Supplement to "An empirical model of non-equilibrium behavior in games"

(*Quantitative Economics*, Vol. 9, No. 1, March 2018, 141–181)

Brendan Kline
Department of Economics, University of Texas at Austin

Appendix A: Model likelihood

Use the notation that y is the entire data set, y_i is the data of subject i, and y_{ig} is the data of subject i in game g. Also, $\tau(i)$ is the strategic behavior rule used by subject i. Suppose that γ_{ig} is the intended decision rule for subject i in game g. Neither $\tau(i)$ nor γ_{ig} is observed by the econometrician. Then the likelihood for observing subjects $i = 1, 2, \dots, N$ to take actions in games $g = 1, 2, \dots, G$ is

$$\begin{aligned}
\log L(y|\theta) &= \sum_{i=1}^{N} \log L(y_i|\theta) \\
&= \sum_{i=1}^{N} \log\left(\sum_{r=1}^{R} P\big(y_i|\tau(i)=r, \theta\big)P\big(\tau(i)=r|\theta\big)\right) \\
&= \sum_{i=1}^{N} \log\left(\sum_{r=1}^{R}\left(\prod_{g=1}^{G} P\big(y_{ig}|\tau(i)=r, \theta\big)\right)\pi(r)\right) \\
&= \sum_{i=1}^{N} \log\left(\sum_{r=1}^{R}\left(\prod_{g=1}^{G}\left(\sum_{k} P\big(y_{ig}|\tau(i)=r, \gamma_{ig}=k, \theta\big)P\big(\gamma_{ig}=k|\tau(i)=r, \theta\big)\right)\right)\pi(r)\right) \\
&= \sum_{i=1}^{N} \log\left(\sum_{r=1}^{R}\left(\prod_{g=1}^{G}\left(\sum_{k} P\big(y_{ig}|\tau(i)=r, \gamma_{ig}=k, \theta\big)\Lambda_r(k)\right)\right)\pi(r)\right),
\end{aligned}$$

where θ collects all of the parameters of the model. The sum over k corresponds to the sum over the decision rules that subjects might use, per Assumption 2.1. It remains to derive the form of $P(y_{ig}|\tau(i)=r, \gamma_{ig}=k, \theta)$ from the model specification.

For $k = s_{\text{unanch}}$, for some $s \in \mathcal{U}$,

$$P\big(y_{ig} \le t|\tau(i)=r, \gamma_{ig}=s_{\text{unanch}}, \theta\big) = F_{gs_{\text{unanch}}}(t),$$

Brendan Kline: brendan.kline@austin.utexas.edu

https://doi.org/10.3982/QE647

where $F_{gs_{\text{unanch}}}(\cdot)$ is the cumulative distribution function of a random variable with density $\zeta^s_{1g}(\cdot)$ with respect to the appropriate dominating measure on Σ^s_{1g}, per Section 2.2.2.

For $k \in \mathcal{M}$, and letting m_{ig} be a binary variable to indicate whether subject i incorrectly computes the action associated with decision rule k in game g, which is not observed by the econometrician,

$$\begin{aligned} P\big(y_{ig} \leq t \mid \tau(i) = r, \gamma_{ig} = k, \theta\big) &= P\big(y_{ig} \leq t \mid \tau(i) = r, \gamma_{ig} = k, m_{ig} = 1, \theta\big) \\ &\quad \times P\big(m_{ig} = 1 \mid \tau(i) = r, \gamma_{ig} = k, \theta\big) \\ &\quad + P\big(y_{ig} \leq t \mid \tau(i) = r, \gamma_{ig} = k, m_{ig} = 0, \theta\big) \\ &\quad \times P\big(m_{ig} = 0 \mid \tau(i) = r, \gamma_{ig} = k, \theta\big) \\ &= F_{rgk}(t)\Delta_r + 1\big[t \geq c_{1g}(k)\big](1 - \Delta_r), \end{aligned}$$

where $F_{rgk}(\cdot)$ is the cumulative distribution function of computational mistakes on $[\alpha_{Lg}(1), \alpha_{Ug}(1)] \cap [c_{1g}(k) - \mathrm{P}_r(\alpha_{Ug}(1) - \alpha_{Lg}(1)), c_{1g}(k) + \mathrm{P}_r(\alpha_{Ug}(1) - \alpha_{Lg}(1))]$, per Section 2.3.

Appendix B: Sufficient conditions for point identification except for the magnitude of computational mistakes

This section establishes sufficient conditions for point identification of all unknown parameters except for those related to the *magnitude* of computational mistakes, under weaker conditions than used by Theorem 4.1. The result does still allow that individuals might make computational mistakes. This can be interpreted as a *partial identification* result, showing that some but not necessarily all of the parameters are point identified. Alternatively, this can be interpreted as a *point identification* result, showing that a model without computational mistakes (or even a model with computational mistakes with *known* magnitudes of computational mistakes) is point identified.

The identification result in this section uses a different definition of observational equivalence of strategic behavior types. Essentially, the alternative definition treats the magnitude of computational mistakes as irrelevant and is similar to Definition 1, except the last condition involving P is dropped. There is a corresponding definition of point identification, which ignores the magnitude of computational mistakes.

Definition 3 (Observational Equivalence of Strategic Behavior Types, Ignoring the Magnitude of Computational Mistakes). The quantities $\Theta_1 = (\Lambda_1, \Delta_1, \mathrm{P}_1)$ and $\Theta_2 = (\Lambda_2, \Delta_2, \mathrm{P}_2)$ are observationally equivalent, ignoring the magnitude of computational mistakes, if

(i) it holds that $\Lambda_1 = \Lambda_2$,
(ii) it holds that $\Delta_1 1[\sum_{k \in \mathcal{M}} \Lambda_1(k) > 0] = \Delta_2 1[\sum_{k \in \mathcal{M}} \Lambda_2(k) > 0]$.

Definition 4 (Point Identification of Model Parameters, Ignoring the Magnitude of Computational Mistakes). The model parameters are point identified, ignoring the magnitude of computational mistakes, if for any specifications $\{\Theta_{0r}, \pi_0(r)\}_{r=1}^{\tilde{R}_0}$ and $\{\Theta_{1r}, \pi_1(r)\}_{r=1}^{\tilde{R}_1}$ of the model parameters that satisfy the assumptions and also are such that

(i) both specifications $\{\Theta_{0r}, \pi_0(r)\}_{r=1}^{\tilde{R}_0}$ and $\{\Theta_{1r}, \pi_1(r)\}_{r=1}^{\tilde{R}_1}$ generate the observable data,

(ii) it holds that $\pi_0(\cdot) > 0$ and $\pi_1(\cdot) > 0$,

(iii) the strategic behavior rules Θ_{0r} and $\Theta_{0r'}$ are not observationally equivalent, ignoring the magnitude of computational mistakes for all $r \neq r'$, and Θ_{1r} and $\Theta_{1r'}$ are not observationally equivalent, ignoring the magnitude of computational mistakes for all $r \neq r'$,

then $\tilde{R}_0 = \tilde{R} = \tilde{R}_1$ and there is a permutation ϕ of $\{1, 2, \ldots, \tilde{R}\}$ such that for each $r = 1, 2, \ldots, \tilde{R}$ it holds that $\pi_0(r) = \pi_1(\phi(r))$ and Θ_{0r} is observationally equivalent, ignoring the magnitude of computational mistakes, to $\Theta_{1\phi(r)}$.

The main difference between the sufficient conditions of this section and the sufficient conditions of Section 4 is that Assumption 4.1 is dropped in favor of the weaker Assumption B.1. Moreover, Assumption 4.2 is dropped entirely.

Assumption B.1 (Conditions on the Games). *The data set includes at least* $2R - 1$ *games, such that each game* g *of those* $2R - 1$ *games satisfies all of the following three conditions*:

(i) *It holds that* $\Omega_{1g} > 0$.

(ii) *For each* $k \in \mathcal{M}$ *and* $k' \in \mathcal{M}$ *such that* $k \neq k'$, $c_{1g}(k) \neq c_{1g}(k')$.

(iii) *For each* $k \in \mathcal{M}$ *and* $s \in \mathcal{U}$ *such that* Σ_{1g}^s *is a finite set,* $c_{1g}(k) \notin \Sigma_{1g}^s$.

The data set includes at least $2R - 1$ *games, such that each game* g *of those* $2R - 1$ *games satisfies the following condition*:

(iv) *For each* $s \in \mathcal{U}$, $R_{1g}(s, s, \overline{\rho}) > 0$.

Assumption 4.1 requires that the same games satisfy all of the conditions stated in Assumption 4.1, whereas Assumption B.1 allows that some games satisfy Conditions (i), (ii), and (iii), and other games satisfy Condition (iv). However, it is allowed that the set of games satisfying Conditions (i), (ii), and (iii) arbitrarily overlaps with the set of games satisfying Condition (iv).

The next assumption disallows certain "knife-edge" cases and requires additional notation. Use the notation that $\mathcal{M}(r)$ is the rth smallest element of $\mathcal{M}$, with Nash equilibrium the largest element by convention, and that $\mathcal{U}(r)$ is the rth smallest element of $\mathcal{U}$.

Assumption B.2 (No Knife-Edge Strategic Behavior Rules). *There are* $\tilde{R}$ *strategic behavior rules used in the population, with* $\pi(r) > 0$ *for* $r = 1, 2, \ldots, \tilde{R}$. *For each* $r' \neq r$,

(i) *it holds that* $((1-\Delta_r)\Lambda_r(\mathcal{M}(1)),\ldots,(1-\Delta_r)\Lambda_r(\mathcal{M}(|\mathcal{M}|)))\neq((1-\Delta_{r'})\Lambda_{r'}(\mathcal{M}(1)),\ldots,(1-\Delta_{r'})\Lambda_{r'}(\mathcal{M}(|\mathcal{M}|)))$ *and* $(\Lambda_r(\mathcal{U}(1)),\ldots,\Lambda_r(\mathcal{U}(|\mathcal{U}|)))\neq(\Lambda_{r'}(\mathcal{U}(1)),\ldots,\Lambda_{r'}(\mathcal{U}(|\mathcal{U}|)))$,

(ii) *it holds that* $\pi(r)\neq\pi(r')$.

Condition (i) rules out the knife-edge case that strategic behavior rules r and r', despite being distinct, are such that $((1-\Delta_r)\Lambda_r(\mathcal{M}(1)),\ldots,(1-\Delta_r)\Lambda_r(\mathcal{M}(|\mathcal{M}|)))=((1-\Delta_{r'})\Lambda_{r'}(\mathcal{M}(1)),\ldots,(1-\Delta_{r'})\Lambda_{r'}(\mathcal{M}(|\mathcal{M}|)))$ or $(\Lambda_r(\mathcal{U}(1)),\ldots,\Lambda_r(\mathcal{U}(|\mathcal{U}|)))=(\Lambda_{r'}(\mathcal{U}(1)),\ldots,\Lambda_{r'}(\mathcal{U}(|\mathcal{U}|)))$. Condition (ii) rules out the knife-edge case that two strategic behavior rules are used with the same probability.

Theorem B.1. *Under Assumptions* 2.1, 2.2, B.1, *and* B.2, *the parameters of the model are point identified in the sense of Definition* 4.

Appendix C: Identification of the selection rule on unanchored strategic reasoning

It is possible to point identify the selection rule on unanchored strategic reasoning (introduced in Section 2.2.2) under suitable restrictions on the class of admissible selection rules. Per the discussion in Section 2.7, the discussion focuses without loss of generality on player 1 in the game. Specifically, consider player 1 in game g and the problem of identifying the function $\psi_{1g}(\cdot)$ that characterizes the selection rule for player 1 in game g. As in the empirical application, suppose that Σ^s_{1g} is Lebesgue measurable with nonzero and finite measure, for all $s\in\mathcal{U}$. Suppose also that the class of admissible selection rules is such that the selection rule has the derivatives used in the following analysis. Based on $\psi_{1g}(\cdot)$, for any $s\in\mathcal{U}$, the selection rule from using s steps of unanchored strategic reasoning as player 1 in game g has the ordinary density $\frac{\psi_{1g}(\cdot)}{\Psi_{1g}(\Sigma^s_{1g})}$ on Σ^s_{1g}. Per similar arguments as used to establish Lemma D.4(iv), based on the set $U_{1g}(s)$ from condition (v) of Assumption 4.1, the observed ordinary density of the data at an action a taken within any interval subset of $U_{1g}(s)$ is $d_{sg}\psi_{1g}(a)$, where $d_{sg}=\sum_{r=1}^{R}\sum_{s'=0,s'\in\mathcal{U}}^{s}\frac{1}{\Psi_{1g}(\Sigma^{s'}_{1g})}\Lambda_r(s'_{\text{unanch}})\pi(r)$. Hence, if a positive fraction of subjects use a strategic behavior rule that uses s or fewer steps of unanchored strategic reasoning with positive probability, $\psi_{1g}(a)$ is identified up to positive (and unknown) scale $d_{sg}>0$, for all a in that interval subset of $U_{1g}(s)$ where the density exists. Under sufficient restrictions on the class of admissible $\psi_{1g}(\cdot)$, which are equivalent to restrictions on the class of admissible selection rules, this suffices to point identify the entire $\psi_{1g}(\cdot)$ function. Intuitively, these restrictions must be such that knowledge of the identified properties of $\psi_{1g}(\cdot)$ on a subset of the domain is enough to "extrapolate" to knowledge of the entire $\psi_{1g}(\cdot)$ function. Specifically, the following text discusses using the information contained in the identified quantity $\frac{\psi'_{1g}(\cdot)}{\psi_{1g}(\cdot)}$ on a subset of the domain.

For example, suppose that $\psi_{1g}(\cdot)$ is known by the econometrician to be the density of a normal distribution with unknown mean μ_{1g} and unknown variance $\sigma^2_{1g}>0$. As already discussed in Section 2.2.2, because that distribution is unimodal with mode μ_{1g}, the resulting selection rule tends to be biased toward actions around μ_{1g} and biased against actions away from μ_{1g}. The "degree" of that bias is measured by the variance σ^2_{1g}, where relatively large σ^2_{1g} result in relatively small biases, since relatively large σ^2_{1g} result in $\psi_{1g}(\cdot)$ that approaches a constant function. Since $\psi_{1g}(a)$ is identified up to positive (and unknown) scale for all a in that interval subset of $U_{1g}(s)$ where the density exists, also $\psi'_{1g}(a)$ is identified up to the same positive (and unknown) scale for all a in that same interval subset of $U_{1g}(s)$ where the density exists, based on any $s \in \mathcal{U}$ such that the corresponding $d_{sg} > 0$ as discussed above. It is a simple exercise to establish, based on the functional form of the normal distribution, that $\sigma^2_{1g} \frac{\psi'_{1g}(a)}{\psi_{1g}(a)} = \mu_{1g} - a$. Hence, based on the above identification of $\frac{\psi'_{1g}(a)}{\psi_{1g}(a)}$ at two distinct points $a = a^{(1)}$ and $a = a^{(2)}$, where the positive (and unknown) scale cancels in the ratio, it is possible to identify μ_{1g} and σ^2_{1g} by solving the system of equations $\sigma^2_{1g} \frac{\psi'_{1g}(a^{(1)})}{\psi_{1g}(a^{(1)})} = \mu_{1g} - a^{(1)}$ and $\sigma^2_{1g} \frac{\psi'_{1g}(a^{(2)})}{\psi_{1g}(a^{(2)})} = \mu_{1g} - a^{(2)}$ for the unknown μ_{1g} and σ^2_{1g}. Hence, under this restriction on the class of admissible $\psi_{1g}(\cdot)$, it is possible to identify the entire $\psi_{1g}(\cdot)$ function.

Appendix D: Proof of point identification

Use the notation that $\mathcal{M}(r)$ is the rth smallest element of $\mathcal{M}$ with Nash equilibrium the largest element by convention, $\mathcal{U}(r)$ is the rth smallest element of $\mathcal{U}$, $U_g(s) = U_{1g}(s)$, $R_g(s, s', \varepsilon) = R_{1g}(s, s', \varepsilon)$, and $\Omega_g = \alpha_{Ug}(1) - \alpha_{Lg}(1)$. Additionally,

$$M_g(k, \varepsilon, \mathrm{P}_r) = \begin{cases} \displaystyle\int_{c_{1g}(k)-\varepsilon\Omega_g}^{c_{1g}(k)+\varepsilon\Omega_g} \omega_{1g, c_{1g}(k), \mathrm{P}_r}(a)\, da & \text{if } \mathrm{P}_r > 0, \\ 1 & \text{if } \mathrm{P}_r = 0. \end{cases}$$

Let the set of nonzero unique values of $\{\mathrm{P}_r 1[\Delta_r > 0] 1[\sum_{k\in\mathcal{M}} \Lambda_r(k) > 0] \times 1[\pi(r) > 0]\}_{r=1}^R$ together with $\overline{\rho}$ be $\{\tilde{\mathrm{P}}_w\}_{w=1}^W$, and without loss of generality assume that $0 \le \tilde{\mathrm{P}}_1 < \tilde{\mathrm{P}}_2 < \cdots < \tilde{\mathrm{P}}_W$ and that $1 \le W \le R + 1$. By Assumption 2.2, $\tilde{\mathrm{P}}_W = \overline{\rho}$.

For any decision rule $k \in \mathcal{M}$, let $C_g(k, \varepsilon)$ be the event that a subject takes an action weakly within $\varepsilon\Omega_g$ of the action predicted by decision rule k in game g, but excluding the action exactly predicted by decision rule k in game g. For any decision rule $k \in \mathcal{M}$, let $C_g(k)$ be the event that a subject takes the action exactly predicted by decision rule k in game g. Note that $C_g(k) \neq C_g(k, 0)$.

Use the generic notation that P_θ refers to the distribution of observables based on strategic behavior rule θ and that $P_{g,\theta}$ refers to the distribution of observables based on strategic behavior rule θ in game g. By some abuse of notation, let $P_{g,\theta}$ be the $(|\mathcal{M}| + |\mathcal{U}| + W|\mathcal{M}|) \times 1$ vector:

(i) The first $|\mathcal{M}|$ rows are $(P_{g,\theta}(C_g(\mathcal{M}(1))), \ldots, P_{g,\theta}(C_g(\mathcal{M}(|\mathcal{M}|))))$.

(ii) The next $|\mathcal{U}|$ rows are $(P_{g,\theta}(U_g(\mathcal{U}(1))), \ldots, P_{g,\theta}(U_g(\mathcal{U}(|\mathcal{U}|))))$.

(iii) The final $W|\mathcal{M}|$ rows are $(P_{g,\theta}(C_g(\mathcal{M}(1), \tilde{\mathrm{P}}_1)), \ldots, P_{g,\theta}(C_g(\mathcal{M}(1), \tilde{\mathrm{P}}_W)), P_{g,\theta}(C_g(\mathcal{M}(2), \tilde{\mathrm{P}}_1)), \ldots)$.

Use the notation that $\bigotimes^n b = \underbrace{b \otimes b \otimes \cdots \otimes b}_{n \text{ times}}$, for $n \in \mathbb{N}$.

Lemma D.1. *The following claims are true*:

(i) *For a game g that satisfies Assumption* 4.1(i) *and for $\rho_i > 0$, the density $\omega_{jg,c,\rho_i}(a)$ has discontinuities at, and only at,* $\min\{\alpha_{Ug}(j), c + \rho_i(\alpha_{Ug}(j) - \alpha_{Lg}(j))\}$ *and* $\max\{\alpha_{Lg}(j), c - \rho_i(\alpha_{Ug}(j) - \alpha_{Lg}(j))\}$.

(ii) *For a game g that satisfies the conditions of Assumption* 4.2 *and for* $\mathrm{P}_r > 0$, *for any $k \in \mathcal{M}$ and $0 < \varepsilon < \overline{\rho}$, $M_g(k, \varepsilon, \mathrm{P}_r)$ has a kink at, and only at, $\varepsilon = \mathrm{P}_r$.*

(iii) *For a game g that satisfies conditions* (i) *and* (iv) *of Assumption* 4.1, *for any $k \in \mathcal{M}$, $M_g(k, \varepsilon, \mathrm{P}_r) = M_g(k, \mathrm{P}_r, \mathrm{P}_r)$ if $\varepsilon \geq \mathrm{P}_r$ and $M_g(k, \varepsilon_1, \mathrm{P}_r) < M_g(k, \varepsilon_2, \mathrm{P}_r)$ if $0 \leq \varepsilon_1 < \varepsilon_2 \leq \mathrm{P}_r$.*

Proof. Because the game g satisfies Assumption 4.1(i) and $\rho_i > 0$, the density $\omega_{jg,c,\rho_i}(a)$ does not involve dividing by zero and, therefore, is well defined.

(i) Because $\xi(\cdot)$ is continuous on $[-1, 1]$, discontinuities in $\omega_{jg,c,\rho_i}(a)$ can occur only at a such that the argument of $\xi(\cdot)$ in the definition of $\omega_{jg,c,\rho_i}(a)$ is either -1 or 1. Therefore, discontinuities can occur only at $a = \min\{\alpha_{Ug}(j), c + \rho_i(\alpha_{Ug}(j) - \alpha_{Lg}(j))\}$ and $a = \max\{\alpha_{Lg}(j), c - \rho_i(\alpha_{Ug}(j) - \alpha_{Lg}(j))\}$. Moreover, by assumption, ξ is bounded away from zero on $[-1, 1]$, but equals zero off $[-1, 1]$, and, therefore, indeed $\omega_{jg,c,\rho_i}(a)$ does have discontinuities at the claimed points.

(ii) By part (i), the integrand in $M_g(k, \varepsilon, \mathrm{P}_r) = \int_{c_{1g}(k)-\varepsilon\Omega_g}^{c_{1g}(k)+\varepsilon\Omega_g} \omega_{1g,c_{1g}(k),\mathrm{P}_r}(a)\,da$ has discontinuities at, and only at, $a = \min\{\alpha_{Ug}(1), c_{1g}(k) + \mathrm{P}_r(\alpha_{Ug}(1) - \alpha_{Lg}(1))\}$ and $a = \max\{\alpha_{Lg}(1), c_{1g}(k) - \mathrm{P}_r(\alpha_{Ug}(1) - \alpha_{Lg}(1))\}$. Because the game g satisfies the conditions of Assumption 4.2, $0 < \mathrm{P}_r\Omega_g < \overline{\rho}\Omega_g < \alpha_{Ug}(1) - c_{1g}(k)$ or $0 < \mathrm{P}_r\Omega_g < \overline{\rho}\Omega_g < c_{1g}(k) - \alpha_{Lg}(1)$ by Assumption 4.1(iv). Therefore, either $\min\{\alpha_{Ug}(1), c_{1g}(k) + \mathrm{P}_r(\alpha_{Ug}(1) - \alpha_{Lg}(1))\} = c_{1g}(k) + \mathrm{P}_r\Omega_g$ or $\max\{\alpha_{Lg}(1), c_{1g}(k) - \mathrm{P}_r(\alpha_{Ug}(1) - \alpha_{Lg}(1))\} = c_{1g}(k) - \mathrm{P}_r\Omega_g$. Therefore, $M_g(k, \varepsilon, \mathrm{P}_r)$ has a kink at $\varepsilon = \mathrm{P}_r$. Moreover, there can be no other kinks in $M_g(k, \varepsilon, \mathrm{P}_r)$ for any $k \in \mathcal{M}$ and $0 < \varepsilon < \overline{\rho}$, by Assumption 4.2. That follows because any other kink would be located at $\varepsilon = \frac{c_{1g}(k)-\alpha_{Lg}(1)}{\Omega_g}$ or $\varepsilon = \frac{\alpha_{Ug}(1)-c_{1g}(k)}{\Omega_g}$. But by Assumption 4.2 evaluated at $s = 0$, such ε would equal either 0 or 1 under Assumption 4.2(iii) or 4.2(iv), or would be weakly greater than $\overline{\rho}$ under Assumption 4.2(i). However, $0 < \varepsilon < \overline{\rho}$ and by Assumption 4.1(iv), $\overline{\rho} < 1$.

(iii) Note that $M_g(k, \varepsilon, \mathrm{P}_r) = \int_{c_{1g}(k)-\varepsilon\Omega_g}^{c_{1g}(k)+\varepsilon\Omega_g} \omega_{1g,c_{1g}(k),\mathrm{P}_r}(a)\,da$, where the integrand is 0 for $a > \min\{\alpha_{Ug}(1), c_{1g}(k) + \mathrm{P}_r\Omega_g\}$ and $a < \max\{\alpha_{Lg}(1), c_{1g}(k) - \mathrm{P}_r\Omega_g\}$. Therefore, $M_g(k, \varepsilon, \mathrm{P}_r) = \int_{\max\{c_{1g}(k)-\varepsilon\Omega_g,\max\{\alpha_{Lg}(1),c_{1g}(k)-\mathrm{P}_r\Omega_g\}\}}^{\min\{c_{1g}(k)+\varepsilon\Omega_g,\min\{\alpha_{Ug}(1),c_{1g}(k)+\mathrm{P}_r\Omega_g\}\}} \omega_{1g,c_{1g}(k),\mathrm{P}_r}(a)\,da$. Therefore, since the bounds of integration are $[\max\{\alpha_{Lg}(1), c_{1g}(k) - \mathrm{P}_r\Omega_g\}, \min\{\alpha_{Ug}(1), c_{1g}(k) + \mathrm{P}_r\Omega_g\}]$ for $\varepsilon \geq \mathrm{P}_r$, it follows that $M_g(k, \varepsilon, \mathrm{P}_r) = M_g(k, \mathrm{P}_r, \mathrm{P}_r)$ if $\varepsilon \geq \mathrm{P}_r$. Since the bounds of integration are $[\max\{\alpha_{Lg}(1), c_{1g}(k) - \varepsilon\Omega_g\}, \min\{\alpha_{Ug}(1), c_{1g}(k) + \varepsilon\Omega_g\}]$ for $\varepsilon \leq \mathrm{P}_r$, and the integrand is positive over that range for all $\varepsilon \leq \mathrm{P}_r$, and by Assumption 4.1(iv), either the lower bound equals $c_{1g}(k) - \varepsilon\Omega_g$ or the upper bound equals $c_{1g}(k) + \varepsilon\Omega_g$, which both depend nontrivially on ε by Assumption 4.1(i), it follows that $M_g(k, \varepsilon_1, \mathrm{P}_r) < M_g(k, \varepsilon_2, \mathrm{P}_r)$ if $0 \leq \varepsilon_1 < \varepsilon_2 \leq \mathrm{P}_r$. □

Lemma D.2. *Let $R \in \mathbb{N}$ and $m \in \mathbb{N}$ satisfy $m \geq R - 1$. Let $C(m, n) = \sum_{p=0}^{m} n^p$. Let $\gamma_{p,n}(\cdot) : \mathbb{R}^n \to \mathbb{R}^{n^p}$ be defined by $\gamma_{p,n}(z) = \bigotimes^p z$. Let $\Gamma_{m,n}(\cdot) : \mathbb{R}^n \to \mathbb{R}^{C(m,n)}$ be defined by $\Gamma_{m,n}(z) =$*

$(1, \gamma_{1,n}(z), \dots, \gamma_{m,n}(z))$. *Thus,* $\Gamma_{m,n}(z)$ *gives all monomials of the argument vector* z*, of order between* 0 *and* m*, in ascending order (i.e., the order* 0 *monomial in the first row, then order* 1 *monomials in the next rows, etc.). Suppose* $b_1, \dots, b_R \in \mathbb{R}^n$ *are distinct. Let* $B^* = (\Gamma_{m,n}(b_1)\ \Gamma_{m,n}(b_2)\ \cdots\ \Gamma_{m,n}(b_R)) \in \mathbb{R}^{C(m,n)\times R}$*. Then* B^* *has full column rank.*

Proof. The following argument establishes that since $b_k \neq b_l$ for $k \neq l$, there exists a $t \in \mathbb{R}^n$ such that $t'b_k \neq t'b_l$ for all $k \neq l$. Let $\mathcal{D}(t) = \{(k,l) : t'b_k = t'b_l, k \neq l\}$. Let $t_0 \in \mathbb{R}^n$. If $|\mathcal{D}(t_0)| = 0$, then the claim is established. Otherwise, for some k^* and l^* such that $k^* \neq l^*$, $t_0'b_{k^*} = t_0'b_{l^*}$. By slightly perturbing t_0 in the element of t_0 corresponding to the element where b_{k^*} and b_{l^*} are not equal (which must exist since $b_{k^*} \neq b_{l^*}$), there exists $t_1 \in \mathbb{R}^n$ such that $t_1'b_{k^*} \neq t_1'b_{l^*}$. If the perturbation is sufficiently small, then $t_1'(b_k - b_l) \approx t_0'(b_k - b_l)$ uniformly for all k and l. Therefore, for any (k,l) such that $t_0 b_k \neq t_0 b_l$, also $t_1 b_k \neq t_1 b_l$. Therefore, $|\mathcal{D}(t_1)| < |\mathcal{D}(t_0)|$. Similarly, it is possible to perturb t_1 to construct t_2 such that $|\mathcal{D}(t_2)| < |\mathcal{D}(t_1)|$ if $|\mathcal{D}(t_1)| > 0$. Necessarily, this process terminates at $t \in \mathbb{R}^n$ such that $t'b_k \neq t'b_l$ for all $k \neq l$.

Then there is an $(m+1) \times C(m,n)$ matrix T such that TB^* has full column rank. The matrix T is defined constructively, using the notation that $z \in \mathbb{R}^n$ is a free variable. For each integer $p \in \{0, 1, \dots, m\}$, row $p+1$ of T has $C(p-1,n)$ leading zeros, then is equal to $\bigotimes^p t'$, and then has trailing zeros. Therefore, row $p+1$ of $T\Gamma_{m,n}(z)$ is $(t'z)^p$, since $(t'z)^p = \bigotimes^p (t'z) = \bigotimes^p t' \bigotimes^p z$. In particular, for $p = 0$, use the convention that $(t'z)^0 = 1$. So since the first element of $\Gamma_{m,n}(z)$ is 1, the first row of T has a 1 along the diagonal and is equal to 0 everywhere else. Since $t'z$ is the sum of n terms, there are n^p terms in the series expansion of $(t'z)^p$. Therefore, the last nonzero term in row $p+1$ is in column $C(p-1,n) + n^p = C(p,n)$. Hence, as claimed, T has $C(m,n)$ columns.

By construction of T, the element of TB^* in row $p+1$ and column c is $(t'b_c)^p$. Therefore, TB^* is a Vandermonde matrix of dimension $(m+1) \times R$, in terms of the powers of $(t'b_c)$ for $c = 1, \dots, R$. Since $m+1 \geq R$, in particular one submatrix of TB^* is the Vandermonde matrix of dimension $R \times R$. Since $t'b_c \neq t'b_{c'}$ for $c \neq c'$ by construction of t, this Vandermonde matrix is based on distinct parameters, which implies that the square Vandermonde submatrix is nonsingular. So TB^* contains an $R \times R$ nonsingular submatrix. Since TB^* is $(m+1) \times R$, this implies that TB^* has full column rank. Because of the general result on the rank of products of matrices, $R = \operatorname{rank}(TB^*) \leq \min\{\operatorname{rank}(T), \operatorname{rank}(B^*)\}$, so B^* has full column rank. □

Lemma D.3. *Let* $\tilde{\mathrm{P}} = \{\tilde{\mathrm{P}}_w\}_{w=1}^W$ *be a set of possible magnitudes of computational mistakes with* $\tilde{\mathrm{P}}_1 < \tilde{\mathrm{P}}_2 < \cdots < \tilde{\mathrm{P}}_W$*. Based on* $\tilde{\mathrm{P}}$*, define vector-valued mappings* η_1*,* η_2*, and* η_3 *of the strategic behavior rules* $\Theta = (\Lambda, \Delta, \mathrm{P})$*.*

(i) *Define* $\eta_1(\Theta) = ((1-\Delta)\Lambda(\mathcal{M}(1)), \dots, (1-\Delta)\Lambda(\mathcal{M}(|\mathcal{M}|)))$*. So* η_1 *gives the vector of* $(1-\Delta)\Lambda(k)$ *for decision rules* $k \in \mathcal{M}$*.*

(ii) *Define* $\eta_2(\Theta) = (\Lambda(\mathcal{U}(1)), \dots, \Lambda(\mathcal{U}(|\mathcal{U}|)))$*. So* η_2 *gives the vector of* $\Lambda(k)$ *for decision rules* $k \in \mathcal{U}$*.*

(iii) *Define* $\eta_3(\Theta) = (\Delta\Lambda(\mathcal{M}(1))1[\mathrm{P} = \tilde{\mathrm{P}}_1], \dots, \Delta\Lambda(\mathcal{M}(1))1[\mathrm{P} = \tilde{\mathrm{P}}_W], \Delta\Lambda(\mathcal{M}(2)) \times 1[\mathrm{P} = \tilde{\mathrm{P}}_1], \dots)$*. So* η_3 *gives* $\Delta\Lambda(k)1[\mathrm{P} = \tilde{\mathrm{P}}_w]$ *for decision rules* $k \in \mathcal{M}$ *and* $w = 1, 2, \dots, W$*.*

Let $\eta^*(\Theta) = (\eta_1(\Theta), \eta_2(\Theta), \eta_3(\Theta))$ *and* $\eta^{**}(\Theta) = (\eta_1(\Theta), \eta_2(\Theta))$.

Suppose Θ_1 *and* Θ_2 *are two strategic behavior rules such that* $\mathrm{P}_1 \in \tilde{\mathrm{P}}$ *and* $\mathrm{P}_2 \in \tilde{\mathrm{P}}$. *If* $\eta^*(\Theta_1) = \eta^*(\Theta_2)$, *then* $\Lambda_1 = \Lambda_2$, $\Delta_1 1[\sum_{k\in\mathcal{M}} \Lambda_1(k) > 0] = \Delta_2 1[\sum_{k\in\mathcal{M}} \Lambda_2(k) > 0]$ *and* $\mathrm{P}_1 1[\Delta_1 > 0]1[\sum_{k\in\mathcal{M}} \Lambda_1(k) > 0] = \mathrm{P}_2 1[\Delta_2 > 0]1[\sum_{k\in\mathcal{M}} \Lambda_2(k) > 0]$.

Suppose Θ_1 *and* Θ_2 *are two strategic behavior rules. If* $\eta^{**}(\Theta_1) = \eta^{**}(\Theta_2)$, *then* $\Lambda_1 = \Lambda_2$ *and* $\Delta_1 1[\sum_{k\in\mathcal{M}} \Lambda_1(k) > 0] = \Delta_2 1[\sum_{k\in\mathcal{M}} \Lambda_2(k) > 0]$.

Proof. Suppose that $\eta^{**}(\Theta_1) = \eta^{**}(\Theta_2)$. It is immediate from the definition of η_1 that $(1 - \Delta_1)\Lambda_1(k) = (1 - \Delta_2)\Lambda_2(k)$ for any decision rule $k \in \mathcal{M}$. Also, it is immediate from the definition of η_2 that $\Lambda_1(k) = \Lambda_2(k)$ for any decision rule $k \in \mathcal{U}$. Necessarily, $1 = \sum_k \Lambda(k) = \sum_{k\in\mathcal{M}} \Lambda(k) + \sum_{k\in\mathcal{U}} \Lambda(k)$. Therefore, it must be that $\sum_{k\in\mathcal{M}} \Lambda_1(k) = \sum_{k\in\mathcal{M}} \Lambda_2(k)$ since $\sum_{k\in\mathcal{U}} \Lambda_1(k) = \sum_{k\in\mathcal{U}} \Lambda_2(k)$. Therefore, since $\sum_{k\in\mathcal{M}}(1 - \Delta_1)\Lambda_1(k) = \sum_{k\in\mathcal{M}}(1 - \Delta_2)\Lambda_2(k)$, it must be that $\Delta_1 1[\sum_{k\in\mathcal{M}} \Lambda_1(k) > 0] = \Delta_2 1[\sum_{k\in\mathcal{M}} \Lambda_2(k) > 0]$. Suppose that for all $k \in \mathcal{M}$ it holds that $\Lambda_1(k) = 0$. Then, since $\Delta_2 < 1$, it must be that $\Lambda_2(k) = 0$ for all $k \in \mathcal{M}$ by definition of η_1. So, in that case, $\Lambda_1(k) = \Lambda_2(k)$ for all $k \in \mathcal{M}$. If there is $k^* \in \mathcal{M}$ such that $\Lambda_1(k^*) > 0$, then since $\Delta_1 < 1$, it must be that $\Lambda_2(k^*) > 0$ by definition of η_1. In that case, it must indeed be that $\Delta_1 = \Delta_2$ since $1[\sum_{k\in\mathcal{M}} \Lambda_1(k) > 0] = 1[\sum_{k\in\mathcal{M}} \Lambda_2(k) > 0] = 1$. So then, by definition of η_1, it must be that $\Lambda_1(k) = \Lambda_2(k)$ for all $k \in \mathcal{M}$. So again, in that case, $\Lambda_1(k) = \Lambda_2(k)$ for all $k \in \mathcal{M}$.

Now suppose in addition that $\eta^*(\Theta_1) = \eta^*(\Theta_2)$. If $\Delta_1 = \Delta_2 > 0$ and $\sum_{k\in\mathcal{M}} \Lambda_1(k) = \sum_{k\in\mathcal{M}} \Lambda_2(k) > 0$, note that $\Delta_1 \sum_{k\in\mathcal{M}} \Lambda_1(k)1[\mathrm{P}_1 = \tilde{\mathrm{P}}_w]$ (or, respectively, $\Delta_2 \sum_{k\in\mathcal{M}} \Lambda_2(k) \times 1[\mathrm{P}_2 = \tilde{\mathrm{P}}_w]$) is nonzero if and only if $\mathrm{P}_1 = \tilde{\mathrm{P}}_w$ (or $\mathrm{P}_2 = \tilde{\mathrm{P}}_w$). Therefore, by definition of η_3, it must be that $\mathrm{P}_1 1[\Delta_1 > 0]1[\sum_{k\in\mathcal{M}} \Lambda_1(k) > 0] = \mathrm{P}_2 1[\Delta_2 > 0]1[\sum_{k\in\mathcal{M}} \Lambda_2(k) > 0]$. □

Lemma D.4. *The following claims are true.*

(i) *Suppose that* $k \in \mathcal{M}$. *In a game* g *that satisfies Assumptions* 4.1(ii) *and* 4.1(iii), *or a game* g *that satisfies Assumptions* B.1(ii) *and* B.1(iii), *it holds that*

$$P_{rg}\big(C_g(k)\big) = (1 - \Delta_r)\Lambda_r(k).$$

(ii) *Suppose that* $k \in \mathcal{M}$. *Suppose that* $0 < \varepsilon$. *In a game* g *that satisfies Assumption* 4.1(i), *or equivalently a game* g *that satisfies Assumption* B.1(i), *it holds that*

$$P_{rg}\big(C_g(k, \varepsilon)\big) = \sum_{k'\neq k} P_{rg}\big(C_g(k, \varepsilon)|\gamma_g = k'\big)\Lambda_r\big(k'\big) + \Delta_r M_g(k, \varepsilon, \mathrm{P}_r)\Lambda_r(k).$$

(iii) *Suppose that* $k \in \mathcal{M}$. *Suppose that* $0 < \varepsilon \le \overline{\rho}$, *where* $\overline{\rho}$ *arises from Assumption* 4.1. *In a game* g *that satisfies Assumptions* 4.1(i) *and* 4.1(ii), *it holds that*

$$P_{rg}\big(C_g(k, \varepsilon)\big) = \sum_{s\in\mathcal{U}} P_{rg}\big(C_g(k, \varepsilon)|\gamma_g = s_{\text{unanch}}\big)\Lambda_r(s_{\text{unanch}}) + \Delta_r M_g(k, \varepsilon, \mathrm{P}_r)\Lambda_r(k).$$

(iv) *Suppose that* $s \in \mathcal{U}$. *It holds that*

$$P_{rg}\big(U_g(s)\big) = \sum_{0\le s'\le s, s'\in\mathcal{U}} R_g\big(s, s', \overline{\rho}\big)\Lambda_r\big(s'_{\text{unanch}}\big).$$

Proof. (i) By the law of total probability,

$$P_{rg}\big(C_g(k)\big) = \sum_{k'} P_{rg}\big(C_g(k)|\gamma_g = k'\big)P_{rg}\big(\gamma_g = k'\big).$$

Under Assumptions 4.1(ii) and 4.1(iii), or Assumptions B.1(ii) and B.1(iii), there are no decision rules $k' \neq k$ that take the action associated with decision rule k with positive probability, so $P_{rg}(C_g(k)) = P_{rg}(C_g(k)|\gamma_g = k)\Lambda_r(k)$. Additionally, a subject who uses strategic behavior rule r and decision rule k will actually take the action predicted by decision rule k with probability $1 - \Delta_r$, since with probability Δ_r it makes a computational mistake and takes an action according to the density on a nondegenerate interval since $\mathrm{P}_r > 0$ by assumption when $\Delta_r > 0$. So $P_{rg}(C_g(k)|\gamma_g = k) = 1 - \Delta_r$.

(ii) A subject who uses strategic behavior rule r and intends to use decision rule k in game g and who makes a computational mistake will take an action that is distributed according to $\xi(\cdot)$, translated to the interval with radius $\mathrm{P}_r(\alpha_{Ug}(1) - \alpha_{Lg}(1))$ centered at the action predicted by decision rule k and intersected with the action space. Therefore,

$$\begin{aligned} P_{rg}\big(C_g(k,\varepsilon)\big) &= \sum_{k'} P_{rg}\big(C_g(k,\varepsilon)|\gamma_g = k'\big)P_{rg}\big(\gamma_g = k'\big) \\ &= \sum_{k'\neq k} P_{rg}\big(C_g(k,\varepsilon)|\gamma_g = k'\big)\Lambda_r\big(k'\big) + \Delta_r M_g(k,\varepsilon,\mathrm{P}_r)\Lambda_r(k). \end{aligned}$$

By Assumption 4.1(i), $\alpha_{Lg}(1) < \alpha_{Ug}(1)$, as long as $\mathrm{P}_r > 0$. This last expression does not involve dividing by zero in the definition of $\omega_{1g,c_{1g}(k),\mathrm{P}_r}(\cdot)$ that appears as the integrand in $M_g(k,\varepsilon,\mathrm{P}_r)$. The condition that $\mathrm{P}_r > 0$ is assumed in Section 2.3 when $\Delta_r > 0$. Otherwise, if $\mathrm{P}_r = 0$, then $\Delta_r = 0$ and the expression is still correct.

(iii) Since g is a game that additionally satisfies Assumption 4.1(ii) and $\varepsilon \leq \overline{\rho}$, then $P_{rg}(C_g(k,\varepsilon)|\gamma_g = k') = 0$ for any decision rule $k' \in \mathcal{M}$.

(iv) By construction, the only time $U_g(s)$ happens (with positive probability) is from subjects who use s' steps of unanchored strategic reasoning for some $0 \leq s' \leq s$ with $s' \in \mathcal{U}$, so it follows that

$$\begin{aligned} P_{rg}\big(U_g(s)\big) &= \sum_{k'} P_{rg}\big(U_g(s)|\gamma_g = k'\big)P_{rg}\big(\gamma_g = k'\big) \\ &= \sum_{0\leq s'\leq s, s'\in\mathcal{U}} R_g\big(s,s',\overline{\rho}\big)\Lambda_r\big(s'_{\text{unanch}}\big). \end{aligned}$$

□

Lemma D.5. *Suppose Assumptions* 2.2 *and* 4.1. *Suppose that the econometrician allows the possibility of computational mistakes. Suppose that g is a game that satisfies Assumptions* 4.1(i), 4.1(ii), 4.1(iv), *and* 4.2. *Then $\{\tilde{\mathrm{P}}_w\}_{w=1}^{W-1}$ is identified by the locations of the kinks in $\{P_g(C_g(k,\varepsilon))\}_{k\in\mathcal{M}}$ as a function of ε, for $0 < \varepsilon < \overline{\rho}$.*

Proof. Suppose that $k \in \mathcal{M}$. For $0 < \varepsilon \leq \overline{\rho}$, the probability of the event $C_g(k, \varepsilon)$ in game g is, using the result of Lemma D.4(ii) and Assumption 4.1(i),

$$\begin{aligned} P_g\big(C_g(k,\varepsilon)\big) &= \sum_{r=1}^{R} P_{rg}\big(C_g(k,\varepsilon)\big)\pi(r) \\ &= \sum_{r=1}^{R}\Big(\sum_{k'\neq k} P_{rg}\big(C_g(k,\varepsilon)|\gamma_g = k'\big)\Lambda_r\big(k'\big)\Big)\pi(r) \\ &\quad + \sum_{r=1}^{R}\big(\Delta_r M_g(k,\varepsilon,\mathrm{P}_r)\Lambda_r(k)\big)\pi(r). \end{aligned}$$

Since g is a game that satisfies Assumption 4.1(ii), it follows that $P_{rg}(C_g(k,\varepsilon)|\gamma_g = k') = 0$ for all such decision rules $k' \in \mathcal{M}$ with $k' \neq k$, since $\varepsilon \leq \overline{\rho}$. Therefore,

$$\begin{aligned} P_g\big(C_g(k,\varepsilon)\big) &= \sum_{r=1}^{R}\Big(\sum_{s\in\mathcal{U}} P_{rg}\big(C_g(k,\varepsilon)|\gamma_g = s_{\text{unanch}}\big)\Lambda_r(s_{\text{unanch}})\Big)\pi(r) \\ &\quad + \sum_{r=1}^{R}\big(\Delta_r M_g(k,\varepsilon,\mathrm{P}_r)\Lambda_r(k)\big)\pi(r). \end{aligned}$$

Since g satisfies Assumption 4.2(i), for any $s \in \mathcal{U}$, $P_{rg}(C_g(k,\varepsilon)|\gamma_g = s_{\text{unanch}})$ is a differentiable function of ε, for all $0 < \varepsilon < \overline{\rho}$. Under Assumptions 4.2(i), 4.2(iii), or 4.2(iv), $P_{rg}(C_g(k,\varepsilon)|\gamma_g = s_{\text{unanch}}) = \int_{C_g(k,\varepsilon)} \zeta^s_{1g}(a)\, d\mu(a; \Sigma^s_{1g}) = \int_{c_{1g}(k)-\varepsilon\Omega_{1g}}^{c_{1g}(k)+\varepsilon\Omega_{1g}} \zeta^s_{1g}(a)\, d\mu(a)$, where $\mu(\cdot)$ is Lebesgue measure, since Σ^s_{1g} cannot be a finite set under these conditions and, therefore, by the condition in footnote 10 is Lebesgue measurable with nonzero and finite measure, is differentiable in ε. Under Assumption 4.2(ii), $P_{rg}(C_g(k,\varepsilon)|\gamma_g = s_{\text{unanch}}) = 0$ for all $0 < \varepsilon < \overline{\rho}$.

Suppose that r is such that $\pi(r) > 0$ and $\Delta_r > 0$. Suppose that r uses at least one $k^*_r \in \mathcal{M}$ with positive probability. So it holds that $\Delta_r\Lambda_r(k^*_r)\pi(r) > 0$. Therefore, there is a kink in $P_g(C_g(k^*_r,\varepsilon))$ at $\varepsilon = \mathrm{P}_r$ since there is a kink in $M_g(k^*_r,\varepsilon,\mathrm{P}_r)$ at $\varepsilon = \mathrm{P}_r$ by Lemma D.1(ii). This uses the fact that $\mathrm{P}_r < \overline{\rho}$ for all r by Assumption 2.2, whereas the above expression for $P_g(C_g(k,\varepsilon))$ is valid for all $\varepsilon \leq \overline{\rho}$, so that the location of all relevant kinks is indeed identified. Moreover, there can be no other kinks in $M_g(k,\varepsilon,\mathrm{P}_r)$ for any $k \in \mathcal{M}$ and $0 < \varepsilon < \overline{\rho}$, by Lemma D.1(ii). Consequently, the list of nonzero unique values corresponding to $\{\mathrm{P}_r 1[\Delta_r > 0]1[\sum_{k\in\mathcal{M}} \Lambda_r(k) > 0]1[\pi(r) > 0]\}_r$ is identified by the list of the locations of the kinks in $\{P_g(C_g(k,\varepsilon))\}_{k\in\mathcal{M}}$ as a function of ε, for $0 < \varepsilon < \overline{\rho}$. □

Lemma D.6. *For each game g, define the following quantities.*

(i) *Let Q_{2g} be the $|\mathcal{U}| \times |\mathcal{U}|$ matrix that has an element in row r and column c that equals $R_g(\mathcal{U}(r), \mathcal{U}(c), \overline{\rho})$.*

(ii) *Let Q_{3g} be the $(W|\mathcal{M}|) \times |\mathcal{U}|$ matrix that has an element in row r and column c that equals the probability in game g of the event $C_g(\mathcal{M}(\lceil \frac{r}{W} \rceil), \tilde{\mathrm{P}}_{\mathrm{mod}(r-1,W)+1})$ according* *to the distribution of actions used by subjects who use* $\mathcal{U}(c)$ *steps of unanchored strategic reasoning.*

(iii) *For each* $k \in \mathcal{M}$*, let* Q_{4gk} *be the* $W \times W$ *matrix that has an element in row r and column c that equals* $M_g(k, \tilde{\mathrm{P}}_r, \tilde{\mathrm{P}}_c)$. *Then let* Q_{4g} *be the* $(W|\mathcal{M}|) \times (W|\mathcal{M}|)$ *matrix that has* $(Q_{4g\mathcal{M}(1)}, \dots, Q_{4g\mathcal{M}(|\mathcal{M}|)})$ *along the diagonal.*

Then let

$$Q_g = \begin{pmatrix} I_{|\mathcal{M}|\times|\mathcal{M}|} & 0 & 0 \\ 0 & Q_{2g} & 0 \\ 0 & Q_{3g} & Q_{4g} \end{pmatrix}.$$

For any game g satisfying Assumptions 4.1(i)–4.1(v), $P_{g,\theta} = Q_g \eta^*(\theta)$ *and* Q_g *is nonsingular.*

For any game g satisfying Assumption 4.1(v)*, or equivalently any game g satisfying Assumption* B.1(iv), Q_{2g} *is nonsingular.*

For any game g satisfying Assumptions 4.1(ii) *and* 4.1(iii)*, or any game g satisfying Assumptions* B.1(ii) *and* B.1(iii)*, the first* $|\mathcal{M}|$ *rows of* $P_{g,\theta}$ *are equal to the first* $|\mathcal{M}|$ *rows of* $Q_g\eta^*(\theta)$.

For any game g, rows $|\mathcal{M}| + 1$ *through* $|\mathcal{M}| + |\mathcal{U}|$ *of* $P_{g,\theta}$ *are equal to rows* $|\mathcal{M}| + 1$ *through* $|\mathcal{M}| + |\mathcal{U}|$ *of* $Q_g\eta^*(\theta)$.

For any game g satisfying Assumptions 4.1(i) *and* 4.1(ii)*, the last* $|W|\mathcal{M}$ *rows of* $P_{g,\theta}$ *are equal to the last* $W|\mathcal{M}|$ *rows of* $Q_g\eta^*(\theta)$.

Proof. Since $R_g(s, s', \overline{\rho}) = 0$ for $s' > s$ by construction, it follows that Q_{2g} is lower triangular. Since g is a game that satisfies Assumption 4.1(v), the diagonal elements are nonzero, implying that Q_{2g} is nonsingular.

By the following arguments, for a game g that satisfies Assumptions 4.1(i) and 4.1(iv), Q_{4gk} is nonsingular for each $k \in \mathcal{M}$. First consider the case that the econometrician allows the possibility of computational mistakes. Apply repeated elementary row operations: for rows $r \geq 2$ (if indeed $W \geq 2$), starting with row W and then moving to the next higher row, subtract row $r - 1$ from row r and substitute the result into row r. The resulting matrix $\tilde{\tilde{Q}}_{4gk}$ has element in row $r \geq 2$ and column c that equals $M_g(k, \tilde{\mathrm{P}}_r, \tilde{\mathrm{P}}_c) - M_g(k, \tilde{\mathrm{P}}_{r-1}, \tilde{\mathrm{P}}_c)$. For a game g that satisfies Assumptions 4.1(i) and 4.1(iv), by Lemma D.1(iii), this difference is 0 if $r - 1 \geq c$ and is strictly positive if $r \leq c$. Therefore, row $r \geq 2$ has $r - 1$ leading zeros and then positive elements. In row 1 and column c, the element is $M_g(k, \tilde{\mathrm{P}}_1, \tilde{\mathrm{P}}_c) > 0$. Therefore, for a game g that satisfies Assumptions 4.1(i) and 4.1(iv), $\tilde{\tilde{Q}}_{4gk}$ is an upper-diagonal matrix with nonzero elements along the diagonal, so is nonsingular. Therefore, Q_{4gk} is nonsingular for a game g that satisfies Assumptions 4.1(i) and 4.1(iv), and, therefore, the matrix Q_{4g} has full rank if g satisfies Assumptions 4.1(i) and 4.1(iv). Second, consider the case that the econometrician does not allow computational mistakes. In that case, $W = 1$ and $\tilde{\mathrm{P}}_1 = 0$, so $Q_{4gk} = 1$ has full rank.

Then Q_g is nonsingular since all of the diagonal matrices are nonsingular.

The first block of $|\mathcal{M}|$ rows of $\eta^*(\theta)$ gives the vector of $((1-\Delta)\Lambda(\mathcal{M}(1)),\ldots,(1-\Delta)\Lambda(\mathcal{M}(|\mathcal{M}|)))$. Therefore, since g is a game that satisfies Assumptions 4.1(ii) and 4.1(iii), the first block of $\mathcal{M}$ rows of $Q_g\eta^*(\theta)$ is indeed the first block of $\mathcal{M}$ rows of $P_{g,\theta}$ by Lemma D.4(i) (and, similarly, the same would be true if g were a game satisfying Assumptions B.1(ii) and B.1(iii)). The second block of $|\mathcal{U}|$ rows of $\eta^*(\theta)$ gives the vector of $(\Lambda(\mathcal{U}(1)),\ldots,\Lambda(\mathcal{U}(|\mathcal{U}|)))$. Therefore, by Lemma D.4(iv), by definition, it follows that the second block of $|\mathcal{U}|$ rows of $Q_g\eta^*(\theta)$ is indeed the second block of $|\mathcal{U}|$ rows of $P_{g,\theta}$. Finally, the last block of $W|\mathcal{M}|$ rows of $\eta^*(\theta)$ gives the vector of $(\Delta\Lambda(\mathcal{M}(1))\times 1[\mathrm{P}=\tilde{\mathrm{P}}_1],\ldots,\Delta\Lambda(\mathcal{M}(1))1[\mathrm{P}=\tilde{\mathrm{P}}_W],\Delta\Lambda(\mathcal{M}(2))1[\mathrm{P}=\tilde{\mathrm{P}}_1],\ldots)$. Also, the last block of $W|\mathcal{M}|$ rows of $P_{g,\theta}$ is $(P_{g,\theta}(C_g(\mathcal{M}(1),\tilde{\mathrm{P}}_1)),\ldots,P_{g,\theta}(C_g(\mathcal{M}(1),\tilde{\mathrm{P}}_W)),P_{g,\theta}(C_g(\mathcal{M}(2),\tilde{\mathrm{P}}_1)),\ldots)$. Therefore, it follows from Lemma D.4(iii), and the fact that g is a game that satisfies Assumptions 4.1(i) and 4.1(ii) and the definition of Q_{3g}, that indeed the last block of $W|\mathcal{M}|$ rows of $Q_g\eta^*(\theta)$ is indeed the last block of $|W|\mathcal{M}$ rows of $P_{g,\theta}$. □

Proof of Theorem 4.1. Using the game g that satisfies the conditions of Assumption 4.2, and Lemma D.5, it is possible to identify $\{\tilde{\mathrm{P}}_w\}_{w=1}^W$.

Let $\mathcal{G}$ be a subset of $\{1,2,\ldots,G\}$ with $|\mathcal{G}|\geq 2R-1$ games that satisfy the conditions of Assumption 4.1. Let $\mathcal{G}(p)$ be the pth smallest element of $\mathcal{G}$. Let $\mathcal{G}_p=\{\mathcal{G}(1),\ldots,\mathcal{G}(p)\}$. Let $Q_{\mathcal{G}}^{(0)}=1$ and $Q_{\mathcal{G}}^{(p)}=Q_{\mathcal{G}(1)}\otimes\cdots\otimes Q_{\mathcal{G}(p)}$. Let $Q_{\mathcal{G}}$ be the block diagonal matrix with the blocks along the diagonal equal to $Q_{\mathcal{G}}^{(0)},\ldots,Q_{\mathcal{G}}^{(|\mathcal{G}|)}$; $Q_{\mathcal{G}}$ is nonsingular as long as each diagonal block is nonsingular. So since $Q_{\mathcal{G}(p)}$ is nonsingular for all p by Lemma D.6, which implies that $Q_{\mathcal{G}}^{(p)}$ is nonsingular by the algebra of the Kronecker product, then $Q_{\mathcal{G}}$ is nonsingular.

Let $P_{\mathcal{G},\theta,p}\equiv P_{\mathcal{G}(1),\theta}\otimes\cdots\otimes P_{\mathcal{G}(p),\theta}$. Since actions are independent across games, $P_{\mathcal{G},\theta,p}$ gives the joint distribution of the events $C(\cdot)$, $U(\cdot)$, and $C(\cdot,\cdot)$ across games $\mathcal{G}_p$. Let $P_{\mathcal{G},\theta}=(1,P_{\mathcal{G},\theta,1},\ldots,P_{\mathcal{G},\theta,|\mathcal{G}|})$. Let $\eta^*(\theta)^{(0)}=1$ and $\eta^*(\theta)^{(p)}=\eta^*(\theta)\otimes\cdots\otimes\eta^*(\theta)$ be the p-times Kronecker products. Let $\overline{\eta}^*(\theta)=(1,\eta^*(\theta)^{(1)},\ldots,\eta^*(\theta)^{(|\mathcal{G}|)})$.

Then, using the results of Lemma D.6, it follows from the algebra of the Kronecker product that $P_{\mathcal{G},\theta,p}\equiv P_{\mathcal{G}(1),\theta}\otimes\cdots\otimes P_{\mathcal{G}(p),\theta}=(Q_{\mathcal{G}(1)}\eta^*(\theta))\otimes\cdots\otimes(Q_{\mathcal{G}(p)}\eta^*(\theta))=(Q_{\mathcal{G}(1)}\otimes\cdots\otimes Q_{\mathcal{G}(p)})(\eta^*(\theta)\otimes\cdots\otimes\eta^*(\theta))=Q_{\mathcal{G}}^{(p)}\eta^*(\theta)^{(p)}$. Also $P_{\mathcal{G},\theta}=Q_{\mathcal{G}}\overline{\eta}^*(\theta)$.

Let the true parameters of the data generating process be $\Theta_{01},\ldots,\Theta_{0\tilde{R}_0}$ and $\pi_0(1),\ldots,\pi_0(\tilde{R}_0)$, where $\tilde{R}_0\leq R$ is the number of strategic behavior rules that are used in the population and where Θ_{0r} is not observationally equivalent to $\Theta_{0r'}$ for all $r\neq r'$ per Definition 1. So, by construction, $\pi_0(\cdot)>0$. Then, by the above, it follows that $P_{\mathcal{G},\Theta_{0r}}=Q_{\mathcal{G}}\overline{\eta}^*(\Theta_{0r})$ for each r. Let $Y_0^*=(\overline{\eta}^*(\Theta_{01})\ \cdots\ \overline{\eta}^*(\Theta_{0\tilde{R}_0}))$. Since no pair of strategic behavior rules are observationally equivalent, by Lemma D.3, the columns of Y_0^* are distinct. Then $P_{\mathcal{G},0}=Q_{\mathcal{G}}Y_0^*\pi_0$, where $P_{\mathcal{G},0}$ is the observed joint distribution of actions in games $\mathcal{G}$.

Suppose that there were an observationally equivalent specification of the parameters $\Theta_{1\cdot}$ and $\pi_1(\cdot)$, with corresponding Y_1^*, such that $P_{\mathcal{G},0}=Q_{\mathcal{G}}Y_1^*\pi_1$, where again by construction no columns of Y_1^* correspond to a rule r such that $\pi_1(r)=0$ and no pair of strategic behavior rules are observationally equivalent. Let $\overline{Y}^*$ collect the unique columns of $(Y_0^*\ Y_1^*)$. Similarly, let $\overline{\pi}$ be the corresponding differences between π_0 and π_1.

If column c of $\overline{Y}^*$ exists in both Y_0^* and Y_1^* as columns c_0 and c_1, respectively, then set $\overline{\pi}_c = \pi_0(c_0) - \pi_1(c_1)$. If column c of $\overline{Y}^*$ exists only in Y_0^* as column c_0, then set $\overline{\pi}_c = \pi_0(c_0)$, and if column c of $\overline{Y}^*$ exists only in Y_1^* as column c_1, then set $\overline{\pi}_c = -\pi_1(c_1)$. Then $0 = Q_{\mathcal{G}}\overline{Y}^*\overline{\pi}$. By Lemma D.2, since the number of columns of $\overline{Y}^*$ is at most $2R$ and since $|\mathcal{G}| \geq 2R-1$, $\overline{Y}^*$ has full column rank and, therefore, $Q_{\mathcal{G}}\overline{Y}^*$ has full column rank since $Q_{\mathcal{G}}$ is nonsingular, so $\overline{\pi} = 0$. Therefore, any strategic behavior rules that appear in specifications 0 and 1 are used with equal probability, and there are no strategic behavior rules used only in specifications 0 and 1, since no elements of π_0 and π_1 are equal to zero by construction.

Therefore, Y_0^* and Y_1^* contain exactly the same columns, up to permuting the order of the columns, and, the probabilities of the corresponding strategic behavior rules are also equal across specifications. Note, in particular, that this implies that the set of $\eta^*(\Theta_{0r})$ for $r = 1, 2, \dots, \tilde{R}$ and the set of $\eta^*(\Theta_{1r})$ for $r = 1, 2, \dots, \tilde{R}$ are equal up to permutations of the labels. Since η^* is injective in the sense of Lemma D.3, the two specifications of the parameters are the same up to observational equivalence in Definition 1 (up to permutations of the labels), so the parameters are point identified in the sense of Definition 2. □

Proof of Theorem B.1. Let $\mathcal{G}_{\mathcal{M}}$ be a subset of $\{1, 2, \dots, G\}$ with at least $|\mathcal{G}_{\mathcal{M}}| \geq 2R-1$ games that satisfy the first set of conditions of Assumption B.1. Let $\mathcal{G}_{\mathcal{M}}(p)$ be the pth smallest element of $\mathcal{G}_{\mathcal{M}}$. Let $\mathcal{G}_{p,\mathcal{M}} = \{\mathcal{G}_{\mathcal{M}}(1), \dots, \mathcal{G}_{\mathcal{M}}(p)\}$. Let $Q_{\mathcal{G}_{\mathcal{M}}}^{(0)} = 1$ and let $Q_{\mathcal{G}_{\mathcal{M}}}^{(p)} = I_{|\mathcal{M}|\times|\mathcal{M}|} \otimes \cdots \otimes I_{|\mathcal{M}|\times|\mathcal{M}|}$ be the p-times Kronecker product of $I_{|\mathcal{M}|\times|\mathcal{M}|}$. Let $Q_{\mathcal{G}_{\mathcal{M}}}$ be the block diagonal matrix with the blocks along the diagonal equal to $Q_{\mathcal{G}_{\mathcal{M}}}^{(0)}, \dots, Q_{\mathcal{G}_{\mathcal{M}}}^{(|\mathcal{G}_{\mathcal{M}}|)}$; $Q_{\mathcal{G}_{\mathcal{M}}}$ is nonsingular as long as each diagonal block is nonsingular. So since $Q_{\mathcal{G}_{\mathcal{M}}}^{(p)}$ is nonsingular by the algebra of the Kronecker product, $Q_{\mathcal{G}_{\mathcal{M}}}$ is nonsingular.

Let $\mathcal{G}_{\mathcal{U}}$ be a subset of $\{1, 2, \dots, G\}$ with at least $|\mathcal{G}_{\mathcal{U}}| \geq 2R-1$ games that satisfy the second set of conditions of Assumption B.1. Let $\mathcal{G}_{\mathcal{U}}(p)$ be the pth smallest element of $\mathcal{G}_{\mathcal{U}}$. Let $\mathcal{G}_{p,\mathcal{U}} = \{\mathcal{G}_{\mathcal{U}}(1), \dots, \mathcal{G}_{\mathcal{U}}(p)\}$. Let $Q_{\mathcal{G}_{\mathcal{U}}}^{(0)} = 1$ and $Q_{\mathcal{G}_{\mathcal{U}}}^{(p)} = Q_{2\mathcal{G}_{\mathcal{U}}(1)} \otimes \cdots \otimes Q_{2\mathcal{G}_{\mathcal{U}}(p)}$. Let $Q_{\mathcal{G}_{\mathcal{U}}}$ be the block diagonal matrix with the blocks along the diagonal equal to $Q_{\mathcal{G}_{\mathcal{U}}}^{(0)}, \dots, Q_{\mathcal{G}_{\mathcal{U}}}^{(|\mathcal{G}_{\mathcal{U}}|)}$; $Q_{\mathcal{G}_{\mathcal{U}}}$ is nonsingular as long as each diagonal block is nonsingular. So since $Q_{2\mathcal{G}_{\mathcal{U}}(p)}$ is nonsingular for all p by Lemma D.6, which implies that $Q_{\mathcal{G}_{\mathcal{U}}}^{(p)}$ is nonsingular by the algebra of the Kronecker product, $Q_{\mathcal{G}_{\mathcal{U}}}$ is nonsingular.

Let $P_{\mathcal{G}_{\mathcal{M}}(p),\theta,\mathcal{M}}$ be the first $|\mathcal{M}|$ rows of $P_{\mathcal{G}_{\mathcal{M}}(p),\theta}$. Let $P_{\mathcal{G}_{\mathcal{M}},\theta,p,\mathcal{M}} \equiv P_{\mathcal{G}_{\mathcal{M}}(1),\theta,\mathcal{M}} \otimes \cdots \otimes P_{\mathcal{G}_{\mathcal{M}}(p),\theta,\mathcal{M}}$. Since the actions in the games are independent across games, $P_{\mathcal{G}_{\mathcal{M}},\theta,p,\mathcal{M}}$ gives the joint distribution of the events $C(\cdot)$ across games $\mathcal{G}_{p,\mathcal{M}}$. Let $P_{\mathcal{G}_{\mathcal{M}},\theta,\mathcal{M}} = (1, P_{\mathcal{G}_{\mathcal{M}},\theta,1,\mathcal{M}}, \dots, P_{\mathcal{G}_{\mathcal{M}},\theta,|\mathcal{G}_{\mathcal{M}}|,\mathcal{M}})$. Let $\eta^*_{\mathcal{M}}(\theta)$ be the first $|\mathcal{M}|$ rows of $\eta^*(\theta)$. Let $\eta^*_{\mathcal{M}}(\theta)^{(0)} = 1$ and $\eta^*_{\mathcal{M}}(\theta)^{(p)} = \eta^*_{\mathcal{M}}(\theta) \otimes \cdots \otimes \eta^*_{\mathcal{M}}(\theta)$ be the p-times Kronecker product. Let $\overline{\eta}^*_{\mathcal{M}}(\theta) = (1, \eta^*_{\mathcal{M}}(\theta)^{(1)}, \dots, \eta^*_{\mathcal{M}}(\theta)^{(|\mathcal{G}_{\mathcal{M}}|)})$.

Let $P_{\mathcal{G}_{\mathcal{U}}(p),\theta,\mathcal{U}}$ be rows $|\mathcal{M}|+1$ through $|\mathcal{M}|+|\mathcal{U}|$ of $P_{\mathcal{G}_{\mathcal{U}}(p),\theta}$. Let $P_{\mathcal{G}_{\mathcal{U}},\theta,p,\mathcal{U}} \equiv P_{\mathcal{G}_{\mathcal{U}}(1),\theta,\mathcal{U}} \otimes \cdots \otimes P_{\mathcal{G}_{\mathcal{U}}(p),\theta,\mathcal{U}}$. Since the actions in the games are independent across games, $P_{\mathcal{G}_{\mathcal{U}},\theta,p,\mathcal{U}}$ gives the joint distribution of the events $U(\cdot)$ across games $\mathcal{G}_{p,\mathcal{U}}$. Let $P_{\mathcal{G}_{\mathcal{U}},\theta,\mathcal{U}} = (1, P_{\mathcal{G}_{\mathcal{U}},\theta,1,\mathcal{U}}, \dots, P_{\mathcal{G}_{\mathcal{U}},\theta,|\mathcal{G}_{\mathcal{U}}|,\mathcal{U}})$. Let $\eta^*_{\mathcal{U}}(\theta)$ be rows $|\mathcal{M}|+1$ through $|\mathcal{M}|+|\mathcal{U}|$ of

$\eta^*(\theta)$. Let $\eta^*_{\mathcal{U}}(\theta)^{(0)} = 1$ and $\eta^*_{\mathcal{U}}(\theta)^{(p)} = \eta^*_{\mathcal{U}}(\theta) \otimes \cdots \otimes \eta^*_{\mathcal{U}}(\theta)$ be the p-times Kronecker product. Let $\overline{\eta}^*_{\mathcal{U}}(\theta) = (1, \eta^*_{\mathcal{U}}(\theta)^{(1)}, \ldots, \eta^*_{\mathcal{U}}(\theta)^{(|\mathcal{G}_{\mathcal{U}}|)})$.

Then, using the results of Lemma D.6, it follows from the algebra of the Kronecker product that $P_{\mathcal{G}_{\mathcal{M}},\theta,p,\mathcal{M}} \equiv P_{\mathcal{G}_{\mathcal{M}}(1),\theta,\mathcal{M}} \otimes \cdots \otimes P_{\mathcal{G}_{\mathcal{M}}(p),\theta,\mathcal{M}} = (I_{|\mathcal{M}|\times|\mathcal{M}|}\eta^*_{\mathcal{M}}(\theta)) \otimes \cdots \otimes (I_{|\mathcal{M}|\times|\mathcal{M}|}\eta^*_{\mathcal{M}}(\theta)) = (I_{|\mathcal{M}|\times|\mathcal{M}|} \otimes \cdots \otimes I_{|\mathcal{M}|\times|\mathcal{M}|})(\eta^*_{\mathcal{M}}(\theta) \otimes \cdots \otimes \eta^*_{\mathcal{M}}(\theta)) = Q^{(p)}_{\mathcal{G}_{\mathcal{M}}}\eta^*_{\mathcal{M}}(\theta)^{(p)}$. Also $P_{\mathcal{G}_{\mathcal{M}},\theta,\mathcal{M}} = Q_{\mathcal{G}_{\mathcal{M}}}\overline{\eta}^*_{\mathcal{M}}(\theta)$.

Similarly, using the results of Lemma D.6, it follows from the algebra of the Kronecker product that $P_{\mathcal{G}_{\mathcal{U}},\theta,p,\mathcal{U}} \equiv P_{\mathcal{G}_{\mathcal{U}}(1),\theta,\mathcal{U}} \otimes \cdots \otimes P_{\mathcal{G}_{\mathcal{U}}(p),\theta,\mathcal{U}} = (Q_{2\mathcal{G}_{\mathcal{U}}(1)}\eta^*_{\mathcal{U}}(\theta)) \otimes \cdots \otimes (Q_{2\mathcal{G}_{\mathcal{U}}(p)}\eta^*_{\mathcal{U}}(\theta)) = (Q_{2\mathcal{G}_{\mathcal{U}}(1)} \otimes \cdots \otimes Q_{2\mathcal{G}_{\mathcal{U}}(p)})(\eta^*_{\mathcal{U}}(\theta) \otimes \cdots \otimes \eta^*_{\mathcal{U}}(\theta)) = Q^{(p)}_{\mathcal{G}_{\mathcal{U}}}\eta^*_{\mathcal{U}}(\theta)^{(p)}$. Also $P_{\mathcal{G}_{\mathcal{U}},\theta,\mathcal{U}} = Q_{\mathcal{G}_{\mathcal{U}}}\overline{\eta}^*_{\mathcal{U}}(\theta)$.

Then let $\tilde{P}_{\mathcal{G}_{\mathcal{M}},\mathcal{G}_{\mathcal{U}},\theta} = (P_{\mathcal{G}_{\mathcal{M}},\theta,\mathcal{M}}, P_{\mathcal{G}_{\mathcal{U}},\theta,\mathcal{U}})$. Let $\overline{\eta}^*_{\mathcal{M},\mathcal{U}}(\theta) = (\overline{\eta}^*_{\mathcal{M}}(\theta), \overline{\eta}^*_{\mathcal{U}}(\theta))$ and let $Q_{\mathcal{G}_{\mathcal{M}},\mathcal{G}_{\mathcal{U}}}$ be the partioned matrix with $(Q_{\mathcal{G}_{\mathcal{M}}}, Q_{\mathcal{G}_{\mathcal{U}}})$ along the diagonal.

Let the true parameters of the data generating process be $\Theta_{01}, \ldots, \Theta_{0\tilde{R}_0}$ and $\pi_0(1), \ldots, \pi_0(\tilde{R}_0)$, where $\tilde{R}_0 \leq R$ is the number of strategic behavior rules that are used in the population and Θ_{0r} is not observationally equivalent ignoring the magnitude of computational mistakes to $\Theta_{0r'}$ for all $r \neq r'$ per Definition 3. So, by construction, $\pi_0(\cdot) > 0$. Then, by the above, it follows that $P_{\mathcal{G}_{\mathcal{M}},\Theta_{0r},\mathcal{M}} = Q_{\mathcal{G}_{\mathcal{M}}}\overline{\eta}^*_{\mathcal{M}}(\Theta_{0r})$ and $P_{\mathcal{G}_{\mathcal{U}},\Theta_{0r},\mathcal{U}} = Q_{\mathcal{G}_{\mathcal{U}}}\overline{\eta}^*_{\mathcal{U}}(\Theta_{0r})$. Let $Y^*_{0,\mathcal{M}} = (\overline{\eta}^*_{\mathcal{M}}(\Theta_{01}) \cdots \overline{\eta}^*_{\mathcal{M}}(\Theta_{0\tilde{R}_0}))$ and $Y^*_{0,\mathcal{U}} = (\overline{\eta}^*_{\mathcal{U}}(\Theta_{01}) \cdots \overline{\eta}^*_{\mathcal{U}}(\Theta_{0\tilde{R}_0}))$. By Assumption B.2, the columns of $Y^*_{0,\mathcal{M}}$ are distinct and the columns of $Y^*_{0,\mathcal{U}}$ are distinct. Then $P_{\mathcal{G}_{\mathcal{M}},0,\mathcal{M}} = Q_{\mathcal{G}_{\mathcal{M}}}Y^*_{0,\mathcal{M}}\pi_0$, where $P_{\mathcal{G}_{\mathcal{M}},0,\mathcal{M}}$ is the observed joint distribution of actions in games $\mathcal{G}_{\mathcal{M}}$, and $P_{\mathcal{G}_{\mathcal{U}},0,\mathcal{U}} = Q_{\mathcal{G}_{\mathcal{U}}}Y^*_{0,\mathcal{U}}\pi_0$, where $P_{\mathcal{G}_{\mathcal{U}},0,\mathcal{U}}$ is the observed joint distribution of actions in games $\mathcal{G}_{\mathcal{U}}$.

Suppose that there were an observationally equivalent specification of the parameters $\Theta_{1\cdot}$ and $\pi_1(\cdot)$, with corresponding $Y^*_{1,\mathcal{M}}$ and $Y^*_{1,\mathcal{U}}$, such that $P_{\mathcal{G}_{\mathcal{M}},0,\mathcal{M}} = Q_{\mathcal{G}_{\mathcal{M}}}Y^*_{1,\mathcal{M}}\pi_1$ and $P_{\mathcal{G}_{\mathcal{U}},0,\mathcal{U}} = Q_{\mathcal{G}_{\mathcal{U}}}Y^*_{1,\mathcal{U}}\pi_1$, where again by construction the columns of $Y^*_{1,\mathcal{M}}$ and the columns of $Y^*_{1,\mathcal{U}}$ are distinct, and no columns correspond to a rule r such that $\pi_1(r) = 0$. By the same arguments as finishes the proof of Theorem 4.1, since $|\mathcal{G}_{\mathcal{M}}| \geq 2R - 1$ and $|\mathcal{G}_{\mathcal{U}}| \geq 2R - 1$, $(\pi(r), (1-\Delta_r)\Lambda_r(\mathcal{M}(1)), \ldots, (1-\Delta_r)\Lambda_r(\mathcal{M}(|\mathcal{M}|)))$ and $(\pi(r), \Lambda_r(\mathcal{U}(1)), \ldots, \Lambda_r(\mathcal{U}(|\mathcal{U}|)))$ are point identified up to permutations of the labels in the sense that the values of those two quantities must be equal across specifications of the parameters, up to permutations of the labels. And then, since $\pi(r)$ and $\pi(r')$ are distinct for $r' \neq r$ by Assumption B.2, it is possible to point identify $(\pi(r), (1-\Delta_r)\Lambda_r(\mathcal{M}(1)), \ldots, (1-\Delta_r)\Lambda_r(\mathcal{M}(|\mathcal{M}|)), \Lambda_r(\mathcal{U}(1)), \ldots, \Lambda_r(\mathcal{U}(|\mathcal{U}|)))$, in the sense that that quantity must be equal across specifications of the parameters, up to permutations of the labels, by "piecing together" the two point identification results on $(\pi(r), (1-\Delta_r)\Lambda_r(\mathcal{M}(1)), \ldots, (1-\Delta_r)\Lambda_r(\mathcal{M}(|\mathcal{M}|)))$ and $(\pi(r), \Lambda_r(\mathcal{U}(1)), \ldots, \Lambda_r(\mathcal{U}(|\mathcal{U}|)))$.

Note, in particular, that this implies that the set of $\eta^{**}(\Theta_{0r})$ for $r = 1, 2, \ldots, \tilde{R}$ and the set of $\eta^{**}(\Theta_{1r})$ for $r = 1, 2, \ldots, \tilde{R}$ are equal up to permutations of the labels. Since η^{**} is injective in the sense of Lemma D.3, the two specifications of the parameters are the same up to observational equivalence in Definition 3 (up to permutations of the labels), so the parameters are point identified in the sense of Definition 4. □

Appendix E: Verifying model assumptions in the empirical application

This establishes that the sufficient conditions for point identification are satisfied in the empirical application. The same approach would be taken in any empirical application.

First, it is necessary to specify the sets $\mathcal{A}$ and $\mathcal{U}$ from Assumption 2.1. Overall, based on visually inspecting the figures from Section 5.2 and Appendix F, it appears that there is essentially no subject who uses three or more steps of anchored strategic reasoning, basically the standard finding in experimental game theory. Therefore, Assumption 2.1 is maintained with $\mathcal{A} = \{1_{\text{anch}}, 2_{\text{anch}}\}$. Further, Assumption 2.1 is maintained with $\mathcal{U} = \{0_{\text{unanch}}, 1_{\text{unanch}}\}$, largely because there are not enough games in this data set such that the predictions of 1 and 2 steps of unanchored strategic reasoning differ sufficiently to guarantee point identification of the model with a larger set for $\mathcal{U}$, given the conditions in Assumptions 4.1 or B.1. See below for further discussion of Assumptions 4.1 or B.1.

Second, Assumption 2.2 states that the model of computational mistakes is correct and, therefore, is directly assumed by the econometrician. Specifically, the empirical application rules out computational mistakes. Because computational mistakes are ruled out, $\overline{\rho} = 0$.

Third, verifying Assumption 4.1 (or, by similar steps, Assumption B.1) requires inspecting Table 1 and checking which games satisfy the conditions in Assumption 4.1 (or the weaker conditions in Assumption B.1):

- Assumption 4.1(i) requires that the game has a nondegenerate action space. Obviously, all games in this data set satisfy this.

- Assumption 4.1(ii) requires that the game is such that the actions associated with the strategies in $\mathcal{M}$ (in this application, 1 and 2 steps of anchored strategic reasoning, and Nash equilibrium) are all distinct. It is easy to directly verify by inspecting Table 1 that games 1, 2, 3, 9, 10, 11, 12, 13, 14, 15, and 16 satisfy this condition. More generally, the condition requires that if computational mistakes were to be allowed, then those actions would need to separated from each other by a sufficient magnitude.

- Assumption 4.1(iii) requires that if a certain number of steps of unanchored strategic reasoning in $\mathcal{U}$ (in this application, 0 and 1 steps) predict a finite set of actions, then those actions are distinct from the predictions of the steps of anchored strategic reasoning in $\mathcal{A}$ and Nash equilibrium. Since no game is such that 0 or 1 steps of unanchored strategic reasoning predict a finite set of actions, this condition is satisfied in all games in the data set.

- Assumption 4.1(iv) requires that the game be such that the actions associated with the strategies in $\mathcal{M}$ (in this application, 1 and 2 steps of anchored strategic reasoning and Nash equilibrium) are not on *both* endpoints of the action space. Since the action spaces are all intervals, it is not possible for any given action to be on both endpoints, so all games in this data set satisfy this. More generally, the condition requires that if computational mistakes were to be allowed, then those actions would be required to be separated from *at least one of* the endpoints of the action space by a sufficient magnitude.

- Assumption 4.1(v) requires that the game be such that, for each $s \in \mathcal{U}$, there are actions used by s steps of unanchored strategic reasoning that are *not used* by s' steps of unanchored strategic reasoning (for each $s' \in \mathcal{U}$ with $s' > s$) *nor* used by the strategies in $\mathcal{M}$. In this application, that means there must be actions used by 0 steps of unanchored strategic reasoning, but not used by 1 step of unanchored strategic reasoning, nor used by 1 or 2 steps of anchored strategic reasoning, nor used by Nash equilibrium. Also this means there must be actions used by 1 step of unanchored strategic reasoning but not used by 1 or 2 steps of anchored strategic reasoning, nor used by Nash equilibrium. It is easy to directly verify by inspecting Table 1 that games 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, and 16 satisfy this condition. More generally, the condition requires that if computational mistakes were to be allowed, it would be necessary that these actions are not just different from the actions used by the strategies in $\mathcal{M}$, but also separated from the actions used by the strategies in $\mathcal{M}$ by a sufficient magnitude.

Therefore, games 2, 3, 9, 10, 11, 12, 14, 15, and 16 satisfy all these conditions, a total of 9 games, and, therefore, Assumption 4.1 is satisfied for any $R \leq 5$.

Finally, verifying Assumption 4.2 requires establishing that at least one game satisfies the extra condition in Assumption 4.2 among the games satisfying Assumptions 4.1(i), 4.1(ii), and 4.1(iv), or, in other words, in this application among games 1, 2, 3, 9, 10, 11, 12, 13, 14, 15, and 16. But recall from above that $\overline{\rho} = 0$ since computational mistakes are ruled out in the empirical application. In that case, note that logically either Assumption 4.2(i) or 4.2(ii) must be true, since the singleton $c_{1g}(k)$ must either be a subset or disjoint from any given set. Therefore, Assumption 4.2 is clearly satisfied for all games satisfying Assumptions 4.1(i), 4.1(ii), and 4.1(iv).

Note that even if computational mistakes were to be allowed, this assumption can be easily verified as true for sufficiently small $\overline{\rho}$ (maximum magnitude of computational mistakes). For example, consider game $g = 2$. Verifying Assumption 4.2 holds for game $g = 2$ and sufficiently small $\overline{\rho}$ requires simply verifying the following based on inspecting Table 1:

- For $k = 1_{\text{anch}}$ and $s = 0_{\text{unanch}}$, notice that $c_{1g}(1_{\text{anch}}) = 150$ is in the interior of $\Sigma^0_{1g} = [100, 900]$, so clearly $[c_{1g}(1_{\text{anch}}) - \overline{\rho}\Omega_{1g}, c_{1g}(1_{\text{anch}}) + \overline{\rho}\Omega_{1g}]$ is a subset of Σ^0_{1g} for small enough $\overline{\rho}$.

- For $k = 1_{\text{anch}}$ and $s = 1_{\text{unanch}}$, notice that $c_{1g}(1_{\text{anch}}) = 150$ is in the interior of $\Sigma^1_{1g} = [100, 250]$, so clearly again $[c_{1g}(1_{\text{anch}}) - \overline{\rho}\Omega_{1g}, c_{1g}(1_{\text{anch}}) + \overline{\rho}\Omega_{1g}]$ is a subset of Σ^1_{1g} for small enough $\overline{\rho}$.

- For $k = 2_{\text{anch}}$ and $s = 0_{\text{unanch}}$, notice that $c_{1g}(2_{\text{anch}}) = 175$ is in the interior of $\Sigma^0_{1g} = [100, 900]$, so clearly $[c_{1g}(2_{\text{anch}}) - \overline{\rho}\Omega_{1g}, c_{1g}(2_{\text{anch}}) + \overline{\rho}\Omega_{1g}]$ is a subset of Σ^0_{1g} for small enough $\overline{\rho}$.

- For $k = 2_{\text{anch}}$ and $s = 1_{\text{unanch}}$, notice that $c_{1g}(2_{\text{anch}}) = 175$ is in the interior of $\Sigma^1_{1g} = [100, 250]$, so clearly again $[c_{1g}(2_{\text{anch}}) - \overline{\rho}\Omega_{1g}, c_{1g}(2_{\text{anch}}) + \overline{\rho}\Omega_{1g}]$ is a subset of Σ^1_{1g} for small enough $\overline{\rho}$.

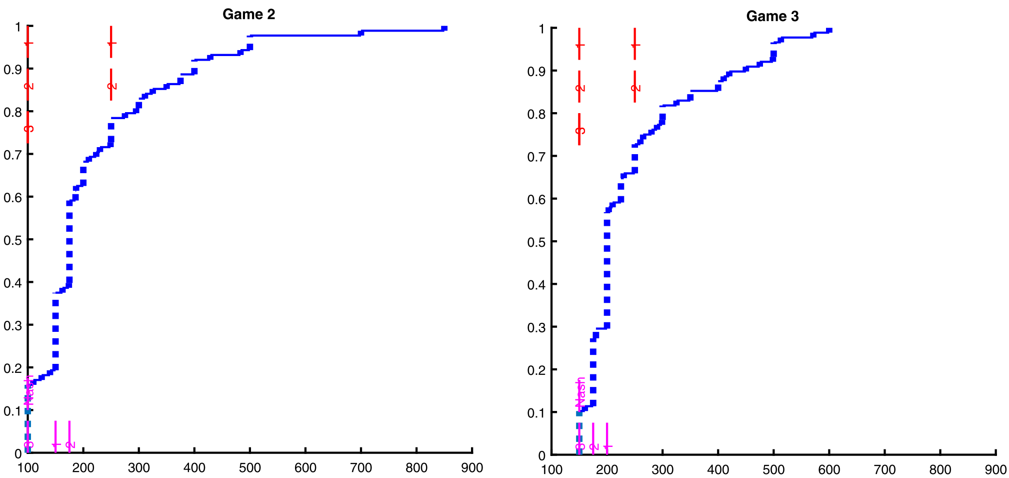
- For $k = \text{NE}$ and $s = 0_{\text{unanch}}$, notice that $c_{1g}(\text{NE}) = 100 = \alpha_{Lg}(1)$ is on the lower bound of $\Sigma^0_{1g} = [100, 900]$, so clearly $[c_{1g}(\text{NE}), c_{1g}(\text{NE}) + \overline{\rho}\Omega_{1g}]$ is a subset of Σ^0_{1g} for small enough $\overline{\rho}$.

- For $k = \text{NE}$ and $s = 1_{\text{unanch}}$, notice that $c_{1g}(\text{NE}) = 100 = \alpha_{Lg}(1)$ is on the lower bound of $\Sigma^1_{1g} = [100, 250]$, so clearly again $[c_{1g}(\text{NE}), c_{1g}(\text{NE}) + \overline{\rho}\Omega_{1g}]$ is a subset of Σ^1_{1g} for small enough $\overline{\rho}$.

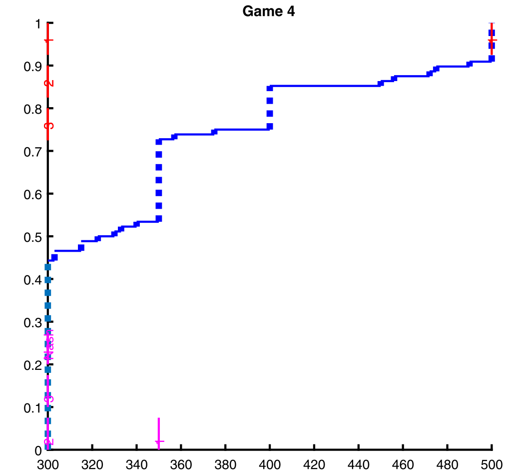
More generally, establishing Assumptions 4.1 and 4.2 can be accomplished by a computerized algorithm that takes as inputs the information in Table 1 and replicates the steps of verifying the assumptions just described. Further, note that verifying Assumption B.1 follows similar steps to verifying Assumption 4.1, since the assumptions are similar. Assumption B.2 rules out the described knife-edge situations, and is directly assumed by the econometrician. Finally, note that establishing these assumptions concerns the structure of the games, and, therefore, the experiment can be designed to ensure that the conditions are indeed satisfied before conducting the experiment.

Appendix F: Additional empirical cumulative distribution functions from the empirical application

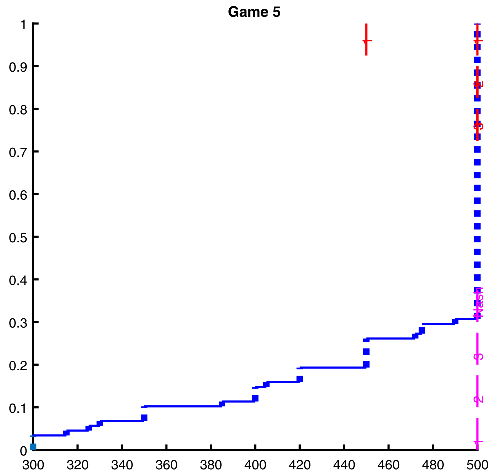
The following figures are empirical cumulative distribution functions of actions taken by subjects in games 2–16, as in Section 5.2.



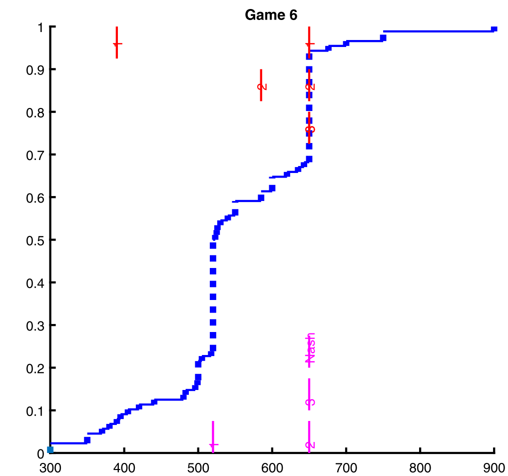
(a) Game 2 (b) Game 3



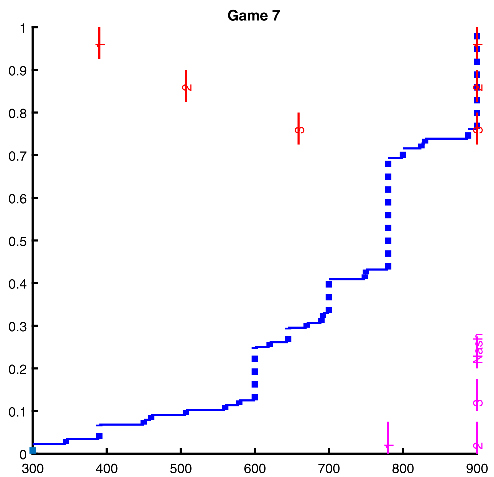
(c) Game 4



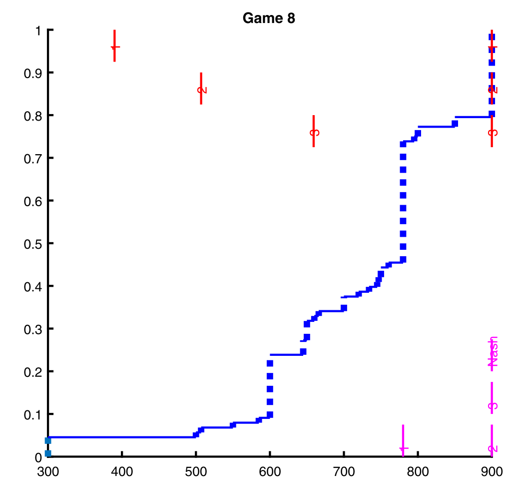
(d) Game 5



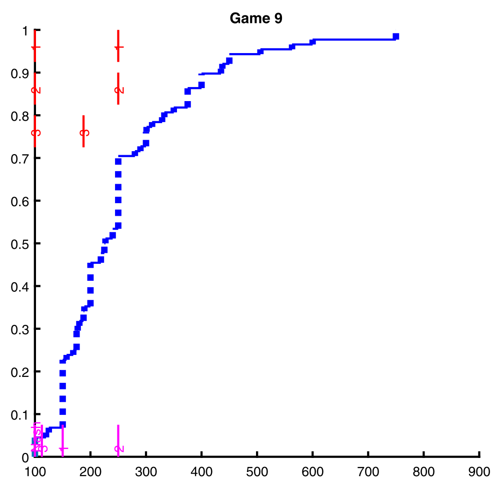
(e) Game 6



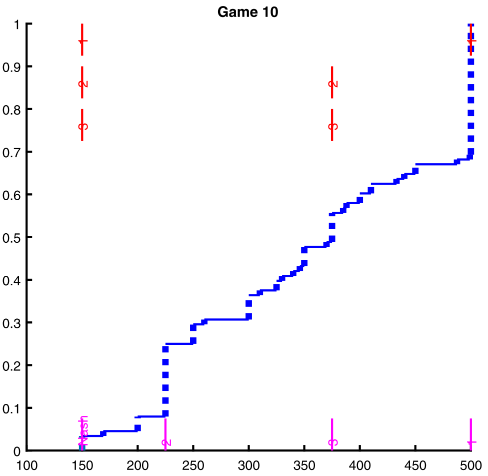
(f) Game 7



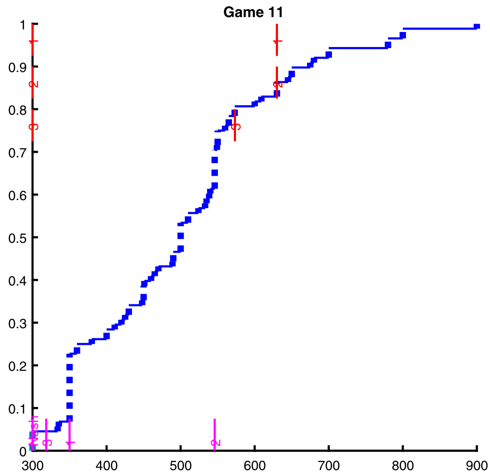
(g) Game 8



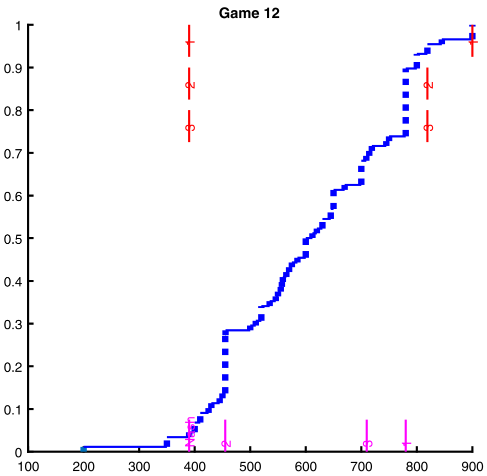
(h) Game 9



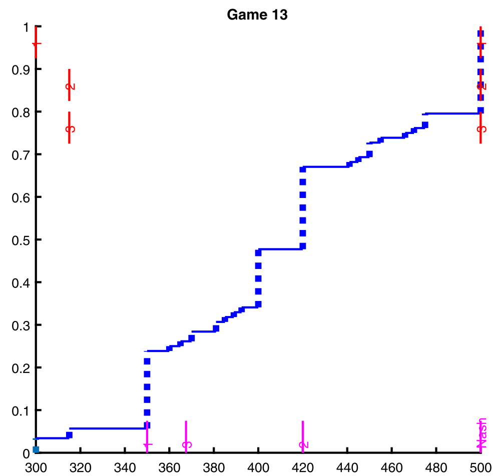
(i) Game 10



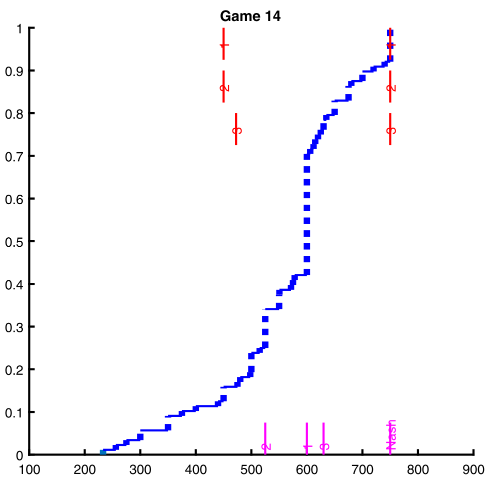
(j) Game 11



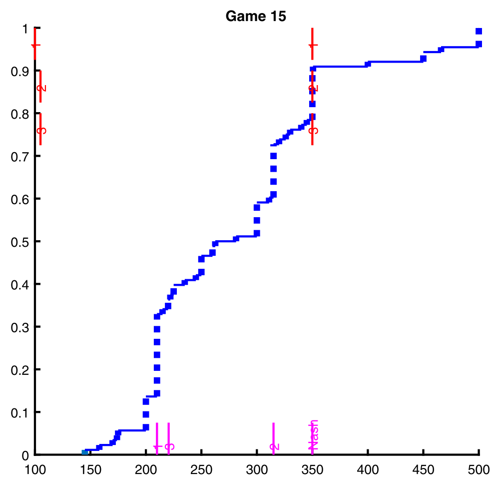
(k) Game 12



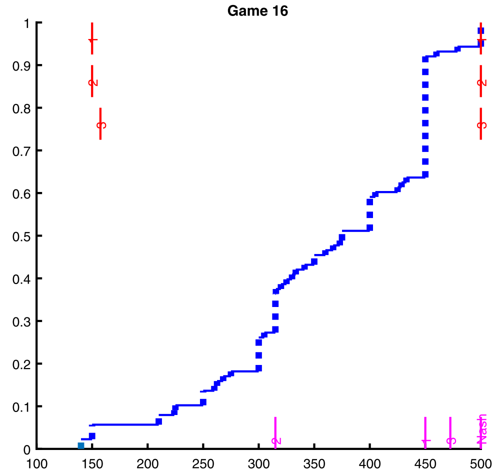
(l) Game 13



(m) Game 14



(n) Game 15



(o) Game 16

Appendix G: Estimates of the model allowing computational mistakes

Table 4 reports estimates of the model allowing uniformly distributed computational mistakes. It is assumed that $P_r = \frac{2.5}{200}$ for all rules r.[1] The results are almost identical to the model not allowing computational mistakes, and estimates of Δ_r are close to zero for all r.

Table 4. Estimates allowing computational mistakes.

	Λ					Probability of	
	Anchored Reasoning		Unanchored Reasoning			Type	Mistake
	1	2	0	1	Nash		
r	$\Lambda_r(1_{\text{anch}})$	$\Lambda_r(2_{\text{anch}})$	$\Lambda_r(0_{\text{unanch}})$	$\Lambda_r(1_{\text{unanch}})$	$\Lambda_r(\text{NE})$	$\pi(r)$	Δ_r
1	0.10 (0.08, 0.12)	0.04 (0.03, 0.06)	0.49 (0.37, 0.55)	0.31 (0.25, 0.42)	0.07 (0.03, 0.10)	0.44 (0.39, 0.55)	0.00 (0.00, 0.00)
2	0.70 (0.59, 0.77)	0.00 (0.00, 0.00)	0.15 (0.09, 0.25)	0.11 (0.06, 0.18)	0.04 (0.02, 0.06)	0.20 (0.14, 0.29)	0.00 (0.00, 0.00)
3	0.21 (0.04, 0.31)	0.44 (0.40, 0.79)	0.10 (0.00, 0.19)	0.20 (0.00, 0.32)	0.05 (0.00, 0.09)	0.15 (0.10, 0.23)	0.07 (0.00, 0.12)
4	0.05 (0.01, 0.08)	0.04 (0.00, 0.07)	0.05 (0.00, 0.08)	0.40 (0.33, 0.51)	0.46 (0.41, 0.60)	0.14 (0.08, 0.26)	0.00 (0.00, 0.00)
5	0.09 (0.00, 0.16)	0.89 (0.86, 1.00)	0.00 (0.00, 0.00)	0.02 (0.00, 0.04)	0.00 (0.00, 0.00)	0.06 (0.00, 0.08)	0.00 (0.00, 0.00)

Note: See notes to Table 3.

[1] Theorem 4.1 establishes that the magnitude of the computational mistakes P_r is identified if $\Delta_r > 0$. That is required because if $\Delta_r = 0$ for some rule r, then that rule does not make computational mistakes, so P_r has no observable implications. This is not a concern based on Theorem B.1, which applies when P_r are known by the econometrician. The estimates of Δ_rs are very close to 0, which makes identification and estimation of the corresponding P_rs very tenuous. Indeed, precisely because of the small values of the Δ_rs, the P_rs are essentially irrelevant.

Co-editor Rosa L. Matzkin handled this manuscript.

Manuscript received 8 December, 2015; final version accepted 18 January, 2017; available online 21 February, 2017.