SUPPLEMENT TO "IDENTIFICATION AND ESTIMATION OF A DISCRETE GAME OF COMPLETE INFORMATION"

BY PATRICK BAJARI, HAN HONG, AND STEPHEN P. RYAN

APPENDIX B: TECHNICAL APPENDIX

B.1. *Identification at Infinity in Two-by-Two Games*

FOR THE PURPOSE OF ILLUSTRATION, in this appendix we specialize the arguments for identification to the two-by-two game.

Identifying the Mean Payoff Functions

Given a linear specification of the mean utility functions, let $f_i(\tau, x) = x_i^\tau \beta_i^\tau$ for $\tau = (2 - i)Bk + (i - 1)jR$. The following assumption requires a rich support of the covariates.

ASSUMPTION 9: *For each* $i = 1, 2$, $j = T, B$, *and* $k = L, R$, *there exists a set* $\mathcal{T}_i^{j(2-i)+k(i-1)}$ *of covariates* x *such that* $\lim_{\|x\| \to \infty, x \in \mathcal{T}_i^{j(2-i)+k(i-1)}} P[a_i = j^{2-i} k^{i-1} | x] = 1$.

Assumption 9 requires that for each player i and for each of player i's strategies, we can shift the covariates x along a dimension such that action a_i is a dominant strategy for player i with probability arbitrarily close to 1. For example, for $i = 2$ and $k = L$, Assumption 9 requires that along a path of $\|x\| \to \infty$, $x \in \mathcal{T}_2^L$, $P[a_2 = L|x] \to 1$, or

$$P[x_2^{TR}\beta_2^{TR} + \epsilon_2(TR) < 0, x_2^{BR}\beta_2^{BR} + \epsilon_2 < 0(BR)] \longrightarrow 1.$$

Assumptions 1, 2, and 9 allow us to identify the mean utilities along these paths. The next assumption requires that we can extrapolate knowledge of the deterministic utilities along this path to other values of x on its support.

ASSUMPTION 10: *For each* $i = 1, 2$, $j = T, B$, $k = L, R$, *and* $x \in \mathcal{T}_i^{j(2-i)+k(i-1)}$, *there exists some* $L_0 > 0$ *such that*

$$\inf_{L \geq L_0} \min \operatorname{eig} E\left[x_i^\tau x_i^{\tau\prime} | x \in \mathcal{T}_{3-i}^{j(i-1)+k(2-i)}, \|x\| \geq L\right] > 0.$$

Assumption 10 requires that the linear deterministic payoff functions $x_i^\tau \beta_i^\tau$ can be extrapolated from the path $\|x\| \to \infty$, $x \in \mathcal{T}_{3-i}^{j(i-1)+k(2-i)}$, to the full support of x.

THEOREM 4: *Under Assumptions* 1, 2, 9, *and* 10, $f_i((2 - i)Bk + (i - 1)jR, x)$ *is identified for all* $i = 1, 2$, $j = T, B$, *and* $k = L, R$.

PROOF: For $i = 1, 2$, $k = L, R$, and $j = T, B$, denote by $\bar{P}((2-i)Bk + (i-1)jR|x)$ the unconditional probabilities $P(\epsilon_i(\tau) + x_i^\tau \beta_i^\tau \geq 0|x)$. The data do not directly identify this probability, but only identify the conditional probabilities:

$$P\big(a_i = (2-i)B + (i-1)R | a_{3-i} = (2-i)k + (i-1)j, x\big).$$

However, because of Assumption 9,

$$\lim_{\|x\| \to \infty, x \in \mathcal{T}_{3-i}^{(i-1)j+(2-i)k}} \big[P\big(a_i = (2-i)B + (i-1)R |$$

$$a_{3-i} = (2-i)k + (i-1)j, x\big) - \bar{P}((2-i)Bk + (i-1)jR, x)\big] = 0.$$

This implies that $\bar{P}((2-i)Bk + (i-1)jR, x)$ and hence $f_i((2-i)Bk + (i-1)jR, x)$ can be identified along the path of $\|x\| \to \infty$, $x \in \mathcal{T}_{3-i}^{(i-1)j+(2-i)k}$. Because the cumulative distribution function of $\epsilon_i(\tau)$ is strictly increasing, the linear index $x_i^\tau \beta_i^\tau$ is identified along this path. Assumption 10 further identifies the coefficient parameters β_i^τ. According to Assumption 10, for every $\beta_i^\tau \neq \beta_{i0}^\tau$, there exists a set of x_i^τ with positive probabilities such that $x_i^\tau \beta_i^\tau \neq x_i^\tau \beta_{i0}^\tau$, which implies identification of β_i^τ.   Q.E.D.

A special case of Assumption 9 is when $\epsilon = (\epsilon_i(jk))$, $i = 1, 2$, $j = T, B$, and $k = L, R$, has finite support but the support for x_i^τ for $i = 1, 2$ and all τ is either larger or infinite. Denote by $\bar{U}$ an upper bound of the absolute value of the support of $\epsilon_i(jk)$ for all i, k, and j. Then a sufficient condition for Assumption 9 to hold is that for all i and τ, $P[x_i^{\tau'} \beta_i^\tau > 2\bar{U}] > 0$ and $P[x_i^\tau \beta_i^\tau < -2\bar{U}] > 0$. Then we do not need the requirement that $\|x\| \to \infty$. The sets $\mathcal{T}_i^{j(2-i)+k(i-1)}$ can be defined as $\mathcal{T}_i^{B(2-i)+R(i-1)} = \{x : x_i^\tau \beta_i^\tau > 2\bar{U}, \forall \tau\}$ and $\mathcal{T}_i^{T(2-i)+L(i-1)} = \{x : x_i^\tau \beta_i^\tau < -2\bar{U}, \forall \tau\}$. In this case, a sufficient condition for Assumption 10 to hold is that for all $i = 1$, $j = \{T, B\}$ and $i = 2$, $k = \{L, R\}$, the matrices $E[xx'|x \in \mathcal{T}_{3-i}^{(i-1)j+(2-i)k}]$ are positive definite and finite.

Identifying the Equilibrium Selection Mechanism

Given that the deterministic utility components are identified in Theorem 4, the next goal is to identify the equilibrium selection mechanism. The equilibrium selection probabilities are only needed when there are three equilibria, which can be either (TL, BR, mix) or (BL, TR, mix). The mixing probabilities for these two cases are

$$P_m(R; x, \epsilon) = \frac{f_1(BL, x) + \epsilon_1(BL)}{f_1(BL, x) - f_1(BR, x) + \epsilon_1(BL) - \epsilon_1(BR)},$$

$$P_m(L; x, \epsilon) = 1 - P_m(R; x, \epsilon)$$

and

$$P_m(B; x, \epsilon) = \frac{f_2(TR, x) + \epsilon_2(TR)}{f_2(TR, x) - f_2(BR, x) + \epsilon_2(TR) - \epsilon_2(BR)},$$

$$P_m(T; x, \epsilon) = 1 - P_m(B; x, \epsilon).$$

In the ideal case where there are no error terms, $\epsilon_1(BL) = \epsilon_1(BR) = \epsilon_2(TR) = \epsilon_2(BR) = 0$, all of $P_m(R), P_m(L), P_m(T)$, and $P_m(B)$ are functions of the known deterministic payoffs. Define the observed equilibrium selection probabilities as $\rho(TL, x), \rho(BR, x), 1 - \rho(TL, x) - \rho(BR, x)$ in the case of (TL, BR, mix) and as $\rho(BL, x), \rho(TR, x), 1 - \rho(BL, x) - \rho(TR, x)$ in the case of (BL, TR, mix), where the dependence on covariates x is made explicit. Then for those values of x where (TL, BR, mix) is realized,

$$P(TL|x) = \rho(TL, x) + (1 - \rho(TL, x) - \rho(BR, x))P_m(T)P_m(L),$$

$$P(TR|x) = (1 - \rho(TL, x) - \rho(BR, x))P_m(T)P_m(R),$$

$$P(BL|x) = (1 - \rho(TL, x) - \rho(BR, x))P_m(B)P_m(L).$$

These are three equations that identify the two unknown variables $\rho(TL, x)$ and $\rho(BR, x)$. Similarly, for values of x such that (BL, TR, mix) is realized,

$$P(BL|x) = \rho(BL, x) + (1 - \rho(BL, x) - \rho(TR, x))P_m(B)P_m(L),$$

$$P(BR|x) = (1 - \rho(BL, x) - \rho(TR, x))P_m(B)P_m(R),$$

$$P(TL|x) = (1 - \rho(BL, x) - \rho(TR, x))P_m(T)P_m(L)$$

are the three equations that overidentify the two unknown variables $\rho(BL, x)$ and $\rho(TR, x)$.

In the presence of the unobservable error terms ϵ, additional identification assumptions need to be imposed to isolate the effects of the error terms.

ASSUMPTION 11: *The equilibrium selection probabilities depend only on the utility indexes,*

$$\rho(x, \epsilon) = \rho\big(u_i((2 - i)Bk + (i - 1)jR, x) \; \forall i, j, k\big),$$

where $\rho(x, \epsilon) = [\rho(TL; x, \epsilon), \rho(BR; x, \epsilon), \rho(BL; x, \epsilon), \rho(TR; x, \epsilon)]$. In addition, the equilibrium selection probabilities are scale invariant with respect to the utility indexes. For all $\alpha > 0$, $\rho(\alpha u_i((2 - i)Bk + (i - 1)jR, x) \; \forall i, j, k) = \rho(u_i((2 - i)Bk + (i - 1)jR, x) \; \forall i, j, k).$

This assumption rules out the possibility that $\rho(x, \epsilon)$ might depend on x and ϵ nonseparably, independent of the latent utility indexes. It also requires that the equilibrum selection probabilities only depend on the relative but not absolute scales of the latent utilities.

The scale invariance assumption, supplemented by the next support condition on the observables and unobservables, allows us to identify the equilib-

rium selection probabilities from the variations in the covariates x. In particular, Assumption 11 implies that the determinants for the equilibrium selection probabilities are the same as the determinants for the mixing probabilities. It allows for a rich class of equilibrium selection mechanisms, but does exclude some important ones. For example, it allows for the Pareto efficient equilibrium to be selected with a larger probability and for this probability to depend on the relative efficiency level. This restriction follows from the intuition that if all payoffs were scaled by a constant, we would not expect the distribution over outcomes to change. However, it does not allow this probability to depend on how much more efficient the efficient equilibrium is compared to the inefficient ones in absolute terms. It also rules out equilibrium selection probabilities that depend independently on some of the observed covariates but not on other observed covariates or the error terms. This potentially limits the antitrust implications of the model, because firms concerned with avoiding the suspicions of antitrust investigators might want to choose selection rules which depend on some variables that are easier to communicate but not others.

ASSUMPTION 12: There exists a set $\mathcal{T}$ such that for all $\epsilon > 0$,

$$(20) \qquad \lim_{|x| \to \infty, x \in \mathcal{T}} P\left(\frac{f_i((2-i)Bk + (i-1)jR, x)}{u_i((2-i)Bk + (i-1)jR, x, \epsilon)} > 1 - \eta \right) = 1$$

for all $i = 1, 2$, $j = T, B$, and $k = L, R$, and that for all $\Lambda = R, B, T, L$,

$$\lim_{|x| \to \infty, x \in \mathcal{T}} P\left(\frac{P_m(\Lambda; x, \epsilon)}{P_m(\Lambda; x, 0)} > 1 - \epsilon \right) = 1.$$

This assumption is satisfied if ϵ has finite support but x has infinite support.

THEOREM 5: *Under Assumptions 1–12, the equilibrium selection probabilities*

$$\rho\big(u_i((2-i)Bk + (i-1)jR, x) \ \forall i, j, k\big)$$

are all identified from the observed choice probabilities.

PROOF: Assumptions 1–10 identify the payoff functions $f_i((2-i)Bk + (i-1)jR, x)$ for all i, j, k. Using Assumption 12, we can approximate the mixing probabilities with arbitrary precision by using larger and larger values of the covariates x. This allows us to recover the equilibrium selection probabilities with arbitrary precision at very large values of the covariates x. By Assumption 11, the equilibrium selection probabilities with smaller values of the latent utility indexes are obtained by extrapolation along the remote sections of a ray that emanates from the origin and goes through the latent utility indexes.   *Q.E.D.*

B.2. *Monte Carlo*

To demonstrate the performance of our estimator in small samples, we conduct a Monte Carlo experiment using a simple entry game with two players.

Each player has the profit function

$$\pi_i(a) = 1(a_i = 1)\{\beta_1 x_{i1} + \beta_2 x_{2i} + \epsilon_i(a)\},$$

with the observable covariates defined by $x_{i1} \sim N(1, 1)$ and $x_{2i} = n(a)$, where $N(\mu, \sigma^2)$ is the normal distribution with mean μ and variance σ^2, and $n(a)$ is the number of competitors a firm faces given action profile a. The idiosyncratic error term, which is different for each player for each action profile a, is drawn independently from the standard normal distribution. The choice of unit variance in the random shock satisfies the need for a scale normalization, and assigning payoffs of zero to not entering the market satisfies the location normalization. x_{1i} represents variability in profits to firm i from entering that market, and x_{2i} captures the effects of having a competitor. The true payoff parameters are $\beta_1 = 2$ and $\beta_2 = -10$.

The distributions of the covariates were chosen such that when payoffs are evaluated at their means, it is optimal for only one of the two firms to enter the market. Under these circumstances, the set of equilibria in this game, denoted by $\mathcal{E}$, has three elements: two pure strategies characterized by one firm or the other entering the market, and one mixed strategy where firms enter with some probability. We specify that the probability of equilibrium $\pi_i \in \mathcal{E}$ being played is

$$\Pr(\pi_i) = \frac{\exp(\theta_1 \text{MIXED}_i)}{\sum_{\pi_j \in \mathcal{E}} \exp(\theta_1 \text{MIXED}_j)},$$

where MIXED_i is an indicator variable equal to 1 if equilibrium π_i is in mixed strategies. When $\theta_1 = 0$, one of the three equilibria is picked with equal chance. As that parameter tends to either negative or positive infinity, the mixed strategy is played with probability approaching 0 or 1, respectively. The true selection parameter is $\theta_1 = 1$.

Our game has three unknown parameters: β_1, β_2, and θ_1. The game generates moments from the probabilities of observing the four possible combinations of entry choices. Only three of these moments are linearly independent, as the probabilities must sum to 1, implying that our model is exactly identified. We generate 500 samples of size $n = 25, 50, 100, 200$, and 400 to assess the finite sample properties of our estimator. We set the number of importance games per observation to be equal to the sample size, and generated new importance games for each observation and each replication. Asymptotic errors for each run are calculated using the optimal weighting matrix from a two-step generalized method of moments (GMM) procedure. The Laplace-type estimator of Chernozhukov and Hong (2003) is used to recover the parameters. The results of our Monte Carlo are reported in Table IX.

The results are encouraging even in the smallest sample sizes. The payoff parameters are tightly estimated near their true values, while the mixed strategy

TABLE IX

MONTE CARLO RESULTS

	Mean	Median	Std. Dev.	Mean Bias	Median Bias	Mean AD	Median AD	Mean ASE	Median ASE
$N = 25$									
β_1	1.982	2.017	0.491	−0.018	0.017	0.346	0.218	0.624	0.595
β_2	−9.952	−10	0.916	0.048	0	0.643	0.399	2.22	1.249
θ_1	−1.743	−0.977	3.959	−2.743	−1.977	3.421	2.398	2.644	2.346
$N = 50$									
β_1	2.048	2.039	0.401	0.048	0.039	0.265	0.113	0.588	0.574
β_2	−9.962	−10	0.787	0.038	0	0.534	0.28	1.232	1.148
θ_1	−0.62	0.526	3.395	−1.62	−0.474	2.04	0.588	2.545	2.124
$N = 100$									
β_1	2.057	2.036	0.327	0.057	0.036	0.221	0.095	0.544	0.536
β_2	−10.016	−10.012	0.663	−0.016	−0.012	0.442	0.213	1.086	1.071
θ_1	0.14	0.611	2.297	−0.86	−0.389	1.122	0.498	2.218	1.882
$N = 200$									
β_1	2.049	2.037	0.245	0.049	0.037	0.166	0.063	0.452	0.434
β_2	−9.981	−10.003	0.543	0.019	−0.003	0.347	0.114	0.915	0.887
θ_1	0.664	0.779	0.598	−0.336	−0.221	0.461	0.303	1.548	1.187
$N = 400$									
β_1	2.044	2.041	0.224	0.044	0.041	0.166	0.105	0.274	0.261
β_2	−10.013	−10.026	0.496	−0.013	−0.026	0.362	0.231	0.579	0.566
θ_1	0.732	0.756	0.377	−0.268	−0.244	0.356	0.281	0.275	0.257

The true parameter vector is $\beta_1 = 2$, $\beta_2 = -10$, and $\theta_1 = 1$. Each sample size was evaluated 500 times. AD = Absolute Deviations, ASE = Asymptotic Standard Errors.

shifter is estimated with considerably lower precision. There is a distinct downward bias in the estimates of the equilibrium selection parameter that shrinks as the sample size grows. The median bias in all parameters is much better than the mean bias, implying that the mean bias is largely driven by occasional extreme outliers.

The standard deviation of the estimates of all three parameters shrinks as the sample size increases, as do the mean and median absolute deviations. Significantly, the decrease in the standard deviation for the payoff parameters is close to $\sqrt{n}$, as theory implies. The rate of convergence of the equilibrium selection parameter is much more dramatic as the sample size increases, largely because this parameter is not precisely estimated at smaller sample sizes.

We include the means and medians of the asymptotic standard errors as well as the standard deviations of the Monte Carlo results. The asymptotic errors are usually comparable to the Monte Carlo standard errors, which are calculated by looking at the variance of the parameter estimates across the 500 replications for each sample size. The asymptotic errors tend to overstate the variance of the payoff parameters and understate the variance of the selection

parameter relative to the Monte Carlo errors. Their magnitude decreases at an increasing rate as the sample size grows, also similar to the Monte Carlo errors. Overall, our results suggest that the asymptotic errors are good small-sample approximations to their Monte Carlo counterparts.

The precision of the estimated payoff coefficients relative to the equilibrium selection parameter follows from the intuition that the payoff-relevant covariates define the thresholds at which firms are willing to enter a market, and thus enter the likelihood of every observation directly. On the other hand, the equilibrium selection parameter enters the estimating moments in a more subtle manner. This parameter is identified using coordination failures between firms due to mixed strategy equilibrium.

To illustrate, suppose that all payoffs, including idiosyncratic shocks, are observed by the econometrician. For some realizations of the covariates, the model will predict two pure strategies, with one or the other of the firms entering the market, and a single mixed strategy. If the mixed strategy equilibrium is played, there is a chance of either no firms entering the market or both firms entering the market. It is only when these mistakes are observed that the econometrician is certain that the mixed strategy is played. Behind this is a subtle and complex relationship between variables, as the probability of observing a mistake is a function of both θ_1, which controls how often a mixed strategy occurs, and the payoffs of the game, which determine the probability of observing a mistake conditional on playing a mixed strategy.

This interplay illustrates a more general point, which is that although the parameters are identified, in small samples the estimation of some parameters may depend on a relatively small subset of outcomes. Note well that this is true even in the extreme case when the payoff functions, including the idiosyncratic shocks, are known with certainty, since the model itself generates probabilistic outcomes through both the equilibrium selection mechanism and the random nature of mixed strategies. In light of this, the results here are very positive, as we are able to recover estimates of the true parameters with acceptable precision in moderate sample sizes.

There is one caveat to our procedure that researchers have to address in practice. In each Monte Carlo simulation, we know the true parameters of the game and we are able to generate importance games using them. With real data, of course, them parameters are initially unknown. The importance sampler can generate imprecise parameter estimates with poor initial guesses, so it is necessary to derive starting parameters from a separate source. Below we use a related game of private information to generate initial starting values. Parametric identification is another difficult empirical issue. In the Monte Carlo example, we are able to calculate the rank condition explicitly at the true parameter values and find it to be nonsingular. This is an overly strong condition for local identification, but is not sufficient for global identification, which is difficult to obtain.

B.3. *Rank Conditions in the Monte Carlo Example*

As an illustration, we explicitly analyze the rank condition for identification, which requires that the Jacobian matrix is invertible everywhere, in the context of the Monte Carlo simulation example. The Jacobian matrix is formed by taking derivatives of the outcome probabilities with respect to the parameters:

$$
A = \begin{bmatrix}
\dfrac{\partial P(1,1|x,\beta_1,\beta_1,\theta_1)}{\partial \beta_1} \\[2mm]
\dfrac{\partial P(1,0|x,\beta_1,\beta_1,\theta_1)}{\partial \beta_1} \\[2mm]
\dfrac{\partial P(0,0|x,\beta_1,\beta_1,\theta_1)}{\partial \beta_1}
\end{bmatrix}
$$

$$
\begin{bmatrix}
\dfrac{\partial P(1,1|x,\beta_1,\beta_1,\theta_1)}{\partial \beta_2} & \dfrac{\partial P(1,1|x,\beta_1,\beta_1,\theta_1)}{\partial \theta_1} \\[2mm]
\dfrac{\partial P(1,0|x,\beta_1,\beta_1,\theta_1)}{\partial \beta_2} & \dfrac{\partial P(1,0|x,\beta_1,\beta_1,\theta_1)}{\partial \theta_1} \\[2mm]
\dfrac{\partial P(0,1|x,\beta_1,\beta_1,\theta_1)}{\partial \beta_2} & \dfrac{\partial P(1,1|x,\beta_1,\beta_1,\theta_1)}{\partial \theta_1}
\end{bmatrix}.
$$

Despite the simple structure of the Monte Carlo setup, the observed outcome probabilities,

$$
P(1,1|x,\beta_1,\beta_1,\theta_1), \quad P(1,0|x,\beta_1,\beta_1,\theta_1), \quad \text{and}
$$
$$
P(0,0|x,\beta_1,\beta_1,\theta_1),
$$

are highly complex functions of the model parameters, and the calculation of A is nontrivial.[14] We investigate the nonsingularity of the Jacobian matrix A. Because this is a conditional model, for each vector of parameters $(\beta_1, \beta_2, \theta_1)$ we look for a covariate x that gives a nonsingular Jacobian matrix A. Due to the computation cost of numerically evaluating an integral in the Jacobian matrix, it is prohibitively time consuming to verify nonsingularity for an exhaustive search over the parameter space. Therefore, we take a grid of size 6 with interval length 0.3 symmetrically around each of the true parameter values. For each combination of the parameter values in the grid, we evaluate the smallest absolute value of the eigenvalues of the Jacobian A at 25 values of the covariates x. The maximum over the covariates x of the smallest absolute eigenvalue

---

[14]We derive these relationships in supplementary material which is available upon request. Calculating the numerical values of A also requires numerically integrating a one-dimensional integral, which we evaluate by simulation. Since the derivatives are expressed analytically, we do not need to choose a step size to compute the numerical derivatives.
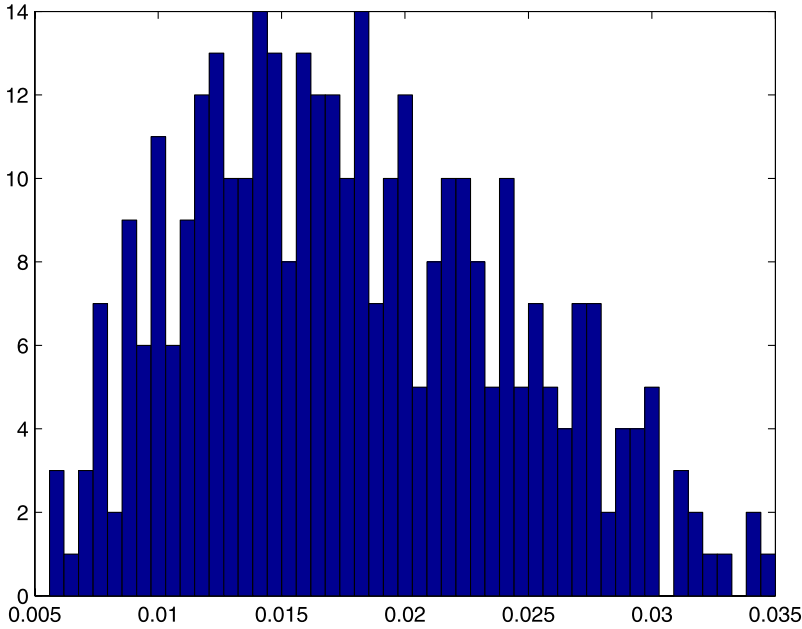
FIGURE 1.—Distribution of absolute eigenvalues in the Monte Carlo example.

of the Jacobian matrix A is recorded and plotted in Figure 1. The histogram of the minimum absolute eigenvalues suggests that they are bounded away from zero and, as a result, that the rank condition is satisfied. While this is not definitive analytical proof that the Jacobian is invertible everywhere (this requires demonstrating strong conditions on the Jacobian, for example that it is a P-matrix, which is infeasible given the complexity of our system), as formally required for identification, it does suggest that our problem is identified in a neighborhood of the true parameters.

B.4. *Maximal Number of Nash Equilibria*

When considering the nonparametric identification of the equilibrium selection mechanism, knowledge of generic properties of the set of Nash equilibria proves useful. Here we briefly review results in the literature on the maximum number of equilibria to normal form games of the class considered here.

Solutions to normal form games can be characterized using polynomial equalities and inequalities. Therefore, before considering games, we review some important results on the solutions of a system polynomials. Let $F = \{f_i(x)\}_{i=1}^n$ be the system of n polynomials of n variables. We are looking for the set of all common roots of this system. A polynomial $f_i(x) = \sum_{j=1}^J a_j x_1^{e_1^{ij}} x_2^{e_2^{ij}} \cdots x_n^{e_n^{ij}} = \sum_{j=1}^J a_j \prod_{k=1}^n x_k^{e_k^{ij}}$. The powers e_k^{ij} are integers in gen-

eral: index i refers to the equation number, index j refers to the number of monomials in polynomial i, and index k refers to the specific variable x_k. The points $e^{ij} = (e_1^{ij}, \ldots, e_n^{ij})$ form the finite sets $E_i = (e^{ij}, j = 1, \ldots, J)$ and indicate which monomial terms appear in f_i. For instance, in the polynomial $f_i(x_1, x_2) = x_1^2 x_2^3 + 2x_1^2$, the support set is $E_i = \{(2, 3), (2, 0)\}$.

The collection of sets $E = (E_1, E_2, \ldots, E_n)$ is called the support of the system of polynomials. The convex hulls $\mathrm{Conv}(E_i)$ are called Newton polytopes of f_i. For example, the Newton polytope of the polynomial $f(x_1, x_2) = x_1 x_2 + x_1 + x_2 + 1$ is the unit square with vertices in $(1, 1)$, $(1, 0)$, $(0, 1)$, and $(0, 0)$.

The degree of the polynomial i is $d_i = \max_j \sum_{k=1}^{n} e_k^{ij}$. One of the most important theorems describing the behavior of zeros of F in the complex space $\mathbb{C}^n$ is Bézout's theorem, which says that the total number of common complex roots of F is at most $\prod_{i=1}^{n} d_i$. Bézout's theorem provides an upper bound to the number of common roots in the system, giving little information on the polynomials that are sparse. In fact, for sparse systems, the number of common roots of the polynomial system can be significantly less than the bound set by this theorem. A universal and powerful tool for root counting in case of sparse polynomials is Bernstein's theorem.

Let P_i be Newton polytopes of equations $f_i(x)$ in the system F defined previously. The mixed volume of the system of polytopes is defined as

$$(21) \qquad \mathcal{M}(P_1, \ldots, P_n) = \sum_{S \subseteq \{1, \ldots, n\}} (-1)^{|S|} \mathrm{Vol}\left(\sum_{i \in S} P_i\right),$$

where S are all subsets of $\{1, 2, \ldots, n\}$, $|S|$ is the cardinality (number of elements) of a particular subset, and $\mathrm{Vol}(\cdot)$ is the conventional geometric volume. The sum of the polytopes is defined for two polytopes P and Q as $P + Q = \{p + q \mid p \in P, q \in Q\}$.

THEOREM 6—(Bernstein): *The number of common roots in the system F is equal to the mixed volume of the n Newton polytopes of this system.*

This is an extremely powerful result because the mixed volume is easy to compute. A general problem with complex roots though is that they are not invariant with respect to the group of polynomial transformations of F. For example, if the polynomial $f(x)$ has degree d and thus has d distinct complex roots, then the polynomial $f(x)^2$ can have $2d$ distinct complex roots. This is not the case with the real roots of a system of polynomials and thus power transformations have no effect on the number of distinct real roots. This is captured by Khovanskii's theorem, which sets the upper bound on the number of common real roots of the polynomial system which does not depend on the degrees of the equations in the system.

THEOREM 7—(Khovanskii): *If m is the number of all monomials in F (or equivalently $m = |E| = \sum_{i=1}^{n} J - i$—cardinality of the support) and there are n polynomials in F, then the maximum number of real solutions of the system is $2^{\binom{m}{2}}(n+1)^m$.*

In many cases, however, the so-called Kouchnirenko conjecture holds: if the number of terms in f_i is at most m_i, then the number of isolated real roots is at most $\prod_{i=1}^{n}(m_i - 1)$. This conjecture is violated for some generic (although quite complex) counterexamples.

Consider an arbitrary N-person game where player i has n_i strategies. Using Lagrangian multiplier techniques, it can be reduced to a system of $n + \sum_i n_i$ polynomial equations with $n + \sum_i n_i$ unknowns. Let $x_k^{(i)}$ denote strategy k of player i, let $\xi^{(i)}_{j_1, j_2, \ldots, j_{i-1}, k, j_{i+1}, \ldots, j_N}$ be the payoff function, representing the payoff of player i when she plays the pure strategy k and the other players are playing $j_1, \ldots, j_N$, and let $\pi^{(i)}$ be the expected payoff of player i. Let $\lambda_{k0}^{(i)}$ be the Lagrange multiplier for the constraint $x_k^{(i)} \geq 0$ and let $\widetilde{\lambda}^{(i)}$ be the Lagrange multiplier for the constraint $\sum_{k=1}^{n_i} x_k^{(i)} = 1$. The Lagrangian for bidder i can be written as

$$\mathcal{L}^{(i)} = \sum_{k=1}^{n_i} x_k^{(i)} \sum_{j_{-i}} \xi^{(i)}_{j_1, j_2, \ldots, j_{i-1}, k, j_{i+1}, \ldots, j_N} x_{j_1}^{(1)} \cdots x_{j_{i-1}}^{(i-1)} x_{j_{i+1}}^{(i+1)} \cdots x_{j_N}^{(N)}$$
$$- \sum_{k=1}^{n_i} \lambda_{k0}^{(i)} x_k^{(i)} + \widetilde{\lambda}^{(i)} \left(1 - \sum_{k=1}^{n_i} x_k^{(i)} \right).$$

The first order condition for the Lagrangian and the complementary slackness conditions for the multipliers $\lambda_{k0}^{(i)}$ are

$$(22) \qquad \frac{\partial \mathcal{L}^{(i)}}{\partial x_k^{(i)}} = \sum_{j_{-i}} \xi^{(i)}_{j_1, j_2, \ldots, j_{i-1}, k, j_{i+1}, \ldots, j_N} x_{j_1}^{(1)} \cdots x_{j_{i-1}}^{(i-1)} x_{j_{i+1}}^{(i+1)} \cdots x_{j_N}^{(N)} - \lambda_{k0}^{(i)} - \lambda^{(i)}$$
$$= 0,$$

$$(23) \qquad 1 - \sum_{k=1}^{n_i} x_k^{(i)} = 0,$$

$$(24) \qquad x_k^{(i)} \lambda_{k0}^{(i)} = 0, \quad k = 1, \ldots, n_i.$$

If we multiply the first equation by x_k^i, it reduces to

$$(25) \qquad x_k^{(i)} \sum_{j_1, j_2, \ldots, j_{i-1}, j_{i+1}, \ldots, j_N} \xi^{(i)}_{j_1, j_2, \ldots, j_{i-1}, k, j_{i+1}, \ldots, j_N}$$
$$\times x_{j_1}^{(1)} \cdots x_{j_{i-1}}^{(i-1)} x_{j_{i+1}}^{(i+1)} \cdots x_{j_N}^{(N)} - x_k^{(i)} \widetilde{\lambda}^{(i)} = 0,$$

$$(26) \qquad 1 - \sum_{k=1}^{n_i} x_k^{(i)} = 0,$$

$$(27) \qquad x_k^{(i)} \lambda_{k0}^{(i)} = 0, \quad k = 1, \ldots, n_i.$$

Summing the first equation over all k, we obtain $\pi^{(i)} = \widetilde{\lambda}^{(i)}$. Then each bidder is characterized by the system of equations

$$(28) \qquad x_k^{(i)} \left(\pi^i - \sum_{j_1, j_2, \ldots, j_{i-1}, j_{i+1}, \ldots, j_N} \xi_{j_1, j_2, \ldots, j_{i-1}, k, j_{i+1}, \ldots, j_N}^{(i)} \right.$$

$$\left. \times x_{j_1}^{(1)} \cdots x_{j_{i-1}}^{(i-1)} x_{j_{i+1}}^{(i+1)} \cdots x_{j_N}^{(N)} \right) = 0,$$

$$(29) \qquad \sum_{k=1}^{n_i} x_k^{(i)} - 1 = 0, \quad k = 1, \ldots, n_i, \quad i = 1, \ldots, N.$$

Thus, for each player we have $n_i + 1$ equations and $n_i + 1$ unknown parameters (n_i mixed strategies and the expected payoff). The individual equation has $\prod_{j \neq i} n_j + 1$ terms (the number of strategies of the other players when the strategy of player i is fixed, plus the expected payoff of player i). In addition, the linear equations limiting the mixed strategies to the simplex have $n_i + 1$ terms each. So in total there are $\sum_i n_i + N$ equations and unknowns. The total number of terms is $\sum_i \prod_{j \neq i} n_j + \sum_i n_i + 2N$. If we consider purely mixed strategies, then $x_k^{(i)} > 0$, and thus the system can be rewritten as

$$(30) \qquad \sum_{j_{-i}} \left(\xi_{j_1, j_2, \ldots, j_{i-1}, k, j_{i+1}, \ldots, j_N}^{(i)} - \xi_{j_1, j_2, \ldots, j_{i-1}, n_i, j_{i+1}, \ldots, j_N}^{(i)} \right)$$

$$\times x_{j_1}^{(1)} \cdots x_{j_{i-1}}^{(i-1)} x_{j_{i+1}}^{(i+1)} \cdots x_{j_N}^{(N)} = 0,$$

$$(31) \qquad k = 1, \ldots, n_i - 1, \quad i = 1, \ldots, N.$$

This system has $n_i - 1$ unknowns for player i and $\sum_i n_i - N$ unknowns in total. The number of terms for each equation is $\prod_{j \neq i} n_j$, according to the number of strategies of the rival players when the strategy of the given player is fixed. The total number of terms is then given by the sum $\sum_i (\prod_{j \neq i} n_j)(n_i - 1)$.

McKelvey and McLennan (1996) directly applied Bernstein's theorem to the given system of equations and expressed the number of solutions in terms of the mixed volume of Newton polytopes for the case of totally mixed solutions (the case with possible pure strategies needs specific consideration for each payoff structure). We can also apply Khovanskii's result to this system as fol-

TABLE X

TABULATION OF KOUCHNIRENKO'S FORMULA

	Number of Strategies (k)				
N	2	3	4	5	6
2	3	16	729	65536	9765625
3	27	262,144	38,443,359,375	3.65203E+16	1.44884E+23
4	2401	2.08827E+11	3.90919E+21	3.12426E+33	4.45419E+46
5	759,375	1.07374E+19	1.25344E+36	8.01109E+55	6.40832E+77
6	887,503,681	4.03445E+28	1.50578E+54	7.4656E+83	5.2603E+116

lows. First, by Kouchnirenko's conjecture, the maximum number of solutions to this system is

$$(32) \qquad \prod_{i=1}^{N}\left(\prod_{j\neq i} n_j - 1\right)^{n_i-1},$$

which gives an approximate upper bound on the number of solutions. An exact application of Khovanskii's formula with $m = \sum_i (\prod_{j\neq i} n_j)(n_i - 1)$ gives the maximum number of solutions as $2^{m!/(2(m-2)!)}(\sum_i n_i - N + 1)^m$.

For a particular case when k equals the number of strategies, Kouchnirenko's formula gives $(k^{N-1} - 1)^{N(k-1)}$ for the number of equilibria, while Khovanskii's bound is

$$2^{\binom{Nk^{(N-1)}(k-1)}{2}}[N(k-1)+1]^{Nk^{(N-1)}(k-1)}.$$

The number of moments with N players when each player has k strategies is $k^N - 1$. The corresponding numbers of moments are tabulated in Table XI. They are significantly smaller than Kouchnirenko's bounds.

Dept. of Economics, University of Minnesota, 4-127 Hanson Hall, 1925 Fourth Street South, Minneapolis, MN 55455, U.S.A. and NBER; bajari@econ.umn.edu,

TABLE XI

TABULATION OF THE NUMBER OF AVAILABLE MOMENTS

	Number of Strategies (k)				
N	2	3	4	5	6
2	3	8	15	24	35
3	7	26	63	124	215
4	15	80	255	624	1295
5	31	242	1023	3124	7775
6	63	728	4095	15,624	46,655

Dept. of Economics, Stanford University, 579 Serra Mall, Stanford, CA 94305, U.S.A.; hanhong@stanford.edu,

and

Dept. of Economics, Massachusetts Institute of Technology, 50 Memorial Drive, E52-391, Cambridge, MA 02142, U.S.A.; sryan@mit.edu.