Supplemental Appendix to

MAXIMUM LIKELIHOOD ESTIMATION AND UNIFORM INFERENCE WITH SPORADIC IDENTIFICATION FAILURE

By

Donald W. K. Andrews and Xu Cheng

October 2011

COWLES FOUNDATION DISCUSSION PAPER NO. 1824

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Supplemental Appendices

for

Maximum Likelihood Estimation and Uniform Inference with Sporadic Identification Failure

Donald W. K. Andrews
Cowles Foundation for Research in Economics
Yale University

Xu Cheng
Department of Economics
University of Pennsylvania

First Draft: August, 2007
Revised: October 14, 2011

11. Outline

This Supplement includes three Supplemental Appendices (denoted C, D, and E) to the paper "Maximum Likelihood Estimation and Uniform Inference with Sporadic Identification Failure." Supplemental Appendix C provides additional numerical results to those provided in the main paper for both the smooth transition autoregressive (STAR) model and the nonlinear binary choice model. Supplemental Appendix D verifies Assumptions S1-S4, B1, B2, C6, V1, and V2 for the nonlinear binary choice model. Supplemental Appendix E does likewise for the STAR model.

We let AC1 abbreviate the paper Andrews and Cheng (2007) "Estimation and Inference with Weak, Semi-strong, and Strong Identification."

12. Supplemental Appendix C: Numerical Results

Table S-1 compares the finite-sample ($n = 500$) coverage probabilities of the null-imposed robust CI's for β in the STAR model with true and estimated values of ζ. (See the end of the STAR-model numerical-results section in the main paper for further discussion.)

Figures S-1 and S-2 report asymptotic and finite-sample ($n = 500$) densities of the estimators for β and π in the STAR model when $\pi_0 = -3.0$. Figures S-3 to S-6 report asymptotic and finite-sample (n=500) densities of the t and QLR statistics for β and π in the STAR model when $\pi_0 = -1.5$. Figures S-7 and S-8 report CP's of nominal 0.95 standard and robust $|t|$ and QLR CI's for β and π in the STAR model when $\pi_0 = -3.0$.

Figures S-9 to S-16 are analogous to Figures S-1 to S-8 but for the binary choice model. The true values of π considered are $\pi_0 = 1.5$ and $\pi_0 = 2.0$.

Table S-1. Finite-Sample Coverage Probabilities of Null-Imposed Robust CI's for β in the STAR Model with True and Estimated Values of ζ, $n = 500$, $\pi_0 = -1.5$[38]

b	0	1	2	3	4	5	6	7	8	9	10	11	12
t_ζ_0	0.939	0.950	0.946	0.947	0.948	0.947	0.944	0.946	0.949	0.949	0.950	0.947	0.956
$t\_\widehat{\zeta}$	0.936	0.951	0.946	0.947	0.947	0.947	0.949	0.944	0.947	0.947	0.947	0.947	0.957
QLR_ζ_0	0.923	0.932	0.930	0.925	0.923	0.924	0.916	0.921	0.923	0.926	0.929	0.932	0.933
$QLR\_\widehat{\zeta}$	0.920	0.935	0.927	0.924	0.926	0.919	0.915	0.908	0.915	0.922	0.925	0.924	0.929

[38] The simulation is conducted with the null value of b and the true value of π imposed so that the

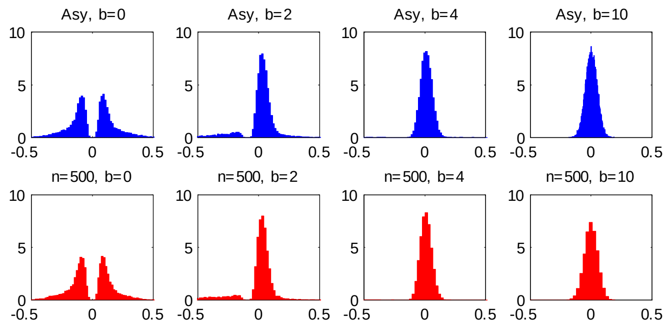

Figure S-1. Asymptotic and Finite-Sample ($n = 500$) Densities of the Estimator of β in the STAR Model when $\pi_0 = -3.0$.

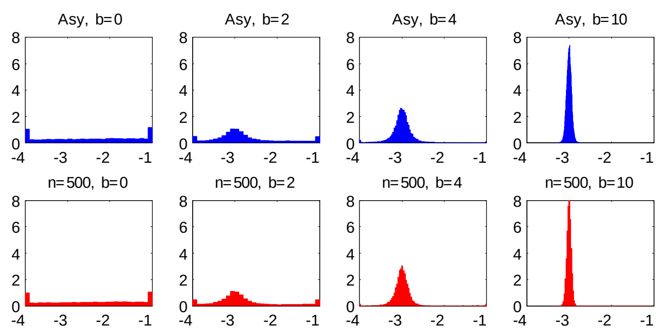

Figure S-2. Asymptotic and Finite-Sample ($n = 500$) Densities of the Estimator of π in the STAR Model when $\pi_0 = -3.0$.

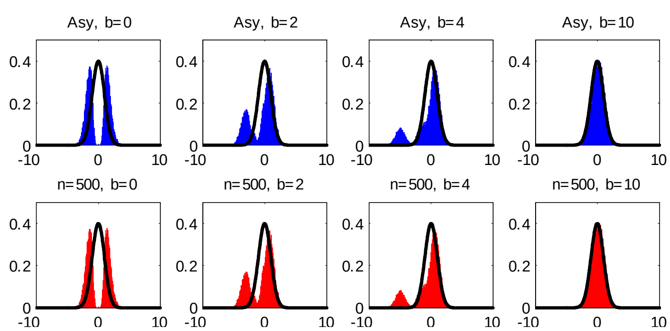

Figure S-3. Asymptotic and Finite-Sample ($n = 500$) Densities of the t Statistic for β in the STAR Model when $\pi_0 = -1.5$ and the Standard Normal Density (Black Line).

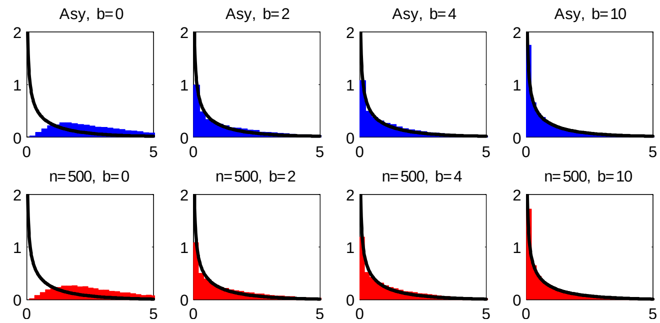

Figure S-4. Asymptotic and Finite-Sample (n=500) Densities of the QLR Statistic for β in the STAR Model when $\pi_0 = -1.5$ and the χ^2_1 Density (Black Line).

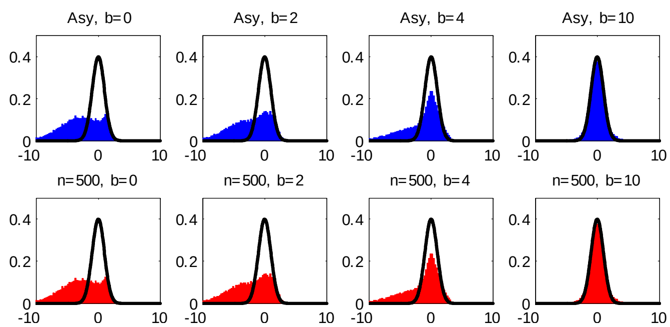

Figure S-5. Asymptotic and Finite-Sample ($n = 500$) Densities of the t Statistic for π in the STAR Model when $\pi_0 = -1.5$ and the Standard Normal Density (Black Line).

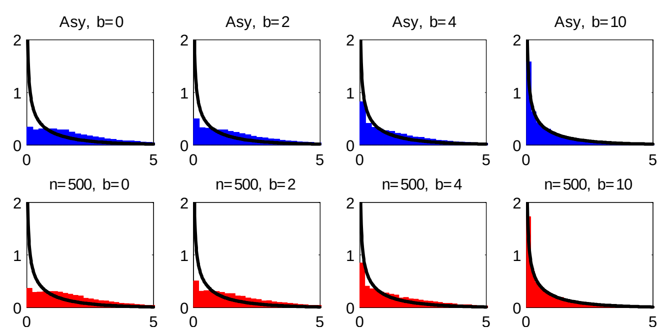

Figure S-6. Asymptotic and Finite-Sample (n=500) Densities of the QLR Statistic for π in the STAR Model when $\pi_0 = -1.5$ and the χ^2_1 Density (Black Line).

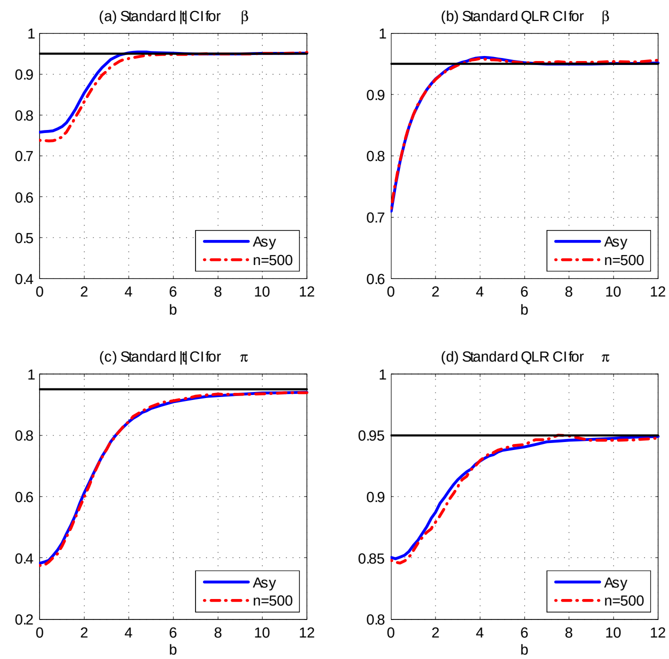

Figure S-7. Coverage Probabilities of Standard $|t|$ and QLR CI's for β and π in the STAR Model when $\pi_0 = -3.0$.

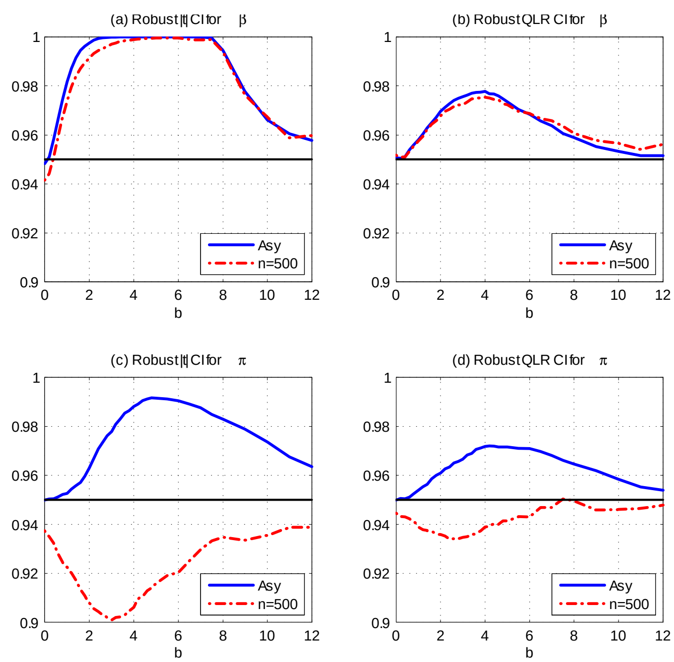

Figure S-8. Coverage Probabilities of Robust $|t|$ and QLR CI's for β and π in the STAR Model when $\pi_0 = -3.0$, $\kappa = 2.5$, $D = 1$, and $s(x) = \exp(-2x)$.

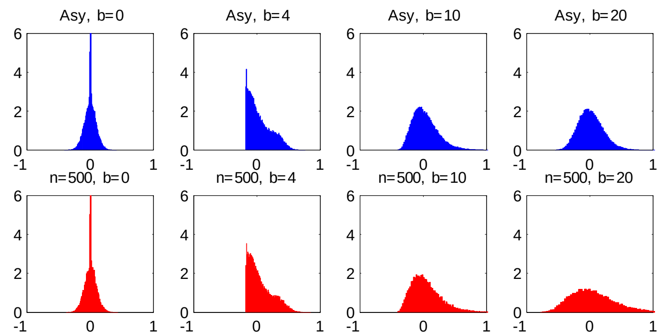

Figure S-9. Asymptotic and Finite-Sample (n=500) Densities of the Estimator of β in the Binary Choice Model when $\pi_0 = 2.0$.

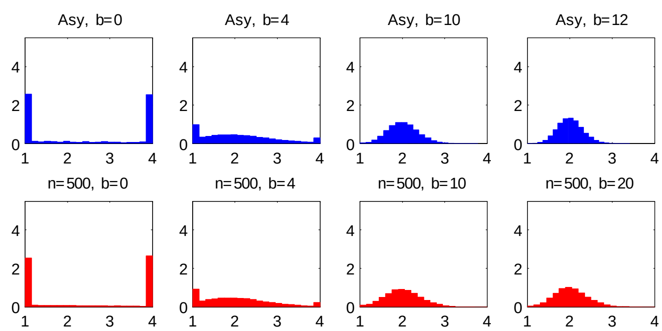

Figure S-10. Asymptotic and Finite-Sample (n=500) Densities of the Estimator of π in the Binary Choice Model when $\pi_0 = 2.0$.

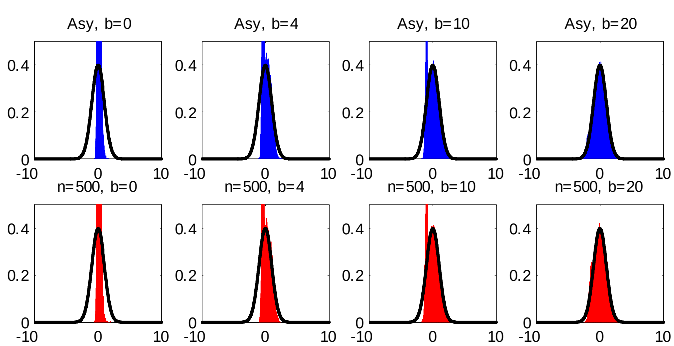

Figure S-11. Asymptotic and Finite-Sample (n=500) Densities of the t Statistic for π in the Binary Choice Model when $\pi_0 = 1.5$ and the Standard Normal Density (Black Line).

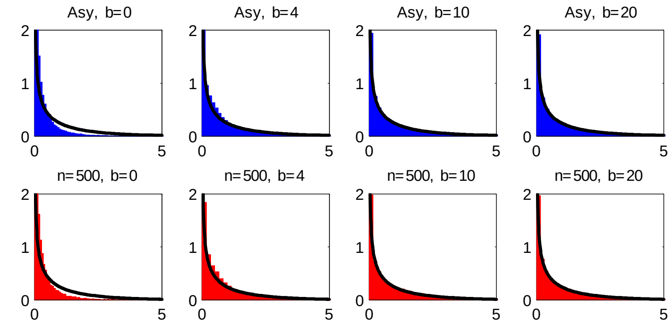

Figure S-12. Asymptotic and Finite-Sample (n=500) Densities of the QLR Statistic for π in the Binary Choice Model when $\pi_0 = 1.5$ and the χ^2_1 Density (Black Line).

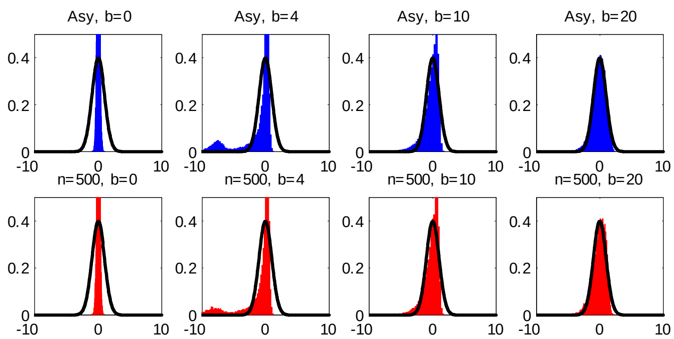

Figure S-13. Asymptotic and Finite-Sample (n=500) Densities of the t Statistic for β in the Binary Choice Model when $\pi_0 = 1.5$ and the Standard Normal Density (Black Line).

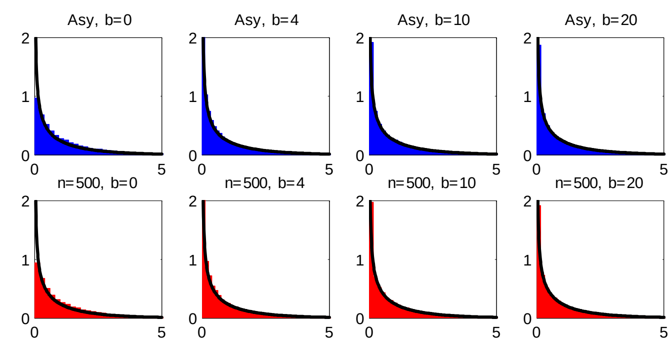

Figure S-14. Asymptotic and Finite-Sample (n=500) Densities of the QLR Statistic for β in the Binary Choice Model when $\pi_0 = 1.5$ and the χ^2_1 Density (Black Line).

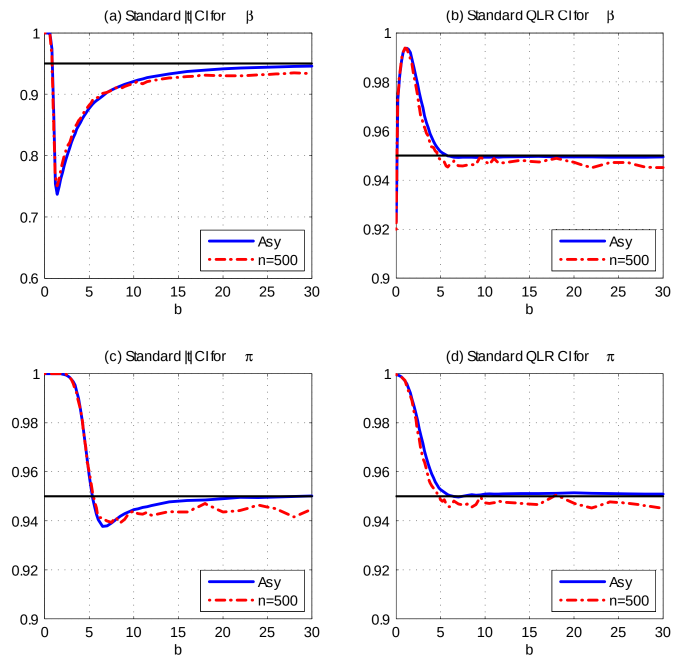

Figure S-15. Coverage Probabilities of Standard $|t|$ and QLR CI's for β and π in the Binary Choice Model when $\pi_0 = 2.0$.

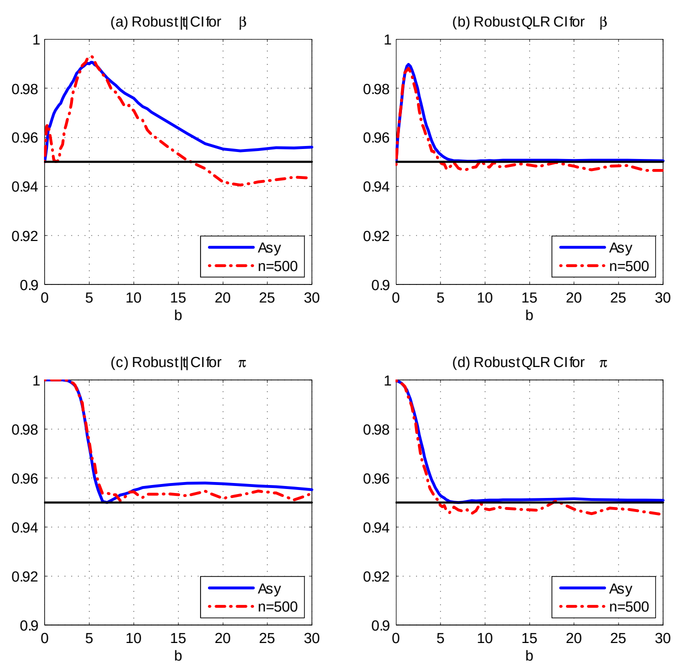

Figure S-16. Coverage Probabilities of Robust $|t|$ and QLR CI's for β and π in the Binary Choice Model when $\pi_0 = 2.0$, $\kappa = 1.5$, $D = 1$, and $s(x) = \exp(-x/2)$.

13. Supplemental Appendix D: Nonlinear Binary Choice Model, Verification of Assumptions

We start by deriving the formulae for the key quantities specified in (3.33). Next, we verify Assumptions S1-S4. Then, we verify Assumptions B1 and B2. Finally, we verify the remaining Assumptions C6, V1, and V2. (Note that Assumption C7 is verified in Section 3.5.)

13.1. Derivation of Key Quantities

Here we calculate the key quantities $\Omega(\pi_1, \pi_2; \gamma_0)$, $H(\pi; \gamma_0)$, $J(\gamma_0)$, and $V(\gamma_0)$ that are specified in (3.33).

By (2.4),

$$\begin{aligned} E_{\gamma_0}(Y_i - L_i(\theta_0)|X_i, Z_i) &= 0 \text{ a.s. and} \\ E_{\gamma_0}((Y_i - L_i(\theta_0))^2|X_i, Z_i) &= L_i(\theta_0)(1 - L_i(\theta_0)) \text{ a.s.} \end{aligned} \tag{13.1}$$

For γ_0 with $\beta_0 = 0$, we have $g_i(\psi_0, \pi) = g_i(\theta_0)$, $L_i(\psi_0, \pi) = L_i(\theta_0)$, $L_i'(\psi_0, \pi) = L_i'(\theta_0)$, and $w_{j,i}(\psi_0, \pi) = w_{j,i}(\theta_0)$ for $j = 1, 2$, $\forall \pi \in \Pi$. In consequence,

$$\begin{aligned} \Omega(\pi_1, \pi_2; \gamma_0) &= S_\psi V^\dagger((\psi_0, \pi_1), (\psi_0, \pi_2); \gamma_0) S_\psi' \\ &= E_{\gamma_0} \frac{L_i'^2(\theta_0)}{L_i(\theta_0)(1 - L_i(\theta_0))} d_{\psi,i}(\pi_1) d_{\psi,i}(\pi_2)', \end{aligned} \tag{13.2}$$

where $S_\psi = [I_{d_\psi} : 0_{d_\psi \times d_\pi}]$, the first equality holds by Lemma 9.1, and the second equality holds by independence across i of $\{W_i : i \leq n\}$ and (13.1).

Now, we have

$$\begin{aligned} \rho_{\psi\psi}(W_i, \psi_0, \pi) &= [w_{1,i}^2(\theta_0)(Y_i - L_i(\theta_0))^2 + w_{2,i}(\theta_0)(Y_i - L_i(\theta_0))] d_{\psi,i}(\pi) d_{\psi,i}(\pi)' \text{ and} \\ H(\pi; \gamma_0) &= E_{\gamma_0} \rho_{\psi\psi}(W_i, \psi_0, \pi) = E_{\gamma_0} \frac{L_i'^2(\theta_0)}{L_i(\theta_0)(1 - L_i(\theta_0))} d_{\psi,i}(\pi) d_{\psi,i}(\pi)', \end{aligned} \tag{13.3}$$

where the first equality uses (3.22), the second equality holds by Lemma 9.1, and the

asymptotic CP is 0.95 for all b values, which serves as a good benchmark. The finite-sample CP's in Table S-1 sometimes differ noticeably from 0.95 due to the small scale of the simulation, i.e., only 1000 simulations repetitions are employed to compute the CP's, as described in footnote 28.

third equality uses (13.1).

In addition, we have

$$\begin{aligned} V(\gamma_0) &= V^{\dagger}(\theta_0, \theta_0; \gamma_0) = Var_{\gamma_0}(\rho_\theta^{\dagger}(W_i, \theta_0)) \\ &= E_{\gamma_0} w_{1,i}^2(\theta_0)(Y_i - L_i(\theta_0))^2 d_i(\pi_0) d_i(\pi_0)' \\ &= E_{\gamma_0} \frac{L_i'^2(\theta_0)}{L_i(\theta_0)(1 - L_i(\theta_0))} d_i(\pi_0) d_i(\pi_0)', \end{aligned} \tag{13.4}$$

where the first equality holds by (3.20) and the second equality holds by independence across i of $\{W_i : i \leq n\}$ and (13.1).

Next, we have

$$\begin{aligned} J(\gamma_0) &= E_{\gamma_0} \rho_{\theta\theta}^{\dagger}(W_i, \theta_0) \\ &= E_{\gamma_0}[w_{1,i}^2(\theta_0)(Y_i - L_i(\theta_0))^2 + w_{2,i}(\theta_0)(Y_i - L_i(\theta_0))] d_i(\pi) d_i(\pi)' \\ &= E_{\gamma_0} \frac{L_i'^2(\theta_0)}{L_i(\theta_0)(1 - L_i(\theta_0))} d_i(\pi) d_i(\pi)', \end{aligned} \tag{13.5}$$

where the first equality holds by Lemma 9.1, the second equality holds using (3.23), and the third equality holds by (13.1).

The matrix $K(\pi; \gamma_0)$ is derived in Section 13.6 below.

13.2. Verification of Assumptions S1 and S2

Given that $\{W_i : i \geq 1\}$ are i.i.d. under γ_0 $\forall \gamma_0 \in \Gamma$, Assumption S1 holds with $q = 2 + \delta$ for $\delta > 0$.

Assumption S2(i) holds with

$$\rho(W_i, \theta) = -[Y_i \log L_i(\theta) + (1 - Y_i) \log(1 - L_i(\theta))]. \tag{13.6}$$

When $\beta = 0$, $L_i(\theta) = L(\beta h(X_i, \pi) + Z_i'\zeta)$ does not depend on π and, hence, $\rho(W_i, \theta)$ does not depend on π. This verifies Assumption S2(ii).

To verify Assumptions S2(iii) and S2(iv), we have

$$E_{\gamma_0}(\rho(W_i, \theta) | X_i, Z_i) = -[L_i(\theta_0) \log L_i(\theta) + (1 - L_i(\theta_0)) \log(1 - L_i(\theta))] \tag{13.7}$$

because $E_{\gamma_0}(Y_i | X_i, Z_i) = L_i(\theta_0)$ by (2.4). Now we view $E_{\gamma_0}(\rho(W_i, \theta) | X_i, Z_i)$ as a function of $L_i(\theta)$. The first- and second-order derivatives of $E_{\gamma_0}(\rho(W_i,\theta)|X_i,Z_i)$ wrt $L_i(\theta)$ are

$$\begin{aligned}\frac{\partial}{\partial L_i(\theta)}E_{\gamma_0}(\rho(W_i,\theta)|X_i,Z_i) &= \frac{L_i(\theta)-L_i(\theta_0)}{L_i(\theta)(1-L_i(\theta))} \text{ and} \\ \frac{\partial^2}{\partial L_i^2(\theta)}E_{\gamma_0}(\rho(W_i,\theta)|X_i,Z_i) &= \frac{L_i(\theta_0)+L_i^2(\theta)-2L_i(\theta)L_i(\theta_0)}{L_i^2(\theta)(1-L_i(\theta))^2}, \end{aligned} \tag{13.8}$$

see (13.45) below. The second-order derivative is positive for all $\theta \in \Theta$ because its numerator is greater than $(L_i(\theta_0)-L_i(\theta))^2 \geq 0$. When $L_i(\theta)=L_i(\theta_0)$, the first-order derivative is 0. Hence, $E_{\gamma_0}(\rho(W_i,\theta)|X_i,Z_i)$, viewed as a function of $L_i(\theta)$, has a unique global minima at $L_i(\theta_0)$. Because $L'(u)>0$, $E_{\gamma_0}\rho(W_i,\theta)$ is minimized at θ if and only if $P_{\gamma_0}(g_i(\theta)=g_i(\theta_0))=1$.

When $\beta_0=0$, $g_i(\theta)-g_i(\theta_0)=\beta h(X_i,\pi)+(\zeta-\zeta_0)'Z_i$. Because $P_{\gamma_0}(a'(h(X_i,\pi),Z_i)=0)<1$ for all $a\in R^{d_\zeta+1}$ with $a\neq 0$ (by the definition of Φ^* in (3.32)), $P_{\gamma_0}(g_i(\theta)-g_i(\theta_0)=0)=1$ if and only if $\beta=0$ and $\zeta=\zeta_0$. This implies Assumption S2(iii).

When $\beta_0\neq 0$, $g_i(\theta)-g_i(\theta_0)=\beta h(X_i,\pi)-\beta_0 h(X_i,\pi_0)+(\zeta-\zeta_0)'Z_i$. Because $P_{\gamma_0}(a'(h(X_i,\pi),h(X_i,\pi_0),Z_i)=0)<1$ for all $a\in R^{d_\zeta+2}$ with $a\neq 0$ and $\pi\neq\pi_0$, $P_{\gamma_0}(g_i(\theta)-g_i(\theta_0)=0)<1$ when $\pi\neq\pi_0$. When $\pi=\pi_0$, $g_i(\theta)-g_i(\theta_0)=(\beta-\beta_0)h(X_i,\pi v)+(\zeta-\zeta_0)'Z_i$. Because $P_{\gamma_0}(a'(h(X_i,\pi),Z_i)=0)<1$ for all $a\in R^{d_\zeta+1}$ with $a\neq 0$, $P_{\gamma_0}(g_i(\theta)-g_i(\theta_0)=0)=1$ if and only if $\zeta=\zeta_0$, $\beta=\beta_0$, and $\pi=\pi_0$. This verifies Assumption S2(iv).

Assumption S2(v) holds because $\Psi(\pi)$ does not depend on π and Ψ, Π, and Θ are all compact. Assumption S2(vi) holds automatically because $\Psi(\pi)$ does not depend on π.

13.3. Verification of Assumption S3(ii)

Assumption S3(i) is verified in the text of the paper. Here we verify Assumption S3(ii). We use the following generic results in the calculations below. Let $A=aa'$, where $a=(a_1',...,a_p')'\in R^{d_a}$ and $a_1,...,a_p$ are vectors (possibly of different dimensions). Then,

$$||A|| = \left(\sum_{j=1}^{p}\sum_{k=1}^{p}||a_j a_k'||^2\right)^{1/2} = \sum_{j=1}^{p}||a_j||^2, \tag{13.9}$$

where the first equality holds by the definition of $||A||$ and the second equality holds because $||ab'||=||a||\cdot||b||$ for vectors a and b. Similarly, let $A^*=a^*a^{*\prime}$, where $a_1^*,...,a_p^*$

are sub-vectors of a^* that are conformable with $a_1, ..., a_p$. Then,

$$\begin{aligned} ||A - A^*|| &= ||aa' - a^*a^{*\prime}|| \leq ||a(a-a^*)'|| + ||(a-a^*)a^{*\prime}|| \\ &= (||a|| + ||a^*||)||a - a^*|| \leq \sum_{j=1}^{p}(||a_j|| + ||a_j^*||)\sum_{k=1}^{p}||a_k - a_k^*||, \end{aligned} \tag{13.10}$$

where the first inequality holds by triangle inequality, the second equality holds because $||ab'|| = ||a|| \cdot ||b||$, and the last inequality holds because $(x^2+y^2)^{1/2} \leq x + y$ for non-negative scalars x and y.

Define $v_{1,i}(\theta) = w_{1,i}(\theta)(Y_i - L_i(\theta))$, $v_{2,i}(\theta) = w_{2,i}(\theta)(Y_i - L_i(\theta))$, and $\overline{\beta} = \max\{b_1, b_2\}$. Below, let $\theta_1, \theta_2 \in \Theta$ with $||\theta_1 - \theta_2|| \leq \delta$ for $\delta > 0$.

By the triangle inequality, we have

$$\begin{aligned} &||\rho_{\psi\psi}(W_i, \theta_1) - \rho_{\psi\psi}(W_i, \theta_2)|| \\ \leq\ & \left(||v_{1,i}^2(\theta_1) - v_{1,i}^2(\theta_2)|| + ||v_{2,i}(\theta_1) - v_{2,i}(\theta_2)||\right) \cdot ||d_{\psi,i}(\pi_1)d_{\psi,i}(\pi_1)'|| \\ &+ \left(||v_{1,i}^2(\theta_2)|| + ||v_{2,i}(\theta_2)||\right) \cdot ||d_{\psi,i}(\pi_1)d_{\psi,i}(\pi_1)' - d_{\psi,i}(\pi_2)d_{\psi,i}(\pi_2)'||. \end{aligned} \tag{13.11}$$

Note that

$$\begin{aligned} ||v_{1,i}^2(\theta_1) - v_{1,i}^2(\theta_2)|| &= ||v_{1,i}(\theta_1) - v_{1,i}(\theta_2)|| \cdot ||v_{1,i}(\theta_1) + v_{1,i}(\theta_2)||, \text{ where} \\ ||v_{1,i}(\theta_1) - v_{1,i}(\theta_2)|| &\leq ||w_{1,i}(\theta_1) - w_{1,i}(\theta_2)|| \cdot ||Y_i - L_i(\theta_1)|| + ||w_{1,i}(\theta_2)|| \cdot ||L_i(\theta_1) - L_i(\theta_2)|| \\ &\leq \left(M_1(W_i) + \overline{w}_{1,i}\overline{L}_i'(\overline{h}_i + ||Z_i|| + \overline{\beta} \cdot \overline{h}_{\pi,i})\right)\delta, \text{ and} \\ ||v_{1,i}(\theta_1) + v_{1,i}(\theta_2)|| &\leq 2\overline{w}_{1,i}, \end{aligned} \tag{13.12}$$

where the first inequality follows from the triangle inequality, the second inequality holds by (i) $||w_{1,i}(\theta_1) - w_{1,i}(\theta_2)|| \leq M_1(W_i)\delta$, (ii) $||Y_i - L_i(\theta)|| \leq 1$, and (iii) $||L_i(\theta_1) - L_i(\theta_2)|| \leq \overline{L}_i'(\overline{h}_i + ||Z_i|| + \overline{\beta} \cdot \overline{h}_{\pi,i})\delta$ by a mean-value expansion of $L_i(\theta) = L(g_i(\theta))$ wrt θ, and the third inequality follows from the triangle inequality and $||Y_i - L_i(\theta)|| \leq 1$. Similarly,

$$\begin{aligned} &||v_{2,i}(\theta_1) - v_{2,i}(\theta_2)|| \leq \left(M_2(W_i) + \overline{w}_{2,i}\overline{L}_i'(\overline{h}_i + ||Z_i|| + \overline{\beta} \cdot \overline{h}_{\pi,i})\right)\delta, \\ &||v_{1,i}^2(\theta_2)|| \leq \overline{w}_{1,i}^2, \text{ and } ||v_{2,i}(\theta_2)|| \leq \overline{w}_{2,i}. \end{aligned} \tag{13.13}$$

Applying the inequality in (13.9) with $a = d_{\psi,i}(\pi_1) = (h(X_i, \pi_1), Z_i')'$, we have

$$||d_{\psi,i}(\pi_1)d_{\psi,i}(\pi_1)'|| \leq \overline{h}_i^2 + ||Z_i||^2. \tag{13.14}$$

Applying the inequality in (13.10) with $a = d_{\psi,i}(\pi_1)$, $a^* = d_{\psi,i}(\pi_2)$, $||a_1 - a_1^*|| \leq \overline{h}_{\pi,i}||\pi_1 - \pi_2||$, and $||a_2 - a_2^*|| = 0$, we have

$$||d_{\psi,i}(\pi_1)d_{\psi,i}(\pi_1)' - d_{\psi,i}(\pi_2)d_{\psi,i}(\pi_2)'|| \leq 2(\overline{h}_i + ||Z_i||)\overline{h}_{\pi,i}||\pi_1 - \pi_2||. \tag{13.15}$$

Equations (13.11)-(13.15) combine to yield

$$||\rho_{\psi\psi}(W_i, \theta_1) - \rho_{\psi\psi}(W_i, \theta_2)|| \leq M_\psi(W_i)\delta, \tag{13.16}$$

where

$$\begin{aligned} M_\psi(W_i) = &\Big[2\overline{w}_{1,i}\left(M_1(W_i) + \overline{w}_{1,i}\overline{L}_i'(\overline{h}_i + ||Z_i|| + \overline{\beta} \cdot \overline{h}_{\pi,i})\right) + M_2(W_i) \\ &+\overline{w}_{2,i}\overline{L}_i'(\overline{h}_i + ||Z_i|| + \overline{\beta} \cdot \overline{h}_{\pi,i})\Big] (\overline{h}_i^2 + ||Z_i||^2) + 2\left(\overline{w}_{1,i}^2 + \overline{w}_{2,i}\right)(\overline{h}_i + ||Z_i||)\overline{h}_{\pi,i}. \end{aligned} \tag{13.17}$$

To show $||\rho_{\theta\theta}^\dagger(\theta_1) - \rho_{\theta\theta}^\dagger(\theta_2)|| \leq M_{\theta\theta}(W_i)\delta$ for some function $M_{\theta\theta}(W_i)$, the calculation is the same as that above with $d_{\psi,i}(\pi)$ replaced by $d_i(\pi)$. The inequalities in (13.14) and (13.15) become

$$\begin{aligned} &||d_i(\pi_1)d_i(\pi_1)'|| \leq \overline{h}_i^2 + ||Z_i||^2 + \overline{h}_{\pi,i}^2 \text{ and} \\ &||d_i(\pi_1)d_i(\pi_1)' - d_i(\pi_2)d_i(\pi_2)'|| \leq 2(\overline{h}_i + ||Z_i|| + \overline{h}_{\pi,i})(\overline{h}_{\pi,i} + \overline{h}_{\pi\pi,i}) \cdot ||\pi_1 - \pi_2||. \end{aligned} \tag{13.18}$$

By the same arguments as those used in (13.11)-(13.17), we have

$$\begin{aligned} M_{\theta\theta}(W_i) = &\Big[2\overline{w}_{1,i}\left(M_1(W_i) + \overline{w}_{1,i}\overline{L}_i'(\overline{h}_i + ||Z_i|| + \overline{\beta} \cdot \overline{h}_{\pi,i})\right) \\ &+M_2(W_i) + \overline{w}_{2,i}\overline{L}_i'(\overline{h}_i + ||Z_i|| + \overline{\beta} \cdot \overline{h}_{\pi,i})\Big] \times (\overline{h}_i^2 + ||Z_i||^2 + \overline{h}_{\pi,i}^2) \\ &+2\left(\overline{w}_{1,i}^2 + \overline{w}_{2,i}\right)(\overline{h}_i + ||Z_i|| + \overline{h}_{\pi,i})(\overline{h}_{\pi,i} + \overline{h}_{\pi\pi,i}). \end{aligned} \tag{13.19}$$

Next, we show $||\rho_\theta^\dagger(\theta_1) - \rho_\theta^\dagger(\theta_2)|| \leq M_\theta(W_i)\delta$ for some function $M_\theta(W_i)$. To this end, note that

$$\begin{aligned}
&||\rho_\theta^\dagger(\theta_1) - \rho_\theta^\dagger(\theta_2)|| \\
&\leq ||v_{1,i}(\theta_1) - v_{1,i}(\theta_2)|| \cdot (\overline{h}_i + ||Z_i|| + \overline{h}_{\pi,i}) + ||v_{1,i}(\theta_2)|| \cdot (\overline{h}_{\pi,i} + \overline{h}_{\pi\pi,i})\delta, \quad (13.20)
\end{aligned}$$

where $||v_{1,i}(\theta_1) - v_{1,i}(\theta_2)||$ satisfies the inequality in (13.12) and $||v_{1,i}(\theta_2)|| \leq \overline{w}_{1,i}$. Hence,

$$M_\theta(W_i) = \left(M_1(W_i) + \overline{w}_{1,i}\overline{L}_i'(\overline{h}_i + ||Z_i|| + \overline{\beta} \cdot \overline{h}_{\pi,i})\right)(\overline{h}_i + ||Z_i|| + \overline{h}_{\pi,i}) + \overline{w}_{1,i}(\overline{h}_{\pi,i} + \overline{h}_{\pi\pi,i}). \quad (13.21)$$

Next, we show $||\varepsilon(W_i, \theta_1) - \varepsilon(W_i, \theta_2)|| \leq M_\varepsilon(W_i)\delta$ for some function $M_\varepsilon(W_i)$. To this end, note that

$$\begin{aligned}
||\varepsilon(W_i, \theta_1) - \varepsilon(W_i, \theta_2)|| &\leq 2||v_{1,i}(\theta_1)h_\pi(X_i, \pi_1) - v_{1,i}(\theta_2)h_\pi(X_i, \pi_2)|| \\
&\quad + ||v_{1,i}(\theta_1)h_{\pi\pi}(X_i, \pi_1) - v_{1,i}(\theta_2)h_{\pi\pi}(X_i, \pi_2)|| \\
&\leq 2||v_{1,i}(\theta_1) - v_{1,i}(\theta_2)||\overline{h}_{\pi,i} + 2||v_{1,i}(\theta_2)||\overline{h}_{\pi\pi,i}\delta \quad (13.22) \\
&\quad + ||v_{1,i}(\theta_1) - v_{1,i}(\theta_2)||\overline{h}_{\pi\pi,i} + ||v_{1,i}(\theta_2)||M_h(W_i)\delta,
\end{aligned}$$

where $||v_{1,i}(\theta_1) - v_{1,i}(\theta_2)||$ satisfies the inequality in (13.12), $||v_{1,i}(\theta_2)|| \leq \overline{w}_{1,i}$, the first inequality follows from a mean-value expansion of $h_\pi(X_i, \pi)$ wrt π and the second inequality follows from $||h_{\pi\pi}(X_i, \pi_1) - h_{\pi\pi}(X_i, \pi_2)|| \leq M_h(X_i) \cdot ||\pi_1 - \pi_2||$. By (13.22), we have

$$M_\varepsilon(W_i) = \left(M_1(W_i) + \overline{w}_{1,i}\overline{L}_i'(\overline{h}_i + ||Z_i|| + \overline{\beta} \cdot \overline{h}_{\pi,i})\right)(2\overline{h}_{\pi,i} + \overline{h}_{\pi\pi,i}) + \overline{w}_{1,i}(2\overline{h}_{\pi\pi,i} + M_h(W_i)). \quad (13.23)$$

Hence, Assumption S3(ii) holds with

$$M_1(W_i) = M_\psi(W_i) + M_{\theta\theta}(W_i) \text{ and } M_2(W_i) = M_\theta(W_i) + M_\varepsilon(W_i). \quad (13.24)$$

13.4. Verification of Assumption S3(iii)

The condition $E_{\gamma_0}M_2(W_i)^q \leq C_1$ for some $C_1 < \infty$ holds if $E_{\gamma_0}M_\theta(W_i)^q \leq C_2$ and $E_{\gamma_0}M_\varepsilon(W_i)^q \leq C_2$ for some $C_2 < \infty$. Because $\overline{L}_i' \leq \overline{w}_{1,i}$, $E_{\gamma_0}M_\theta(W_i)^q \leq C_2$ and $E_{\gamma_0}M_\varepsilon(W_i)^q \leq C_2$ hold provided, for some $C < \infty$, (i) $E_{\gamma_0}M_1^q(W_i)(\overline{h}_i^q + ||Z_i||^q + \overline{h}_{\pi,i}^q + \overline{h}_{\pi\pi,i}^q) \leq C$, (ii) $E_{\gamma_0}\overline{w}_{1,i}^{2q}(\overline{h}_i^q + ||Z_i||^q + \overline{h}_{\pi,i}^q)(\overline{h}_i^q + ||Z_i||^q + \overline{h}_{\pi,i}^q + \overline{h}_{\pi\pi,i}^q) \leq C$, and (iii) $E_{\gamma_0}\overline{w}_{1,i}^q(\overline{h}_{\pi,i}^q + \overline{h}_{\pi\pi,i}^q + M_h(W_i)^q) \leq C$. Condition (i) holds by conditions in (3.32) using

Hölder's inequality to give $E_{\gamma_0} M_1^q(W_i)\overline{h}_i^q \leq (E_{\gamma_0} M_1^{4q/3})^{3/4}(E_{\gamma_0}\overline{h}_i^{4q})^{1/4} \leq C$ and likewise with $||Z_i||$, $\overline{h}_{\pi,i}$, and $\overline{h}_{\pi\pi,i}$ in place of $\overline{h}_i$. Condition (ii) holds by $E_{\gamma_0}\overline{w}_{1,i}^{2q}\overline{h}_i^q||Z_i||^q \leq (E_{\gamma_0}\overline{w}_{1,i}^{4q})^{1/2}(E_{\gamma_0}\overline{h}_i^{4q})^{1/4}(E_{\gamma_0}||Z_i||^{4q})^{1/4} \leq C$ and likewise with $||Z_i||^q$ and $\overline{h}_{\pi,i}^q$ in place of $\overline{h}_i^q$ and $\overline{h}_i^q$, $\overline{h}_{\pi,i}^q$, and $\overline{h}_{\pi\pi,i}^q$ in place of $||Z_i||^q$. Condition (iii) holds by $E_{\gamma_0}\overline{w}_{1,i}^q M_h(W_i)^q \leq (E_{\gamma_0}\overline{w}_{1,i}^{4q})^{1/4}(E_{\gamma_0} M_h(W_i)^{4q/3})^{3/4} \leq C$ and likewise with $\overline{h}_{\pi,i}^q$ and $\overline{h}_{\pi\pi,i}^q$ in place of $M_h(W_i)^q$.

The condition $E_{\gamma_0} M_1(W_i) \leq C_1$ for some $C_1 < \infty$ holds if $E_{\gamma_0} M_\psi(W_i) \leq C_2$ and $E_{\gamma_0} M_{\theta\theta}(W_i) \leq C_2$ for some $C_2 < \infty$. Because $\overline{L}_i' \leq \overline{w}_{1,i}$, $E_{\gamma_0} M_\psi(W_i) \leq C_2$ and $E_{\gamma_0} M_{\theta\theta}(W_i) \leq C_2$ hold provided, for some $C < \infty$, (i) $E_{\gamma_0} M_1(W_i)\overline{w}_{1,i}(\overline{h}_i^2 + ||Z_i||^2 + \overline{h}_{\pi,i}^2) \leq C$, (ii) $E_{\gamma_0}\overline{w}_{1,i}^3(\overline{h}_i + ||Z_i|| + \overline{h}_{\pi,i})(\overline{h}_i^2 + ||Z_i||^2 + \overline{h}_{\pi,i}^2) \leq C$, (iii) $E_{\gamma_0} M_2(W_i)(\overline{h}_i^2 + ||Z_i||^2 + \overline{h}_{\pi,i}^2) \leq C$, (iv) $E_{\gamma_0}\overline{w}_{1,i}\overline{w}_{2,i}(\overline{h}_i + ||Z_i|| + \overline{h}_{\pi,i})(\overline{h}_i^2 + ||Z_i||^2 + \overline{h}_{\pi,i}^2) \leq C$, (v) $E_{\gamma_0}\left(\overline{w}_{1,i}^2 + \overline{w}_{2,i}\right)(\overline{h}_i + ||Z_i|| + \overline{h}_{\pi,i})(\overline{h}_{\pi,i} + \overline{h}_{\pi\pi,i}) \leq C$. Condition (i) holds by conditions in (3.32) using the Cauchy-Schwarz inequality and $q > 2$ to give $E_{\gamma_0} M_1(W_i)\overline{w}_{1,i}\overline{h}_i^2 \leq (E_{\gamma_0} M_1(W_i)^2)^{1/2}(E_{\gamma_0}\overline{w}_{1,i}^4)^{1/4}(E_{\gamma_0}\overline{h}_i^8)^{1/4} \leq C$ and likewise with $||Z_i||^2$ and $\overline{h}_{\pi,i}^2$ in place of $\overline{h}_i^2$. Condition (ii) holds by $E_{\gamma_0}\overline{w}_{1,i}^3\overline{h}_i\overline{h}_{\pi,i}^2 \leq (E_{\gamma_0}\overline{w}_{1,i}^6)^{1/2}(E_{\gamma_0}\overline{h}_i^4)^{1/4}(E_{\gamma_0}\overline{h}_{\pi,i}^8)^{1/4} \leq C$ and likewise with $||Z_i||$ and $\overline{h}_{\pi,i}$ in place of $\overline{h}_i$ and with $\overline{h}_i^2$ and $||Z_i||^2$ in place of $\overline{h}_{\pi,i}^2$. Condition (iii) holds by $E_{\gamma_0} M_2(W_i)\overline{h}_i^2 \leq (E_{\gamma_0} M_2(W_i)^{4/3})^{3/4}(E_{\gamma_0}\overline{h}_i^8)^{1/4} \leq C$ and likewise with $||Z_i||^2$ and $\overline{h}_{\pi,i}^2$ in place of $\overline{h}_i^2$. Condition (iv) holds by $E_{\gamma_0}\overline{w}_{1,i}\overline{w}_{2,i}\overline{h}_i\overline{h}_{\pi,i}^2 \leq (E_{\gamma_0}\overline{w}_{1,i}^8)^{1/8}(E_{\gamma_0}\overline{w}_{2,i}^2)^{1/2}(E_{\gamma_0}\overline{h}_i^8)^{1/8}(E_{\gamma_0}\overline{h}_{\pi,i}^8)^{1/4} \leq C$ and likewise with $||Z_i||$ and $\overline{h}_{\pi,i}$ in place of $\overline{h}_i$ and with $\overline{h}_i^2$ and $||Z_i||^2$ in place of $\overline{h}_{\pi,i}^2$. Condition (v) holds by $E_{\gamma_0}\overline{w}_{2,i}\overline{h}_i\overline{h}_{\pi,i} \leq (E_{\gamma_0}\overline{w}_{2,i}^2)^{1/2}(E_{\gamma_0}\overline{h}_i^4)^{1/4}(E_{\gamma_0}\overline{h}_{\pi,i}^4)^{1/4} \leq C$ and likewise with $\overline{w}_{1,i}^2$ in place of $\overline{w}_{2,i}$, $||Z_i||$ and $\overline{h}_{\pi,i}$ in place of $\overline{h}_i$, and $\overline{h}_{\pi\pi,i}$ in place of $\overline{h}_{\pi,i}$.

By (13.6),

$$E_{\gamma_0}\sup_{\theta\in\Theta}|\rho(W_i,\theta)|^{1+\delta} \leq E_{\gamma_0}(\sup_{\theta\in\Theta}|\log L_i(\theta)| + \sup_{\theta\in\Theta}|\log(1 - L_i(\theta))|)^{1+\delta} \leq C, \qquad (13.25)$$

for some $C < \infty$, where the first inequality holds because Y_i is 0 or 1 and the second inequality holds by conditions in (3.32).

By (3.22),

$$E_{\gamma_0}\sup_{\theta\in\Theta}||\rho_{\psi\psi}(W_i,\theta)||^{1+\delta} \leq E_{\gamma_0}(\overline{w}_{1,i}^2 + \overline{w}_{2,i})^{1+\delta}\sup_{\pi\in\Pi}||d_{\psi,i}(\pi)d_{\psi,i}(\pi)'||^{1+\delta} \leq C \qquad (13.26)$$

for some $C < \infty$, where the first inequality holds by $|Y_i - L_i(\theta)| \leq 1$ and the triangle inequality and the second inequality holds (13.14) and conditions in (3.32). Similarly, we can show $E_{\gamma_0}\sup_{\theta\in\Theta}||\rho_{\theta\theta}^\dagger(W_i,\theta)||^{1+\delta} \leq C$ with $d_{\psi,i}(\pi)$ in (13.26) replaced by $d_i(\pi)$

and (13.14) replaced by (13.18).

By (3.23),

$$E_{\gamma_0} \sup_{\theta \in \Theta} ||\rho_\theta^\dagger(W_i, \theta)||^q \leq E_{\gamma_0} (\overline{w}_{1,i} \sup_{\pi \in \Pi} ||d_i(\pi)||)^q$$
$$\leq E_{\gamma_0} \overline{w}_{1,i}^q (\overline{h}_i + ||Z_i|| + \overline{h}_{\pi,i})^q \leq C \tag{13.27}$$

for some $C < \infty$, where the first inequality holds because $|Y_i - L_i(\theta)| \leq 1$, the second inequality holds because $||d_i(\pi)|| \leq ||h(X_i, \pi)|| + ||Z_i|| + ||h_\pi(X_i, \pi)||$, and the third inequality holds by conditions in (3.32).

By (3.23),

$$E_{\gamma_0} \sup_{\theta \in \Theta} ||\varepsilon(W_i, \theta)||^q \leq E_{\gamma_0} \overline{w}_{1,i}^q (2\overline{h}_{\pi,i} + \overline{h}_{\pi\pi,i})^q \leq C \tag{13.28}$$

for some $C < \infty$, where the the first inequality follows from $|Y_i - L_i(\theta)| \leq 1$ and the second inequality holds by conditions in (3.32).

This completes the verification of Assumption S3(iii).

13.5. Verification of Assumptions S3(iv) and S3(v)

To verify Assumption S3(iv), we apply the LIE and obtain

$$E_{\gamma_0} \rho_{\psi\psi}(W_i, \theta) = E_{\gamma_0} [w_{1,i}^2(\theta) e_{1,i}(\theta) + w_{2,i}(\theta) e_{2,i}(\theta)] d_{\psi,i}(\pi) d_{\psi,i}(\pi)', \text{ where} \tag{13.29}$$
$$e_{1,i}(\theta) = E_{\gamma_0}((Y_i - L_i(\theta))^2 | X_i, Z_i) \text{ and } e_{2,i}(\theta) = E_{\gamma_0}(Y_i - L_i(\theta) | X_i, Z_i).$$

When $\beta_0 = 0$, $g_i(\psi_0, \pi) = Z_i'\zeta_0$ and $L_i(\psi_0, \pi) = L(g_i(\psi_0, \pi)) = L(Z_i'\zeta_0)$, $\forall \pi \in \Pi$. By (2.4),

$$e_{1,i}(\psi_0, \pi) = L(Z_i'\zeta_0)(1 - L(Z_i'\zeta_0)) \text{ and } e_{2,i}(\psi_0, \pi) = 0. \tag{13.30}$$

Hence, when $\beta_0 = 0$,

$$E_{\gamma_0} \rho_{\psi\psi}(W_i, \psi_0, \pi) = E_{\gamma_0} \frac{L'^2(Z_i'\zeta_0)}{L(Z_i'\zeta_0)(1 - L(Z_i'\zeta_0))} d_{\psi,i}(\pi) d_{\psi,i}(\pi)'. \tag{13.31}$$

The quantity $E_{\gamma_0} \rho_{\psi\psi}(W_i, \psi_0, \pi)$ is continuous in π on Π by the DCT using (13.16), (13.17), and the discussion following (13.24). Hence, $\lambda_{\min}(E_{\gamma_0} \rho_{\psi\psi}(W_i, \psi_0, \pi))$ also is continuous on the compact set Π and attains its minimum at some point $\pi_{\min} \in \Pi$. Its minimum is zero only if the positive semi-definite matrix $E_{\gamma_0} \rho_{\psi\psi}(W_i, \psi_0, \pi_{\min})$ is not positive definite. The latter is ruled out by the fact that $L'^2(Z_i'\zeta_0)/(L(Z_i'\zeta_0)(1-L(Z_i'\zeta_0)))$ is positive a.s. and the condition in (3.32) that $P_{\gamma_0}(a'(h(X_i,\pi),Z_i)=0)<1$, $\forall \pi \in \Pi$, $\forall a \in R^{d_\zeta+1}$ with $a \neq 0$. Thus, $\inf_{\pi\in\Pi}\lambda_{\min}(E_{\gamma_0}\rho_{\psi\psi}(W_i,\psi_0,\pi))>0$ when $\beta_0=0$ and the first part of Assumption S3(iv) holds.

As in (13.29)-(13.31), we can show

$$E_{\gamma_0}\rho^\dagger_{\theta\theta}(W_i,\theta_0)=E_{\gamma_0}\frac{L_i'^2(\theta_0)}{L_i(\theta_0)(1-L_i(\theta_0))}d_i(\pi_0)d_i(\pi_0)' \tag{13.32}$$

by replacing (ψ_0,π) with θ_0 and $d_{\psi,i}(\pi)$ with $d_i(\pi_0)$ in the arguments above. Because $L_i'(\theta_0)>0$ and $0<L_i(\theta_0)<1$, $E_{\gamma_0}\rho^\dagger_{\theta\theta}(W_i,\theta_0)$ is positive definite because $E_{\gamma_0}d_i(\pi_0)d_i(\pi_0)'$ is positive definite as specified in (3.32). Hence, the second part of Assumption S3(iv) holds.

By (13.4) and (13.32), $V^\dagger(\theta_0,\theta_0;\gamma_0)=E_{\gamma_0}\rho^\dagger_{\theta\theta}(W_i,\theta_0)$. Hence, $V^\dagger(\theta_0,\theta_0;\gamma_0)$ is positive definite.

13.6. Verification of Assumption S4

Because $m(W_i,\theta)=\rho_\psi(W_i,\theta)$ by Lemma 9.1,

$$\begin{aligned}E_{\gamma_0}m(W_i,\theta) &= E_{\gamma_0}\rho_\psi(W_i,\theta)=E_{\gamma_0}w_{1,i}(\theta)(Y_i-L_i(\theta))d_{\psi,i}(\pi)\\ &= E_{\gamma_0}w_{1,i}(\theta)(L_i(\theta_0)-L_i(\theta))d_{\psi,i}(\pi),\end{aligned} \tag{13.33}$$

where $\gamma_0=(\beta_0,\zeta_0,\pi_0,\phi_0)$, the second equality holds by (3.22), and the third equality holds by iterated expectations and (2.4). In (13.33), $E_{\gamma_0}m(W_i,\theta)$ depends on β_0 only through $L_i(\theta_0)$. Hence,

$$\begin{aligned}K(\theta;\gamma_0) &= (\partial/\partial\beta_0)E_{\gamma_0}w_{1,i}(\theta)(L_i(\theta_0)-L_i(\theta))d_{\psi,i}(\pi)\\ &= E_{\gamma_0}w_{1,i}(\theta)L_i'(\theta_0)h(X_i,\pi_0)d_{\psi,i}(\pi),\end{aligned} \tag{13.34}$$

where the first equality holds because the observations are identically distributed and the second equality holds by an exchange of E and ∂ because $E_{\gamma_0}\sup_{\theta\in\Theta,\theta_0\in\Theta_0}||w_{1,i}(\theta)L_i'(\theta_0)$ $h(X_i,\pi_0)d_{\psi,i}(\pi)||<\infty$ by conditions in (3.32) and $(\partial/\partial\beta_0)g_i(\theta_0)=h(X_i,\pi_0)$. Hence, Assumption S4(i) holds.

Now we show that Assumptions S4(ii) holds with

$$K(\pi;\gamma_0) = K(\psi_0,\pi;\gamma_0) = E_{\gamma_0} w_{1,i}(\psi_0,\pi) L_i'(\theta_0) h(X_i,\pi_0) d_{\psi,i}(\pi). \tag{13.35}$$

Define $a_i(\theta,\theta_0) = w_{1,i}(\theta) L_i'(\theta_0) h(X_i,\pi_0) d_{\psi,i}(\pi)$. It suffices to show that $E_{\gamma_n} a_i(\theta,\theta^*) \to E_{\gamma_0} a_i(\theta,\theta^*)$ uniformly over $(\theta,\theta^*) \in \Theta \times \Theta^*$ as $\gamma_n \to \gamma_0$ and $E_{\gamma_0} a_i(\theta,\theta^*)$ is continuous in (θ,θ^*). The continuity holds by the continuity of $a_i(\theta,\theta^*)$ in (θ,θ^*), $E_{\gamma_0} \sup_{(\theta,\theta^*)\in\Theta\times\Theta^*} ||a_i(\theta,\theta^*)|| < \infty$ by conditions in (3.32), and the dominated convergence theorem. By Lemma 9.3, the uniform convergence follows from the pointwise convergence and the equicontinuity of $E_{\gamma_0} a_i(\theta,\theta^*)$ in (θ,θ^*) over $\gamma_0 \in \Gamma$. The pointwise convergence $E_{\gamma_n} a_i(\theta,\theta^*) \to E_{\gamma_0} a_i(\theta,\theta^*)$ holds because (i) the expectations $E_{\gamma_n} a_i(\theta,\theta^*)$ and $E_{\gamma_0} a_i(\theta,\theta^*)$ depend on ϕ_n and ϕ_0, respectively, but not on θ_n and θ_0, (ii) $\phi_n \to \phi_0$ implies convergence in distribution by the metric on Φ^*, and (iii) the $L^{1+\delta}$ boundedness of $a_i(\theta,\theta^*)$, i.e., $E_{\gamma_0} ||a_i(\theta,\theta^*)||^{1+\delta} \le C < \infty$ for any $\gamma_0 \in \Gamma$. Equicontinuity holds because for any (θ_1,θ_1^*) and (θ_2,θ_2^*) with $||(\theta_1,\theta_1^*) - (\theta_2,\theta_2^*)|| \le \delta$,

$$\begin{aligned}
& E_{\gamma_0} ||a_i(\theta_1,\theta_1^*) - a_i(\theta_2,\theta_2^*)|| \\
\le\ & E_{\gamma_0} ||w_{1,i}(\theta_1) - w_{1,i}(\theta_2)|| \cdot ||L_i'(\theta_1^*) h(X_i,\pi_1^*) d_{\psi,i}(\pi_1)|| \\
& + E_{\gamma_0} ||w_{1,i}(\theta_2)|| \cdot ||L_i'(\theta_1^*) h(X_i,\pi_1^*) d_{\psi,i}(\pi_1) - L_i'(\theta_2^*) h(X_i,\pi_2^*) d_{\psi,i}(\pi_2)|| \\
\le\ & E_{\gamma_0} M_1(W_i) \overline{L}_i' \overline{h}_i \sup_{\pi\in\Pi} ||d_{\psi,i}(\pi)|| \delta \qquad (13.36) \\
& + E_{\gamma_0} \overline{w}_{1,i} \left[\left(\overline{L}_i'' \overline{h}_i + \overline{L}_i \overline{h}_{\pi,i} \right) \sup_{\pi\in\Pi} ||d_{\psi,i}(\pi)|| + \overline{L}_i' \overline{h}_i || \sup_{\pi\in\Pi} (\partial/\partial\pi') d_{\psi,i}(\pi)|| \right] \delta \le C\delta
\end{aligned}$$

for some $C < \infty$ for all $\gamma_0 \in \Gamma$, where the first inequality holds by the triangle inequality, the second inequality follows from $||w_{1,i}(\theta_1) - w_{1,i}(\theta_2)|| \le M_1(W_i)\delta$ and a mean-value expansion of $L_i'(\theta_1^*) h(X_i,\pi_1^*) d_{\psi,i}(\pi_1)$ wrt (θ_1,θ_1^*) around (θ_2,θ_2^*), and the third inequality holds by the Cauchy-Schwarz inequality and conditions in (3.32). This completes the verification of Assumption S4.

13.7. Verification of Assumptions B1 and B2

Given the definitions in Section 3.2, Assumptions B1(i) and B1(iii) follow immediately. Assumption B1(ii) holds by taking $\delta < \min\{b_1^*, b_2^*\}$ and $\mathcal{Z}^0 = int(\mathcal{Z})$.

Given the definitions in Sections 3.2, the true parameter space Γ is of the form in (2.6). Thus, Assumption B2(i) holds immediately. Assumption B2(ii) follows from the form of $\mathcal{B}^*$ given in (2.9). Assumption B2(iii) follows from the form of $\mathcal{B}^*$ and the fact that Θ^* is a product space and $\Phi^*(\theta_0)$ does not depend on β_0. Hence, the true parameter space Γ satisfies Assumption B2.

13.8. Verification of Assumption C6

Assumption C6 holds by Lemma 3.2 under Assumptions S1-S3 and C6†. We now verify Assumption C6. Assumption C6†(i) holds because β is a scalar. To verify Assumption C6†(ii), we have

$$\rho_\beta(W_i,\theta) = w_{1,i}(\theta)(Y_i - L_i(\theta))h(X_i,\pi) \text{ and } \rho_\zeta(W_i,\theta) = w_{1,i}(\theta)(Y_i - L_i(\theta))Z_i. \quad (13.37)$$

When $\beta_0 = 0$,

$$\begin{aligned} \rho^*_\psi(W_i,\psi_0,\pi_1,\pi_2) &= w_{1,i}(\psi_0)(Y_i - L_i(\psi_0))h_{Z,i}(\pi_1,\pi_2), \text{ where} \\ w_{1,i}(\psi_0) &= \frac{L'(Z_i'\zeta_0)}{L(Z_i'\zeta_0)(1-L(Z_i'\zeta_0))}, \; L_i(\psi_0) = L(Z_i'\zeta_0), \text{ and} \\ h_{Z,i}(\pi_1,\pi_2) &= (h(X_i,\pi_1), h(X_i,\pi_2), Z_i')'. \end{aligned} \quad (13.38)$$

The covariance matrix in Assumption C6†(ii) is

$$\begin{aligned} \Omega_G(\pi_1,\pi_2;\gamma_0) &= Cov_{\gamma_0}(\rho^*_\psi(W_i,\psi_0,\pi_1,\pi_2), \rho^*_\psi(W_i,\psi_0,\pi_1,\pi_2)) \\ &= E_{\gamma_0} w^2_{1,i}(\psi_0)(Y_i - L_i(\psi_0))^2 h_{Z,i}(\pi_1,\pi_2)h_{Z,i}(\pi_1,\pi_2)' \\ &= E_{\gamma_0}\frac{L'^2(Z_i'\zeta_0)}{L(Z_i'\zeta_0)(1-L(Z_i'\zeta_0))} h_{Z,i}(\pi_1,\pi_2)h_{Z,i}(\pi_1,\pi_2)', \end{aligned} \quad (13.39)$$

where the first equality holds because the observations are independent and identically distributed, the second equality follows from $E\rho^*_\psi(W_i,\psi_0,\pi_1,\pi_2) = 0$, which in turn holds by the LIE and (2.4), and the third equality holds by (13.1). Because $L'(Z_i'\zeta_0) > 0$ and $0 < L(Z_i'\zeta_0) < 1$, $\Omega_G(\pi_1,\pi_2;\gamma_0)$ is positive definite because $P(a'h_{Z,i}(\pi_1,\pi_2) = 0) < 1$ for all $a \in R^{d_z+2}$ with $a \neq 0$ by the conditions in (3.32).

13.9. Verification of Assumptions V1 and V2

Here we verify Assumptions V1 (scalar β) and V2, which are stated in Appendix B of the main paper.

For the binary choice model, the matrices $J(\gamma_0)$ $(= V(\gamma_0))$ and $\widehat{J}_n(\theta)$ $(= \widehat{V}_n(\theta))$ are defined in (3.35) and (5.18), respectively. Define

$$J(\theta;\gamma_0) = E_{\gamma_0} \frac{L_i'^2(\theta)}{L_i(\theta)(1 - L_i(\theta))} d_i(\pi) d_i(\pi)'. \tag{13.40}$$

Under $\{\gamma_n\} \in \Gamma(\gamma_0)$, $\sup_{\theta\in\Theta} ||\widehat{J}_n(\theta) - J(\theta;\gamma_0)|| \to_p 0$ and $J(\theta;\gamma_0)$ is continuous in θ on Θ by the uniform law of large numbers in Lemma 9.3, where the smoothness and moment conditions hold by conditions in (3.32). In addition, $J(\theta_0;\gamma_0) = J(\gamma_0)$. This verifies Assumption V1(i) and V1(ii) (for scalar β).

To verify Assumption V1(iii), note that

$$\Sigma(\theta;\gamma_0) = J^{-1}(\theta;\gamma_0) \text{ and } \Sigma(\pi;\gamma_0) = J^{-1}(\psi_0, \pi;\gamma_0). \tag{13.41}$$

Hence, it suffices to show that (i) $\lambda_{\min}(J(\psi_0, \pi;\gamma_0)) > 0$ and (ii) $\lambda_{\max}(J(\psi_0, \pi;\gamma_0)) < \infty$ for all $\pi \in \Pi$. Property (i) holds by essentially the same argument as in the paragraph following (13.31) with $d_i(\pi)$ in place of $d_{\psi,i}(\pi)$ using the condition in (3.32) that $E_{\gamma_0} d_i(\pi) d_i(\pi)'$ is positive definite $\forall \pi \in \Pi$. Positive definiteness of $E_{\gamma_0} d_i(\pi) d_i(\pi)'$ implies the same for $E_{\gamma_0}[L'^2(Z_i'\zeta_0)/(L(Z_i'\zeta_0)(1 - L(Z_i'\zeta_0)))] d_i(\pi) d_i(\pi)'$ because the latter is well-defined and $L'^2(Z_i'\zeta_0)/(L(Z_i'\zeta_0)(1 - L(Z_i'\zeta_0)))$ is positive a.s. Property (ii) holds by the moment conditions in (3.32). This completes the verification of Assumption V1(iii).

Assumptions V1(i) and V1(ii) hold not only under $\{\gamma_n\} \in \Gamma(\gamma_0, 0, b)$, but also under $\{\gamma_n\} \in \Gamma(\gamma_0, \infty, \omega_0)$ in this example. This and $\widehat{\theta}_n \to_p \theta_0$ under $\{\gamma_n\} \in \Gamma(\gamma_0, \infty, \omega_0)$, which holds by Lemma 5.3 of AC1, imply that Assumption V2 holds. Among the assumptions employed in Lemma 5.3 of AC1, Assumptions B1, B2, and C7 are verified directly, Assumptions A, B3, and C1-C5 hold by Lemma 9.1 under Assumptions B1, B2, and S1-S4, and Assumption C6 holds by Lemma 3.2 under Assumptions S1-S3 and C6†.

13.10. Calculation of Partial Derivatives

Here we calculate the partial derivatives of $\rho(W_i,\theta)$ wrt θ. Let L abbreviate $L(g_i(\theta))$. The first-order derivative wrt θ is

$$\begin{aligned}\rho_\theta(W_i,\theta) &= -\left[\frac{Y_i}{L}-\frac{1-Y_i}{1-L}\right]L'\frac{\partial}{\partial\theta}g_i(\theta)\\ &= -\frac{Y_i-L}{L(1-L)}L'\frac{\partial}{\partial\theta}g_i(\theta)=w_{1,i}(\theta)(Y_i-L)B(\beta)d_i(\pi),\ \text{where}\\ w_{1,i}(\theta) &= \frac{-L'}{L(1-L)}. \end{aligned} \tag{13.42}$$

Now we calculate the second-order derivatives. To this end, we have

$$\begin{aligned}&\frac{\partial}{\partial\theta'}\left[\frac{Y_i}{L}-\frac{1-Y_i}{1-L}\right]=\left[\frac{-Y_i}{L^2}+\frac{-(1-Y_i)}{(1-L)^2}\right]L'\frac{\partial}{\partial\theta'}g_i(\theta)\\ &= -\left[\frac{Y_i(1-L)^2+(1-Y_i)L^2}{L^2(1-L)^2}\right]L'\frac{\partial}{\partial\theta'}g_i(\theta)\\ &= -\left[\frac{Y_i-2Y_iL+L^2}{L^2(1-L)^2}\right]L'\frac{\partial}{\partial\theta'}g_i(\theta)\\ &= -\frac{(Y_i-L)^2}{L^2(1-L)^2}L'\frac{\partial}{\partial\theta'}g_i(\theta),\\ &\frac{\partial}{\partial\theta'}L'=L''\frac{\partial}{\partial\theta'}g_i(\theta),\ \text{and}\ \frac{\partial^2}{\partial\theta\partial\theta'}g_i(\theta)=D_i(\theta).\end{aligned} \tag{13.43}$$

Hence,

$$\begin{aligned}\rho_{\theta\theta}(W_i,\theta) &= \left[\frac{(Y_i-L)^2}{L^2(1-L)^2}(L')^2-\frac{Y_i-L}{L(1-L)}L''\right]\left(\frac{\partial}{\partial\theta}g_i(\theta)\frac{\partial}{\partial\theta'}g_i(\theta)\right)\\ &\quad -\frac{Y_i-L}{L(1-L)}L'\frac{\partial^2}{\partial\theta\partial\theta'}g_i(\theta)\\ &= [w_{1,i}^2(Y_i-L)^2+w_{2,i}(Y_i-L)]B(\beta)d_i(\pi)d_i(\pi)'B(\beta)\\ &\quad +w_{1,i}(Y_i-L)D_i(\theta),\ \text{where}\\ w_{1,i}(\theta) &= \frac{-L'}{L(1-L)}\ \text{and}\ w_{2,i}(\theta)=\frac{-L''}{L(1-L)}.\end{aligned} \tag{13.44}$$

Lastly, we calculate the derivatives in (13.8). Let $L=L_i(\theta)$ and $L_0=L_i(\theta_0)$. We have

$$\begin{aligned}
\text{FOC} &= -\frac{L_0}{L} + \frac{1-L_0}{1-L} = \frac{L-L_0}{L(1-L)} \text{ and} \\
\text{SOC} &= \frac{L(1-L)-(L-L_0)(1-2L)}{L^2(1-L)^2} \\
&= \frac{L-L^2-(L-L_0-2L^2+2LL_0)}{L^2(1-L)^2} \\
&= \frac{L_0+L^2-2LL_0}{L^2(1-L)^2} > \frac{(L_0-L)^2}{L^2(1-L)^2} > 0. \qquad (13.45)
\end{aligned}$$

14. Supplemental Appendix E: STAR Example, Verification of Assumptions

14.1. Verification of Assumptions S1 and S2

Assumption S1 holds by Assumption STAR1(ii).

Assumption S2(i) holds with

$$\rho(W_t, \theta) = U_t^2(\theta)/2, \text{ where } U_t(\theta) = Y_t - X_t'\zeta - X_t'\beta \cdot m(Z_t, \pi). \tag{14.46}$$

The residual $U_t(\theta)$ is twice continuously differentiable in θ for both the logistic function and the exponential function. When $\beta = 0$, $U_t(\theta) = Y_t - X_t'\zeta$, which does not depend on π. This verifies Assumption S2(ii).

To verify Assumptions S2(iii) and S2(iv), we have

$$\begin{aligned} E_{\gamma_0}\rho(W_t, \theta) &= E_{\gamma_0}\left[Y_t - X_t'\zeta - X_t'\beta \cdot m(Z_t, \pi)\right]^2 \\ &= E_{\gamma_0}\left(U_t - X_t'(\zeta - \zeta_0) - X_t'\left[\beta m(Z_t, \pi) - \beta_0 m(Z_t, \pi_0)\right]\right)^2 \\ &= E_{\gamma_0}U_t^2 + E_{\gamma_0}[X_t'(\zeta - \zeta_0) + X_t'(\beta m(Z_t, \pi) - \beta_0 m(Z_t, \pi_0))]^2. \end{aligned} \tag{14.47}$$

To verify Assumption S2(iii), we need that when $\beta_0 = 0$,

$$E_{\gamma_0}\rho(W_t, \psi, \pi) - E_{\gamma_0}\rho(W_t, \psi_0, \pi) = E_{\gamma_0}[X_t'(\zeta - \zeta_0) + X_t'\beta m(Z_t, \pi)]^2 > 0 \tag{14.48}$$

$\forall \psi \neq \psi_0$ and $\forall \pi \in \Pi$. The inequality in (14.48) holds unless

$$P_{\gamma_0}((X_t' + X_t' m(Z_t, \pi))a = 0) = 1, \tag{14.49}$$

where $a = ((\zeta - \zeta_0)', \beta')'$. By Assumption STAR2(i), (14.49) does not hold for any $a \neq 0$. Hence, the inequality in (14.48) holds $\forall \psi \neq \psi_0$. This completes the verification of Assumption S2(iii).

To verify Assumption S2(iv), we need that when $\beta_0 \neq 0$,

$$\begin{aligned} & E_{\gamma_0}\rho(W_t, \theta) - E_{\gamma_0}\rho(W_t, \theta_0) \\ = & E_{\gamma_0}[X_t'(\zeta - \zeta_0) + X_t'\beta m(Z_t, \pi) - X_t'\beta_0 m(Z_t, \pi_0)]^2 > 0 \end{aligned} \tag{14.50}$$

$\forall\theta \neq \theta_0$. The inequality in (14.50) holds unless

$$P_{\gamma_0}\left(X_t'(\zeta - \zeta_0) + X_t'\beta m(Z_t, \pi) - X_t'\beta_0 m(Z_t, \pi_0) = 0\right) = 1 \tag{14.51}$$

for some $\theta \neq \theta_0$. Because $\beta_0 \neq 0$, Assumption STAR2(i) implies that (14.51) does not hold for any $\pi \neq \pi_0$. When $\pi = \pi_0$, (14.51) becomes

$$P_{\gamma_0}\left(X_t'(\zeta - \zeta_0) + X_t'(\beta - \beta_0)m(Z_t, \pi_0) = 0\right) = 1. \tag{14.52}$$

Because (14.49) does not hold for any $a \neq 0$ for any $\pi \in \Pi$, (14.52) cannot hold for $(\beta, \zeta) \neq (\beta_0, \zeta_0)$. This completes the verification of Assumption S2(iv).

Assumption S2(v) holds by Assumption STAR5(ii). Assumption S2(vi) holds because Ψ does not depend on π.

14.2. Verification of Assumption S3(i)

Now we verify Assumption S3 (vector β). In the STAR model, Z_t is an element of X_t and the function $\rho(\omega, \theta)$ takes the form in (3.19) with

$$\begin{aligned} a(X_t, \beta) &= X_t'\beta \in R,\ h(X_t, \pi) = m\left(Z_t, \pi\right) \in R, \text{ and} \\ \rho^*(W_t, a(X_t, \beta)h(X_t, \pi), \zeta) &= [Y_t - X_t'\zeta - a(X_t, \beta)h(X_t, \pi)]^2/2. \end{aligned} \tag{14.53}$$

By Lemma 3.1, we verify Assumption S3(i) by showing that Assumption S3* holds.

To verify Assumption S3*(i), we have

$$\rho'(W_t, a(X_t, \beta_0)h(X_t, \pi_0), \zeta_0) = -[Y_t - X_t'\zeta_0 - a(X_t, \beta_0)h(X_t, \pi_0)] = -U_t. \tag{14.54}$$

Note that $\rho'(\cdot)$ and $\rho''(\cdot)$ in Assumption S3* are partial derivatives of $\rho^*(\cdot)$ wrt $a(X_t, \beta)$ $h(X_t, \pi)$. Assumption S3*(i) holds immediately by Assumption STAR1(i).

To verify Assumption S3*(ii), we first derive the terms that appear in it. By (14.53),

$$\begin{aligned} &\rho''(W_t, a(X_t, \beta)h(X_t, \pi), \zeta) = 1, \\ &h(X_t, \pi) = m(Z_t, \pi), h_\pi(X_t, \pi) = m_\pi(Z_t, \pi),\ h_{\pi\pi}(X_t, \pi) = m_{\pi\pi}(Z_t, \pi), \\ &a_\beta(X_t, \beta) = X_t,\ a_{\beta\beta}(X_t, \beta) = 0. \end{aligned} \tag{14.55}$$

Assumption S3*(ii) holds because $E_{\gamma_0} \sup_{\pi \in \Pi}(|m(Z_t, \pi)| + ||m_\pi(Z_t, \pi)||) \cdot (|m(Z_t, \pi)| +$

$||m_\pi(Z_t,$
$\pi)|| + ||m_{\pi\pi}(Z_t,\pi)||) \cdot ||X_t||^2 \leq C$ for some $C < \infty$ by Assumption STAR2(iii) and the Cauchy-Schwarz inequality.

This completes the verification of Assumption S3(i).

14.3. Verification of Assumption S3(ii)

Next, we verify Assumption S3(ii). We first show some generic results that are used in the calculation below. Let $A = aa'$, where $a = [a_1', ..., a_m']' \in R^{d_a}$ and $a_1, ..., a_m$ are sub-vectors of a. Similarly, $A^* = a^* a^{*\prime}$ and $a_1^*, ..., a_m^*$ are sub-vectors of a^*. Then,

$$\begin{aligned} ||A - A^*|| = ||aa' - a^* a^{*\prime}|| &\leq \sum_{i=1}^m \sum_{j=1}^m ||a_i a_j' - a_i^* a_j^{*\prime}|| \\ &\leq \sum_{i=1}^m \sum_{j=1}^m \left(||a_i a_j' - a_i a_j^{*\prime}|| + ||a_i a_j^{*\prime} - a_i^* a_j^{*\prime}|| \right) \\ &\leq \sum_{i=1}^m (||a_i|| + ||a_i^*||) \sum_{j=1}^m ||a_i - a_i^*||, \end{aligned} \tag{14.56}$$

where the first inequality holds by the inequality $(x^2 + y^2)^{1/2} \leq x + y$ for non-negative scalars x and y, the second inequality holds by the triangle inequality, and the third inequality holds by the inequality $||AB|| \leq ||A|| \cdot ||B||$ for matrices A and B.

By (7.12),

$$\begin{aligned} &||\rho_{\psi\psi}(W_t, \theta_1) - \rho_{\psi\psi}(W_t, \theta_2)|| \leq ||d_{\psi,t}(\pi_1) d_{\psi,t}(\pi_1)' - d_{\psi,t}(\pi_2) d_{\psi,t}(\pi_2)'|| \\ &\leq 4||X_t|| \cdot ||X_t' m(Z_t, \pi_1) - X_t' m(Z_t, \pi_2)|| \\ &\leq 4||X_t||^2 \cdot \sup_{\pi \in \Pi} ||m_\pi(Z_t, \pi)|| \cdot ||\pi_1 - \pi_2||, \end{aligned} \tag{14.57}$$

where the first inequality holds by applying the inequality in (14.56) to $a = d_{\psi,t}(\pi_1) = (X_t' m(Z_t, \pi_1), X_t')'$ and $a^* = d_{\psi,t}(\pi_2) = (X_t' m(Z_t, \pi_2), X_t')'$ and the second inequality holds by a mean-value expansion of $m(Z_t, \pi)$ wrt π.

Applying the arguments in (14.57) to $\rho_{\theta\theta}^\dagger(W_t, \theta^+)$ with $a = (X_t' m(Z_t, \pi_1), X_t', \omega_1' X_t m_\pi(Z_t,$

$\pi_1)')'$ and $a^* = (X_t'm(Z_t, \pi_2), X_t', \omega_2' X_t m_\pi(Z_t, \pi_2)')'$ yields

$$||\rho^\dagger_{\theta\theta}(W_t, \theta_1^+) - \rho^\dagger_{\theta\theta}(W_t, \theta_2^+)|| \leq 2||X_t||^2(2 + \sup_{\pi\in\Pi} ||m_\pi(Z_t, \pi)||) \times \quad (14.58)$$
$$\left(\sup_{\pi\in\Pi}(||m_\pi(Z_t, \pi)|| + ||m_{\pi\pi}(Z_t, \pi)||) \cdot ||\pi_1 - \pi_2|| + \sup_{\pi\in\Pi} ||m_\pi(Z_t, \pi)|| \cdot ||\omega_1 - \omega_2||\right).$$

Therefore, the function $M_1(W_t)$ in Assumption S3(ii) takes the form

$$M_1(W_t) = 4||X_t||^2 \cdot \sup_{\pi\in\Pi} ||m_\pi(Z_t, \pi)||$$
$$+2||X_t||^2(2 + \sup_{\pi\in\Pi} ||m_\pi(Z_t, \pi)||) \cdot \sup_{\pi\in\Pi}(2||m_\pi(Z_t, \pi)|| + ||m_{\pi\pi}(Z_t, \pi)||). \quad (14.59)$$

The form of $M_1(W_t)$ is used in the verification of Assumption S3(iii) below.

Next, we show the form of $M_2(W_t)$ in Assumption S3(ii) (vector β). By (7.12),

$$||\rho_\psi(W_t, \theta_1) - \rho_\psi(W_t, \theta_2)|| = ||U_t(\theta_1) d_{\psi,t}(\pi_1) - U_t(\theta_2) d_{\psi,t}(\pi_2)||$$
$$\leq |U_t(\theta_1) - U_t(\theta_2)| \cdot ||d_{\psi,t}(\pi_2)|| + |U_t(\theta_1)| \cdot ||d_{\psi,t}(\pi_1) - d_{\psi,t}(\pi_2)||, \quad (14.60)$$

where the inequality holds by the triangle inequality and $||aB|| = |a| \cdot ||B||$ when a is a scalar.

Let $\overline{\beta} = \sup_{\theta\in\Theta} ||\beta||$ and $\overline{\zeta} = \sup_{\theta\in\Theta} ||\zeta||$.

Note that in (14.60), the terms concerning $U_t(\theta)$ satisfy

$$|U_t(\theta_1) - U_t(\theta_2)| \leq \sup_{\theta\in\Theta} ||\frac{\partial}{\partial\theta'} U_t(\theta)|| \cdot ||\theta_1 - \theta_2||$$
$$\leq (2||X_t|| + ||X_t|| \cdot \overline{\beta} \cdot \sup_{\pi\in\Pi} ||m_\pi(Z_t, \pi)||) \cdot ||\theta_1 - \theta_2||,$$
$$|U_t(\theta_1)| \leq ||Y_t|| + ||X_t||\overline{\zeta} + ||X_t||\overline{\beta}. \quad (14.61)$$

The terms concerning $d_{\psi,t}(\pi)$ satisfy

$$||d_{\psi,t}(\pi_2)|| \leq 2||X_t|| \text{ and}$$
$$||d_{\psi,t}(\pi_1) - d_{\psi,t}(\pi_2)|| \leq ||X_t|| \cdot \sup_{\pi\in\Pi} ||m_\pi(Z_t, \pi)|| \cdot ||\pi_1 - \pi_2||. \quad (14.62)$$

The inequalities in (14.60)-(14.62) imply that

$$
\begin{aligned}
&||\rho_\psi(W_t,\theta_1)-\rho_\psi(W_t,\theta_2)|| \le M_\psi(W_t)\cdot||\theta_1-\theta_2||, \text{ where}\\
&M_\psi(W_t)=2||X_t||^2(2+\overline{\beta}\cdot\sup_{\pi\in\Pi}||m_\pi(Z_t,\pi)||)\\
&+(|Y_t|+||X_t||\overline{\zeta}+||X_t||\overline{\beta})\cdot||X_t||\cdot\sup_{\pi\in\Pi}||m_\pi(Z_t,\pi)||. \qquad (14.63)
\end{aligned}
$$

Similarly, (7.14) gives

$$
\begin{aligned}
&||\rho_\theta^\dagger(W_t,\theta_1^+)-\rho_\theta^\dagger(W_t,\theta_2^+)||=||U_t(\theta_1^+)d_t(\pi_1,\omega_1)-U_t(\theta_2^+)d_t(\pi_2,\omega_2)|| \qquad (14.64)\\
&\le|U_t(\theta_1^+)-U_t(\theta_2^+)|\cdot||d_t(\pi_2,\omega_2)||+|U_t(\theta_1^+)|\cdot||d_t(\pi_1,\omega_1)-d_t(\pi_2,\omega_2)||.
\end{aligned}
$$

In (14.64), the terms concerning $U_t(\theta^+)$ satisfy that

$$
\begin{aligned}
&U_t(\theta^+)=Y_t-X_t'\zeta-||\beta||\omega'X_t\cdot m(Z_t,\pi),\\
&|U_t(\theta_1^+)|\le||Y_t||+||X_t||\overline{\zeta}+||X_t||\overline{\beta},\\
&\frac{\partial}{\partial\theta^{+\prime}}U(\theta^+)=-(\omega'X_tm(Z_t,\pi),||\beta||X_t'm(Z_t,\pi),X_t',||\beta||\omega'X_tm_\pi(Z_t,\pi)'),\\
&|U_t(\theta_1^+)-U_t(\theta_2^+)|\le\sup_{\theta^+\in\Theta^+}||\frac{\partial}{\partial\theta^+}U(\theta^+)||\cdot||\theta_1^+-\theta_2^+||\\
&\le\left(2+\overline{\beta}\cdot\left(\sup_{\pi\in\Pi}||m_\pi(Z_t,\pi)||+1\right)\right)\cdot||X_t||\cdot||\theta_1^+-\theta_2^+||. \qquad (14.65)
\end{aligned}
$$

In (14.64), the terms concerning $d_t(\pi,\omega)$ satisfy

$$
\begin{aligned}
&||d_t(\pi,\omega)||\le||X_t||(2+\sup_{\pi\in\Pi}||m_\pi(Z_t,\pi)||) \text{ and}\\
&||d_t(\pi_1,\omega_1)-d_t(\pi_2,\omega_2)||\le||X_t||\cdot\left(\sup_{\pi\in\Pi}||m_\pi(Z_t,\pi)||+\sup_{\pi\in\Pi}||m_{\pi\pi}(Z_t,\pi)||\right)\cdot||\pi_1-\pi_2||\\
&+||X_t||\cdot\sup_{\pi\in\Pi}||m_\pi(Z_t,\pi)||\cdot||\omega_1-\omega_2||. \qquad (14.66)
\end{aligned}
$$

By (14.64)-(14.66),

$$
\begin{aligned}
&||\rho_\theta^\dagger(W_t,\theta_1^+)-\rho_\theta^\dagger(W_t,\theta_2^+)||\le M_\rho(W_t)\cdot||\theta_1^+-\theta_2^+||, \text{ where}\\
&M_\rho(W_t)=[2+\overline{\beta}\cdot(\sup_{\pi\in\Pi}||m_\pi(Z_t,\pi)||+1)]\cdot||X_t||^2\cdot(2+\sup_{\pi\in\Pi}||m_\pi(Z_t,\pi)||)\\
&+\left(||Y_t||+||X_t||\overline{\zeta}+||X_t||\overline{\beta}\right)\cdot||X_t||\cdot\sup_{\pi\in\Pi}(2||m_\pi(Z_t,\pi)||+||m_{\pi\pi}(Z_t,\pi)||) \qquad (14.67)
\end{aligned}
$$

Another term in Assumption S3(ii) is $||\varepsilon(W_t, \theta_1^+) - \varepsilon(W_t, \theta_2^+)||$, which satisfies

$$\begin{aligned}
&||\varepsilon(W_t, \theta_1^+) - \varepsilon(W_t, \theta_2^+)|| \\
&\leq |U_t(\theta_1^+) - U_t(\theta_2^+)| \cdot ||X_t|| \cdot \sup_{\pi \in \Pi}(2||m_\pi(Z_t, \pi) + ||m_{\pi\pi}(Z_t, \pi)||) \\
&\quad + |U_t(\theta_1^+)| \cdot ||X_t|| \cdot \sup_{\pi \in \Pi}(||3m_{\pi\pi}(Z_t, \pi) + M_{\pi\pi}(Z_t)) \cdot ||\theta_1^+ - \theta_2^+||, \qquad (14.68)
\end{aligned}$$

where $M_{\pi\pi}(Z_t)$ is as in Assumption STAR2. This and the inequalities in (14.65) imply that

$$\begin{aligned}
&||\varepsilon(W_t, \theta_1^+) - \varepsilon(W_t, \theta_2^+)|| \leq M_\varepsilon(W_t) \cdot ||\theta_1^+ - \theta_2^+||, \text{ where } M_\varepsilon(W_t) = \\
&\left(2 + \overline{\beta} \cdot \left(\sup_{\pi \in \Pi} ||m_\pi(Z_t, \pi)|| + 1\right)\right) \cdot ||X_t||^2 \cdot \sup_{\pi \in \Pi}(2||m_\pi(Z_t, \pi)|| + ||m_{\pi\pi}(Z_t, \pi)||) \\
&+ \left(|Y_t| + ||X_t||\overline{\zeta} + ||X_t||\overline{\beta}\right) \cdot ||X_t|| \cdot (\sup_{\pi \in \Pi} ||3m_{\pi\pi}(Z_t, \pi)|| + M_{\pi\pi}(Z_t)). \qquad (14.69)
\end{aligned}$$

Equations (14.63), (14.67), and (14.69) yield that Assumption S3(ii) holds with

$$M_2(W_t) = M_\psi(W_t) + M_\rho(W_t) + M_\varepsilon(W_t). \qquad (14.70)$$

14.4. Verification of Assumption S3(iii)

In the verification of Assumption S3(iii) below, we use

$$\begin{aligned}
E_{\gamma_0} \sup_{\theta \in \Theta} |U_t(\theta)|^{2q} &= E_{\gamma_0} \sup_{\theta \in \Theta} |Y_t - X_t'\zeta - X_t'\beta \cdot m(Z_t, \pi)|^{2q} \\
&\leq C_1 E_{\gamma_0}(|Y_t| + ||X_t||)^{2q} \leq C_2 \qquad (14.71)
\end{aligned}$$

for some $C_1, C_2 < \infty$, where the first inequality holds because the parameter spaces of ζ and β are bounded and $|m(Z_t, \pi)| \in [0, 1]$ and the second inequality holds by Holder's inequality and Assumptions STAR1(ii) and STAR2(iii). Because the value of $U_t(\theta)$ does not change when θ is reparameterized as θ^+, (14.71) is equivalent to

$$E_{\gamma_0} \sup_{\theta^+ \in \Theta^+} |U_t(\theta^+)|^{2q} \leq C \qquad (14.72)$$

for some $C < \infty$.

By (14.46),

$$E_{\gamma_0} \sup_{\theta \in \Theta} |\rho(W_t, \theta)|^{1+\delta} = \frac{1}{2^{1+\delta}} E_{\gamma_0} \sup_{\theta \in \Theta} |U_t(\theta)|^{2(1+\delta)} \leq C \tag{14.73}$$

for some $C < \infty$ by (14.71).

By (7.14),

$$\begin{aligned} E_{\gamma_0} \sup_{\theta^+ \in \Theta^+} ||\rho_\theta^\dagger(W_t, \theta^+)||^q &\leq E_{\gamma_0} \sup_{\theta^+ \in \Theta^+} |U_i(\theta^+)|^{2q} E_{\gamma_0} \sup_{\theta^+ \in \Theta^+} ||d_t(\pi, \omega)||^{2q} \\ &\leq C_1 E_{\gamma_0} \sup_{\pi \in \Pi} (2||X_t|| + ||X_t|| \cdot ||m_\pi(Z_t, \pi)||)^{2q} \leq C_2 \end{aligned} \tag{14.74}$$

for some $C_1, C_2 < \infty$, where the first inequality holds by the Cauchy-Schwarz inequality, the second inequality holds by (14.72) and $||AB|| \leq ||A|| \cdot ||B||$, and the third inequality holds by Holder's inequality and Assumptions STAR1(ii) and STAR2(iii).

In the calculation of $E_{\gamma_0} \sup_{\theta \in \Theta} ||\rho_{\psi\psi}(W_i, \theta)||^{1+\delta}$ and $E_{\gamma_0} \sup_{\theta \in \Theta} ||\rho_{\theta\theta}^\dagger(W_i, \theta)||^{1+\delta}$ below, we use the following inequality. Let $A = aa'$, where $a = [a_1', ..., a_m'] \in R^{d_a}$ and $a_1, ..., a_m$ are sub-vectors of a. Then,

$$||A|| \leq \sum_{i=1}^m \sum_{j=1}^m ||a_i a_j'|| \leq \left(\sum_{i=1}^m ||a_i|| \right)^2, \tag{14.75}$$

by arguments analogous to those in (14.56).

By (7.12),

$$E_{\gamma_0} \sup_{\theta \in \Theta} ||\rho_{\psi\psi}(W_i, \theta)||^{1+\delta} = E_{\gamma_0} \sup_{\theta \in \Theta} ||d_{\psi,t}(\pi) d_{\psi,t}(\pi)'||^{1+\delta} \leq E_{\gamma_0} (2||X_t||)^{2(1+\delta)} \leq C \tag{14.76}$$

for some $C < \infty$, where the first inequality holds by (14.75) with $a = (X_t' m(Z_t, \pi), X_t')'$ and the second inequality holds by Assumptions STAR1(ii) and STAR2(iii).

Similarly, by (7.14),

$$\begin{aligned} &E_{\gamma_0} \sup_{\theta^+ \in \Theta^+} ||\rho_{\theta\theta}^\dagger(W_i, \theta^+)||^{1+\delta} = E_{\gamma_0} \sup_{\theta^+ \in \Theta^+} ||d_t(\pi, \omega) d_t(\pi, \omega)'||^{1+\delta} \\ &\leq E_{\gamma_0} \sup_{\pi \in \Pi} (2||X_t|| + ||X_t|| \cdot ||m_\pi(Z_t, \pi)||)^{2(1+\delta)} \leq C \end{aligned} \tag{14.77}$$

for some $C < \infty$, where the first inequality holds by (14.75) with $a = (X_t' m(Z_t, \pi), X_t', \omega' X_t m_\pi(Z_t, \pi))'$ and the second inequality holds by Assumption STAR2.

By (7.14),

$$
\begin{aligned}
&E_{\gamma_0} \sup_{\theta^+ \in \Theta^+} ||\varepsilon(W_i, \theta^+)||^q \leq E_{\gamma_0} \sup_{\theta^+ \in \Theta^+} ||U_t(\theta^+)||^{2q} \\
&\times E_{\gamma_0} \sup_{\pi \in \Pi} (2||X_t|| \cdot ||m_\pi(Z_t, \pi)|| + ||X_t|| \cdot ||m_{\pi\pi}(Z_t, \pi)||)^{2q} \leq C \qquad (14.78)
\end{aligned}
$$

for some $C < \infty$, where the first inequality holds by the Cauchy-Schwarz inequality and the inequality $||A|| \leq \sum_{i,j} ||A_{i,j}||$ for any matrix A, where $A_{i,j}$ denotes an element of A, and the second inequality follows from (14.72), Holder's inequality, and Assumptions STAR1(ii) and STAR2(iii).

Finally, $E_{\gamma_0}(M_1(W_t) + M_2(W_t)^q) \leq C$ for some $C < \infty$ by Holder's inequality, (14.59), (14.63), (14.67), (14.69), (14.70), and Assumptions STAR(ii) and STAR2(iii).

This completes the verification of Assumption S3(iii) (vector β).

14.5. Verification of Assumptions S3(iv) and S3(v)

To verify Assumption S3(iv), note that

$$
\begin{aligned}
E_{\gamma_0} \rho_{\psi\psi}(W_t, \psi_0, \pi) &= E_{\gamma_0} d_{\psi,t}(\pi) d_{\psi,t}(\pi)' \text{ and} \\
E_{\gamma_0} \rho^{\dagger}_{\theta\theta}(W_t, \theta_0) &= E_{\gamma_0} d_t(\pi_0, \omega_0) d_t(\pi_0, \omega_0)'. \qquad (14.79)
\end{aligned}
$$

For any $\lambda = (\lambda_1, \lambda_2) \neq 0$, $\lambda_1, \lambda_2 \in R^{d_\beta}$, and $\forall \pi \in \Pi$,

$$
\lambda' E_{\gamma_0} d_{\psi,t}(\pi) d_{\psi,t}(\pi)' \lambda = E_{\gamma_0} \left(\lambda_1' X_t m(Z_t, \pi) + \lambda_2' X_t\right)^2 > 0, \qquad (14.80)
$$

where the inequality holds by Assumption STAR2(i). This implies that $E_{\gamma_0} d_{\psi,t}(\pi) d_{\psi,t}(\pi)'$ is positive definite $\forall \pi \in \Pi$.

For any $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4) \neq 0$, $\lambda_1, \lambda_2 \in R^{d_\beta}$, $\lambda_3, \lambda_4 \in R$, $\forall \omega$ with $||\omega|| = 1$ and $\forall \pi \in \Pi$,

$$
\begin{aligned}
&\lambda' E_{\gamma_0} d_t(\pi, \omega) d_t(\pi, \omega)' \lambda \qquad (14.81) \\
&= E_{\gamma_0} (\lambda_1' X_t m(Z_t, \pi) + \lambda_2' X_t + \lambda_3 \omega' X_t m_{\pi,1}(Z_t, \pi) + \lambda_4 \omega' X_t m_{\pi,2}(Z_t, \pi))^2 > 0,
\end{aligned}
$$

where the inequality holds by Assumption STAR2(ii) with $a = (\lambda_1, \lambda_2, \lambda_3 \omega, \lambda_4 \omega)$. Note that $\lambda \neq 0$ implies that $a \neq 0$. The inequality in (14.81) implies that $E_{\gamma_0} \rho^{\dagger}_{\theta\theta}(W_t, \theta_0)$ is positive definite $\forall \gamma_0 \in \Gamma$.

To verify Assumption S3(v), note that $\forall m \neq 0$,

$$Cov_{\phi}(\rho_{\theta}^{\dagger}(W_t,\theta_0),\rho_{\theta}^{\dagger}(W_{t+m},\theta_0)) = E_{\gamma_0}U_tU_{t+m}d_t(\pi_0,\omega_0)d_{t+m}(\omega_0,\pi_0)' = 0 \tag{14.82}$$

by Assumption STAR1(i). This yields that

$$\begin{aligned} V^{\dagger}(\theta_0,\theta_0;\gamma_0) &= Cov_{\phi_0}U_t^2 d_t(\pi_0,\omega_0)d_t(\pi_0,\omega_0)' \\ &= E_{\gamma_0}U_t^2 d_t(\pi_0,\omega_0)d_t(\pi_0,\omega_0)', \end{aligned} \tag{14.83}$$

where the second equality uses $E_{\gamma_0}U_t d_t(\pi_0,\omega_0) = 0$ by Assumption STAR1(i). The matrix $E_{\gamma_0}U_t^2 d_t(\pi_0,\omega_0)d_t(\pi_0,\omega_0)'$ is positive definite by the argument in (14.81) with $d_{\psi,t}(\pi)$ replaced by $U_t d_t(\pi,\omega)$ and using $E_{\gamma_0}(U_t^2|\mathcal{F}_{t-1}) = \sigma^2 > 0$.

14.6. Verification of Assumption S4

To verify Assumption S4, we have

$$\begin{aligned} E_{\gamma_0}\rho_{\psi}(W_t,\theta) &= -E_{\gamma_0}U_t(\theta)d_{\psi,t}(\pi) \\ &= -E_{\gamma_0}(U_t + X_t'(\zeta_0-\zeta) + X_t'[\beta_0 m(Z_t,\pi_0) - \beta m(Z_t,\pi)])d_{\psi,t}(\pi) \text{ and} \\ K(\theta;\gamma_0) &= -E_{\gamma_0}d_{\psi,t}(\pi)X_t'm(Z_t,\pi_0) \\ &= -E_{\gamma_0}d_{\psi,t}(\pi)d_{\psi,t}(\pi_0)'\cdot S_{\beta}'. \end{aligned} \tag{14.84}$$

where $S_{\beta} = [I_{d_\beta} : 0] \in R^{d_\beta \times (2d_\beta)}$.

Assumption S4(i) holds with $K(\theta;\gamma_0)$ in (14.84) by the moment conditions in Assumption STAR2(iii). To verify Assumption S4(ii), we need to show that $E_{\gamma_0}d_{\psi,t}(\pi)d_{\psi,t}(\pi_0)'$ is continuous in π,π_0, and ϕ. Continuity in π and π_0 follows from the the continuity of $m(Z_t,\pi)$ in π and the moment conditions in Assumption STAR2(iii). Continuity in ϕ holds because $\phi_n \to \phi_0$ under d_Φ implies weak convergence of (Y_t,Y_{t+m}) for all $t,m \geq 1$, which in turn implies the convergence of $E_{\gamma_n}d_{\psi,t}(\pi)d_{\psi,t}(\pi_0)'$ to $E_{\gamma_0}d_{\psi,t}(\pi)d_{\psi,t}(\pi_0)'$ by the moment conditions in Assumption STAR2(iii).

The continuity in π,π_0, and ϕ holds uniformly over $\pi \in \Pi$ by Lemma 9.2 using (i) the pointwise convergence above, (ii) the fact that $E_{\gamma_0}d_{\psi,t}(\pi)d_{\psi,t}(\pi_0)'$ is differentiable in π and the partial derivative is bounded over $\pi \in \Pi$, and (iii) the compactness of Π. This completes the verification of Assumption S4.

14.7. Verification of Assumptions B1 and B2

Now we verify Assumptions B1 and B2. Assumptions B1(i) and B1(iii) hold by Assumptions STAR5(i) and STAR5(ii) immediately. Assumption B1(ii) holds with $\mathcal{Z}^0 = int(\mathcal{Z}_0)$ by Assumptions STAR4(iv) and STAR5(iii). Assumption B2(i) holds immediately because the true parameter space Γ is of the form in (2.6) and Γ is assumed to be compact. Assumption B2(ii) holds by Assumption STAR4(ii). Assumption B2(iii) holds by Assumption STAR4(iv) and the form of the true parameter space in (7.10).

14.8. Verification of Assumptions C6 and C7

Assumption C6 is implied by Assumption STAR3(i).

Now we verify Assumption C7 with $H(\pi;\gamma_0)$ and $K(\pi;\gamma_0)$ given in (7.9). By the matrix Cauchy-Schwarz inequality in Tripathi (1999),

$$K(\pi;\gamma_0)'H^{-1}(\pi;\gamma_0)K(\pi;\gamma_0) \leq E_{\gamma_0}X_tX_t'm^2(Z_t,\pi_0). \tag{14.85}$$

The matrix "$\leq$" holds as an equality if and only if $X_tm(Z_t,\pi_0)a+(X_t', X_t'm(Z_t,\pi))c = 0$ with probability 1 for some $a \in R^{d_\beta}$ and $c \in R^{2d_\beta}$ with $(a',c')' \neq 0$. The "$\leq$" holds as an equality uniquely at $\pi = \pi_0$ by Assumption STAR2(i).

Proof of Lemma 7.1. We prove Lemma 7.1 by verifying Assumption C6† and using Lemma 3.2. Note that

$$\begin{aligned}
\rho_\beta(W_t,\psi_0,\pi) &= U_tX_tm(Z_t,\pi), \\
\rho_\zeta(W_t,\psi_0,\pi) &= U_tX_t, \\
\rho_\psi^*(W_i,\psi_0,\pi_1,\pi_2) &= U_td_\psi^*(\pi_1,\pi_2), \text{ where} \\
d_\psi^*(\pi_1,\pi_2) &= (X_t'm(Z_t,\pi_1), X_t'm(Z_t,\pi_2), X_t')'.
\end{aligned} \tag{14.86}$$

The matrix $\Omega_G(\pi_1,\pi_2;\gamma_0)$ that appears in Assumption C6† takes the form

$$\Omega_G(\pi_1,\pi_2;\gamma_0) = E_{\gamma_0}U_t^2d_\psi^*(\pi_1,\pi_2)d_\psi^*(\pi_1,\pi_2)' \tag{14.87}$$

by Assumption STAR1(i). Assumption C6†(ii) holds by Assumption STAR2(i) and $E_{\gamma_0}(U_t^2|\mathcal{F}_{t-1}) = \sigma^2 > 0$ using arguments analogous to those in (14.81). $\square$

14.9. Verification of Assumptions V1 (vector β) and V2

Here we verify Assumptions V1 (vector β) and V2, which are stated in Appendix B of the main paper.

In the STAR model, Assumption V1(i) holds with

$$
\begin{aligned}
&J(\theta^+;\gamma_0) = E_{\gamma_0} d_t(\pi,\omega)\, d_t(\pi,\omega)' \text{ and} \\
&V(\theta^+;\gamma_0) = E_{\gamma_0} U_t^2 d_t(\pi,\omega)\, d_t(\pi,\omega)' \qquad (14.88) \\
&+E_{\gamma_0}[X_t'(\zeta_0-\zeta) + X_t'(||\beta_0||\omega_0 m(Z_t,\pi_0) - ||\beta||\omega m(Z_t,\pi))]^2 d_t(\pi,\omega) d_t(\pi,\omega)',
\end{aligned}
$$

by the uniform law of large numbers in Lemma 9.3.

Assumption V1(ii) holds by the continuity of $m(z,\pi)$ and $m_\pi(z,\pi)$ in π and Assumption STAR2(iii).

To verify Assumption V1(iii), note that $\Sigma(\pi,\omega;\gamma_0)$ takes the form

$$
\begin{aligned}
&\Sigma(\pi,\omega;\gamma_0) \qquad (14.89) \\
&= \left(E_{\gamma_0} d_t(\pi,\omega)\, d_t(\pi,\omega)'\right)^{-1} E_{\gamma_0} U_t^2 d_t(\pi,\omega)\, d_t(\pi,\omega)' \left(E_{\gamma_0} d_t(\pi,\omega)\, d_t(\pi,\omega)'\right)^{-1}.
\end{aligned}
$$

Given that $E_{\gamma_0} d_t(\pi,\omega)\, d_t(\pi,\omega)'$ and $E_{\gamma_0} U_t^2 d_t(\pi,\omega)\, d_t(\pi,\omega)'$ are both positive definite, $\Sigma(\pi,\omega;\gamma_0)$ is positive definite $\forall \pi \in \Pi$ and $\forall \omega$ with $||\omega|| = 1$.

Because the determinant of $E_{\gamma_0} d_t(\pi,\omega)\, d_t(\pi,\omega)'$ is bounded away from 0 as a function of (π,ω) $\forall \gamma_0 \in \Gamma$ and $||E_{\gamma_0} d_t(\pi,\omega)\, d_t(\pi,\omega)'|| \leq C_1$ for some $C_1 < \infty$ $\forall \gamma_0 \in \Gamma$ by Assumption STAR2(iii), we have $||\left(E_{\gamma_0} d_t(\pi,\omega)\, d_t(\pi,\omega)'\right)^{-1}|| \leq C_2$ for some $C_2 < \infty$. Hence, $||\Sigma(\pi,\omega;\gamma_0)|| \leq C$ $\forall \pi \in \Pi$ and $\forall \omega$ with $||\omega|| = 1$. This completes the verification of Assumption V1(iii).

Assumption V1(iv) holds by Assumption STAR3(ii).

Assumptions V1(i) and V1(ii) hold not only under $\{\gamma_n\} \in \Gamma(\gamma_0, 0, b)$, but also under $\{\gamma_n\} \in \Gamma(\gamma_0, \infty, \omega_0)$ in this example. This and $\widehat{\theta}_n \to_p \theta_0$ under $\{\gamma_n\} \in \Gamma(\gamma_0, \infty, \omega_0)$, which holds by Lemma 5.3 of AC1, imply that Assumption V2 holds. Regarding the assumptions employed in Lemma 5.3 of AC1, Assumptions B1, B2, C6, and C7 are verified above and Assumptions A, B3, and C1-C4 hold by Lemma 9.1 under Assumptions B1, B2, and S1-S4. $\square$

References

Andrews, D. W. K., Cheng, X., 2007. Estimation and inference with weak, semi-strong, and strong identification. Cowles Foundation Discussion Paper No. 1773R, Yale University.

Tripathi, G. T., 1999. A matrix extension of the Cauchy-Schwarz inequality. Economics Letters 63, 1-3.