Supplement to

"INFERENCE FOR PARAMETERS DEFINED BY MOMENT INEQUALITIES USING GENERALIZED MOMENT SELECTION"

BY

DONALD W. K. ANDREWS and GUSTAVO SOARES

COWLES FOUNDATION PAPER NO. 1291s

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

2010

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

SUPPLEMENT TO "INFERENCE FOR PARAMETERS DEFINED BY MOMENT INEQUALITIES USING GENERALIZED MOMENT SELECTION"

(*Econometrica*, Vol. 78, No. 1, January 2010, 119–157)

BY DONALD W. K. ANDREWS AND GUSTAVO SOARES

THIS SUPPLEMENT CONTAINS the proofs of Theorems 1–4 and Lemma 2 of the main paper in Sections S1–S3; the verification of Assumptions GMS1, GMS3, GMS6, and GMS7 for $\varphi^{(5)}$ in Section S4; the specification of the mean vectors used in the simulation results in Section S5; some Monte Carlo simulation results for a missing data model and an interval-outcome linear regression model in Section S6; and some Monte Carlo simulation results that compare the null rejection probabilities of subsampling with and without recentering in Section S7. A separate file provides the GAUSS code used to compute the simulation results.

S1. PROOF OF THE ASYMPTOTIC SIZE RESULT FOR GMS

In this section, we reiterate and prove Theorem 1 of the main paper. In addition, at the end of this section, the proof of Lemma 1 is given.

THEOREM 1: *Suppose Assumptions* 1–3, GMS1, *and* GMS2 *hold and* $0 < \alpha < 1/2$. *Then the nominal level* $1-\alpha$ GMS CS *based on* $T_n(\theta)$ *satisfies the following statements*:

(a) $\text{AsyCS} \geq 1-\alpha$,

(b) $\text{AsyCS} = 1-\alpha$ *if Assumptions* GMS3, GMS4, *and* 7 *also hold*.

(c) $\text{AsyMaxCP} = 1$ *if* $v = 0$ (*i.e.*, *no moment equalities appear*) *and Assumption* M *also holds*.

The following lemmas are used in the proof of Theorem 1. The first lemma uses the following notation. Suppose $\pi = (\pi_1, \pi_2) \in R^p_{+,\infty} \times R^q_{[\pm\infty]}$, where $\pi_2 = (\pi_{2,1}, \pi_{2,2})$ and $\pi_{2,2} = \text{vech}_*(\Omega_{\pi_{2,2}})$ for some $k \times k$ correlation matrix $\Omega_{\pi_{2,2}}$. Given π, define $\pi^*_{1,j} = \infty$ if $\pi_{1,j} > 0$ and $\pi^*_{1,j} = 0$ if $\pi_{1,j} = 0$ for $j = 1, \ldots, p$, and let $\pi^*_1 = (\pi^*_{1,1}, \ldots, \pi^*_{1,p})'$. Define $\pi^* = (\pi^*_1, \pi_2)$ and let $c_{\pi^*}(1-\alpha)$ denote the $1-\alpha$ quantile of $S(\Omega^{1/2}_{\pi_{2,2}} Z^* + (\pi^*_1, 0_v), \Omega_{\pi_{2,2}})$, where $Z^* \sim N(0_k, I_k)$ and, by definition, if $\pi^*_{1,j} = \infty$, then the jth element of $\Omega^{1/2}_{\pi_{2,2}} Z^* + (\pi^*_1, 0_v)$ equals ∞ for $j = 1, \ldots, p$. Because there is a one-to-one mapping between $(\theta, F) \in \mathcal{F}$ and $\gamma \in \Gamma$, we write E_F and E_γ interchangeably.

LEMMA 2: *Suppose Assumptions* 1–3, GMS1, *and* GMS2 *hold and* $0 < \alpha < 1/2$. *Let* $\{\gamma_{n,h} = (\gamma_{n,h,1}, \gamma_{n,h,2}, \gamma_{n,h,3}) : n \geq 1\}$ *be a sequence of points in* Γ *that satisfies* (i) $n^{1/2}\gamma_{n,h,1} \to h_1$ *for some* $h_1 \in R^p_{+,\infty}$, (ii) $\kappa_n^{-1} n^{1/2} \gamma_{n,h,1} \to \pi_1$ *for some*

 DOI: 10.3982/ECTA7502

$\pi_1 \in R^p_{+,\infty}$, *and* (iii) $\gamma_{n,h,2} \to h_2$ *for some* $h_2 \in R^q_{[\pm\infty]}$. *Let* $h = (h_1, h_2)$, $\pi = (\pi_1, \pi_2)$, *and* $\pi_2 = h_2$. *Then* (a)–(c) *hold*:

(a) $\widehat{c}_n(\theta_{n,h}, 1-\alpha) \geq c^*_n$ *a.s. for all* n *for a sequence of random variables* $\{c^*_n : n \geq 1\}$ *that satisfies* $c^*_n \to_p c_{\pi^*}(1-\alpha)$ *under* $\{\gamma_{n,h} : n \geq 1\}$, *where* $\gamma_{n,h,2} = (\theta_{n,h}, \gamma_{n,h,2,2})$.

(b) $\liminf_{n\to\infty} P_{\gamma_{n,h}}(T_n(\theta_{n,h}) \leq \widehat{c}_n(\theta_{n,h}, 1-\alpha)) \geq 1-\alpha$.

(c) *For any subsequence* $\{w_n : n \geq 1\}$ *of* $\{n\}$, *the results of parts* (a) *and* (b) *hold with* w_n *in place of* n *provided conditions* (i)–(iii) *above hold with* w_n *in place of* n.

LEMMA 3: *Suppose Assumptions* 1–3, 7, *and* GMS1–GMS4 *hold and* $0 < \alpha < 1/2$. *Let* (θ^*, F^*) *be an element of* $\mathcal{F}$ *for which Assumption* 7 *applies*, *let* γ^* *be the value in* Γ *that corresponds to* $(\theta^*, F^*) \in \mathcal{F}$, *and let* $h^* = (h^*_1, h^*_2)$ *be defined by* $h^*_1 = (h^*_{1,1}, \ldots, h^*_{1,p})'$, *where* $h^*_{1,j} = \infty$ *if* $\gamma^*_{1,j} > 0$ *and* $h^*_{1,j} = 0$ *if* $\gamma^*_{1,j} = 0$ *for* $j = 1, \ldots, p$, *and* $h^*_2 = \gamma^*_2$. *Let* $c_{h^*}(1-\alpha)$ *denote the* $1-\alpha$ *quantile of the distribution of* $S(\Omega^{1/2}_{h^*_{2,2}} Z^* + (h^*_1, 0_v), \Omega_{h^*_{2,2}})$. *When the true distribution is determined by* γ^* *for all* n, *the following results hold*:

(a) $\widehat{c}_n(\theta^*, 1-\alpha) \to_p c_{h^*}(1-\alpha)$.

(b) $\lim_{n\to\infty} P_{\gamma^*}(T_n(\theta^*) \leq \widehat{c}_n(\theta^*, 1-\alpha)) = 1-\alpha$.

PROOF OF THEOREM 1: First, we prove part (a). Let $\mathrm{CP}_n(\gamma) = P_\gamma(T_n(\theta) \leq \widehat{c}_n(\theta, 1-\alpha))$, where $\gamma = (\gamma_1, \gamma_2, \gamma_3)$, $\gamma_2 = (\gamma_{2,1}, \gamma_{2,2})$, and $\gamma_{2,1} = \theta$. Let $\{\gamma^*_n = (\gamma^*_{n,1}, \gamma^*_{n,2}, \gamma^*_{n,3}) \in \Gamma : n \geq 1\}$ be a sequence such that $\liminf_{n\to\infty} \mathrm{CP}_n(\gamma^*_n) = \liminf_{n\to\infty} \inf_{\gamma\in\Gamma} \mathrm{CP}_n(\gamma)$ (= AsyCS). Such a sequence always exists. Let $\{u_n : n \geq 1\}$ be a subsequence of $\{n\}$ such that $\lim_{n\to\infty} \mathrm{CP}_{u_n}(\gamma^*_{u_n})$ exists and equals $\liminf_{n\to\infty} \mathrm{CP}_n(\gamma^*_n) = $ AsyCS. Such a subsequence always exists.

Let $\gamma^*_{n,1,j}$ denote the jth component of $\gamma^*_{n,1}$ for $j = 1, \ldots, p$. Either

(i) $\limsup_{n\to\infty} u_n^{1/2} \gamma^*_{u_n,1,j} < \infty$ or

(ii) $\limsup_{n\to\infty} u_n^{1/2} \gamma^*_{u_n,1,j} = \infty$.

If (i) holds, then for some subsequence $\{w_n\}$ of $\{u_n\}$,

$$\text{(S1.1)} \qquad \begin{aligned} &\kappa^{-1}_{w_n} w_n^{1/2} \gamma^*_{w_n,1,j} \to 0, \\ &w_n^{1/2} \gamma^*_{w_n,1,j} \to h_{1,j} \quad \text{for some } h_{1,j} \in R_+. \end{aligned}$$

If (ii) holds, then either

(ii)(a) $\limsup_{n\to\infty} \kappa^{-1}_{u_n} u_n^{1/2} \gamma^*_{u_n,1,j} < \infty$ or

(ii)(b) $\limsup_{n\to\infty} \kappa^{-1}_{u_n} u_n^{1/2} \gamma^*_{u_n,1,j} = \infty$.

If (ii)(a) holds, then for some subsequence $\{w_n\}$ of $\{u_n\}$,

$$\text{(S1.2)} \qquad \begin{aligned} &\kappa^{-1}_{w_n} w_n^{1/2} \gamma^*_{w_n,1,j} \to \pi_{1,j} \quad \text{for some } \pi_{1,j} \in R_+, \\ &w_n^{1/2} \gamma^*_{w_n,1,j} \to h_{1,j}, \quad \text{where} \quad h_{1,j} = \infty. \end{aligned}$$

If (ii)(b) holds, then for some subsequence $\{w_n\}$ of $\{u_n\}$,

$$\text{(S1.3)} \qquad \kappa_{w_n}^{-1} w_n^{1/2} \gamma_{w_n,1,j}^* \to \pi_{1,j}, \quad \text{where} \quad \pi_{1,j} = \infty,$$
$$w_n^{1/2} \gamma_{w_n,1,j}^* \to h_{1,j}, \quad \text{where} \quad h_{1,j} = \infty.$$

In addition, for some subsequence $\{w_n\}$ of $\{u_n\}$,

$$\text{(S1.4)} \qquad \gamma_{w_n,2}^* \to h_2 \quad \text{for some } h_2 \in \mathrm{cl}(\varGamma_2).$$

By taking successive subsequences over the p components of $\gamma_{u_n,1}^*$ and $\gamma_{u_n,2}^*$, we find that there exists a subsequence $\{w_n\}$ of $\{u_n\}$ such that for each $j = 1, \ldots, p$, exactly one of the cases (S1.1)–(S1.3) applies and (S1.4) holds. In consequence, conditions (i)–(iii) of Lemma 2 hold. Hence,

$$\text{(S1.5)} \qquad \liminf_{n\to\infty} \mathrm{CP}_{w_n}(\gamma_{w_n}^*) \geq 1 - \alpha$$

by Lemma 2. Since $\lim_{n\to\infty} \mathrm{CP}_{u_n}(\gamma_{u_n}^*) = \mathrm{AsyCS}$ and $\{w_n\}$ is a subsequence of $\{u_n\}$, we have $\lim_{n\to\infty} \mathrm{CP}_{w_n}(\gamma_{w_n}^*) = \mathrm{AsyCS}$. This and (S1.5) yield the result of part (a).

Part (b) follows from part (a) and Lemma 3(b) because $\mathrm{AsyCS} \leq \lim_{n\to\infty} P_{\gamma^*}(T_n(\theta^*) \leq \widehat{c}_n(\theta^*, 1-\alpha))$.

Now, we prove part (c) of the theorem. By assumption, $v = 0$. Under Assumption M, the sequence of constant true values $\{(\theta^*, F^*) \in \mathcal{F} : n \geq 1\}$ satisfies $n^{1/2} E_{F^*} m_j(W_i, \theta^*)/\sigma_{F^*,j}(\theta^*) \to \infty$ for $j = 1, \ldots, p$. Let $\gamma^* = (\gamma_1^*, \gamma_2^*, F^*) \in \varGamma$ correspond to $(\theta^*, F^*) \in \mathcal{F}$, where $\gamma_2^* = (\theta^*, \gamma_{2,2}^*)$. Define $h^* = (\infty^p, \gamma_2^*)$. As in the proof of part (b) of Lemma 2 below, we have $T_n(\theta^*) \to_d J_{h^*}$ under $\{\gamma^*\}$, where J_{h^*} is the distribution of $S(Z_{h_{2,2}^*} + (h_1^*, 0_v), \varOmega_{h_{2,2}^*})$, where $Z_{h_{2,2}^*} \sim N(0_k, \varOmega_{h_{2,2}^*})$. (Note that J_{h^*} is the same as $J_{h_1^*, 0_p}$ defined in the text.) Furthermore, $J_{h^*}(x) = 1$ for $x \geq 0$ because $S(Z_{h_{2,2}^*} + \infty^p, \varOmega_{h_{2,2}^*}) = S(\infty^p, \varOmega_{h_{2,2}^*}) = 0$ by Assumption 3. Using these results, we obtain

$$\text{(S1.6)} \qquad \mathrm{AsyMaxCP} \geq \limsup_{n\to\infty} P_{\gamma^*}(T_n(\theta^*) \leq \widehat{c}_n(\theta^*, 1-\alpha))$$
$$\geq \limsup_{n\to\infty} P_{\gamma^*}(T_n(\theta^*) \leq 0) = J_{h^*}(0) = 1,$$

where the first inequality follows from the definition of AsyMaxCP and the second inequality holds by Assumption 1(c). *Q.E.D.*

PROOF OF LEMMA 2: First, suppose $c_{\pi^*}(1-\alpha) = 0$. In this case, define $c_n^* = 0$ and we have

$$\text{(S1.7)} \qquad c_n(\widehat{\theta}_{n,h}, 1-\alpha) \geq c_n^* \to_p c_{\pi^*}(1-\alpha),$$

where the inequality holds by Assumption 1(c), which establishes part (a) for this case.

Next, suppose $c_{\pi^*}(1-\alpha) > 0$. For $(\xi, \Omega) \in R^k \times \Psi$, let $\varphi^*(\xi, \Omega)$ denote the k-vector whose jth element is

$$\text{(S1.8)} \qquad \varphi_j^*(\xi, \Omega) = \begin{cases} \varphi_j(\xi, \Omega), & \text{if } \pi_{1,j} = 0 \text{ and } j = 1, \ldots, p, \\ \infty, & \text{if } \pi_{1,j} > 0 \text{ and } j = 1, \ldots, p, \\ 0, & \text{if } j = p+1, \ldots, k. \end{cases}$$

By construction,

$$\text{(S1.9)} \qquad \varphi^*(\xi_n(\theta_{n,h}), \widehat{\Omega}_n(\theta_{n,h})) \geq \varphi(\xi_n(\theta_{n,h}), \widehat{\Omega}_n(\theta_{n,h})) \quad \text{a.s. } [Z^*] \text{ for all } n.$$

Let c_n^* denote the $1-\alpha$ quantile of the df of $S(\widehat{\Omega}_n^{1/2}(\theta_{n,h})Z^* + \varphi^*(\xi_n(\theta_{n,h}), \widehat{\Omega}_n(\theta_{n,h})), \widehat{\Omega}_n(\theta_{n,h}))$, where Z^* is random and $(\xi_n(\theta_{n,h}), \widehat{\Omega}_n(\theta_{n,h}))$ is fixed. Then $\widehat{c}_n(\theta_{n,h}, 1-\alpha) \geq c_n^*$ a.s. for all n by (S1.9) and Assumption 1(a).

We now show that $c_n^* \to_p c_{\pi^*}(1-\alpha) > 0$. By definition, $\gamma_{n,h,1} = (\gamma_{n,h,1,1}, \ldots, \gamma_{n,h,1,p})' \in R_+^p$ satisfies

$$\text{(S1.10)} \qquad \sigma_{F_{n,h},j}^{-1}(\theta_{n,h}) E_{F_{n,h}} m_j(W_i, \theta_{n,h}) - \gamma_{n,h,1,j} = 0 \quad \text{for} \quad j = 1, \ldots, p.$$

Also by definition, Γ is such that under any sequence $\{\gamma_{n,h} = (\gamma_{n,h,1}, (\theta_{n,h}, \text{vech}_*(\Omega_{n,h})), F_{n,h}) : n \geq 1\}$ of parameters in Γ that satisfies $n^{1/2}\gamma_{n,h,1} \to h_1$ and $(\theta_{n,h}, \text{vech}_*(\Omega_{n,h})) \to h_2 = (h_{2,1}, h_{2,2})$ for some $h = (h_1, h_2) \in R_{+,\infty}^p \times R_{[\pm\infty]}^q$, we have

$$\text{(S1.11)} \qquad \begin{aligned} &\text{(vii)} && A_n = (A_{n,1}, \ldots, A_{n,k})' \to_d Z_{h_{2,2}} \sim N(0_k, \Omega_{h_{2,2}}) \quad \text{as} \quad n \to \infty, \\ & && \text{where} \quad A_{n,j} = n^{1/2}\big(\overline{m}_{n,j}(\theta_{n,h}) - E_{F_{n,h}}\overline{m}_{n,j}(\theta_{n,h})\big)/\sigma_{F_{n,h},j}(\theta_{n,h}), \\ &\text{(viii)} && \widehat{\sigma}_{n,j}(\theta_{n,h})/\sigma_{F_{n,h},j}(\theta_{n,h}) \to_p 1 \quad \text{as} \quad n \to \infty \quad \text{for} \quad j = 1, \ldots, k, \\ &\text{(ix)} && \widehat{D}_n^{-1/2}(\theta_{n,h})\widehat{\Sigma}_n(\theta_{n,h})\widehat{D}_n^{-1/2}(\theta_{n,h}) \to_p \Omega_{h_{2,2}} \quad \text{as} \quad n \to \infty, \quad \text{and} \\ &\text{(x)} && \text{conditions (vii)–(ix) hold for all subsequences } \{w_n\} \\ & && \text{in place of } \{n\}. \end{aligned}$$

Now, under $\{\gamma_{n,h} : n \geq 1\}$, we have

$$\text{(S1.12)} \qquad \begin{aligned} &\kappa_n^{-1} n^{1/2} \widehat{D}_n^{-1/2}(\theta_{n,h}) \overline{m}_n(\theta_{n,h}) \\ &\quad = \kappa_n^{-1} \widehat{D}_n^{-1/2}(\theta_{n,h}) D^{1/2}(\theta_{n,h}, F_{n,h}) \big(A_n + \big(n^{1/2}\gamma_{n,h,1}, 0_v\big)\big) \\ &\quad = o_p(1) + (I_k + o_p(1)) \big(\kappa_n^{-1} n^{1/2} \gamma_{n,h,1}, 0_v\big) \to_p (\pi_1, 0_v), \end{aligned}$$

where the first equality holds by the definitions of $\gamma_{n,h,1}$ and A_n in (S1.10) and (S1.11), $D(\theta_{n,h}, F_{n,h}) = \text{Diag}\{\sigma_{F_{n,h},1}^2(\theta_{n,h}), \ldots, \sigma_{F_{n,h},k}^2(\theta_{n,h})\}$ in the second line, the second equality holds using $\kappa_n \to \infty$ and conditions (vii) and (viii) of (S1.11), which apply by conditions (i) and (iii) of the lemma, and the convergence holds using condition (ii) of the lemma. This and condition (ix) of (S1.11) yield that under $\{\gamma_{n,h} : n \geq 1\}$,

$$(\xi_n(\theta_{n,h}), \widehat{\Omega}_n(\theta_{n,h})) \to_p \big((\pi_1, 0_v), \Omega_{\pi_{2,2}}\big). \tag{S1.13}$$

For notational simplicity, let Ω_0 denote $\Omega_{\pi_{2,2}}$. We now show that $\varphi^*(\xi, \Omega) \to \varphi((\pi_1, 0_v), \Omega_0)$ for any sequence (ξ, Ω) for which $(\xi, \Omega) \to ((\pi_1, 0_v), \Omega_0)$. If $\pi_{1,j} = 0$ and $j \leq p$, then

$$\varphi_j^*(\xi, \Omega) = \varphi_j(\xi, \Omega) \to \varphi_j((\pi_1, 0_v), \Omega_0) = 0 \tag{S1.14}$$

as $(\xi, \Omega) \to ((\pi_1, 0_v), \Omega_0)$, where the first equality holds by (S1.8), the convergence holds by Assumption GMS1(a), and the last equality holds by Assumption GMS1(b). If $\pi_{1,j} > 0$, then $\varphi_j^*(\xi, \Omega) = \infty = \varphi_j((\pi_1, 0_v), \Omega_0)$ by the definition in (S1.8) and Assumption GMS3. For $j = p+1, \ldots, k$, $\varphi_j^*(\xi, \Omega) = 0 = \varphi_j((\pi_1, 0_v), \Omega_0)$ by the definition in (S1.8) and Assumption GMS1(c). These results, (S1.14), and Assumption 1(d) give, for $x > 0$, as $(\xi, \Omega) \to ((\pi_1, 0_v), \Omega_0)$,

$$\begin{aligned}
&S\big(\Omega^{1/2} Z^* + \varphi^*(\xi, \Omega), \Omega\big) \\
&\quad \to S\big(\Omega_0^{1/2} Z^* + \varphi^*((\pi_1, 0_v), \Omega_0), \Omega_0\big) \quad \text{a.s. } [Z^*], \\
&1\big(S\big(\Omega^{1/2} Z^* + \varphi^*(\xi, \Omega), \Omega\big) \leq x\big) \\
&\quad \to 1\big(S\big(\Omega_0^{1/2} Z^* + \varphi^*((\pi_1, 0_v), \Omega_0), \Omega_0\big) \leq x\big) \quad \text{a.s. } [Z^*], \\
&P\big(S\big(\Omega^{1/2} Z^* + \varphi^*(\xi, \Omega), \Omega\big) \leq x\big) \\
&\quad \to P\big(S\big(\Omega_0^{1/2} Z^* + \varphi^*((\pi_1, 0_v), \Omega_0), \Omega_0\big) \leq x\big).
\end{aligned} \tag{S1.15}$$

The third convergence result of (S1.15) holds by the second result and the bounded convergence theorem. The second convergence result of (S1.15) follows from the first result provided

$$\begin{aligned}
P\big(S\big(Z + \varphi^*((\pi_1, 0_v), \Omega_0), \Omega_0\big) = x\big) &= P\big(S(Z + (\pi_1^*, 0_v), \Omega_0) = x\big) \\
&= 0,
\end{aligned} \tag{S1.16}$$

where $Z = \Omega_0^{1/2} Z^* \sim N(0_k, \Omega_0)$. The first equality in (S1.16) holds because $[(\pi_1^*, 0_v)]_j = \infty = \varphi_j^*((\pi_1, 0_v), \Omega_0)$ if $\pi_{1,j}^* = \infty$ by definition and $[(\pi_1^*, 0_v)]_j = 0 = \varphi_j((\pi_1, 0_v), \Omega_0) = \varphi_j^*((\pi_1, 0_v), \Omega_0)$ if $\pi_{1,j}^* = \pi_{1,j} = 0$ using Assumption GMS1(b) and (S1.8). The second equality in (S1.16) holds because the df of $S(Z + (\pi_1^*, 0_v), \Omega_0)$ is continuous and strictly increasing for $x > 0$ by Assumption 2(a) and (b) unless $v = 0$ and $\pi_1^* = \infty^p$. The latter does not hold because, if $v = 0$ and $\pi_1^* = \infty^p$, then $S(Z + (\pi_1^*, 0_v), \Omega_0) = 0$ by Assumptions 1(c) and 3 and $c_{\pi^*}(1-\alpha) = 0$, which contradicts the assumption that $c_{\pi^*}(1-\alpha) > 0$.

In sum, (S1.15) shows that $P(S(\Omega^{1/2}Z^* + \varphi^*(\xi, \Omega), \Omega) \leq x)$ is a continuous function of (ξ, Ω) at $((\pi_1, 0_v), \Omega_0)$. This, (S1.13), and Slutsky's theorem combine to give, under $\{\gamma_{n,h} : n \geq 1\}$,

$$\begin{aligned} L_n(x) &= P\big(S\big(\widehat{\Omega}_n^{1/2}(\theta_{n,h})Z^* + \varphi^*(\xi_n(\theta_{n,h}), \widehat{\Omega}_n(\theta_{n,h})), \widehat{\Omega}_n(\theta_{n,h})\big) \leq x\big) \\ &\to_p P\big(S\big(\Omega_0^{1/2}Z^* + \varphi^*((\pi_1, 0_v), \Omega_0), \Omega_0\big) \leq x\big) = L(x) \end{aligned} \tag{S1.17}$$

for all $x > 0$, where $P(\cdot)$ denotes conditional probability given $(\xi_n(\theta_{n,h}), \widehat{\Omega}_n(\theta_{n,h}))$ in (S1.17) and hence is a random probability. By definition, c_n^* is the $1-\alpha$ quantile of $L_n(x)$ and $c_{\pi^*}(1-\alpha)$ is the $1-\alpha$ quantile of $L(x)$. By Lemma 5 of Andrews and Guggenberger (2010b) (AG1), given that (S1.17) holds for all x in a neighborhood of $c_{\pi^*}(1-\alpha) > 0$ and that $L(x)$ is continuous and strictly increasing at $x = c_{\pi^*}(1-\alpha)$ (see the previous paragraph), we have $c_n^* \to_p c_{\pi^*}(1-\alpha)$. This completes the proof of part (a).

Part (b) is proved as follows. First, conditions (i) and (ii) of the lemma imply that if $\pi_{1,j} > 0$, then $h_{1,j} = \infty$ and $\pi_{1,j}^* = \infty$, and if $\pi_{1,j} = 0$, then $h_{1,j} \geq 0$ and $\pi_{1,j}^* = 0$. Thus, we have

$$\begin{aligned} &\pi_1^* \leq h_1, \\ &S\big(\Omega_{h_{2,2}}^{1/2}Z^* + (\pi_1^*, 0_v), \Omega_{h_{2,2}}\big) \geq S\big(\Omega_{h_{2,2}}^{1/2}Z^* + (h_1, 0_v), \Omega_{h_{2,2}}\big), \\ &c_{\pi^*}(1-\alpha) \geq c_h(1-\alpha), \end{aligned} \tag{S1.18}$$

where $c_h(1-\alpha)$ denotes the $1-\alpha$ quantile of $S(\Omega_{h_{2,2}}^{1/2}Z^* + (h_1, 0_v), \Omega_{h_{2,2}})$, the second inequality holds by the first inequality and Assumption 1(a), and the third inequality holds by the second.

Second, by the verification of Assumption B0 in Andrews and Guggenberger (2009b) (AG4), we have

$$T_n(\theta_{n,h}) \to_d J_h \quad \text{under } \{\gamma_{n,h}\}, \tag{S1.19}$$

where J_h is the distribution of $S(\Omega_{h_{2,2}}^{1/2}Z^* + (h_1, 0_v), \Omega_{h_{2,2}})$. This result is obtained by using Assumption 1(b) to write

$$T_n(\theta_{n,h}) = S\big(\widehat{D}_n^{-1/2}(\theta_{n,h})n^{1/2}\overline{m}_n(\theta_{n,h}), \widehat{D}_n^{-1/2}(\theta_{n,h})\widehat{\Sigma}_n(\theta_{n,h})\widehat{D}_n^{-1/2}(\theta_{n,h})\big). \tag{S1.20}$$

If any element of h_1 equals ∞, then it can be shown using (S1.11) that the corresponding element of $\widehat{D}_n^{-1/2}(\theta_{n,h})n^{1/2}\overline{m}_n(\theta_{n,h})$ diverges in probability to ∞. Hence, $\widehat{D}_n^{-1/2}(\theta_{n,h})n^{1/2}\overline{m}_n(\theta_{n,h})$ does not converge in distribution to a proper finite random vector and the continuous mapping theorem cannot be applied to obtain the asymptotic distribution of the right-hand side of (S1.20). The verification of Assumption B0 in AG4 avoids this problem by (i) considering a transformation of $\widehat{D}_n^{-1/2}(\theta_{n,h})n^{1/2}\overline{m}_n(\theta_{n,h})$ that converges in distribution even if some elements of h_1 equal ∞, (ii) writing the right-hand side of (S1.20) as a continuous function of this transformation, and (iii) applying the continuous mapping theorem to the transformation.

We now have

$$
\begin{aligned}
\text{(S1.21)}\quad & \liminf_{n\to\infty} P_{\gamma_{n,h}}\big(T_n(\theta_{n,h}) \le \widehat{c}_n(\theta_{n,h}, 1-\alpha)\big) \\
& \ge \liminf_{n\to\infty} P_{\gamma_{n,h}}\big(T_n(\theta_{n,h}) \le c_n^*\big) \\
& \ge J_h\big(c_{\pi^*}(1-\alpha)-\big),
\end{aligned}
$$

where $J_h(x-)$ denotes the limit from the left of $J_h(\cdot)$ at x, the first inequality holds because $\widehat{c}_n(\theta_{n,h}, 1-\alpha) \ge c_n^*$ a.s., and the second inequality holds by part (a) of the lemma and (S1.19).

Suppose $c_{\pi^*}(1-\alpha) > 0$. Then

$$
\text{(S1.22)}\quad J_h\big(c_{\pi^*}(1-\alpha)-\big) = J_h\big(c_{\pi^*}(1-\alpha)\big) \ge 1-\alpha
$$

and part (b) of the lemma holds, where the equality holds because $J_h(x)$ is continuous for all $x > 0$ by Assumption 2(a) and the inequality holds by (S1.18).

Next, suppose $c_{\pi^*}(1-\alpha) = 0$. This implies that $c_h(1-\alpha) = 0$ by (S1.18) and Assumption 1(c). The conditions $c_h(1-\alpha) = 0$ and $0 < \alpha < 1/2$ are consistent with Assumption 2(c) only if $v = 0$. Given $v = 0$, under $\{\gamma_{n,h} : n \ge 1\}$, we have

$$
\begin{aligned}
\text{(S1.23)}\quad & P_{\gamma_{n,h}}\big(T_n(\theta_{n,h}) \le 0\big) \\
& = P_{\gamma_{n,h}}\big(n^{1/2}\overline{m}_{n,j}(\theta_{n,h})/\sigma_{F_{n,h},j}(\theta_{n,h}) \ge 0 \text{ for all } j = 1, \dots, p\big) \\
& = P_{\gamma_{n,h}}\big(A_{n,j} + n^{1/2}\gamma_{n,h,1,j} \ge 0 \text{ for all } j = 1, \dots, p\big) \\
& \to P\big(\big[\Omega_{h_{2,2}}^{1/2} Z^*\big]_j + h_{1,j} \ge 0 \text{ for all } j = 1, \dots, p\big) \\
& = P\big(S\big(\Omega_{h_{2,2}}^{1/2} Z^* + h_1, \Omega_{h_{2,2}}\big) \le 0\big) \\
& = J_h(0) \ge J_h\big(c_h(1-\alpha)\big) \ge 1-\alpha,
\end{aligned}
$$

where the first equality holds by Assumptions 1(b) and 3, the second equality and the convergence hold by (S1.11), the third equality holds by Assumption 3, the fourth equality holds by the definition of J_h, the first inequality holds because $c_h(1-\alpha) \ge 0$ (note that $c_h(1-\alpha) = 0$ here, but the argument in (S1.23) is applied below to a case in which one only knows that $c_h(1-\alpha) \ge 0$), and the second inequality holds by the definition of $c_h(1-\alpha)$. This completes the proof of part (b).

The proof of part (c) is the same as that for parts (a) and (b) with w_n in place of n. *Q.E.D.*

PROOF OF LEMMA 3: Conditions (i)–(iii) of Lemma 2 hold with $\gamma_{n,h} = \gamma^*$ for all n, $h = h^*$, and $\pi = h^*$ because $\kappa_n^{-1} n^{1/2} \to \infty$ by Assumption GMS4. Each element of π_1 is either zero or infinity. Thus, the vector π^* that depends on π and is defined preceding Lemma 2 equals π. Now, (S1.13) in the proof of Lemma 2 applies with $\theta_{n,h} = \theta^*$, $\pi_1 = h_1^*$, and $\Omega_{\pi_{2,2}} = \Omega_{h_{2,2}^*}$.

Equation (S1.14) applies (with the first quantity on the left-hand side deleted) for all $j = 1, \ldots, p$ for which $\pi_{1,j} = 0$. In addition, we have that as $(\xi, \Omega) \to ((\pi_1, 0_v), \Omega_{h_{2,2}^*})$, $\varphi_j(\xi, \Omega) \to \infty$ a.s. $[Z^*]$ for all $j = 1, \ldots, p$ for which $\pi_{1,j} = \infty$ by Assumption GMS3. Given these results, (S1.15) and (S1.16) hold with $\varphi^*(\xi, \Omega)$ and $S(\Omega_0^{1/2} Z^* + \varphi^*((\pi_1, 0_v), \Omega_0), \Omega_0)$ replaced by $\varphi(\xi, \Omega)$ and $S(\Omega_0^{1/2} Z^* + \varphi((h_1^*, 0_v), \Omega_{h_{2,2}^*}), \Omega_{h_{2,2}^*})$, respectively, and the second equality in (S1.16) holds because $c_{h^*}(1-\alpha) = c_{\pi^*}(1-\alpha) > 0$. (The case $c_{h^*}(1-\alpha) = 0$ does not occur because the df of $S(\Omega_{h_{2,2}^*}^{1/2} Z^* + (h_1^*, 0_v), \Omega_{h_{2,2}^*})$ at $x < 0$ is zero by Assumption 1(c), the df at $x = c_{h^*}(1-\alpha) = 0$ is zero by continuity (Assumption 7), the latter implies that the df is less than $1-\alpha$ for $x > 0$, and the latter implies that $c_{h^*}(1-\alpha) > 0$.) The remainder of the proof of part (a) is the same as that given in (S1.17) but with $\widehat{c}_n(\theta^*, 1-\alpha)$ in place of c_n^*.

To prove part (b), we note that the asymptotic distribution of $T_n(\theta^*)$ is $S(\Omega_{h_{2,2}^*}^{1/2} Z^* + (h_1^*, 0_v), \Omega_{h_{2,2}^*})$ under $\{\gamma^* : n \geq 1\}$ by the verification of Assumption B0 in AG1, see (S1.19) and the discussion following it. The df of $S(\Omega_{h_{2,2}^*}^{1/2} Z^* + (h_1^*, 0_v), \Omega_{h_{2,2}^*})$ is continuous and strictly increasing at $c_{h^*}(1-\alpha) > 0$ by Assumption 2(a) and (b) unless $v = 0$ and $h_1^* = \infty^p$. The latter does not hold by the argument given in the proof of Lemma 2 just below (S1.16) because $c_{h^*}(1-\alpha) > 0$. These results and $\widehat{c}_n(\theta^*, 1-\alpha) \to_p c_{h^*}(1-\alpha)$ establish part (b). *Q.E.D.*

We now restate and prove the following lemma.

LEMMA 1: *The functions* $S_1(m, \Sigma)$–$S_3(m, \Sigma)$ *satisfy Assumptions* 1–6 *with* $\Psi = \Psi_1$ *for* $S_1(m, \Sigma)$ *and* $S_3(m, \Sigma)$, *and with* $\Psi = \Psi_2$ for $S_2(m, \Sigma)$.

PROOF: Assumptions 1–4 hold by Lemma 1 of AG4. For S_1, Assumption 5(a) holds by the same arguments as for Assumption 2 given in the proof of Lemma 1 of AG4. Assumption 5(b) holds with a nonstrict inequality by Assumption 1(a) and the fact that $Z + (m_1, 0_v)$ is stochastically strictly increasing in $m_1 \in R_{+,\infty}^p$. Assumption 5(b) holds for S_1 with a strict inequality because $S_1(Z + (m_1^*, 0_v), \Omega)$ is strictly stochastically less than $S_1(Z + (m_1, 0_v), \Omega)$ on R_+ for $m_1 < m_1^*$. Assumption 6 holds immediately for S_1 with $\chi = 2$.

For S_2, Assumption 5(a)(i) and (ii) hold by the same arguments as for Assumption 2(a) and (b) given in the proof of Lemma 1 of AG4. Assumption 5(b) holds for S_2 by the same argument as for S_1. Assumption 6 holds immediately for S_2 with $\chi = 2$. The verification of Assumptions 1–6 for S_3 is essentially the same as that for S_1. *Q.E.D.*

S2. PROOFS OF RESULTS FOR LOCAL ALTERNATIVES

In this section, we restate and prove the following two theorems.

THEOREM 2: *Under Assumptions* 1–5, LA1, *and* LA2, *the following statements hold*:

(a) $\lim_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > \widehat{c}_n(\theta_{n,*}, 1-\alpha)) = 1 - J_{h_1,\lambda}(c_{\pi_1}(\varphi, 1-\alpha))$ *provided Assumptions* GMS2, GMS3, LA4, *and* LA5 *hold*.

(b) $\lim_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > c_{n,b}(\theta_{n,*}, 1-\alpha)) = 1 - J_{h_1,\lambda}(c_{g_1,0_d}(1-\alpha))$ *provided Assumption* LA6 *holds*.

(c) $\lim_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > c(\widehat{\Omega}_n(\theta_{n,*}), 1-\alpha)) = 1 - J_{h_1,\lambda}(c_{0_p,0_d}(1-\alpha))$.

THEOREM 3: *Under Assumptions* 1–5, LA1–LA4, LA6, GMS2–GMS3, *and* GMS5–GMS6, *the following assertions are valid*:

(a) $\liminf_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > \widehat{c}_n(\theta_{n,*}, 1-\alpha)) \geq \lim_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > c_{n,b}(\theta_{n,*}, 1-\alpha))$ *with strict inequality whenever* $g_{1,j} < \infty$ *and* $\pi_{1,j} = \infty$ *for some* $j = 1, \ldots, p$ *and* $c_{g_1,0_d}(1-\alpha) > 0$.

(b) $\liminf_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > \widehat{c}_n(\theta_{n,*}, 1-\alpha)) \geq \lim_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > c(\widehat{\Omega}_n(\theta_{n,*}), 1-\alpha))$ *with strict inequality whenever* $\pi_{1,j} = \infty$ *for some* $j = 1, \ldots, p$.

(c) $\lim_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > c_{n,b}(\theta_{n,*}, 1-\alpha)) \geq \lim_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > c(\widehat{\Omega}_n(\theta_{n,*}), 1-\alpha))$ *with strict inequality whenever* $g_1 > 0_p$, *where Assumptions* GMS2, GMS3, GMS5, GMS6, *and* LA4 *are not needed for this result*.

Theorem 2 follows immediately from Lemmas 4 and 5 below. Theorem 3(a) and (c) do likewise from Lemmas 5–8. Theorem 3(b) follows from Lemmas 6 and 7, where one takes $g_1 = 0_p$ in Lemma 7 and one notes that $c_{0_p,0_d}(1-\alpha) > 0$ by Assumption 2(c) and $\alpha \in (0, 1/2)$.

In each of Lemmas 4–8, the parameter space $\mathcal{F}$ for (θ, F) is assumed to satisfy

$$
\text{(S2.1)} \quad
\begin{aligned}
&\text{(i)} && \theta \in \Theta, \\
&\text{(ii)} && E_F m_j(W_i, \theta) \geq 0 \quad \text{for} \quad j = 1, \ldots, p, \\
&\text{(iii)} && E_F m_j(W_i, \theta) = 0 \quad \text{for} \quad j = p+1, \ldots, k, \\
&\text{(iv)} && \{W_i : i \geq 1\} \text{ are i.i.d. under } F, \\
&\text{(v)} && \sigma^2_{F,j}(\theta) = \mathrm{Var}_F(m_j(W_i, \theta)) \in (0, \infty) \quad \text{for} \quad j = 1, \ldots, k, \\
&\text{(vi)} && \mathrm{Corr}_F(m(W_i, \theta)) \in \Psi, \\
&\text{(vii)} && E_F|m_j(W_i, \theta)/\sigma_{F,j}(\theta)|^{2+\delta} \leq M \quad \text{for} \quad j = 1, \ldots, k.
\end{aligned}
$$

In the lemmas that involve subsampling, it is assumed that $b \to \infty$ and $b/n \to 0$ as $n \to \infty$. Let $Z \sim N(0_k, \Omega_0)$.

LEMMA 4: *Under Assumptions* 1–3, 5(a), GMS2, GMS3, LA1, LA2, LA4, *and* LA5, *the following equations hold*:

(a) $T_n(\theta_{n,*}) \to_d S(Z + (h_1, 0_v) + \Pi_0\lambda, \Omega_0) \sim J_{h_1,\lambda}$.

(b) $\widehat{c}_n(\theta_{n,*}, 1-\alpha) \to_p c_{\pi_1}(\varphi, 1-\alpha)$.

(c) $\lim_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > \widehat{c}_n(\theta_{n,*}, 1-\alpha)) = 1 - J_{h_1,\lambda}(c_{\pi_1}(\varphi, 1-\alpha))$.

LEMMA 5: *Under Assumptions* 1–3, 5(a), LA1, LA2, *and* LA6, *the following results hold*:

(a) $T_n(\theta_{n,*}) \to_d S(Z + (h_1, 0_v) + \Pi_0\lambda, \Omega_0) \sim J_{h_1,\lambda}$.

(b) $T_b(\theta_{n,*}) \to_d S(Z + (g_1, 0_v), \Omega_0) \sim J_{g_1,0_d}$.

(c) $c_{n,b}(\theta_{n,*}, 1-\alpha) \to c_{g_1,0_d}(1-\alpha)$.

(d) $\lim_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > c_{n,b}(\theta_{n,*}, 1-\alpha)) = 1 - J_{h_1,\lambda}(c_{g_1,0_d}(1-\alpha))$.

LEMMA 6: *Under Assumptions* 1, 4, 5(a), LA1, *and* LA2, *the following results hold*:

(a) $T_n(\theta_{n,*}) \to_d S(Z + (h_1, 0_v) + \Pi_0\lambda, \Omega_0) \sim J_{h_1,\lambda}$.

(b) $c(\widehat{\Omega}_n(\theta_{n,*}), 1-\alpha) \to c_{0_p,0_d}(1-\alpha)$.

(c) $\lim_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > c(\widehat{\Omega}_n(\theta_{n,*}), 1-\alpha)) = 1 - J_{h_1,\lambda}(c_{0_p,0_d}(1-\alpha))$.

The next lemma uses the following notation. Let $g_1 = (g_{1,1}, \ldots, g_{1,p})'$ be as in Assumption LA6. Let $\pi^{**}_{1,j} = \infty$ if $\pi_{1,j} = \infty$ and let $\pi^{**}_{1,j} = 0$ if $\pi_{1,j} < \infty$ for $j = 1, \ldots, p$. As defined, $\pi^{**}_1 = (\pi^{**}_{1,1}, \ldots, \pi^{**}_{1,p})' \le h_1$. Let $\pi^{**} = (\pi^{**}_1, \pi^{**}_2) \in R^p_{+,\infty} \times R^q_{[\pm\infty]}$, where $\pi^{**}_2 = (\pi^{**}_{2,1}, \pi^{**}_{2,2})$, $\pi^{**}_{2,1} = \theta_0$, where θ_0 is as in Assumption LA1(a), and $\pi^{**}_{2,2} = \mathrm{vech}_*(\Omega_0)$ for the $k \times k$ correlation matrix $\Omega_0 = \Omega(\theta_0, F_0)$ determined by Assumption LA1(a). Let $c_{\pi^{**}_1}(1-\alpha)$ denote the $1-\alpha$ quantile of $S(\Omega_0^{1/2} Z^* + (\pi^{**}_1, 0_v), \Omega_0)$, where $Z^* \sim N(0_k, I_k)$ and, by definition, if $\pi^{**}_{1,j} = \infty$, then the jth element of $\Omega_0^{1/2} Z^* + (\pi^{**}_1, 0_v)$ equals ∞ for $j = 1, \ldots, p$.

LEMMA 7: *Under Assumptions* 1–3, 5(a), LA1–LA4, LA6, GMS2, GMS3, GMS5, *and* GMS6, *the following statements hold*:

(a) *If* $c_{\pi^{**}_1}(1-\alpha) > 0$, $\widehat{c}_n(\theta_{n,*}, 1-\alpha) \le c^{**}_n$ *a.s. for all* n *for a sequence of random variables* $\{c^{**}_n : n \ge 1\}$ *that satisfies* $c^{**}_n \to_p c_{\pi^{**}_1}(1-\alpha)$.

(b) $\liminf_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > \widehat{c}_n(\theta_{n,*}, 1-\alpha)) \ge 1 - J_{h_1,\lambda}(c_{\pi^{**}_1}(1-\alpha))$.

(c) $c_{g_1,0_d}(1-\alpha) \ge c_{\pi^{**}_1}(1-\alpha)$ *with strict inequality whenever* $g_{1,j} < \infty$ *and* $\pi_{1,j} = \infty$ *for some* $j = 1, \ldots, p$ *and* $c_{g_1,0_d}(1-\alpha) > 0$.

(d) $1 - J_{h_1,\lambda}(c_{\pi^{**}_1}(1-\alpha)) \ge 1 - J_{h_1,\lambda}(c_{g_1,0_d}(1-\alpha))$ *with strict inequality whenever* $g_{1,j} < \infty$ *and* $\pi_{1,j} = \infty$ *for some* $j = 1, \ldots, p$ *and* $c_{g_1,0_d}(1-\alpha) > 0$.

LEMMA 8: *Under Assumptions* 1–5, LA1–LA3, *and* LA6, *assertions* (a)–(d) *are valid*:

(a) $c_{0_p,0_d}(1-\alpha) \ge c_{g_1,0_d}(1-\alpha)$.

(b) $J_{h_1,\lambda}(c_{g_1,0_d}(1-\alpha)) \geq J_{h_1,\lambda}(c_{0_p,0_d}(1-\alpha))$.
(c) $c_{0_p,0_d}(1-\alpha) > c_{g_1,0_d}(1-\alpha)$ *unless* $g_1 = 0_p$.
(d) $1 - J_{h_1,\lambda}(c_{g_1,0_d}(1-\alpha)) > 1 - J_{h_1,\lambda}(c_{0_p,0_d}(1-\alpha))$ *unless* $g_1 = 0_p$.

PROOF OF LEMMA 4: To prove part (a), by using element-by-element mean-value expansions about $\theta = \theta_n$ and Assumptions LA1 and LA2, we obtain

$$\begin{aligned} &D^{-1/2}(\theta_{n,*}, F_n) E_{F_n} m(W_i, \theta_{n,*}) \\ &\quad = D^{-1/2}(\theta_n, F_n) E_{F_n} m(W_i, \theta_n) + \Pi(\widetilde{\theta}_n, F_n)(\theta_{n,*} - \theta_n), \\ &n^{1/2} D^{-1/2}(\theta_{n,*}, F_n) E_{F_n} m(W_i, \theta_{n,*}) \to (h_1, 0_v) + \Pi_0 \lambda, \end{aligned} \tag{S2.2}$$

where $D(\theta, F) = \mathrm{Diag}\{\sigma^2_{F,1}(\theta), \ldots, \sigma^2_{F,k}(\theta)\}$, $\widetilde{\theta}_n$ may differ across rows of $\Pi(\widetilde{\theta}_n, F_n)$, $\widetilde{\theta}_n$ lies between $\theta_{n,*}$ and θ_n, $\widetilde{\theta}_n \to \theta_0$, and $\Pi(\widetilde{\theta}_n, F_n) \to \Pi_0$.

Next, under $\{(\theta_n, F_n) \in \mathcal{F} : n \geq 1\}$ as in Assumption LA1, we have

$$\begin{aligned} &\text{(i)} \quad A^0_n = (A^0_{n,1}, \ldots, A^0_{n,k})' \to_d Z \sim N(0_k, \Omega_0) \quad \text{as} \quad n \to \infty, \\ &\qquad \text{where} \quad A^0_{n,j} = n^{1/2}\big(\overline{m}_{n,j}(\theta_{n,*}) - E_{F_n}\overline{m}_{n,j}(\theta_{n,*})\big)/\sigma_{F_n,j}(\theta_{n,*}), \\ &\text{(ii)} \quad \widehat{\sigma}_{n,j}(\theta_{n,*})/\sigma_{F_n,j}(\theta_{n,*}) \to_p 1 \quad \text{as} \quad n \to \infty \quad \text{for} \quad j = 1, \ldots, k, \\ &\text{(iii)} \quad \widehat{D}^{-1/2}_n(\theta_{n,*}) \widehat{\Sigma}_n(\theta_{n,*}) \widehat{D}^{-1/2}_n(\theta_{n,*}) \to_p \Omega_0 \quad \text{as} \quad n \to \infty, \end{aligned} \tag{S2.3}$$

where result (i) holds by the Cramér–Wold device and the Liapounov triangular array CLT for rowwise i.i.d. random variables with mean 0 and variance 1 using condition (iv) of (S2.1) and Assumptions LA1(a) and LA1(c), and results (ii) and (iii) hold by standard arguments using a weak law of large numbers for rowwise i.i.d. random variables with variance 1 by condition (iv) of (S2.1) and Assumptions LA1(a) and LA1(c). Note that results (i)–(iii) of (S2.3) do not hold by (S1.11) because the functions are evaluated at $\theta_{n,*}$ but the true value is θ_n.

For the same reason as described above following (S1.20), to obtain the asymptotic distribution of $T_n(\theta_{n,*})$, we use the same type of argument as in the verification of Assumption B0 in AG4. Let $G(\cdot)$ be a strictly increasing continuous df on R, such as the standard normal df. Using (S2.2) and (S2.3), for $j = 1, \ldots, k$, we have

$$\begin{aligned} G^0_{n,j} &= G\big(\widehat{\sigma}^{-1}_{n,j}(\theta_{n,*}) n^{1/2} \overline{m}_{n,j}(\theta_{n,*})\big) \\ &= G\big(\widehat{\sigma}^{-1}_{n,j}(\theta_{n,*}) \sigma_{F_n,j}(\theta_{n,*}) \big[A^0_{n,j} + n^{1/2} \sigma^{-1}_{F_n,j}(\theta_{n,*}) E_{F_n} m_j(W_i, \theta_{n,*})\big]\big), \\ G^0_{n,j} &\begin{cases} \to_p 1, & \text{if } j \leq p \text{ and } h_{1,j} = \infty, \\ \to_d G(Z_j + h_{1,j} + \Pi'_{0,j}\lambda), & \text{if } j \leq p \text{ and } h_{1,j} < \infty, \\ \to_d G(Z_j + \Pi'_{0,j}\lambda), & \text{if } j = p+1, \ldots, k, \end{cases} \end{aligned} \tag{S2.4}$$

$$G_n^0 = (G_{n,1}^0, \dots, G_{n,k}^0)$$
$$\to_d G_\infty^0 = (G(Z_1 + h_{1,1} + \Pi_{0,1}'\lambda), \dots, G(Z_k + \Pi_{0,k}'\lambda))',$$

where $Z = (Z_1, \dots, Z_k)'$ and $Z_j + h_{1,j} + \Pi_{0,j}'\lambda = \infty$ by definition if $h_{1,j} = \infty$. Now, the same argument as in the verification of Assumption B0 in AG4 gives

$$T_n(\theta_{n,*}) \to_d S(Z + (h_1, 0_v) + \Pi_0\lambda, \Omega_0) \sim J_{h_1,\lambda}. \tag{S2.5}$$

In short, the idea behind the argument is to write the right-hand side of (S1.20) as a continuous function of G_n^0 and $\widehat{D}_n^{-1/2}(\theta_{n,*})\widehat{\Sigma}_n(\theta_{n,*})\widehat{D}_n^{-1/2}(\theta_{n,*})$ and apply the continuous mapping theorem. This completes the proof of part (a).

To prove part (b), by the mean-value expansions in (S2.2), Assumptions LA1(a), LA2, and LA4, and $\kappa_n \to \infty$, we obtain

$$\kappa_n^{-1} n^{1/2} D^{-1/2}(\theta_{n,*}, F_n) E_{F_n} m(W_i, \theta_{n,*}) \to (\pi_1, 0_v). \tag{S2.6}$$

This leads to

$$\begin{aligned} \kappa_n^{-1} n^{1/2} \widehat{D}_n^{-1/2}(\theta_{n,*}) \overline{m}_n(\theta_{n,*}) &= \kappa_n^{-1} \widehat{D}_n^{-1/2}(\theta_{n,*}) D^{1/2}(\theta_{n,*}, F_n) \\ &\quad \times \big(A_n^0 + n^{1/2} D^{-1/2}(\theta_{n,*}, F_n) E_{F_n} m(W_i, \theta_{n,*})\big) \\ &\to_p (\pi_1, 0_v), \end{aligned} \tag{S2.7}$$

where the equality holds by the definition of A_n^0 in (S2.3), and the convergence holds by (S2.6), conditions (i) and (ii) of (S2.3), and $\kappa_n \to \infty$. Equation (S2.7) and condition (iii) of (S2.3) yield

$$(\xi_n(\theta_{n,*}), \widehat{\Omega}_n(\theta_{n,*})) \to_p ((\pi_1, 0_v), \Omega_0). \tag{S2.8}$$

For $j = 1, \dots, k$, as $(\xi, \Omega) \to ((\pi_1, 0_v), \Omega_0)$,

$$\varphi_j(\xi, \Omega) \to \varphi_j((\pi_1, 0_v), \Omega_0), \tag{S2.9}$$

because by Assumption LA5(a), $\pi_1 \in C(\varphi)$, which yields (S2.9) by Assumption GMS3 if $\pi_{1,j} = \infty$ and yields (S2.9) by the definition of $C(\varphi)$ otherwise.

Assumption 1(d) and (S2.9) give that for x in a neighborhood of $c_{\pi_1}(\varphi, 1 - \alpha)$, as $(\xi, \Omega) \to ((\pi_1, 0_v), \Omega_0)$,

$$\begin{aligned} & S\big(\Omega^{1/2} Z^* + \varphi(\xi, \Omega), \Omega\big) \\ &\quad \to S\big(\Omega_0^{1/2} Z^* + \varphi((\pi_1, 0_v), \Omega_0), \Omega_0\big) \quad \text{a.s. } [Z^*], \\ & 1\big(S\big(\Omega^{1/2} Z^* + \varphi(\xi, \Omega), \Omega\big) \le x\big) \\ &\quad \to 1\big(S\big(\Omega_0^{1/2} Z^* + \varphi((\pi_1, 0_v), \Omega_0), \Omega_0\big) \le x\big) \quad \text{a.s. } [Z^*], \end{aligned} \tag{S2.10}$$

$$P\big(S\big(\Omega^{1/2}Z^* + \varphi(\xi, \Omega), \Omega\big) \le x\big)$$
$$\to P\big(S\big(\Omega_0^{1/2}Z^* + \varphi((\pi_1, 0_v), \Omega_0), \Omega_0\big) \le x\big).$$

The third convergence result of (S2.10) holds by the second result and the bounded convergence theorem, and the second convergence result of (S2.10) follows from the first result provided $P(S(\Omega_0^{1/2}Z^* + \varphi((\pi_1, 0_v), \Omega_0), \Omega_0) = x) = 0$, which holds by Assumption LA5(b).

Given (S2.8) and (S2.10), the remainder of the proof of part (b) is the same as that given in the paragraph containing (S1.17) using Lemma 5 of AG1.

Part (c) of the lemma holds by parts (a) and (b) and Assumption LA5(b). *Q.E.D.*

PROOF OF LEMMA 5: Part (a) holds by Lemma 4(a).

To prove part (b), by the mean-value expansions in (S2.2), Assumptions LA1(a), LA2, and LA6, and $b/n \to 0$, we obtain

$$b^{1/2}D^{-1/2}(\theta_{n,*}, F_n)E_{F_n}m(W_i, \theta_{n,*}) \to (g_1, 0_v). \tag{S2.11}$$

Using (S2.11) and an analogous argument to that given in the proof of Lemma 4(a) with $n^{1/2}$ replaced by $b^{1/2}$ in (S2.4), we have

$$T_b(\theta_{n,*}) \to_d S(Z + (g_1, 0_v), \Omega_0) \sim J_{g_1,0_d}, \tag{S2.12}$$

which proves part (b).

To establish Lemma 5(c) and (d), we apply Lemma 5 of AG1. We verify conditions (i)–(iii) of Lemma 5 of AG1 as follows. Lemma 5(a) of the present paper implies condition (ii). To verify condition (i), Lemma 5(b) of the present paper and identical distributions for $\{W_i : i \le n\}$ imply that

$$E_{F_n}U_{n,b}(\theta_{n,*}, x) = P_{F_n}(T_b(\theta_{n,*}) \le x) \to J_{g_1,0_d}(x) \tag{S2.13}$$

for all continuity points x of $J_{g_1,0_d}$. In addition, $\mathrm{Var}_{F_n}(U_{n,b}(\theta_{n,*}, x)) \to 0$ by a U-statistic inequality of Hoeffding, as in Politis, Romano, and Wolf (1999, p. 44), using the i.i.d. property of $\{W_i : i \le n\}$ and the boundedness of $U_{n,b}(\theta_{n,*}, x)$. This and (S2.13) give

$$U_{n,b}(\theta_{n,*}, x) \to_p J_{g_1,0_d}(x) \tag{S2.14}$$

for all continuity points x of $J_{g_1,0_d}$, which verifies condition (i) of Lemma 5 of AG1.

To verify condition (iii) of Lemma 5 of AG1, we need to show

$$J_{g_1,0_d}\big(c_{g_1,0_d}(1-\alpha) + \varepsilon\big) > 1 - \alpha \quad \text{for all } \varepsilon > 0. \tag{S2.15}$$

When $v = 0$ and $g_1 = \infty^p$, $S(Z + (g_1, 0_v), \Omega_0) = S(\infty^p, \Omega_0) = 0$ using Assumption 3. In consequence, $J_{g_1,0_d}(x) = 1$ for all $x \ge 0$, $c_{g_1,0_d}(1-\alpha) = 0$, and

(S2.15) holds for $\alpha > 0$. Now, suppose $v \geq 1$ or $g_1 \neq \infty^p$. Then, by Assumption 2(b), $J_{g_1,0_d}(x)$ is strictly increasing for $x > 0$. Using this, we have (i) if $c_{g_1,0_d}(1-\alpha) > 0$, then $J_{g_1,0_d}(x)$ is strictly increasing at $x = c_{g_1,0_d}(1-\alpha)$ and (S2.15) holds, (ii) if $c_{g_1,0_d}(1-\alpha) = 0$, then $J_{g_1,0_d}(0) \geq 1-\alpha$ (by the definition of $c_{g_1,0_d}(1-\alpha)$), (iii) if $c_{g_1,0_d}(1-\alpha) = 0$ and $J_{g_1,0_d}(0) \geq 1-\alpha$, then $J_{g_1,0_d}(x) > 1-\alpha$ for all $x > 0$ and (S2.15) holds (otherwise, $J_{g_1,0_d}(x) = 1-\alpha$ for some $x > 0$ and $J_{g_1,0_d}(x/2) = 1-\alpha$ since $J_{g_1,0_d}$ is nondecreasing, which contradicts the fact that $J_{g_1,0_d}(x)$ is strictly increasing for $x > 0$). Hence, (S2.15) holds.

Lemma 5 of AG1 establishes Lemma 5(c) of the present paper and shows that $\lim_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > c_{n,b}(\theta_{n,*}, 1-\alpha)) \in [1 - J_{h_1,\lambda}(c_{g_1,0_d}(1-\alpha)), 1 - J_{h_1,\lambda}(c_{g_1,0_d}(1-\alpha)-)]$. If $c_{g_1,0_d}(1-\alpha) > 0$, then by Assumption 5(a)(i), $J_{h_1,\lambda}$ is continuous at $c_{g_1,0_d}(1-\alpha)$ and the result of Lemma 5(d) holds. Assumption 1(c) implies that $c_{0_p,0_d}(1-\alpha) \geq 0$. The conditions $c_{g_1,0_d}(1-\alpha) = 0$ and $0 < \alpha < 1/2$ are consistent with Assumption 2(c) only if $v = 0$. Given $v = 0$ and $c_{g_1,0_d}(1-\alpha) = 0$, we use the argument given in (S1.23) to establish Lemma 5(d) with $\theta_{n,h}$, $P_{\gamma_{n,h}}$, $S(\Omega_{h_{2,2}}^{1/2} Z^* + h_1, \Omega_{h_{2,2}})$, J_h, and $c_h(1-\alpha)$ replaced by $\theta_{n,*}$, P_{F_n}, $S(Z + h_1 + \Pi_0\lambda, \Omega_0)$, $J_{h_1,\lambda}$, and $c_{h_1,\lambda}(1-\alpha)$, respectively. *Q.E.D.*

PROOF OF LEMMA 6: Part (a) holds by Lemma 4(a) because Assumptions 2 and 3 are not used in the proof of Lemma 4(a).

By standard arguments using a weak law of large numbers for rowwise i.i.d. triangular arrays and Assumption LA1(c), we have

$$\begin{aligned}
\text{(S2.16)} \quad & D^{-1/2}(\theta_{n,*}, F_n)\widehat{\Sigma}_n(\theta_{n,*})D^{-1/2}(\theta_{n,*}, F_n) \\
&\quad - D^{-1/2}(\theta_{n,*}, F_n)\,\mathrm{Var}_{F_n}(m(W_i, \theta_{n,*}))D^{-1/2}(\theta_{n,*}, F_n) \to_p 0_{k\times k}, \\
& \widehat{D}_n^{-1/2}(\theta_{n,*})D^{1/2}(\theta_{n,*}, F_n) - I_k \to_p 0_{k\times k}.
\end{aligned}$$

In consequence, $\widehat{\Omega}_n(\theta_{n,*}) - \Omega(\theta_{n,*}, F_n) \to_p 0_{k\times k}$. This and Assumptions LA1(a) and LA1(c) give $\widehat{\Omega}_n(\theta_{n,*}) \to_p \Omega_0$. The latter and Assumption 4(b) yield $c(\widehat{\Omega}_n(\theta_{n,*}), 1-\alpha) \to_p c(\Omega_0, 1-\alpha)$. This establishes Lemma 6(b) because $c(\Omega_0, 1-\alpha) = c_{0_p,0_d}(1-\alpha)$ by definition.

If $c_{0_p,0_d}(1-\alpha) > 0$, Lemma 6(c) holds by parts (a) and (b) and Assumption 5(a)(i). Assumption 1(c) implies that $c_{0_p,0_d}(1-\alpha) \geq 0$. The conditions $c_{0_p,0_d}(1-\alpha) = 0$ and $0 < \alpha < 1/2$ are consistent with Assumption 2(c) only if $v = 0$ (because $c_{0_p,0_d}(1-\alpha)$ is the $1-\alpha$ quantile of $S(Z, \Omega_0)$). Given $v = 0$ and $c_{0_p,0_d}(1-\alpha) = 0$, Lemma 6(c) holds by the same argument as used to prove Lemma 5(d) when $c_{g_1,0_d}(1-\alpha) = 0$. *Q.E.D.*

PROOF OF LEMMA 7: First we prove part (a). By assumption, $c_{\pi_1^{**}}(1-\alpha) > 0$. For $(\xi, \Omega) \in R^k \times \Psi$, let $\varphi^{**}(\xi, \Omega)$ denote the k-vector whose jth element is

$$\text{(S2.17)} \quad \varphi_j^{**}(\xi, \Omega) = \begin{cases} 0, & \text{if } \pi_{1,j} < \infty \text{ and } j = 1, \ldots, p, \\ \varphi_j(\xi, \Omega), & \text{if } \pi_{1,j} = \infty \text{ and } j = 1, \ldots, p, \\ 0, & \text{if } j = p+1, \ldots, k. \end{cases}$$

By Assumption GMS6, $\varphi_j(\xi,\Omega) \ge 0$ for $j \le p$. Hence, $\varphi^{**}(\xi,\Omega) \le \varphi(\xi,\Omega)$ and

$$\varphi^{**}(\xi_n(\theta_n), \widehat{\Omega}_n(\theta_n)) \le \varphi(\xi_n(\theta_n), \widehat{\Omega}_n(\theta_n)) \quad \text{a.s. } [Z^*] \text{ for all } n. \tag{S2.18}$$

Let c_n^{**} denote the $1-\alpha$ quantile of the conditional df of $S(\widehat{\Omega}_n^{1/2}(\theta_n)Z^* + \varphi^{**}(\xi_n(\theta_n), \widehat{\Omega}_n(\theta_n)), \widehat{\Omega}_n(\theta_n))$ given $(\xi_n(\theta_n), \widehat{\Omega}_n(\theta_n))$. Then $\widehat{c}_n(\theta_n, 1-\alpha) \le c_n^{**}$ a.s. for all n by (S2.18) and Assumption 1(a). By the same argument as in the proof of Lemma 4, (S2.8) holds.

For $j = 1, \ldots, p$ with $\pi_{1,j} < \infty$, $\varphi_j^{**}(\xi,\Omega) = 0$. For $j = 1, \ldots, p$ with $\pi_{1,j} = \infty$, as $(\xi,\Omega) \to ((\pi_1, 0_v), \Omega_0)$, we have $\varphi_j(\xi,\Omega) \to \infty$ a.s. $[Z^*]$ by Assumption GMS3. These results can be written as

$$\varphi_j^{**}(\xi,\Omega) \to \pi_{1,j}^{**} \quad \text{a.s. } [Z^*] \tag{S2.19}$$

for $j = 1, \ldots, p$ by the definition of $\pi_{1,j}^{**}$.

Assumption 1(d) and (S2.19) give that for x in a neighborhood of $c_{\pi_1^{**}}(1-\alpha)$, as $(\xi,\Omega) \to ((\pi_1, 0_v), \Omega_0)$,

$$\begin{aligned} &S\big(\Omega^{1/2}Z^* + \varphi^{**}(\xi,\Omega), \Omega\big) \to S\big(\Omega^{1/2}Z^* + (\pi_1^{**}, 0_v), \Omega_0\big) \quad \text{a.s. } [Z^*], \\ &1\big(S\big(\Omega^{1/2}Z^* + \varphi^{**}(\xi,\Omega), \Omega\big) \le x\big) \\ &\quad \to 1\big(S\big(\Omega^{1/2}Z^* + (\pi_1^{**}, 0_v), \Omega_0\big) \le x\big) \quad \text{a.s. } [Z^*], \\ &P\big(S\big(\Omega^{1/2}Z^* + \varphi^{**}(\xi,\Omega), \Omega\big) \le x\big) \\ &\quad \to P\big(S\big(\Omega^{1/2}Z^* + (\pi_1^{**}, 0_v), \Omega_0\big) \le x\big). \end{aligned} \tag{S2.20}$$

The third convergence result of (S2.20) holds by the second result and the bounded convergence theorem. The second convergence result of (S2.20) follows from the first result provided $P(S(\Omega^{1/2}Z^* + (\pi_1^{**}, 0_v), \Omega_0) = x) = 0$, which holds because $c_{\pi_1^{**}}(1-\alpha) > 0$ for the same reason that the second equality in (S1.16) holds.

Given (S2.8) and (S2.20), the remainder of the proof of part (a) is the same as that given in the paragraph containing (S1.17) using Lemma 5 of AG1.

Now we prove part (b). If $c_{\pi_1^{**}}(1-\alpha) > 0$, part (b) of the lemma holds by part (a), Lemma 4(a) (i.e., $T_n(\theta_{n,*}) \to_d J_{h_1,\lambda}$), and Assumption 5(a)(i).

Next, we prove part (b) for the case where $c_{\pi_1^{**}}(1-\alpha) = 0$. We have

$$\liminf_{n\to\infty} P_{\gamma_n}(T_n(\theta_n) \le c_n^{**}) \ge \liminf_{n\to\infty} P_{\gamma_n}(T_n(\theta_n) \le 0) \tag{S2.21}$$

because $c_n^{**} \ge 0$ by Assumption 1(c). By the definition of π_1^{**}, we have $\pi_1^{**} \le h_1$. As in (S1.18), this implies that $c_h(1-\alpha) \le c_{\pi^{**}}(1-\alpha)$ and hence $c_h(1-\alpha) = 0$ using Assumption 1(c). The conditions $c_h(1-\alpha) = 0$ and $0 < \alpha < 1/2$ are consistent with Assumption 2(c) only if $v = 0$. Given $v = 0$, we use the same argument as given in (S1.23) with $\gamma_{n,h}$, $A_{n,j}$, $h_{1,j}$, J_h, and $c_h(1-\alpha)$ replaced by γ_n, $A^0_{n,j}$, $h_{1,j}+\Pi'_{0,j}\lambda$, $J_{h_1,\lambda}$, and $c_{h_1,\lambda}(1-\alpha)$, respectively, where A^0_n is defined in (S2.3), to show that the right-hand side in (S2.21) is greater than or equal to $1-\alpha$. This completes the proof of part (b).

When the inequality is not strict, part (c) holds because (i) $\pi^{**}_1 \geq g_1$, which holds because if $\pi_{1,j}=\infty$, then $\pi^{**}_{1,j}=\infty$, and if $\pi_{1,j}<\infty$ then $g_{1,j}=0$ by Assumptions LA4 and GMS5, and $\pi^{**}_{1,j}=0$ by definition; (ii) $S(\Omega_0^{1/2}Z^*+(\pi^{**}_1,0_v),\Omega_0) \leq S(Z+(g_1,0_v),\Omega_0)$ a.s. by (i) and Assumption 1(a), and (iii) the corresponding quantiles satisfy $c_{\pi^{**}_1}(1-\alpha) \leq c_{g_1,0_d}(1-\alpha)$ by (ii).

Next, we show part (c) holds with a strict inequality when $c_{g_1,0_d}(1-\alpha)>0$, $g_{1,j}<\infty$, and $\pi_{1,j}=\infty$ for some $j=1,\ldots,p$. The latter implies that $\pi^{**}_1>g_1$. Given $\pi^{**}_1>g_1$ and $c_{g_1,0_d}(1-\alpha)>0$, Assumption 5(b) implies that

$$
\begin{aligned}
\text{(S2.22)}\quad & P\big(S(Z+(\pi^{**}_1,0_v),\Omega) \leq c_{g_1,0_d}(1-\alpha)\big) \\
& > P\big(S(Z+(g_1,0_v),\Omega) \leq c_{g_1,0_d}(1-\alpha)\big) \geq 1-\alpha,
\end{aligned}
$$

where $Z \sim N(0_k,\Omega)$. If "$v=0$ and $\pi^{**}_1=\infty^p$" does not hold, then the df of $S(Z+(\pi^{**}_1,0_v),\Omega)$ is strictly increasing for $x>0$ by Assumption 2(b). This and (S2.22) imply that $c_{\pi^{**}_1}(1-\alpha)<c_{g_1,0_d}(1-\alpha)$. If $v=0$ and $\pi^{**}_1=\infty^p$, then $S(Z+(\pi^{**}_1,0_v),\Omega)=S(Z+\infty^p,\Omega)=0$ by Assumption 1(c) and $c_{\pi^{**}_1}(1-\alpha)=0<c_{g_1,0_d}(1-\alpha)$, and the proof of part (c) is complete.

Part (d) follows immediately from part (c) when the inequality is not strict. When $c_{g_1,0_d}(1-\alpha)>0$, $g_{1,j}<\infty$, and $\pi_{1,j}=\infty$ for some $j=1,\ldots,p$, part (c) holds with a strict inequality. The latter, $c_{\pi^{**}_1}(1-\alpha)\geq 0$ (which holds by Assumption 1(c)), and $J_{h_1,\lambda}(x)$ is strictly increasing for $x>0$ (which holds by Assumption 5(a)(ii) because the caveat in Assumption 5(a)(ii) that "$v=0$ and $\ell=\infty^p$ does not occur" holds by Assumption LA3) imply that part (d) holds with a strict inequality. *Q.E.D.*

PROOF OF LEMMA 8: Part (a) holds because for $0_p \leq g_1 \in R^p_{+,\infty}$, we have

$$
\text{(S2.23)}\quad S(Z+(0_p,0_v),\Omega_0) \geq S(Z+(g_1,0_v),\Omega_0)
$$

by Assumption 1(a). Part (b) follows from part (a). To prove Lemma 8(c), note that $c_{0_p,0_d}(1-\alpha)>0$ by Assumption 2(c) and $\alpha\in(0,1/2)$. This, Assumptions 2(a) and 5(b), and $g_1>0_p$ imply that

$$
\begin{aligned}
\text{(S2.24)}\quad 1-\alpha &= P\big(S(Z+(0_p,0_v),\Omega_0) \leq c_{0_p,0_d}(1-\alpha)\big) \\
& < P\big(S(Z+(g_1,0_v),\Omega_0) \leq c_{0_p,0_d}(1-\alpha)\big),
\end{aligned}
$$

where $Z\sim N(0_k,\Omega_0)$. The latter and Assumption 2(a) prove part (c).

Lemma 8(d) holds by part (c), $c_{g_1,0_d}(1-\alpha)\geq 0$ (which holds by Assumption 1(c)), and Assumption 5(a)(ii) (because the caveat in Assumption 5(a)(ii) that "$v=0$ and $\ell=\infty^p$ does not occur" holds by Assumption LA3). *Q.E.D.*

S3. PROOF OF RESULT FOR DISTANT ALTERNATIVES

In this section, we restate and prove Theorem 4.

THEOREM 4: *Under Assumptions* 1, 3, 6, *and* DA, *the following results hold*:

(a) $\lim_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > \widehat{c}_n(\theta_{n,*}, 1-\alpha)) = 1$ *provided Assumption* GMS7 *holds*.

(b) $\lim_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > c_{n,b}(\theta_{n,*}, 1-\alpha)) = 1$.

(c) $\lim_{n\to\infty} P_{F_n}(T_n(\theta_{n,*}) > c(\widehat{\Omega}_n(\theta_{n,*}), 1-\alpha)) = 1$.

PROOF: It suffices to show that for any subsequence $\{t_n\}$ of $\{n\}$ there exists a sub-subsequence $\{s_n\}$ such that $\lim_{n\to\infty} P_{F_{s_n}}(T_{s_n}(\theta_{n,*}) > c_{1-\alpha}) = 1$, where $c_{1-\alpha} = \widehat{c}_n(\theta_{n,*}, 1-\alpha)$, $c_{n,b}(\theta_{n,*}, 1-\alpha)$, or $c(\widehat{\Omega}_n(\theta_{n,*}), 1-\alpha)$. We can take the subsequence $\{s_n\}$ to be such that $m^*_{s_n,j}/\beta_{s_n} \to e_j$ for some $e_j \in [-1, \infty]$ for $j = 1, \dots, k$ because $\{m^*_{n,j}/\beta_n : n \geq 1\}$ is a sequence of points in the set $[-1, \infty]$ by the definition of β_n. For notational simplicity, we establish the former result with s_n replaced by n and by a subsequence argument we assume without loss of generality (wlog) that

$$m^*_{n,j}/\beta_n \to e_j \quad \text{for some } e_j \in [-1, \infty] \text{ for } j = 1, \dots, k. \tag{S3.1}$$

The following equation is used in the proofs of parts (a)–(c). We have

$$\begin{aligned}
&\big(n^{1/2}\beta_n\big)^{-\chi} T_n(\theta_{n,*}) \\
&\quad = \big(n^{1/2}\beta_n\big)^{-\chi} \\
&\qquad \times S\big(\widehat{D}_n^{-1/2}(\theta_{n,*}) n^{1/2} \overline{m}_n(\theta_{n,*}), \widehat{D}_n^{-1/2}(\theta_{n,*}) \widehat{\Sigma}_n(\theta_{n,*}) \widehat{D}_n^{-1/2}(\theta_{n,*})\big) \\
&\quad = \big(n^{1/2}\beta_n\big)^{-\chi} S\big(\widehat{D}_n^{-1/2}(\theta_{n,*}) D^{1/2}(\theta_{n,*}, F_n)\big(A_n^0 + n^{1/2} m_n^*\big), \Omega_1 + o_p(1)\big) \\
&\quad = S(o_p(1) + m_n^*/\beta_n, \Omega_1 + o_p(1)) \\
&\quad \to_p S(e, \Omega_1) > 0,
\end{aligned} \tag{S3.2}$$

where A_n^0 is defined in (S2.3), $m_n^* = (m^*_{n,1}, \dots, m^*_{n,k})'$, $e = (e_1, \dots, e_k)'$, the first equality uses Assumption 1(b), the second equality holds by the definitions of A_n^0, m_n^*, and $D(\theta_{n,*}, F_n)$ and by (S2.3) (with Ω_1 in place of Ω_0 and with Assumption DA(b) used in place of Assumption LA1(a) in the proof of (S2.3)), the third equality holds by Assumptions 6 and DA(a) and (S2.3) (with the same adjustments as above), the convergence holds by Assumption 1(d) and (S3.1), and the inequality holds by Assumption 3 because for some $j^* \leq k$ the j^*th element of e, e_{j^*}, has absolute value equal to 1 and is negative if $j^* \leq p$, which implies that $e_{j^*} < 0$ if $j^* \leq p$ and $e_{j^*} \neq 0$ if $j^* \geq p+1$.

We prove part (b) first. By another subsequence argument, we can assume wlog that $\lim_{n\to\infty} b^{1/2}\beta_n$ exists and (S3.1) holds. We consider two cases:

(i) $\lim_{n\to\infty} b^{1/2}\beta_n = \infty$ and (ii) $\lim_{n\to\infty} b^{1/2}\beta_n \in [0,\infty)$. When case (i) holds, the same argument used to show (S3.2) gives

$$\text{(S3.3)} \qquad \left(b^{1/2}\beta_n\right)^{-\chi} T_b(\theta_{n,*}) \to_p S(e, \Omega_1),$$

where β_n appears, not β_b, because $m_n^*/\beta_n \to e$ under $\{F_n : n \geq 1\}$. Equation (S3.3) and $b/n \to 0$ imply that $T_b^\dagger(\theta_{n,*}) = (n^{1/2}\beta_n)^{-\chi} T_b(\theta_{n,*}) \to_p 0$.

Define $U_{n,b}^\dagger(\theta_{n,*}, x)$ as $U_{n,b}(\theta_{n,*}, x)$ is defined but with $T_{n,b,j}^\dagger(\theta_{n,*}) = (n^{1/2} \times \beta_n)^{-\chi} T_{n,b,j}(\theta_{n,*})$ in place of $T_{n,b,j}(\theta_{n,*})$. Using the result of the previous paragraph, we have $E_{F_n} U_{n,b}^\dagger(\theta_{n,*}, x) = P_{F_n}(T_b^\dagger(\theta_{n,*}) \leq x) \to 0$ for $x < 0$ and $\to 1$ for $x > 0$. In addition, $\mathrm{Var}_{F_n}(U_{n,b}^\dagger(\theta_{n,*}, x)) \to 0$ by Hoeffding's U-statistic inequality for bounded i.i.d. random variables; see Politis, Romano, and Wolf (1999, p. 44). Hence, $U_{n,b}^\dagger(\theta_{n,*}, x) \to_p 0$ for $x < 0$ and $\to_p 1$ for $x > 0$. This and Lemma 5(a) of AG1 imply that $c_{n,b}^\dagger(\theta_{n,*}, 1-\alpha) \to_p 0$, where $c_{n,b}^\dagger(\theta_{n,*}, 1-\alpha)$ is the $1-\alpha$ quantile of the rescaled subsample statistics $\{T_{n,b,j}^\dagger(\theta_{n,*}) : j = 1, \ldots, q_n\}$. The latter result and (S3.2) give

$$\begin{aligned}
\text{(S3.4)} \qquad & P_{F_n}(T_n(\theta_{n,*}) > c_{n,b}(\theta_{n,*}, 1-\alpha)) \\
&= P_{F_n}\left(\left(n^{1/2}\beta_n\right)^{-\chi} T_n(\theta_{n,*}) > \left(n^{1/2}\beta_n\right)^{-\chi} c_{n,b}(\theta_{n,*}, 1-\alpha)\right) \\
&= P_{F_n}\left(\left(n^{1/2}\beta_n\right)^{-\chi} T_n(\theta_{n,*}) > c_{n,b}^\dagger(\theta_{n,*}, 1-\alpha)\right) \\
&\to P(S(e, \Omega_1) > 0) = 1,
\end{aligned}$$

where the second equality holds because $c_{n,b}^\dagger(\theta_{n,*}, 1-\alpha) = (n^{1/2}\beta_n)^{-\chi} c_{n,b}(\theta_{n,*}, 1-\alpha)$ by the scale equivariance of quantiles and the last equality holds because $S(e, \Omega_1) > 0$ by (S3.2).

Next, suppose case (ii) holds. Then the same argument as used to show (S3.2) but with $(n^{1/2}\beta_n)^{-\chi}$ deleted gives

$$\text{(S3.5)} \qquad T_b(\theta_{n,*}) = S\left(O_p(1) + \left(b^{1/2}\beta_n\right)\beta_n^{-1} m_n^*, \Omega_1 + o_p(1)\right) = O_p(1),$$

where the second equality uses Assumption 1(a). Hence, $T_b^\dagger(\theta_{n,*}) = (n^{1/2} \times \beta_n)^{-\chi} T_b(\theta_{n,*}) \to_p 0$. Given this, the remainder of the proof is the same as in case (i).

Next, we prove part (c). We have $\widehat{\Omega}_n(\theta_{n,*}) \to_p \Omega_1$ because (S2.16) of the paper holds by the argument given for (S2.16) but using condition (vii) of (2.2) and Assumption DA(b). This and Assumption 4(b) imply that $c(\widehat{\Omega}_n(\theta_{n,*}), 1-\alpha) \to_p c(\Omega_1, 1-\alpha)$. Combining the latter with (S3.2) and (S3.4), with $c_{n,b}(\theta_{n,*}, 1-\alpha)$ replaced by $c(\widehat{\Omega}_n(\theta_{n,*}), 1-\alpha)$ in (S3.4), gives the desired result.

Finally, we prove part (a). By (S3.2) and the first equality of (S3.4) with $c_{n,b}(\theta_{n,*}, 1-\alpha)$ replaced by $\widehat{c}_n(\theta_{n,*}, 1-\alpha)$, it suffices to show that $(n^{1/2}\beta_n)^{-\chi}\widehat{c}_n(\theta_{n,*}, 1-\alpha) = o_p(1)$.

Let A_n^0 be defined as in (S2.3). We have

$$\begin{aligned}
\text{(S3.6)}\quad & \big(n^{1/2}\beta_n\big)^{-1}\kappa_n^{-1}n^{1/2}\widehat{D}_n^{-1/2}(\theta_{n,*})\overline{m}_n(\theta_{n,*}) \\
&= \big(n^{1/2}\beta_n\big)^{-1}\kappa_n^{-1}\big(\widehat{D}_n^{-1/2}(\theta_{n,*})D^{1/2}(\theta_{n,*},F_n)\big)\big(A_n^0+n^{1/2}m_n^*\big) \\
&= o_p(1)+\kappa_n^{-1}(m_n^*/\beta_n)(1+o_p(1)),
\end{aligned}$$

where the second equality uses (S2.3) with Ω_0 replaced by Ω_1 (using Assumption DA(b)), $\kappa_n \to \infty$, and $n^{1/2}\beta_n \to \infty$. By the definition of β_n, $m_{n,j}^*/\beta_n \in [-1,\infty)$ for $j=1,\dots,k$ for all n. By a subsequence argument, wlog we assume $\kappa_n^{-1}m_{n,j}^*/\beta_n \to \eta_j \in [0,\infty]$ for $j=1,\dots,k$. This and (S3.6) give

$$\begin{aligned}
\text{(S3.7)}\quad & \big(n^{1/2}\beta_n\big)^{-1}\xi_n(\theta_{n,*}) \to_p \eta=(\eta_1,\dots,\eta_k)' \in R_{+,\infty}^p \times R^v, \\
& \Phi_{n,1} \equiv \big(n^{1/2}\beta_n\big)^{-1}\big(\min\{\xi_{n,1}(\theta_{n,*}),0\},\dots,\min\{\xi_{n,p}(\theta_{n,*}),0\},0,\dots,0\big)' \\
& \quad \to_p 0_k,
\end{aligned}$$

where $\xi_n(\theta_{n,*})=(\xi_{n,1}(\theta_{n,*}),\dots,\xi_{n,k}(\theta_{n,*}))'$.

Using Assumption 6, we have

$$\begin{aligned}
\text{(S3.8)}\quad & \big(n^{1/2}\beta_n\big)^{-\chi}S\big(\widehat{\Omega}_n^{1/2}(\theta_{n,*})Z^*+\varphi(\xi_n(\theta_{n,*}),\widehat{\Omega}_n(\theta_{n,*})),\widehat{\Omega}_n(\theta_{n,*})\big) \\
&= S\big(\big(n^{1/2}\beta_n\big)^{-1}\big[\widehat{\Omega}_n^{1/2}(\theta_{n,*})Z^*+\varphi(\xi_n(\theta_{n,*}),\widehat{\Omega}_n(\theta_{n,*}))\big],\widehat{\Omega}_n(\theta_{n,*})\big) \\
&\le S(\Phi_{n,2}Z^*+\Phi_{n,1},\widehat{\Omega}_n(\theta_{n,*})),
\end{aligned}$$

where $\Phi_{n,2} \equiv (n^{1/2}\beta_n)^{-1}\widehat{\Omega}_n^{1/2}(\theta_{n,*})$ ($\in R^{k\times k}$) and the inequality holds by Assumptions 1(a) and GMS7. We have $\Phi_{n,2}=o_p(1)$ by (S2.3) and Assumption DA(a). Let $\widetilde{c}_n$ denote the $1-\alpha$ quantile of $S(\Phi_{n,2}Z^*+\Phi_{n,1},\widehat{\Omega}_n(\theta_{n,*}))$ in (S3.8).

By (S3.8), $(n^{1/2}\beta_n)^{-\chi}\widehat{c}_n(1-\alpha) \le \widetilde{c}_n$. Hence, it suffices to show that $\widetilde{c}_n = o_p(1)$. To do so, we use a similar argument to that in (S1.15). For $x>0$, as $(\xi,\Omega_a,\Omega_b) \to (0_k,0_{k\times k},\Omega_1)$, we have

$$\begin{aligned}
\text{(S3.9)}\quad & S\big(\Omega_a^{1/2}Z^*+\xi,\Omega_b\big) \to S(0_k,\Omega_1)=0 \quad \text{a.s. } [Z^*], \\
& 1\big(S\big(\Omega_a^{1/2}Z^*+\xi,\Omega_b\big)\le x\big) \to 1(0\le x) \quad \text{a.s. } [Z^*], \\
& P\big(S\big(\Omega_a^{1/2}Z^*+\xi,\Omega_b\big)\le x\big) \to 1,
\end{aligned}$$

where the equality in the first line uses Assumption 3, the second convergence result follows from the first result for $x>0$, and the third convergence result holds by the second result and the bounded convergence theorem. The third result of (S3.9), $(\Phi_{n,1}, \Phi_{n,2}, \widehat{\Omega}_n(\theta_{n,*})) \to_p (0_k, 0_{k\times k}, \Omega_1)$ (which uses (S3.7)), and Slutsky's theorem give

$$P\big(S(\Phi_{n,2}Z^* + \Phi_{n,1}, \widehat{\Omega}_n(\theta_{n,*})) \le x\big) \to_p 1 \quad \text{for all } x > 0, \tag{S3.10}$$

where $P(\cdot)$ denotes the distribution of Z^* conditional on $(\Phi_{n,1}, \Phi_{n,2}, \widehat{\Omega}_n(\theta_{n,*}))$. By Assumption 1(c), the probability limit in (S3.10) is zero for all $x < 0$. These results and Lemma 5(a) of AG1 imply that $\widetilde{c}_n \to_p 0$, where $\widetilde{c}_n$ is the $1-\alpha$ quantile of the (random) df in (S3.10). This completes the proof of part (a). *Q.E.D.*

S4. VERIFICATION OF ASSUMPTIONS GMS1, GMS3, GMS6, AND GMS7 FOR $\varphi^{(5)}$

We now verify Assumptions GMS1, GMS3, GMS6, and GMS7 for $\varphi^{(5)}$. Assumption GMS1(b) holds for $\varphi^{(5)}$ if $c_j(\xi, \Omega) = 1$ whenever the jth element of ξ equals 0 by the definition of $\varphi^{(5)}$. If the jth element of ξ equals zero, $c_j(\xi, \Omega)$ does not enter the criterion function $S(-c\cdot\xi, \Omega) - \eta(|c|)$. In consequence, the criterion function is minimized by taking $c_j(\xi, \Omega) = 1$ because $\eta(\cdot)$ is strictly increasing. Hence, Assumption GMS1(b) holds for $\varphi^{(5)}$.

We show Assumption GMS1(a) holds for $\varphi^{(5)}$ (provided S satisfies Assumption 1(d)) by showing that if $(\xi_{[r]}, \Omega_{[r]}) \to (\xi, \Omega)$ as $r\to\infty$ and $\xi_j = 0$, then $c_j(\xi_{[r]}, \Omega_{[r]}) = 1$ for r sufficiently large. By Assumption 1(d), S is continuous at (ξ, Ω). Hence, $\lim_{r\to\infty} S(-c\cdot\xi_{[r]}, \Omega_{[r]}) \to S(-c\cdot\xi, \Omega)$ as $r\to\infty$. The limit $S(-c\cdot\xi, \Omega)$ does not depend on c_j because $\xi_j = 0$. Given $\varepsilon > 0$, there exists an r^* sufficiently large that $|S(-c\cdot\xi_{[r]}, \Omega_{[r]}) - S(-c\cdot\xi, \Omega)| \le \varepsilon$ for all $c \in \mathcal{C}$ and all $r \ge r^*$. Hence, the first term of the selection criterion, $S(-c\cdot\xi, \Omega)$, is reduced by at most ε if c_j is changed from 1 to 0, where $c = (c_1, \ldots, c_k)'$. On the other hand, the second term of the selection criterion, $-\eta(|c|)$, is increased by $\eta(|c|+1) - \eta(|c|) > 0$. Taking $\varepsilon < \inf_{c\in\mathcal{C}}(\eta(|c|+1) - \eta(|c|))$ implies that the selection criterion is minimized by $c_j(\xi_{[r]}, \Omega_{[r]}) = 1$ for all $r \ge r^*$. Hence, Assumption GMS1(a) holds for $\varphi^{(5)}$.

Next we verify Assumption GMS3 for $\varphi^{(5)}$ for all functions S for which $S(-c\cdot\xi, \Omega) \to \infty$ as $(\xi, \Omega) \to (\xi_*, \Omega_*)$ whenever $c_j = 1$. For any $c_* \in \mathcal{C}$ with $c_{*\ell} = 0$ for all ℓ such that $\xi_{*\ell} = \infty$ we have $S(-c_*\cdot\xi, \Omega) \le K$ as $(\xi, \Omega) \to (\xi_*, \Omega_*)$ for some $K < \infty$ by Assumption 1(d). Hence, some $c_* = (c_{*1}, \ldots, c_{*k})' \in \mathcal{C}$ with $c_{*j} = 0$ is selected over any $c = (c_1, \ldots, c_k)' \in \mathcal{C}$ with $c_j = 1$ as $(\xi, \Omega) \to (\xi_*, \Omega_*)$. This gives $c_j(\xi, \Omega) = 0$ and $\varphi_j(\xi, \Omega) = \infty$ (using the definition of $\varphi^{(5)}$) as $(\xi, \Omega) \to (\xi_*, \Omega_*)$.

Assumptions GMS6 and GMS7 hold immediately for $\varphi^{(5)}$ by its definition.

S5. MEAN VECTORS FOR SIMULATION RESULTS

The main paper provides simulation results concerning the finite-sample average power of PA, subsampling, and GMS tests. Here we define the sets of mean vectors, denoted $\mathcal{M}_p(\Omega)$, for which the average power is calculated for $p = 2$, 4, and 10 and Ω_{Neg}, Ω_{Zero}, and Ω_{Pos}. Each element of $\mathcal{M}_p(\Omega)$ is in the alternative hypothesis.

For $p = 2$, the μ vectors considered are

$$\text{(S5.1)} \quad \begin{aligned} \mathcal{M}_2(\Omega_{\text{Neg}}) &= \{(-1.001, 0), (-1.804, 1), (-2.303, 2), (-2.309, 3), \\ &\qquad (-2.309, 4), (-2.309, 7), (-.5165, -.5165)\}, \\ \mathcal{M}_2(\Omega_{\text{Zero}}) &= \{(-2.309, 0), (-2.309, 1), (-2.309, 2), (-2.309, 3), \\ &\qquad (-2.309, 4), (-2.309, 7), (-1.6263, -1.6263)\}, \\ \mathcal{M}_2(\Omega_{\text{Pos}}) &= \mathcal{M}_k(\Omega_{\text{Zero}}) \text{ except the last vector is } (-2.0040, -2.0040). \end{aligned}$$

The finite-sample power envelope (for known Ω) at each of these μ vectors is .75.

For $p = 4$, $\mathcal{M}_4(\Omega)$ includes 24 elements and is of the form

$$\text{(S5.2)} \quad \begin{aligned} \mathcal{M}_4(\Omega) = \big\{&(-\mu_1, -\mu_1, 1, 1), (-\mu_2, -\mu_2, 2, 2), (-\mu_3, -\mu_3, 3, 3), \\ &(-\mu_4, -\mu_4, 4, 4), (-\mu_5, -\mu_5, 7, 7), (-\mu_6, -\mu_6, 1, 7), \\ &(-\mu_7, -\mu_7, 2, 7), (-\mu_8, -\mu_8, 3, 7), (-\mu_9, -\mu_9, 4, 7), \\ &(-\mu_{10}, 1, 1, 1), (-\mu_{11}, 2, 2, 2), (-\mu_{12}, 3, 3, 3), \\ &(-\mu_{13}, 4, 4, 4), (-\mu_{14}, 7, 7, 7), (-\mu_{15}, 1, 1, 7), \\ &(-\mu_{16}, 2, 2, 7), (-\mu_{17}, 3, 3, 7), (-\mu_{18}, 4, 4, 7), \\ &(-\mu_{19}, -\mu_{19}, 0, 0), (-\mu_{20}, 0, 0, 0), (-\mu_{21}, 25, 25, 25), \\ &(-\mu_{22}, -\mu_{22}, 25, 25), (-\mu_{23}, -\mu_{23}, -\mu_{23}, 25), \\ &(-\mu_{24}, -\mu_{24}, -\mu_{24}, -\mu_{24})\big\}, \end{aligned}$$

where μ_j depends on Ω and is such that the finite-sample power envelope (for known Ω) is .80 at each element of $\mathcal{M}_4(\Omega)$.

The μ vectors in $\mathcal{M}_4(\Omega_{\text{Neg}})$ are defined by (S5.2) and the following values: $\mu_1 = -.5505$, $\mu_j = -.5526$ for $j = 2, \ldots, 5$, $\mu_6 = -.5505$, $\mu_j = -.5526$ for $j = 7, 8, 9$, $\mu_{10} = -1.8814$, $\mu_{11} = -2.4283$, $\mu_j = -2.4705$ for $j = 12, 13, 14, 17, 18, 21$, $\mu_{15} = -1.8814$, $\mu_{16} = -2.4283$, $\mu_{19} = -.3176$, $\mu_{20} = -.8624$, $\mu_{22} = -.5526$, $\mu_{23} = -.2607$, and $\mu_{24} = -.1756$.

The μ vectors in $\mathcal{M}_4(\Omega_{\text{Zero}})$ are defined by (S5.2) and the following values: $\mu_j = 1.7388$ for $j = 1, \ldots, 9, 19, 22$, $\mu_j = -2.4705$ for $j = 10, \ldots, 18, 20, 21$, $\mu_{23} = 1.4242$, and $\mu_{24} = 1.2350$.

The μ vectors in $\mathcal{M}_4(\Omega_{\text{Pos}})$ are defined by (S5.2) and the following values: $\mu_j = 2.4047$ for $j = 1, \ldots, 9, 19, 22$, $\mu_j = -2.4705$ for $j = 10, \ldots, 18, 20, 21$, $\mu_{23} = 2.2628$, and $\mu_{24} = -2.1293$.

For $p = 10$, $\mathcal{M}_{10}(\Omega)$ includes 40 vectors and is of the form:

$$
\begin{aligned}
\text{(S5.3)} \quad \mathcal{M}_{10}(\Omega) = \big\{ & (-\mu_1, -\mu_1, 1, \dots, 1), (-\mu_2, -\mu_2, 2, \dots, 2), \\
& (-\mu_3, -\mu_3, 3, \dots, 3), (-\mu_4, -\mu_4, 4, \dots, 4), \\
& (-\mu_5, -\mu_5, 7, \dots, 7), (-\mu_6, -\mu_6, 1, 1, 1, 7, \dots, 7), \\
& (-\mu_7, -\mu_7, 2, 2, 2, 7, \dots, 7), (-\mu_8, -\mu_8, 3, 3, 3, 7, \dots, 7), \\
& (-\mu_9, -\mu_9, 4, 4, 4, 7, \dots, 7), \\
& (-\mu_{10}, -\mu_{10}, -\mu_{10}, -\mu_{10}, 1, \dots, 1), \\
& (-\mu_{11}, -\mu_{11}, -\mu_{11}, -\mu_{11}, 2, \dots, 2), \\
& (-\mu_{12}, -\mu_{12}, -\mu_{12}, -\mu_{12}, 3, \dots, 3), \\
& (-\mu_{13}, -\mu_{13}, -\mu_{13}, -\mu_{13}, 4, \dots, 4), \\
& (-\mu_{14}, -\mu_{14}, -\mu_{14}, -\mu_{14}, 7, \dots, 7), \\
& (-\mu_{15}, -\mu_{15}, -\mu_{15}, -\mu_{15}, 1, 1, 1, 7, 7, 7), \\
& (-\mu_{16}, -\mu_{16}, -\mu_{16}, -\mu_{16}, 2, 2, 2, 7, 7, 7), \\
& (-\mu_{17}, -\mu_{17}, -\mu_{17}, -\mu_{17}, 3, 3, 3, 7, 7, 7), \\
& (-\mu_{18}, -\mu_{18}, -\mu_{18}, -\mu_{18}, 4, 4, 4, 7, 7, 7), (-\mu_{19}, 1, \dots, 1), \\
& (-\mu_{20}, 2, \dots, 2), (-\mu_{21}, 3, \dots, 3), (-\mu_{22}, 4, \dots, 4), \\
& (-\mu_{23}, 7, \dots, 7), (-\mu_{24}, 1, 1, 1, 7, \dots, 7), \\
& (-\mu_{25}, 2, 2, 2, 7, \dots, 7), (-\mu_{26}, 3, 3, 3, 7, \dots, 7), \\
& (-\mu_{27}, 4, 4, 4, 7, \dots, 7), (-\mu_{28}, -\mu_{28}, 0, \dots, 0), \\
& (-\mu_{29}, -\mu_{29}, -\mu_{29}, -\mu_{29}, 0, \dots, 0), (-\mu_{30}, 0, \dots, 0), \\
& (-\mu_{31}, 25, \dots, 25), (-\mu_{32}, -\mu_{32}, 25, \dots, 25), \\
& (-\mu_{33}, -\mu_{33}, -\mu_{33}, 25, \dots, 25), \\
& (-\mu_{34}, -\mu_{34}, -\mu_{34}, -\mu_{34}, 25, \dots, 25), \\
& (-\mu_{35}, -\mu_{35}, -\mu_{35}, -\mu_{35}, -\mu_{35}, 25, \dots, 25), \\
& (-\mu_{36}, \dots, -\mu_{36}, 25, 25, 25, 25), \\
& (-\mu_{37}, \dots, -\mu_{37}, 25, 25, 25), (-\mu_{38}, \dots, -\mu_{38}, 25, 25), \\
& (-\mu_{39}, \dots, -\mu_{39}, 25), (-\mu_{40}, \dots, -\mu_{40}) \big\}.
\end{aligned}
$$

The μ vectors in $\mathcal{M}_{10}(\Omega_{\text{Neg}})$ are defined by (S5.3) and the following values: $\mu_j = .6016$ for $j = 1, \dots, 9$, $\mu_j = .3475$ for $j = 10, \dots, 18$, $\mu_{19} = 1.9847$, $\mu_{20} = 2.5835$, $\mu_j = 2.6817$ for $j = 21, 22, 23, 26, 27, 31$, $\mu_{24} = 1.9847$, $\mu_{25} = 2.5835$,

$\mu_{28} = .5341$, $\mu_{29} = .3322$, $\mu_{30} = 1.1551$, $\mu_{32} = .6016$, $\mu_{33} = .4195$, $\mu_{34} = .3475$, $\mu_{35} = .2985$, $\mu_{36} = .2674$, $\mu_{37} = .2430$, $\mu_{38} = .2254$, $\mu_{39} = .2106$, and $\mu_{40} = .1993$.

The μ vectors in $\mathcal{M}_{10}(\Omega_{\text{Zero}})$ are defined by (S5.3) and the following values: $\mu_j = 1.8927$ for $j = 1, \dots, 9, 28, 32$, $\mu_j = 1.3360$ for $j = 10, \dots, 18, 29, 34$, $\mu_j = 2.6817$ for $j = 19, \dots, 27, 30, 31$, $\mu_{33} = 1.5463$, $\mu_{35} = 1.1963$, $\mu_{36} = 1.0893$, $\mu_{37} = 1.0099$, $\mu_{38} = .9465$, $\mu_{39} = .8882$, and $\mu_{40} = .8440$.

The μ vectors in $\mathcal{M}_{10}(\Omega_{\text{Pos}})$ are defined by (S5.3) and the following values: $\mu_j = 2.6227$ for $j = 1, \dots, 9$, $\mu_j = 2.4676$ for $j = 10, \dots, 18$, $\mu_j = 2.6817$ for $j = 19, \dots, 27$, $\mu_{29} = 2.6227$, $\mu_{30} = 2.6817$, $\mu_{31} = 2.6817$, $\mu_{32} = 2.6227$, $\mu_{33} = 2.5401$, $\mu_{34} = 2.4676$, $\mu_{35} = 2.4005$, $\mu_{36} = 2.3140$, $\mu_{37} = 2.2846$, $\mu_{38} = 2.2565$, $\mu_{39} = 2.2343$, and $\mu_{40} = 2.2066$.

For $p = 10$, the finite-sample power envelope (for known Ω) at each of the μ vectors is .85.

S6. MONTE CARLO EXPERIMENTS I

S6.1. *Introduction*

In this section, we report some additional simulation results. We use simulation to investigate the finite-sample properties of GMS CS's and to compare them to some other methods in the literature. We consider the coverage probabilities (CP's) of the CS's for points in and not in the identified set. For points on the boundary of the identified set and for which all inequalities are binding (i.e., hold as equalities), the CP's should be close to the nominal level $1 - \alpha$. For points on the boundary of the identified set and for which some inequality is not binding, the CP's should be greater than or equal to $1 - \alpha$. Probabilities for these points indicate the nonsimilarity on the boundary of the CS's. For points in the interior of the identified set, the CP's should be greater than $1 - \alpha$. For points that are not in the identified set, the CP's should be less than $1 - \alpha$—the smaller, the better.

We consider two very simple models. The first is a particular case of the missing-data model considered in Imbens and Manski (2004) (IM). In this model, there is one parameter, two moment inequalities, and no moment equalities. We consider the GMS CS based on the MMM test statistic (i.e., the test function S_1) with the MMSC function $\varphi^{(5)}$ with $\eta(x) = x$ and $\kappa_n = (2.01 \ln \ln n)^{1/2}$ (i.e., the HQIC MMSC procedure).[1] We compare the GMS CS to the CS introduced by IM for this model (see IM for its definition) and to the subsampling CS based on the MMM test statistic, and with subsample size $b = [n^{1/2}]$, the integer part of $n^{1/2}$. (We mention, but do not report, results for other values of b.) Rosen (2008) provided a comparison of the finite-sample properties of his proposed CS with that of IM.

[1]We take $\widehat{\sigma}^2_{n,j}(\theta) = n^{-1} \sum_{i=1}^n m_j^2(W_i, \theta)$ for $j = 1, \dots, k$, rather than $\widehat{\sigma}^2_{n,j}(\theta) = n^{-1} \sum_{i=1}^n (m_j(W_i, \theta) - \overline{m}_{n,j}(\theta))^2$. Results for the latter are similar.

The second model considered is the interval-outcome regression model of Manski and Tamer (2002). In this model, there are two parameters, two moment inequalities, and no moment equalities. We compare the same GMS and subsampling procedures as defined above. (The IM CS does not apply to this model.)

All results reported are for CS's with nominal level .95. For both models, we report results for $n = 100$, 500, and 1000. We take the number of simulation repetitions, R, to be 20,000 for the GMS and IM CS's and 5000 for the subsampling CS's.[2] The reported CP's are the relative frequencies of coverage over the R repetitions.

S6.2. *Missing-Data Model*

In this model, $W_i = (Y_i, D_i)$ are i.i.d. for $i = 1, \ldots, n$ with $Y_i \sim U[0, 1]$, $D_i \sim$ Bern[.85], and Y_i and D_i independent. The observations are $\{(Y_i D_i, D_i) : i \leq n\}$. Thus, Y_i is not observed when $D_i = 0$. The parameter of interest is $\theta = EY_i$. Given that data are missing, the parameter θ is not identified. Two moment inequality functions used to bound θ are

$$\begin{pmatrix} m_1(W_i, \theta) \\ m_2(W_i, \theta) \end{pmatrix} = \begin{pmatrix} \theta - Y_i D_i \\ (1-\theta) - (1 - Y_i) D_i \end{pmatrix}. \tag{S6.1}$$

When θ is the true parameter, we have $Em_1(W_i, \theta) = \theta - EY_i D_i \geq \theta - EY_i = 0$, where the inequality holds because $Y_i \geq 0$ and $D_i \leq 1$ and $Em_2(W_i, \theta) = (1-\theta) - E(1-Y_i)D_i \geq (1-\theta) - E(1-Y_i) = 0$, where the inequality holds because $1 - Y_i \geq 0$ and $D_i \leq 1$. Hence, the functions in (S6.1) satisfy two moment inequalities.

For the data-generating process above, the identified set $[\theta_L, \theta_U]$ is [.425, .575].[3] We consider the CP's of the CS's for the values $\theta_L = .425$, $\theta = .5$, and $\theta_H = .575$, which are in the identified set, and for the values $.9 \times \theta_L$ and $1.1 \times \theta_U$, which are not in the identified set.

Table S.I reports the CP's of the GMS, IM, and subsampling CS's with nominal level 95%. The table shows that for θ values in the identified set, the performance of the GMS and IM CS's is excellent for all sample sizes. Probabilities for the GMS CS for boundary θ points range from .948 to .951. In contrast, the subsampling CS overcovers by a noticeable amount for all sample sizes. Probabilities for the subsampling CS for boundary θ points range from .971 to .990. (This overcoverage is a finite-sample phenomenon because the subsampling asymptotic CP is .95 at both boundaries.) All three CS's cover $\theta = .5$, which

[2]The subsampling CS's are more computationally intensive than the moment selection CS's. Only 5000 repetitions are used for the moment selection results of Table III for θ not in the identified set.

[3]The identified set is determined by $\theta_L - EY_i D_i = 0$ (i.e., $\theta_L = .5 \times .85 = .425$) and $(1 - \theta_U) - E(1 - Y_i)D_i = 0$ (i.e., $\theta_U = 1 - E(1 - Y_i)ED_i = 1 - .5 \times .85 = .575$).

TABLE S.I

MISSING-DATA MODEL: FINITE-SAMPLE COVERAGE PROBABILITIES OF NOMINAL 95% CONFIDENCE INTERVALS

		Coverage Probabilities for θ Values				
		θ Values in Identified Set			θ Values Not in Identified Set	
n	Type of Confidence Interval	$\theta_L = .425$	$\theta = .5$	$\theta_H = .575$	$.9 \times \theta_L$	$1.1 \times \theta_H$
100	GMS	.951	1.0	.948	.619	.445
	Imbens/Manski	.946	1.0	.951	.637	.439
	Subsampling	.981	1.0	.990	.791	.667
500	GMS	.951	1.0	.951	.095	.010
	Imbens/Manski	.949	1.0	.950	.094	.007
	Subsampling	.975	1.0	.971	.145	.032
1 000	GMS	.951	1.0	.949	.006	.000
	Imbens/Manski	.953	1.0	.950	.005	.000
	Subsampling	.972	1.0	.971	.008	.000

lies in the interior of the identified set and is far from either boundary, with probability 1. This is in accord with the asymptotic results.

For θ points not in the identified set, we want the CP of a CS to be as close to zero as possible. (A lower CP for such points translates into a shorter and more informative CS.) Table S.I shows that the GMS CS covers points not in the identified set with substantially lower probability than the subsampling CS when $n = 100$ (viz., .619 versus .791 and .445 versus .667) and with slightly lower probability for $n = 500$ (viz., .095 versus .145 and .010 versus .030). This is consistent with the asymptotic power comparisons given in the main paper. For points not in the identified set, the GMS and IM CS's have comparable CP's. For $n = 500$ and 1000, the CP's of all three CS's are sufficiently low that the differences between them are small.

As has been reported in other scenarios, subsampling CP's are sensitive to the choice of the subsample size b. Additional simulation results not reported here show that for smaller b, the subsampling CP's for θ in the identified set become slightly closer to the nominal level, while for larger subsample sizes, they become closer to 1. For θ not in the identified set, smaller b reduces the subsampling CP's slightly and larger b increases them slightly.

S6.3. *Interval-Outcome Regression Model*

This model is a regression model with unobserved dependent variable Y_i,

$$Y_i = \theta_1 + Z_i\theta_2 + U_i, \tag{S6.2}$$

where (Z_i, U_i) are i.i.d. for $i = 1, \ldots, n$, $Z_i \sim N(1, 1)$, $U_i \sim N(0, 1)$, and $\theta = (\theta_1, \theta_2)$. The observations are $\{(Y_i^L, Y_i^H, Z_i) : i \leq n\}$, where Y_i^L equals the integer part of Y_i, Y_i^H equals the smallest integer greater than or equal to Y_i, and so $Y_i^L \leq Y_i \leq Y_i^H$ a.s. The parameter θ is not identified because Y_i is not observed. The two moment inequality functions are

$$\begin{pmatrix} m_1(W_i, \theta) \\ m_2(W_i, \theta) \end{pmatrix} = \begin{pmatrix} \theta_1 + Z_i\theta_2 - Y_i^L \\ (Y_i^H - \theta_1 - Z_i\theta_2)Z_i^2 \end{pmatrix}. \tag{S6.3}$$

When θ is the true parameter value, we have $Em_1(W_i, \theta) = \theta_1 + Z_i\theta_2 - EY_i^L \geq \theta_1 + Z_i\theta_2 - EY_i = 0$ and $Em_2(W_i, \theta) = E(Y_i^H - \theta_1 - Z_i\theta_2)Z_i^2 \geq E(Y_i - \theta_1 - Z_i\theta_2)Z_i^2 = 0$.[4] Thus, the functions in (S6.3) satisfy two moment inequalities.

We consider the case where the true parameter is $\theta = (1, 1)$. In this case, the identified set consists of the (θ_1, θ_2) values that satisfy[5]

$$\theta_1 + \theta_2 \geq 1.5 \quad \text{and} \quad 2\theta_1 + 4\theta_2 \leq 7. \tag{S6.4}$$

We consider the CP's of the CS's for the θ values $(-.5, 2)$, $(1.5, 0)$, $(1, 1.25)$, and $(1, 1)$, which are all in the identified set. The point $(-.5, 2)$ is on the boundary of the identified set with both moment inequalities binding; $(1.5, 0)$ and $(1, 1.25)$ are on the boundary of the identified set with only one inequality binding in each case; and $(1, 1)$ is in the interior of the identified set. We also consider CP's of the CS's for the θ values $(1.35, 0)$ and $(1, 1.375)$, which are not in the identified set. The point $(1.35, 0)$ violates the first inequality in (S6.4) and satisfies the second. The reverse is true for the point $(1, 1.375)$.

Table S.II reports CP's for the interval-outcome regression model. Table S.II shows that the GMS CS performs very well at $\theta = (-.5, 2)$ (at which both inequalities are binding) and at $\theta = (1.5, 0)$ (at which only the first inequality is binding) with CP's ranging between .948 and .953. Its CP's at $\theta = (1, 1.25)$ (at which only the second inequality is binding) are somewhat higher, with CP's ranging between .957 and .963. Overcoverage in this case is not necessarily a finite-sample phenomenon because the CS is not asymptotically similar on the boundary of the identified set. For points on the boundary of the identified set, the CP's of the subsampling CS are not quite as good as for the GMS CS. At $\theta = (-.5, 2)$ they vary between .940 and .963; at $\theta = (1.5, 0)$, they vary between .972 and .986; at $\theta = (1, 1.25)$ they are comparable to those of the GMS CS. Both CS's cover the point $\theta = (1, 1)$ (which is in the interior of the identi-

[4]In the second moment function, $Y_i^H - \theta_1 - Z_i\theta_2$ is multiplied by Z_i^2 to avoid perfect colinearity with $\theta_1 + Z_i\theta_2 - Y_i^L$ since $Y_i^H = Y_i^L + 1$ by definition. We do not consider optimal choices of moment functions for this model because such choices are not known and the results are only illustrative anyway.

[5]The identified set is determined by $\theta_1 + EZ_i\theta_2 - EY_i^L \geq 0$ (i.e., $\theta_1 + \theta_2 \geq 1.5$), where $EY_i^L \approx 1.5$ (by numerical calculation), and by $E(Y_i^H - \theta_1 - Z_i\theta_2)Z_i^2 \geq 0$ (i.e., $EY_i^H Z_i^2 - EZ_i^2\theta_1 - EZ_i^3\theta_2 \geq 0$), where $EY_i^H Z_i^2 \approx 7$ (by numerical calculation), $EZ_i^2 = 2$, and $EZ_i^3 = 4$.

TABLE S.II

INTERVAL-OUTCOME REGRESSION MODEL: FINITE-SAMPLE COVERAGE PROBABILITIES OF NOMINAL 95% CONFIDENCE SETS FOR (θ_1, θ_2)

		Coverage Probabilities for (θ_1, θ_2) Values					
		(θ_1, θ_2) in Identified Set				(θ_1, θ_2) Not in Identified Set	
n	Type of Confidence Interval	$(-.5, 2)$	$(1.5, 0)$	$(1, 1.25)$	$(1, 1)$	$(1.35, 0)$	$(1, 1.375)$
100	GMS	.953	.948	.963	1.0	.719	.626
	Subsampling	.963	.986	.962	1.0	.851	.638
500	GMS	.953	.950	.959	1.0	.236	.061
	Subsampling	.944	.980	.962	1.0	.375	.067
1000	GMS	.951	.951	.957	1.0	.056	.002
	Subsampling	.940	.972	.957	1.0	.106	.002

fied set and not close to a boundary) with probability 1. This is expected given that the asymptotic CP is 1.

Next, we consider θ points not in the identified set. Table S.II shows that the GMS CS has noticeably lower CP at $\theta = (1.35, 0)$ than the subsampling CS (viz., .719 versus .851 and .236 versus .375). For $\theta = (1, 1.375)$, the two CS's have comparable CP's. These results are consistent with the power results given in the main paper which show that the GMS test has higher power at some points and equal power at other points compared to the subsampling test.

Similar comments regarding the sensitivity of the subsampling results to b apply in this model as in the missing-data model.

In sum, the simulation results of this section are in accord with the asymptotic results. They show that the GMS CS has advantages relative to the subsampling CS. The GMS CS has CP's that are (i) closer to the nominal level and less nonsimilar on the boundary of the identified set, and (ii) lower for points outside the identified set.

S7. MONTE CARLO EXPERIMENT II: SUBSAMPLING WITH AND WITHOUT RECENTERING

In this section, we provide finite-sample comparisons of the maximum null rejection probability (MNRP) over different null mean vectors of recentered and non-recentered subsampling tests based on the QLR test statistic (i.e., S_2). (MNRP is defined precisely below.) In short, the recentered subsampling test is found to out-perform the non-recentered subsampling test by a substantial margin in terms of the closeness of the nominal level and the finite-sample MNRP of the tests.

We consider the case in which no equalities arise (i.e., $v = 0$) and the number of inequalities, p, is 2 or 4. For given θ, the null hypothesis is

$H_0\!:\!Em(W_i,\theta)\geq 0_p$ for some given moment functions $m(W_i,\theta)$ and the alternative hypothesis is that H_0 does not hold. The finite-sample properties of tests of H_0 depend on $m(W_i,\theta)$ only through (i) $\mu = Em(W_i,\theta)$, (ii) $\Omega = \mathrm{Corr}(m(W_i,\theta))$, and (iii) the distribution of the mean zero, variance I_p random vector $Z^\dagger = \mathrm{Var}^{-1/2}(m(W_i,\theta))(m(W_i,\theta) - Em(W_i,\theta))$. We consider the case in which $Z^\dagger \sim N(0_p, I_p)$. We consider three representative correlation matrices Ω_{Neg}, Ω_{Zero}, and Ω_{Pos}, which exhibit negative, zero, and positive correlations, respectively. Specifically, Ω_{Zero} equals I_p for $p=2$ and 4. The matrices Ω_{Neg} and Ω_{Pos} are Toeplitz matrices with correlations on the diagonals given by a $p-1$ vector ρ. For $p=2$, $\rho=-.9$ for Ω_{Neg} and $\rho=.5$ for Ω_{Pos}. For $p=4$, $\rho=(-.9,.7,-.5)$ for Ω_{Neg} and $\rho=(.9,.7,.5)$ for Ω_{Pos}.

By MNRP, we mean the maximum null rejection probability over all mean vectors in H_0 (i.e., all $\mu = Em(W_i,\theta)\in R_+^p$, where $R_+ = \{x\in R\!:\!x\geq 0\}$) under the assumption of normally distributed moment inequalities (i.e., $Z^\dagger \sim N(0_p, I_p)$) and given the correlation matrix Ω_{Neg}, Ω_{Zero}, or Ω_{Pos}.

For the subsampling test without recentering, the subsample size is $m=n^{1/2}$. (Better results for this test were not found by taking other values of m.) For the recentered subsampling test, the subsample size is $m=.75n^{2/3}$. (This choice is based on the simulation results reported in the main paper.) For the subsampling test without recentering, the simulation results are based on 5000 subsamples and 5000 simulation repetitions. For the recentered subsampling test, the simulation results are based on 2500 subsamples and 2500 simulation repetitions.

Table S.III reports the MNRP results for subsampling with and without re-

TABLE S.III

FINITE-SAMPLE MNRP'S OF NOMINAL .05 TESTS[a]

Number of Moment Inequalities	Sample Size n	Critical Value	MNRP[b] Ω_{Neg}	Ω_{Zero}	Ω_{Pos}
			.0	.0	.0
2	250	Sub/Recenter	.050	.050	.051
	250	Sub/NoRecenter	.027	.020	.018
	1000	Sub/NoRecenter	.037	.027	.025
	5000	Sub/NoRecenter	.044	.039	.037
4	250	Sub/Recenter	.046	.047	.051
	250	Sub/NoRecenter	.038	.030	.020
	1000	Sub/NoRecenter	.036	.032	.028
	5000	Sub/NoRecenter	.044	.041	.036

[a]Tests are based on the quasi-likelihood ratio test statistic combined with recentered subsampling (Sub/Recenter) and non-recentered subsampling (Sub/NoRecenter) critical values for sample sizes $n=250$, 1000, and 5000, three correlation matrices Ω_{Neg}, Ω_{Zero}, and Ω_{Pos}, and 2 and 4 moment inequalities.

[b]MNRP denotes the maximum null rejection probability over mean vectors in H_0 under the assumption of normally distributed moment inequalities and given the correlation matrix Ω_{Neg}, Ω_{Zero}, or Ω_{Pos}.

centering. The results show that subsampling without recentering leads to substantial underrejection of the null hypothesis for sample sizes $n = 250$ and 1000. For example, for $p = 2$ and $n = 250$, its MNRP ranges from .018 to .027 for nominal .05 tests. For $p = 2$ and $n = 1000$, it ranges from .025 to .037. Even for a sample size of 5000, the subsampling test without recentering underrejects the null hypothesis somewhat. In contrast, the subsampling test with recentering has good MNRP values for a sample size as small as 250. For $p = 2$ and $n = 250$, its MNRP ranges from .050 to .051.

REFERENCES

ANDREWS, D. W. K., AND P. GUGGENBERGER (2009b): "Validity of Subsampling and 'Plug-in Asymptotic' Inference for Parameters Defined by Moment Inequalities," *Econometric Theory*, 25, 669–709. [6]

——— (2010b): "Asymptotic Size and a Problem With Subsampling and With the *m* Out of *n* Bootstrap," *Econometric Theory*, 26 (forthcoming). [6]

IMBENS, G., AND C. F. MANSKI (2004): "Confidence Intervals for Partially Identified Parameters," *Econometrica*, 72, 1845–1857. [23]

MANSKI, C. F., AND E. TAMER (2002): "Inference on Regressions With Interval Data on a Regressor or Outcome," *Econometrica*, 70, 519–546. [24]

ROSEN, A. M. (2008): "Confidence Sets for Partially Identified Parameters That Satisfy a Finite Number of Moment Inequalities," *Journal of Econometrics*, 146, 107–117. [23]

Cowles Foundation for Research in Economics, Dept. of Economics, Yale University, P.O. Box 208281, Yale Station, New Haven, CT 06520-8281, U.S.A.; Donald.andrews@yale.edu

and

Dept. of Economics, Yale University, P.O. Box 208268, New Haven, CT 06520-8268, U.S.A.; gsoares@aya.yale.edu.

Manuscript received May, 2007; final revision received September, 2009.