

Online Appendix: Estimation and Inference with a (Nearly) Singular Jacobian

Sukjin Han
Department of Economics
University of Texas at Austin
sukjin.han@austin.utexas.edu

Adam McCloskey
Department of Economics
University of Colorado at Boulder
adam.mccloskey@colorado.edu

First Draft: February 15, 2014

This Draft: November 26, 2018

Abstract

This online appendix provides a proof of the claim that the quantity $\|\tilde{\eta}(\hat{\mu}_n)\|$ diverges in the context of the Threshold Crossing Model made in Section 4 of [Han and McCloskey \(2018\)](#) as well as a set of primitive conditions and corresponding verification of the high level assumptions made in [Han and McCloskey \(2018\)](#) for the Threshold Crossing Model.

A Assumption Verifications for Threshold Crossing Example

Before proceeding to verify the assumptions imposed for the Threshold Crossing Model example, we provide the details for the claim that $\|\tilde{\eta}(\hat{\mu}_n)\|$ diverges for $\hat{\mu}_n^s = (\hat{\mu}_{n,3}, \hat{\mu}_{n,4})$ made in the continuation of Example 2.3 in Section 4 of Han and McCloskey (2018) (HM18 henceforth).

Proof $\|\tilde{\eta}(\hat{\mu}_n)\|$ diverges in Example 2.3: Note that

$$\begin{aligned}
\tilde{\eta}_n(\hat{\mu}_n) &= \sqrt{n}\mathcal{S}(\hat{\mu}_n) \left[h_3(\zeta_n, \hat{\pi}_n) - h_3(\zeta_n, \pi_n) + \frac{C_3(h_3(\hat{\mu}_n), \hat{\zeta}_{1,n}; \hat{\pi}_n)}{C_1(h_3(\hat{\mu}_n), \hat{\zeta}_{1,n}; \hat{\pi}_n)} (\hat{\pi}_n - \pi_n) \right] \\
&= \sqrt{n}\mathcal{S}(\hat{\mu}_n) \left[\frac{\zeta_{3,n}(\zeta_{1,n} - 1)(\zeta_{1,n} - \zeta_{3,n})}{(\zeta_{1,n} - \zeta_{3,n}\hat{\pi}_n + \zeta_{1,n}\zeta_{3,n}\hat{\pi}_n)(\zeta_{1,n} - \zeta_{3,n}\pi_n + \zeta_{1,n}\zeta_{3,n}\pi_n)} \right. \\
&\quad \left. + \frac{C_3(h_3(\hat{\mu}_n), \hat{\zeta}_{1,n}; \hat{\pi}_n)}{C_1(h_3(\hat{\mu}_n), \hat{\zeta}_{1,n}; \hat{\pi}_n)} \right] (\hat{\pi}_n - \pi_n) \\
&= \sqrt{n}\mathcal{S}(\hat{\mu}_n) \left[\frac{\zeta_{3,n}(\zeta_{1,n} - 1)(\zeta_{1,n} - \zeta_{3,n})}{(\zeta_{1,n} - \zeta_{3,n}\hat{\pi}_n + \zeta_{1,n}\zeta_{3,n}\hat{\pi}_n)(\zeta_{1,n} - \zeta_{3,n}\pi_n + \zeta_{1,n}\zeta_{3,n}\pi_n)} \right. \\
&\quad \left. - \frac{\hat{\zeta}_{3,n}(\hat{\zeta}_{1,n} - 1)(\hat{\zeta}_{1,n} - \hat{\zeta}_{3,n})}{(\hat{\zeta}_{1,n} - \hat{\zeta}_{3,n}\hat{\pi}_n + \hat{\zeta}_{1,n}\hat{\zeta}_{3,n}\hat{\pi}_n)^2} \right] (\hat{\pi}_n - \pi_n) \\
&= \sqrt{n}\mathcal{S}(\hat{\mu}_n) \left[\frac{\tilde{\eta}_n^N(\hat{\mu}_n)}{\tilde{\eta}_n^D(\hat{\mu}_n)} \right] (\hat{\pi}_n - \pi_n),
\end{aligned}$$

where

$$\begin{aligned}
\tilde{\eta}_n^N(\hat{\mu}_n) &= \zeta_{3,n}(\zeta_{1,n} - 1)(\zeta_{1,n} - \zeta_{3,n})(\hat{\zeta}_{1,n} - \hat{\zeta}_{3,n}\hat{\pi}_n + \hat{\zeta}_{1,n}\hat{\zeta}_{3,n}\hat{\pi}_n) \\
&\quad - \hat{\zeta}_{3,n}(\hat{\zeta}_{1,n} - 1)(\hat{\zeta}_{1,n} - \hat{\zeta}_{3,n})(\zeta_{1,n} - \zeta_{3,n}\pi_n + \zeta_{1,n}\zeta_{3,n}\pi_n) \\
&= \zeta_{3,n}(\zeta_{1,n} - 1)(\zeta_{1,n} - \zeta_{3,n})[(\hat{\zeta}_{1,n} - \hat{\zeta}_{3,n}\hat{\pi}_n + \hat{\zeta}_{1,n}\hat{\zeta}_{3,n}\hat{\pi}_n) - (\zeta_{1,n} - \zeta_{3,n}\pi_n + \zeta_{1,n}\zeta_{3,n}\pi_n)] \\
&\quad + [\zeta_{3,n}(\zeta_{1,n} - 1)(\zeta_{1,n} - \zeta_{3,n}) - \hat{\zeta}_{3,n}(\hat{\zeta}_{1,n} - 1)(\hat{\zeta}_{1,n} - \hat{\zeta}_{3,n})](\zeta_{1,n} - \zeta_{3,n}\pi_n + \zeta_{1,n}\zeta_{3,n}\pi_n) \\
&= \zeta_{3,n}^2(\zeta_{1,n} - 1)^2(\zeta_{1,n} - \zeta_{3,n})(\hat{\pi}_n - \pi_n) + O_p(n^{-1/2}) = O_p(n^{-1/2}\|\beta_n\|^{-1})
\end{aligned}$$

with the final two equalities resulting from Lemma A.1 of HM18 and a mean value expansion of the term $\hat{\zeta}_{3,n}(\hat{\zeta}_{1,n} - 1)(\hat{\zeta}_{1,n} - \hat{\zeta}_{3,n})$, and

$$\tilde{\eta}_n^D(\hat{\mu}_n) = (\zeta_{1,n} - \zeta_{3,n}\hat{\pi}_n + \zeta_{1,n}\zeta_{3,n}\hat{\pi}_n + O_p(n^{-1/2}))^2(\zeta_{1,n} - \zeta_{3,n}\pi_n + \zeta_{1,n}\zeta_{3,n}\pi_n) = O_p(1)$$

by Lemma A.1 of HM18. Noting that both $\mathcal{S}(\hat{\mu}_n)$ and $\tilde{\eta}_n^D(\hat{\mu}_n)^{-1}$ are also $O_p(1)$ by Lemma A.1 of HM18, we may combine the expressions for $\tilde{\eta}_n(\hat{\mu}_n)$, $\mathcal{S}(\hat{\mu}_n)$, $\tilde{\eta}_n^N(\hat{\mu}_n)$ and $\tilde{\eta}_n^D(\hat{\mu}_n)$ to conclude that $\|\tilde{\eta}_n(\hat{\mu}_n)\| = \|O_p(n^{-1/2}\|\beta_n\|^{-1})\sqrt{n}(\hat{\pi}_n - \pi_n)\| = \|O_p(n^{-1/2}\|\beta_n\|^{-2})\| \rightarrow \infty$, according to Lemma A.1 of HM18. ■

We now proceed to verify the imposed assumptions for the Threshold Crossing Model example. Hereafter, [Andrews and Cheng \(2013a\)](#) and [Han and Vytlačil \(2017\)](#) are abbreviated as AC13 and HV17. The supplemental material for AC12, AC13 and AC14, [Andrews and Cheng \(2012, 2013b, 2014\)](#), are abbreviated as AC12supp, AC13supp and AC14supp. The working paper version of AC13 is abbreviated as ACMLwp. And “with respect to” is abbreviated as “w.r.t.”

A.1 Assumptions for Threshold Crossing Models

The assumptions in the main text of the current paper and the assumptions in AC12 on objects involving the transformed parameter θ are verified under assumptions introduced in this section. The assumptions in AC12 are verified by verifying those in AC13.

Assumption TC1: $\{W_i = (Y_i, D_i, Z_i) : i \geq 1\}$ is an i.i.d. sequence.

Assumption TC2: (i) $Z \perp (\varepsilon, \nu)$;

(ii) F_ε and F_ν are known marginal distributions of ε and ν , respectively, that are strictly increasing and absolutely continuous with respect to the Lebesgue measure such that $E[\varepsilon] = E[\nu] = 0$ and $Var(\varepsilon) = Var(\nu) = 1$;

(iii) $(\varepsilon, \nu)' \sim F_{\varepsilon\nu}(\varepsilon, \nu) = C(F_\varepsilon(\varepsilon), F_\nu(\nu); \pi)$ where $C : (0, 1)^2 \rightarrow (0, 1)$ is a copula known up to a scalar parameter $\pi \in \Pi$ such that $C(u_1, u_2, ; \pi)$ is three-times differentiable in $(u_1, u_2, \pi) \in (0, 1)^2 \times \Pi$;

(iv) The copula $C(u_1, u_2, ; \pi)$ satisfies

$$C(u_1|u_2; \pi) \prec_S C(u_1|u_2; \pi') \text{ for any } \pi < \pi', \quad (\text{A.1})$$

where “ \prec_S ” is a stochastic ordering defined in HV17 (Definition 3.2);

(v) $(1, Z)$ does not lie in a proper linear subspace of \mathbb{R}^2 wp1;

(vi) Θ^* is compact and convex.

Given the form of h in (3.8) of HM18 with $c_4(\zeta)$ set equal to zero, we write $\pi = \pi_3$ in this assumption and below. The conditions in TC2 are sufficient for (global) identification of θ when $\beta \neq 0$. The argument is similar to that in HV17, except that the condition for the parameter space TC2(vi) is stronger than that in HV17.

For the next assumption, define $\Theta_\delta^* \equiv \{\theta \in \Theta^* : |\beta| < \delta\}$ for some $\delta > 0$.

Assumption TC3: (i) $\Theta \equiv \Theta_{-\pi} \times \Pi$, and $\Theta_{-\pi}$ and Π are compact and simply connected;

(ii) $\text{int}(\Theta) \supset \Theta^*$;

(iii) For some $\delta > 0$, $\Theta \supset \{\beta \in \mathbb{R}^{d_\beta} : |\beta| < \delta\} \times \mathcal{Z}^0 \times \Pi \supset \Theta_\delta^*$ for some non-empty open set $\mathcal{Z}^0 \subset \mathbb{R}^{d_\mu - d_\pi}$ and Π .

(iv) $h^{-1}(\mathcal{Z}^0 \times \Pi) = \mathcal{Z}^0 \times \Pi$ for some non-empty open set $\mathcal{Z}^0 \subset \mathbb{R}^{d_\mu - d_\pi}$.

As is typical, Assumption TC3(i)-(ii) will be satisfied by a proper choice of the optimization parameter space. For concreteness, we define

$$\Theta^* \equiv \{\theta = (\beta, \zeta, \pi_1, \pi_2, \pi_3) \in [-0.98, 0.98] \times [0.01, 0.99] \times [0.01, 0.99] \times [0.01, 0.99] \times [-0.99, 0.99] : 0.01 \leq \beta + \zeta \leq 0.99\} \quad (\text{A.2})$$

and

$$\Theta \equiv \{\theta = (\beta, \zeta, \pi_1, \pi_2, \pi_3) \in [-0.98 - \epsilon, 0.98 + \epsilon] \times [0.01 - \epsilon, 0.99 + \epsilon] \times [0.01 - \epsilon, 0.99 + \epsilon] \times [0.01 - \epsilon, 0.99 + \epsilon] \times [-0.99 - \epsilon, 0.99 + \epsilon] : 0.01 - \epsilon \leq \beta + \zeta \leq 0.99 + \epsilon\} \quad (\text{A.3})$$

for some $\epsilon > 0$ so that TC3(i)-(ii) is clearly satisfied for small enough ϵ . Given the definition (A.2), TC4 below also holds if we define the parameter space $\Phi^*(\theta)$ of $\phi \equiv \phi_1$ as

$$\Phi^*(\theta) = \Phi^* \equiv [0.01, 0.99]. \quad (\text{A.4})$$

TC3(iii) is satisfied by setting

$$\mathcal{Z}^0 \equiv (0.01 - \delta, 0.99 + \delta)^3$$

for $\delta < \epsilon/2$. For TC3(iv), let $\tilde{h}^{-1}(\zeta, \pi) = (h_1^{-1}(\zeta, \pi), h_2^{-1}(\zeta, \pi), h_3^{-1}(\zeta, \pi))$, the first three elements of (3.12) of HM18. Note that $h_4(\zeta, \pi) = \pi$ (i.e., $\pi_3 = \pi$) and for any given $\pi \in \Pi$, $\tilde{h}^{-1}(\mathcal{Z}^0, \pi)$ does not depend on π . Thus, we may set $\mathcal{Z}^0 = \tilde{h}^{-1}(\mathcal{Z}^0, \pi)$ for any $\pi \in \Pi$, noting that \mathcal{Z}^0 must be a non-empty open set by the continuity of the first three elements of $h(\cdot)$. The latter follows from TC2(iii) and (3.8) of HM18 after setting $c_1(\zeta) = \zeta_1$, $c_2(\zeta) = \zeta_2$ and $c_3(\zeta) = \zeta_3$.

- Assumption TC4:** (i) Γ is compact and $\Gamma = \{\gamma = (\theta, \phi) : \theta \in \Theta^*, \phi \in \Phi^*(\theta)\}$;
(ii) $\forall \delta > 0, \exists \gamma = (\beta, \mu, \phi) \in \Gamma$ with $0 < |\beta| < \delta$;
(iii) $\forall \gamma = (\beta, \mu, \phi) \in \Gamma$ with $0 < |\beta| < \delta$ for some $\delta > 0$, $\gamma_a = (a\beta, \mu, \phi) \in \Gamma \forall a \in [0, 1]$.

Assumption TC4(ii) guarantees that the true parameter space includes a region where weak identification occurs and TC4(iii) ensures that Γ is consistent with the existence of $K(\theta; \gamma)$, defined later.

- Assumption TC5:** (i) $C(u_1, u_2, ; \pi)$ is bounded away from zero over $(0, 1)^2 \times \Pi$;
(ii) $0 < \phi_1 \equiv \Pr_\gamma[Z = 1] < 1 \forall \gamma \in \Gamma$.

Lemma A.1. *TC5 and TC2(iii) imply the following: for $(y, d, z) \in \{0, 1\}^3, \forall \gamma = (\theta, \phi) \in \Gamma$, and $\forall \gamma = (\theta, \phi) \in \Gamma$,*

- (i) *the first, second, and third order derivatives of $\mathbf{p}_{y,d,z}(\theta)$ are bounded over Θ ;*
(ii) *$\mathbf{p}_{y,d,z}(\theta)$ is bounded away from zero over Θ and $0 < \phi_1 < 1$;*
(iii) *$\bar{h}(\theta)$ is three-times differentiable on Θ ;*
(iv) *$p_{y,d,z}(\theta) \equiv \mathbf{p}_{y,d,z}(\bar{h}(\theta))$ is three-times differentiable on Θ and the first, second, and third*

order derivatives of $p_{y_d,z}(\theta)$ are bounded over Θ ;
(v) $p_{y_d,z}(\theta)$ is bounded away from zero over Θ .

Proof of Lemma A.1: (i) holds by TC2(iii), the fact that the domain Θ is compact by TC3(i), and the definitions of $\mathbf{p}_{y_d,z}(\theta)$. (ii) immediately holds by TC5. For (iii), given (3.9) of HM18, TC2(iii) and TC3(i) imply that $h(\mu)$ is three-times differentiable in μ and hence $\bar{h}(\theta) = (\beta, h(\mu))$ is three-times differentiable in θ . Next, (iv) holds by (i), (iii), and the chain rule, and (v) trivially holds by (ii).

A.2 Verification of Assumptions in the Main Text

Assumptions CF, ID, Jac, and Reg3 are verified in the main text. Assumption Reg1 is satisfied with $\bar{\mathbf{g}}_n(\theta) = \hat{\xi}_n - \mathbf{g}(\theta)$, where each element $\mathbf{p}_{y_d,z}(\theta)$ of the vector $\mathbf{g}(\theta)$ is continuously differentiable by TC2(iii). For Assumption H, H(i) holds since its sufficient conditions that Θ is bounded and h is continuous hold by S2(v), verified below, and by Proposition 3.1 of HM18, respectively. H(ii) is also trivially satisfied by TC3(i). For Reg2, $\text{rank}(h_\pi^s(\mu)) = 1$ if $h^s(\pi)$ contains $h_2(\pi)$, $h_3(\pi)$ or $h_4(\pi)$ and $\text{rank}(h_\pi^s(\mu)) = 0$ otherwise, as can be seen from the form of h in (3.8) of HM18 upon setting $c_1(\zeta) = \zeta_1$, $c_2(\zeta) = \zeta_2$, $c_3(\zeta) = \zeta_3$ and $c_4(\zeta) = 0$.

A.3 Verification of Assumptions in Andrews and Cheng (2013)

In this section, given our transformed parameter θ and associated transformed objects, we verify the regularity conditions for the asymptotic theory of the ML estimator $\hat{\theta}_n$ in AC13. Specifically, we show that Assumptions TC1–TC5 are sufficient for Assumptions S1–S4, B1, B2, C6, C7, V1, and V2 of AC13. Then, under Assumptions B1 and B2, Assumptions S1–S3 of AC13 imply Assumptions A, B3, C1–C4, C8, and D1–D3 of AC12; see Lemma 9.1 in ACMLwp. Maintaining the same labels of AC13, below we rewrite the assumptions of AC13 before verifying them. Note that in our stylized threshold crossing model, β is scalar. Therefore we do not consider Assumptions S3* and V1* of AC13 which apply to the vector β case.

Assumption S1: $\forall \gamma_0 \in \Gamma$, $\{W_i : i \geq 1\}$ is an i.i.d. sequence and the constant q (that appears in Assumption S3 below) equals $2 + \delta$ for some $\delta > 0$.

Assumption S2: (i) For some function $\rho(w, \theta) \in \mathbb{R}$, $Q_n(\theta) = n^{-1} \sum_{i=1}^n \rho(W_i, \theta)$, where $\rho(w, \theta)$ is twice continuously differentiable in θ on an open set containing $\Theta^* \forall w \in \mathcal{W}$.

(ii) $\rho(w, \theta)$ does not depend on π when $\beta = 0 \forall w \in \mathcal{W}$.

(iii) $\forall \gamma_0 \in \Gamma$ with $\beta_0 = 0$, $E_{\gamma_0} \rho(W_i, \psi, \pi)$ is uniquely minimized by $\psi_0 \forall \pi \in \Pi$.

(iv) $\forall \gamma_0 \in \Gamma$ with $\beta_0 \neq 0$, $E_{\gamma_0} \rho(W_i, \theta)$ is uniquely minimized by θ_0 .

(v) $\Psi(\pi)$ is compact $\forall \pi \in \Pi$, and Π and Θ are compact.

(vi) $\forall \epsilon > 0$, $\exists \delta > 0$ such that $d_H(\Psi(\pi_1), \Psi(\pi_2)) < \epsilon \forall \pi_1, \pi_2 \in \Pi$ with $|\pi_1 - \pi_2| < \delta$, where $d_H(\cdot, \cdot)$ is the Hausdorff metric.

Verification of S2(i): By TC2(iii), $p_{y,d,z}(\theta)$ is twice continuously differentiable in θ . Then, since $p_{y,d,z}(\theta) \equiv p_{y,d,z}(\bar{h}(\theta))$ is twice continuously differentiable by Lemma A.1, so is $\rho(w, \theta) = -\sum_{y,d,z=0,1} \mathbf{1}_{ydz}(w) \log p_{y,d,z}(\theta)$. ■

Verification of S2(ii): It is easy to see from (2.5)–(2.6) of HM18 that, when $\beta = 0$, $p_{y,d,0}(\theta) = p_{y,d,1}(\theta)$ for all θ and (y, d) , which implies that $p_{y,d,0}(\bar{h}(\theta)) = p_{y,d,1}(\bar{h}(\theta))$ for all θ . Therefore

$$\begin{aligned} p_{11,1}(\theta) &= p_{11,0}(\theta) = \zeta_3, \\ p_{10,1}(\theta) &= p_{10,0}(\theta) = \zeta_2, \\ p_{01,1}(\theta) &= p_{01,0}(\theta) = \zeta_1 - \zeta_3, \end{aligned} \tag{A.5}$$

where the second equality in each equation is from (7.1)–(7.2) of HM18. Therefore $p_{y,d,z}(\theta)$ does not depend on π when $\beta = 0$, and hence $\rho(w, \theta) = -\sum_{y,d,z=0,1} \mathbf{1}_{ydz}(w) \log p_{y,d,z}(\theta)$ does not depend on π . ■

Verification of S2(iii): When $\beta_0 = 0$, for $\psi \neq \psi_0$ and for a given π ,

$$\begin{aligned} E_{\gamma_0} \rho(W_i, \psi, \pi) - E_{\gamma_0} \rho(W_i, \psi_0, \pi) &= -\sum_{y,d,z=0,1} p_{y,d,z}(\psi_0, \pi_0) \phi_{z,0} \log \frac{p_{y,d,z}(\psi, \pi)}{p_{y,d,z}(\psi_0, \pi)} \\ &\geq -\log \sum_{y,d,z=0,1} p_{y,d,z}(\psi_0, \pi_0) \phi_{z,0} \frac{p_{y,d,z}(\psi, \pi)}{p_{y,d,z}(\psi_0, \pi)} \\ &= -\log \sum_{y,d,z=0,1} p_{y,d,z}(\psi, \pi) \phi_{z,0} \\ &= 0, \end{aligned}$$

where the last equality holds since $\sum_{y,d} p_{y,d,1}(\theta) = \sum_{y,d} p_{y,d,0}(\theta) = 1$ and $\phi_{0,0} = 1 - \phi_{1,0}$, and the second-to-last equality holds since

$$p_{y,d,z}(\psi_0, \pi_0) = p_{y,d,z}(\psi_0, \pi) \equiv p_{y,d}^0 \tag{A.6}$$

when $\beta_0 = 0$, as in (A.5). Notationally, $p_{11} = \zeta_3$, $p_{10} = \zeta_2$, and $p_{01} = \zeta_1 - \zeta_3$. The Jensen's inequality is strict if there exist $(y, d, z) \in \{0, 1\}^3$ such that

$$\frac{p_{y,d,z}(\psi, \pi)}{p_{y,d,z}(\psi_0, \pi)} \neq 1.$$

Under TC2, this condition can be readily shown to hold by a slight modification of the identification proof of Theorem 4.1 in HV17, which is omitted here for brevity. ■

Verification of S2(iv): For $\theta \neq \theta_0$,

$$Q_0(\theta) - Q_0(\theta_0) = -\sum_{y,d,z=0,1} p_{y,d,z}(\theta_0) \phi_{z,0} \log \frac{p_{y,d,z}(\theta)}{p_{y,d,z}(\theta_0)}$$

$$\begin{aligned}
&> -\log \sum_{y,d,z=0,1} p_{yd,z}(\theta) \phi_{z,0} \\
&= 0,
\end{aligned}$$

where the Jensen's inequality is strict because there exist $(y, d, z) \in \{0, 1\}^3$ such that

$$\frac{p_{yd,z}(\theta)}{p_{yd,z}(\theta_0)} \neq 1$$

by Theorem 4.1 in HV17 under TC2. ■

Verification of S2(v): By TC3(i), Π is compact and the parameter space is the same before and after the transformation. Also, $\Theta = \bar{h}^{-1}(\Theta)$ is compact since Θ is compact and Assumption H(i) holds. For compactness of $\Psi(\pi)$, first note that, for a given $\pi \in \Pi$, $\bar{h}_{-\pi}(\cdot, \pi)$, which is $\bar{h}(\cdot, \pi)$ except the last element, is a homeomorphism. This is because $\Theta_{-\pi}$ is simply connected, $\bar{h}_{-\pi}(\cdot, \pi)$ is continuous, and $\Psi(\pi)$ is bounded since Θ is bounded. Then,

$$\Theta_{-\pi} = \Theta_{-\pi}(\pi) \equiv \bar{h}_{-\pi}(\Psi(\pi), \pi)$$

where the first equality is because the dependence parameter π does not restrict the space of the remaining elements of θ (or by TC3(i)), and thus $\Psi(\pi) = \bar{h}_{-\pi}^{-1}(\Theta_{-\pi}, \pi)$. Therefore $\Psi(\pi)$ is compact since $\Theta_{-\pi}$ is compact and $\bar{h}_{-\pi}(\cdot, \pi)$ is proper. ■

Verification of S2(vi): The space of $\psi = (\beta, \zeta)$ is continuous in π since $\Psi(\pi) = \bar{h}_{-\pi}^{-1}(\Theta_{-\pi}, \pi)$, where $\bar{h}_{-\pi}^{-1}(\Theta_{-\pi}, \pi)$ is continuous in π by (3.12) of HM18 and TC2(iii). ■

Let $\rho_\theta(w, \theta)$ and $\rho_{\theta\theta}(w, \theta)$ denote the first and second order partial derivatives of $\rho(w, \theta)$ w.r.t. θ , respectively. Also, let $\rho_\psi(w, \theta)$ and $\rho_{\psi\psi}(w, \theta)$ denote the first and second order partial derivatives of $\rho(w, \theta)$ w.r.t. ψ , respectively. Recall

$$B(\beta) \equiv \begin{bmatrix} I_{d_\psi} & 0_{d_\psi \times 1} \\ 0_{1 \times d_\psi} & \beta \end{bmatrix} \in \mathbb{R}^{d_\theta \times d_\theta}.$$

For $\beta \neq 0$, let

$$\begin{aligned}
B^{-1}(\beta)\rho_\theta(w, \theta) &\equiv \rho_\theta^\dagger(w, \theta), \\
B^{-1}(\beta)\rho_{\theta\theta}(w, \theta)B^{-1}(\beta) &\equiv \rho_{\theta\theta}^\dagger(w, \theta) + r(w, \theta),
\end{aligned} \tag{A.7}$$

where $\rho_{\theta\theta}^\dagger(w, \theta)$ is symmetric and $\rho_\theta^\dagger(w, \theta)$, $\rho_{\theta\theta}^\dagger(w, \theta)$, and $r(w, \theta)$ satisfy Assumption S3 below¹; see below for actual expressions of these terms. Next, define

$$V^\dagger(\theta_1, \theta_2; \gamma_0) \equiv Cov_{\gamma_0} \left(\rho_\theta^\dagger(W_i, \theta_1), \rho_\theta^\dagger(W_i, \theta_2) \right).$$

¹The remainder term $r(w, \theta)$ and related conditions in S3 are slightly more general than conditions on $\beta^{-1}\varepsilon(w, \theta)$ and related conditions in AC13.

Let $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ denote the maximum and minimum eigenvalues, respectively, of a square matrix A .

In this example of a threshold crossing model, define $D_{\theta}p_{y,d,z}^{\dagger}(\theta) \equiv B^{-1}(\beta)D_{\theta}p_{y,d,z}(\theta)$ so that

$$\begin{aligned}\rho_{\theta}(w, \theta) &= - \sum_{y,d,z=0,1} \mathbf{1}_{ydz}(w) \frac{1}{p_{y,d,z}(\theta)} D_{\theta}p_{y,d,z}(\theta), \\ \rho_{\theta\theta}(w, \theta) &= - \sum_{y,d,z=0,1} \mathbf{1}_{ydz}(w) \left[-\frac{1}{p_{y,d,z}(\theta)^2} D_{\theta}p_{y,d,z}(\theta) D_{\theta}p_{y,d,z}(\theta)' + \frac{1}{p_{y,d,z}(\theta)} D_{\theta\theta}p_{y,d,z}(\theta) \right], \\ \rho_{\theta}^{\dagger}(w, \theta) &= - \sum_{y,d,z=0,1} \mathbf{1}_{ydz}(w) \frac{1}{p_{y,d,z}(\theta)} D_{\theta}p_{y,d,z}^{\dagger}(\theta), \\ \rho_{\theta\theta}^{\dagger}(w, \theta) &= \rho_{\theta}^{\dagger}(w, \theta) \rho_{\theta}^{\dagger}(w, \theta)' = \sum_{y,d,z=0,1} \mathbf{1}_{ydz}(w) \frac{1}{p_{y,d,z}(\theta)^2} D_{\theta}p_{y,d,z}^{\dagger}(\theta) D_{\theta}p_{y,d,z}^{\dagger}(\theta)', \\ r(w, \theta) &= - \sum_{y,d,z=0,1} \mathbf{1}_{ydz}(w) \frac{1}{p_{y,d,z}(\theta)} B^{-1}(\beta) D_{\theta\theta}p_{y,d,z}(\theta) B^{-1}(\beta).\end{aligned}$$

Suppressing the argument (ζ_1, ζ_3, π) in h_3 and its derivatives, and suppressing the argument (ζ_1, ζ_2, π) in h_2 and its derivatives, note that from (7.1)–(7.2) of HM18,

$$D_{\theta}p_{11,0}(\theta) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad D_{\theta}p_{10,0}(\theta) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad D_{\theta}p_{01,0}(\theta) = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \quad D_{\theta}p_{00,0}(\theta) = \begin{bmatrix} 0 \\ -1 \\ -1 \\ 0 \\ 0 \end{bmatrix},$$

$$\begin{aligned}D_{\theta}p_{11,1}(\theta) &= \begin{bmatrix} C_2(h_3, \zeta_1 + \beta; \pi) \\ C_2(h_3, \zeta_1 + \beta; \pi) + C_1(h_3, \zeta_1 + \beta; \pi) h_{3,\zeta_1} \\ 0 \\ C_1(h_3, \zeta_1 + \beta; \pi) h_{3,\zeta_3} \\ C_{\pi}(h_3, \zeta_1 + \beta; \pi) + C_1(h_3, \zeta_1 + \beta; \pi) h_{3,\pi} \end{bmatrix} \\ &= \begin{bmatrix} C_2(h_3, \zeta_1 + \beta; \pi) \\ C_2(h_3, \zeta_1 + \beta; \pi) + C_1(h_3, \zeta_1 + \beta; \pi) h_{3,\zeta_1} \\ 0 \\ C_1(h_3, \zeta_1 + \beta; \pi) h_{3,\zeta_3} \\ \beta \{ C_{\pi 2}(h_3, \zeta_1 + \beta^{\dagger}; \pi) + C_{12}(h_3, \zeta_1 + \beta^{\dagger}; \pi) h_{3,\pi} \} \end{bmatrix}, \quad (\text{A.8})\end{aligned}$$

where $0 \leq |\beta^{\dagger}| \leq \beta$. The last equality is derived using a mean value expansion and the fact that $C_{\pi}(h_3, \zeta_1; \pi) + C_1(h_3, \zeta_1; \pi) h_{3,\pi} = 0$, obtained by differentiating $C(h_3, \zeta_1; \pi) = \zeta_3$

w.r.t. π . Furthermore,

$$\begin{aligned}
D_{\theta}p_{10,1}(\theta) &= \begin{bmatrix} -C_2(h_2, \zeta_1 + \beta; \pi) \\ h_{2,\zeta_1} - C_2(h_2, \zeta_1 + \beta; \pi) - C_1(h_2, \zeta_1 + \beta; \pi) h_{2,\zeta_1} \\ h_{2,\zeta_2} - C_1(h_2, \zeta_1 + \beta; \pi) h_{2,\zeta_2} \\ 0 \\ h_{2,\pi} - C_{\pi}(h_2, \zeta_1 + \beta; \pi) - C_1(h_2, \zeta_1 + \beta; \pi) h_{2,\pi} \end{bmatrix} \\
&= \begin{bmatrix} -C_2(h_2, \zeta_1 + \beta; \pi) \\ h_{2,\zeta_1} - C_2(h_2, \zeta_1 + \beta; \pi) - C_1(h_2, \zeta_1 + \beta; \pi) h_{2,\zeta_1} \\ h_{2,\zeta_2} - C_1(h_2, \zeta_1 + \beta; \pi) h_{2,\zeta_2} \\ 0 \\ -\beta \{C_{\pi_2}(h_2, \zeta_1 + \beta^{\dagger\dagger}; \pi) + C_{12}(h_2, \zeta_1 + \beta^{\dagger\dagger}; \pi) h_{2,\pi}\} \end{bmatrix}, \quad (\text{A.9})
\end{aligned}$$

where $0 \leq |\beta^{\dagger\dagger}| \leq \beta$ and the last equality is derived using a mean value expansion and the fact that $h_{2,\pi} - C_{\pi}(h_2, \zeta_1; \pi) - C_1(h_2, \zeta_1; \pi) h_{2,\pi} = 0$. Finally,

$$D_{\theta}p_{01,1}(\theta) = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} - D_{\theta}p_{11,1}(\theta), \quad D_{\theta}p_{00,1}(\theta) = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} - D_{\theta}p_{10,1}(\theta).$$

Also, note that for all (y, d) ,

$$D_{\theta\theta}p_{yd,0}(\theta) = 0 \quad (\text{A.10})$$

and

$$D_{\theta\theta}p_{01,1}(\theta) = -D_{\theta\theta}p_{11,1}(\theta), \quad D_{\theta\theta}p_{00,1}(\theta) = -D_{\theta\theta}p_{10,1}(\theta). \quad (\text{A.11})$$

Now, for $z = 0$,

$$D_{\theta}p_{yd,z}^{\dagger}(\theta) = D_{\theta}p_{yd,z}(\theta) \quad (\text{A.12})$$

and, for $z = 1$,

$$D_{\theta}p_{11,1}^{\dagger}(\theta) = \begin{bmatrix} C_2(h_3, \zeta_1 + \beta; \pi) \\ C_2(h_3, \zeta_1 + \beta; \pi) + C_1(h_3, \zeta_1 + \beta; \pi) h_{3,\zeta_1} \\ 0 \\ C_1(h_3, \zeta_1 + \beta; \pi) h_{3,\zeta_3} \\ C_{\pi_2}(h_3, \zeta_1 + \beta^{\dagger}; \pi) + C_{12}(h_3, \zeta_1 + \beta^{\dagger}; \pi) h_{3,\pi} \end{bmatrix}, \quad (\text{A.13})$$

$$D_{\theta} p_{10,1}^{\dagger}(\theta) = \begin{bmatrix} -C_2(h_2, \zeta_1 + \beta; \pi) \\ h_{2,\zeta_1} - C_2(h_2, \zeta_1 + \beta; \pi) - C_1(h_2, \zeta_1 + \beta; \pi) h_{2,\zeta_1} \\ h_{2,\zeta_2} - C_1(h_2, \zeta_1 + \beta; \pi) h_{2,\zeta_2} \\ 0 \\ -C_{\pi_2}(h_2, \zeta_1 + \beta^{\dagger\dagger}; \pi) - C_{12}(h_2, \zeta_1 + \beta^{\dagger\dagger}; \pi) h_{2,\pi} \end{bmatrix}, \quad (\text{A.14})$$

and expressions for the remaining two derivatives can be derived analogously.

Note that

$$\begin{aligned} \rho_{\psi}(w, \theta) &= - \sum_{y,d,z=0,1} \mathbf{1}_{ydz}(w) \frac{1}{p_{yd,z}(\theta)} D_{\psi} p_{yd,z}(\theta), \\ \rho_{\psi\psi}(w, \theta) &= - \sum_{y,d,z=0,1} \mathbf{1}_{ydz}(w) \left[-\frac{1}{p_{yd,z}(\theta)^2} D_{\psi} p_{yd,z}(\theta) D_{\psi} p_{yd,z}(\theta)' + \frac{1}{p_{yd,z}(\theta)} D_{\psi\psi} p_{yd,z}(\theta) \right], \end{aligned}$$

where, with $\psi = (\beta, \zeta) = (\beta, \zeta_1, \zeta_2, \zeta_3)$,

$$D_{\psi} p_{11,0}(\theta) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad D_{\psi} p_{10,0}(\theta) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad D_{\psi} p_{01,0}(\theta) = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \quad D_{\psi} p_{00,0}(\theta) = \begin{bmatrix} 0 \\ -1 \\ -1 \\ 0 \end{bmatrix},$$

$$D_{\psi} p_{11,1}(\theta) = \begin{bmatrix} C_2(h_3, \zeta_1 + \beta; \pi) \\ C_2(h_3, \zeta_1 + \beta; \pi) + C_1(h_3, \zeta_1 + \beta; \pi) h_{3,\zeta_1} \\ 0 \\ C_1(h_3, \zeta_1 + \beta; \pi) h_{3,\zeta_3} \end{bmatrix},$$

$$D_{\psi} p_{10,1}(\theta) = \begin{bmatrix} -C_2(h_2, \zeta_1 + \beta; \pi) \\ h_{2,\zeta_1} - C_2(h_2, \zeta_1 + \beta; \pi) - C_1(h_2, \zeta_1 + \beta; \pi) h_{2,\zeta_1} \\ h_{2,\zeta_2} - C_1(h_2, \zeta_1 + \beta; \pi) h_{2,\zeta_2} \\ 0 \end{bmatrix},$$

and

$$D_{\psi} p_{01,1}(\theta) = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - D_{\psi} p_{11,1}(\theta), \quad D_{\psi} p_{00,1}(\theta) = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \end{bmatrix} - D_{\psi} p_{10,1}(\theta).$$

Also, for all (y, d) and θ ,

$$D_{\psi\psi} p_{yd,0}(\theta) = 0 \quad (\text{A.15})$$

and

$$D_{\psi\psi} p_{01,1}(\theta) = -D_{\psi\psi} p_{11,1}(\theta), \quad D_{\psi\psi} p_{00,1}(\theta) = -D_{\psi\psi} p_{10,1}(\theta). \quad (\text{A.16})$$

Assumption S3: (i) (a) $E_{\gamma_0} r(W_i, \theta_0) = 0$; and (b) $\|E_{\gamma_0} r(W_i, \psi_0, \pi)\| \leq C |\pi - \pi_0| \forall \gamma_0 \in \Gamma$ with $0 < |\beta_0| < \delta$ for some $\delta > 0$.

(ii) (a) For all $\delta > 0$ and some function $M_1(w) : \mathcal{W} \rightarrow \mathbb{R}_+$, $\|\rho_{\psi\psi}(w, \theta_1) - \rho_{\psi\psi}(w, \theta_2)\| + \|\rho_{\theta\theta}^\dagger(w, \theta_1) - \rho_{\theta\theta}^\dagger(w, \theta_2)\| \leq M_1(w)\delta, \forall \theta_1, \theta_2 \in \Theta$ with $\|\theta_1 - \theta_2\| \leq \delta, \forall w \in \mathcal{W}$; and (b) for all $\delta > 0$ and some function $M_2(w) : \mathcal{W} \rightarrow \mathbb{R}_+$, $\|\rho_{\theta}^\dagger(w, \theta_1) - \rho_{\theta}^\dagger(w, \theta_2)\| + \|r(w, \theta_1) - r(w, \theta_2)\| \leq M_2(w)\delta, \forall \theta_1, \theta_2 \in \Theta$ with $\|\theta_1 - \theta_2\| \leq \delta, \forall w \in \mathcal{W}$.

(iii) $E_{\gamma_0} \sup_{\theta \in \Theta} \left\{ |\rho(W_i, \theta)|^{1+\delta} + \|\rho_{\psi\psi}(W_i, \theta)\|^{1+\delta} + \|\rho_{\theta\theta}^\dagger(W_i, \theta)\|^{1+\delta} + M_1(W_i) + \|\rho_{\theta}^\dagger(W_i, \theta)\|^q + \|r(W_i, \theta)\|^q + M_2(W_i)^q \right\} \leq C$ for some $\delta > 0 \forall \gamma_0 \in \Gamma$, where q is as in Assumption S1.

(iv) (a) $\lambda_{\min}(E_{\gamma_0} \rho_{\psi\psi}(W_i, \psi_0, \pi)) > 0 \forall \pi \in \Pi$ when $\beta_0 = 0$; and (b) $E_{\gamma_0} \rho_{\theta\theta}^\dagger(W_i, \theta_0)$ is positive definite $\forall \gamma_0 \in \Gamma$.

(v) $V^\dagger(\theta_0, \theta_0; \gamma_0)$ is positive definite $\forall \gamma_0 \in \Gamma$.

Verification of S3(i)(a): Note that

$$E_{\gamma_0} r(W_i, \theta_0) = - \sum_{y,d,z=0,1} \phi_{z,0} B^{-1}(\beta_0) D_{\theta\theta} p_{y,d,z}(\theta_0) B^{-1}(\beta_0) = 0$$

by (A.10) and (A.11) since $\beta_0 \neq 0$. ■

Verification of S3(i)(b): Using (A.10) and (A.11),

$$\begin{aligned} E_{\gamma_0} r(W_i, \psi_0, \pi) &= \sum_{y,d=0,1} p_{y,d,1}(\theta_0) \phi_{1,0} B^{-1}(\beta_0) \frac{D_{\theta\theta} p_{y,d,1}(\psi_0, \pi)}{p_{y,d,1}(\psi_0, \pi)} B^{-1}(\beta_0) \\ &= \phi_{1,0} B^{-1}(\beta_0) \left[\frac{p_{11,1}(\theta_0)}{p_{11,1}(\psi_0, \pi)} D_{\theta\theta} p_{11,1}(\psi_0, \pi) + \frac{p_{01,1}(\theta_0)}{p_{01,1}(\psi_0, \pi)} D_{\theta\theta} p_{01,1}(\psi_0, \pi) \right. \\ &\quad \left. + \frac{p_{10,1}(\theta_0)}{p_{10,1}(\psi_0, \pi)} D_{\theta\theta} p_{10,1}(\psi_0, \pi) + \frac{p_{00,1}(\theta_0)}{p_{00,1}(\psi_0, \pi)} D_{\theta\theta} p_{00,1}(\psi_0, \pi) \right] B^{-1}(\beta_0) \\ &= \phi_{1,0} B^{-1}(\beta_0) \left[\left(\frac{p_{11,1}(\theta_0)}{p_{11,1}(\psi_0, \pi)} - \frac{p_{01,1}(\theta_0)}{p_{01,1}(\psi_0, \pi)} \right) D_{\theta\theta} p_{11,1}(\psi_0, \pi) \right. \\ &\quad \left. + \left(\frac{p_{10,1}(\theta_0)}{p_{10,1}(\psi_0, \pi)} - \frac{p_{00,1}(\theta_0)}{p_{00,1}(\psi_0, \pi)} \right) D_{\theta\theta} p_{10,1}(\psi_0, \pi) \right] B^{-1}(\beta_0) \\ &= \phi_{1,0} B^{-1}(\beta_0) \left[\left(\frac{(\zeta_{10} + \beta_0)(p_{11,1}(\theta_0) - p_{11,1}(\psi_0, \pi))}{p_{11,1}(\psi_0, \pi)(\zeta_{10} + \beta_0 - p_{11,1}(\psi_0, \pi))} \right) D_{\theta\theta} p_{11,1}(\psi_0, \pi) \right. \\ &\quad \left. + \left(\frac{(1 - \zeta_{10} - \beta_0)(p_{10,1}(\theta_0) - p_{10,1}(\psi_0, \pi))}{p_{10,1}(\psi_0, \pi)(1 - \zeta_{10} - \beta_0 - p_{10,1}(\psi_0, \pi))} \right) D_{\theta\theta} p_{10,1}(\psi_0, \pi) \right] B^{-1}(\beta_0) \end{aligned} \tag{A.17}$$

where the last equality uses $p_{01,1}(\theta) = \zeta_1 + \beta - p_{11,1}(\theta)$ and $p_{00,1}(\theta) = 1 - \zeta_1 - \beta - p_{10,1}(\theta)$. Apply the mean value theorem to $p_{11,1}(\theta_0) - p_{11,1}(\psi_0, \pi)$ w.r.t. π :

$$p_{11,1}(\psi_0, \pi_0) - p_{11,1}(\psi_0, \pi) = \frac{\partial p_{11,1}(\psi_0, \pi^\dagger)}{\partial \pi} (\pi_0 - \pi)$$

$$= \frac{\partial^2 p_{11,1}(\beta^\dagger, \zeta_0, \pi^\dagger)}{\partial \pi \partial \beta} (\pi_0 - \pi) \beta_0, \quad (\text{A.18})$$

where π^\dagger is between π_0 and π and $0 \leq |\beta^\dagger| \leq |\beta_0|$. The second equality holds by another mean value expansion of $\frac{\partial p_{11,1}(\psi_0, \pi^\dagger)}{\partial \pi}$ w.r.t. β_0 around $\beta_0 = 0$ and the fact that $\left. \frac{\partial p_{11,1}(\beta, \zeta_0, \pi^\dagger)}{\partial \pi} \right|_{\beta=0} = 0$ since

$$C_\pi(h_3(\pi), \zeta_1; \pi) + C_1(h_3(\pi), \zeta_1; \pi) h_{3,\pi}(\pi) = 0$$

for all (ζ_1, ζ_3, π) . Similarly, using mean value expansions,

$$p_{10,1}(\psi_0, \pi_0) - p_{10,1}(\psi_0, \pi) = \frac{\partial^2 p_{10,1}(\beta^{\dagger\dagger}, \zeta_0, \pi^{\dagger\dagger})}{\partial \pi \partial \beta} (\pi_0 - \pi) \beta_0 \quad (\text{A.19})$$

for some $\pi^{\dagger\dagger}$ between π_0 and π and $0 \leq |\beta^{\dagger\dagger}| \leq |\beta_0|$. Therefore, combining (A.17)–(A.19),

$$\begin{aligned} \|E_{\gamma_0} r(W_i, \psi_0, \pi)\| &\leq |c_1| \|B^{-1}(\beta_0) \beta_0 D_{\theta\theta} p_{11,1}(\psi_0, \pi) B^{-1}(\beta_0)\| |\pi_0 - \pi| \\ &\quad + |c_2| \|B^{-1}(\beta_0) \beta_0 D_{\theta\theta} p_{10,1}(\psi_0, \pi) B^{-1}(\beta_0)\| |\pi_0 - \pi| \end{aligned}$$

where c_1 and c_2 are collections of all other terms, whose norms are bounded by (7.1)–(7.2) of HM18 and Lemma A.1. Also $\|B^{-1}(\beta_0) \beta_0\|$ is bounded for $0 < |\beta_0| < \delta$. Note that $\|D_{\theta\theta} p_{11,1}(\psi_0, \pi) B^{-1}(\beta_0)\|$ and $\|D_{\theta\theta} p_{10,1}(\psi_0, \pi) B^{-1}(\beta_0)\|$ can be shown to be bounded for $0 < |\beta_0| < \delta$ by differentiating (A.13) and (A.14) w.r.t. θ , respectively, and applying Lemma A.1. ■

Verification of S3(ii)(a): Generically, for $A = aa'$ where $a = (a_1, \dots, a_p) \in \mathbb{R}^{d_a}$ and a_1, \dots, a_p are vectors,

$$\|A\| \leq \sum_{j=1}^p \|a_j\|^2,$$

and for $A^* = a^* a^{* \prime}$

$$\begin{aligned} \|A - A^*\| &\leq \|a(a - a^*)'\| + \|(a - a^*)a^{* \prime}\| \leq (\|a\| + \|a^*\|) \|a - a^*\| \\ &\leq \sum_{j=1}^p (\|a_j\| + \|a_j^*\|) \sum_{j=1}^p \|a_j - a_j^*\|. \end{aligned}$$

Applying this result to the last inequality below,

$$\begin{aligned} &\|\rho_{\psi\psi}(w, \theta_1) - \rho_{\psi\psi}(w, \theta_2)\| \\ &\leq \sum_{y,d,z=0,1} \left\| \frac{D_\psi p_{yd,z}(\theta_1) D_\psi p_{yd,z}(\theta_1)'}{p_{yd,z}(\theta_1)^2} - \frac{D_\psi p_{yd,z}(\theta_2) D_\psi p_{yd,z}(\theta_2)'}{p_{yd,z}(\theta_2)^2} \right\| \end{aligned}$$

$$\begin{aligned}
& + \sum_{y,d,z=0,1} \left\| \frac{D_{\psi\psi} p_{y,d,z}(\theta_1)}{p_{y,d,z}(\theta_1)} - \frac{D_{\psi\psi} p_{y,d,z}(\theta_2)}{p_{y,d,z}(\theta_2)} \right\| \\
& \leq \sum_{y,d,z=0,1} \left(\left\| \frac{D_{\psi} p_{y,d,z}(\theta_1)}{p_{y,d,z}(\theta_1)} \right\| + \left\| \frac{D_{\psi} p_{y,d,z}(\theta_2)}{p_{y,d,z}(\theta_2)} \right\| \right) \left\| \frac{D_{\psi} p_{y,d,z}(\theta_1)}{p_{y,d,z}(\theta_1)} - \frac{D_{\psi} p_{y,d,z}(\theta_2)}{p_{y,d,z}(\theta_2)} \right\| \\
& + \sum_{y,d,z=0,1} \left\| \frac{D_{\psi\psi} p_{y,d,z}(\theta_1)}{p_{y,d,z}(\theta_1)} - \frac{D_{\psi\psi} p_{y,d,z}(\theta_2)}{p_{y,d,z}(\theta_2)} \right\| \\
& \leq \sum_{y,d,z=0,1} \sum_{j=1}^{d_\psi} \left(\left| \frac{D_{\psi_j} p_{y,d,z}(\theta_1)}{p_{y,d,z}(\theta_1)} \right| + \left| \frac{D_{\psi_j} p_{y,d,z}(\theta_2)}{p_{y,d,z}(\theta_2)} \right| \right) \sum_{j=1}^{d_\psi} \left| \frac{D_{\psi_j} p_{y,d,z}(\theta_1)}{p_{y,d,z}(\theta_1)} - \frac{D_{\psi_j} p_{y,d,z}(\theta_2)}{p_{y,d,z}(\theta_2)} \right| \\
& + \sum_{y,d,z=0,1} \sum_{j,k=1}^{d_\psi} \left| \frac{D_{\psi_j \psi_k} p_{y,d,z}(\theta_1)}{p_{y,d,z}(\theta_1)} - \frac{D_{\psi_j \psi_k} p_{y,d,z}(\theta_2)}{p_{y,d,z}(\theta_2)} \right|,
\end{aligned}$$

where $|\mathbf{1}_{ydz}(w)| \leq 1$ is used in the first inequality. Applying the mean value theorem to the differential terms,

$$\begin{aligned}
\left| \frac{D_{\psi_j} p_{y,d,z}(\theta_1)}{p_{y,d,z}(\theta_1)} - \frac{D_{\psi_j} p_{y,d,z}(\theta_2)}{p_{y,d,z}(\theta_2)} \right| & \leq \left\| D_\theta \left\{ \frac{D_{\psi_j} p_{y,d,z}(\theta^\dagger)}{p_{y,d,z}(\theta^\dagger)} \right\} \right\| \|\theta_1 - \theta_2\|, \\
\left| \frac{D_{\psi_j \psi_k} p_{y,d,z}(\theta_1)}{p_{y,d,z}(\theta_1)} - \frac{D_{\psi_j \psi_k} p_{y,d,z}(\theta_2)}{p_{y,d,z}(\theta_2)} \right| & \leq \left\| D_\theta \left\{ \frac{D_{\psi_j \psi_k} p_{y,d,z}(\theta^{\dagger\dagger})}{p_{y,d,z}(\theta^{\dagger\dagger})} \right\} \right\| \|\theta_1 - \theta_2\|,
\end{aligned}$$

where θ^\dagger and $\theta^{\dagger\dagger}$ lie between θ_1 and θ_2 (element-wise). By Lemma A.1, $\sup_\theta \left| \frac{D_{\psi_j} p_{y,d,z}(\theta)}{p_{y,d,z}(\theta)} \right| < c_1$, $\sup_\theta \left| D_{\theta_k} \left\{ \frac{D_{\psi_j} p_{y,d,z}(\theta)}{p_{y,d,z}(\theta)} \right\} \right| < c_2$ and $\sup_\theta \left| D_{\theta_l} \left\{ \frac{D_{\psi_j \psi_k} p_{y,d,z}(\theta)}{p_{y,d,z}(\theta)} \right\} \right| < c_3$ for some positive constants c_1 , c_2 and c_3 , and therefore combining the inequalities,

$$\begin{aligned}
\|\rho_{\psi\psi}(w, \theta_1) - \rho_{\psi\psi}(w, \theta_2)\| & \leq \sum_{y,d,z=0,1} \sum_{j=1}^{d_\psi} 2c_1 \sum_{j=1}^{d_\psi} \sum_{k=1}^{d_\theta} c_2 \|\theta_1 - \theta_2\| \\
& + \sum_{y,d,z=0,1} \sum_{j,k=1}^{d_\psi} \sum_{l=1}^{d_\theta} c_3 \|\theta_1 - \theta_2\|. \tag{A.20}
\end{aligned}$$

Similarly,

$$\begin{aligned}
& \left\| \rho_{\theta\theta}^\dagger(w, \theta_1) - \rho_{\theta\theta}^\dagger(w, \theta_2) \right\| \\
& \leq \sum_{y,d,z=0,1} \left\| \frac{D_\theta p_{y,d,z}^\dagger(\theta_1) D_\theta p_{y,d,z}^\dagger(\theta_1)' - D_\theta p_{y,d,z}^\dagger(\theta_2) D_\theta p_{y,d,z}^\dagger(\theta_2)'}{p_{y,d,z}(\theta_1)^2} - \frac{D_\theta p_{y,d,z}^\dagger(\theta_2) D_\theta p_{y,d,z}^\dagger(\theta_2)'}{p_{y,d,z}(\theta_2)^2} \right\| \\
& \leq \sum_{y,d,z=0,1} \left(\left\| \frac{D_\theta p_{y,d,z}^\dagger(\theta_1)}{p_{y,d,z}(\theta_1)} \right\| + \left\| \frac{D_\theta p_{y,d,z}^\dagger(\theta_2)}{p_{y,d,z}(\theta_2)} \right\| \right) \left\| \frac{D_\theta p_{y,d,z}^\dagger(\theta_1)}{p_{y,d,z}(\theta_1)} - \frac{D_\theta p_{y,d,z}^\dagger(\theta_2)}{p_{y,d,z}(\theta_2)} \right\| \\
& \leq \sum_{y,d,z=0,1} \sum_{j=1}^{d_\theta} \left(\left| \frac{D_{\theta_j} p_{y,d,z}^\dagger(\theta_1)}{p_{y,d,z}(\theta_1)} \right| + \left| \frac{D_{\theta_j} p_{y,d,z}^\dagger(\theta_2)}{p_{y,d,z}(\theta_2)} \right| \right) \sum_{j=1}^{d_\theta} \left| \frac{D_{\theta_j} p_{y,d,z}^\dagger(\theta_1)}{p_{y,d,z}(\theta_1)} - \frac{D_{\theta_j} p_{y,d,z}^\dagger(\theta_2)}{p_{y,d,z}(\theta_2)} \right|
\end{aligned}$$

and by Lemma A.1, $\sup_{\theta} \left| \frac{D_{\theta_j} p_{y d, z}^{\dagger}(\theta)}{p_{y d, z}(\theta)} \right| < c_4$ and $\sup_{\theta} \left| D_{\theta_k} \left\{ \frac{D_{\theta_j} p_{y d, z}^{\dagger}(\theta)}{p_{y d, z}(\theta)} \right\} \right| < c_5$ for some positive constants c_4 and c_5 , and therefore by applying the mean value theorem as above,

$$\left\| \rho_{\theta\theta}^{\dagger}(w, \theta_1) - \rho_{\theta\theta}^{\dagger}(w, \theta_2) \right\| \leq \sum_{y, d, z=0,1} \sum_{j=1}^{d_{\theta}} 2c_4 \sum_{j, k=1}^{d_{\theta}} c_5 \|\theta_1 - \theta_2\|. \quad (\text{A.21})$$

By combining (A.20) and (A.21), we have the desired result. ■

Verification of S3(ii)(b): For bounding $\|r(w, \theta_1) - r(w, \theta_2)\|$, the proof is very similar to the one above with $\left\| \rho_{\theta\theta}^{\dagger}(w, \theta_1) - \rho_{\theta\theta}^{\dagger}(w, \theta_2) \right\|$. Bounding $\left\| \rho_{\theta}^{\dagger}(w, \theta_1) - \rho_{\theta}^{\dagger}(w, \theta_2) \right\|$ can also be done analogously. ■

Verification of S3(iii): First, $M_1(w)$ is finite and does not depend on w , as can be seen from the verification of S3(ii)(a). Now, since $|\mathbf{1}_{ydz}(w)| \leq 1$

$$\begin{aligned} E_{\gamma_0} \sup_{\theta \in \Theta} |\rho(W_i, \theta)|^{1+\delta} &\leq E_{\gamma_0} \left(\sum_{y, d, z=0,1} \sup_{\theta \in \Theta} |\mathbf{1}_{ydz}(w) \cdot \log p_{y d, z}(\theta)| \right)^{1+\delta} \\ &\leq \left(\sum_{y, d, z=0,1} \sup_{\theta \in \Theta} |\log p_{y d, z}(\theta)| \right)^{1+\delta}, \end{aligned}$$

which is bounded since $p_{y d, z}(\theta)$ is bounded away from zero for any $\theta \in \Theta$ and $(y, d, z) \in \{0, 1\}$ by Lemma A.1. Next,

$$\begin{aligned} &E_{\gamma_0} \sup_{\theta \in \Theta} \|\rho_{\psi\psi}(W_i, \theta)\|^{1+\delta} \\ &\leq E_{\gamma_0} \left(\sum_{y, d, z=0,1} \sup_{\theta \in \Theta} \left\| \mathbf{1}_{ydz}(w) \left[-\frac{1}{p_{y d, z}(\theta)^2} D_{\psi} p_{y d, z}(\theta) D_{\psi} p_{y d, z}(\theta)' + \frac{1}{p_{y d, z}(\theta)} D_{\psi\psi} p_{y d, z}(\theta) \right] \right\| \right)^{1+\delta} \\ &\leq \left(\sum_{y, d, z=0,1} C \sup_{\theta \in \Theta} \{ \|D_{\psi} p_{y d, z}(\theta) D_{\psi} p_{y d, z}(\theta)'\| + \|D_{\psi\psi} p_{y d, z}(\theta)\| \} \right)^{1+\delta} \end{aligned}$$

by Lemma A.1, where $\|D_{\psi} p_{y d, z}(\theta) D_{\psi} p_{y d, z}(\theta)'\| \leq \sum_{j=1}^{d_{\psi}} \|D_{\psi_j} p_{y d, z}(\theta)\|^2$, which is bounded by Lemma A.1, and similarly for $\|D_{\psi\psi} p_{y d, z}(\theta)\|$. Similar arguments to those used in the verification of S3(i)(b) and S3(ii)(a) provide the desired result for the remaining four terms in the assumption. ■

Verification of S3(iv)(a): Note that, when $\beta_0 = 0$,

$$E_{\gamma_0} \rho_{\psi\psi}(W_i, \psi_0, \pi) = \sum_{y, d, z=0,1} p_{y d, z}(\theta_0) \phi_{z,0} \left[\frac{D_{\psi} p_{y d, z}(\psi_0, \pi) D_{\psi} p_{y d, z}(\psi_0, \pi)'}{p_{y d, z}(\psi_0, \pi)^2} - \frac{D_{\psi\psi} p_{y d, z}(\psi_0, \pi)}{p_{y d, z}(\psi_0, \pi)} \right]$$

$$\begin{aligned}
&= \sum_{y,d,z=0,1} \phi_{z,0} \left[\frac{D_\psi p_{yd,z}(\psi_0, \pi) D_\psi p_{yd,z}(\psi_0, \pi)'}{p_{yd}^0} - D_{\psi\psi} p_{yd,z}(\psi_0, \pi) \right] \\
&= \sum_{y,d,z=0,1} \phi_{z,0} \frac{D_\psi p_{yd,z}(\psi_0, \pi) D_\psi p_{yd,z}(\psi_0, \pi)'}{p_{yd}^0}
\end{aligned}$$

where the second equality is by (A.6), and the third equality is by (A.15) and (A.16). Let $M_{yd,z} \equiv D_\psi p_{yd,z}(\psi_0, \pi) D_\psi p_{yd,z}(\psi_0, \pi)'$ and $\tilde{M}_{yd,z} \equiv M_{yd,z}/p_{yd}^0$ so that

$$E_{\gamma_0} \rho_{\psi\psi}(W_i, \psi_0, \pi) = \phi_{1,0} \sum_{y,d=0,1} \tilde{M}_{yd,1} + \phi_{0,0} \sum_{y,d=0,1} \tilde{M}_{yd,0}. \quad (\text{A.22})$$

Let $h_3(\pi) \equiv h_3(\zeta_{10}, \zeta_{30}; \pi)$ and $h_2(\pi) \equiv h_2(\zeta_{10}, \zeta_{20}; \pi)$. Note that when $\beta_0 = 0$, the $D_\psi p_{yd,z}(\psi_0, \pi)$ terms can be expressed as

$$D_\psi p_{11,0} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad D_\psi p_{10,0} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad D_\psi p_{01,0} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \quad D_\psi p_{00,0} = \begin{bmatrix} 0 \\ -1 \\ -1 \\ 0 \end{bmatrix},$$

$$D_\psi p_{11,1} = \begin{bmatrix} C_2(h_3(\pi), \zeta_1; \pi) \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

$$D_\psi p_{10,1} = \begin{bmatrix} -C_2(h_2(\pi), \zeta_1; \pi) \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

and

$$D_\psi p_{01,1} = \begin{bmatrix} 1 - C_2(h_3(\pi), \zeta_1; \pi) \\ 1 \\ 0 \\ -1 \end{bmatrix}, \quad D_\psi p_{00,1} = \begin{bmatrix} -1 + C_2(h_2(\pi), \zeta_1; \pi) \\ -1 \\ -1 \\ 0 \end{bmatrix},$$

where, in $D_\psi p_{11,1}$ and $D_\psi p_{10,1}$,

$$C_2(h_3, \zeta_1; \pi) + C_1(h_3, \zeta_1; \pi) h_{3,\zeta_1} = 0, \quad (\text{A.23})$$

$$C_1(h_3, \zeta_1; \pi) h_{3,\zeta_3} = 1, \quad (\text{A.24})$$

$$h_{2,\zeta_1} - C_2(h_2, \zeta_1; \pi) - C_1(h_2, \zeta_1; \pi) h_{2,\zeta_1} = 0, \quad (\text{A.25})$$

$$h_{2,\zeta_2} - C_1(h_2, \zeta_1; \pi) h_{2,\zeta_2} = 1, \quad (\text{A.26})$$

by differentiating the objects in (7.1)–(7.2) of HM18 w.r.t. ζ_1, ζ_2 and ζ_3 and (A.5). Let $c \equiv C_2(h_3(\pi), \zeta_{10}; \pi)$ and $\tilde{c} \equiv C_2(h_2(\pi), \zeta_{10}; \pi)$ for notational simplicity. Then,

$$\begin{aligned}
M_{11,1} &= \begin{bmatrix} c^2 & 0 & 0 & c \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ c & 0 & 0 & 1 \end{bmatrix}, \quad M_{10,1} = \begin{bmatrix} \tilde{c}^2 & 0 & -\tilde{c} & 0 \\ 0 & 0 & 0 & 0 \\ -\tilde{c} & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\
M_{01,1} &= \begin{bmatrix} (1-c)^2 & 1-c & 0 & c-1 \\ 1-c & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ c-1 & -1 & 0 & 1 \end{bmatrix}, \quad M_{00,1} = \begin{bmatrix} (1-\tilde{c})^2 & 1-\tilde{c} & 1-\tilde{c} & 0 \\ 1-\tilde{c} & 1 & 1 & 0 \\ 1-\tilde{c} & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\
M_{11,0} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad M_{10,0} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\
M_{01,0} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}, \quad M_{00,0} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.
\end{aligned}$$

By Weyl (1912),

$$\lambda_{\min}(A+B) \geq \lambda_{\min}(A) + \lambda_{\min}(B) \quad (\text{A.27})$$

for symmetric matrices A and B . Thus, for (A.22),

$$\lambda_{\min}(E_{\gamma_0} \rho_{\psi\psi}(W_i, \psi_0, \pi)) \geq \lambda_{\min}\left(\phi_{1,0} \sum_{y,d=0,1} \tilde{M}_{yd,1}\right) + \lambda_{\min}\left(\phi_{0,0} \sum_{y,d=0,1} \tilde{M}_{yd,0}\right).$$

The second term on the right hand side satisfies $\lambda_{\min}\left(\phi_{0,0} \sum_{y,d=0,1} \tilde{M}_{yd,0}\right) \geq \phi_{0,0} \sum_{y,d=0,1} \lambda_{\min}\left(\tilde{M}_{yd,0}\right) = 0$ by (A.27), the above expressions for the $M_{yd,0}$'s and since $\lambda_{\min}\left(\tilde{M}_{yd,0}\right) = \lambda_{\min}(M_{yd,0}) = 0$ because $p_{yd}^0 > 0$ for all (y, d) by Lemma A.1(v). The first term on the right hand side satisfies $\lambda_{\min}\left(\phi_{1,0} \sum_{y,d=0,1} \tilde{M}_{yd,1}\right) \geq \phi_{1,0} \lambda_{\min}\left(\left\{\tilde{M}_{11,1} + \tilde{M}_{01,1} + \tilde{M}_{00,1}\right\}\right)$ by (A.27) and since $\lambda_{\min}\left(\tilde{M}_{10,1}\right) = \lambda_{\min}(M_{10,1}) = 0$. Now we prove $\lambda_{\min}\left(\tilde{M}_{11,1} + \tilde{M}_{01,1} + \tilde{M}_{00,1}\right) > 0$, which then implies that $\lambda_{\min}(E_{\gamma_0} \rho_{\psi\psi}(W_i, \psi_0, \pi)) > 0$ as desired since $\phi_{1,0} > 0$ by TC5(ii). Under TC5(i) and by Lemma A.1(v), let $a \equiv p_{11}^0/p_{01}^0$ and $b \equiv p_{11}^0/p_{00}^0$ for simplicity. Then,

$\tilde{M}_{11,1} + \tilde{M}_{01,1} + \tilde{M}_{00,1} = (M_{11,1} + aM_{01,1} + bM_{00,1})/p_{11}^0$ and

$$M \equiv M_{11,1} + aM_{01,1} + bM_{00,1} \\ = \begin{bmatrix} a(1-c)^2 + b(1-\tilde{c})^2 + c^2 & a(1-c) + b(1-\tilde{c}) & b(1-\tilde{c}) & -a(1-c) + c \\ a(1-c) + b(1-\tilde{c}) & a+b & b & -a \\ b(1-\tilde{c}) & b & b & 0 \\ -a(1-c) + c & -a & 0 & a+1 \end{bmatrix}.$$

Then one can easily show the following: For the k -th leading principal minor $|M_k|$ and determinant $|M|$ of M ,

$$|M_1| = a(1-c)^2 + b(1-\tilde{c})^2 + c^2 > 0,$$

$$|M_2| = ab[(1-c) + (1-\tilde{c})]^2 + (a+b)c^2 > 0,$$

$$|M_3| = abc\tilde{c}^2 > 0,$$

$$|M| = ab[a(2c-1)^2 + b(\tilde{c}-1)^2] > 0,$$

and therefore M is positive definite and so is M/p_{11}^0 , i.e., $\lambda_{\min}(\tilde{M}_{11,1} + \tilde{M}_{01,1} + \tilde{M}_{00,1}) > 0$.

■

Verification of S3(iv)(b): We divide this proof into two cases: (i) $\beta_0 \neq 0$ and (ii) $\beta_0 = 0$.

Case (i): Note that by S3(i)(a),

$$E_{\gamma_0} B^{-1}(\beta_0) \rho_{\theta\theta}(W_i, \theta_0) B^{-1}(\beta_0) = E_{\gamma_0} \rho_{\theta\theta}^\dagger(W_i, \theta_0).$$

First, note that $E_{\gamma_0} \rho_{\theta\theta}(w, \theta_0)$ is positive definite by the information matrix equality with the fact that the information matrix is nonsingular by the identification result of HV17. Thus, for a nonzero vector $a \in \mathbb{R}^{d_\theta}$, $a' E_{\gamma_0} \rho_{\theta\theta}(w, \theta_0) a > 0$, which implies that, for a nonzero vector $\tilde{a} \in \mathbb{R}^{d_\theta}$, $\tilde{a}' E_{\gamma_0} \rho_{\theta\theta}^\dagger(w, \theta_0) \tilde{a} = \tilde{a}' B^{-1}(\beta_0) E_{\gamma_0} \rho_{\theta\theta}(w, \theta_0) B^{-1}(\beta_0) \tilde{a} > 0$. Therefore $E_{\gamma_0} \rho_{\theta\theta}^\dagger(w, \theta_0)$ is positive definite.

Case (ii): First note that by (A.12)–(A.14) and (A.23)–(A.26), we can express $D_\theta p_{y_{d,z}}^\dagger(\psi_0, \pi)$'s as follows when $\beta_0 = 0$,

$$D_\theta p_{11,0}^\dagger = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad D_\theta p_{10,0}^\dagger = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad D_\theta p_{01,0}^\dagger = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \quad D_\theta p_{00,0}^\dagger = \begin{bmatrix} 0 \\ -1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad (\text{A.28})$$

$$D_{\theta}p_{11,1}^{\dagger} = \begin{bmatrix} C_2(h_3(\pi), \zeta_{10}; \pi) \\ 0 \\ 0 \\ 1 \\ C_{\pi_2}(h_3(\pi), \zeta_{10}; \pi) + C_{12}(h_3(\pi), \zeta_{10}; \pi) h_{3,\pi}(\zeta_{10}, \zeta_{30}, \pi) \end{bmatrix}, \quad (\text{A.29})$$

$$D_{\theta}p_{10,1}^{\dagger} = \begin{bmatrix} -C_2(h_2(\pi), \zeta_{10}; \pi) \\ 0 \\ 1 \\ 0 \\ -C_{\pi_2}(h_2(\pi), \zeta_{10}; \pi) - C_{12}(h_2(\pi), \zeta_{10}; \pi) h_{2,\pi}(\zeta_{10}, \zeta_{20}, \pi) \end{bmatrix}, \quad (\text{A.30})$$

and

$$D_{\theta}p_{01,1}^{\dagger} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} - D_{\theta}p_{11,1}^{\dagger}, \quad D_{\theta}p_{00,1}^{\dagger} = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} - D_{\theta}p_{10,1}^{\dagger}. \quad (\text{A.31})$$

The remaining arguments are similar to those used to verify S3(iv)(a): Let $M_{y d, z}^{\dagger} \equiv D_{\theta}p_{y d, z}^{\dagger}(\theta_0) \times D_{\theta}p_{y d, z}^{\dagger}(\theta_0)'$ and $\tilde{M}_{y d, z}^{\dagger} \equiv M_{y d, z}^{\dagger}/p_{y d}^0$. Then,

$$E_{\gamma_0} \rho_{\theta\theta}^{\dagger}(W_i, \theta_0) = E_{\gamma_0} \rho_{\theta}^{\dagger}(W_i, \theta_0) \rho_{\theta}^{\dagger}(W_i, \theta_0)' = \phi_{1,0} \sum_{y,d=0,1} \tilde{M}_{y d, 1}^{\dagger} + \phi_{0,0} \sum_{y,d=0,1} \tilde{M}_{y d, 0}^{\dagger}. \quad (\text{A.32})$$

For notational simplicity, let $c \equiv C_2(h_3(\pi_0), \zeta_{10}; \pi_0)$ and $\tilde{c} \equiv C_2(h_2(\pi_0), \zeta_{10}; \pi_0)$. Also let $d \equiv C_{\pi_2}(h_3(\pi_0), \zeta_{10}; \pi_0) + C_{12}(h_3(\pi_0), \zeta_{10}; \pi_0) h_{3,\pi}(\zeta_{10}, \zeta_{30}, \pi_0)$ and $\tilde{d} \equiv C_{\pi_2}(h_2(\pi_0), \zeta_{10}; \pi_0) + C_{12}(h_2(\pi_0), \zeta_{10}; \pi_0) h_{2,\pi}(\zeta_{10}, \zeta_{20}, \pi_0)$. Therefore,

$$M_{11,1}^{\dagger} = \begin{bmatrix} c^2 & 0 & 0 & c & cd \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ c & 0 & 0 & 1 & d \\ cd & 0 & 0 & d & d^2 \end{bmatrix}, \quad M_{01,1}^{\dagger} = \begin{bmatrix} (1-c)^2 & 1-c & 0 & c-1 & (c-1)d \\ 1-c & 1 & 0 & -1 & -d \\ 0 & 0 & 0 & 0 & 0 \\ c-1 & -1 & 0 & 1 & d \\ (c-1)d & -d & 0 & d & d^2 \end{bmatrix},$$

$$\begin{aligned}
M_{10,1}^\dagger &= \begin{bmatrix} \tilde{c}^2 & 0 & -\tilde{c} & 0 & \tilde{c}\tilde{d} \\ 0 & 0 & 0 & 0 & 0 \\ -\tilde{c} & 0 & 1 & 0 & -\tilde{d} \\ 0 & 0 & 0 & 0 & 0 \\ \tilde{c}\tilde{d} & 0 & -\tilde{d} & 0 & \tilde{d}^2 \end{bmatrix}, & M_{00,1}^\dagger &= \begin{bmatrix} (1-\tilde{c})^2 & 1-\tilde{c} & 1-\tilde{c} & 0 & (\tilde{c}-1)\tilde{d} \\ 1-\tilde{c} & 1 & 1 & 0 & -\tilde{d} \\ 1-\tilde{c} & 1 & 1 & 0 & -\tilde{d} \\ 0 & 0 & 0 & 0 & 0 \\ (\tilde{c}-1)\tilde{d} & -\tilde{d} & -\tilde{d} & 0 & \tilde{d}^2 \end{bmatrix}, \\
M_{11,0}^\dagger &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, & M_{01,0}^\dagger &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \\
M_{10,0}^\dagger &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, & M_{00,0}^\dagger &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.
\end{aligned}$$

By Lemma A.1, in analogy to the verification of S3(iv)(a), since $\sum_{y,d=0,1} \lambda_{\min}(\tilde{M}_{y,d,0}^\dagger) = \lambda_{\min}(\tilde{M}_{00,1}^\dagger) = 0$, we consider the rest of the sum in (A.32) and apply (A.27). Let $a \equiv p_{11}^0/p_{01}^0$ and $b \equiv p_{11}^0/p_{10}^0$. Then, $\tilde{M}_{11,1}^\dagger + \tilde{M}_{01,1}^\dagger + \tilde{M}_{10,1}^\dagger = (M_{11,1}^\dagger + aM_{01,1}^\dagger + bM_{10,1}^\dagger)/p_{11}^0$ and

$$\begin{aligned}
M^\dagger &\equiv M_{11,1}^\dagger + aM_{01,1}^\dagger + bM_{10,1}^\dagger \\
&= \begin{bmatrix} a(1-c)^2 + b\tilde{c}^2 + c^2 & a(1-c) & -b\tilde{c} & -a(1-c) + c & a(c-1)d + b\tilde{c}\tilde{d} + cd \\ a(1-c) & a & 0 & -a & -ad \\ -b\tilde{c} & 0 & b & 0 & -b\tilde{d} \\ -a(1-c) + c & -a & 0 & a+1 & (a+1)d \\ a(c-1)d + b\tilde{c}\tilde{d} + cd & -ad & -b\tilde{d} & (a+1)d & (a+1)d^2 + b\tilde{d}^2 \end{bmatrix}.
\end{aligned}$$

For the k -th leading principal minor $|M_k^\dagger|$ of M^\dagger ,

$$|M_1^\dagger| = a(1-c)^2 + b\tilde{c}^2 + c^2 > 0,$$

$$|M_2^\dagger| = ab\tilde{c}^2 + ac^2 > 0,$$

$$|M_3^\dagger| = abc^2 > 0,$$

$$|M_4^\dagger| = a^2b(1-c)^2 + abc^2 > 0,$$

$$|M_5^\dagger| = |M^\dagger| = ab \left\{ a^2(1+(1-c)^2)d^2 + b^2\tilde{c}^2\tilde{d}^2 + c^2(d^2 + b\tilde{d}^2) + a \left(((1-c)^2 + c^2)d^2 + b\tilde{c}^2\tilde{d}^2 \right) \right\} > 0.$$

Therefore, $\tilde{M}_{01,1}^\dagger + \tilde{M}_{10,1}^\dagger + \tilde{M}_{11,1}^\dagger$ is positive definite and by (A.27), we can easily show that $\lambda_{\min}(E_{\gamma_0} \rho_{\theta\theta}^\dagger(W_i, \theta_0)) > 0$. ■

Verification of S3(v): Recall

$$V^\dagger(\theta_1, \theta_2; \gamma_0) \equiv \text{Cov}_{\gamma_0} \left(\rho_\theta^\dagger(W_i, \theta_1), \rho_\theta^\dagger(W_i, \theta_2) \right).$$

But

$$\begin{aligned} \text{Cov}_{\gamma_0} \left(\rho_\theta^\dagger(W_i, \theta_0), \rho_\theta^\dagger(W_i, \theta_0) \right) &= E_{\gamma_0} \rho_\theta^\dagger(W_i, \theta_0) \rho_\theta^\dagger(W_i, \theta_0)' \\ &= E_{\gamma_0} \rho_{\theta\theta}^\dagger(W_i, \theta_0), \end{aligned} \quad (\text{A.33})$$

where the first equality is by $E_{\gamma_0} \rho_\theta^\dagger(W_i, \theta_0) = B^{-1}(\beta_0) E_{\gamma_0} \rho_\theta(w, \theta_0) = 0$ and the second equality is by the definition of $\rho_\theta^\dagger(W_i, \theta)$ and $\rho_{\theta\theta}^\dagger(W_i, \theta)$. Since $E_{\gamma_0} \rho_{\theta\theta}^\dagger(W_i, \theta_0)$ is positive definite from S3(iv)(b), we have the desired result. ■

Define the $d_\psi \times d_\beta$ matrix-valued function

$$K(\theta; \gamma_0) \equiv \frac{\partial}{\partial \beta_0'} E_{\gamma_0} \rho_\psi(W_i, \theta) \quad (\text{A.34})$$

with domain $\Theta_\delta \times \Gamma_0$, where $\Theta_\delta \equiv \{\theta \in \Theta : |\beta| < \delta\}$ and

$$\Gamma_0 \equiv \{\gamma_a = (a\beta, \zeta, \pi, \phi) \in \Gamma : \gamma = (\beta, \zeta, \pi, \phi) \in \Gamma \text{ with } |\beta| < \delta \text{ and } a \in [0, 1]\}$$

for some $\delta > 0$.

Assumption S4: (i) $K(\theta; \gamma_0)$ exists $\forall (\theta, \gamma_0) \in \Theta_\delta \times \Gamma_0$.

(ii) $K(\theta; \gamma^*)$ is continuous in (θ, γ^*) at $(\theta, \gamma^*) = ((\psi_0, \pi), \gamma_0)$ uniformly over $\pi \in \Pi \forall \gamma_0 \in \Gamma$ with $\beta_0 = 0$, where ψ_0 is a subvector of γ_0 .

Verification of S4(i): Note that

$$\begin{aligned} K(\theta; \gamma_0) &\equiv \frac{\partial}{\partial \beta_0} E_{\gamma_0} \rho_\psi(W_i, \theta) \\ &= - \frac{\partial}{\partial \beta_0} \sum_{y,d,z=0,1} \frac{p_{yd,z}(\theta_0) \phi_{z,0}}{p_{yd,z}(\theta)} D_\psi p_{yd,z}(\theta) \\ &= - \sum_{y,d,z=0,1} \frac{\partial p_{yd,z}(\theta_0)}{\partial \beta_0} \frac{\phi_{z,0}}{p_{yd,z}(\theta)} D_\psi p_{yd,z}(\theta), \end{aligned}$$

where $\frac{\partial p_{yd,z}(\theta_0)}{\partial \beta_0}$ is the first element of $D_{\psi_0} p_{yd,z}(\theta_0)$ for all (y, d, z) , whose expressions are above. ■

Verification of S4(ii): For

$$K(\pi; \gamma_0) \equiv K(\psi_0, \pi; \gamma_0) = - \sum_{y,d,z=0,1} \frac{\partial p_{y,d,z}(\theta_0)}{\partial \beta_0} \frac{\phi_{z,0}}{p_{y,d,z}(\psi_0, \pi)} D_{\psi} p_{y,d,z}(\psi_0, \pi),$$

let $a_{y,d,z}(\pi, \theta_0, \phi_{1,0}) \equiv \frac{\partial p_{y,d,z}(\theta_0)}{\partial \beta_0} \frac{\phi_{z,0}}{p_{y,d,z}(\psi_0, \pi)} D_{\psi} p_{y,d,z}(\psi_0, \pi)$ since $\phi_{0,0} = 1 - \phi_{1,0}$. Note that $a_{y,d,z}(\pi, \theta_0, \phi_{1,0})$ is continuous in its arguments by Lemma A.1(iv). We can show that $a_{y,d,z}(\pi, \theta_0, \phi_{1,0})$ is continuous uniformly in $\pi \in \Pi$ by applying the uniform convergence result in Lemma 9.2 of ACMLwp to $a_{y,d,z}(\pi, \theta_n, \phi_{1,n}) - a_{y,d,z}(\pi, \theta_0, \phi_{1,0})$, using (i) the pointwise convergence (i.e., pointwise continuity) above, (ii) $a_{y,d,z}(\pi, \theta_0, \phi_{1,0})$'s differentiability in π with derivatives bounded over $\pi \in \Pi$ by Lemma A.1 and (iii) the compactness of Π (B1(iii) below). ■

Next, we impose conditions on the parameter spaces Θ and Γ . Recall $\Theta_{\delta}^* \equiv \{\theta \in \Theta^* : |\beta| < \delta\}$, where Θ^* is the true parameter space for θ . The “optimization parameter space” Θ satisfies:

Assumption B1 (AC13): (i) $\text{int}(\Theta) \supset \Theta^*$.

(ii) For some $\delta > 0$, $\Theta \supset \{\beta \in \mathbb{R}^{d_{\beta}} : |\beta| < \delta\} \times \mathcal{Z}^0 \times \Pi \supset \Theta_{\delta}^*$ for some non-empty open set $\mathcal{Z}^0 \subset \mathbb{R}^{d_{\mathcal{Z}}}$ and Π .

(iii) Π is compact.

The following general results are useful in verifying B1 and B2 below: for a continuous function f , (i) if a set A is compact, then $f(A)$ is compact and (ii) $f^{-1}(\text{int}(A)) \subset \text{int}(f^{-1}(A))$ for any set A in the range of f , where the latter is necessary and sufficient for continuity. Also note that by definition, for a proper function f , if B is compact, then $f^{-1}(B)$ is compact. Lastly, for a function f , if $A \subset B$ then $f(A) \subset f(B)$.

Verification of B1: TC3(ii) implies B1(i) since

$$\text{int}(\Theta) = \text{int}(\bar{h}^{-1}(\Theta)) \supset \bar{h}^{-1}(\text{int}(\Theta)) \supset \bar{h}^{-1}(\Theta^*) = \Theta^*,$$

where the first \supset is by the continuity of \bar{h} and the second \supset is by TC3(ii) and \bar{h}^{-1} being a function. For B1(ii), first note that given TC3(iii),

$$\bar{h}^{-1}(\Theta) \supset \bar{h}^{-1}(\{\beta \in \mathbb{R}^{d_{\beta}} : \|\beta\| < \delta\} \times \mathcal{Z}^0 \times \Pi) \supset \bar{h}^{-1}(\Theta_{\delta}^*).$$

But $\bar{h}^{-1}(\Theta) = \Theta$ and

$$\begin{aligned} \bar{h}^{-1}(\Theta_{\delta}^*) &= \{\theta \in \Theta^* : \bar{h}(\theta) \in \Theta_{\delta}^*\} \\ &= \{\theta \in \Theta^* : \bar{h}(\theta) \in \Theta^*, |\bar{h}_1(\theta)| < \delta\} \\ &= \{\theta \in \Theta^* : \theta \in \Theta^*, |\beta| < \delta\} \\ &= \Theta_{\delta}^*, \end{aligned}$$

where the third equality is by \bar{h} being a homeomorphism and $\bar{h}_1(\theta) = \beta$ being the first element of \bar{h} . Also, with $B_\delta \equiv \{\beta \in \mathbb{R}^{d_\beta} : |\beta| < \delta\}$,

$$\begin{aligned} \bar{h}^{-1}(B_\delta \times \mathcal{Z}^0 \times \Pi) &= \{\theta \in \Theta^* : \bar{h}(\theta) \in B_\delta \times \mathcal{Z}^0 \times \Pi\} \\ &= B_\delta \times \{\mu \in \mathcal{M}^* : h(\mu) \in \mathcal{Z}^0 \times \Pi\} \\ &= B_\delta \times h^{-1}(\mathcal{Z}^0 \times \Pi) \\ &\equiv B_\delta \times \mathcal{Z}^0 \times \Pi, \end{aligned}$$

where $\mathcal{M}^* = \{\mu \in \mathbb{R}^{d_\mu} : \theta = (\beta, \mu) \text{ for some } \theta \in \Theta^*\}$, the second equality holds since $\bar{h}(\theta) = (\beta, h(\mu))$ and the last equality holds by TC3(iv). Lastly, B1(iii) holds by TC3(i). ■

Assumption B2 (AC13): (i) Γ is compact and $\Gamma = \{\gamma = (\theta, \phi) : \theta \in \Theta^*, \phi \in \Phi^*(\theta)\}$.

(ii) $\forall \delta > 0, \exists \gamma = (\beta, \zeta, \pi, \phi) \in \Gamma$ with $0 < \|\beta\| < \delta$.

(iii) $\forall \gamma = (\beta, \zeta, \pi, \phi) \in \Gamma$ with $0 < \|\beta\| < \delta$ for some $\delta > 0, \gamma_a = (a\beta, \zeta, \pi, \phi) \in \Gamma \forall a \in [0, 1]$.

Verification of B2: Consider B2(i). Under TC4(i), define $\Phi^*(\theta)$ as $\Phi^*(\theta) \equiv \Phi^*(\bar{h}(\theta))$. Since Γ is compact, Θ^* and $\Phi^*(\theta)$ are compact for $\theta \in \Theta^*$. Thus, $\Theta^* = \bar{h}^{-1}(\Theta^*)$ is compact by the properness of \bar{h} . Also given (A.4), we have

$$\Phi^*(\theta) \equiv \Phi^*(\bar{h}(\theta)) = \Phi^* = [0.01, 0.99],$$

which is compact. And therefore Γ is also compact. Next, TC4(ii) implies B2(ii). This is because, $\forall \delta > 0$, for $\gamma = (\beta, \mu, \phi)$ that satisfies TC4(ii), let γ in B2(ii) be $\gamma = (\beta, h^{-1}(\mu), \phi)$, which is in Γ since $(\beta, \mu) \in \Theta^*$ implies $(\beta, h^{-1}(\mu)) = \bar{h}^{-1}(\beta, \mu) \in \Theta^*$. To show that TC4(iii) implies B2(iii), note that for any $\gamma = (\beta, \zeta, \pi, \phi) \in \Gamma$ with $0 < |\beta| < \delta$ for some $\delta > 0, \gamma = (\beta, h(\zeta, \pi), \phi) \in \Gamma$. By TC4(iii), this implies that $\gamma_a = (a\beta, h(\zeta, \pi), \phi) \in \Gamma \forall a \in [0, 1]$. Therefore, $\gamma_a = (a\beta, h^{-1}(h(\zeta, \pi)), \phi) \in \Gamma$. ■

Define a “weighted non-central chi-square” process $\{\xi(\pi; \gamma_0, b) : \pi \in \Pi\}$ by

$$\xi(\pi; \gamma_0, b) \equiv -\frac{1}{2} (G(\pi; \gamma_0) + K(\pi; \gamma_0)b)' H^{-1}(\pi; \gamma_0) (G(\pi; \gamma_0) + K(\pi; \gamma_0)b),$$

where $G(\pi; \gamma_0)$ is defined such that $G_n(\cdot) \Rightarrow G(\cdot; \gamma_0)$, where “ \Rightarrow ” denotes weak convergence, with

$$G_n(\pi) \equiv n^{-1/2} \sum_{i=1}^n (\rho_\psi(W_i; \psi_{0,n}, \pi) - E_{\gamma_n} \rho_\psi(W_i; \psi_{0,n}, \pi))$$

and

$$H(\pi; \gamma_0) \equiv E_{\gamma_0} \rho_{\psi\psi}(W_i; \psi_0, \pi).$$

Assumption C6: Each sample path of the stochastic process $\{\xi(\pi; \gamma_0, b) : \pi \in \Pi\}$ in some

set $A(\gamma_0, b)$ with $\Pr_{\gamma_0}(A(\gamma_0, b)) = 1$ is minimized over Π at a unique point (which may depend on the sample path), denoted $\pi^*(\gamma_0, b)$, $\forall \gamma_0 \in \Gamma$ with $\beta_0 = 0$, $\forall b$ with $\|b\| < \infty$.

In Assumption C6, $\pi^*(\gamma_0, b)$ is random. The following is a primitive sufficient condition for Assumption C6 for the case where β is scalar. Let $\rho_\psi(w, \theta) \equiv (\rho_\beta(w, \theta)', \rho_\zeta(w, \theta)')$. When $\beta = 0$, $\rho_\zeta(w, \theta)'$ does not depend on π by Assumption S2(ii) and is denoted by $\rho_\zeta(w, \psi)'$. For $\beta_0 = 0$, define

$$\rho_\psi^*(W_i, \psi_0, \pi_1, \pi_2)' \equiv (\rho_\beta(W_i, \psi_0, \pi_1)', \rho_\beta(W_i, \psi_0, \pi_2)', \rho_\zeta(W_i, \psi_0)')',$$

$$\Omega_G(\pi_1, \pi_2; \psi_0) \equiv Cov_{\gamma_0}(\rho_\psi^*(W_i, \psi_0, \pi_1, \pi_2)', \rho_\psi^*(W_i, \psi_0, \pi_1, \pi_2)').$$

Assumption C6[†]: (i) $d_\beta = 1$

(ii) $\Omega_G(\pi_1, \pi_2; \gamma_0)$ is positive definite $\forall \pi_1, \pi_2 \in \Pi$ with $\pi_1 \neq \pi_2$, $\forall \gamma_0 \in \Gamma$ with $\beta_0 = 0$.

Note that Assumptions S1–S3 and C6[†] imply C6; see Lemma 3.1 of AC13.

Verification of C6[†](ii): Noting that $D_\zeta p_{y_{d,z}}(\psi_0, \pi)$ does not depend on π when $\beta_0 = 0$ so that we may denote it $D_\zeta p_{y_{d,z}}(\psi_0)$, define

$$D_\psi p_{y_{d,z}}^*(\psi_0, \pi_1, \pi_2) \equiv (D_\beta p_{y_{d,z}}(\psi_0, \pi_1)', D_\beta p_{y_{d,z}}(\psi_0, \pi_2)', D_\zeta p_{y_{d,z}}(\psi_0)')'.$$

Then

$$\begin{aligned} \Omega_G(\pi_1, \pi_2; \psi_0) &= E_{\gamma_0} \rho_\psi^*(W_i, \psi_0, \pi_1, \pi_2) \rho_\psi^*(W_i, \psi_0, \pi_1, \pi_2)' \\ &= \sum_{y,d,z=0,1} \frac{\phi_{z,0}}{p_{yd}^0} D_\psi p_{y_{d,z}}^*(\psi_0, \pi_1, \pi_2) D_\psi p_{y_{d,z}}^*(\psi_0, \pi_1, \pi_2)', \end{aligned}$$

where the second equality follows from (A.6) and $D_\psi p_{y_{d,z}}^*(\psi_0, \pi_1, \pi_2)$ can be expressed as

$$D_\psi p_{11,0}^* = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad D_\psi p_{10,0}^* = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad D_\psi p_{01,0}^* = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \quad D_\theta p_{00,0}^* = \begin{bmatrix} 0 \\ 0 \\ -1 \\ -1 \\ 0 \end{bmatrix},$$

$$D_\psi p_{11,1}^* = \begin{bmatrix} C_2(h_3(\zeta_{10}, \zeta_{30}, \pi_1), \zeta_{10}; \pi_1) \\ C_2(h_3(\zeta_{10}, \zeta_{30}, \pi_2), \zeta_{10}; \pi_2) \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

$$D_\psi p_{10,1}^* = \begin{bmatrix} -C_2(h_2(\zeta_{10}, \zeta_{20}, \pi_1), \zeta_{10}; \pi_1) \\ -C_2(h_2(\zeta_{10}, \zeta_{20}, \pi_2), \zeta_{10}; \pi_2) \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

and

$$D_\psi p_{01,1}^* = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - D_\psi p_{11,1}^*, \quad D_\psi p_{00,1}^* = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 0 \\ 0 \end{bmatrix} - D_\psi p_{10,1}^*,$$

using (A.23)-(A.26). The remaining arguments are similar to those used in the verification of S3(iv)(a): Let $M_{y,d,z}^* \equiv D_\psi p_{y,d,z}^*(\psi_0, \pi_1, \pi_2) \times D_\psi p_{y,d,z}^*(\psi_0, \pi_1, \pi_2)'$ and $\tilde{M}_{y,d,z}^* \equiv M_{y,d,z}^*/p_{y,d}^0$. Then,

$$\Omega_G(\pi_1, \pi_2; \psi_0) = \phi_{1,0} \sum_{y,d=0,1} \tilde{M}_{y,d,1}^* + \phi_{0,0} \sum_{y,d=0,1} \tilde{M}_{y,d,0}^*. \quad (\text{A.35})$$

Let $c \equiv C_2(h_3(\zeta_{10}, \zeta_{30}, \pi_1), \zeta_{10}; \pi_1)$, $\tilde{c} \equiv C_2(h_3(\zeta_{10}, \zeta_{30}, \pi_2), \zeta_{10}; \pi_2)$, $d \equiv C_2(h_2(\zeta_{10}, \zeta_{20}, \pi_1), \zeta_{10}; \pi_1)$, and $\tilde{d} \equiv C_2(h_2(\zeta_{10}, \zeta_{20}, \pi_2), \zeta_{10}; \pi_2)$ for notational simplicity. Then,

$$M_{11,1}^* = \begin{bmatrix} c^2 & c\tilde{c} & 0 & 0 & c \\ c\tilde{c} & \tilde{c}^2 & 0 & 0 & \tilde{c} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ c & \tilde{c} & 0 & 0 & 1 \end{bmatrix}, \quad M_{01,1}^* = \begin{bmatrix} (1-c)^2 & (1-c)(1-\tilde{c}) & 1-c & 0 & -(1-c) \\ (1-c)(1-\tilde{c}) & (1-\tilde{c})^2 & 1-\tilde{c} & 0 & -(1-\tilde{c}) \\ 1-c & 1-\tilde{c} & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ -(1-c) & -(1-\tilde{c}) & -1 & 0 & 1 \end{bmatrix},$$

$$M_{10,1}^* = \begin{bmatrix} d^2 & d\tilde{d} & 0 & d & 0 \\ d\tilde{d} & \tilde{d}^2 & 0 & \tilde{d} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ d & \tilde{d} & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad M_{00,1}^* = \begin{bmatrix} (1-d)^2 & (1-d)(1-\tilde{d}) & 1-d & 1-d & 0 \\ (1-d)(1-\tilde{d}) & (1-\tilde{d})^2 & 1-\tilde{d} & 1-\tilde{d} & 0 \\ 1-d & 1-\tilde{d} & 1 & 1 & 0 \\ 1-d & 1-\tilde{d} & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$M_{11,0}^* = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad M_{01,0}^* = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$M_{10,0}^* = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad M_{00,0}^* = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

By Lemma A.1 and similar arguments to those used to verify S3(iv)(a), since $\sum_{y,d=0,1} \lambda_{\min}(\tilde{M}_{y^d,0}^*) = \lambda_{\min}(\tilde{M}_{00,1}^*) = 0$, we consider the rest of the sum in (A.35) and apply (A.27). Let $a \equiv p_{01}^0/p_{10}^0$ and $b \equiv p_{01}^0/p_{11}^0$. Then, $\tilde{M}_{01,1}^* + \tilde{M}_{10,1}^* + \tilde{M}_{11,1}^* = (M_{01,1}^* + aM_{10,1}^* + bM_{11,1}^*)/p_{01}^0$, and

$$M^* \equiv M_{01,1}^* + aM_{10,1}^* + bM_{11,1}^* = \begin{bmatrix} ad^2 + (1-c)^2 + bc^2 & ad\tilde{d} + (1-c)(1-\tilde{c}) + bc\tilde{c} & 1-c & ad & -(1-c) + bc \\ ad\tilde{d} + (1-c)(1-\tilde{c}) + bc\tilde{c} & a\tilde{d}^2 + (1-\tilde{c})^2 + b\tilde{c}^2 & 1-\tilde{c} & a\tilde{d} & -(1-\tilde{c}) + b\tilde{c} \\ 1-c & 1-\tilde{c} & 1 & 0 & -1 \\ ad & a\tilde{d} & 0 & a & 0 \\ -(1-c) + bc & -(1-\tilde{c}) + b\tilde{c} & -1 & 0 & 1+b \end{bmatrix}.$$

For the k -th leading principal minor $|M_k^*|$ and determinant $|M^*|$ of M^* ,

$$|M_1^*| = ad^2 + (1-c)^2 > 0,$$

$$|M_2^*| = a \left\{ \tilde{d}(1-c) - d(1-\tilde{c}) \right\}^2 + b \left\{ c(1-\tilde{c}) - \tilde{c}^2(1-c) \right\}^2 + ab(\tilde{d}c - d\tilde{c})^2 > 0,$$

$$|M_3^*| = ab(\tilde{d}c - d\tilde{c})^2 + ab(\tilde{d}c + d\tilde{c})^2 + 4bc\tilde{c}(1-c)(1-\tilde{c}) > 0,$$

$$|M_4^*| = a \left\{ a\tilde{d}^2(1-c)^2 + (1-c)^2(1-\tilde{c})^2 \right\} + ab \left\{ (1-\tilde{c})^2c^2 + bc^2\tilde{c}^2 + ad^2\tilde{c}^2 + (1-c)^2\tilde{c}^2 \right\} > 0,$$

$$|M^*| = ab \left[a(\tilde{d}c - d\tilde{c})^2 + \{c(1-\tilde{c}) - \tilde{c}(1-c)\}^2 + a \left\{ (\tilde{d}(1-c) - d\tilde{c})^2 + (1-b)d^2\tilde{c}^2 \right\} + a^2d^2\tilde{d}^2 + (1-c)^2(1-\tilde{c})^2 + b(1-\tilde{c})^2c^2 + b^2c^2\tilde{c}^2 + b\tilde{c}^2(1-c)^2 + 2ad^2c(1-c) \right] > 0.$$

Therefore, $\tilde{M}_{01,1}^* + \tilde{M}_{10,1}^* + \tilde{M}_{11,1}^*$ is positive definite and by (A.27), we can easily show that $\lambda_{\min}(\Omega_G(\pi_1, \pi_2; \psi_0)) > 0$. ■

Define a non-stochastic function $\{\eta(\pi; \gamma_0) : \pi \in \Pi\}$ by

$$\eta(\pi; \gamma_0) \equiv -\frac{1}{2}K(\pi; \gamma_0)'H^{-1}(\pi; \gamma_0)K(\pi; \gamma_0).$$

Assumption C7: The non-stochastic function $\eta(\pi; \gamma_0)$ is uniquely minimized over $\pi \in \Pi$

at $\pi_0 \forall \gamma_0 \in \Gamma$ with $\beta_0 = 0$.

For $\beta_0 = 0$, by (A.15)–(A.16) we can write

$$\begin{aligned} K(\pi; \gamma_0) &= - \sum_{y,d,z=0,1} \frac{\phi_{z,0}}{p_{yd}^0} \frac{\partial p_{yd,z}(\theta_0)}{\partial \beta_0} D_\psi p_{yd,z}(\psi_0, \pi), \\ H(\pi; \gamma_0) &= \sum_{y,d,z=0,1} \frac{\phi_{z,0}}{p_{yd}^0} D_\psi p_{yd,z}(\psi_0, \pi) D_\psi p_{yd,z}(\psi_0, \pi)'. \end{aligned}$$

Note that we can partition $H(\pi)$ and $K(\pi)$, suppressing γ_0 , as

$$H(\pi) = \begin{bmatrix} H_{11}(\pi) & H_{12}(\pi) \\ H_{21}(\pi) & H_{22} \end{bmatrix} \begin{matrix} \} d_\beta \\ \} d_\zeta \end{matrix} \quad \text{and} \quad K(\pi) = \begin{pmatrix} K_1(\pi) \\ K_2 \end{pmatrix} \begin{matrix} \} d_\beta \\ \} d_\zeta \end{matrix},$$

and note that $K(\pi_0) = [-H_{11}(\pi_0) : -H_{21}(\pi_0)]'$ by the expressions for $K(\pi; \gamma_0)$ and $H(\pi; \gamma_0)$.

Verification of C7: We first show that, for any $\pi \in \Pi$,

$$\eta(\pi) \geq \eta(\pi_0).$$

For matrices A and B , let $A \leq B$ denote $B - A$ being p.s.d. Then we can show that

$$K(\pi)' H^{-1}(\pi) K(\pi) \leq H_{11}(\pi_0) = K(\pi_0)' H^{-1}(\pi_0) K(\pi_0), \quad (\text{A.36})$$

where the inequality is an application of the matrix Cauchy-Schwarz inequality (Proposition A.1 below) and the equality holds because $K(\pi_0) = [-H_{11}(\pi_0) : -H_{21}(\pi_0)]'$; see below for the proof. Lastly, the weak inequality in (A.36) holds as an equality if and only if $\rho_\beta(W_i, \psi_0, \pi_0)a + \rho_\psi(W_i, \psi_0, \pi)'b = 0$ with probability 1 for some $a \in \mathbb{R}$ and $b \in \mathbb{R}^{d_\psi}$ with $(a, b)' \neq 0$. Let $D_\beta p_{yd,z}^0 \equiv D_\beta p_{yd,z}(\psi_0, \pi_0)$ and $D_\psi p_{yd,z}(\pi) \equiv D_\psi p_{yd,z}(\psi_0, \pi)$ for simplicity. Then, when $\beta_0 = 0$

$$\rho_\beta(W_i, \psi_0, \pi_0)a + \rho_\psi(W_i, \psi_0, \pi)'b = \sum_{y,d,z=0,1} \frac{\mathbf{1}_{ydz}(W_i)}{p_{yd}^0} [D_\beta p_{yd,z}^0 a + D_\psi p_{yd,z}(\pi)'b].$$

But, it is easy to see that a $(1 + d_\psi) \times 8$ matrix (suppressing π in $D_\psi p_{yd,z}(\pi)$ and letting $h_{3,0} \equiv h_3(\pi_0)$ and $h_{2,0} \equiv h_2(\pi_0)$)

$$\begin{aligned} & \begin{bmatrix} D_\beta p_{11,1}^0 & D_\beta p_{10,1}^0 & D_\beta p_{01,1}^0 & D_\beta p_{00,1}^0 & D_\beta p_{11,0}^0 & D_\beta p_{10,0}^0 & D_\beta p_{01,0}^0 & D_\beta p_{00,0}^0 \\ D_\psi p_{11,1} & D_\psi p_{10,1} & D_\psi p_{01,1} & D_\psi p_{00,1} & D_\psi p_{11,0} & D_\psi p_{10,0} & D_\psi p_{01,0} & D_\psi p_{00,0} \end{bmatrix} \\ = & \begin{bmatrix} C_2(h_{3,0}; \zeta_{10}, \pi_0) & -C_2(h_{2,0}, \zeta_{10}; \pi_0) & 1 - C_2(h_{3,0}, \zeta_{10}; \pi_0) & -1 + C_2(h_{2,0}, \zeta_{10}; \pi_0) & 0 & 0 & 0 & 0 \\ C_2(h_3; \zeta_{10}, \pi_0) & -C_2(h_2, \zeta_{10}; \pi_0) & 1 - C_2(h_3, \zeta_{10}; \pi_0) & -1 + C_2(h_2, \zeta_{10}; \pi_0) & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 \end{bmatrix} \end{aligned}$$

has full row rank (i.e., rank of $1 + d_\psi$) except when $\pi = \pi_0$, since

$$\begin{aligned} C_2(h_3(\pi), \zeta_{10}; \pi) &\neq C_2(h_3(\pi_0), \zeta_{10}; \pi_0) \\ C_2(h_2(\pi), \zeta_{10}; \pi) &\neq C_2(h_2(\pi_0), \zeta_{10}; \pi_0) \end{aligned}$$

for $\pi \neq \pi_0$. This can be shown by modifying the proof of Lemmas 3.1 and 4.1 of HV17 under Assumption TC2, which yields

$$\partial C_2(h_3(\pi), \zeta_1; \pi) / \partial \pi = C_{\pi 2}(h_3(\pi), \zeta_1; \pi) + C_{12}(h_3(\pi), \zeta_1; \pi) h_{3,\pi}(\pi) < 0$$

and

$$C_{\pi 2}(h_2(\pi), \zeta_1; \pi) + C_{12}(h_2(\pi), \zeta_1; \pi) h_{2,\pi}(\pi) < 0.$$

In fact, h_2 or h_3 can be seen as u_1^* in Lemma 4.1 of HV17. Therefore, there is no (a, b') with $(a, b') \neq 0$ such that $D_\beta p_{y,d,z}^0 a + D_\psi p'_{y,d,z}(\pi) b = 0$ for all $(y, d, z) \in \{0, 1\}^3$, which implies that there is no (a, b') with $(a, b') \neq 0$ such that $\rho_\beta(W_i, \psi_0, \pi_0) a + \rho_\psi(W_i, \psi_0, \pi) b = 0$ with probability 1. In other words, the equality holds uniquely at $\pi = \pi_0$ so that for any $\pi \neq \pi_0$, $\Pr[c'(\rho_\beta(W_i, \psi_0, \pi_0), \rho_\psi(W_i, \psi_0, \pi))' = 0] < 1$ for all $c \in \mathbb{R}^{d_\beta + d_\psi}$ with $c \neq 0$ and thus the inequality in (A.36) is strict. ■

Proposition A.1. *Let $x \in \mathbb{R}^p$ and $y \in \mathbb{R}^q$ be random vectors such that $E \|x\|^2 < \infty$, $E \|y\|^2 < \infty$, and Eyy' is nonsingular. Then*

$$(Exy')(Eyy')^{-1}(Eyx') \leq Exx'.$$

For our verification proof, taking $x = \rho_\beta(W_i, \psi_0, \pi_0)$ and $y = \rho_\psi(W_i, \psi_0, \pi)$, we have

$$\begin{aligned} E_{\gamma_0} yy' &= H(\pi), \\ E_{\gamma_0} xx' &= H_{11}(\pi_0), \\ -E_{\gamma_0} xy' &= -(E_{\gamma_0} yx')' = K(\pi). \end{aligned}$$

Proof of $H_{11}(\pi_0) = K(\pi_0)'H^{-1}(\pi_0)K(\pi_0)$: Define a 4×4 block-diagonalizing matrix

$$A(r) = \begin{bmatrix} 1 & -H_{12}(r)H_{22}^{-1} \\ 0_3 & I_3 \end{bmatrix}.$$

Then,

$$\begin{aligned} K(r_0)'H^{-1}(r_0)K(r_0) &= K(r_0)'A(r) [A(r)H(r_0)A(r)]^{-1} A(r)K(r_0) \\ &= (-1)^2 [H_{11}(r_0) : H_{21}(r_0)'] A(r) [A(r)H(r_0)A(r)]^{-1} A(r) \begin{bmatrix} H_{11}(r_0) \\ H_{21}(r_0) \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
&= [H_{11}(r_0) - H_{12}(r_0)H_{22}^{-1}H_{21}(r_0) : H_{21}(r_0)'] \begin{bmatrix} H_{11}^*(r_0)^{-1} & 0 \\ 0 & H_{22}^{-1} \end{bmatrix} \\
&\times \begin{bmatrix} H_{11}(r_0) - H_{12}(r_0)H_{22}^{-1}H_{21}(r_0) \\ H_{21}(r_0) \end{bmatrix} \\
&= [1 : H_{21}(r_0)'H_{22}^{-1}] \begin{bmatrix} H_{11}(r_0) - H_{12}(r_0)H_{22}^{-1}H_{21}(r_0) \\ H_{21}(r_0) \end{bmatrix} \\
&= H_{11}(r_0),
\end{aligned}$$

where the second equality is due to the fact that $K(r_0) = [-H_{11}(r_0) : -H_{21}(r_0)']'$ and $H_{11}^*(r_0)$ is implicitly defined. We also use the symmetricity of $H(r)$ in this derivation. ■

Define the following quantities that arise in the asymptotic distribution of $\hat{\theta}_n$ and the test statistics we consider. Letting $S_\psi \equiv [I_{d_\psi} : 0_{d_\psi \times 1}]$ denote the $d_\psi \times d_\theta$ selector matrix that selects ψ out of θ :

$$\begin{aligned}
\Omega(\pi_1, \pi_2; \gamma_0) &\equiv S_\psi V^\dagger((\psi_0, \pi_1), (\psi_0, \pi_2); \gamma_0) S_\psi', \\
J(\theta; \gamma_0) &\equiv E_{\gamma_0} \rho_{\theta\theta}^\dagger(W_i; \theta), \\
V(\theta; \gamma_0) &= V^\dagger(\theta, \theta; \gamma_0),
\end{aligned}$$

and

$$\begin{aligned}
J(\gamma_0) &\equiv J(\theta_0; \gamma_0), \\
V(\gamma_0) &\equiv V(\theta_0; \gamma_0).
\end{aligned}$$

Note that

$$J(\gamma_0) = V(\gamma_0)$$

by (A.33). Define

$$\Sigma(\theta; \gamma_0) \equiv J^{-1}(\theta; \gamma_0) V(\theta; \gamma_0) J^{-1}(\theta; \gamma_0)$$

and

$$\Sigma(\pi; \gamma_0) \equiv \Sigma(\psi_0, \pi; \gamma_0).$$

Assumption V1: (i) $\hat{J}_n = \hat{J}_n(\hat{\theta}_n)$ and $\hat{V}_n = \hat{V}_n(\hat{\theta}_n)$ for some (stochastic) $d_\theta \times d_\theta$ matrix-valued functions $\hat{J}_n(\theta)$ and $\hat{V}_n(\theta)$ on Θ that satisfy $\sup_{\theta \in \Theta} \|\hat{J}_n(\theta) - J(\theta; \gamma_0)\| \rightarrow_p 0$ and $\sup_{\theta \in \Theta} \|\hat{V}_n(\theta) - V(\theta; \gamma_0)\| \rightarrow_p 0$ under $\{\gamma_n\} \in \Gamma(\gamma_0, 0, b)$ with $\|b\| < \infty$.

(ii) $J(\theta; \gamma_0)$ and $V(\theta; \gamma_0)$ are continuous in θ on $\Theta \forall \gamma_0 \in \Gamma$ with $\beta_0 = 0$.

(iii) $\lambda_{\min}(\Sigma(\pi; \gamma_0)) > 0$ and $\lambda_{\max}(\Sigma(\pi; \gamma_0)) < \infty \forall \pi \in \Pi, \forall \gamma_0 \in \Gamma$ with $\beta_0 = 0$.

Verification of V1(i): We define the following:

$$\hat{J}_n(\theta) \equiv \frac{1}{n} \sum_{i=1}^n \rho_{\theta\theta}^\dagger(W_i, \theta) = \frac{1}{n} \sum_{i=1}^n \rho_\theta^\dagger(W_i, \theta) \rho_\theta^\dagger(W_i, \theta)'$$

$$= \frac{1}{n} \sum_{i=1}^n \sum_{y,d,z=0,1} \mathbf{1}_{ydz}(W_i) \frac{D_{\theta} p_{y,d,z}^{\dagger}(\theta) D_{\theta} p_{y,d,z}^{\dagger}(\theta)'}{p_{y,d,z}(\theta)^2},$$

where $D_{\theta} p_{y,d,z}^{\dagger}(\theta_0)$ are defined above. Also,

$$\begin{aligned} \widehat{V}_n(\theta) &\equiv \frac{1}{n} \sum_{i=1}^n \rho_{\theta}^{\dagger}(W_i, \theta) \rho_{\theta}^{\dagger}(W_i, \theta)' \\ &= \frac{1}{n} \sum_{i=1}^n \sum_{y,d,z=0,1} \mathbf{1}_{ydz}(W_i) \frac{D_{\theta} p_{y,d,z}^{\dagger}(\theta) D_{\theta} p_{y,d,z}^{\dagger}(\theta)'}{p_{y,d,z}(\theta)^2} = \widehat{J}_n(\theta). \end{aligned}$$

The rest of the proof follows from the uniform law of large numbers in Lemma 9.3 of ACMLwp with Assumptions S1 and S3 and Θ being compact. ■

Verification of V1(ii): The continuity follows from the fact that the first and second derivatives of $p_{y,d,z}(\theta)$ are continuous by Lemma A.1(vi). ■

Verification of V1(iii): Note that

$$\Sigma(\pi; \gamma_0) = J^{-1}(\psi_0, \pi; \gamma_0) V(\psi_0, \pi; \gamma_0) J^{-1}(\psi_0, \pi; \gamma_0) = V^{-1}(\psi_0, \pi; \gamma_0)$$

since $V(\psi_0, \pi; \gamma_0) = J(\psi_0, \pi; \gamma_0)$. This is because

$$\begin{aligned} V(\psi_0, \pi; \gamma_0) &= Cov_{\gamma_0} \left(\rho_{\theta}^{\dagger}(W_i, \psi_0, \pi), \rho_{\theta}^{\dagger}(W_i, \psi_0, \pi) \right) = E_{\gamma_0} \rho_{\theta}^{\dagger}(W_i; \psi_0, \pi) \rho_{\theta}^{\dagger}(W_i; \psi_0, \pi)' \\ &= E_{\gamma_0} \rho_{\theta\theta}^{\dagger}(W_i; \theta) \end{aligned}$$

where the last equality holds since $\rho_{\theta\theta}^{\dagger}(w, \theta) = \rho_{\theta}^{\dagger}(w, \theta) \rho_{\theta}^{\dagger}(w, \theta)'$, and the second-to-last equality holds since

$$\begin{aligned} E_{\gamma_0} \rho_{\theta}^{\dagger}(W_i, \psi_0, \pi) &= - \sum_{y,d,z=0,1} \phi_{z,0} D_{\theta} p_{y,d,z}^{\dagger}(\psi_0, \pi) \\ &= - \sum_{y,d=0,1} \phi_{0,0} D_{\theta} p_{y,d,0}^{\dagger}(\psi_0, \pi) - \sum_{y,d=0,1} \phi_{1,0} D_{\theta} p_{y,d,1}^{\dagger}(\psi_0, \pi) \\ &= 0. \end{aligned}$$

Now, for the first part of V1(iii), note that since each element of the vectors in (A.28)–(A.31) are bounded by TC2(iii) and B2(i), the elements of the matrix

$$V(\psi_0, \pi; \gamma_0) = E_{\gamma_0} \rho_{\theta}^{\dagger}(W_i; \psi_0, \pi) \rho_{\theta}^{\dagger}(W_i; \psi_0, \pi)' = \sum_{y,d,z=0,1} \frac{\phi_{z,0}}{p_{y,d,0}} D_{\theta} p_{y,d,z}^{\dagger}(\psi_0, \pi) D_{\theta} p_{y,d,z}^{\dagger}(\psi_0, \pi)'$$

are bounded. For a $d \times d$ matrix A , $\sum_{i=1}^d |\lambda_i| \leq \sum_{i,j=1}^d |A_{ij}|$ where the λ_i 's are A 's eigenvalues and the A_{ij} 's are A 's elements. Therefore, $\lambda_{\max}(V(\psi_0, \pi; \gamma_0)) < \infty$. This implies that $\lambda_{\min}(V^{-1}(\psi_0, \pi; \gamma_0)) > 0$. By Lemma A.1, the proof of the second part is similar to the

proofs of S3(iv)(b) and S3(v) and we can show that $\lambda_{\min}(V(\psi_0, \pi; \gamma_0)) > 0$, which implies that $\lambda_{\max}(V^{-1}(\psi_0, \pi; \gamma_0)) < \infty$. ■

References

- Andrews, D. W. K., Cheng, X., 2012. Supplement to ‘estimation and inference with weak, semi-strong and strong identification’. *Econometrica Supplementary Material*. [A](#)
- Andrews, D. W. K., Cheng, X., 2013a. Maximum likelihood estimation and uniform inference with sporadic identification failure. *Journal of Econometrics* 173, 36–56. [A](#)
- Andrews, D. W. K., Cheng, X., 2013b. Supplemental appendices for ‘Maximum likelihood estimation and uniform inference with sporadic identification failure’, *Journal of Econometrics Supplement*. [A](#)
- Andrews, D. W. K., Cheng, X., 2014. Supplementary material on ‘GMM estimation and uniform subvector inference with possible identification failure’, *Econometric Theory Supplement*. [A](#)
- Han, S., McCloskey, A., 2018. Estimation and inference with a (nearly) singular jacobian, Unpublished Manuscript, Department of Economics, University of Colorado. ([document](#)), [A](#)
- Han, S., Vytlačil, E., 2017. Identification in a generalization of bivariate probit models with dummy endogenous regressors. *Journal of Econometrics* 199, 63–73. [A](#)
- Weyl, H., 1912. Das asymptotische verteilungsgesetz der eigenwerte linearer partieller differentialgleichungen (mit einer anwendung auf die theorie der hohlraumstrahlung). *Mathematische Annalen* 71 (4), 441–479. [A.3](#)