:*

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

2. $\mathbf{P}\{y \in K|\underline{x}\} \leq \mathbf{T}_{Y_\theta|\underline{x}}(K)$ $\forall K \subset \mathcal{Y}$.

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

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

As discussed in Appendix B.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 B.2, one can show that the result in equation (B.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

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

that is, the set of ω for which Y_θ intersects both K_1 and K_2 has probability zero, then the inequality $\mathbf{P}\{y \in K_1 \cup K_2|\underline{x}\} \leq \mathbf{P}\{Y_\theta \cap (K_1 \cup K_2) \neq \emptyset|\underline{x}\}$ does not add any information beyond that provided by the inequalities $\mathbf{P}\{y \in K_1|\underline{x}\} \leq \mathbf{P}\{Y_\theta \cap K_1 \neq \emptyset|\underline{x}\}$ and $\mathbf{P}\{y \in K_2|\underline{x}\} \leq \mathbf{P}\{Y_\theta \cap K_2 \neq \emptyset|\underline{x}\}$. Therefore, prior knowledge of some properties of the game can be very helpful in eliminating unnecessary inequalities. For example, in a Bresnahan and Reiss entry model with 4 players, if the number of entrants is identified, the number of inequalities to be verified reduces from 65,536 to at most 100. Theorem B.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 B.3 *Take $\theta \in \Theta$ and let Assumptions 3.1 and 3.2 in BMM and Assumption B.1 hold. Consider a partition of Ω into sets $\Omega^1, \ldots, \Omega^M$ of positive probability. Let $\mathcal{Y}_i$*

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

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

Proof. Beresteanu, Molchanov, and Molinari (2008, Theorem 5.1). ■

A simple Corollary of Theorem B.3, the proof of which is omitted, is the following:

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

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

$\underline{x} - a.s.$ 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.[5] 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 (B.3) shows that actually the sharp identification region can be obtained by checking only five inequalities which have to hold for $\underline{x} - a.s.$, given by inequalities (B.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 of BMM shows that CT's approach does not yield the sharp identification region when mixed strategies are allowed for.

When no prior knowledge of the game such as, for example, that required in Theorem B.3 is available, it is still possible to use the insight in equation (B.11) to determine which inequalities yield the sharp identification region, by decomposing $\mathcal{Y}$ into subsets such that Y_θ does not jointly

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

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 B.5 *Let Assumptions 3.1 and 3.2 in BMM and Assumption B.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 $\mathbf{P}\left(\tilde{\mathcal{X}}\right) > 0$, there exist $t^1, t^2 \in \mathcal{Y}$ with*

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

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

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

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

Proof. Beresteanu, Molchanov, and Molinari (2008, Theorems 5.2 and 5.3) ■
These results show that the extra inequalities matter in general, compared to those used by ABJ, and CT, to fully characterize Y_θ and determine if $y \in \mathrm{Sel}(Y_\theta)$. In fact, the assumptions of Theorem B.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 one, and it has at least three different equilibria. The assumptions of Theorem B.5-(b) are satisfied whenever (1) there are regions of the unobservables of positive probability where two different outcomes can result from equilibrium strategy profiles; and (2) the probability that the cardinality of Y_θ is greater than one exceeds the probability that each of these two outcomes is not a unique equilibrium. It is easy to see that these assumptions are not satisfied in a two player entry game where players are allowed only to play pure strategies, but they are satisfied in the four player, two type game described in Section B.1.

C Extensions to Other Solution Concepts

While in Section 3 of BMM and Appendix B 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.

C.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

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

Omitting the explicit reference to its dependence on $\underline{x}$ and ε, R_θ is the set of level-1 rational strategy profiles of the game. By similar arguments to what we used above, this is a random closed set in $\Sigma(\mathcal{Y})$. Figure 4 plots this set against the possible realizations of $\varepsilon_1, \varepsilon_2$, in a simple two player simultaneous move, complete information, static game of entry 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, θ_2 are assumed to be negative.

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

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

with

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

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

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

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

$k = 1, \ldots, 4$, and therefore it follows that $\mathbb{E}(Q(R_\theta))$ is equal to $\mathbb{E}\left(Q\left(\tilde{R}_\theta\right)\right)$, with $\tilde{R}_\theta$ restricted to be the set of level-1 rational pure strategies. Hence, by Theorem B.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) suggest to obtain the sharp identification region as the set of parameter values that return value zero for the objective function of a linear programming problem. For the general case in which ε may have a continuous distribution, AT apply the same insight of CT and characterize an outer identification region through eight moment inequalities similar to those in equation (B.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 (B.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, θ_2. Figure 5 plots the upper contours of Θ_I, Θ_O^{CT}, and Θ_O^{ABJ} in a simple example with $(\varepsilon_1, \varepsilon_2) \stackrel{iid}{\sim} N(0,1)$ and $\Theta = [-5, 0]^2$. The data is generated with $\theta_1^\star = -1.15$, $\theta_2^\star = -1.4$, and using a selection mechanism which picks outcome $(0,0)$ for 40% of $\omega : \varepsilon(\omega) \in [0, -\theta_1^\star] \times [0, -\theta_2^\star]$, outcome $(1,1)$ for 10% of $\omega : \varepsilon(\omega) \in [0, -\theta_1^\star] \times [0, -\theta_2^\star]$, and each of outcome $(1,0)$ and $(0,1)$ for 25% of $\omega : \varepsilon(\omega) \in [0, -\theta_1^\star] \times [0, -\theta_2^\star]$. Hence, the observed distribution is $\mathbf{P}(y) = [0.5048\ 0.2218\ 0.1996\ 0.0738]'$. Our methodology allows us to obtain significantly lower upper contours compared to AT (and CT) and ABJ. The upper bounds on θ_1, θ_2 resulting from the projections of Θ_O^{ABJ}, Θ_O^{CT} and Θ_I are, respectively, $(-0.02, -0.02)$, $(-0.15, -0.26)$, and $(-0.54, -0.61)$.

C.2 Objective Correlated Equilibria

Suppose that players play correlated equilibria, a notion introduced by Aumann (1974). A correlated equilibrium can be interpreted as the distribution of play instructions given by some "trusted authority" to the players. Each player is given her instruction privately but does not know the instruction received by others. The distribution of instructions is common knowledge across all players. Then a correlated joint strategy $\gamma \in \Delta^{\kappa_{\mathcal{Y}}-1}$, 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

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

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

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

In our simple two player simultaneous move, complete information, static game of entry, $\mathcal{Y}_j = \{0,1\}$, $j = 1,2$, $\mathcal{Y} = \{(0,0), (1,0), (0,1), (1,1)\}$. Omitting again the covariates, we assume that players' payoffs are given by $\pi_j = y_j\left(y_{-j}\theta_j + \varepsilon_j\right)$, 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 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 $Q_\theta(\omega)$ resulting from the unique Nash equilibrium in these regions. For $\omega \in \Omega$ such that $\varepsilon(\omega) \in [0, -\theta_1] \times [0, -\theta_2]$, $C_\theta(\omega)$ is given by a polytope with five vertices, three of which are implied by Nash equilibria, see Calvó-Armengol (2006).[6] 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 7 and Table 2 report Θ_I, Θ_O^{CT}, and Θ_O^{ABJ} in a simple example with $(\varepsilon_1, \varepsilon_2) \overset{iid}{\sim} N(0,1)$ and $\Theta = [-5, 0]^2$. In the figure, Θ_O^{ABJ} is given by the union of the yellow, red and black areas, and Θ_O^{CT} by the union of the red and black areas. Θ_I is the black region. The data is generated with $\theta_1^\star = -1.15, \theta_2^\star = -1.4$, and using a selection mechanism which picks each of outcome $(0,0)$ and $(1,1)$ for 10% of $\omega : \varepsilon(\omega) \in [0, -\theta_1^\star] \times [0, -\theta_2^\star]$, and each of outcome $(1,0)$ and $(0,1)$ for 40% of $\omega : \varepsilon(\omega) \in [0, -\theta_1^\star] \times [0, -\theta_2^\star]$. Hence, the observed distribution is $\mathbf{P}(y) = [0.26572\ 0.34315\ 0.36531\ 0.02582]'$. Also in this case Θ_I is smaller than Θ_O^{CT} and Θ_O^{ABJ}, although the reduction in the size of the identification region is less pronounced than in the case where mixed strategy Nash equilibrium is the solution concept.

[6] These vertices are

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

D Multinomial Choice Models with Interval Regressors Data

This Section of the Supplement applies the methodology introduced in Section 2 of BMM to provide a tractable characterization of the sharp identification region of the parameters θ characterizing random utility models of multinomial choice, when only interval information is available on regressors. In doing so, we extend the seminal contribution of Manski and Tamer (2002), who considered the same inferential problem in the case of binary choice models. For these models, Manski and Tamer (2002) provided a tractable characterization of the sharp identification region, and proposed set estimators which are consistent with respect to the Hausdorff distance. However, their characterization of the sharp identification region does not easily extend to models in which the agents face more than two choices, as we illustrate below.

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

To simplify the exposition, we assume that the random utility is linear, and that w, z_k and ε_k, $k = 0, \ldots, \kappa_{\mathcal{C}} - 1$, 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 $x_k = [1 \quad z_k \quad w]$ and $\theta_k = [\alpha_k \quad \delta \quad \beta_k]'$. 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 $\theta = \left[\{\alpha_k\}_{k=1}^{\kappa_{\mathcal{C}}-1} \quad \delta \quad \{\beta_k\}_{k=1}^{\kappa_{\mathcal{C}}-1}\right]' \in \Theta$ be the vector of parameters of interest, with Θ the parameter space. We denote $\varepsilon^k = \varepsilon_k - \varepsilon_0$, $k \in \mathcal{C}$, and $\varepsilon = \left[\left\{\varepsilon^k\right\}_{k=1}^{\kappa_{\mathcal{C}}-1}\right]$. Under these assumptions, if the econometrician observes a random sample of choices, socioeconomic characteristics, and alternatives' attributes, the parameter vector θ is point identified.

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

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

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

This construction is based on the observation that if the agent chooses alternative 1, this implies that $\varepsilon^1 > -x_1\theta \geq -x_{1U}\theta$. On the other hand, $\varepsilon^1 > -x_{1L}\theta \geq -x_1\theta$ implies that the agent chooses alternative 1.[7] 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

(D.1)

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

The lower bound on $\mathbf{P}\left(y = m \middle| \underline{x}\right)$ in equation (D.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 of BMM, Θ_O is just an outer region for θ, and is not sharp in general. Appendix B.2 provides further insights to explain the lack of sharpness of Θ_O.[8]

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

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

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

(D.2)

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

This is the set of choices associated with a specific value of θ and realization of $\underline{x}$ and ε, which are optimal for some combination of $x_k \in [x_{kL}, x_{kU}]$, $k \in \mathcal{C}$, and therefore form the set of model's

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

[8] Appendix B.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 (D.1).

predictions. As we did in Section 3 of BMM, 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 of BMM.

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 of BMM. 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 $k = 0, \ldots, \kappa_{\mathcal{C}} - 1$ 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 $[q(m(\omega))]_k = 1(m(\omega) = k)$, $k = 0, \ldots, \kappa_{\mathcal{C}} - 1$, with $1(\cdot)$ the indicator function of the event in parenthesis. Clearly, the random point $q(m)$ is an element of the unit simplex in the space of dimension $\kappa_{\mathcal{C}}$, denoted $\Delta^{\kappa_{\mathcal{C}}-1}$. Because M_θ is a random closed set in $\mathcal{C}$, the set resulting from repeating the above construction for each $m \in \mathrm{Sel}(M_\theta)$ and given by

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

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

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

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

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

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

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

Notice that the set $Q(M_\theta)$ assumes at most a finite number of values, and its realizations lie in the subsets of the vertices of $\Delta^{\kappa_C - 1}$. The conditional Aumann expectation of $Q(M_\theta)$ is given by the weighted Minkowski sum of the possible realizations of 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 B.1, a candidate θ belongs to Θ_I as defined in equation (D.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 either equal to zero or to one. Hence, Θ_I can be obtained through a finite set of moment inequalities which have to hold for $\underline{x} - a.s.$

References

Andrews, D. W. K., S. T. Berry, and P. Jia (2004): "Confidence Regions for Parameters in Discrete Games with Multiple Equilibria, with an Application to Discount Chain Store Location," mimeo.

Aradillas-Lopez, A., and E. Tamer (2008): "The Identification Power of Equilibrium in Simple Games," *Journal of Business and Economic Statistics*, 26(3), 261–310.

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

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

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

——— (2010): "Sharp Identification Regions in Models with Convex Moment Predictions," mimeo.

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

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

Bresnahan, T. F., and P. C. Reiss (1990): "Entry in Monopoly Markets," *Review of Economic Studies*, 57, 531–553.

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

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

Ciliberto, F., and E. Tamer (2009): "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 77, 1791–1828.

Galichon, A., and M. Henry (2006): "Inference in Incomplete Models," mimeo.

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

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

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

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

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

E Tables and Figures

Table 1: Projections of Θ_O^{ABJ}, Θ_O^{CT} and Θ_I, and reduction in bounds width compared to ABJ. Four player, two type entry game with pure strategy Nash equilibrium as solution concept.

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

Table 2: Projections of Θ_O^{ABJ}, Θ_O^{CT} and Θ_I, reduction in bounds width (in parentheses), and area of the identification regions compared to ABJ. Two player entry game with correlated equilibrium as solution concept.

	True Values	Projections of:		
		Θ_O^{ABJ}	Θ_O^{CT}	Θ_I
$\theta_1^{\star}$	-1.15	$[-4.475, -0.485]$	$[-4.475, -0.585]$	$[-4.125, -0.595]$
			(2.5%)	(11.5%)
$\theta_2^{\star}$	-1.40	$[-4.585, -0.625]$	$[-4.585, -0.725]$	$[-4.425, -0.735]$
			(2.4%)	(6.8%)
Approximate Reduction in Total Area Compared to Θ_O^{ABJ}			(7.9%)	(23.1%)

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

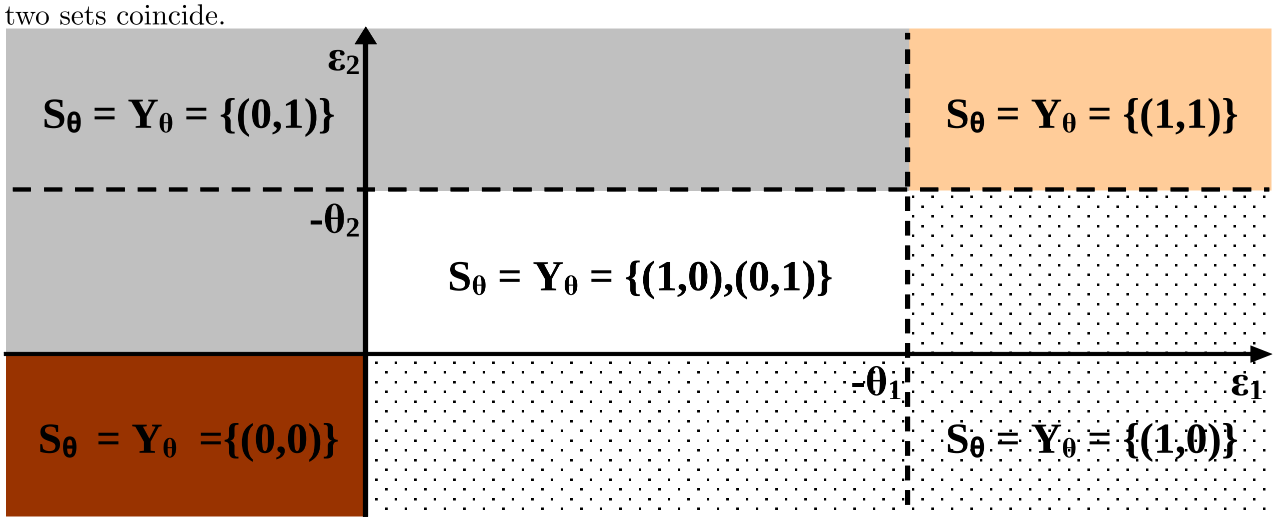

Figure 2: The random set of pure strategy NE outcomes as a function of $\varepsilon_1, \varepsilon_2$ in a four player, two type entry game.

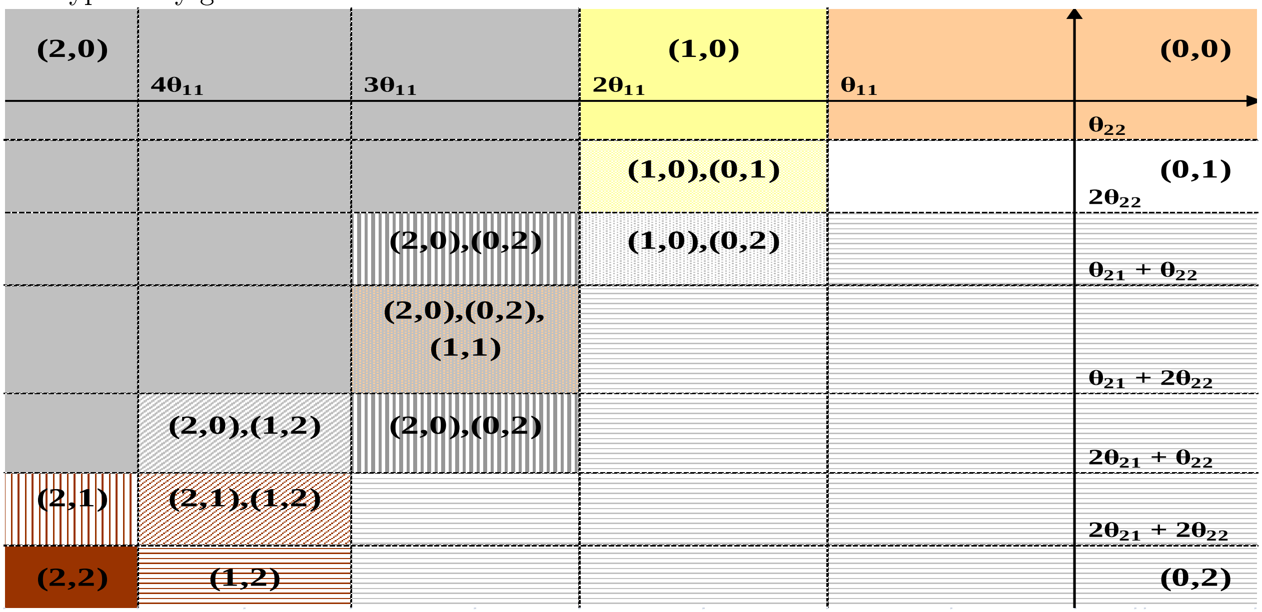

Figure 3: Identification regions in a four player, two type entry game with pure strategy Nash equilibrium as solution concept.

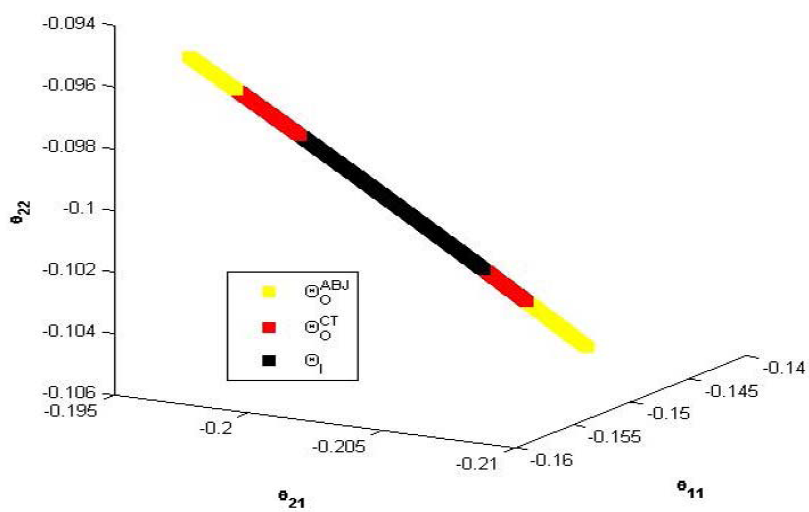

Figure 4: The random set of level-1 rational profiles as a function of $\varepsilon_1, \varepsilon_2$ in a two player entry game.

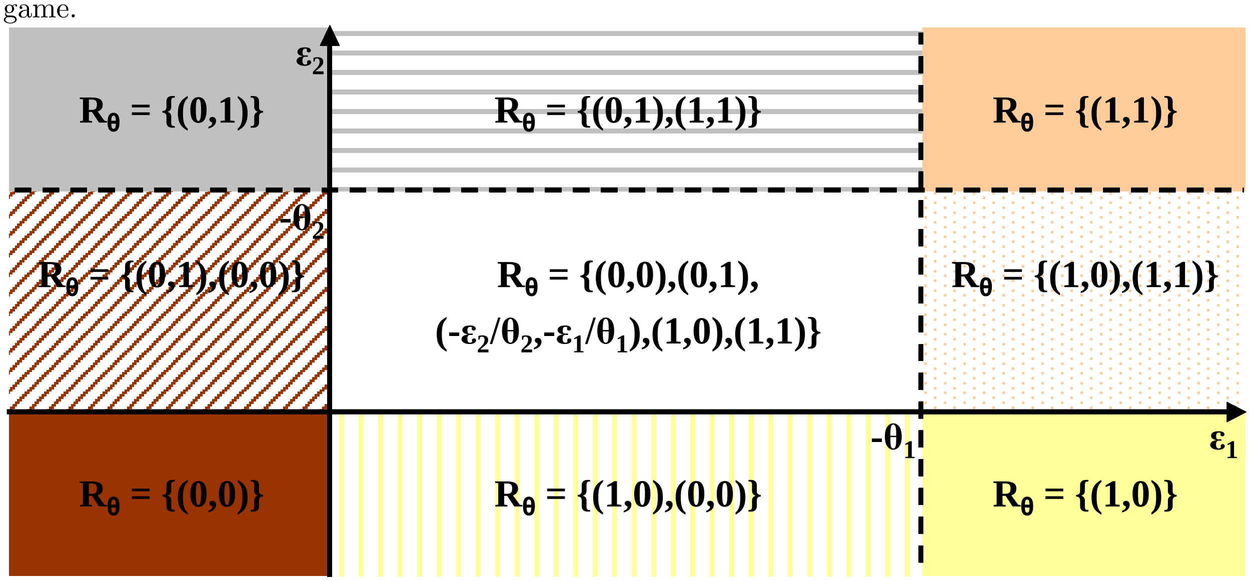

Figure 5: Upper contours of the identification regions in a two player entry game with level-1 rationality as solution concept.

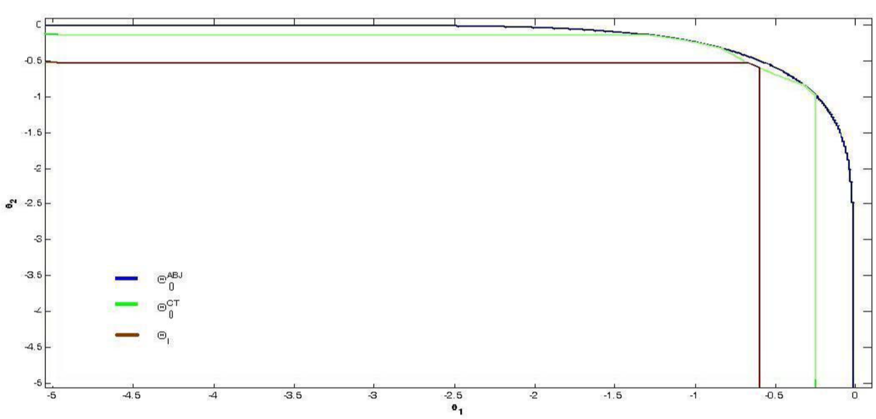

Figure 6: The random set of correlated equilibria as a function of ε_1,ε_2 in a two player entry game. The correlated equilibria $\gamma_1, \gamma_2, \gamma_3$ are defined in Section C.2.

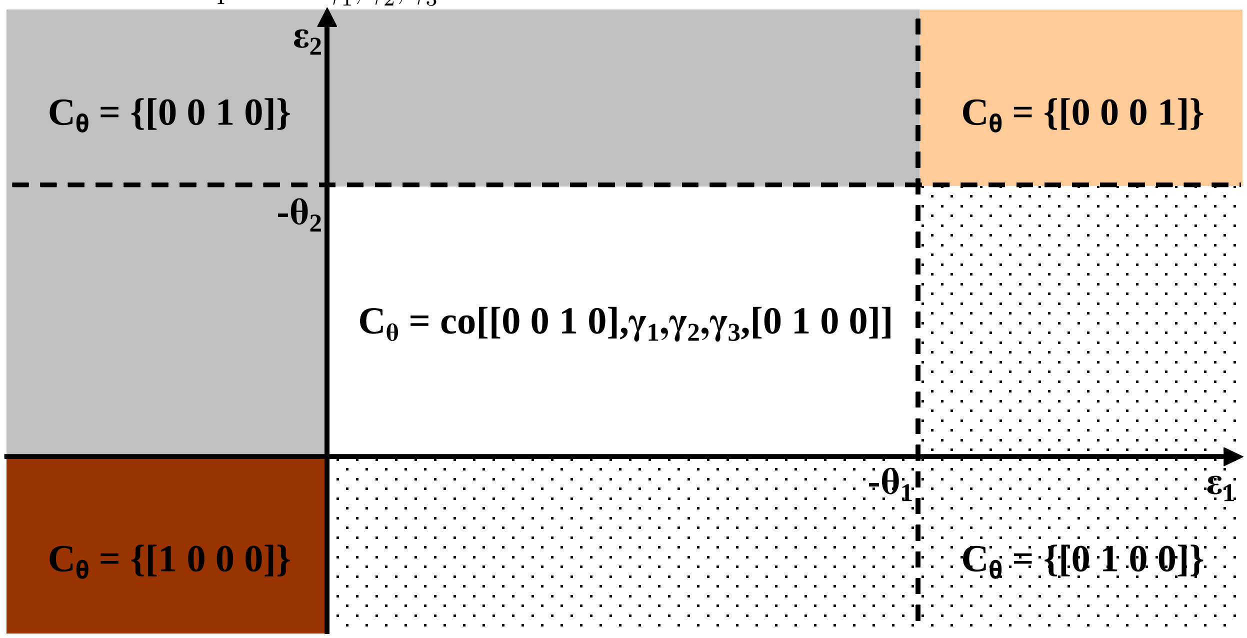

Figure 7: Identification regions in a two player entry game with correlated equilibrium as solution concept.