Online Appendix for "Estimation and Inference in Games of Incomplete Information with Unobserved Heterogeneity and Large State Space"

Yanqin Fan* Shuo Jiang† Xuetao Shi‡

July 28, 2024

Abstract

This document provides proofs and additional details for the results in the authors' paper: "Estimation and Inference in Games of Incomplete Information with Unobserved Heterogeneity and Large State Space". Appendix B collects the proofs. Appendix C collects illustration of Xiao (2018)'s CCP estimator and discussion on Step-2 identification in the General Game. Appendix D provides more details on the games used in the simulation.

*Department of Economics, University of Washington; e-mail: fany88@uw.edu

†MOE Key Lab of Econometrics, WISE, Department of Statistics and Data Science at School of Economics, Xiamen University; e-mail: jiangshuo@xmu.edu.cn

‡School of Economics, University of Sydney; e-mail: xuetao.shi@sydney.edu.au

Appendix B Proofs

In this appendix, we provide proofs for the results in Fan et al. (2024).

B.1 Proofs of Results in Section 2

Proof of Lemma 2.1: Under Assumption 2.1-2.5, CCPs are identified up to a label swapping, which enables us to identify expected payoffs (via $F^{-1}(\cdot)$) and set up system: $G(\pi) \equiv \overline{\pi} - \Gamma\pi$. Under Assumption 2.6, the systems corresponding to $c \neq c_0$ has no solution. Note that system corresponding to c_0 always has a solution, as π_0 generates the system corresponding to c_0. Under the condition that Γ_{c_0} has full column rank, this π_0 is uniquely determined by $\overline{\pi}_{c_0}$ and Γ_{c_0}. □

B.2 Proofs of Results in Section 3

Lemma B.1 is used to prove Theorem 3.1. Lemmas B.2 and B.3 are introduced to prove Theorem 3.2.

Lemma B.1. *Let the result of Lemma 2.1 and the following assumptions hold for the Simple Game. (i) For any $\pi \in \Pi$, $G_n(\pi) = G(\pi) + O_p\left(n^{-1/2}\right)$. (ii) For any $c \in \mathscr{C}_n$, $W_n(c) = W(c) + o_p(1)$ with $W(c)$ being positive definite. (iii) For any $c \in \mathscr{C}_n$, $\min_{\pi\in\Pi} \|G_{n,c}(\pi)\|^2_{W_n(c)} \xrightarrow{p} \min_{\pi\in\Pi} \|G_c(\pi)\|^2_{W(c)}$. Then it holds that $\widehat{c} = c_0$ wp $\to 1$ and $\widehat{\pi} \xrightarrow{p} \pi_0$ for any $l_1 \in \{l_\pi, l_\pi + 1, .., l\}$, $\alpha_1 \in (0,1]$, $\lambda \in (-1,0)$, and $\Delta \in \{1, .., l - l_1\}$.*

Proof of Lemma B.1: For $s = 1, ..., S$, let $sc_0^s \in \mathscr{SC}^s$ denote the sub-selection vector whose first $2l_s$ elements are the same as those of c_0. We first show that $\Pr\left(sc_0^s \in \mathscr{SC}_n^s\right) \to 1$ for $s = 1, ..., S$. This implies that $\Pr\left(c_0 \in \mathscr{C}_n\right) \to 1$, because $c_0 \in \mathscr{C}_n$ occurs if and only if $sc_0^S \in \mathscr{SC}_n^S$ occurs.

We first prove the result for $s = 1$. By Lemma 2.1, we have that $G_{c_0}(\pi_0) = \mathbf{0}$. Together with Assumption (i) in the lemma, we can obtain that $\|G_{n,c_0}(\pi_0)\|^2 = O_p\left(n^{-1}\right)$ and $\left\|G_{n,sc_0^1}(\pi_0)\right\|^2 = O_p\left(n^{-1}\right)$. Therefore, it holds that $J_n\left(sc_0^1\right) \leq \left\|G_{n,sc_0^1}(\pi_0)\right\|^2 = O_p\left(n^{-1}\right)$. For any $sc^1 \in \mathscr{SC}^1$, if $J_n\left(sc^1\right) \leq n^\lambda$, then $sc^1 \in \mathscr{SC}_n^1$ occurs. And for any $\alpha_1 > 0$, $\mathscr{SC}_n^1$ can contain sc^1's such that $J_n\left(sc^1\right) > n^\lambda$. Thus, we obtain that for any $\lambda > -1$,

$$1 \geq \Pr\left(sc_0^1 \in \mathscr{SC}_n^1\right) \geq \Pr\left(J_n\left(sc_0^1\right) \leq n^\lambda\right) \to 1.$$

Hence, we obtain that $\Pr\left(sc_0^1 \in \mathscr{SC}_n^1\right) \to 1$.

$\Pr\left(sc_0^2 \in \mathscr{SC}_n^2 \mid sc_0^2 \in \mathscr{SC}^2\right) \to 1$ and $\Pr\left(sc_0^2 \in \mathscr{SC}^2\right) \to 1$ together imply that $\Pr\left(sc_0^2 \in \mathscr{SC}_n^2\right) \to 1$. The convergence of $\Pr\left(sc_0^2 \in \mathscr{SC}_n^2 \mid sc_0^2 \in \mathscr{SC}^2\right) \to 1$ follows from the same argument as $\Pr\left(sc_0^1 \in \mathscr{SC}_n^1\right) \to 1$; and $sc_0^1 \in \mathscr{SC}_n^1$ implies that $sc_0^2 \in \mathscr{SC}^2$. Thus, we have that $\Pr\left(sc_0^2 \in \mathscr{SC}_n^2\right) \to 1$. Applying the same argument sequentially, we can obtain that $\Pr\left(sc_0^s \in \mathscr{SC}_n^s\right) \to 1$ for $s = 1, ..., S$. By the definition of $\mathscr{C}_n$, elements in $\mathscr{C}_n$ select all possible combinations of the last $2\left(l - l_S\right)$ moments allowed by $\mathscr{C}$. The event $sc_0^S \in \mathscr{SC}_n^S$ occurs if and only if $c_0 \in \mathscr{C}_n$ occurs. Thus, we have that $\Pr\left(c_0 \in \mathscr{C}_n\right) \to 1$.

Next, we show that if $c_0 \in \mathscr{C}_n$, then $\Pr\left(\widehat{c} = c_0\right) \to 1$. For any $c \in \mathscr{C}_n$ and $c \neq c_0$, Assumption (iii) implies that

$$\min_{\pi \in \Pi} \left\|G_{n,c}\left(\pi\right)\right\|_{W_n(c)}^2 \xrightarrow{p} \min_{\pi \in \Pi} \left\|G_c\left(\pi\right)\right\|_{W(c)}^2 > 0, \tag{B.1}$$

where the inequality follows from Lemma 2.1 and Assumption (ii). Again by Lemma 2.1, we have that $G_{c_0}\left(\pi_0\right) = \mathbf{0}$. Then by Assumption (i), it holds that $G_{n,c_0}\left(\pi_0\right) = o_p\left(1\right)$. Therefore, we obtain that

$$\min_{\pi \in \Pi} \left\|G_{n,c_0}\left(\pi\right)\right\|_{W_n(c)}^2 \leq \left\|G_{n,c_0}\left(\pi_0\right)\right\|_{W_n(c)}^2 = o_p\left(1\right). \tag{B.2}$$

(B.1) and (B.2) together imply that $\Pr\left(\widehat{c} = c_0 \mid c_0 \in \mathscr{C}_n\right) \to 1$. Combining with $\Pr\left(c_0 \in \mathscr{C}_n\right) \to 1$, we conclude that $\Pr\left(\widehat{c} = c_0\right) \to 1$.

It remains to show that $\widehat{\pi} \xrightarrow{p} \pi_0$. Define $\breve{\pi} \equiv \arg\min_{\pi \in \Pi} \left\|G_{n,c_0}\left(\pi\right)\right\|_{W_n(c_0)}^2$. We have $\breve{\pi}$ converges in probability to its population counterpart π_0 by the standard argument for consistency of a GMM estimator provided that Lemma 2.1 holds. Thus for any $\epsilon > 0$ and $\delta > 0$, we can find N_1 such that $\Pr\left(\left\|\breve{\pi} - \pi_0\right\| > \frac{\epsilon}{2}\right) < \frac{\delta}{2}$ for $n \geq N_1$. Because $\widehat{c} = c_0$ implies $\widehat{\pi} = \breve{\pi}$, we have $\Pr\left(\widehat{\pi} = \breve{\pi}\right) \geq \Pr\left(\widehat{c} = c_0\right) \to 1$. Thus for the given $\epsilon > 0$ and $\delta > 0$, we can find N_2 such that $\Pr\left(\left\|\widehat{\pi} - \breve{\pi}\right\| > \frac{\epsilon}{2}\right) < \frac{\delta}{2}$ for $n \geq N_2$. Combing the above two results, we have that for any given $\epsilon > 0$ and $\delta > 0$, there exists $N \equiv \max\left\{N_1, N_2\right\}$ such that for $n > N$,

$$\begin{aligned} & \Pr\left(\left\|\widehat{\pi} - \pi_0\right\| > \epsilon\right) = \Pr\left(\left\|\widehat{\pi} - \breve{\pi} + \breve{\pi} - \pi_0\right\| > \epsilon\right) \\ \leq \; & \Pr\left(\left\|\widehat{\pi} - \breve{\pi}\right\| > \frac{\epsilon}{2}\right) + \Pr\left(\left\|\breve{\pi} - \pi_0\right\| > \frac{\epsilon}{2}\right) \leq \delta. \end{aligned}$$

Therefore, the result $\widehat{\pi} \xrightarrow{p} \pi_0$ holds. □

Proof of Theorem 3.1: We prove the theorem by verifying the conditions in Lemma B.1. Lemma 2.1 holds by Assumptions 2.1-2.6. Assumption (ii) in Lemma

B.1 is implied by Assumption 3.2. It remains to the show that Assumptions (i) and (iii) in Lemma B.1 are satisfied.

Let $\mathbf{p}$ be the vector that stores the CCPs of all players on all observed and latent states ($\mathbf{z} \in \left\{\mathbf{z}^1, \ldots, \mathbf{z}^l\right\}$ and $k \in \{A, B\}$) and $\widehat{\mathbf{p}}$ be its consistent estimator obtained via the eigendecomposition method. Because the estimated coefficient matrix in the Simple Game is a smooth function of $\widehat{\mathbf{p}}$, we write it in the form of $\Gamma_n \equiv \mathbf{e}\left(\widehat{\mathbf{p}}\right)$ for some smooth function $\mathbf{e}\left(\cdot\right)$. Similarly, the estimated expected payoff vector can be written in the form of $\overline{\pi}_n \equiv \mathbf{F}^{-1}\left(\widehat{\mathbf{p}}\right)$, where $\mathbf{F}^{-1}\left(\widehat{\mathbf{p}}\right)$ stacks $F^{-1}\left(p_{1\mathbf{z}k}\right)$ for any $\mathbf{z} \in \left\{\mathbf{z}^1, \ldots, \mathbf{z}^l\right\}$ and $k \in \{A, B\}$. Thus the sample moment functions can be written as $G_n\left(\pi\right) = \mathbf{F}^{-1}\left(\widehat{\mathbf{p}}\right) - \mathbf{e}\left(\widehat{\mathbf{p}}\right)\pi$. By differentiability of $\mathbf{F}^{-1}\left(\cdot\right)$ and $\mathbf{e}\left(\cdot\right)$, given any π, the mean value expansion of $G_n\left(\pi\right)$ with respect to $\mathbf{p}$ gives

$$\sqrt{n}\left(G_n\left(\pi\right) - G\left(\pi\right)\right) = \sqrt{n} D_{\mathbf{p}^*}\left(\pi\right)\left(\widehat{\mathbf{p}} - \mathbf{p}\right),$$

where $D_{\mathbf{p}^*}\left(\pi\right) = \frac{\partial \mathbf{F}^{-1}\left(\mathbf{p}^*\right)}{\partial \mathbf{p}'} - \frac{\partial\left(\mathbf{e}\left(\mathbf{p}^*\right)\pi\right)}{\partial \mathbf{p}'}$ and $\mathbf{p}^*$ is a point between $\widehat{\mathbf{p}}$ and $\mathbf{p}$. Under Assumptions 2.1-2.6, Lemma C.2 in Xiao (2018) implies that that $\widehat{\mathbf{p}}$ is consistent and $\sqrt{n}\left(\widehat{\mathbf{p}} - \mathbf{p}\right)$ converges to a normal distribution. Under Assumption 2.1 (i) and (ii), we have that $D_{\mathbf{p}^*}\left(\pi\right) \xrightarrow{p} D_{\mathbf{p}}\left(\pi\right) \equiv \frac{\partial \mathbf{F}^{-1}\left(\mathbf{p}\right)}{\partial \mathbf{p}'} - \frac{\partial\left(\mathbf{e}\left(\mathbf{p}\right)\pi\right)}{\partial \mathbf{p}'}$, where $D_{\mathbf{p}}$ is bounded for any $\pi \in \Pi$. Therefore, it holds that $\sqrt{n}\left(G_n\left(\pi\right) - G\left(\pi\right)\right) = O_p\left(1\right)$, which verifies Assumption (i) in Lemma B.1.

For Assumption (iii) in Lemma B.1, it suffices to prove that for any $c \in \mathscr{C}_n$

$$\max_{\pi \in \Pi}\left|G_{n,c}\left(\pi\right)^\top W_n\left(c\right) G_{n,c}\left(\pi\right) - G_c\left(\pi\right)^\top W\left(c\right) G_c\left(\pi\right)\right| = o_p\left(1\right),$$

where max is used rather than sup because Π is compact by Assumption 3.1 and both $G_{n,c}\left(\cdot\right)$ and $G_c\left(\cdot\right)$ are continuous in π. The triangular inequality provides that

$$\begin{aligned} & \max_{\pi \in \Pi}\left|G_{n,c}\left(\pi\right)^\top W_n\left(c\right) G_{n,c}\left(\pi\right) - G_c\left(\pi\right)^\top W\left(c\right) G_c\left(\pi\right)\right| \\ \leq & \max_{\pi \in \Pi}\left|G_{n,c}\left(\pi\right)^\top W_n\left(c\right) G_{n,c}\left(\pi\right) - G_{nc}\left(\pi\right)^\top W\left(c\right) G_{nc}\left(\pi\right)\right| \\ & + \max_{\pi \in \Pi}\left|G_{n,c}\left(\pi\right)^\top W\left(c\right) G_{n,c}\left(\pi\right) - G_c\left(\pi\right)^\top W\left(c\right) G_c\left(\pi\right)\right|. \end{aligned}$$

By the property of the matrix infinity norm,[1] the first term on the right hand side of the inequality is bounded above by $\max_{\pi \in \Pi}\left\|c\right\|_0\left\|G_{n,c}\left(\pi\right)\right\|^2\left\|W_n\left(c\right) - W\left(c\right)\right\|_\infty$. Π is bounded. Moreover, Γ_n and $\overline{\pi}_n$ are both $O_p\left(1\right)$ by the continuity of $\mathbf{e}\left(\cdot\right)$ and

[1]For a generic matrix E, $\left\|E\right\|_\infty$ denotes the matrix infinity norm that equals the maximum absolute row sum of matrix E.

$\mathbf{F}^{-1}(\cdot)$ and the consistency of $\widehat{\mathbf{p}}$. Therefore, we have $\max_{\pi\in\Pi} \|G_{n,c}(\pi)\|^2 = O_p(1)$. By Assumption 3.2, it holds that $\|W_n(c) - W(c)\|_\infty = o_p(1)$. Thus, the first term is $o_p(1)$. For the second term, it holds that

$$\begin{aligned}
& \max_{\pi\in\Pi} \left| G_{n,c}(\pi)^\top W(c) G_{n,c}(\pi) - G_c(\pi)^\top W(c) G_c(\pi) \right| \\
= & \max_{\pi\in\Pi} \left| G_{n,c}(\pi)^\top W(c) \left(G_{n,c}(\pi) - G_c(\pi) + G_c(\pi)\right) - G_c(\pi)^\top W(c) G_c(\pi) \right| \\
\leq & \max_{\pi\in\Pi} \left| G_{n,c}(\pi)^\top W(c) \left(G_{n,c}(\pi) - G_c(\pi)\right) \right| \\
& + \max_{\pi\in\Pi} \left| G_{n,c}(\pi)^\top W(c) G_c(\pi) - G_c(\pi)^\top W(c) G_c(\pi) \right| \\
\leq & \max_{\pi\in\Pi} \left\| G_{n,c}(\pi)^\top W(c) \right\|_\infty \max_{\pi\in\Pi} \|G_{n,c}(\pi) - G_c(\pi)\| \\
& + \max_{\pi\in\Pi} \|G_{n,c}(\pi) - G_c(\pi)\| \max_{\pi\in\Pi} \|W(c) G_c(\pi)\|_\infty \\
= & o_p(1),
\end{aligned}$$

where the first and second inequality follow from the triangular inequality and applying matrix norm, and last equality holds by Γ_n and $\overline{\pi}_n$ being $O_p(1)$ and the compactness of Π.

Hence, we have verified all the conditions in Lemma B.1. The claimed theorem holds. □

Lemma B.2. *Let $\alpha_s^\star \equiv |\mathscr{SC}_n^s| / |\mathscr{SC}^s|$. Given l_1, α_1, λ, and Δ, assume that there exists some $s^\dagger \in \{2, .., S\}$ independent of l, such that $\alpha_s^\star = \alpha$ for all $s = s^\dagger, .., S$. Then both the time complexity and space complexities of the MMS procedure are linear in l.*

Proof of Lemma B.2: We show the time complexity result by counting the elementary operations (EOs). Note that the values of l_1, α_1, λ, and Δ are independent of l. In Step s, the algorithm computes $J_n(sc^s)$ for every $sc^s \in \mathscr{SC}^s$; and for each sc^s, computing $J_n(sc^s)$ is a quadratic programming problem with fixed number of unknowns and constraint (from Π). By definition, $\|G_{n,sc^s}(\cdot)\|^2 = \|G_{n,sc^{s-1}}(\cdot)\|^2 + \left\|G_{n,sc^{s\backslash(s-1)}}(\cdot)\right\|^2$, where $sc^{s\backslash(s-1)}$ is a sub-selection vector that selects the moments selected by sc^s but not sc^{s-1}. Because the function $\|G_{n,sc^{s-1}}(\cdot)\|^2$ is stored in Step $(s-1)$, the number of EOs in computing $\|G_{n,sc^s}(\cdot)\|^2$ only depends on that in computing $\left\|G_{n,sc^{s\backslash(s-1)}}(\cdot)\right\|^2$, which depends solely on Δ. Therefore, the number of EOs for computing $J_n(sc^s)$ for every $sc^s \in \mathscr{SC}^s$ is $\beta\,|\mathscr{SC}^s|$, where $\beta < \infty$ is a constant independent of l. Next, the algorithm sorts $J_n(sc^s)$. The sorting process involves $|\mathscr{SC}^s|^2$ number of EOs. The comparison between $J_n(sc^s)$ and n^λ in each step takes place $|\mathscr{SC}^s|$ times at most. Thus, there are $(\beta+1)|\mathscr{SC}^s|+|\mathscr{SC}^s|^2$ number of EOs in Step s for $s=1,...,S$. The number of EOs in Step $(S+1)$ can be computed in the same way, except that the algorithm only searches for the minimum of $|\mathscr{C}_n|$ values rather than sorts them. Thus, $(\beta+1)|\mathscr{C}_n|$ number of EOs are needed. In total, there are $\sum_{s=1}^{S}\left(\gamma|\mathscr{SC}^s|+|\mathscr{SC}^s|^2\right)+\gamma|\mathscr{C}_n|$ number of EOs, with $\gamma\equiv\beta+1$.

In Step s for $s=1,...,S$, the input set $\mathscr{SC}^s$ has cardinality $2^{l_s-1}\prod_{i=1}^{s-1}\alpha_i^\star$. For $s=1,..,S$, $l_s=l_1+(s-1)\Delta$. We have $\left|\mathscr{SC}^1\right|=2^{l_1-1}$ and $|\mathscr{SC}^s|=2^{l_s-1}\prod_{i=1}^{s-1}\alpha_i^\star\leq 2^{l_1+(s-1)\Delta-1}$ for $s=2,...,s^\dagger$. Therefore,

$$\sum_{s=1}^{s^\dagger}\left(\gamma|\mathscr{SC}^s|+|\mathscr{SC}^s|^2\right)\leq\sum_{s=1}^{s^\dagger}\left(\gamma 2^{l_1+(s-1)\Delta-1}+4^{l_1+(s-1)\Delta-1}\right)\equiv T_1.$$

For $s=s^\dagger,..,S$, $\alpha_s^\star=\alpha=2^{-\Delta}$. Thus, for $s=s^\dagger+1,...,S$, it holds that $|\mathscr{SC}^s|\leq 2^{l_s-1}\prod_{i=s^\dagger}^{s-1}\alpha_i^\star=2^{l_1+(s-1)\Delta-1}\left(2^{-\Delta}\right)^{\left(s-s^\dagger\right)}=2^{l_1+\left(s^\dagger-1\right)\Delta-1}$. As a result, we have that

$$\begin{aligned}\sum_{s=s^\dagger+1}^{S}\left(\gamma|\mathscr{SC}^s|+|\mathscr{SC}^s|^2\right)&\leq\sum_{s=s^\dagger+1}^{S}\left(\gamma 2^{l_1+\left(s^\dagger-1\right)\Delta-1}+4^{l_1+\left(s^\dagger-1\right)\Delta-1}\right)\\&\leq\left(S-s^\dagger\right)\left(\gamma 2^{l_1+\left(s^\dagger-1\right)\Delta-1}+4^{l_1+\left(s^\dagger-1\right)\Delta-1}\right)\equiv T_2,\end{aligned}$$

where the second inequality holds because the summands are independent of s. The input set $\mathscr{C}_n$ in Step $(S+1)$ has cardinality $2^{l-1}\prod_{i=1}^{S}\alpha_i^\star$. Let $\Delta_f\equiv l-[l_1+(S-1)\Delta]$. Because $\alpha_s^\star=\alpha=2^{-\Delta}$ for $s=s^\dagger,..,S$, $2^{l-1}\prod_{i=1}^{S}\alpha_i^\star\leq 2^{l_1+\left(s^\dagger-2\right)\Delta+\Delta_f-1}$. We obtain that

$$\gamma|\mathscr{C}_n|\leq\gamma 2^{l_1+\left(s^\dagger-2\right)\Delta+\Delta_f-1}\leq\gamma 2^{l_1+\left(s^\dagger-1\right)\Delta-1}\equiv T_3,$$

where the inequality holds because $\Delta_f\leq\Delta$. The number of EOs is no more than $T_1+T_2+T_3$. Since l_π, $l_1<l$, and Δ are fixed, T_1 and T_3 are constants of l. In fact, only S depends on l. Because $S<\frac{l-l_1}{\Delta}+1$ by definition, we obtain that

$$T_2<\left(\frac{l-l_1}{\Delta}+1-s^\dagger\right)\left(\gamma 2^{l_1+\left(s^\dagger-1\right)\Delta-1}+4^{l_1+\left(s^\dagger-1\right)\Delta-1}\right).$$

Since $s^\dagger$ is a constant by assumption, T_2 grows at most linearly in l. Hence, we have shown that the number of EOs needed to perform the MMS procedure is linear in l. The first claim of the lemma follows.

Since in each step, the algorithm stores $|\mathscr{SC}^s|$ number of values, the space complexity is $\sum_{s=1}^{S}|\mathscr{SC}^s|$. Following from the similar calculation as above, we obtain that the space complexity of the MMS procedure is also linear in l. □

Lemma B.3. *Define $\mathcal{Q} \equiv [0,1]^{6l}$ as the set of CCPs constituting the moment functions $G(\cdot)$. Let Assumptions 2.1-2.6 and 3.1 hold. With probability approaching one as $n \to \infty$, for all possible CCPs in $\mathcal{Q}$ except for a set of Lebesgue measure zero, both the time complexity and space complexity of the MMS procedure are linear in l.*

Proof of Lemma B.3: We prove the lemma by showing that the assumption in Lemma B.2 holds with probability approaching one as $n \to \infty$. The assumption holds if $J_n^\alpha > n^\lambda$ at Step s for every $s = s^\dagger, .., S$, because $\alpha_s^\star = \alpha$ whenever $J_n^\alpha > n^\lambda$. Since $\lambda < 0$, $n^\lambda \to 0$ as $n \to \infty$. For any $sc \in \mathbb{R}^{2l}$ consisting of zeros and ones, let $G_{sc}(\cdot)$ denote the moment functions selected by sc. The proof for Theorem 3.1 shows that under Assumptions 2.1-2.6 and 3.1, Assumption (iii) in Lemma B.1 holds. Following the similar argument, it holds that $J_n(sc) = \min_{\pi\in\Pi} \|G_{n,sc}(\pi)\|^2 \xrightarrow{p} \min_{\pi\in\Pi} \|G_{sc}(\pi)\|^2$. Therefore, if $\min_{\pi\in\Pi} \|G_{sc}(\pi)\|^2 \neq 0$, then $\Pr\left(J_n(sc) > n^\lambda\right) \to 1$ as $n \to \infty$.

For any $sc^s \in \mathscr{SC}^s$, $G_{sc^s}(\pi)$ contains l_s number of moments. By (2.10) in Fan et al. (2024), $G_{sc^s}(\pi) \equiv \overline{\pi}_{sc^s} - \Gamma_{sc^s}\pi$, where $\overline{\pi}_{sc^s} \in \mathbb{R}^{l_s}$, $\Gamma_{sc^s} \in \mathbb{R}^{l_s\times l_\pi}$, and $l_\pi = 4$. $\min_{\pi\in\Pi} \|G_{sc^s}(\pi)\|^2 \neq 0$ if and only if $\operatorname{rank}([\overline{\pi}_{sc^s}, \Gamma_{sc^s}]) > \operatorname{rank}(\Gamma_{sc^s})$. Define $\mathcal{P} \equiv \left\{\mathfrak{p} : \mathfrak{p} \in [0,1]^{2l_s}\right\}$ as the set of CCPs constituting Γ_{sc^s} defined in (2.9). Elements in Γ_{sc^s} can be presented as a map $M : \mathcal{P} \to [0,1]^{l_s\times l_\pi}$. Define another map $R(\mathfrak{p}) \equiv \operatorname{rank}(M(\mathfrak{p}))$ for every $\mathfrak{p} \in \mathcal{P}$. We aim to show that at any Step s such that $l_s > l_\pi$, the set $\mathcal{P}\backslash\mathcal{P}_r$, where $\mathcal{P}_r \equiv \{\mathfrak{p} \in \mathcal{P} : R(\mathfrak{p}) = l_\pi\}$, has Lebesgue measure zero in $\mathcal{P}$.

Partition $\Gamma_{sc^s}^\top$ as $[\Gamma_{sc^s}^1, \Gamma_{sc^s}^2]$, where $\Gamma_{sc^s}^1 \in \mathbb{R}^{l_\pi\times l_\pi}$ and $\Gamma_{sc^s}^2 \in \mathbb{R}^{l_\pi\times(l_s-l_\pi)}$. Define a set $\mathcal{P}^1 \equiv \left\{\mathfrak{p}^1 : \mathfrak{p}^1 \in [0,1]^{2l_\pi}\right\}$. Similar to Γ_{sc^s}, $\Gamma_{sc^s}^1$ can be presented as a map: $M^1 : \mathcal{P}^1 \to [0,1]^{l_\pi\times l_\pi}$. Let $\det^\dagger(\mathfrak{p}^1) \equiv \det(M^1(\mathfrak{p}^1))$ for every $\mathfrak{p}^1 \in \mathcal{P}^1$. For $\mathcal{P}_r^1 \equiv \left\{\mathfrak{p}^1 \in \mathcal{P}^1 : \det^\dagger(\mathfrak{p}^1) \neq 0\right\}$, we first to show that the set $\mathcal{P}^1\backslash\mathcal{P}_r^1$ has Lebesgue measure zero in $\mathcal{P}^1$. By the Leibniz formula, $\det^\dagger(\mathfrak{p}^1)$ defines a polynomial function on $[0,1]^{2l_\pi}$. We have that $\det^\dagger(\mathfrak{p}^{1*}) \neq 0$ for some

$$\mathfrak{p}^{1*} \equiv \left[p_2\left(\mathbf{z}^1, k_1\right), p_3\left(\mathbf{z}^1, k_1\right), \ldots, p_2\left(\mathbf{z}^{l_\pi}, k_{l_\pi}\right), p_3\left(\mathbf{z}^{l_\pi}, k_{l_\pi}\right)\right]^\top,$$

where $k_1, \ldots, k_{l_\pi}$ are the latent states selected by sc^s. In algebraic geometry, the set $\mathcal{P}^1\backslash\mathcal{P}_r^1$ is defined as a proper subvariety, and must be of Lebesgue measure zero in $\mathcal{P}^1$. For more discussion on the algebraic geometry, especially the algebraic variety, see Cox et al. (2013). In fact, $\mathcal{P}_r^1$ is generic in $\mathcal{P}^1$. We explore the genericity result for the rest of the proof.

By the property of generic sets, it holds that

$$\mathcal{P}_r^1 \times [0,1]^{2(l_s - l_\pi)} \equiv \left\{ \left(\mathfrak{p}^1, \mathfrak{p}^2\right) : \mathfrak{p}^1 \in \mathcal{P}_r^1, \mathfrak{p}^2 \in [0,1]^{2(l_s - l_\pi)} \right\}$$

is generic in $\mathcal{P}$. As a result, $\mathcal{P}_t \equiv \left\{ (\mathfrak{p}^1, \mathfrak{p}^2) \in [0,1]^{2l_s} : \det^\dagger (\mathfrak{p}^1) \neq 0 \right\}$ is generic in $\mathcal{P}$. Since $\det(\Gamma_{sc^s}^1) \neq 0$ is equivalent to $\text{rank}(\Gamma_{sc^s}^1) = l_\pi$, $\text{rank}(\Gamma_{sc^s}) = l_\pi$ holds if $\det(\Gamma_{sc^s}^1) \neq 0$. We have that $\mathcal{P}_t \subset \mathcal{P}_r$. Hence, $\mathcal{P}_r$ is generic in $\mathcal{P}$, which implies that $\mathcal{P} \backslash \mathcal{P}_r$ has Lebesgue measure zero in $\mathcal{P}$.

Next, we show that the set of CCPs constituting $G_{sc^s}(\cdot)$ such that $\text{rank}([\overline{\pi}_{sc^s}, \Gamma_{sc^s}]) = l_\pi + 1$ is generic in $\mathcal{Q}_s \equiv \left\{ \mathfrak{q} : \mathfrak{q} \in [0,1]^{3l_s} \right\}$. Assumption 2.1 assumes that $F : \mathbb{R} \to [0,1]$ is absolutely continuous, where $F(\cdot)$ is the distribution function. A function is said to have the Luzin's property if the image of any Lebesgue null set has again Lebesgue measure zero. By Chapter 7 in Saks (1937), for functions with bounded variation, absolute continuity is equivalent to Luzin's property. In consequence, it suffices to show that the set

$$\mathcal{PE}_r \equiv \left\{ (\pi, \mathfrak{p}) : \pi \in \mathbb{R}^{l_s}, \mathfrak{p} \in [0,1]^{2l_s}, \text{rank}([\pi, M(\mathfrak{p})]) = l_\pi + 1 \right\}$$

is generic in $\mathcal{PE} \equiv \left\{ (\pi, \mathfrak{p}) : \pi \in \mathbb{R}^{l_s}, \mathfrak{p} \in [0,1]^{2l_s} \right\}$. The proof follows the same argument as proving that $\mathcal{P}_r$ is generic in $\mathcal{P}$. Thus, we obtain that the set of CCPs such that $\text{rank}([\overline{\pi}_{sc^s}, \Gamma_{sc^s}]) > \text{rank}(\Gamma_{sc^s})$ is generic in $\mathcal{Q}_s$. Because $\text{rank}([\overline{\pi}_{sc^s}, \Gamma_{sc^s}]) > \text{rank}(\Gamma_{sc^s})$ is equivalent to $\min_{\pi \in \Pi} \|G_{sc^s}(\pi)\|^2 \neq 0$, and $\Pr\left(J_n(sc^s) > n^\lambda\right) \to 1$ as $n \to \infty$ if $\min_{\pi \in \Pi} \|G_{sc^s}(\pi)\|^2 \neq 0$ for any $sc^s \in \mathscr{SC}^s$ and $sc^s \neq sc_0^s$, we conclude that the set of CCPs such that $\alpha_s^\star \neq \alpha$ wp $\to 1$ as $n \to \infty$ for any Step s with $s = s^\dagger, .., S$ has Lebesgue measure zero in $\mathcal{Q}$, the set of all possible CCPs constituting $G(\cdot)$. $\square$

Proof of Theorem 3.2: Define Π^I as the parameter space of player 1's payoffs on z_1^1 and some latent state such that Assumptions 2.1-2.6 hold. Since the space of payoffs for all three players is a Cartesian product of spaces for each player on each observed and latent states, it suffices to show that player 1's "exceptional" payoffs have zero Lebesgue measure in Π^I. Recall the definition of $\mathcal{Q}$ in Lemma B.3. Define $\mathcal{Q}^I \subseteq \mathcal{Q}$ such that elements in $\mathcal{Q}^I$ are CCPs in the Simple Game for all three players holding $z_1 = z_1^1$ and that Assumptions 2.1-2.6 are satisfied. By definition, elements in $\mathcal{Q}^I$ can identify the true payoffs of player 1 under the correct matching. Lemma B.3 implies that for any s with $l_s > l_\pi$ and $sc^s \neq sc_0^s$, the set of CCPs such that $\text{rank}([\overline{\pi}_{sc^s}, \Gamma_{sc^s}]) > \text{rank}(\Gamma_{sc^s})$ does not hold has measure zero in $\mathcal{Q}$. Since $\mathcal{Q}^I \subseteq \mathcal{Q}$ has non-zero Lebesgue measure, the set of such CCPs also has measure zero in $\mathcal{Q}^I$.

Define a map $\mathcal{G}:\mathcal{Q}^I\rightarrow\Pi^I$ that maps the CCPs to the true payoff. The map exists because elements in $\mathcal{Q}^I$ identify the true payoff. By Lemma 7.25 in Walter (1987), Luzin's property holds if the mapping $\mathcal{G}(\mathfrak{p})$ is differentiable in $\mathfrak{p}$. Therefore, it suffices to prove that $\mathcal{G}(\mathfrak{p})$ is differentiable in $\mathfrak{p}$. The map $\mathcal{G}$ is the solution to $\overline{\pi}_{c_0}-\Gamma_{c_0}\pi=\mathbf{0}$ by Lemma 2.1, where both $\overline{\pi}_{c_0}$ and Γ_{c_0} are functions of $\mathfrak{p}$. In consequence, $\mathcal{G}(\mathfrak{p})=\Gamma_{c_0}^{+}\overline{\pi}_{c_0}$, where $\Gamma_{c_0}^{+}$ is the Moore-Penrose pseudo-inverse of Γ_{c_0}. We show that $\Gamma_{c_0}^{+}$ and $\overline{\pi}_{c_0}$ are both differentiable in $p_i(\mathbf{z},k)$ for each i, $\mathbf{z}\in\mathscr{Z}$, and $k\in\mathscr{K}$. Without loss of generality, we can just focus on $p_1(\mathbf{z}^1,A)$. Because $F(\cdot)$ is continuously differentiable and $f(\cdot)$ is positive everywhere in $\mathbb{R}$, the inverse function theorem provides that $F^{-1}(\cdot)$ is continuously differentiable in $p_1(\mathbf{z}^1,A)$ on $[0,1]$. Thus, $\overline{\pi}_{c_0}$ is differentiable in $p_1(\mathbf{z}^1,A)$. Since Γ_{c_0} has full column rank, $\Gamma_{c_0}^{+}=\left(\Gamma_{c_0}^{\top}\Gamma_{c_0}\right)^{-1}\Gamma_{c_0}^{\top}$. By the definition of Γ in (2.9) in Fan et al. (2024), Γ_{c_0} is differentiable in $p_1(\mathbf{z}^1,A)$. Hence, $\mathcal{G}(\mathfrak{p})$ is differentiable in $\mathfrak{p}$. The claimed result in the theorem holds. □

B.3 Proofs of Results in Section 4

Theorem 4.1 is proved based upon Lemmas B.4-B.7.

Lemma B.4. *Let Assumptions 2.1-2.6 and the following assumptions hold for the Simple Game. For any parameter sequence $\xi_n\in\Xi_R(\xi)$ with any $\xi\in\Xi_R$: (i) $\sqrt{n}\left(G_n(\pi)-G(\pi)\right)\xrightarrow{d}N(0,\Omega(\pi))$ with some positive definite covariance matrix $\Omega(\pi)$ for any $\pi\in\Pi$; (ii) for any $c\in\mathscr{C}_n$, $W_n(c)=W(c)+o_p(1)$ with $W(c)$ being positive definite; and (iii) $W(c_0)=\Omega_0^{-1}$ for Ω_0 being the asymptotic variance of $\sqrt{n}\left(G_{n,c_0}(\pi_0)-G_{c_0}(\pi_0)\right)$. Then $\limsup_{n\to\infty}\sup_{\xi\in\Xi_R}\Pr_\xi\left(T_n>\chi^2_{[l-l_\pi+l_R],1-\alpha}\right)=\alpha$.*

Proof of Lemma B.4: First, we show that $\limsup_{n\to\infty}\sup_{\xi\in\Xi_R}\Pr_\xi\left(c_0\in\mathscr{C}_n\right)\to1$. Second, we prove that $\limsup_{n\to\infty}\sup_{\xi\in\Xi_R}\Pr_\xi\left(T_n>\chi^2_{[l-l_\pi+l_R],1-\alpha}\right)=\alpha$ if $c_0\in\mathscr{C}_n$.

For $s=1,...,S$, let $sc_0^s\in\mathscr{SC}^s$ denote the sub-selection vector whose first $2l_s$ elements are the same as c_0. We first show that $\limsup_{n\to\infty}\sup_{\xi\in\Xi_R}\Pr_\xi\left(sc_0^1\in\mathscr{SC}^1_n\right)\to 1$. By Assumptions 2.1-2.6, Lemma 2.1 holds, and we have that $G_{c_0}(\pi_0)=\mathbf{0}$. Together with Assumption (i), we can obtain that for any $\xi\in\Xi_R$ and the parameter sequence $\{\xi_n\}\in\Xi_R(\xi)$, $\left\|G_{n,c_0}(\pi_0)\right\|^2=O_p\left(n^{-1}\right)$ and $\left\|G_{n,sc_0^1}(\pi_0)\right\|^2=O_p\left(n^{-1}\right)$. Therefore, it holds that $J_n\left(sc_0^1\right)\leq\left\|G_{n,sc_0^1}(\pi_0)\right\|^2=O_p\left(n^{-1}\right)$ for any $\xi\in\Xi_R$ and the parameter sequence $\{\xi_n\}\in\Xi_R(\xi)$. We can further obtain that for any $\lambda>-1$,

$$1\geq\limsup_{n\to\infty}\sup_{\xi\in\Xi_R}\Pr_\xi\left(sc_0^1\in\mathscr{SC}^1_n\right)\geq\limsup_{n\to\infty}\sup_{\xi\in\Xi_R}\Pr_\xi\left(J_n\left(sc_0^1\right)\leq n^\lambda\right)\to1,$$

where the second inequality holds for any $\alpha_1 > 0$, because $\mathscr{S}\mathscr{C}_n^1$ can contain sc^1's such that $J_n(sc) > n^\lambda$. Thus, we obtain that $\limsup_{n\to\infty} \sup_{\xi\in\Xi_R} \Pr_\xi \left(sc_0^1 \in \mathscr{S}\mathscr{C}_n^1\right) \to 1$.

The convergence $\limsup_{n\to\infty} \sup_{\xi\in\Xi_R} \Pr_\xi \left(sc_0^2 \in \mathscr{S}\mathscr{C}_n^2\right) \to 1$ holds if

$$\limsup_{n\to\infty} \sup_{\xi\in\Xi_R} \Pr_\xi \left(sc_0^2 \in \mathscr{S}\mathscr{C}_n^2 \mid sc_0^2 \in \mathscr{S}\mathscr{C}^2\right) \to 1$$

and $\limsup_{n\to\infty} \sup_{\xi\in\Xi_R} \Pr_\xi \left(sc_0^2 \in \mathscr{S}\mathscr{C}^2\right) \to 1$. The former holds by the same proof as above and the latter holds because $sc_0^2 \in \mathscr{S}\mathscr{C}^2$ occurs if and only if $sc_0^1 \in \mathscr{S}\mathscr{C}_n^1$ occurs. We can obtain that $\limsup_{n\to\infty} \sup_{\xi\in\Xi_R} \Pr_\xi \left(sc_0^S \in \mathscr{S}\mathscr{C}_n^S\right) \to 1$ by sequentially applying this argument. Because the event $sc_0^S \in \mathscr{S}\mathscr{C}_n^S$ occurs if and only if $c_0 \in \mathscr{C}_n$ occurs, it holds that $\limsup_{n\to\infty} \sup_{\xi\in\Xi_R} \Pr_\xi (c_0 \in \mathscr{C}_n) \to 1$.

In the next step, we show that $\limsup_{n\to\infty} \sup_{\xi\in\Xi_R} \Pr_\xi \left(T_n > \chi^2_{[l-l_\pi+l_R],1-\alpha}\right) = \alpha$ if the event $c_0 \in \mathscr{C}_n$ occurs,. We prove the result in two cases: $l_R < l_\pi$ and $l_R = l_\pi$. When $l_R < l_\pi$, the null space of R has dimension $l_\pi - l_R$. Let Ψ be a $l_\pi \times (l_\pi - l_R)$ matrix storing a basis of the null space. Then there exists $\pi_f \in R^{l_\pi - l_R}$ and μ such that any π satisfying H_0 can be written as $\pi = \Psi\pi_f + \mu$. By imposing H_0 on the sample moment functions, we obtain that

$$G_{n,c}(\pi) = G_{n,c}(\Psi\pi_f + \mu) = \overline{\pi}_{n,c} - \Gamma_{n,c}\mu - \Gamma_{n,c}\Psi\pi_f.$$

Define $\widehat{\pi}_f(c) \equiv \arg\min_{\pi_f} \|G_{n,c}(\Psi\pi_f + \mu)\|^2_{W_n(c)}$ for any $c \in \mathscr{C}_n$, where $W_n(c) = W(c) + o_p(1)$ by Assumption (ii). Define $\pi_f^*(c) \equiv \text{plim}_{n\to\infty}\widehat{\pi}_f(c)$. For each given $c \in \mathscr{C}_n$, we have the solution $\widehat{\pi}_f(c)$ to the minimum-distance problem with a corresponding "pseudo-true" value $\pi_f^*(c)$ defined as its probability limit. The "pseudo-true" value at c_0 delivers the true value as $\pi_0 = \Psi\pi_f^*(c_0) + \mu$ for any parameter sequence $\xi_n \in \Xi_R(\xi)$. Our test statistic T_n is equivalent to $\min_{c\in\mathscr{C}_n} n \|G_{n,c}(\Psi\widehat{\pi}_f(c) + \mu)\|^2_{W_n(c)}$. We aim to show that the test based upon the test statistic T_n and the critical value $\chi^2_{[l-l_\pi+l_R],1-\alpha}$ controls the asymptotic size when $c_0 \in \mathscr{C}_n$. For this purpose we derive its asymptotic distribution under drifting model parameter sequences. To simplify the discussion, we omit "under any $\xi \in \Xi_R$ and the parameter sequence $\{\xi_n\} \in \Xi_R(\xi)$" with the understanding that all the derivations are for any $\xi \in \Xi_R$ and the parameter sequence $\{\xi_n\} \in \Xi_R(\xi)$.

Apply mean value expansion of $G_{n,c_0}(\Psi\widehat{\pi}_f(c_0) + \mu)$ at $\pi_f^*(c_0)$. We can obtain that

$$G_{n,c_0}(\Psi\widehat{\pi}_f(c_0) + \mu) = G_{n,c_0}\left(\Psi\pi_f^*(c_0) + \mu\right) - \Gamma_{n,c_0}\Psi\left(\widehat{\pi}_f(c_0) - \pi_f^*(c_0)\right). \qquad \text{(B.3)}$$

Let $a_{n,c_0} \equiv \Gamma_{n,c_0}\Psi$. By construction, $\widehat{\pi}_f(c_0)$ satisfies the following first order condition:

$$a_{n,c_0}^\top W_n(c_0) G_{n,c_0}(\Psi\widehat{\pi}_f(c_0) + \mu) = 0.$$

Multiply both sides of (B.3) by $a_{n,c_0}^\top W_n(c_0)$. We obtain that

$$
\begin{aligned}
& a_{n,c_0}^\top W_n(c_0)\, G_{n,c_0}(\Psi\widehat{\pi}_f(c_0)+\mu) \\
= \;& a_{n,c_0}^\top W_n(c_0)\, G_{n,c_0}\left(\Psi\pi_f^*(c_0)+\mu\right) - a_{n,c_0}^\top W_n(c_0)\, a_{n,c_0}\left(\widehat{\pi}_f(c_0)-\pi_f^*(c_0)\right).
\end{aligned}
$$

By Assumption 2.6, Γ_{c_0} has full row rank. Together with Assumption (i), we obtain that Γ_{n,c_0} has full row rank with probability approaching one for any parameter sequence. Thus, $a_{n,c_0}^\top W_n(c_0)\, a_{n,c_0}$ is invertible with probability approaching one. After rearrangement, we have that

$$
\begin{aligned}
\left(\widehat{\pi}_f(c_0)-\pi_f^*(c_0)\right) &= \left(a_{n,c_0}^\top W_n(c_0)\, a_{n,c_0}\right)^{-1} a_{n,c_0}^\top W_n(c_0)\, G_{n,c_0}\left(\Psi\pi_f^*(c_0)+\mu\right) \\
&\quad - \left(a_{n,c_0}^\top W_n(c_0)\, a_{n,c_0}\right)^{-1} a_{n,c_0}^\top W_n(c_0)\, G_{n,c_0}(\Psi\widehat{\pi}_f(c_0)+\mu) \\
&= \left(a_{n,c_0}^\top W_n(c_0)\, a_{n,c_0}\right)^{-1} a_{n,c_0}^\top W_n(c_0)\, G_{n,c_0}(\pi_0), \qquad \text{(B.4)}
\end{aligned}
$$

where the last equality follows from the first order condition. Multiply both sides of (B.4) with a_{n,c_0}, we get that

$$
\begin{aligned}
& a_{n,c_0}\left(\widehat{\pi}_f(c_0)-\pi_f^*(c_0)\right) \\
= \;& a_{n,c_0}\left(a_{n,c_0}^\top W_n(c_0)\, a_{n,c_0}\right)^{-1} a_{n,c_0}^\top W_n(c_0)\, G_{n,c_0}(\pi_0). \qquad \text{(B.5)}
\end{aligned}
$$

Combining (B.3) with (B.5), we have that

$$
\begin{aligned}
& \sqrt{n}\left(G_{n,c_0}(\Psi\widehat{\pi}_f(c_0)+\mu) - G_{n,c_0}(\pi_0)\right) \\
= \;& -a_{n,c_0}\sqrt{n}\left(\widehat{\pi}_f(c_0)-\pi_f^*(c_0)\right) \\
= \;& -a_{n,c_0}\left(a_{n,c_0}^\top W_n(c_0)\, a_{n,c_0}\right)^{-1} a_{n,c_0}^\top W_n(c_0)\sqrt{n}G_{n,c_0}(\pi_0).
\end{aligned}
$$

Let I_l be the identify matrix of dimension l. It then holds that

$$
\begin{aligned}
& \sqrt{n}G_{n,c_0}(\Psi\widehat{\pi}_f(c_0)+\mu) \\
= \;& \left(I_l - a_{n,c_0}\Psi\left(a_{n,c_0}^\top W_n(c_0)\, a_{n,c_0}\right)^{-1} a_{n,c_0}^\top W_n(c_0)\right)\sqrt{n}G_{n,c_0}(\pi_0).
\end{aligned}
$$

Because $\Gamma_{n,c_0} = \Gamma_{c_0} + o_p(1)$ by Assumption (i), we have $a_{n,c_0} = \Gamma_{c_0}\Psi + o_p(1) \equiv a_{c_0} + o_p(1)$. Together with $W_n(c_0) = W(c_0) + o_p(1) = \Omega_0^{-1} + o_p(1)$ by Assumptions (i) and (iii), we have that

$$
\left(I_l - a_{n,c_0}\left(a_{n,c_0}^\top W_n(c_0)\, a_{n,c_0}\right)^{-1} a_{n,c_0}^\top W_n(c_0)\right) = \left(I_l - a_{c_0}\left(a_{c_0}^\top\Omega_0^{-1}a_{c_0}\right)^{-1} a_{c_0}^\top\Omega_0^{-1}\right) + o_p(1).
$$

Since $\pi_0 = \Psi\pi_f^*(c_0) + \mu$ under the null hypothesis and $\sqrt{n}G_{n,c_0}(\pi_0) = O_p(1)$ by Assumption (i), we obtain that

$$
\begin{aligned}
& \sqrt{n}G_{n,c_0}(\Psi\widehat{\pi}_f(c_0)+\mu) \\
= \;& \left(I_l - a_{c_0}\left(a_{c_0}^\top\Omega_0^{-1}a_{c_0}\right)^{-1} a_{c_0}^\top\Omega_0^{-1}\right)\sqrt{n}G_{n,c_0}(\pi_0) + o_p(1). \qquad \text{(B.6)}
\end{aligned}
$$

By Assumptions (ii) and (iii), Ω_0^{-1} is symmetric and positive definite. As a result it admits a Cholesky decomposition: $\Omega_0^{-1} = A^\top A$. Thus it holds that

$$\begin{aligned} & nG_{n,c_0}\left(\Psi\widehat{\pi}_f\left(c_0\right)+\mu\right)^\top W_n\left(c_0\right) G_{nc_0}\left(\Psi\widehat{\pi}_f\left(c_0\right)+\mu\right) \\ = \; & nG_{n,c_0}\left(\Psi\widehat{\pi}_f\left(c_0\right)+\mu\right)^\top \Omega_0^{-1} G_{n,c_0}\left(\Psi\widehat{\pi}_f\left(c_0\right)+\mu\right) + o_p\left(1\right) \\ = \; & \left(A\sqrt{n}G_{n,c_0}\left(\Psi\widehat{\pi}_f\left(c_0\right)+\mu\right)\right)^\top \left(A\sqrt{n}G_{n,c_0}\left(\Psi\widehat{\pi}_f\left(c_0\right)+\mu\right)\right) + o_p\left(1\right). \end{aligned}$$

Since $\Omega_0 = A^{-1}\left(A^\top\right)^{-1}$, Assumption (i) and Lemma 2.1 imply that

$$\sqrt{n}G_{n,c_0}\left(\pi_0\right) \xrightarrow{d} N\left(0,\Omega_0\right) \overset{d}{=} A^{-1}Z,$$

where $Z \sim N\left(0, I_l\right)$. Together with (B.6), we obtain that

$$\begin{aligned} A\sqrt{n}G_{n,c_0}\left(\Psi\widehat{\pi}_f\left(c_0\right)+\mu\right) & \xrightarrow{d} A\left(I_l - a_{c_0}\left(a_{c_0}^\top\Omega_0^{-1}a_{c_0}\right)^{-1}a_{c_0}^\top\Omega_0^{-1}\right)A^{-1}Z \\ & = \left(I_l - Aa_{c_0}\left(a_{c_0}^\top\Omega_0^{-1}a_{c_0}\right)^{-1}a_{c_0}^\top A^\top\right)Z. \end{aligned}$$

It is easy to verify that the matrix in front of Z in the above expression is symmetric and idempotent. It has the rank $l - l_\pi + l_R$ verified by the following derivation:

$$\begin{aligned} & \operatorname{rank}\left[\left(I_l - Aa_{c_0}\left(a_{c_0}^\top\Omega_0^{-1}a_{c_0}\right)^{-1}a_{c_0}^\top A^\top\right)\right] \\ = \; & \operatorname{trace}\left[\left(I_l - Aa_{c_0}\left(a_{c_0}^\top\Omega_0^{-1}a_{c_0}\right)^{-1}a_{c_0}^\top A^\top\right)\right] \\ = \; & \operatorname{trace}\left(I_l\right) - \operatorname{trace}\left(Aa_{c_0}\left(a_{c_0}^\top\Omega_0^{-1}a_{c_0}\right)^{-1}a_{c_0}^\top A^\top\right) \\ = \; & \operatorname{trace}\left(I_l\right) - \operatorname{trace}\left(a_{c_0}^\top A^\top A a_{c_0}\left(a_{c_0}^\top\Omega_0^{-1}a_{c_0}\right)^{-1}\right) \\ = \; & l - l_\pi + l_R, \end{aligned}$$

where the first equality follows from the property of idempotent matrix and third equality follows from the invariance property of trace under cyclic permutations. Thus, we obtain that for any $\xi \in \Xi_R$ and the parameter sequence $\{\xi_n\} \in \Xi_R\left(\xi\right)$,

$$\min_{R\pi=r}\left\|\sqrt{n}G_{n,c_0}\left(\pi\right)\right\|^2_{W_n(c_0)} = n\left\|G_{n,c_0}\left(\Psi\widehat{\pi}_f\left(c_0\right)+\mu\right)\right\|^2_{W_n(c_0)} \xrightarrow{d} \chi^2_{[l-l_\pi+l_R]}.$$

By the definition of test statistic in (4.2) in Fan et al. (2024), it holds that $T_n \leq \min_{R\pi=r}\left\|\sqrt{n}G_{n,c_0}\left(\pi\right)\right\|^2_{W_n(c_0)}$. Therefore, T_n is asymptotically stochastically dominated by $\chi^2_{[l-l_\pi+l_R]}$. Moreover, for some $\xi \in \Xi_R$ and the parameter sequence $\{\xi_n\} \in \Xi_R\left(\xi\right)$ such that each element in $\lim_{n\to\infty}\sqrt{n}G_{n,c}\left(\pi\right)$ is infinite for every $c \neq c_0$, $T_n \xrightarrow{d} \chi^2_{[l-l_\pi+l_R]}$. Thus, if $c_0 \in \mathscr{C}_n$, using the $(1-\alpha)$-th quantile of $\chi^2_{[l-(l_\pi-l_R)]}$ denoted as $\chi^2_{[l-(l_\pi-l_R)],1-\alpha}$, achieves asymptotic size control:

$$\limsup_{n\to\infty}\sup_{\xi\in\Xi_R}\Pr_\xi\left(T_n > \chi^2_{[l-l_\pi+l_R],1-\alpha} \mid c_0 \in \mathscr{C}_n\right) = \alpha.$$

The proof for the case where $l_R = l_\pi$ is the same with $\pi = R^{-1}r$. Therefore, combining with $\limsup_{n\to\infty}\sup_{\xi\in\Xi_R}\Pr_\xi\left(c_0\in\mathscr{C}_n\right)\to 1$, the lemma holds. $\square$

Lemma B.5. *Let Lemma 2.1 and the following assumptions hold for the Simple Game: (i) for any* $\pi\in\Pi$, $G_n(\pi)=G(\pi)+O_p\left(n^{-1/2}\right)$*; (ii) for any* $c\in\mathscr{C}_n$, $W_n(c)=W(c)+o_p(1)$ *with* $W(c)$ *being positive definite; and (iii) for any* $c\in\mathscr{C}_n$, $\min_{\pi\in\Pi}\|G_{n,c}(\pi)\|^2_{W_n(c)}\xrightarrow{p}\min_{\pi\in\Pi}\|G_c(\pi)\|^2_{W(c)}$. *Then it holds that for any* $\xi\notin\Xi_R$, $\lim_{n\to\infty}\Pr_\xi\left(T_n>\chi^2_{[l-l_\pi+l_R],1-\alpha}\right)=1$.

Proof of Lemma B.5: When $\xi\notin\Xi_R$, it is straightforward to see that for all $c\in\mathscr{C}$, $\min_{R\pi=r}\|\sqrt{n}G_{n,c}(\pi)\|^2_{W_n(c)}$ diverges to infinity under Lemma 2.1 and Assumption (i)-(iii). Since $\chi^2_{[l-l_\pi+l_R],1-\alpha}$ is finite, it holds that $\lim_{n\to\infty}\Pr_\xi\left(T_n>\chi^2_{[l-l_\pi+l_R],1-\alpha}\right)=1$. Thus, the lemma follows. $\square$

Proof of Theorem 4.1: We prove the theorem by verifying the conditions in Lemmas B.4 and B.5. Lemma 2.1 holds by Assumptions 2.1-2.6. It suffices to show that Assumptions (i)-(iii) in Lemma B.4 hold for the first part of the theorem and Assumptions (i)-(iii) in Lemma B.5 hold for the second part of the theorem.

Firstly we show that the moment functions for the Simple Game admits an asymptotic linear representation under drifting sequences. Without loss of generality, we prove it for player 1. The proof of asymptotic linear representation builds on two lemmas provided after this proof. We first focus on finding the asymptotic linear representation for the first observed state $\mathbf{z}=\mathbf{z}^1$, denoted as $G_{n\mathbf{z}^1}(\pi)$, and then stack the asymptotic linear representations of $G_{n\mathbf{z}}(\pi)$ for $\mathbf{z}=\mathbf{z}^1,\ldots,\mathbf{z}^l$ in the end to obtain the asymptotic linear representation for $G_n(\pi)$. Because we have fixed the player and the observed state variable, from now on we will suppress the subscript i, z_i, and $\mathbf{z}$ when there is no confusion. With slight abuse of notation, define the 6×1 vector of equilibrium CCPs and its estimator as:

$$\begin{aligned} p_{\mathbf{z}^1} &\equiv \left[p_1\left(\mathbf{z}^1,k\right),p_1\left(\mathbf{z}^1,k'\right),\ldots,p_3\left(\mathbf{z}^1,k\right),p_3\left(\mathbf{z}^1,k'\right)\right]^\top \equiv \left[p_1,\ldots,p_6\right]^\top \text{ and} \\ \widehat{p}_{\mathbf{z}^1} &\equiv \left[\widehat{p}_1,\ldots,\widehat{p}_6\right]^\top. \end{aligned} \tag{B.7}$$

Let $\overline{\pi}_{\mathbf{z}^1}\equiv\left[F^{-1}\left(p_1\left(\mathbf{z}^1,k\right)\right),F^{-1}\left(p_1\left(\mathbf{z}^1,k'\right)\right)\right]^\top$ be the 2×1 equilibrium expected payoff vector, and define $\overline{\pi}_{n\mathbf{z}^1}$ as the estimated expected payoff vector. The matrix storing the true joint probabilities of opponents' actions is $\Gamma_{\mathbf{z}^1}=\left[\mathbf{p}_{-1}\left(\mathbf{z}^1,k\right)^\top,\quad \mathbf{p}_{-1}\left(\mathbf{z}^1,k'\right)^\top\right]^\top$. The estimated matrix for opponents' actions is denoted as $\Gamma_{n\mathbf{z}^1}$. Then $G_{n\mathbf{z}^1}(\pi)=$

$\overline{\pi}_{n\mathbf{z}^1} - \Gamma_{n\mathbf{z}^1}\pi$. Let the 8×1 vector $a_{\mathbf{z}^1}$ denote the free joint probabilities in the contingency table (conditional on $\mathbf{z} = \mathbf{z}^1$), and let the 9×1 vector $q_{\mathbf{z}^1}$ denote the free unconditional joint probabilities that generate the contingency table on $\mathbf{z}^1$:[2]

$$a_{\mathbf{z}^1} \equiv [a_1, \ldots, a_8]^\top \text{ and } q_{\mathbf{z}^1} \equiv [q_1, \ldots, q_9]^\top \text{, where} \tag{B.8}$$

$$\begin{aligned}
&a_1 \equiv \Pr\left([d_1, d_2, d_3] = [1,1,1] \mid \mathbf{z}^1\right), \ \ a_2 \equiv \Pr\left([d_1, d_2, d_3] = [0,1,1] \mid \mathbf{z}^1\right), \\
&a_3 \equiv \Pr\left([d_1, d_2, d_3] = [1,0,1] \mid \mathbf{z}^1\right), \ \ a_4 \equiv \Pr\left([d_1, d_2] = [1,1] \mid \mathbf{z}^1\right), \\
&a_5 \equiv \Pr\left([d_1, d_2] = [0,1] \mid \mathbf{z}^1\right), \ \ a_6 \equiv \Pr\left([d_1, d_3] = [1,1] \mid \mathbf{z}^1\right), \\
&a_7 \equiv \Pr\left([d_1, d_3] = [0,1] \mid \mathbf{z}^1\right), \ \ a_8 \equiv \Pr\left(d_1 = 1 \mid \mathbf{z}^1\right) \text{ and}
\end{aligned} \tag{B.9}$$

$$\begin{aligned}
&q_1 \equiv \Pr\left([d_1, d_2, d_3] = [1,1,1], \mathbf{z} = \mathbf{z}^1\right), \ldots, \\
&q_8 \equiv \Pr\left(d_1 = 1, \mathbf{z} = \mathbf{z}^1\right), \ \ q_9 \equiv \Pr\left(\mathbf{z} = \mathbf{z}^1\right).
\end{aligned} \tag{B.10}$$

For $i = 1, ..., 8$, it holds that $a_i = q_i / q_9$. Note that $q_{\mathbf{z}^1}$ includes three probabilities for the joint actions of three players on $\mathbf{z}^1$, two probabilities for the joint actions of player 1 and player 2 on $\mathbf{z}^1$, and two probabilities for the joint actions of player 1 and player 3 on $\mathbf{z}^1$, one probability for the action of player 1 on $\mathbf{z}^1$ and the probability of $\mathbf{z} = \mathbf{z}^1$. The estimators $\widehat{q}_{\mathbf{z}^1}$ and $\widehat{a}_{\mathbf{z}^1}$ are calculated as

$$\widehat{q}_{\mathbf{z}^1} = \frac{1}{n} \sum_{m=1}^{n} \eta_{m\mathbf{z}^1} \equiv [\widehat{q}_1, \ldots, \widehat{q}_9]^\top \text{ and } \widehat{a}_{\mathbf{z}^1} = \left[\frac{\widehat{q}_1}{\widehat{q}_9}, \ldots, \frac{\widehat{q}_8}{\widehat{q}_9}\right]^\top \equiv [\widehat{a}_1, \ldots, \widehat{a}_8]^\top, \tag{B.11}$$

where $\eta_{m\mathbf{z}^1}$ is defined as

$$\eta_{m\mathbf{z}^1} \equiv \left[\mathbb{1}\left([d_{1m}, d_{2m}, d_{3m}] = [1,1,1], \mathbf{z}_m = \mathbf{z}^1\right), \ldots, \mathbb{1}\left(d_{1m} = 1, \mathbf{z}_m = \mathbf{z}^1\right), \mathbb{1}\left(\mathbf{z}_m = \mathbf{z}^1\right)\right]^\top. \tag{B.12}$$

By Lemmas B.6 and B.7 below, the second order Taylor expansion of $G_{n\mathbf{z}^1}(\pi)$ at $p_{\mathbf{z}^1}$ gives:

$$\sqrt{n}\left(G_{n\mathbf{z}^1}(\pi) - G_{\mathbf{z}^1}(\pi)\right) = \sqrt{n} D_{p\mathbf{z}^1}(\pi)\left(\widehat{p}_{\mathbf{z}^1} - p_{\mathbf{z}^1}\right) + \sum_j \sum_k \frac{H_{p^*-j,k}(\pi)}{2} \sqrt{n}\left(\widehat{p}_j - p_j\right)\left(\widehat{p}_k - p_k\right),$$

[2]Note that in Appendix C.1 although there are in total 14 probabilities in the contingency tables that generate CCPs via eigendecomposition, only 8 of them are free (not linear combinations of other probabilities): $a_1, \ldots, a_8$.

where $D_{p\mathbf{z}^1}(\pi) \equiv \begin{bmatrix} dp_{11}, \ldots, dp_{16} \\ dp_{21}, \ldots, dp_{26} \end{bmatrix}_{2\times 6}$, in which

$$\begin{aligned}
dp_{11} &= \frac{1}{F'\left(F^{-1}\left(p_1\left(\mathbf{z}^1, k\right)\right)\right)}, \; dp_{22} = \frac{1}{F'\left(F^{-1}\left(p_1\left(\mathbf{z}^1, k'\right)\right)\right)}, \\
dp_{12} &= dp_{14} = dp_{16} = dp_{21} = dp_{23} = dp_{25} = 0, \\
dp_{13} &= \left(p_3\left(\mathbf{z}^1, k\right), \left(1 - p_3\left(\mathbf{z}^1, k\right)\right), -p_3\left(\mathbf{z}^1, k\right), -\left(1 - p_3\left(\mathbf{z}^1, k\right)\right)\right)\pi, \\
dp_{15} &= \left(p_2\left(\mathbf{z}^1, k\right), -p_2\left(\mathbf{z}^1, k\right), 1 - p_2\left(\mathbf{z}^1, k\right), -\left(1 - p_2\left(\mathbf{z}^1, k\right)\right)\right)\pi, \\
dp_{24} &= \left(p_3\left(\mathbf{z}^1, k'\right), \left(1 - p_3\left(\mathbf{z}^1, k'\right)\right), -p_3\left(\mathbf{z}^1, k'\right), -\left(1 - p_3\left(\mathbf{z}^1, k'\right)\right)\right)\pi, \\
dp_{26} &= \left(p_2\left(\mathbf{z}^1, k'\right), -p_2\left(\mathbf{z}^1, k'\right), 1 - p_2\left(\mathbf{z}^1, k'\right), -\left(1 - p_2\left(\mathbf{z}^1, k'\right)\right)\right)\pi.
\end{aligned}$$

First, $D_{p\mathbf{z}^1}(\pi)$ is of full row rank. Second, by Assumption 4.1 (i), the denominators in dp_{11} and dp_{22} are bounded away from 0, and all other elements in $D_{p\mathbf{z}^1}(\pi)$ are bounded for any $\pi \in \Pi$. By Assumption 3.1, $D_{p\mathbf{z}^1}(\pi)$ is bounded uniformly over Π. There are only two types of non-zero elements in all Hessian vectors: payoff parameters and $-\frac{F''\left(F^{-1}\left(p_1\left(\mathbf{z}^1,k\right)\right)\right)}{\left[F'\left(F^{-1}\left(p_1\left(\mathbf{z}^1,k\right)\right)\right)\right]^3}$ or $-\frac{F''\left(F^{-1}\left(p_1\left(\mathbf{z}^1,k'\right)\right)\right)}{\left[F'\left(F^{-1}\left(p_1\left(\mathbf{z}^1,k'\right)\right)\right)\right]^3}$. Thus by Assumption 3.1 and 4.1 (i), all elements in the Hessian vectors are bounded, i.e., there exists a absolute constant M^* such that for any j and k, it holds that $\sum_j \sum_k \left\| \frac{H_{p^*-j,k}(\pi)}{2} \right\| < M^*$. For large enough n, the following inequality holds:

$$\begin{aligned}
& \left\| \sum_j \sum_k \frac{H_{p^*-j,k}(\pi)}{2} \sqrt{n}\left(\widehat{p}_j - p_j\right)\left(\widehat{p}_k - p_k\right) \right\| \\
\leq \; & \sqrt{n} \sum_j \sum_k \left\| \frac{H_{p^*-j,k}(\pi)}{2} \right\| \left\|\widehat{p}_{\mathbf{z}^1} - p_{\mathbf{z}^1}\right\|^2 \leq M^* \sqrt{n} \left\|\widehat{p}_{\mathbf{z}^1} - p_{\mathbf{z}^1}\right\|^2 = o_p(1).
\end{aligned}$$

This implies that

$$\begin{aligned}
& \sqrt{n}\left(G_{n\mathbf{z}^1}(\pi) - G_{\mathbf{z}^1}(\pi)\right) \\
= \; & D_{p\mathbf{z}^1}(\pi) \sqrt{n}\left(\widehat{p}_{q\mathbf{z}^1} - p_{q\mathbf{z}^1}\right) + o_p(1) \\
= \; & D_{p\mathbf{z}^1}(\pi) D_{a\mathbf{z}^1}\left(\sqrt{n}\left(\widehat{a}_{q\mathbf{z}^1} - a_{q\mathbf{z}^1}\right) + o_p(1)\right) \\
= \; & D_{p\mathbf{z}^1}(\pi) D_{a\mathbf{z}^1}\left(D_{q\mathbf{z}^1}\sqrt{n}\left(\widehat{q}_{q\mathbf{z}^1} - q_{q\mathbf{z}^1}\right) + o_p(1)\right) \\
= \; & \frac{1}{\sqrt{n}} \sum_{m=1}^{n} D_{p\mathbf{z}^1}(\pi) D_{a\mathbf{z}^1} D_{q\mathbf{z}^1}\left(\eta_{m\mathbf{z}^1} - q_{\mathbf{z}^1}\right) + o_p(1) \text{ and}
\end{aligned}$$

$$G_{n\mathbf{z}^1}(\pi) = \frac{1}{n} \sum_{m=1}^{n} \left[D_{p\mathbf{z}^1}(\pi) D_{a\mathbf{z}^1} D_{q\mathbf{z}^1}\left(\eta_{m\mathbf{z}^1} - q_{\mathbf{z}^1}\right) + G_{\mathbf{z}^1}(\pi)\right] + o_p\left(n^{-1/2}\right),$$

where definitions of $D_{a\mathbf{z}^1}$ and $D_{q\mathbf{z}^1}$ can be found in Lemma B.6 and B.7.

Without loss of generality, Lemma B.6 and B.7 prove the result for $\mathbf{z} = \mathbf{z}^1$. The same result holds for $\mathbf{z} = \mathbf{z}^2, \ldots, \mathbf{z}^l$, and such asymptotic linear representation of $G_{n\mathbf{z}}(\pi)$ is available for every observed state $\mathbf{z}$. Stacking the asymptotic linear representation for $\mathbf{z} = \mathbf{z}^1, \ldots, \mathbf{z}^l$, we obtain the asymptotic linear representation for player 1's moment function when holding his exclusive observed state at $z_1 = z_1^1$. Note that on $\mathbf{z}^l$, $\mathbb{1}\left(\mathbf{z} = \mathbf{z}^l\right) = 1 - \left(\sum_{s=1}^{l-1} \mathbb{1}\left(\mathbf{z} = \mathbf{z}^s\right)\right)$ and $\Pr\left(\mathbf{z} = \mathbf{z}^l\right) = 1 - \sum_{s=1}^{l-1} \Pr\left(\mathbf{z} = \mathbf{z}^s\right)$. We drop these two elements when defining $\eta_{m\mathbf{z}^l}$ and $q_{\mathbf{z}^l}$ and delete the corresponding last column of $D_{q\mathbf{z}^l}$, so that elements in vector $\eta_m - q$ are free of each other, where η_m and q are obtained by stacking $\eta_{m\mathbf{z}^s}$ and $q_{\mathbf{z}^s}$ together for $s = 1, \ldots, l$. For $m = 1, \ldots, n$, denote $Q_m \equiv (d_{1m}, d_{2m}, d_{3m}, z_{1m}, z_{2m}, z_{3m})$. The asymptotic linear representation is expressed as

$$G_n(\pi) \;\; = \;\; \frac{1}{n} \sum_{m=1}^{n} \phi\left(Q_m, \theta, \pi\right) + o_p\left(n^{-1/2}\right), \qquad \text{(B.13)}$$

where $\phi\left(Q_m, \theta, \pi\right) = D_p(\pi) D_a D_q \left(\eta_m - q\right) + G(\pi)$. The θ in $\phi\left(Q_m, \theta, \pi\right)$ includes elements in $D_p(\pi)$, D_a, D_q, Γ and q, which are obtained by stacking $D_{p\mathbf{z}}(\pi)$, $D_{a\mathbf{z}}$, $D_{q\mathbf{z}}$, $\Gamma_{\mathbf{z}}$ and $q_{\mathbf{z}}$ for $\mathbf{z} = \mathbf{z}^1, \ldots, \mathbf{z}^l$.

The CCPs, contingency tables, and indicator functions are different across $\mathbf{z}$. This implies that the coefficient matrix in front of $\left[\left[\eta_{m\mathbf{z}^1} - q_{\mathbf{z}^1}\right]^\top, \ldots, \left[\eta_{m\mathbf{z}^l} - q_{\mathbf{z}^l}\right]^\top\right]^\top$ is block diagonal. Moreover, since $D_{p\mathbf{z}^s}(\pi) D_{a\mathbf{z}^s} D_{q\mathbf{z}^s}$ is of rank 2 for $s = 1, \ldots, l$ and any $\pi \in \Pi$, each block has full row rank. Therefore, we obtain that $D_p(\pi) D_a D_q$ has full row rank and $\mathbb{E}\left(\phi\left(Q_m, \theta, \pi\right) \phi^\top\left(Q_m, \theta, \pi\right)\right)$ is nonsingular. Thus, by the Lindeberg Central Limit Theorem, the scaled and demeaned moment function converges in distribution to a normal distribution with a positive definite variance covariance matrix for any parameter sequence $\xi_n \in \Xi_R(\xi)$ with any $\xi \in \Xi_R$. Assumption (i) in Lemma B.4 is verified. This result also implies that Ω_0 in Assumption 4.1 (iv) is nonsingular. Under Assumption 4.1 (iii) and (iv), Assumptions (ii) and (iii) in Lemma B.4 are satisfied. By Lemma B.4, the first part of the theorem follows.

Assumptions (i)-(iii) in Lemma B.5 are the same as Assumptions (i)-(iii) in Lemma B.1. The proof of Theorem 3.1 shows that Assumptions (i)-(iii) in Lemma B.1 are implied by Assumptions 2.1-2.6, 3.1, and 3.2. Thus, the second part of the theorem holds. □

Lemma B.6. *Under Assumptions 2.3 and 4.1, it holds that for any $\xi \in \Xi_R$ and the* *parameter sequence* $\{\xi_n\} \in \Xi_R(\xi)$,

$$\sqrt{n}\left(\widehat{q}_{\mathbf{z}^1} - q_{\mathbf{z}^1}\right) = O_p(1) \text{ and} \tag{B.14}$$

$$\sqrt{n}\left(\widehat{a}_{\mathbf{z}^1} - a_{\mathbf{z}^1}\right) = \sqrt{n} D_{q\mathbf{z}^1}\left(\widehat{q}_{\mathbf{z}^1} - q_{\mathbf{z}^1}\right) + o_p(1), \text{ where} \tag{B.15}$$

$$D_{q\mathbf{z}^1} \equiv \begin{bmatrix} \frac{1}{q_9} & 0 & \dots & 0 & -\frac{q_1}{q_9^2} \\ 0 & \frac{1}{q_9} & \dots & 0 & -\frac{q_2}{q_9^2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{1}{q_9} & -\frac{q_8}{q_9^2} \end{bmatrix}_{8\times 9}$$

in which $\widehat{q}_{\mathbf{z}^1}$, $q_{\mathbf{z}^1}$, $\widehat{a}_{\mathbf{z}^1}$, *and* $a_{\mathbf{z}^1}$ *are defined in (B.8)-(B.11).*

Proof of Lemma B.6: Since each element in $\eta_{m\mathbf{z}^1}$ (defined in (B.12)) is less than or equal to 1, the Lindeberg Condition is satisfied. Under Assumption 2.3, by Lindeberg Central Limit Theorem, (B.14) holds. To obtain (B.15), apply the Taylor Expansion of $\sqrt{n}\left(\widehat{a}_{\mathbf{z}^1} - a_{\mathbf{z}^1}\right)$ around q:

$$\sqrt{n}\left(\widehat{a}_{\mathbf{z}^1} - a_{\mathbf{z}^1}\right) = \sqrt{n} D_{q\mathbf{z}^1}\left(\widehat{q}_{\mathbf{z}^1} - q_{\mathbf{z}^1}\right) + \sqrt{n}\sum_j \sum_k \frac{H_{q^*-j,k}}{2}\left(\widehat{q}_j - q_j\right)\left(\widehat{q}_k - q_k\right),$$

where $D_{q\mathbf{z}^1}$ is defined in the lemma, and $H_{q^*-j,k}$ is the Hessian vector that stores the second order derivatives with respect to q_j and q_k evaluated at $q^*_{\mathbf{z}^1}$, which is a point lying between $q_{\mathbf{z}^1}$ and $\widehat{q}_{\mathbf{z}^1}$. Each element in $H_{q^*-j,k}$ is either $-\frac{1}{q_9^{*2}}$, $\frac{2q_j^*}{q_9^{*3}}$, or 0. Note that

$$\left\|\sqrt{n}\sum_j \sum_k \frac{H_{q^*-j,k}}{2}\left(\widehat{q}_j - q_j\right)\left(\widehat{q}_k - q_k\right)\right\| \leq \sqrt{n}\sum_j \sum_k \left\|\frac{H_{q^*-j,k}}{2}\right\| \left\|\widehat{q}_{\mathbf{z}^1} - q_{\mathbf{z}^1}\right\|^2.$$

For any $\epsilon > 0$, there exists $N_{\epsilon,q} > 0$ such that $\Pr\left(\left\|\widehat{q}_{\mathbf{z}^1} - q_{\mathbf{z}^1}\right\| < \frac{\delta_1}{2}\right) \geq 1 - \epsilon$ for $n \geq N_{\epsilon,q}$ by triangular array Weak Law of Large Numbers (WLLN); and since $q_9 = \Pr\left(\mathbf{z} = \mathbf{z}^1\right) \geq \delta_1 > 0$, we obtain that $\Pr\left(q_9^* \geq \frac{\delta_1}{2}\right) \geq 1 - \epsilon$ for $n \geq N_{\epsilon,q}$. For $q_9^* \geq \frac{\delta_1}{2}$, there exists M_1 such that $\sum_j \sum_k \left\|\frac{H_{q^*-j,k}}{2}\right\| \leq M_1 < \infty$. Therefore, by (B.14) and triangular array WLLN, we have

$$\sqrt{n}\sum_j \sum_k \left\|\frac{H_{q^*-j,k}}{2}\right\| \left\|\widehat{q}_{\mathbf{z}^1} - q_{\mathbf{z}^1}\right\|^2 \leq M_1 \sqrt{n} \left\|\widehat{q}_{\mathbf{z}^1} - q_{\mathbf{z}^1}\right\|^2$$

holds with probability at least $1 - \epsilon$ for n sufficiently large, where $\epsilon > 0$ is arbitrary. Because $\left\|\widehat{q}_{\mathbf{z}^1} - q_{\mathbf{z}^1}\right\| = O_p\left(n^{-1/2}\right)$, (B.15) holds. □

Lemma B.7. *Under Assumption 2.1-2.5 and 4.1, it holds that*

$$\sqrt{n}\left(\widehat{p}_{\mathbf{z}^1}-p_{\mathbf{z}^1}\right)=\sqrt{n}D_{a\mathbf{z}^1}\left(\widehat{a}_{\mathbf{z}^1}-a_{\mathbf{z}^1}\right)+o_p\left(1\right) \tag{B.16}$$

for any $\xi\in\Xi_R$ and the parameter sequence $\{\xi_n\}\in\Xi_R\left(\xi\right)$, where $D_{a\mathbf{z}^1}\equiv\frac{\partial p_{\mathbf{z}^1}}{\partial a_{\mathbf{z}^1}^{\top}}$, and $\widehat{p}_{\mathbf{z}^1}$ and $p_{\mathbf{z}^1}$ are defined in equation (B.7).

Proof of Lemma B.7: Second order Taylor expansion of $\widehat{p}_{\mathbf{z}^1}$ at a provides that

$$\sqrt{n}\left(\widehat{p}_{\mathbf{z}^1}-p_{\mathbf{z}^1}\right)=\sqrt{n}D_{a\mathbf{z}^1}\left(\widehat{a}_{\mathbf{z}^1}-a_{\mathbf{z}^1}\right)+\sqrt{n}\sum_j\sum_k\frac{H_{a^*-j,k}}{2}\left(\widehat{a}_j-a_j\right)\left(\widehat{a}_k-a_k\right),$$

where $H_{a^*-j,k}$ is the Hessian vector that stores second order derivatives with respect to a_j and a_k evaluated at $a^*_{\mathbf{z}^1}$, which is a point between $a_{\mathbf{z}^1}$ and $\widehat{a}_{\mathbf{z}^1}$. Under Assumption 2.5, the eigenvalues in the eigendecomposition are simple, and consequently there exists a neighborhood around the true value of a such that in this neighborhood the eigenvector function of $A^{12}_{1\mathbf{z}^1}\left(A^{12}_{\mathbf{z}^1}\right)^{-1}$ (a matrix whose elements are continuous functions of a) is analytic. CCPs for player 1 are delivered by the eigenvector function, and CCPs for other players are continuously differentiable transformation of player 1's CCPs. Because $\widehat{a}_{\mathbf{z}^1}\xrightarrow{p}a_{\mathbf{z}^1}$, for large enough n, $H^*_{a-j,k}$ is bounded with probability close to 1. By a similar reasoning as in the previous lemma, the claim in the lemma holds. □

Proof of Proposition 4.1: It suffices to show that $W_n^b\left(c\right)\xrightarrow{p}W^b\left(c\right)$, where $W^b\left(c\right)$ is positive definite for any $c\in\mathscr{C}_n$ and $W^b\left(c_0\right)=\Omega_0^{-1}$. Under Assumption 4.1, Γ_{c_0} is of full column rank and $\text{rank}\left(\Gamma_{c_0}\Psi\right)=l_\pi-l_R$. For n sufficiently large, we have $\text{rank}\left(\Gamma_{n,c_0}\Psi\right)=l_\pi-l_R$ and $\arg\min_{\pi_f}\left\|G_{n,c_0}\left(\Psi\pi_f+\mu\right)\right\|^2$ is unique. Therefore, $W_n^b\left(c_0\right)=\left(\Sigma_n^b\left(c_0,\widehat{\pi}_f\left(c_0\right)\right)\right)^{-1}$ for large enough n.

Latent states are matched across bootstrap draws with probability approaching one based on $p_i\left(\mathbf{z},k\right)$ and $p_i\left(\mathbf{z},k'\right)$. Under Assumptions 2.1-2.6, 3.1, and 4.1 (i) and (ii) the moment functions for the Simple Game admits an asymptotic linear representation by the proof of Theorem 4.1. For the vector of functions ϕ defined in (B.13), let ϕ_{c_0} denotes its elements selected by c_0. Thus it holds that

$$G^{(b)}_{n,c_0}\left(\Psi\widehat{\pi}_f\left(c_0\right)+\mu\right)=\frac{1}{n}\sum_{m=1}^n\phi_{c_0}\left(Q_m^{(b)},\theta,\Psi\widehat{\pi}_f\left(c_0\right)+\mu\right)+o_p\left(1\right).$$

Since $\widehat{\pi}_f\left(c_0\right)\xrightarrow{p}\pi^*_f\left(c_0\right)$ and $\pi_0=\Psi\pi^*_f\left(c_0\right)+\mu$ under the null hypothesis, where $\pi^*_f\left(c_0\right)=\arg\min_{\pi\in\Pi}\left\|G_{c_0}\left(\Psi\pi_f+\mu\right)\right\|^2$, we have $\Psi\widehat{\pi}_f\left(c_0\right)+\mu\xrightarrow{p}\pi_0$. Under Assumptions Assumptions 2.1-2.6, 3.1, and 4.1 (i) and (ii), it holds that

$$\sqrt{n}\left(G_{n,c_0}\left(\Psi\pi_f^*(c_0)+\mu\right)-G_{c_0}\left(\Psi\pi_f^*(c_0)+\mu\right)\right)\xrightarrow{d}N(0,\Omega_0)$$

and

$$\begin{aligned}
&\sqrt{n}\left(G_{n,c_0}^{(b)}\left(\Psi\widehat{\pi}_f(c_0)+\mu\right)-G_{n,c_0}\left(\Psi\widehat{\pi}_f(c_0)+\mu\right)\right)\\
=\ &\sqrt{n}\left(\frac{1}{n}\sum_{m=1}^{n}\phi_{c_0}\left(Q_m^{(b)},\theta,\Psi\widehat{\pi}_f(c_0)+\mu\right)-\frac{1}{n}\sum_{m=1}^{n}\phi_{c_0}\left(Q_m,\theta,\Psi\widehat{\pi}_f(c_0)+\mu\right)\right)+o_p^*(1)\\
=\ &\sqrt{n}\left(\frac{1}{n}\sum_{m=1}^{n}\phi_{c_0}\left(Q_m^{(b)},\theta,\pi_0\right)-\frac{1}{n}\sum_{m=1}^{n}\phi_{c_0}\left(Q_m,\theta,\pi_0\right)\right)+o_p^*(1)\xrightarrow{d^*}N(0,\Omega_0),
\end{aligned}$$

where the definitions of $o_p^*(1)$ and $\xrightarrow{d^*}$ can be found in chapter 10 of Hansen (2021). The first equality holds because conditional on data, the difference between $G_{n,c_0}^{(b)}\left(\Psi\widehat{\pi}_f(c_0)+\mu\right)$ and its linear representation is $o_p(1)$, and the difference between $G_{n,c_0}\left(\Psi\widehat{\pi}_f(c_0)+\mu\right)$ and its linear representation is $o_p(1)$. Conditional on data, convergence in distribution holds by Theorem 10.8 in Hansen (2021) as $\|\phi_{c_0}(Q_l,\theta,\pi_0)\|^2$ is uniformly square integrable by the proof of Theorem 4.1. By Theorem 10.13 in Hansen (2021), $\Sigma_n^b(c_0,\widehat{\pi}_f(c_0))=\Omega_0+o_p(1)$ holds if $y_n^{(b)}$ is uniformly square integrable, where

$$y_n^{(b)}\equiv\sqrt{n}\left(\frac{1}{n}\sum_{m=1}^{n}\phi_{c_0}\left(Q_m^{(b)},\theta,\pi_0\right)-\frac{1}{n}\sum_{m=1}^{n}\phi_{c_0}\left(Q_m,\theta,\pi_0\right)\right).$$

Because the uniform square integrability of a vector is implied by each element of the vector being uniformly square integrable, let $y_n^{(b)}(i)$ be its i-th element. It holds that $\mathbb{E}^*\left|y_n^{(b)}(i)\right|^4=\frac{\widehat{\mu}_4(i)-3\widehat{\sigma}^4(i)}{n}+3\widehat{\sigma}^4(i)$, where

$$\begin{aligned}
\widehat{\mu}_4(i)\ &=\ \frac{1}{n}\sum_{m=1}^{n}\left|\phi_{c_0}^{(i)}\left(Q_m^{(b)},\theta,\pi_0\right)-\frac{1}{n}\sum_{m=1}^{n}\phi_{c_0}^{(i)}\left(Q_m,\theta,\pi_0\right)\right|^4\ \text{ and}\\
\widehat{\sigma}^2(i)\ &=\ \frac{1}{n}\sum_{m=1}^{n}\left|\phi_{c_0}^{(i)}\left(Q_m^{(b)},\theta,\pi_0\right)-\frac{1}{n}\sum_{m=1}^{n}\phi_{c_0}^{(i)}\left(Q_m,\theta,\pi_0\right)\right|^2,
\end{aligned}$$

with $\phi_{c_0}^{(i)}\left(Q_m^{(b)},\theta,\pi_0\right)$ and $\phi_{c_0}^{(i)}(Q_m,\theta,\pi_0)$ being the i-th element in $\phi_{c_0}\left(Q_m^{(b)},\theta,\pi_0\right)$ and $\phi_{c_0}(Q_m,\theta,\pi_0)$ respectively. Uniform square integrability of $\left|\phi_{c_0}^{(i)}(Q_m,\theta,\pi_0)\right|^2$ implies that $\frac{\widehat{\mu}_4(i)}{n}=o_p(1)$ and $\widehat{\sigma}^2(i)=O_p(1)$, which imply that $\mathbb{E}^*\left|y_n^{(b)}(i)\right|^4=O_p(1)$. Therefore $y_n^{(b)}$ is uniformly square integrable and $\Sigma_n^b(c_0,\widehat{\pi}_f(c_0))=\Omega_0+o_p(1)$ holds. Ω_0 is positive definite, because the asymptotic variance matrix of

$$\sqrt{n}\left(G_n\left(\Psi\pi_f^*(c_0)+\mu\right)-G\left(\Psi\pi_f^*(c_0)+\mu\right)\right)$$

is positive definite and Ω_0 is its submatrix with the corresponding rows and columns selected by c_0.

It remains to prove that for $c \neq c_0$, $W_n^b(c)$ converges in probability to some positive definite matrix. If $\text{rank}(\Gamma_{n,c}\Psi) < l_\pi - l_R$, then $W_n^b(c) = W_P$, which is positive definite. If $\text{rank}(\Gamma_c\Psi) = l_\pi - l_R$, then $\text{rank}(\Gamma_{n,c}\Psi) = l_\pi - l_R$ and $\arg\min_{\pi_f} \|G_{n,c}(\Psi\pi_f + \mu)\|^2$ is unique for sufficiently large n. By a similar argument as the one for $W_n^b(c_0)$, it holds that $W_n^b(c) = \left(\Sigma_n^b(c, \widehat{\pi}_f(c))\right)^{-1} = \Sigma^{-1}(c) + o_p(1)$, where $\Sigma(c)$ is the asymptotic variance of $\sqrt{n}\left(G_{n,c}\left(\Psi\pi_f^*(c) + \mu\right) - G_c\left(\Psi\pi_f^*(c) + \mu\right)\right)$ with $\pi_f^*(c) = \text{plim}\widehat{\pi}_f(c)$. $\Sigma(c)$ is positive definite, because the asymptotic variance matrix of $\sqrt{n}\left(G_n\left(\Psi\pi_f^*(c) + \mu\right) - G\left(\Psi\pi_f^*(c) + \mu\right)\right)$ is positive definite and $\Sigma(c)$ is its submatrix with the corresponding rows and columns selected by c. The proposition holds. □

B.4 Proofs of Results in Section 5

Proof of Lemma 5.1: To prove (i), note that under Assumptions 5.1-5.2, Step-1 identification is achieved, $G(\pi)$ and $\mathscr{C}^h$ are constructed for $h \in \{1, \ldots, |\Omega_{\mathbf{z}^1}|\}$. In each $\mathscr{C}^h$, as $\text{rank}([\overline{\pi}_c, \Gamma_c]) > \text{rank}(\Gamma_c)$ for any c that selects different latent states Assumption 5.3 (ii), $\mathscr{C}\mathscr{I}^h$ only contains select vectors that select the same latent state. Furthermore, c_0^h is defined as the vector in $\mathscr{C}\mathscr{I}^h$ that selects the most moments, which happens only when all equilibria corresponding to a latent state is selected. Thus c_0^h is unique. $\Gamma_{c_0^h}$ has full column rank implies that the system selected by c_0^h uniquely determines π_0^h. To prove (ii), it remains to identify $|\mathscr{K}|$ and the corresponding $|\mathscr{K}|$ distinct payoff vectors. This could be done by a pairwise comparison of the payoff vectors. Specifically, if the payoff vectors $\left\|\pi_0^{h_1} - \pi_0^{h_2}\right\| = 0$ for $h_1 \neq h_2$, then they correspond to the same latent state (or the same group). The total number of groups gives $|\mathscr{K}|$; the distinct payoff vectors are the $|\mathscr{K}|$ true payoff vectors denoted by $t\pi_0^k$ for $k = 1, \ldots, |\mathscr{K}|$. □

B.5 Proofs of Results in Appendix A

Lemma B.8. *Let the result of Lemma 5.1 and the following assumptions hold for the General Game. (i) For any* $\pi \in \Pi$, $G_n(\pi) = G(\pi) + O_p\left(n^{-1/2}\right)$. *(ii) For any* $c \in \mathscr{C}_n$, $W_n(c) = W(c) + o_p(1)$ *with* $W(c)$ *being positive definite. (iii) For any* $c \in \mathscr{C}_n$, $\min_{\pi\in\Pi} \|G_{n,c}(\pi)\|^2_{W_n(c)} \xrightarrow{p} \min_{\pi\in\Pi} \|G_c(\pi)\|^2_{W(c)}$. *(iv)* $\rho_1(\cdot) > 0$ *is a known strictly increasing function and* $\kappa_{1,n} \to \infty$ *with* $\kappa_{1,n} = o(n)$. *Then it holds that* *for* $h = 1, ..., |\Omega_{\mathbf{z}^1}|$, $\widehat{c}^h = c_0^h$ *with probability approaching one and* $\widehat{\pi}^h \xrightarrow{p} \pi_0^h$ *for any* $l_1 \in \{l_\pi, l_\pi + 1, .., l\}$, $\alpha_1 \in (0, 1]$, $\lambda \in (-1, 0)$, *and* $\Delta \in \{1, .., l - l_1\}$.

Proof of Lemma B.8: We prove the result for $(\widehat{c}^1, \widehat{\pi}^1)$. The superscript 1 in $\mathscr{C}^1$, $\mathscr{C}\mathscr{I}^1$, (c_0^1, π_0^1), $\mathscr{C}_n^1$, and $(\widehat{c}^1, \widehat{\pi}^1)$ are omitted. Other notations that appeared in the proof are the same as the ones used in Section 5.2.1 and Appendix A.2. For $s = 1, ..., S$, let $sc_0^s \in \mathscr{S}\mathscr{C}^s$ denote the sub-selection vector whose first $J \sum_{t=1}^{l_s} |\Omega_{\mathbf{z}^t}|$ elements are the same as those of c_0. First, we show that $\Pr(sc_0^s \in \mathscr{S}\mathscr{C}_n^s) \to 1$ for $s = 1, ..., S$. This implies that $\Pr(c_0 \in \mathscr{C}_n) \to 1$, because $c_0 \in \mathscr{C}_n$ occurs if and only if $sc_0^S \in \mathscr{S}\mathscr{C}_n^S$ occurs.

Lemma 5.1 implies that $G_{c_0}(\pi_0) = \mathbf{0}$. Combined with Assumption (i) in the lemma, we obtain that $\|G_{n,c_0}(\pi_0)\|^2 = O_p(n^{-1})$ and $\left\|G_{n,sc_0^1}(\pi_0)\right\|^2 = O_p(n^{-1})$. Therefore, it holds that

$$J_n\left(sc_0^1\right) = \min_{\pi \in \Pi} \left\|G_{n,sc_0^1}(\pi)\right\|^2 \leq \left\|G_{n,sc_0^1}(\pi_0)\right\|^2 = O_p\left(n^{-1}\right).$$

For any $sc^1 \in \mathscr{S}\mathscr{C}^1$, if $J_n(sc^1) \leq n^\lambda$, then $sc^1 \in \mathscr{S}\mathscr{C}_n^1$ occurs. Therefore, we have that for any $\lambda > -1$,

$$1 \geq \Pr\left(sc_0^1 \in \mathscr{S}\mathscr{C}_n^1\right) \geq \Pr\left(J_n\left(sc_0^1\right) \leq n^\lambda\right) \to 1,$$

which implies that $\Pr\left(sc_0^1 \in \mathscr{S}\mathscr{C}_n^1\right) \to 1$. The proof for $\Pr(sc_0^s \in \mathscr{S}\mathscr{C}_n^s) \to 1$ for $s = 2, ..., S$ follows the same argument as the one in the proof of Lemma B.1. Hence, we obtain that $\Pr(c_0 \in \mathscr{C}_n) \to 1$.

Second, we show that if $c_0 \in \mathscr{C}_n$, then $\Pr(\widehat{c} = c_0) \to 1$. For any $c \in \mathscr{C}_n$ and $c \notin \mathscr{C}\mathscr{I}$, Assumption (iii) in the lemma implies that

$$\min_{\pi \in \Pi} \|G_{n,c}(\pi)\|^2_{W_n(c)} \xrightarrow{p} \min_{\pi \in \Pi} \|G_c(\pi)\|^2_{W(c)} > 0,$$

where the inequality follows from Lemma 5.1 and Assumption (ii). For any $c \in \mathscr{C}$, define

$$J_n^\dagger(c) \equiv \min_{\pi \in \Pi} \|G_{n,c}(\pi)\|^2_{W_n(c)} - \rho_1(\|c\|_0)\,\kappa_{1,n}/n.$$

By Assumption (iv), $\kappa_{1,n} = o(n)$. Thus, for any $c \notin \mathscr{C}\mathscr{I}$,

$$J_n^\dagger(c) = \min_{\pi \in \Pi} \|G_{n,c}(\pi)\|^2_{W_n(c)} - \rho_1(\|c\|_0)\,\kappa_{1,n}/n > 0. \tag{B.17}$$

On the other hand, we have that $G_{c_0}(\pi_0) = \mathbf{0}$ by Lemma 5.1. Then by Assumption (i), it holds that $G_{n,c_0}(\pi_0) = o_p(1)$. Therefore, we obtain that

$$\min_{\pi \in \Pi} \|G_{n,c_0}(\pi)\|^2_{W_n(c_0)} \leq \|G_{n,c_0}(\pi_0)\|^2_{W_n(c_0)} = O_p\left(n^{-1}\right).$$

Together with Assumption (iv), it holds that

$$J_n^{\dagger}(c_0) = \min_{\pi \in \Pi} \|G_{n,c_0}(\pi)\|^2_{W_n(c_0)} - \rho_1 (\|c_0\|_0)\, \kappa_{1,n}/n = o_p(1)\,. \tag{B.18}$$

(B.17) and (B.18) imply that for any $c \notin \mathscr{C}\mathscr{I}$, $\Pr\left(J_n^{\dagger}(c_0) < J_n^{\dagger}(c)\right) \to 1$.

At the same time, for any $c \in \mathscr{C}\mathscr{I}$, we have $\min_{\pi \in \Pi} \|G_{n,c}(\pi)\|^2_{W_n(c)} = O_p(n^{-1})$. Assumption (iv) implies that

$$\min_{\pi \in \Pi} n \|G_{n,c}(\pi)\|^2_{W_n(c)} / \kappa_{1,n} = o_p(1)\,, \tag{B.19}$$

for any $c \in \mathscr{C}\mathscr{I}$. For any $c \in \mathscr{C}\mathscr{I}$ and $c \neq c_0$, Lemma 5.1 implies that $\|c\|_0 < \|c_0\|_0$. Then, by (B.19), we have that

$$n\left[J_n^{\dagger}(c_0) - J_n^{\dagger}(c)\right] = \left[\rho_1(\|c\|_0) - \rho_1(\|c_0\|_0)\right]\kappa_{1,n} + o_p(\kappa_{1,n}) \to -\infty.$$

Therefore, for any $c \in \mathscr{C}\mathscr{I}$ and $c \neq c_0$, $\Pr\left(J_n^{\dagger}(c_0) < J_n^{\dagger}(c)\right) \to 1$. Combined with the previous result, we have that for any $c \neq c_0$, $\Pr\left(J_n^{\dagger}(c_0) < J_n^{\dagger}(c)\right) \to 1$. Thus, if $c_0 \in \mathscr{C}_n$, then $\Pr(\widehat{c} = c_0) \to 1$.

Hence, we have shown that $\Pr(c_0 \in \mathscr{C}_n) \to 1$ and $\Pr(\widehat{c} = c_0) \to 1$ if $c_0 \in \mathscr{C}_n$. We can conclude that $\Pr(\widehat{c} = c_0) \to 1$. The proof for $\widehat{\pi} \xrightarrow{p} \pi_0$ follows the same proof for Lemma B.1. □

Proof of Theorem A.1: The proof is similar to the proof of Theorem 3.1 with Lemma B.1 replaced by Lemma B.8. Assumption (iv) in Lemma B.8 is imposed by Assumption A.3. □

Proof of Theorem A.2: We have that the numbers of mixing components are in an ascending order for $\mathbf{z}^1, \ldots, \mathbf{z}^l$. Let l_e denote the number of observed states where multiple equilibria exist. We have that there is no multiple equilibria on $\mathbf{z}^1, \ldots, \mathbf{z}^{l-l_e}$; and there are multiple equilibria on $\mathbf{z}^{l-l_e+1}, \ldots, \mathbf{z}^l$. Let $s^{\ddagger}$ be the smallest value such that $l_1 + s^{\ddagger}\Delta \leq l - l_e$. Because l_e is not a function of l by Assumption A.4, the numbers of elementary operations from Step $\left(s^{\ddagger}+1\right)$ to Step S is not a function of l either. Therefore, we only need to show that the time complexity of Step 1 to Step $s^{\ddagger}$ is linear in l with probability approaching one as $n \to \infty$ for all payoffs except for a set of Lebesgue measure zero.

Because no multiple equilibria exists on $\mathbf{z}^1, \ldots, \mathbf{z}^{l-l_e}$, all selection vector $c \in \mathscr{C}\mathscr{I}$ share the same first $(l - l_e)\,|\Omega_{\mathbf{z}^1}|$ components. As a result, from Step 1 to Step $s^{\ddagger}$, there is only one sub-selection vector sc^s in each step that shares the same first $[l_1 + (s-1)\Delta] |\Omega_{\mathbf{z}^1}|$ components as selection vectors in $\mathscr{C}\mathscr{I}$. Then, from Step 1 to Step $s^\ddagger$, the MMS procedure works almost the same the MMS procedure for the Simple Game. The only difference is that in step $s = 1, \ldots, s^\ddagger$, the input set $\mathscr{S}\mathscr{C}^s$ has cardinality $2^{|\Omega_{\mathbf{z}^1}|-1} \left(2^{|\Omega_{\mathbf{z}^1}|} - 1\right)^{l_s - 1} \prod_{i=1}^{s-1} \alpha_i^\star$, where $\alpha_s^\star \equiv |\mathscr{S}\mathscr{C}_n^s| / |\mathscr{S}\mathscr{C}^s|$. Since we set $\alpha = \left(2^{|\Omega_{\mathbf{z}^1}|} - 1\right)^{-\Delta}$, following the same proof as Lemma B.2, we have that the time complexity of the MMS procedure is linear in l for the first $s^\ddagger$ steps, assuming that the condition $\alpha_s^\star = \alpha$ for all $s = s^\dagger, .., s^\ddagger$ holds for some $s^\dagger$ independent of l. Applying the same argument in the proofs of Lemma B.3 and of Theorem 3.2, it can be shown that the required condition holds for with probability approaching one as $n \to \infty$ for all payoffs except for a set of Lebesgue measure zero. The result of the space complexity of the MMS procedure follows from the exact same argument. □

Lemma B.9. *Let l_e denote the number of observed states with multiple equilibria and $\left|\overline{\Omega}\right| \equiv \max\left(|\Omega_{\mathbf{z}^1}|, \ldots, |\Omega_{\mathbf{z}^l}|\right)$. Assume that l_e grows with l at a rate slower than $\log_{2^{|\overline{\Omega}|}} [h_e(l)]$ for some polynomial function $h_e(\cdot)$. Then with probability approaching one as $n \to \infty$, for all payoffs except for a set of Lebesgue measure zero, the time complexity of the MMS procedure for the General Game is at most a polynomial function of l.*

Proof of Lemma B.9: Same to the proof of Theorem A.2, since the numbers of mixing components are in an ascending order for $\mathbf{z}^1, \ldots, \mathbf{z}^l$, there is no multiple equilibria on $\mathbf{z}^1, \ldots, \mathbf{z}^{l-l_e}$; and there are multiple equilibria on $\mathbf{z}^{l-l_e+1}, \ldots, \mathbf{z}^l$. Following the proof of Theorem A.2, the time complexity of the MMS procedure when applying to $\mathbf{z}^1, \ldots, \mathbf{z}^{l-l_e}$ is upper bounded by a linear function of $l - l_e$. Denote the function as $h(l - l_e)$. The time complexity of the MMS procedure when applying to $\mathbf{z}^{l-l_e+1}, \ldots, \mathbf{z}^l$ is less than the number of possible combinations of selecting mixing components on $\mathbf{z}^{l-l_e+1}, \ldots, \mathbf{z}^l$: $\prod_{s=1}^{l_e} \left(2^{|\Omega_{\mathbf{z}^{l-l_e+s}}|} - 1\right)$. The total time complexity is therefore upper bounded by

$$h(l - l_e) \times \prod_{s=1}^{l_e} \left(2^{|\Omega_{\mathbf{z}^{l-l_e+s}}|} - 1\right) \leq h(l)\, 2^{|\overline{\Omega}| l_e} \leq h(l)\, h_e(l),$$

where the first equality holds because $h(\cdot)$ is an increasing function and by the definition of $\left|\overline{\Omega}\right|$, and the second inequality holds by the requirement on l_e stated in the lemma. Because $h(l)\, h_e(l)$ is a polynomial function of l, the claimed lemma holds. □

Proof of Theorem A.3: Define $JJ_n(K, \mathbf{S}_K)$ as

$$JJ_n(K, \mathbf{S}_K) = \sum_{j=1}^{K} \sum_{\widehat{\pi}^s \in S_{K,j}} \|\widehat{\pi}^s - \mu_{K,j}\|^2 + \rho_2(K)\,\kappa_{2,n}/n.$$

Then $\left(\widehat{K}, \widehat{\mathbf{S}}_{\widehat{K}}\right) = \arg\min_{K \leq |\Omega_{\mathbf{z}^1}|, \mathbf{S}_K \in \mathfrak{S}_K} JJ_n(K, \mathbf{S}_K)$. Denote $\mathbf{S}^*_{|\mathscr{K}|}$ as the correct partition. First, we show that $\Pr\left(\widehat{\mathbf{S}}_{|\mathscr{K}|} = \mathbf{S}^*_{|\mathscr{K}|}\right) \to 1$, where

$$\widehat{\mathbf{S}}_{|\mathscr{K}|} = \arg \min_{\mathbf{S}_{|\mathscr{K}|} \in \mathfrak{S}_{|\mathscr{K}|}} JJ_n\left(|\mathscr{K}|, \mathbf{S}_{|\mathscr{K}|}\right).$$

Then, we prove that $\Pr\left(\widehat{K} = |\mathscr{K}|\right) \to 1$. At last, we show that $\widehat{t\pi}^k$ with $k = 1, \ldots, \widehat{K}$ are consistent estimators.

Given the cardinality of partition $|\mathscr{K}|$, the term $\rho_2(|\mathscr{K}|)\,\kappa_{2,n}/n$ is the same for all possible partitions and converges to zero by Assumption A.5. Theorem A.1 shows that for $h = 1, ..., |\Omega_{\mathbf{z}^1}|$, we have $\widehat{\pi}^h \xrightarrow{p} \pi_0^h$. As a result, given the correct partition $\mathbf{S}^*_{|\mathscr{K}|}$, all $\widehat{\pi}^h$ within the same set have the same probability limit. Therefore, $JJ_n\left(|\mathscr{K}|, \mathbf{S}^*_{|\mathscr{K}|}\right) \xrightarrow{p} 0$. On the other hand, if a partition is incorrect, then $\left\|\widehat{\pi}^s - \mu_{|\mathscr{K}|,j}\right\|^2$ does not converge in probability to zero for at least one j. Thus, we conclude that $\Pr\left(\widehat{\mathbf{S}}_{|\mathscr{K}|} = \mathbf{S}^*_{|\mathscr{K}|}\right) \to 1$.

Based on the above discussion and the properties of $\rho_2(\cdot)$ and $\kappa_{2,n}$ by Assumption A.5, we have that given $\widehat{\mathbf{S}}_{|\mathscr{K}|}$, $JJ_n\left(|\mathscr{K}|, \widehat{\mathbf{S}}_{|\mathscr{K}|}\right) \xrightarrow{p} 0$. For any K, define $\widehat{\mathbf{S}}_K$ as $\arg\min_{\mathbf{S}_K \in \mathfrak{S}_K} JJ_n(K, \mathbf{S}_K)$. By the definition of $t\pi_0^k$ for $k = 1, \ldots, |\mathscr{K}|$, $t\pi_0^{k_1} \neq t\pi_0^{k_2}$ for $k_1 \neq k_2$. In consequence, if $K < |\mathscr{K}|$, then there must exists some j such that $\widehat{\pi}^s \in \widehat{S}_{K,j}$ have different probability limits. Because $\rho_2(K)\,\kappa_{2,n}/n \to 0$ for any K, we have that $JJ_n\left(K, \widehat{\mathbf{S}}_K\right) \xrightarrow{p} \delta > 0$ for any $K < |\mathscr{K}|$. Thus, $\Pr\left(\widehat{K} < |\mathscr{K}|\right) \to 0$. On the other hand, if $K > |\mathscr{K}|$,

$$n\left[JJ_n\left(|\mathscr{K}|, \widehat{\mathbf{S}}_{|\mathscr{K}|}\right) - JJ_n\left(K, \widehat{\mathbf{S}}_K\right)\right] = \left[\rho_2(|\mathscr{K}|) - \rho_2(K)\right]\kappa_{2,n} + o_p(\kappa_{2,n}) \to -\infty,$$

where the first equality holds because $\widehat{\pi}^h - \pi_0^h = O_p\left(n^{-1/2}\right)$. Therefore, $\Pr\left(\widehat{K} > |\mathscr{K}|\right) \to 0$. Hence, we have $\Pr\left(\widehat{K} = |\mathscr{K}|\right) \to 1$. Together with the previous result that $\Pr\left(\widehat{\mathbf{S}}_{|\mathscr{K}|} = \mathbf{S}^*_{|\mathscr{K}|}\right) \to 1$, we conclude that

$$\Pr\left(\widehat{\mathbf{S}}_{\widehat{K}} = \mathbf{S}^*_{|\mathscr{K}|}\right) = \Pr\left(\widehat{K} = |\mathscr{K}|, \widehat{\mathbf{S}}_{\widehat{K}} = \mathbf{S}^*_{|\mathscr{K}|}\right) \to 1.$$

By definition, $\widehat{t\pi}^j = \frac{1}{\left|\widehat{S}_{\widehat{K},j}\right|}\sum_{\widehat{\pi}^h \in \widehat{S}_{\widehat{K},j}} \widehat{\pi}^h$ for $j = 1, \ldots, \widehat{K}$. Since $\widehat{\pi}^h \xrightarrow{p} \pi_0^h$ for $h = 1, ..., |\Omega_{\mathbf{z}^1}|$ by Theorem A.1, there exist one permutation of $t\pi_0^k$ for $k = 1, \ldots, |\mathscr{K}|$, denoted as $t\pi_0^{(k)}$, such that

$$\Pr\left(\sum_{k=1}^{|\mathscr{K}|}\left\|\widehat{t\pi}^k - t\pi_0^{(k)}\right\| < \delta \mid \widehat{\mathbf{S}}_{\widehat{K}} = \mathbf{S}^*_{|\mathscr{K}|}\right) \to 1$$

for any $\delta > 0$. Because $\Pr\left(\widehat{\mathbf{S}}_{\widehat{K}} = \mathbf{S}^*_{|\mathscr{K}|}\right) \to 1$, it holds that

$$\begin{aligned}\Pr\left(\sum_{k=1}^{|\mathscr{K}|}\left\|\widehat{t\pi}^k - t\pi_0^{(k)}\right\| < \delta\right) &\geq \Pr\left(\sum_{k=1}^{|\mathscr{K}|}\left\|\widehat{t\pi}^k - t\pi_0^{(k)}\right\| < \delta, \widehat{\mathbf{S}}_{\widehat{K}} = \mathbf{S}^*_{|\mathscr{K}|}\right)\\ &= \Pr\left(\sum_{k=1}^{|\mathscr{K}|}\left\|\widehat{t\pi}^k - t\pi_0^{(k)}\right\| < \delta \mid \widehat{\mathbf{S}}_{\widehat{K}} = \mathbf{S}^*_{|\mathscr{K}|}\right)\Pr\left(\widehat{\mathbf{S}}_{\widehat{K}} = \mathbf{S}^*_{|\mathscr{K}|}\right)\\ &\to 1.\end{aligned}$$

Thus, $\widehat{t\pi}^1, \ldots, \widehat{t\pi}^{\widehat{K}}$ are consistent estimators for $t\pi_0^k$ with $k = 1, \ldots, |\mathscr{K}|$. □

Proof of Theorem A.4: The proof is similar to the proof of Theorem 4.1. Assumption A.6 (i) serves the same role as Assumption 4.1 (i) for Theorem 4.1. We omit the superscript s when there is no confusion. Without loss of generality, suppose $\mathscr{C}\mathscr{I}$ has $(q+1)$ elements: $\mathscr{C}\mathscr{I} = \{c_0, c_1, \ldots, c_q\}$. Following similar arguments in the proof of Lemma B.4, we have that any $\xi \in \Xi_R$ and the parameter sequence $\{\xi_n\} \in \Xi_R(\xi)$, $\|\sqrt{n}G_{n,c}(\pi_0)\|^2_{W_n(c)} \xrightarrow{d} \chi^2_{\left[\|c\|_0 - l_\pi + l_R\right]}$ for any $c \in \mathscr{C}\mathscr{I}$. As a result,

$$\min_{R\pi=r}\left\|\sqrt{n}G_{n,c}(\pi)\right\|^2_{W_n(c)} \xrightarrow{d} F_c \overset{d}{\leq} \chi^2_{\left[\|c\|_0 - l_\pi + l_R\right]},$$

where F_c is some tight limiting distribution and $\overset{d}{\leq}$ denotes stochastic dominance. It holds that for any $\xi \in \Xi_R$ and the parameter sequence $\{\xi_n\} \in \Xi_R(\xi)$,

$$T_n \xrightarrow{d} T \overset{d}{\leq} \min\left\{\chi^2_{\left[\|c_0\|_0 - (l_\pi - l_R)\right]}, \ldots, \chi^2_{\left[\|c_q\|_0 - (l_\pi - l_R)\right]}\right\} \overset{d}{\leq} \chi^2_{[Jl-(l_\pi - l_R)]},$$

where the last inequality holds because $\min_{c\in\mathscr{C}\mathscr{I}}\|c\|_0 = Jl$. Thus, using $\chi^2_{[Jl-(l_\pi-l_R)],1-\alpha}$ as the critical value achieves asymptotic size control. □

Lemma B.10. *Under Assumptions 5.3 and A.7, the unique solution to $G_c(\pi) = \mathbf{0}$ for any $c \in \mathscr{C}\mathscr{I}^s$ is the true payoff vector π_0^s.*

Proof of Lemma B.10: By Assumption 5.3 (ii), $G_c(\pi) = \mathbf{0}$ has no solution for any $c \in \mathscr{C}^s$ that selects different latent states. As a result, any $c \in \mathscr{C}\mathscr{I}^s$ must select the same latent state. The payoff vector corresponding to the selected latent state is the solution to the system of linear equations: $G_c(\pi_0^s) = \mathbf{0}$ for any $c \in \mathscr{C}\mathscr{I}^s$. Because $G_c(\pi) = \mathbf{0}$ for any $c \in \mathscr{C}\mathscr{I}^s$ has a unique solution by Assumption A.7, the result in the lemma holds. $\square$

Proof of Theorem A.5: The asymptotic properties of $G_{n,c}(\pi)$ for some given c and π are the same as that of $G_{n,c}(\pi)$ in Theorem 4.1 given the assumptions in Theorem A.4. We omit the superscript s when there is no confusion.

Firstly, by Lemma 5.1 and Assumption A.2, for any $c \notin \mathscr{C}\mathscr{I}$, $\min_\pi \|\sqrt{n}G_{n,c}(\pi)\|^2_{W_n(c)}$ diverges to infinity. Secondly, by Lemma B.10, $G_c(\pi) = \mathbf{0}$ has the unique solution π_0 for any $c \in \mathscr{C}\mathscr{I}$. When $\xi \notin \Xi_R$, for all $c \in \mathscr{C}\mathscr{I}$, $\min_{R\pi=r} \|\sqrt{n}G_{n,c}(\pi)\|^2_{W_n(c)}$ diverges to infinity. Since $\chi^2_{[Jl-l_\pi+l_R],1-\alpha}$ is finite, it holds that $\lim_{n\to\infty} \Pr_\xi\left(T_n > \chi^2_{[Jl-l_\pi+l_R],1-\alpha}\right) = 1$. We conclude the result in the theorem. $\square$

Appendix C Xiao (2018)'s CCP Estimator and Identification in the General Game

In this section, we first present the method developed by Xiao (2018) for Step-1 identification and estimation for both the Simple Game and the General Game. Then we discuss the equivalence between Step-2 identification conditions in the General Game and Aguirregabiria and Mira (2019)'s necessary and sufficient condition.

C.1 Identification of the CCPs in the Simple Game

Recall some notations from Section 2 of the paper. For $i = 1, 2, 3$, $p_i(\mathbf{z}, A) \equiv \Pr(d_i = 1 \mid \mathbf{z}, A)$, $p_i(\mathbf{z}, B) \equiv \Pr(d_i = 1 \mid \mathbf{z}, B)$, $p^A(\mathbf{z}) \equiv \Pr(k = A \mid \mathbf{z})$, and $p^B(\mathbf{z}) \equiv \Pr(k = B \mid \mathbf{z})$. For $k, k' \in \{A, B\}$ and $k \neq k'$, let

$$\mathbf{P}_{i\mathbf{z}} = \begin{bmatrix} p_i(\mathbf{z}, k) & p_i(\mathbf{z}, k') \\ 1 - p_i(\mathbf{z}, k) & 1 - p_i(\mathbf{z}, k') \end{bmatrix} \text{ for } i = 1, 2, 3.$$

Define the vector of mixing weights as $W_{k|\mathbf{z}} = \left(p^k(\mathbf{z}), p^{k'}(\mathbf{z})\right)^\top$. Its diagonal form is written as $D_{k|\mathbf{z}} = \text{diag}\left(W_{k|\mathbf{z}}^\top\right)$. Let the diagonal matrix containing player i's CCPs of choosing action $d_i = 0, 1$ on two latent states be

$$D^i_{1\mathbf{z}} \equiv \text{diag}\left(p_i\left(\mathbf{z}, k\right), p_i\left(\mathbf{z}, k'\right)\right) \text{ and } D^i_{0\mathbf{z}} \equiv \text{diag}\left(1 - p_i\left(\mathbf{z}, k\right), 1 - p_i\left(\mathbf{z}, k'\right)\right).$$

Consider identifying CCPs for player 1 on observed state $\mathbf{z}$. Define the following population contingency tables:

$$\begin{aligned}
A^{12}_{1\mathbf{z}} &\equiv \begin{bmatrix} \Pr\left(\left(d_1, d_2, d_3\right) = (1,1,1) \mid \mathbf{z}\right), & \Pr\left(\left(d_1, d_2, d_3\right) \equiv (1,0,1) \mid \mathbf{z}\right) \\ \Pr\left(\left(d_1, d_2, d_3\right) = (0,1,1) \mid \mathbf{z}\right), & \Pr\left(\left(d_1, d_2, d_3\right) = (0,0,1) \mid \mathbf{z}\right) \end{bmatrix}, \\
A^{12}_{\mathbf{z}} &\equiv \begin{bmatrix} \Pr\left(\left(d_1, d_2\right) = (1,1) \mid \mathbf{z}\right), & \Pr\left(\left(d_1, d_2\right) = (1,0) \mid \mathbf{z}\right) \\ \Pr\left(\left(d_1, d_2\right) = (0,1) \mid \mathbf{z}\right), & \Pr\left(\left(d_1, d_2\right) = (0,0) \mid \mathbf{z}\right) \end{bmatrix}, \\
A^{13}_{\mathbf{z}} &\equiv \begin{bmatrix} \Pr\left(\left(d_1, d_3\right) = (1,1) \mid \mathbf{z}\right), & \Pr\left(\left(d_1, d_3\right) = (1,0) \mid \mathbf{z}\right) \\ \Pr\left(\left(d_1, d_3\right) = (0,1) \mid \mathbf{z}\right), & \Pr\left(\left(d_1, d_3\right) = (0,0) \mid \mathbf{z}\right) \end{bmatrix},
\end{aligned}$$

and $A^1_{\mathbf{z}} \equiv \left[\Pr\left(d_1 = 1 \mid \mathbf{z}\right), \Pr\left(d_1 = 0 \mid \mathbf{z}\right)\right]^\top$. The matrices $A^{12}_{1\mathbf{z}}$, $A^{12}_{\mathbf{z}}$, $A^{13}_{\mathbf{z}}$, and $A^1_{\mathbf{z}}$ are identified from data. By construction, the population contingency tables can be written as products of CCPs and mixing weights:

$$A^{12}_{1\mathbf{z}} = \mathbf{P}_{1\mathbf{z}} D^3_{1\mathbf{z}} D_{k|\mathbf{z}} \mathbf{P}^\top_{2\mathbf{z}}, \; A^{12}_{\mathbf{z}} = \mathbf{P}_{1\mathbf{z}} D_{k|\mathbf{z}} \mathbf{P}^\top_{2\mathbf{z}}, \; A^{13}_{\mathbf{z}} = \mathbf{P}_{1\mathbf{z}} D_{k|\mathbf{z}} \mathbf{P}^\top_{3\mathbf{z}}, \text{ and } A^1_{\mathbf{z}} = \mathbf{P}_{1\mathbf{z}} W_{k|\mathbf{z}}.$$

The CCPs are identified using the eigendecomposition method. First, CCPs for player 1 are identified as the eigenvectors (of the left hand side observable matrix) with column sum being 1: $A^{12}_{1\mathbf{z}} \left(A^{12}_{\mathbf{z}}\right)^{-1} = \mathbf{P}_{1\mathbf{z}} D^3_{1\mathbf{z}} \mathbf{P}^{-1}_{1\mathbf{z}}$ and the vector of mixing weights is identified as $W_{k|\mathbf{z}} = \left(\mathbf{P}_{1\mathbf{z}}\right)^{-1} A^1_{\mathbf{z}}$. Second, given the recovered $\mathbf{P}_{1\mathbf{z}}$, $W_{k|\mathbf{z}}$ (and $D_{k|\mathbf{z}}$), CCPs for players 2 and 3 are identified as $\mathbf{P}_{2\mathbf{z}} = \left(A^{12}_{\mathbf{z}}\right)^\top \left(D^\top_{k|\mathbf{z}} \mathbf{P}^\top_{1\mathbf{z}}\right)^{-1}$ and $\mathbf{P}_{3\mathbf{z}} = \left(A^{13}_{\mathbf{z}}\right)^\top \left(D^\top_{k|\mathbf{z}} \mathbf{P}^\top_{1\mathbf{z}}\right)^{-1}$.

Note that if we change the order of the two columns of the eigenvector matrix $\mathbf{P}_{1\mathbf{z}}$ and eigenvalue matrix $D^3_{1\mathbf{z}}$ at the same time, equation $A^{12}_{1\mathbf{z}} \left(A^{12}_{\mathbf{z}}\right)^{-1} = \mathbf{P}_{1\mathbf{z}} D^3_{1\mathbf{z}} \mathbf{P}^{-1}_{1\mathbf{z}}$ still holds and equations $W_{k|\mathbf{z}} = \left(\mathbf{P}_{1\mathbf{z}}\right)^{-1} A^1_{\mathbf{z}}$, $\mathbf{P}_{2\mathbf{z}} = \left(A^{12}_{\mathbf{z}}\right)^\top \left(D^\top_{k|\mathbf{z}} \mathbf{P}^\top_{1\mathbf{z}}\right)^{-1}$, and $\mathbf{P}_{3\mathbf{z}} = \left(A^{13}_{\mathbf{z}}\right)^\top \left(D^\top_{k|\mathbf{z}} \mathbf{P}^\top_{1\mathbf{z}}\right)^{-1}$ inherit the order of the unobserved states adopted in equation $A^{12}_{1\mathbf{z}} \left(A^{12}_{\mathbf{z}}\right)^{-1} = \mathbf{P}_{1\mathbf{z}} D^3_{1\mathbf{z}} \mathbf{P}^{-1}_{1\mathbf{z}}$. Thus the CCPs for three players are identified up to a common label swapping.

C.2 Identification of the CCPs in the General Game

The General Game could have more than three players and more than two mixing components. In such cases, the identification method developed in Xiao (2018) employs the eigendecomposition using group actions. Following Xiao (2018), we divide the N players into three groups, such that the third group has exactly one player for odd N and two players for even N, and each of the first two groups has $\widetilde{N}$ players. Thus, $N = 2\widetilde{N} + 1$ when N is odd; and $N = 2\widetilde{N} + 2$ when N is even. Player group i is denoted as g_i for $i = 1, 2, 3$. By definition, $\cup_{i=1}^{3} g_i = \{1, ..., N\}$. For each group, we create a group action variable, denoted by d_{g_1}, d_{g_2}, and d_{g_3}. We have $d_{g_1}, d_{g_2} \in \left\{0, ..., (J+1)^{\widetilde{N}} - 1\right\}$, and $d_{g_3} \in \{0, ..., J\}$ if there is one player in group 3 and $d_{g_3} \in \left\{0, ..., (J+1)^2 - 1\right\}$ if there are two players in group 3.

Denote the matrix composed of CCPs for group action d_{g_i} for $i = 1, 2$ on each latent state as

$$\mathbf{P}_{g_i \mathbf{z}} \equiv \left(\Pr\left(d_{g_i} = j \mid \mathbf{z}, k\right)\right)_{j=0,k=1}^{(J+1)^{\widetilde{N}}-1, |\mathscr{K}|}.$$

The assumption needed for identifying CCPs up to a label swapping is stated in the following.

Assumption C.1. *(i)* $N \geq 3$. *(ii)* $(J+1)^{\widetilde{N}} > |\Omega_{\mathbf{z}}|$ *for any* $\mathbf{z}$. *(iii) For each* $\mathbf{z}$, *there exists a partition* $(d_{g_1}, d_{g_2}, d_{g_3})$ *of joint actions* $(d_1, ..., d_N)$ *such that* $\mathbf{P}_{g_1 \mathbf{z}}$ *and* $\mathbf{P}_{g_2 \mathbf{z}}$ *both have full column rank.*

Assumption C.1 guarantees the identification of $|\Omega_{\mathbf{z}}|$ and the equilibrium CCPs (up to a label swapping). Define $A_{d\mathbf{z}}^{g_1 g_2}$ as the joint contingency table for player groups g_1 and g_2 fixing player group g_3's action at $d_{g_3} = d$. Let $A_{\mathbf{z}}^{gi}$ and $A_{\mathbf{z}}^{ig}$ be the joint contingency tables of some generic player group g and individual player i. For each $\mathbf{z}$, CCPs are identified following Xiao (2018). First, under (ii), Lemma 1 in Xiao (2018) applies to identify $|\Omega_{\mathbf{z}}|$ as $\operatorname{rank}\left(A_{\mathbf{z}}^{g_1 g_2}\right)$ for each observed $\mathbf{z}$. Given $|\Omega_{\mathbf{z}}|$, by summing up rows and columns of $A_{\mathbf{z}}^{g_1 g_2}$ (thus collapsing the actions of player group 1 and player group 2), we can create $A_{\mathbf{z}}^{\widetilde{g}_1 \widetilde{g}_2}$ with $\operatorname{rank}\left(A_{\mathbf{z}}^{\widetilde{g}_1 \widetilde{g}_2}\right) = |\Omega_{\mathbf{z}}|$, where $\widetilde{g}_1$ and $\widetilde{g}_2$ denote player group 1 and player group 2 with collapsed actions. Let $\mathbf{P}_{\widetilde{g}_i}$ denote the matrix storing CCPs for player group g_i with collapsed actions (each column correspond to a different ω). Second, we use eigendecomposition to identify $\mathbf{P}_{\widetilde{g}_1 \mathbf{z}}$ as $A_{d\mathbf{z}}^{\widetilde{g}_1 \widetilde{g}_2} \left(A_{\mathbf{z}}^{\widetilde{g}_1 \widetilde{g}_2}\right)^{-1} = \mathbf{P}_{\widetilde{g}_1 \mathbf{z}} D_{d\mathbf{z}}^{g_3} \left(\mathbf{P}_{\widetilde{g}_1 \mathbf{z}}\right)^{-1}$, where $D_{d\mathbf{z}}^{g_3}$ is the diagonal matrix storing the conditional choice probabilities of $d_{g_3} = d$ for all ω on the diagonal. The vector of mixing weights is identified as $W_{\mathbf{z}} = \left(\mathbf{P}_{\widetilde{g}_1 \mathbf{z}}\right)^{-1} A_{\mathbf{z}}^{\widetilde{g}_1}$. Define $D_{\mathbf{z}} = \operatorname{diag}\left(W_{\mathbf{z}}^{\top}\right)$, the equilibrium CCPs for player group 2 with collapsed actions are then identified as $\mathbf{P}_{\widetilde{g}_2 \mathbf{z}} = \left(A_{\mathbf{z}}^{\widetilde{g}_1 \widetilde{g}_2}\right)^{\top} \left(D_{\mathbf{z}}^{\top} \mathbf{P}_{\widetilde{g}_1 \mathbf{z}}^{\top}\right)^{-1}$. For equilibrium CCPs of individual player $i \in g_2 \bigcup g_3$, we obtain $\mathbf{P}_{i\mathbf{z}}^{\top} = \left(\mathbf{P}_{\widetilde{g}_1 \mathbf{z}} D_{\mathbf{z}}\right)^{-1} A_{\mathbf{z}}^{\widetilde{g}_1 i}$. For equilibrium CCPs of individual player $i \in g_1$, we have $\mathbf{P}_{i\mathbf{z}}^{\top} = A_{\mathbf{z}}^{i \widetilde{g}_2} \left(D_{\mathbf{z}} \mathbf{P}_{\widetilde{g}_2 \mathbf{z}}^{\top}\right)^{-1}$.

C.3 $\sqrt{n}$-Consistency and Asymptotic Normality of the CCP Estimator

Both the root-n consistency and the asymptotic normality require the eigenvalues in the eigendecomposition be simple. In the Simple Game, such a condition holds automatically under Assumption 2.5, which implies that all players have different CCPs of choosing each action in different latent states. In the General Game, an additional assumption is needed for the consistency and the asymptotic normality:

Assumption C.2. *There exists a known group action for player group g_3 such that the corresponding equilibrium CCPs of choosing this group action is different across different values of ω.*

Under Assumptions 5.1-5.3 and Assumption C.2, the argument in Xiao (2018) delivers root-n consistency and asymptotic normality of the CCP estimator.

C.4 Further Discussions on Step-2 Identification in the General Game

We mentioned that Assumption 5.3 is equivalent to the necessary and sufficient condition proposed in Proposition 3 of Aguirregabiria and Mira (2019) except that we consider identification for each pair of player and exclusive state separately. In particular, Assumption 5.3 (i) holds if and only if for each $h = 1, ..., |\Omega_{\mathbf{z}^1}|$, there exists some $c^* \in \mathscr{C}\mathscr{I}^h$ such that Γ_{c^*} has full column rank (the condition in Proposition 3 of Aguirregabiria and Mira (2019)). This is due to the following reasoning. On the one hand, if there exists $c^* \in \mathscr{C}\mathscr{I}^h$ such that Γ_{c^*} has full column rank, then $\Gamma_{c_0^h}$ certainly has full column rank. This is because if c^* does not select the most number of rows, adding more rows does not decreases the column rank of a matrix; while if c^* selects the most number of rows then $c_0^* = c_0^h$. On the other hand, if $\Gamma_{c_0^h}$ has full column rank, then there exist $c^* = c_0^h \in \mathscr{C}\mathscr{I}$ such that Γ_{c^*} has full column rank.

Appendix D Additional Details on the Games in the Simulations

In this section we provide more details on the games used in the simulations.

D.1 Identification of Games in the Simulation

For the system of moment functions of Game 1, there are two parameters on each latent state, while the number of correct moments are 18, 27, 64, 100 and 288 respectively from Design 1 to Design 5. Assumption 2.6 for Step-2 identification is verified for the parameter values listed in Section D.2.

For Game 2, let the equilibrium CCP vector be:

$$\mathbf{p}\left(\mathbf{z}, x, k\right) \equiv \left(p_1\left(\mathbf{z}, x, k\right), p_2\left(\mathbf{z}, x, k\right), p_3\left(\mathbf{z}, x, k\right)\right)$$

for $\mathbf{z} \in \mathscr{Z}$ and $x \in \mathscr{X}$, where for $i = 1, 2, 3$, $p_i\left(\mathbf{z}, x, k\right) \equiv \Pr\left(d_i = 1 \mid \mathbf{z}, x, k\right)$. Step-1 identification of equilibrium CCPs and mixing weights makes use of the following system of equations: for $(d_1, d_2, d_3) \in \{0, 1\} \times \{0, 1\} \times \{0, 1\}$,

$$p\left(d_1, d_2, d_3 \mid \mathbf{z}, x\right) = \sum_{k \in \{A, B\}} \left[p^k\left(\mathbf{z}, x\right) \prod_{i=1}^{3} \left(p_i\left(\mathbf{z}, x, k\right)\right)^{d_i} \left(1 - p_i\left(\mathbf{z}, x, k\right)\right)^{1-d_i}\right]$$

and can proceed in exactly the same way as Step-1 for the Simple Game.

Step-2 identification is similar to that of the Simple Game. Let $l_z \equiv \prod_{i=1}^{3} |\mathscr{Z}_i|$ and $\left\{\mathbf{z}^1, \ldots, \mathbf{z}^{l_z}\right\}$ with $\mathbf{z}^t \equiv (z_1^t, z_2^t, z_3^t)$ be the l_z different values in $\mathscr{Z}_1 \times \mathscr{Z}_2 \times \mathscr{Z}_3$. Denote $l_x \equiv |\mathscr{X}|$ and $\mathscr{X} \equiv \{x^1, \ldots, x^{l_x}\}$. We can obtain the following system for player 1 via stacking two latent states on each $(\mathbf{z}, x)$ for $\mathbf{z} \in \left\{\mathbf{z}^1, \ldots, \mathbf{z}^{l_z}\right\}$ and $x \in \mathscr{X}$:

$$G\left(\pi\right) = \left[\overline{\pi}_1^\top\left(x^1\right), \ldots, \overline{\pi}_1^\top\left(x^{l_x}\right)\right]^\top - \left[\Gamma_1^\top\left(x^1\right), \ldots, \Gamma_1^\top\left(x^{l_x}\right)\right]^\top \pi, \text{ where}$$

$$\overline{\pi}_1\left(x\right) \equiv \left[\overline{\pi}_1\left(1, \mathbf{z}^1, x, k_1\right), \overline{\pi}_1\left(1, \mathbf{z}^1, x, k_1'\right), \ldots, \overline{\pi}_1\left(1, \mathbf{z}^{l_z}, x, k_{l_z}\right), \overline{\pi}_1\left(1, \mathbf{z}^{l_z}, x, k_{l_z}'\right)\right]^\top \text{ and}$$

$$\Gamma_1\left(x\right) \equiv \begin{bmatrix} \begin{bmatrix} x & z_1^1\left(p_2\left(\mathbf{z}^1, k_1\right) + p_3\left(\mathbf{z}^1, k_1\right)\right) \end{bmatrix}^\top, \begin{bmatrix} x & z_1^1\left(p_2\left(\mathbf{z}^1, k_1'\right) + p_3\left(\mathbf{z}^1, k_1'\right)\right) \end{bmatrix}^\top, \ldots, \\ \begin{bmatrix} x & z_1^{l_z}\left(p_2\left(\mathbf{z}^{l_z}, k_{l_z}\right) + p_3\left(\mathbf{z}^{l_z}, k_{l_z}\right)\right) \end{bmatrix}^\top, \begin{bmatrix} x & z_1^{l_z}\left(p_2\left(\mathbf{z}^{l_z}, k_{l_z}'\right) + p_3\left(\mathbf{z}^{l_z}, k_{l_z}'\right)\right) \end{bmatrix}^\top \end{bmatrix}^\top.$$

The dimensions for $\overline{\pi}$ and Γ are $2l_zl_x \times 1$ and $2l_zl_x \times 2$ respectively.

Given some selected latent state for $(\mathbf{z}^1, x^1)$, we could match this latent state across all exclusive and common observed states with the parameter spaces of the true selection vectors given by

$$\mathscr{C}^1 = \left\{\left[c_1, ..., c_{l_zl_x}\right]^\top : c_1 = [1, 0] \text{ and } c_t \in \{[1, 0], [0, 1]\} \text{ for } t \in \{2, \ldots, l_zl_x\}\right\} \text{ and}$$

$$\mathscr{C}^2 = \left\{\left[c_1, ..., c_{l_zl_x}\right]^\top : c_1 = [0, 1] \text{ and } c_t \in \{[1, 0], [0, 1]\} \text{ for } t \in \{2, \ldots, l_zl_x\}\right\}.$$

For c belonging to $\mathscr{C}^1$ or $\mathscr{C}^2$, we have $\|c\|_0 = l_z l_x$. Given a parameter space of the true selection vector with the first selected latent state being k, the unique solution to the system $G_{c_0}(\pi) = \mathbf{0}$ gives us $(\beta_{1k}, \delta_{1k})^\top$.

In simulation, $l_z l_x = 24, 36, 54, 81, 162$ respectively for each design specified in Section D.2. The number of unknowns in the system is strictly less than the number of correct moments. The parameter values used in the simulation satisfy Assumption 2.6 for Step-2 identification.

For Games 3-4, let the equilibrium CCP vector be:

$$\mathbf{p}(x,\omega) \equiv (p_1(x,\omega), p_2(x,\omega), p_3(x,\omega), p_4(x,\omega), p_5(x,\omega))$$

for $x \in \mathscr{X}$, where for $i = 1, \ldots, 5$, $p_i(x,\omega) \equiv \Pr(d_i = 1 \mid, x, \omega)$. Step-1 identification of equilibrium CCPs and mixing weights makes use of the following system of equations: for $(d_1, d_2, d_3, d_4, d_5) \in \{0,1\} \times \{0,1\} \times \{0,1\} \times \{0,1\} \times \{0,1\}$,

$$p(d_1, d_2, d_3, d_4, d_5 \mid x) = \sum_{\omega \in \Omega_x} \left[p(\omega|x) \prod_{i=1}^{5} (p_i(x,\omega))^{d_i} (1 - p_i(x,\omega))^{1-d_i} \right]$$

and can proceed in a similar way as Step-1 for the Simple Game. Overidentifying restrictions could be used to improve estimation accuracy.

Step-2 identification is similar to that of the Simple Game. Denote $l_x \equiv |\mathscr{X}|$ and $\mathscr{X} \equiv \{x^1, \ldots, x^{l_x}\}$. We can obtain the following system for player 1 via stacking composite latent variables on each x for $x \in \mathscr{X}$:

$$G(\pi) = \left[\overline{\pi}_1^\top(x^1), \ldots, \overline{\pi}_1^\top(x^{l_x})\right]^\top - \left[\Gamma_1^\top(x^1), \ldots, \Gamma_1^\top(x^{l_x})\right]^\top \pi, \text{ where}$$

$$\overline{\pi}_1(x) \equiv \left[\overline{\pi}_1(1, x, \omega(1,x)), \ldots, \overline{\pi}_1(1, x, \omega(|\Omega_x|, x))\right]^\top \text{ and}$$

$$\Gamma_1(x) \text{ is } \left[\begin{bmatrix} x & p(x, \omega(1,x)) \end{bmatrix}^\top, \ldots, \begin{bmatrix} x & p(x, \omega(|\Omega_x|, x)) \end{bmatrix}^\top\right]^\top \text{ in Game 3 and}$$

$$\left[\begin{bmatrix} x & (1+x^2)(2p(x,\omega(1,x)) - 1) \end{bmatrix}^\top, \ldots, \begin{bmatrix} x & (1+x^2)(2p(x,\omega(|\Omega_x|,x)) - 1) \end{bmatrix}^\top\right]^\top \text{ in Game 4.}$$

The dimensions for $\overline{\pi}$ and Γ are $(2(l_x - 1) + 3) \times 1$ and $(2(l_x - 1) + 3) \times 2$ when only observed state 8 has multiple equilibria. The dimensions for $\overline{\pi}$ and Γ are $(2(l_x - 2) + 6) \times 1$ and $(2(l_x - 2) + 6) \times 2$ when both observed state 8 and observed state 16 have multiple equilibria.

In the simulations $l_x = 18$. The number of unknowns in the system is strictly less than the number of correct moments. The parameter values used in the simulation satisfy Assumption 5.3 for Step-2 identification. Given some selected value of the composite latent variable for x^1, we could try to match its underlying latent state k across all common observed states with the parameter spaces of the true selection vectors specified as follows:

Design 1: 3 mixing components on the 8th observed state:

$$\mathscr{C}^1 = \left\{ \begin{array}{c} [c_1, ..., c_{l_x}]^\top : c_1 \in \{[1,0], [1,1]\} \text{ and } c_t \in \{[1,0], [0,1], [1,1]\} \text{ for } t \neq 8 \text{ and} \\ c_8 \in \{[1,0,0], [0,1,0], [0,0,1], [0,1,1], [1,0,1], [1,1,0], [1,1,1]\} \end{array} \right\}$$

$$\mathscr{C}^2 = \left\{ \begin{array}{c} [c_1, ..., c_{l_x}]^\top : c_1 \in \{[0,1], [1,1]\} \text{ and } c_t \in \{[1,0], [0,1], [1,1]\} \text{ for } t \neq 8 \text{ and} \\ c_8 \in \{[1,0,0], [0,1,0], [0,0,1], [0,1,1], [1,0,1], [1,1,0], [1,1,1]\} \end{array} \right\}$$

Design 2: 3 mixing components on both the 8th and the 16th observed state:

$$\mathscr{C}^1 = \left\{ \begin{array}{c} [c_1, ..., c_{l_x}]^\top : c_1 \in \{[1,0], [1,1]\} \text{ and } c_t \in \{[1,0], [0,1], [1,1]\} \text{ for } t \notin \{8,16\} \text{ and} \\ c_t \in \{[1,0,0], [0,1,0], [0,0,1], [0,1,1], [1,0,1], [1,1,0], [1,1,1]\} \text{ for } t \in \{8,16\} \end{array} \right\}$$

$$\mathscr{C}^2 = \left\{ \begin{array}{c} [c_1, ..., c_{l_x}]^\top : c_1 \in \{[0,1], [1,1]\} \text{ and } c_t \in \{[1,0], [0,1], [1,1]\} \text{ for } t \notin \{8,16\} \text{ and} \\ c_t \in \{[1,0,0], [0,1,0], [0,0,1], [0,1,1], [1,0,1], [1,1,0], [1,1,1]\} \text{ for } t \in \{8,16\} \end{array} \right\}$$

Given a parameter space of the true selection vector with the first selected underlying latent state being k, the unique solution to the system $G_{c_0}(\pi) = \mathbf{0}$ gives us $(\theta_k, \delta_k)^\top$.

D.2 Parameter Values in the Simulation

For Table I and Table II in Section 6 of the main paper, the parameter values in the payoff functions and the values for the observed state variables are set according to Design 1 of Game 1 below. For Table III, the parameter values in the payoff functions and the values for the observed state variables are set according to Designs 1-5 for both games below. Note that for all designs, the parameter values in the payoff functions stay the same for both games while the support of the observed state $\mathbf{z}$ changes.

Game 1: For Design 1, $\mathscr{Z}_1 = \{0.1, 0.8\}$, $\mathscr{Z}_2 = \{0.1, 0.8, 1.5\}$, and $\mathscr{Z}_3 = \{0.1, 0.8, 1.5\}$. For Design 2, $\mathscr{Z}_1 = \{0.1, 0.8, 1.5\}$, $\mathscr{Z}_2 = \{0.1, 0.8, 1.5\}$, and $\mathscr{Z}_3 = \{0.1, 0.8, 1.5\}$. For Design 3, $\mathscr{Z}_1 = \{0.1, 0.8, 1.3, 1.5\}$, $\mathscr{Z}_2 = \{0.1, 0.8, 1.3, 1.5\}$, and $\mathscr{Z}_3 = \{0.1, 0.8, 1.3, 1.5\}$. For Design 4, $\mathscr{Z}_1 = \{0.1, 0.8, 1.3, 1.5\}$, $\mathscr{Z}_2 = \{0.1, 0.8, 1.1, 1.3, 1.5\}$, and $\mathscr{Z}_3 = \{0.1, 0.8, 1.1, 1.3, 1.5\}$. And for Design 5, $\mathscr{Z}_1 = \{0.1, 0.7, 0.8, 1.2, 1.3, 1.4, 1.5\}$, $\mathscr{Z}_2 =$

$\{0.1, 0.7, 0.8, 1.1, 1.3, 1.5\}$, and $\mathscr{Z}_3 = \{0.1, 0.7, 0.8, 1.1, 1.3, 1.5\}$. The parameter values in the payoff functions are set according to Table D.1 below and are the same across Design 1-Design 5.

Table D.1: Parameter values in Game 1

δ_{1A}	δ_{1B}	δ_{2A}	δ_{2B}	δ_{3A}	δ_{3B}	θ_{1A}	θ_{1B}	θ_{2A}	θ_{2B}	θ_{3A}	θ_{3B}
-0.01	-5	-0.02	-5.5	-0.02	-5.5	2.2	0.4	2.5	0.4	2.5	0.4

Game 2: For Design 1, $\mathscr{X} = \{0.4, 0.7, 0.8\}$, $\mathscr{Z}_1 = \{0.5, 1.2\}$, $\mathscr{Z}_2 = \{0.5, 1.2\}$, and $\mathscr{Z}_3 = \{0.5, 1.2\}$. For Design 2, $\mathscr{X} = \{0.4, 0.7, 0.8\}$, $\mathscr{Z}_1 = \{0.5, 0.9, 1.2\}$, $\mathscr{Z}_2 = \{0.5, 1.2\}$, and $\mathscr{Z}_3 = \{0.5, 1.2\}$. For Design 3, $\mathscr{X} = \{0.4, 0.7, 0.8\}$, $\mathscr{Z}_1 = \{0.5, 0.9, 1.2\}$, $\mathscr{Z}_2 = \{0.5, 0.9, 1.2\}$, and $\mathscr{Z}_3 = \{0.5, 1.2\}$. For Design 4, $\mathscr{X} = \{0.4, 0.7, 0.8\}$, $\mathscr{Z}_1 = \{0.5, 0.9, 1.2\}$, $\mathscr{Z}_2 = \{0.5, 0.9, 1.2\}$, and $\mathscr{Z}_3 = \{0.5, 0.9, 1.2\}$. And for Design 5, $\mathscr{X} = \{0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$, $\mathscr{Z}_1 = \{0.5, 0.9, 1.2\}$, $\mathscr{Z}_2 = \{0.5, 0.9, 1.2\}$, and $\mathscr{Z}_3 = \{0.5, 0.9, 1.2\}$. The parameter values in the payoff functions are set according to Table D.1 below and are the same across Design 1-Design 5.

Table D.2: Parameter values in Game 2

δ_{1A}	δ_{1B}	δ_{2A}	δ_{2B}	δ_{3A}	δ_{3B}	β_{1A}	β_{1B}	β_{2A}	β_{2B}	β_{3A}	β_{3B}
-0.1	-2	-0.2	-1.9	-0.3	-1.8	3	0.1	2.6	0.2	2.7	0.1

For Table VI, the parameter values in the payoff functions and the values for the observed state variable are set according to Design 1-2 for Game 3 and Game 4 below.

Game 3: For Design 1, $x \in \mathscr{X} = \{-1.6, -1.5, -1.4, -1.3, -1.2, -0.8, -0.55, -0.45, -0.35, -0.3, -0.25, -0.2, -0.15, -0.1, -0.05, 0.05, 0.1, 0.15\}$, there are multiple equilibria on latent state A for $x = -1.5$ (in the simulation this is designated as the 8th observed state). For Design 2, $x \in \mathscr{X} = \{-1.6, -1.5, -1.4, -1.3, -1.2, -0.8, -0.55, -0.45, -0.35, -0.3, -0.25, -0.2, -0.15, -0.1, -0.05, 0.05, 0.1, 0.15\}$ there are multiple equilibria on latent state A for $x = -1.5$ and $x = -1.4$ (in the simulation these are designated as the 8th and 16th observed states, respectively); $\theta_B = 0.1, \delta_B = -0.2, \theta_A = 1.1, \delta_A = 3.2$ for both designs.

Game 4: For Design 1, $x \in \mathscr{X} = \{-1.25, -1.2, -1.15, -1.1, -1.05, -1, -0.95, -0.9, -0.8, -0.75, -0.7, -0.65, -0.6, -0.5, 0.4, 0.45, 0.55, 0.85\}$, there are multiple equilibria on latent state A for $x = 0.85$ (in the simulation this is designated as the 8th observed state). For Design 2, $x \in \mathscr{X} = \{-1.2, -1.15, -1.1, -1.05, -1, -0.95, -0.9, -0.8, -0.75, -0.7, -0.65, -0.6, -0.5, 0.4, 0.45, 0.5, 0.55, 0.85\}$,there are multiple equilibria on latent state A for $x = 0.85$ and $x = 0.5$ (in the simulation these are designated as the 8th and 16th observed states, respectively); $\theta_B = 2.4, \delta_B = -0.6, \theta_A = 2, \delta_A = 4$ for both designs;

References

Aguirregabiria, V. and P. Mira (2019): "Identification of Games of Incomplete Information with Multiple Equilibria and Common Unobserved Heterogeneity," *Quantitative Economics*, 10, 1659–1701.

Cox, D., J. Little, and D. OShea (2013): *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Springer Science & Business Media.

Fan, Y., S. Jiang, and X. Shi (2024): "Estimation and Inference in Games of Incomplete Information with Unobserved Heterogeneity and Large State Space," Working Paper.

Hansen, B. (2021): *Econometrics*, Princeton University Press, forthcoming.

Saks, S. (1937): *Theory of the Integral*, Polish Mathematical Society, Warszawa.

Walter, R. (1987): *Real and Complex Analysis*, 3rd ed. New York: McGraw-HiII Book Company.

Xiao, R. (2018): "Identification and Estimation of Incomplete Information Games with Multiple Equilibria," *Journal of Econometrics*, 203, 328–343.