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.

 DOI: 10.3982/ECTA8680

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

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 $F(y|\underline{x}) \equiv [\mathbf{P}(y = t^k|\underline{x}), k = 1, \ldots, \kappa_{\mathcal{Y}}]$. 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

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

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

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

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

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

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

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

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 $\{f(\underline{x}_{in}, \theta_n, u),\ u \in S^{\kappa_{\mathcal{Y}}-1},\ i \leq n,\ n \geq 1\}$. Assumption M2(a) holds because for all $u \in S^{\kappa_{\mathcal{Y}}-1}$,

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

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 $f(\underline{x}, \theta, u)$ 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 $\{(u'w(y_{in}) - f(\underline{x}_{in}, \theta_n, u)),\ u \in S^{\kappa_{\mathcal{Y}}-1},\ i \leq n,\ n \geq 1\}$ 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,

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

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

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* ε: $\pi_j(y_j, y_{-j}, x_j, \varepsilon_j; \theta) = \tilde{\pi}_j(y_j, y_{-j}, x_j; \theta) + \varepsilon_j$. *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

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

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*

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

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,

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

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 $\mathrm{co}[Q_\theta(\omega)]$ 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, $S_\theta(\underline{x}, \varepsilon) = \mathrm{vert}(\Sigma(\mathcal{Y}))$, with $\mathrm{vert}(\cdot)$ 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 $\varepsilon_1, \varepsilon_2$. 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 $\{[0\ 1\ 0\ 0]', [0\ 0\ 1\ 0]'\}$.

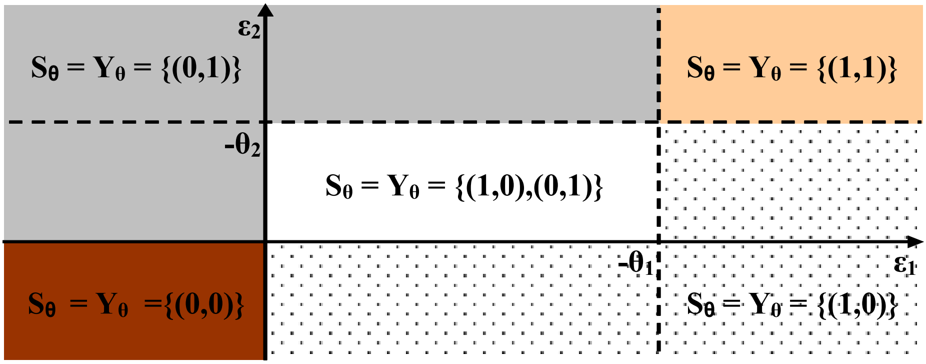

FIGURE S.1.—The random set of pure strategy Nash equilibrium profiles S_θ and the random set of pure strategy Nash equilibrium outcomes Y_θ as a function of $\varepsilon_1, \varepsilon_2$ in a two player entry game. In this simple example, the two sets coincide.

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

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

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

(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

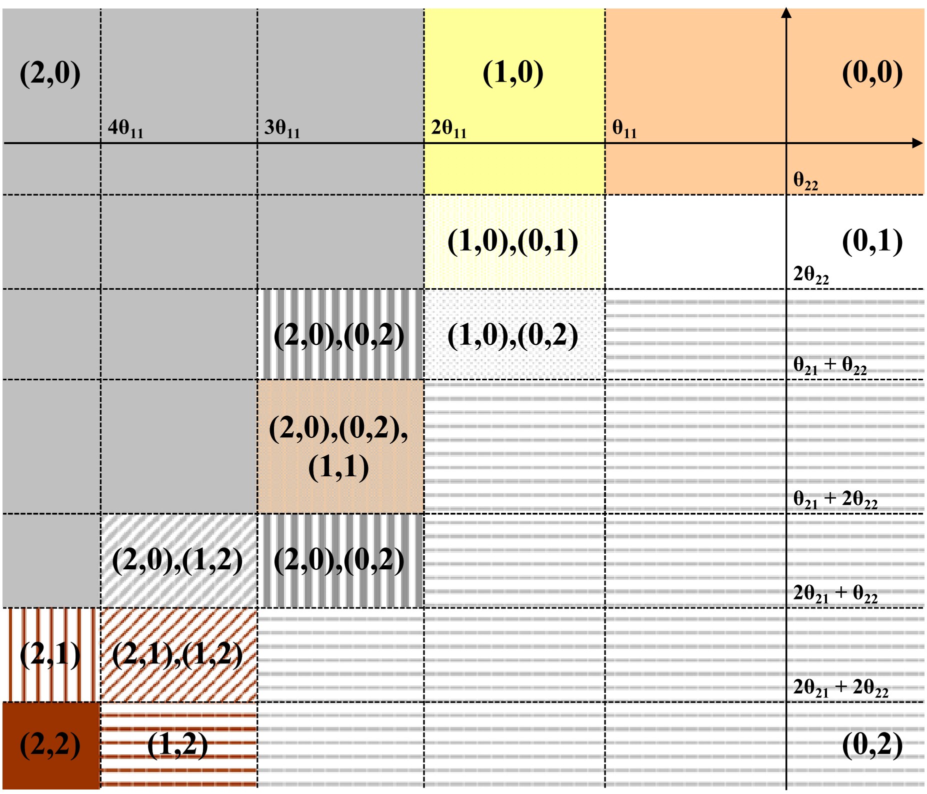

FIGURE S.2.—The random set of pure strategy Nash equilibrium outcomes as a function of $\varepsilon_1, \varepsilon_2$ in a four player, two type entry game.

$\varepsilon_1, \varepsilon_2$. 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

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

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 $\mathbf{P}(y) =$

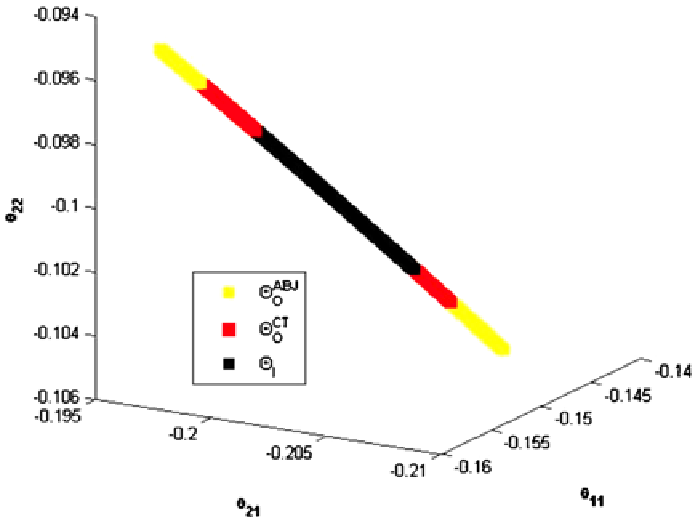

FIGURE S.3.—Identification regions in a four player, two type entry game with pure strategy Nash equilibrium as the solution concept.

[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

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

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

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

[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

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

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:

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

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

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

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

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

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

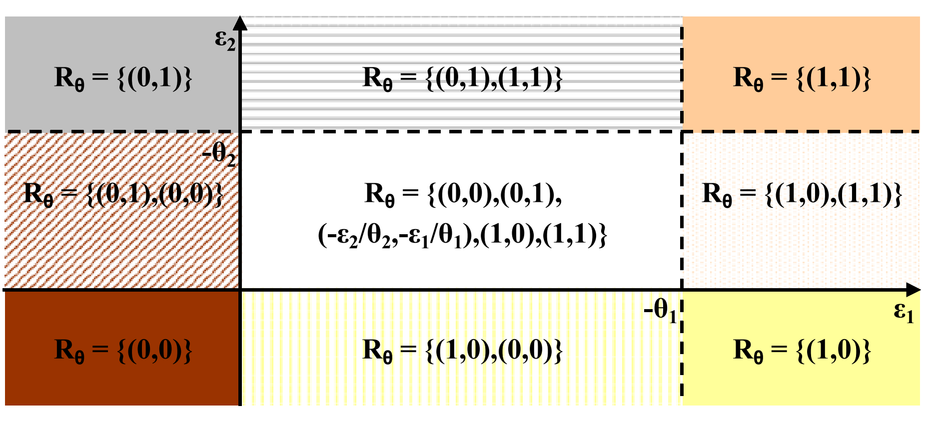

FIGURE S.4.—The random set of level-1 rational profiles as a function of $\varepsilon_1, \varepsilon_2$ in a two player entry game.

with

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

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

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

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

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

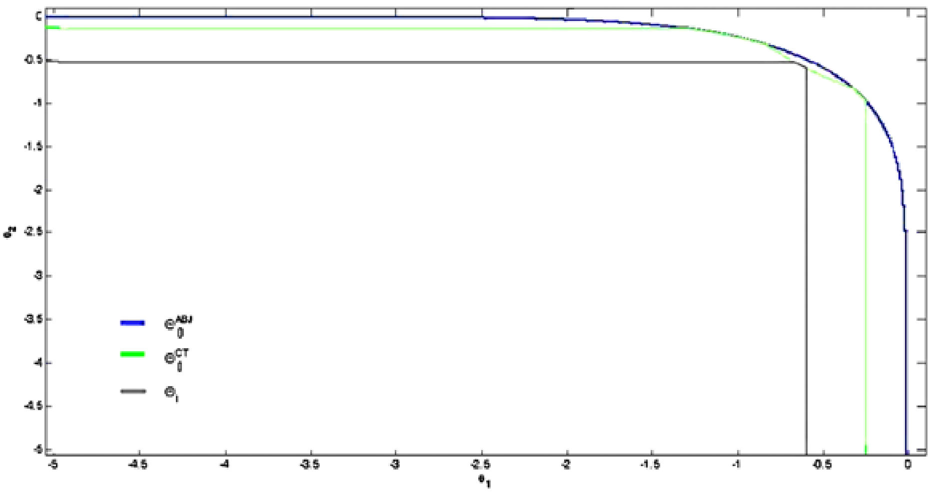

FIGURE S.5.—Upper contours of the identification regions in a two player entry game with level-1 rationality as the solution concept.

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-

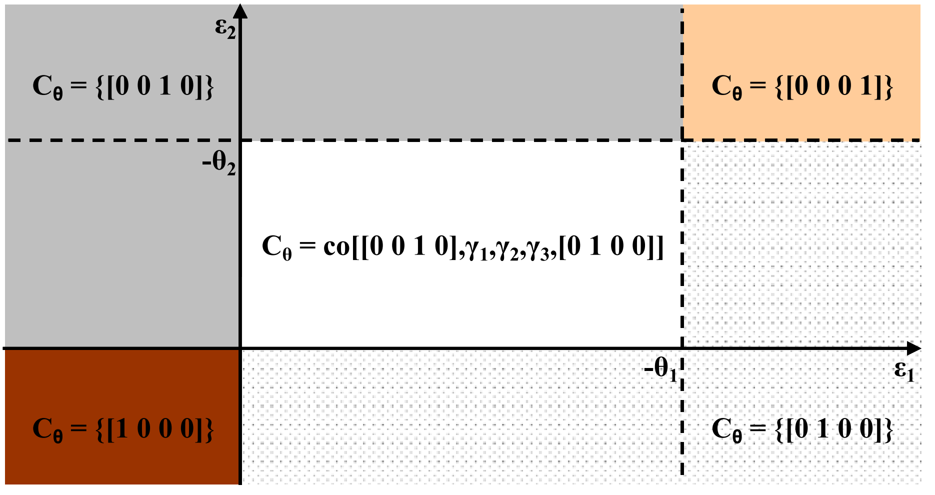

FIGURE S.6.—The random set of correlated equilibria as a function of $\varepsilon_1, \varepsilon_2$ in a two player entry game. The correlated equilibria γ_1, γ_2, and γ_3 are defined in Section E.2.

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

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

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

$$\gamma_0(\omega) = [0 \;\; 0 \;\; 1 \;\; 0]',$$

$$\gamma_1(\omega) = \left[1 \;\; -\frac{\varepsilon_2(\omega)}{\theta_2 + \varepsilon_2(\omega)} \;\; -\frac{\varepsilon_1(\omega)}{\theta_1 + \varepsilon_1(\omega)} \;\; 0\right]' \times \left(1 - \frac{\varepsilon_1(\omega)}{\theta_1 + \varepsilon_1(\omega)} - \frac{\varepsilon_2(\omega)}{\theta_2 + \varepsilon_2(\omega)}\right)^{-1},$$

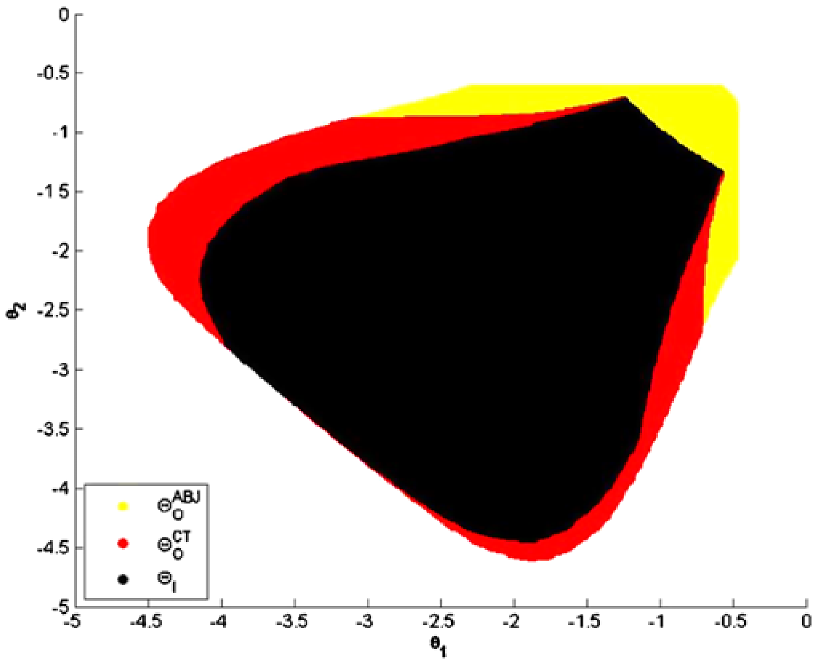

FIGURE S.7.—Identification regions in a two player entry game with correlated equilibrium as the solution concept.

$$\gamma_2(\omega) = \bigg[\bigg(1+\frac{\varepsilon_2(\omega)}{\theta_2}\bigg)\bigg(1+\frac{\varepsilon_1(\omega)}{\theta_1}\bigg) \quad -\frac{\varepsilon_2(\omega)}{\theta_2}\bigg(1+\frac{\varepsilon_1(\omega)}{\theta_1}\bigg)$$

$$-\bigg(1+\frac{\varepsilon_2(\omega)}{\theta_2}\bigg)\frac{\varepsilon_1(\omega)}{\theta_1} \quad \frac{\varepsilon_2(\omega)}{\theta_2}\frac{\varepsilon_1(\omega)}{\theta_1}\bigg]',$$

$$\gamma_3(\omega) = \bigg[0 \quad -\frac{\varepsilon_2(\omega)}{\theta_2+\varepsilon_2(\omega)} \quad -\frac{\varepsilon_1(\omega)}{\theta_1+\varepsilon_1(\omega)} \quad \frac{\varepsilon_1(\omega)}{\theta_1+\varepsilon_1(\omega)}\frac{\varepsilon_2(\omega)}{\theta_2+\varepsilon_2(\omega)}\bigg]'$$

$$\times\bigg(\frac{\varepsilon_1(\omega)}{\theta_1+\varepsilon_1(\omega)}\frac{\varepsilon_2(\omega)}{\theta_2+\varepsilon_2(\omega)} - \frac{\varepsilon_1(\omega)}{\theta_1+\varepsilon_1(\omega)} - \frac{\varepsilon_2(\omega)}{\theta_2+\varepsilon_2(\omega)}\bigg)^{-1},$$

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

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

		Projections		
	True Values	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

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

[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 $m \in \mathrm{Sel}(M_\theta)$ and given by

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

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

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

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

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

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 $\mathrm{co}[Q(M_\theta)]$. 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.