Supplemental Material

for

Inference in Moment Inequality Models That Is Robust to Spurious Precision under Model Misspecification

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

Soonwoo Kwon
Department of Economics
Yale University

First Version: March 12, 2017
Revised: October 9, 2019

Contents

11 Outline

References to sections with section numbers less than 11 refer to sections of the main paper. Similarly, all theorems and lemmas with section numbers less than 11 refer to results in the main paper. BCS abbreviates Bugni, Canay, and Shi (2015). For ease of reference, the assumptions used in the paper and this Supplemental Material are listed in the last section of this Supplemental Material, Section 27.

Section 12 of this Supplemental Material gives lower and upper bounds on the asymptotic power of SPUR1 tests for $n^{-1/2}$-local alternatives and consistency results for these tests. These results have implications for the asymptotic power of SPUR2 tests as well.

Section 13 shows that when the "max" S function is employed, the SPUR test statistic is equivalent to a recentered test statistic, as has been considered in Chernozhukov, Hong, and Tamer (2007) for use with a correctly-specified model.

Section 14 defines the one-sided upper-bound CI $CI_{n,r,UP}(\alpha)$ for $r_F^{\inf}$ introduced in Andrews and Kwon (2019) that is employed by SPUR2 tests and CS's. It also provides some properties of this CI.

Section 15 provides explicit expressions for the bootstrap quantities $sd^*_{anj}(\theta)$ for $a = 1, ..., 4$ that are employed by the EGMS critical values arise in (6.4) and (6.7)–(6.10).

Section 16 discusses extensions of the results of the paper to tests with weighted moment inequalities, to tests without the standard-deviation normalization, and to non-i.i.d. observations.

Section 17 provides additional numerical results concerning the spurious precision of the GMS CS's in Andrews and Soares (2010), as well as the proof of Lemma 17.1, which concerns the spurious precision of these CS's.

Section 18 provides some additional simulation results for the lower/upper bound model considered in Section 8.

Section 19 provides derivations for (8.5) and (8.6), which concern the missing data model.

Section 20 states Lemma 20.1, which gives sufficient conditions for Assumptions NLA and CA, and proves Lemmas 5.1, 5.2, and 20.1.

Sections 21–26 prove the main results of the paper. Section 21 proves Theorem 5.3, which gives the asymptotic distribution of the SPUR test statistic.

Section 22 states Theorem 22.1, which is the key ingredient to the proofs of Theorems 7.1 and 12.1, which provide asymptotic size and power results for SPUR1 and SPUR2 tests and CS's. Theorem 22.1 provides asymptotic null rejection probability (NRP) results, asymptotic $n^{-1/2}$-local power bounds, and consistency results for the nominal level α SPUR1 test $\phi_{n,SPUR1}(\theta_n)$, defined in (4.4), under drifting subsequences of distributions and parameter values. Section 23 proves Lemmas

22.2–22.5, which are used in the proof of Theorem 22.1.

Section 24 proves Theorem 7.1, which shows that the SPUR1 and SPUR2 tests and CS's have correct asymptotic size, using Theorem 22.1. Section 25 proves Theorem 12.1 using Theorem 22.1.

Section 26 proves Theorem 9.1 and establishes rate of convergence results for the set estimator $\widehat{\Theta}_n$ of the MR identified set under correct model specification and misspecification.

Let $o_p^{\Theta}(1)$ and $O_p^{\Theta}(1)$ denote quantities that are $o_p(1)$ and $O_p(1)$, respectively, uniformly over $\theta \in \Theta$.

12 Asymptotic Power

In this section, we give upper and lower bounds on the asymptotic power of SPUR1 tests for $n^{-1/2}$-local alternatives. Bounds on asymptotic local power, rather than precise asymptotic local power, are given due to the complexity of the data-dependent EGMS critical values. Even the bounds involve fairly complicated expressions. We also provide consistency results for these tests under fixed and non-$n^{-1/2}$-local alternatives. The results allow for drifting null hypothesis values, which yield asymptotic false coverage probabilities for SPUR1 CS's. As discussed below, the results have implications for the asymptotic power of SPUR2 tests.

For $\theta \in \Theta$, define

$$j_n(\theta) := \arg\max_{j \leq k} b_{nj}(\theta), \text{ where } b_{nj}(\theta) := n^{1/2}([E_{F_n}\widetilde{m}_j(W,\theta)]_- - r_{F_n}^{\inf}).[19] \tag{12.1}$$

By Lemma 5.2(a),

$$b_{nj_n(\theta)}(\theta) \geq 0 \ \forall \theta \in \Theta. \tag{12.2}$$

We employ the following bootstrap convergence (BC) assumptions, which apply to a drifting sequence of null values $\{\theta_n\}_{n\geq 1}$ and distributions $\{F_n\}_{n\geq 1}$.

Assumption BC.1. $\sup_{\theta\in\Theta} |sd_{anj}^*(\theta) - sd_{aj\infty}(\theta)| \to_p 0$ as $n \to \infty$ for some nonrandom continuous real-valued functions $sd_{aj\infty}(\theta)$ on Θ for $j \leq k$ and $a = 1, 3$.

Define

$$\begin{aligned}\Lambda_{n,F_n}^{*\eta_n} := \Big\{&(\theta, b, b^*, \ell, j^*) \in \Theta_I^{\eta_n}(F_n) \times R^{3k} \times \{1,...,k\} : b_j = n^{1/2}([E_{F_n}\widetilde{m}_j(W,\theta)]_- - r_{F_n}^{\inf}),\\ &b_j^* = (sd_{3j\infty}(\theta)\kappa_n)^{-1} b_j,\ \ell_j = n^{1/2} E_{F_n}\widetilde{m}_j(W,\theta)\ \forall j \leq k,\ j^* := j_n(\theta)\Big\},\end{aligned} \tag{12.3}$$

where $\{\eta_n\}_{n\geq 1}$ is as in Assumption C.8 and $\{\kappa_n\}_{n\geq 1}$ is as in (6.4), (6.7), (6.8), and (6.9). Let

[19] If the arg max is not unique, $j_n(\theta)$ is defined to be the smallest arg max.

$\mathcal{S}(\Theta \times R^{3k}_{[\pm\infty]} \times \{1, ..., k\})$ denote the space of compact subsets of the metric space $(\Theta \times R^{3k+1}_{[\pm\infty]}, d)$, where d is defined following (5.2) with $a_* = d_\theta + 3k + 1$.

Define $\Lambda^{*\eta_{Un}}_{Un,F_n}$ analogously to $\Lambda^{*\eta_n}_{n,F_n}$, but without the elements (ℓ, j^*) and with η_{Un} in place of η_n. Thus, $\Lambda^{*\eta_{Un}}_{Un,F_n}$ contains points $(\theta, b, b^*) \in \Theta^{\eta_{Un}}_I(F_n) \times R^{2k}$. Let $sd_{1j\infty} := sd_{1j\infty}(\theta_\infty)$, where θ_∞ is as in Assumption C.1.

Assumption BC.2. $(sd_{1j\infty}\kappa_n)^{-1} n^{1/2}(E_{F_n}\widetilde{m}_j(W, \theta_n) + r^{\inf}_{F_n}) \to h^*_{j\infty}$ for some $h^*_{j\infty} \in R_{[\pm\infty]}$ $\forall j \leq k$.

Assumption BC.3. $\Lambda^{*\eta_n}_{n,F_n} \to_H \Lambda^*_I$ for some non-empty set $\Lambda^*_I \in \mathcal{S}(\Theta \times R^{3k}_{[\pm\infty]} \times \{1, ..., k\})$ for some constants $\{\eta_n\}_{n\geq1}$ that satisfy $\eta_n \to \infty$ and $\eta_n/\tau_n \to 0$ for $\{\tau_n\}_{n\geq1}$ as in Assumption A.6(ii).

Assumption BC.4. $\Lambda^{*\eta_{Un}}_{Un,F_n} \to_H \Lambda^*_{U,I}$ for some non-empty set $\Lambda^*_{U,I} \in \mathcal{S}(\Theta \times R^{2k}_{[\pm\infty]})$ for constants $\{\eta_{Un}\}_{n\geq1}$ that satisfy $\eta_{Un} \to \infty$ and $\tau_n/\eta_{Un} \to 0$ for $\{\tau_n\}_{n\geq1}$ as in Assumption A.6(ii).

We employ the following assumption on the GMS function $\varphi = (\varphi_1, ..., \varphi_k)'$, which appears in (6.4) and (6.11) and is defined following (6.4).

Assumption A.10. Given the function $\varphi : R^k_{[+\infty]} \times \Psi \to R^k_{[+\infty]}$, there is a function $\varphi^{**} : R^k_{[+\infty]} \to R^k_{[+\infty]}$ that takes the form $\varphi^{**}(\xi) = (\varphi^{**}_1(\xi_1), ..., \varphi^{**}_k(\xi_k))'$ and $\forall j \leq k$, (i) $\varphi^{**}_j(\xi_j) \leq \varphi_j(\xi, \Omega)$ $\forall(\xi, \Omega) \in R^k_{[+\infty]} \times \Psi$, (ii) φ^{**}_j is continuous, and (iii) $\varphi^{**}_j(\xi_j) = 0$ $\forall \xi_j \leq 0$ and $\varphi^{**}_j(\infty) = \infty$.

For example, in the leading case where $\varphi_j(\xi, \Omega) = \infty 1(\xi_j > 1)$ for $j \leq k$, Assumption A.10 holds with $\varphi^{**}_j(\xi_j) = \infty 1(\xi_j \geq 1 + \varepsilon) + ((\xi_j - 1)/(1 + \varepsilon - \xi_j))1(1 \leq \xi_j < 1 + \varepsilon)$ for any $\varepsilon > 0$.

For $\theta \in \Theta$, define a lower bound (wp→1) random variable, $S^*_{Ln,EGMS}(\theta)$, on the EGMS bootstrap statistic $S^*_{n,EGMS}(\theta)$ to be

$$\begin{aligned}
S^*_{Ln,EGMS}(\theta) &:= S\left(T^*_{Ln,EGMS}(\theta) + A^{*\inf}_{Ln,EGMS} 1_k, \widehat{\Omega}_n(\theta)\right), \text{ where} \\
T^*_{Lnj,EGMS}(\theta) &:= \widehat{\nu}^*_{nj}(\theta) + \varphi^*_j(\xi_{nj}(\theta)) \ \forall j \leq k, \\
T^*_{Ln,EGMS}(\theta) &:= (T^*_{L1n,EGMS}, ..., T^*_{Lkn,EGMS})', \\
A^{*\inf}_{Ln,EGMS} &:= \inf_{\theta \in \Theta^{\eta_n}_I(F_n)} \max_{j \leq k} \Big(\chi(\widehat{\nu}^*_{nj}(\theta), n^{1/2} E_{F_n}\widetilde{m}_j(W, \theta)) + 1(j \neq j_n(\theta)) b_{nj}(\theta) \\
&\qquad + 1(j = j_n(\theta))\varphi^*_j(\xi^b_{nj}(\theta))\Big) \qquad (12.4)
\end{aligned}$$

for $1_k := (1, ..., 1)' \in R^k$ and $\chi(\nu, c) := [\nu + c]_- - [c]_-$.

The asymptotic distribution of the lower bound random variable $S^*_{Ln,EGMS}(\theta_n)$ is

$$\begin{aligned}
S^*_{L\infty,EGMS} &:= S\left(T^*_{L\infty,EGMS} + A^{*\inf}_{L\infty,EGMS} 1_k, \Omega_\infty\right), \text{ where} \\
T^*_{Lj\infty,EGMS} &:= G^{m\sigma}_{j\infty} + \varphi^*_j(h^*_{j\infty}) \ \forall j \leq k, \ T^*_{L\infty,EGMS} = (T^*_{L1\infty,EGMS}, ..., T^*_{Lk\infty,EGMS})', \text{ and} \\
A^{*\inf}_{L\infty,EGMS} &:= \inf_{(\theta,b,b^*,\ell,j^*) \in \Lambda^*_I} \max_{j \leq k} \left(\chi(G^{m\sigma}_j(\theta), \ell_j) + 1(j \neq j^*) b_j + 1(j = j^*)\varphi^*_{j^*}(b^*_{j^*})\right) \qquad (12.5)
\end{aligned}$$

for Λ_I^* as in Assumption BC.3.

For $\theta \in \Theta$, define an upper bound (wp→1) random variable, $S_{Un,EGMS}^*(\theta)$, on the EGMS bootstrap statistic $S_{n,EGMS}^*(\theta)$ to be

$$\begin{aligned} S_{Un,EGMS}^*(\theta) &:= S\left(T_{Un,EGMS}^*(\theta) + A_{Un,EGMS}^{*\inf} 1_k, \widehat{\Omega}_n(\theta)\right), \text{ where} \\ T_{Unj,EGMS}^*(\theta) &:= \widehat{\nu}_{nj}^*(\theta) + \varphi_j^{**}(\xi_{nj}(\theta)) \ \forall j \leq k, \\ T_{Un,EGMS}^*(\theta) &:= (T_{U1n,EGMS}^*, ..., T_{Ukn,EGMS}^*)', \text{ and} \\ A_{Un,EGMS}^{*\inf} &:= \inf_{\theta \in \Theta_I^{\eta_{Un}}(F_n)} \min_{j \leq k} \left(-[\widehat{\nu}_{nj}^*(\theta)]_+ + \varphi_j^{**}(\xi_{nj}^b(\theta))\right) \end{aligned} \tag{12.6}$$

for φ_j^{**} as in Assumption A.10. The asymptotic distribution of $S_{Un,EGMS}^*(\theta_n)$ is

$$\begin{aligned} S_{U\infty,EGMS}^* &:= S\left(T_{U\infty,EGMS}^* + A_{U\infty,EGMS}^{*\inf} 1_k, \Omega_\infty\right), \text{ where} \\ T_{Uj\infty,EGMS}^* &:= G_{j\infty}^{m\sigma} + \varphi_j^{**}(h_{j\infty}^*) \ \forall j \leq k, \ T_{U\infty,EGMS}^* = (T_{U1\infty,EGMS}^*, ..., T_{Uk\infty,EGMS}^*)', \\ A_{U\infty,EGMS}^{*\inf} &:= \inf_{(\theta,b,b^*,\ell) \in \Lambda_{U,I}^*} \min_{j \leq k} \left(-[G_j^{m\sigma}(\theta)]_+ + \varphi_j^{**}(b_j^*)\right) \end{aligned} \tag{12.7}$$

for $\Lambda_{U,I}^*$ as in Assumption BC.4.

Let $c_{L\infty,EGMS}(1-\alpha)$ and $c_{U\infty,EGMS}(1-\alpha)$ denote the $1-\alpha$ quantiles of $S_{L\infty,EGMS}^*$ and $S_{U\infty,EGMS}^*$, respectively. For some results, we assume that $S_{U\infty,EGMS}^*$ satisfies the following continuity condition.

Assumption BC.5. The distribution of $S_{U\infty,EGMS}^*$ is continuous at $c_{U\infty,EGMS}(1-\alpha)$.

Theorem 12.1 *For sequences* $\{F_n\}_{n\geq 1}$ *and* $\{\theta_n\}_{n\geq 1}$ *that satisfy Assumptions* A.0–A.4, A.6, BC.1, BC.2, C.1–C.4, C.7–C.9, *and* S.1 *and for* $\alpha \in (0,1)$, *the nominal level* α SPUR1 *test* $\phi_{n,SPUR1}(\theta_n)$ *for testing* $H_0 : \theta_n \in \Theta_I(F_n)$ *satisfies*

(a) $\limsup_{n\to\infty} P_{F_n}(\phi_{n,SPUR1}(\theta_n) = 1) \leq P(S_\infty > c_{L\infty,EGMS}(1-\alpha))$ *provided Assumptions* A.5, BC.3, *and* NLA *hold,*

(b) $\liminf_{n\to\infty} P_{F_n}(\phi_{n,SPUR1}(\theta_n) = 1) \geq P(S_\infty > c_{U\infty,EGMS}(1-\alpha))$ *provided Assumptions* A.10, BC.4, BC.5, *and* NLA *hold, and*

(c) $\liminf_{n\to\infty} P_{F_n}(\phi_{n,SPUR1}(\theta_n) = 1) = 1$ *provided Assumptions* A.10, BC.4, CA, S.2, *and* S.3 *hold.*[20]

Comments. (i). Theorem 12.1(a) and (b) provide upper and lower bounds on the asymptotic power of the SPUR1 test under $n^{-1/2}$-local alternatives.

[20] In Theorem 12.1(a), the constants $\{\eta_n\}_{n\geq 1}$ in Assumptions BC.2 and C.9 are assumed to be the same. For example, one can take $\eta_n := \tau_n^{1/2}$ $\forall n \geq 1$ given τ_n in the definition of $\widehat{\Theta}_n$ in (6.5) and in Assumption A.6(iii). In Theorem 12.1(b) and (c), one can take $\eta_{Un} := \tau_n^2$ $\forall n \geq 1$ to be the constants $\{\eta_{Un}\}_{n\geq 1}$ in Assumption BC.3.

(ii) In Theorem 12.1(a) and (b), the distribution of S_∞, defined in (5.11), and the magnitude of the asymptotic power of the SPUR1 test under $n^{-1/2}$-local alternatives depends on the "noncentrality parameters" $h_{j\infty} := \lim n^{1/2}(E_{F_n}\widetilde{m}_j(W,\theta_n) + r^{\inf}_{F_n}) \in R_{[\pm\infty]}\ \forall j \leq k$ that appear in Assumption C.3. Increasingly negative values of $h_{j\infty}$ lead to greater asymptotic power.

(iii). Theorem 12.1(c) shows that the SPUR1 test is consistent against all alternatives that satisfy Assumption CA.

The results of Theorem 12.1 give the power properties of the SPUR2 test when the model exhibits "large-local" or "global" model misspecification, i.e., when $\{F_n\}_{n\geq 1}$ is such that $n^{1/2}r^{\inf}_{F_n} \to \infty$ (which is Assumption MM in Section 14 below). In this case, the upper bound $\widehat{r}_{n,UP}(\alpha)$ of the CI for $r^{\inf}_{F_n}$ is positive wp→1 by Proposition 14.2(b) below, the level α SPUR2 test equals the level α_2 SPUR1 test wp→1, and the SPUR2 test has the same asymptotic power properties as the level α_2 SPUR1 test.

On the other hand, the asymptotic power of the SPUR2 test is the same as that of the level α_2 GMS test, see Andrews and Soares (2010), when $\{F_n\}_{n\geq 1}$ is such that there exists a sequence $\{\theta^I_n \in \Theta_I(F_n)\}_{n\geq 1}$ for which $n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta^I_n) \to \infty\ \forall j \leq k$ (which is Assumption IS in Section 14 below). This occurs when the model is correctly specified and the identified set contains slack points for which the slackness of the inequalities is of order greater than $n^{-1/2}$. In this case, the upper bound $\widehat{r}_{n,UP}(\alpha)$ equals zero wp→1 by Proposition 14.2(a) below and the level α SPUR2 test equals the level α_2 GMS test wp→1.

The SPUR2 test is consistent against all alternatives that satisfy Assumption CA, because both of the GMS and SPUR1 tests are.

13 Recentered Test Statistics

An alternative to the SPUR test statistic defined in Section 4.1 is a recentered test statistic, such as considered in Chernozhukov, Hong, and Tamer (2007), which is defined to be

$$S_{n,Recen}(\theta) := S_{n,Std}(\theta) - \inf_{\overline{\theta}\in\Theta} S_{n,Std}(\overline{\theta}), \tag{13.1}$$

where $S_{n,Std}(\theta) := S(n^{1/2}\widehat{m}_n(\theta), \widehat{\Omega}_n(\theta))$ is a "standard" test statistic, such as one considered in Andrews and Soares (2010), see (3.1). The MR identified set corresponding to the recentered statistic is the set of θ values that minimize the population version of the recentered statistic.[21] It depends on the choice of test statistic.

[21] The population version of the recentered statistic is $S(E_F\widetilde{m}(W,\theta),\Omega_F(\theta)) - \inf_{\overline{\theta}\in\Theta} S(E_F\widetilde{m}(W,\overline{\theta}),\Omega_F(\theta))$, where $\Omega_F(\theta) := Var_F(\widetilde{m}(W_i,\theta))$.

Chernozhukov, Hong, and Tamer (2007) consider recentered test statistics, but they do not analyze their asymptotic properties under misspecification or under correct specification with drifting sequences of distributions $\{F_n\}_{n\geq 1}$. In consequence, it is not clear whether the application of subsampling to recentered test statistics provides critical values that are uniformly asymptotically valid under misspecification or correct specification.[22]

When $S_{n,Std}(\theta)$ is a test statistic from Andrews and Soares (2010) with the function S equal to S_4, see (4.3), we denote the recentered test statistic by $S_{4n,Recen}(\theta)$. It is easy to show that the MR identified set corresponding to $S_{4n,Recen}(\theta)$ is the same as the MR identified set in Section 2. On the other hand, if one employs a different S function in $S_{Recen,n}(\theta)$, the MR identified set is different.

When the function S employed by the SPUR test statistic $S_n(\theta)$ defined in (4.2) is S_4, we denote the SPUR statistic by $S_{4n}(\theta)$. The following lemma shows that the recentered statistic $S_{4n,Recen}(\theta)$ is identical to the $S_{4n}(\theta)$ SPUR statistic. That is, for the S_4 function, the recentered statistic is not an alternative to the SPUR statistic—it is the same.

Lemma 13.1 *For any $\theta \in \Theta$, $S_{4n,Recen}(\theta) = S_{4n}(\theta)$.*

Proof of Lemma 13.1. By (4.1), $\widehat{r}_n^{\inf} := \inf_{\theta\in\Theta} \max_{j\leq k}[\widehat{m}_{nj}(\theta)]_-$. Hence, for $S = S_4$, $\inf_{\overline{\theta}\in\Theta} S_{n,Std}(\overline{\theta}) = n^{1/2}\widehat{r}_n^{\inf}$. In consequence,

$$\begin{aligned} S_{4n,Recen}(\theta) &= \max_{j\leq k}\left[n^{1/2}\widehat{m}_{nj}(\theta)\right]_- - n^{1/2}\widehat{r}_n^{\inf} \text{ and} \\ S_{4n}(\theta) &= \max_{j\leq k}\left[n^{1/2}\widehat{m}_{nj}(\theta) + n^{1/2}\widehat{r}_n^{\inf}\right]_- . \end{aligned} \tag{13.2}$$

We claim: $S_{4n,Recen}(\theta) > 0$ iff $S_{4n}(\theta) > 0$. This clearly holds if $\widehat{r}_n^{\inf} = 0$, so suppose $\widehat{r}_n^{\inf} > 0$. In this case, $S_{4n,Recen}(\theta) > 0$ iff $-n^{1/2}\widehat{m}_{nj}(\theta) - n^{1/2}\widehat{r}_n^{\inf} > 0$ for some $j \leq k$ iff $S_{4n}(\theta) > 0$, which proves the claim. In addition, $S_{4n}(\theta) \geq 0$ because $[x]_- \geq 0$ for all x, and $S_{4n,Recen}(\theta) \geq 0$ because $\widehat{r}_n^{\inf}$ is the $\inf_{\theta\in\Theta}$ of $\max_{j\leq k}[\widehat{m}_{nj}(\theta)]_-$, which completes the proof. □

For recentered tests based on S not equal to S_4, one can determine the asymptotic distribution of $S_{n,Recen}(\theta_n)$ under suitable drifting sequences $\{\theta_n\}_{n\geq 1}$ and $\{F_n\}_{n\geq 1}$ by altering the proof of Theorem 5.3(b). However, the resulting asymptotic distribution seems problematic because it is not apparent how one can construct a critical value in an EGMS fashion that exploits the analogue of the condition $\max_{j\leq k} b_j \geq 0$, which appears when $S = S_4$.

[22] The reason is that, even under correct specification, the recentering term $\inf_{\overline{\theta}\in\Theta} S_{n,Std}(\overline{\theta})$ has a complicated asymptotic distribution under drifting sequences of distributions (given by $A_\infty^{\inf}(\Lambda)$ in Theorem 5.3(b) when the recentered test is based on S_5 in (4.3)). In consequence, the argument for the correct asymptotic size of the subsampling test based on a test statistic without recentering that is given in Andrews and Guggenberger (2009) does not extend to the case of the subsampling recentered test.

14 Confidence Interval for $\mathbf{r_F^{inf}}$

In this section, we define the one-sided upper-bound CI $CI_{n,r,UP}(\alpha)$ for $r_F^{\inf}$ that is introduced in Andrews and Kwon (2019) and employed by the SPUR2 test and CI in Section 4.2. Define

$$\Delta_{Fj}(\theta) := -E_F \widetilde{m}_j(W,\theta) \text{ for } j \leq k, \ \Delta_F(\theta) := \max_{j \leq k} \Delta_{Fj}(\theta), \text{ and } \Delta_F^{\inf} := \inf_{\theta \in \Theta} \Delta_F(\theta). \tag{14.1}$$

The parameter $\Delta_F^{\inf}$ is the minimum over Θ of the maximum inequality violation over the k moments, where a slack moment inequality yields a negative violation value. We refer to $\Delta_F^{\inf}$ as the minimax violation parameter.

If the model is correctly specified, $\Delta_F^{\inf} \leq 0$ because there exists some $\theta \in \Theta$ for which all of the moment inequalities are satisfied, i.e., $\max_{j \leq k} \Delta_{Fj}^{\inf}(\theta) \leq 0$. If the model is misspecified, $\Delta_F^{\inf} > 0$ because for all $\theta \in \Theta$ some moment inequality is violated, i.e., $\Delta_{Fj}(\theta) > 0$.[23] When $\Delta_F^{\inf} \geq 0$, $r_F^{\inf} = \Delta_F^{\inf}$. When $\Delta_F^{\inf} < 0$, $r_F^{\inf} = 0$. Thus,

$$r_F^{\inf} = \max\{\Delta_F^{\inf}, 0\} \tag{14.2}$$

and $\Delta_F^{\inf}$ provides more information than $r_F^{\inf}$. For this reason, the CI for $r_F^{\inf}$ is obtained from a CI for $\Delta_F^{\inf}$. This yields a CI for $r_F^{\inf}$ that has the feature that it equals $\{0\}$ wp$\to$1 when the model is correctly specified and the identified set contains slack points θ for which the slackness of the inequalities is of order greater than $n^{-1/2}$. In turn, this yields the highly desirable feature of the SPUR2 test that under these circumstances it has the same asymptotic properties as a standard test that assumes correct model specification.

We estimate $\Delta_{Fj}(\theta)$, $\Delta_F(\theta)$, and $\Delta_F^{\inf}$ by

$$\widehat{\Delta}_{nj}(\theta) := -\widehat{m}_{nj}(\theta), \ \widehat{\Delta}_n(\theta) := \max_{j \leq k} \widehat{\Delta}_{nj}(\theta), \text{ and } \widehat{\Delta}_n^{\inf} := \inf_{\theta \in \Theta} \widehat{\Delta}_n(\theta) \tag{14.3}$$

for $j \leq k$, respectively. The nominal level $1-\alpha$ one-sided upper-bound CI for $\Delta_F^{\inf}$ is

$$CI_{n,\Delta,UP}(\alpha) := (-\infty, \widehat{\Delta}_{n,\Delta,UP}^{\inf}(\alpha)], \text{ where } \widehat{\Delta}_{n,\Delta,UP}^{\inf}(\alpha) := \widehat{\Delta}_n^{\inf} - \frac{\widehat{c}_{n,\Delta,UP}(\alpha)}{n^{1/2}} \tag{14.4}$$

and $\widehat{c}_{n,\Delta,UP}(\alpha)$ is a data-dependent EGMS critical value defined below. The nominal $1-\alpha$ one-sided upper-bound CI for $r_F^{\inf}$ is

$$CI_{n,r,UP}(\alpha) := [0, \widehat{r}_{n,UP}(\alpha)], \text{ where } \widehat{r}_{n,UP}(\alpha) := \max\{\widehat{\Delta}_{n,\Delta,UP}^{\inf}(\alpha), 0\}. \tag{14.5}$$

[23] This statement relies on continuity of $\Delta_{Fj}(\theta)$ and compactness of Θ by Assumption A.0.

As defined, $CI_{n,r,UP}(\alpha) = \{0\}$ whenever the CI for $\Delta_F^{\inf}$ indicates the model is correctly specified, i.e., whenever $\widehat{\Delta}_{n,\Delta,UP}^{\inf}(\alpha) \leq 0$.

We have $\Delta_{F_n}^{\inf} \in CI_{n,\Delta,UP}(\alpha)$ iff $n^{1/2}(\widehat{\Delta}_n^{\inf} - \Delta_{F_n}^{\inf}) < \widehat{c}_{n,\Delta,UP}(\alpha)$. Hence, the critical value $\widehat{c}_{n,\Delta,UP}(\alpha)$ is determined using the asymptotic distribution of

$$A_{n,\Delta}^{\inf} := n^{1/2}\left(\widehat{\Delta}_n^{\inf} - \Delta_{F_n}^{\inf}\right). \tag{14.6}$$

The upper-bound EGMS bootstrap critical value $\widehat{c}_{n,\Delta,UP}(\alpha)$ is defined as follows. Let

$$\widehat{e}_{nj}(\theta) := n^{1/2}\left(\widehat{\Delta}_{nj}(\theta) - \widehat{\Delta}_n^{\inf}\right) - sd_{enj}^*(\theta)\kappa_n, \tag{14.7}$$

where $sd_{5nj}^*(\theta) := \max\{Var^*(n^{1/2}(\widehat{\Delta}_{nj}(\theta) - \widehat{\Delta}_n^{\inf}))^{1/2}, 1\}$ for $j \leq k$, $Var^*(\cdot)$ denotes the bootstrap variance defined in (6.1), and $\widehat{\Theta}_n$ is defined in (BC.5). Let

$$\xi_{nj}^e(\theta) := (sd_{5nj}^*(\theta)\kappa_n)^{-1} n^{1/2}\left(\widehat{\Delta}_{nj}(\theta) - \widehat{\Delta}_n^{\inf}\right) \ \forall j \leq k \text{ and } \xi_n^e(\theta) = (\xi_{n1}^e(\theta), ..., \xi_{nk}^e(\theta))', \tag{14.8}$$

where κ_n and $sd_{5nj}^*(\theta)$ are as above. Define

$$\widehat{J}_{ne}(\theta) := \{j \in \{1, ..., k\} : \widehat{\Delta}_{nj}(\theta) \geq \widehat{\Delta}_n(\theta) - sd_{6nj}^*(\theta) n^{-1/2}\kappa_n\}, \tag{14.9}$$

where $\widehat{\Delta}_{nj}(\theta)$ and $\widehat{\Delta}_n(\theta)$ are defined in (14.3) and $sd_{6nj}^*(\theta) := \max\{Var^*(n^{1/2}(\widehat{\Delta}_{nj}(\theta) - \widehat{\Delta}_n(\theta)))^{1/2}, 1\}$ for $j \leq k$. Explicit expressions for $sd_{5nj}^*(\theta)$ and $sd_{6nj}^*(\theta)$ are given in Section 15 below.

The asymptotic distribution of $A_{n,\Delta}^{\inf}$ depends on the set of minimizers of $\Delta_F(\theta)$ over Θ, which is defined by $\Theta_{\min}(F) := \{\theta \in \Theta : \Delta_F(\theta) = \Delta_F^{\inf}\}$. Under Assumption A.0, $\Theta_{\min}(F)$ is non-empty. The critical value $\widehat{c}_{n,\Delta,UP}(\alpha)$ employs the following estimator of $\Theta_{\min}(F)$:

$$\widehat{\Theta}_{\min,n} := \{\theta \in \Theta : \widehat{\Delta}_n(\theta) \leq \widehat{\Delta}_n^{\inf} + \tau_n/n^{1/2}\}, \tag{14.10}$$

where $\{\tau_n\}_{n\geq 1}$ is a sequence of positive constants that satisfies $\tau_n \to \infty$ (and typically $\tau_n/n^{1/2} \to 0$), such as the BIC choice $\tau_n = (\ln n)^{1/2}$.[24]

The upper-bound EGMS bootstrap statistic, $A_{n,\Delta,UP}^{*\inf}$, is defined to be

$$A_{n,\Delta,UP}^{*\inf} := \inf_{\theta \in \widehat{\Theta}_{\min,n}} \min_{j_1 \in \widehat{J}_{ne}(\theta)} \max_{j \leq k} \left(-\widehat{\nu}_{nj}^*(\theta) + 1(j \neq j_1)\widehat{e}_{nj}(\theta) + 1(j = j_1)\varphi_j(\xi_n^e(\theta), \widehat{\Omega}_n(\theta))\right). \tag{14.11}$$

The upper-bound critical value $\widehat{c}_{n,\Delta,UP}(\alpha)$ is the α conditional quantile of $A_{n,\Delta,UP}^{*\inf}$ given $\{W_i\}_{i\leq n}$

[24] More precisely, $\widehat{\Theta}_{\min,n}$ is (and needs to be) an estimator of an asymptotically small expansion of the minimizer set $\Theta_{\min}(F)$, see Andrews and Kwon (2019) for details.

for $\alpha \in (0,1)$. This quantile can be computed by simulation. The form of $A^{*\inf}_{n,\Delta,UP}$ is similar to that of $A^{*\inf}_{n,EGMS}$, but it is not the same. See Andrews and Kwon (2019) for the details behind its specific form.

Proposition 14.1 below shows that $CI_{n,r,UP}(\alpha)$ has correct asymptotic level in a uniform sense with i.i.d. observations under a set of relatively primitive conditions. This result relies on the asymptotic distribution of $A^{\inf}_{n,\Delta}$ for a certain subsequence $\{a_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$.

There always exists a sequence $\{F_n\}_{n\geq 1}$ and a subsequence $\{q_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$ such that

$$\begin{aligned}\liminf_{n\to\infty} \inf_{F\in\mathcal{P}} P_F\left(\Delta^{\inf}_F \in (-\infty, \widehat{\Delta}^{\inf}_{n,UP}(\alpha)]\right) &= \liminf_{n\to\infty} P_{F_n}\left(n^{1/2}(\widehat{\Delta}^{\inf}_n - \Delta^{\inf}_{F_n}) \geq \widehat{c}_{n,\Delta,UP}(\alpha)\right) \\ &= \lim P_{F_{q_n}}(A^{\inf}_{q_n,\Delta} \geq \widehat{c}_{q_n,\Delta,UP}(\alpha)), \quad (14.12)\end{aligned}$$

where the first and second equalities use (14.4) and (14.6), respectively. For the subsequence $\{q_n\}_{n\geq 1}$ in (14.12), let $\{a_n\}_{n\geq 1}$ be a subsequence of $\{q_n\}_{n\geq 1}$ for which $\Lambda_{a_n,\Delta,F_{a_n}} \to_H \Lambda_\Delta$ as $n\to\infty$ for some $\Lambda_\Delta \in \mathcal{S}(\Theta \times R^k_{[\pm\infty]})$. Such a subsequence always exists. The corresponding subsequence of statistics $\{A^{\inf}_{a_n,\Delta} : n \geq 1\}$ has asymptotic distribution $A^{\inf}_{\infty,\Delta}(\Lambda_\Delta)$ defined by

$$A^{\inf}_{\infty,\Delta}(\Lambda_\Delta) := \inf_{(\theta,e)\in\Lambda_\Delta} \max_{j\leq k}\left(-G^m_j(\theta) + \frac{1}{2}\widetilde{m}_j(\theta)G^\sigma_j(\theta) + e_j\right). \quad (14.13)$$

Let $c_{\infty,\Delta}(\alpha)$ denote the α quantile of $A^{\inf}_{\infty,\Delta}(\Lambda_\Delta)$. We impose the following continuity condition on the distribution function of $A^{\inf}_{\infty,\Delta}(\Lambda_\Delta)$ at $c_{\infty,\Delta}(\alpha)$.

Assumption A.7$_\Delta$. $P(A^{\inf}_{\infty,\Delta}(\Lambda_\Delta) = c_{\infty,\Delta}(\alpha)) = 0$.

Assumption A.7$_\Delta$ can be avoided by defining $\widehat{c}_{n,\Delta,UP}(\alpha)$ to be the α conditional quantile of $A^{*\inf}_{n,\Delta,UP}$ given $\{W_i\}_{i\leq n}$ minus a very small constant ζ, such as $\zeta = 10^{-6}$.

Proposition 14.1 *Under Assumptions* A.0–A.6, A.7$_\Delta$, *and* A.8, *for* $\alpha \in (0,1)$, *the nominal* $1-\alpha$ CI $CI_{n,r,UP}(\alpha)$ *satisfies*

$$\liminf_{n\to\infty} \inf_{F\in\mathcal{P}} P_F(r^{\inf}_{F_n} \in CI_{n,r,UP}(\alpha)) \geq 1-\alpha.$$

Comment. Proposition 14.1 follows from Theorem 6.1(a) in Andrews and Kwon (2019), which concerns $CI_{n,\Delta,UP}(\alpha)$, using the definition of $CI_{n,r,UP}(\alpha)$ in (14.5) because $\Delta^{\inf}_{F_n} \in CI_{n,\Delta,UP}(\alpha)$ implies that $r^{\inf}_{F_n} \in CI_{n,r,UP}(\alpha)$.

Next, we show that when the model is correctly specified and the sequence of MR identified sets $\{\Theta_I(F_n)\}_{n\geq 1}$ contains slack points with slackness of order greater than $n^{-1/2}$, defined precisely in Assumption IS below, then $CI_{n,r,UP}(\alpha) = \{0\}$ wp$\to$1. This demonstrates that it is possible to provide evidence that the model is not identifiably misspecified, which is the reverse of evidence provided by a model misspecification test.

We employ the following assumption concerning the MR identified set (IS).

Assumption IS. The sequence $\{F_n\}_{n\geq 1}$ is such that there exists a sequence $\{\theta_n^I \in \Theta_I(F_n)\}_{n\geq 1}$ for which $n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta_n^I) \to \infty \; \forall j \leq k$.

We also show that if the model exhibits "large-local" or "global" model misspecification (MM), then $\widehat{r}_{n,UP}(\alpha) > 0$ wp$\to$1.

Assumption MM. The sequence $\{F_n\}_{n\geq 1}$ is such that $n^{1/2}r_{F_n}^{\inf} \to \infty$.

Proposition 14.2 *Suppose Assumptions* A.0–A.6 *and* A.8 *hold.*

(a) *For sequences* $\{F_n\}_{n\geq 1}$ *that satisfy Assumption* IS, $\liminf_{n\to\infty} P_{F_n}(\widehat{r}_{n,UP}(\alpha) = 0) = 1$.

(b) *For sequences* $\{F_n\}_{n\geq 1}$ *that satisfy Assumption* MM, $\liminf_{n\to\infty} P_{F_n}(\widehat{r}_{n,UP}(\alpha) > 0) = 1$.

Comments. (i). Proposition 14.2(a) is a consequence of Theorem 7.1 in Andrews and Kwon (2019) and $\widehat{r}_{n,UP}(\alpha) := \max\{\widehat{\Delta}_{n,\Delta,UP}^{\inf}(\alpha), 0\}$. Proposition 14.2(b) is a consequence of Theorem 16.1 in the Supplemental Material to Andrews and Kwon (2019) and $\widehat{r}_{n,UP}(\alpha) := \max\{\widehat{\Delta}_{n,\Delta,UP}^{\inf}(\alpha), 0\}$.

(ii). Proposition 14.2(a) implies that the level α misspecification-robust adaptive SPUR2 test has the same power properties as a level α_2 standard GMS test that is designed for correct model specification when the model is correctly specified and Assumption IS holds, where $\alpha = \alpha_1 + \alpha_2$ and $\alpha_1, \alpha_2 > 0$, such as $\alpha = .05$ and $\alpha_2 = .045$.

(iii). Proposition 14.2(b) implies that the level α adaptive SPUR2 test has the same power properties as the level α_2 SPUR1 test when the model is misspecified and Assumption MM holds.

15 Explicit Expressions for $\mathbf{sd}_{\mathbf{anj}}^*(\boldsymbol{\theta})$ for $\mathbf{a = 1, ..., 6}$

Here we provide explicit expressions for the bootstrap quantities $sd_{anj}^*(\theta)$ for $a = 1, ..., 6$ that arise in (6.4), (6.7)–(6.10), (14.7), and (14.9), based on $b = 1, ..., B$ bootstrap samples, which we denote by $sd_{anjB}^*(\theta)$ for $a = 1, ..., 6$. We also provide expressions for the bootstrap statistics $\{S_{nb,EGMS}^*(\theta) : b = 1, ..., B\}$.

Given the definitions of $sd_{anj}^*(\theta)$ for $a = 1, ..., 6$, it suffices to provide explicit expressions for

$$\begin{aligned}
&\widehat{V}_{1nj}^*(\theta) := Var^*(n^{1/2}(\widehat{m}_{nj}(\theta) + \widehat{r}_n(\theta))),\ \widehat{V}_{2nj}^*(\theta) := Var^*(n^{1/2}\widehat{m}_{nj}(\theta)),\\
&\widehat{V}_{3nj}^*(\theta) := Var^*(n^{1/2}([\widehat{m}_{nj}(\theta)]_- - \widehat{r}_n^{\inf})),\ \widehat{V}_{4nj}^*(\theta) := Var^*(n^{1/2}(\widehat{r}_{nj}(\theta) - \widehat{r}_n(\theta))),\\
&\widehat{V}_{5nj}^*(\theta) := Var^*(n^{1/2}(\widehat{\Delta}_{nj}(\theta) - \widehat{\Delta}_n^{\inf})) \text{ and } \widehat{V}_{6nj}^*(\theta) := Var^*(n^{1/2}(\widehat{\Delta}_{nj}(\theta) - \widehat{\Delta}_n(\theta))). \quad (15.1)
\end{aligned}$$

Based on the nonparametric i.i.d. bootstrap in (6.1), let $\{W_{ib}^*\}_{i\leq n}$ denote the b-th bootstrap sample for $b = 1, ..., B$. Then, $\widehat{V}_{1nj}^*(\theta_n)$ and $\widehat{V}_{2nj}^*(\theta_n)$ are simulated using B bootstrap samples via the formulae:

$$\widehat{V}_{1njB}^*(\theta) := nB^{-1}\sum_{b=1}^{B}\left(\frac{\overline{m}_{njb}^*(\theta)}{\widehat{\sigma}_{njb}^*(\theta)} + \max_{J\leq k}\left[\frac{\overline{m}_{nJb}^*(\theta)}{\widehat{\sigma}_{nJb}^*(\theta)}\right]_- - B^{-1}\sum_{c=1}^{B}\left(\frac{\overline{m}_{njc}^*(\theta)}{\widehat{\sigma}_{njc}^*(\theta)} + \max_{J\leq k}\left[\frac{\overline{m}_{nJc}^*(\theta)}{\widehat{\sigma}_{nJc}^*(\theta)}\right]_-\right)\right)^2,$$

$$\widehat{V}_{2njB}^*(\theta) := nB^{-1}\sum_{b=1}^{B}\left(\frac{\overline{m}_{njb}^*(\theta)}{\widehat{\sigma}_{njb}^*(\theta)} - B^{-1}\sum_{c=1}^{B}\frac{\overline{m}_{njc}^*(\theta)}{\widehat{\sigma}_{njc}^*(\theta)}\right)^2,$$

$$\overline{m}_{njb}^*(\theta) := n^{-1}\sum_{i=1}^{n} m_j(W_{ib}^*, \theta), \text{ and } \widehat{\sigma}_{njb}^{*2}(\theta) := n^{-1}\sum_{i=1}^{n}\left(m_j(W_{ib}^*, \theta) - \overline{m}_{njb}^*(\theta)\right)^2. \tag{15.2}$$

The quantity $\widehat{V}_{3nj}^*(\theta)$ is simulated using B bootstrap samples via the formula:

$$\widehat{V}_{3njB}^*(\theta) := nB^{-1}\sum_{b=1}^{B}\left(\left[\frac{\overline{m}_{njb}^*(\theta)}{\widehat{\sigma}_{njb}^*(\theta)}\right]_- - \widehat{r}_{nb}^{*\inf} - B^{-1}\sum_{c=1}^{B}\left(\left[\frac{\overline{m}_{njc}^*(\theta)}{\widehat{\sigma}_{njc}^*(\theta)}\right]_- - \widehat{r}_{nc}^{*\inf}\right)\right)^2, \text{ where}$$

$$\widehat{r}_{nb}^{*\inf} := \inf_{\theta\in\widehat{\Theta}_n}\widehat{r}_{nb}^*(\theta) := \inf_{\theta\in\widehat{\Theta}_n}\max_{j\leq k}\left[\frac{\overline{m}_{njb}^*(\theta)}{\widehat{\sigma}_{njb}^*(\theta)}\right]_-. \tag{15.3}$$

The quantity $\widehat{V}_{4nj}^*(\theta)$ is simulated using B bootstrap samples via the formula:

$$\begin{aligned}\widehat{V}_{4njB}^*(\theta) := nB^{-1}\sum_{b=1}^{B}\Bigg(&\left[\frac{\overline{m}_{njb}^*(\theta)}{\widehat{\sigma}_{njb}^*(\theta)}\right]_- - \max_{J\leq k}\left[\frac{\overline{m}_{nJb}^*(\theta)}{\widehat{\sigma}_{nJb}^*(\theta)}\right]_- \\ &-B^{-1}\sum_{c=1}^{B}\left(\left[\frac{\overline{m}_{njc}^*(\theta)}{\widehat{\sigma}_{njc}^*(\theta)}\right]_- - \max_{J\leq k}\left[\frac{\overline{m}_{nJc}^*(\theta)}{\widehat{\sigma}_{nJc}^*(\theta)}\right]_-\right)\Bigg)^2.\end{aligned} \tag{15.4}$$

The quantities $\widehat{V}_{5nj}^*(\theta)$ and $\widehat{V}_{6nj}^*(\theta)$ are simulated using B bootstrap samples via the formulae:

$$\widehat{V}_{5njB}^*(\theta) := nB^{-1}\sum_{b=1}^{B}\left(\widehat{\Delta}_{njb}^*(\theta) - \widehat{\Delta}_{nb}^{*\inf} - B^{-1}\sum_{c=1}^{B}\left(\widehat{\Delta}_{njc}^*(\theta) - \widehat{\Delta}_{nc}^{*\inf}\right)\right)^2 \text{ and}$$

$$\widehat{V}_{6njB}^*(\theta) := nB^{-1}\sum_{b=1}^{B}\left(\widehat{\Delta}_{njb}^*(\theta) - \widehat{\Delta}_{nb}^*(\theta) - B^{-1}\sum_{c=1}^{B}\left(\widehat{\Delta}_{njc}^*(\theta) - \widehat{\Delta}_{nc}^*(\theta)\right)\right)^2, \text{ where}$$

$$\widehat{\Delta}_{njb}^*(\theta) := -\frac{\overline{m}_{njb}^*(\theta)}{\widehat{\sigma}_{njb}^*(\theta)},\ \widehat{\Delta}_{nb}^*(\theta) := \max_{j\leq k}\widehat{\Delta}_{njb}^*(\theta),\ \widehat{\Delta}_{nb}^{*\inf} := \inf_{\theta\in\widehat{\Theta}_n}\widehat{\Delta}_{nb}^*(\theta), \tag{15.5}$$

$$\overline{m}_{njb}^*(\theta) := n^{-1}\sum_{i=1}^{n} m_j(W_{ib}^*, \theta), \text{ and } \widehat{\sigma}_{njb}^{*2}(\theta) := n^{-1}\sum_{i=1}^{n}\left(m_j(W_{ib}^*, \theta) - \overline{m}_{njb}^*(\theta)\right)^2.$$

By definition, $sd_{anjB}^*(\theta) := (\widehat{V}_{anjB}^*(\theta))^{1/2}$ for $a = 1, ..., 6$.

The bootstrap statistic $S^*_{nb,EGMS}(\theta)$ is simulated for $b = 1, ..., B$ by

$$\begin{aligned}
S^*_{nb,EGMS}(\theta) &:= S\left(T^*_{nb,EGMS}(\theta) + A^{*\inf}_{nb,EGMS}1_k, \widehat{\Omega}_n(\theta)\right), \text{ where} \\
T^*_{njb,EGMS}(\theta) &:= \widehat{\nu}^*_{njb}(\theta) + \varphi_j(\xi_{nB}(\theta), \widehat{\Omega}_n(\theta)), \\
\widehat{\nu}^*_{njb}(\theta) &:= n^{1/2}\left(\frac{\overline{m}^*_{njb}(\theta)}{\widehat{\sigma}^*_{njb}(\theta)} - \widehat{m}_{nj}(\theta)\right), \\
A^{*\inf}_{n,EGMS} &:= \inf_{\theta \in \widehat{\Theta}_n} \min_{j_1 \in \widehat{J}_{nB}(\theta)} \max_{j \leq k}\Big(\widehat{\chi}^*_{njb,EGMS}(\theta) + 1(j \neq j_1)\widehat{b}_{njB,EGMS}(\theta) \\
&\qquad + 1(j = j_1)\varphi_j(\xi^b_{nB}(\theta), \widehat{\Omega}_n(\theta))\Big), \qquad (15.6)
\end{aligned}$$

$\xi_{nB}(\theta)$ is defined in (6.4) with $sd^*_{1njB}(\theta)$ in place of $sd^*_{1nj}(\theta)$, $\widehat{J}_{nB}(\theta)$ is defined in (6.10) with $sd^*_{4njB}(\theta)$ in place of $sd^*_{4nj}(\theta)$, $\widehat{\chi}^*_{njb,EGMS}(\theta)$ is defined in (6.7) with $\widehat{\nu}^*_{njb}(\theta)$ and $sd^*_{2njB}(\theta)$ in place of $\widehat{\nu}^*_{nj}(\theta)$ and $sd^*_{2nj}(\theta)$, respectively, $\widehat{b}_{njB,EGMS}(\theta)$ is defined in (6.8) with $sd^*_{3njB}(\theta)$ in place of $sd^*_{3nj}(\theta)$, and $\xi^b_{nB}(\theta)$ is defined in (6.9) with $sd^*_{3njB}(\theta)$ in place of $sd^*_{3nj}(\theta)$.

The bootstrap critical value is the $1 - \alpha$ sample quantile of $\{S^*_{nb,EGMS}(\theta) : b = 1, ..., B\}$.

16 Extensions

16.1 Weighted Moments

The weights used in the definition of the MR identified set $\Theta_I(F)$ in (2.7) are uniform weights. This follows from the 1_k vector that appears in (2.5) and (2.7). Non-uniform weights $\omega := (\omega_1, ..., \omega_k)'$, where $\omega_j \in [0, \infty)$ for $j \leq k$, can be introduced by replacing 1_k by $\overline{\omega} = (1/\omega_1, ..., 1/\omega_k)'$ in these equations, where $1/0 := \infty$. Equivalently, one can define $r_{Fj}(\theta) := [\omega_j E_F \widetilde{m}_j(W, \theta)]_-$ and $r_F(\theta) = \max_{j \leq k} r_{Fj}(\theta)$ (analogously to (2.6)). The larger is ω_j, the more weight is placed on inequality j and the less inequality j is relaxed in the MR identified set under misspecification. For example, if one believes that some key moment inequalities are correctly specified and one does not want these inequalities to be relaxed under misspecification, then one can set the weights ω_j corresponding to these inequalities to be very large relative to the other weights, such as 1000 versus 1. If $\omega_j = 0$, the jth moment inequality is ignored.

The SPUR1 and SPUR2 tests can be constructed with weights ω. In the definition of $\widehat{r}_{nj}(\theta)$ in (4.1), $\widehat{m}_{nj}(\theta)$ is replaced by $\omega_j \widehat{m}_{nj}(\theta)$, i.e., $\widehat{r}_{nj}(\theta) := [\omega_j \widehat{m}_{nj}(\theta)]_-$. In the definition of the SPUR statistic $S_n(\theta_0)$ in (4.2) and (5.2), $\widehat{r}^{\inf}_n 1_k$ is replaced by $\widehat{r}^{\inf}_n \overline{\omega}$. In the definition of the EGMS critical value, (i) $\widehat{r}_n(\theta)$ is replaced by $\omega_j \widehat{r}_n(\theta)$ in the definitions of $\xi_{nj}(\theta)$ in (6.4) and $sd^*_{1nj}(\theta)$ following (6.4), (ii) $\widehat{m}_{nj}(\theta) + \widehat{r}^{\inf}_n$ is replaced by $\widehat{m}_{nj}(\theta) + \omega_j \widehat{r}^{\inf}_n$ in the definition of $\widehat{\Theta}_n$ in (6.5), (iii) $\widehat{r}^{\inf}_n$ is replaced by $\omega_j \widehat{r}^{\inf}_n$ in the definitions of $\widehat{b}_{nj,EGMS}(\theta)$ in (6.8), $sd^*_{3nj}(\theta)$ following (6.8), and $\xi^b_{nj}(\theta)$

in (6.9), and (iv) $\widehat{r}_{nj}(\theta)$ is defined by $[\omega_j \widehat{m}_{nj}(\theta)]_-$, $\widehat{r}_n(\theta) := \max_{j \le k}[\omega_j \widehat{m}_{nj}(\theta)]_-$, and $sd^*_{4nj}(\theta)$ is defined using these updated definitions in the definition of $\widehat{J}_n(\theta)$ in (6.10), and (v) $A^{*\inf}_{n,EGMS} 1_k$ replaced by $A^{*\inf}_{n,EGMS} \overline{\omega}$ in the definition of $S^*_{n,EGMS}(\theta)$ in (6.3).

For the SPUR2 test, the definition of the CI $CI_{n,\Delta,UP}(\alpha)$ is altered as follows to take account of the weights ω. The definition of the population quantity $\Delta_{Fj}(\theta) := -E_F \widetilde{m}_j(W, \theta)$ in (14.1) is replaced by $\Delta_{Fj}(\theta) := -\omega_j E_F \widetilde{m}_j(W, \theta)$. Correspondingly, the definition of the sample quantity $\widehat{\Delta}_{nj}(\theta) := -\widehat{m}_{nj}(\theta)$ in (14.3) is replaced by $\widehat{\Delta}_{nj}(\theta) := -\omega_j \widehat{m}_{nj}(\theta)$. Given this change, $CI_{n,\Delta,UP}(\alpha)$ is defined as in (14.5), and the critical value $\widehat{c}_{n,\Delta,UP}(\alpha)$ is defined as in (14.7)–(14.11). With the updated definitions of the SPUR1 test and $CI_{n,\Delta,UP}(\alpha)$, the SPUR2 test with weights ω is defined just as in Section 4.2.

The above changes to the definition of the SPUR test statistic to take account of weights ω affect its asymptotic distribution as follows. In the definition of $\Lambda_{n,F}$ in (5.3) and Λ^η_{n,F_n} defined following (5.4), b_j is defined with $\omega_j r^{\inf}_F$ in place of $r^{\inf}_F$. And because $\Lambda_{n,F}$ and Λ^η_{n,F_n} appear in Assumptions C.7 and C.8, respectively, this affects these assumptions and the sets Λ and Λ_I. In the definition of $\Theta^\eta_I(F)$ in (5.4) and in Assumption C.3, $r^{\inf}_F$ is replaced by $\omega_j r^{\inf}_F$. This change in Assumption C.3 effects the definition of $h_{j\infty}$. In Lemma 5.2(a), $b_{nj}(\theta)$ is defined with $r^{\inf}_{F_n}$ replaced by $\omega_j r^{\inf}_{F_n}$. The changes above affect the definitions of $A^{\inf}_n(\Lambda_{n,F_n})$ and $A^{\inf}_\infty(\Lambda)$ in (5.10), but do not require any changes in their expressions given in (5.10).

Provided $\omega_j \in [0, \infty)$ for all $j \le k$ and $\omega_j > 0$ for some $j \le k$, all of the results above concerning the SPUR1 and SPUR2 tests, namely, Theorems 7.1, 9.1, and 12.1, as well as Propositions 14.1 and 14.2, go through for the weighted versions of these tests given the changes above. The tests are invariant to the scale of ω.

16.2 Tests without the Standard-Deviation Normalization

In some scenarios, it may be desirable to define the MR identified set $\Theta_I(F)$ in (2.7) without the standard deviation normalization of the moment functions—i.e., to define $\Theta_I(F)$ with $m(W, \theta)$ in place of $\widetilde{m}(W, \theta)$. For example, in their study of demand based on quasilinear utility, Allen and Rehbeck (2018) do not renormalize their moment inequality functions because the moment functions are denominated in dollars, which makes the interpretation simple. In this paper, a notationally-convenient equivalent way to describe non-normalized moments is to redefine $\sigma^2_{Fj}(\theta)$ in (2.2) to equal 1 $\forall j \le k$, $\forall \theta \in \Theta$. Then, $m(W, \theta) = \widetilde{m}(W, \theta)$. One forms a "non-normalized" test statistic by redefining $\widehat{\sigma}^2_{nj}(\theta)$ in (2.10) to equal 1 $\forall j \le k$, $\forall \theta \in \Theta$. In this case, $\widehat{\Omega}_n(\theta) = \widehat{\Sigma}_n(\theta)$ in (2.11) and $\widehat{\Omega}_n(\theta)$ is a variance matrix, rather than a correlation matrix. Denote the resulting test statistic by $S_{n,non}(\theta)$, where "*non*" stands for non-normalized.

The asymptotic distributions of $S_{n,non} := S_{n,non}(\theta_n)$ and its components, denoted by $T_{n,non}(\theta_n)$ and $A^{\inf}_{n,non}$, are as in Theorem 5.3 with all of its assumptions defined with $\sigma^2_{Fj}(\theta) = \widehat{\sigma}^2_{nj}(\theta) = 1$, which yields $m(W,\theta) = \widetilde{m}(W,\theta)$, and with Assumption C.5 redefined with $\nu^{\sigma}_n(\theta) = G^{\sigma}_j(\theta) = 0$ $\forall\theta \in \Theta$, which yields $G^{\sigma}_{j\infty} = 0$ and $G^{m\sigma}_{j\infty} := G^{m}_{j\infty}$. Thus, the asymptotic distribution of $S_{n,non}$ differs from that of S_n because there is no effect of estimation of the standard deviations, but otherwise is unchanged.

Given this, one defines the EGMS critical values for $S_{n,non}$ as in Section 4.1 and the CI $CI_{n,r,UP}(\alpha)$ as in Section 14, but with $\widehat{\sigma}^{*2}_{nj}(\theta) = 1$, which yields $\nu^{\sigma*}_{nj}(\theta) = 0$ and $\widehat{\nu}^{*}_{nj}(\theta) := n^{1/2}\left(\overline{m}^{*}_{nj}(\theta) - \overline{m}_{nj}(\theta)\right)$, and $\sigma^2_{Fj}(\theta) = 1$, which yields $\widehat{m}_{nj}(\theta) = \overline{m}_{nj}(\theta)$ $\forall j \leq k$, $\forall\theta \in \Theta$.

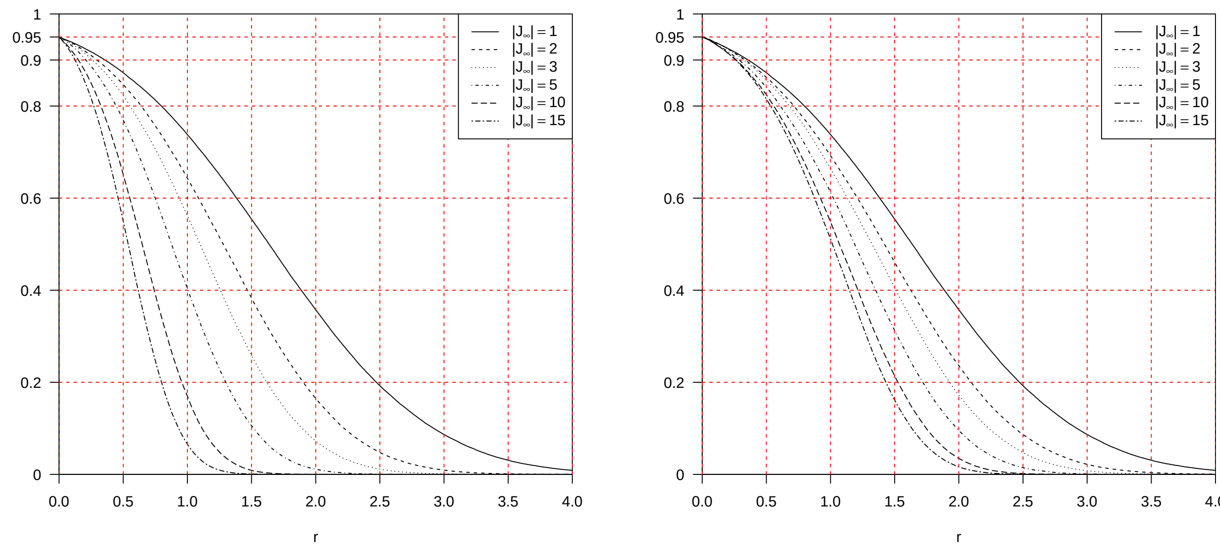
The results of Theorem 7.1 and 12.1 hold for the SPUR1 and SPUR2 tests based on $S_{n,non}$ provided the assumptions imposed in the theorems are modified by taking $\sigma^2_{Fj}(\theta) = \widehat{\sigma}^2_{nj}(\theta) = 1$ and the number of moments finite in Assumption A.3 is reduced to $2 + a$ from $4 + a$. Finally, the results of Theorem 9.1 for the set estimator $\widehat{\Theta}_n$ also hold in the non-normalized case with the same modifications.

Note that weighted moments also can be employed with non-normalized moments. In this case, the changes outlined above for both of these scenarios need to be employed.

16.3 Non-I.I.D. Observations

The basic results in this paper are given under high-level conditions that allow for non-identically distributed and/or clustered observations, as well as time series observations. For example, this is true of Theorem 5.3 and of Theorem 22.1 below, which is the key ingredient to the proofs of Theorems 7.1 and 12.1. In particular, provided the distributions F of the observations are restricted such that Assumptions C.5, C.6, and BC.6 can be verified for suitable subsequences $\{p_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$, the rest of the proofs of the asymptotic size results go through.

For non-i.i.d. observations, the following changes are needed: the nonparametric i.i.d. bootstrap defined in (6.1) needs to be changed (a) for clustered observations to a cluster-level nonparametric i.i.d. bootstrap and (b) for time series observations to a block bootstrap or Markov bootstrap, but (c) for independent non-identically distributed observations does not need to be changed. With these changes, the SPUR1 and SPUR2 tests have correct asymptotic size (under conditions such that Assumptions C.5, C.6, and BC.6 can be verified).



(a) Test Function $S_1(\cdot)$ (equivalently $S_2(\cdot)$)

(b) Test Function $S_4(\cdot)$

Figure 17.1: Maximum Coverage Probabilities for any $\theta \in \Theta$ for a Standard 95% GMS Confidence Set under Model Misspecification Indexed by r: $J_{\#} = 1, 2, \ldots, 15$, $\rho = 0$, and (a) Test Function $S_1(\cdot)$ (equivalently $S_2(\cdot)$) and (b) Test Function $S_4(\cdot)$

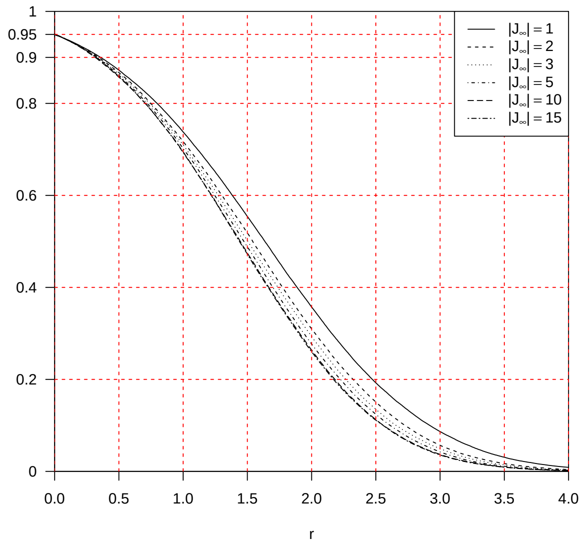
17 Spurious Precision of GMS CS's

This section provides numerical results regarding the spurious precision of GMS CS's that augment those in Figure 3.1. In addition, it provides the asymptotic spurious precision results upon which Figure 3.1 is based.

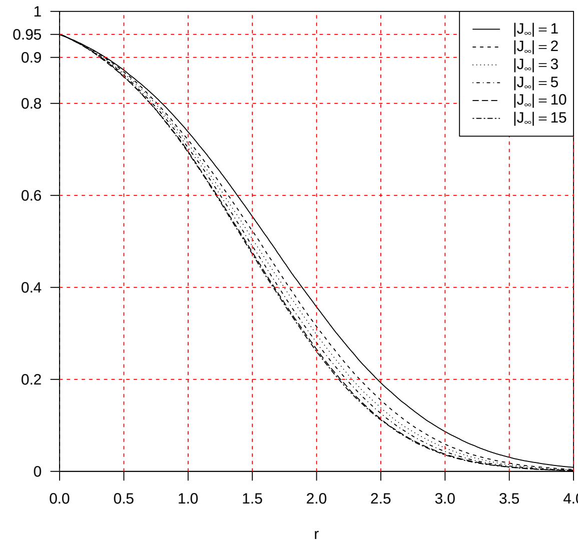
17.1 Numerical Results

Figures 16.1, 16.2, and 16.3 compare the spurious precision of GMS CS's based on different S functions, viz., S_1, S_2, and S_4, for different numbers $J_{\#}$ of violated moment inequalities and different values of the common correlation ρ between the moment functions. Figure 16.1 considers $J_{\#} = 1, 2, 3, 5, 10, 15$ with $\rho = 0$ (in which case $S_1 = S_2$). Figure 16.2 considers the same $J_{\#}$ values with $\rho = .75$ (in which case $S_1 \neq S_2$). Figure 16.3 considers $J_{\#} = 2$ and $\rho = 0, .2, .4, .6, .8, .95$.

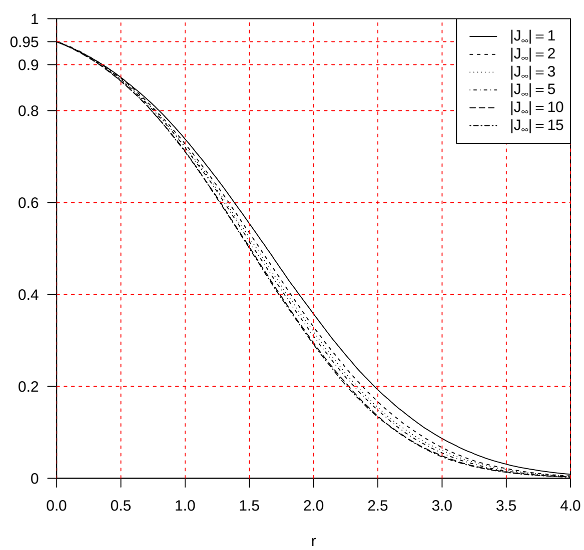
Figure 16.1 shows higher levels of spurious precision for S_1 (and S_2) than S_4 when $J_{\#}$ is large, but little difference for small $J_{\#}$. This is to be expected because the magnitude of spurious precision under model misspecification is inversely related to power under correct model specification. Figure 16.2 exhibits the same patterns as in Figure 16.1, but the differences between the S functions and across $J_{\#}$ values are much smaller when $\rho = .75$ than when $\rho = 0$. The results for the S_1 and S_2



(a) Test Function $S_1(\cdot)$

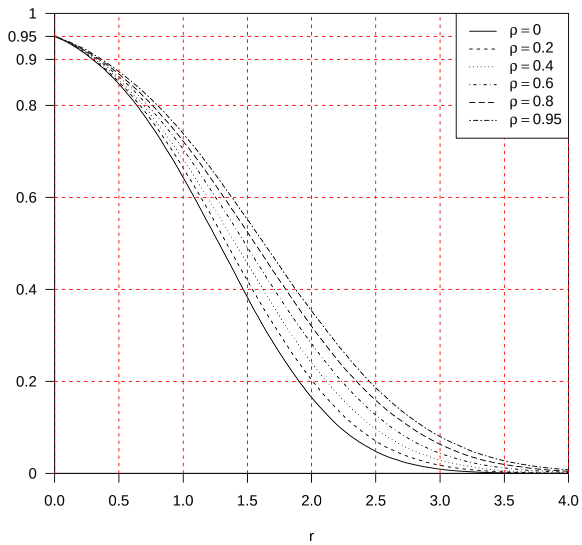


(b) Test Function $S_2(\cdot)$

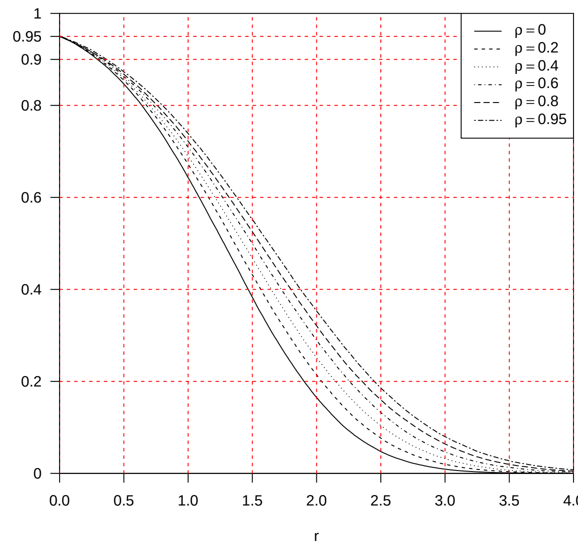


(c) Test Function $S_4(\cdot)$

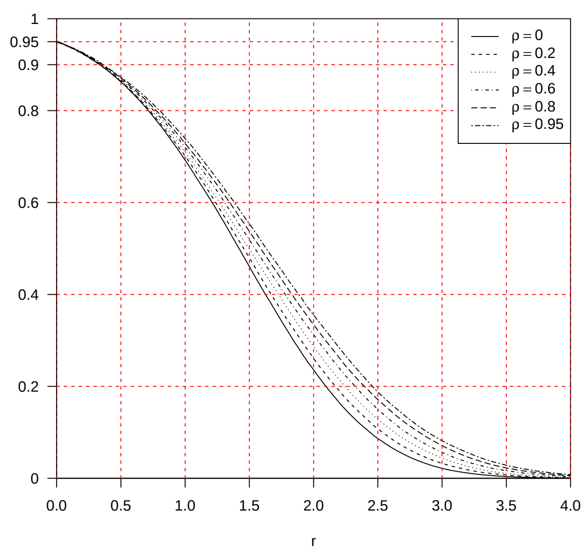
Figure 17.2: Maximum Coverage Probabilities for any $\theta \in \Theta$ for a Standard 95% GMS Confidence Set under Model Misspecification Indexed by r: $J_\# = 1, 2, \ldots, 15$, $\rho = .75$, (a) Test Function $S_1(\cdot)$, (b) Test Function $S_2(\cdot)$, and (c) Test Function $S_4(\cdot)$



(a) Test Function $S_1(\cdot)$



(b) Test Function $S_2(\cdot)$



(c) Test Function $S_4(\cdot)$

Figure 17.3: Maximum Coverage Probabilities for any $\theta \in \Theta$ for a Standard 95% GMS Confidence Set under Model Misspecification Indexed by r: $J_{\#} = 2$, $\rho = 0, .2, \ldots, .95$, and (a) Test Function $S_1(\cdot)$, (b) Test Function $S_2(\cdot)$, and (c) Test Function $S_4(\cdot)$

functions are quite similar. Figure 16.3 shows that, for all three S functions, spurious precision is greatest for $\rho = 0$ and least for $\rho = .95$, but the differences across ρ values are not huge in the case considered where $J_{\#} = 2$.

17.2 Asymptotic Results for Spurious Precision

Here, we provide an expression for the maximum asymptotic coverage probability for any $\theta \in \Theta$ for a standard GMS CS. The standard test statistic is of the form

$$S_{n,Std}(\theta) := S\left(n^{1/2}\widehat{m}_n(\theta), \widehat{\Omega}_n(\theta)\right), \tag{17.1}$$

where $S(m, \Omega)$ is a test function defined as in Andrews and Soares (2010) with $m \in R^k$ and $\Omega \in \Psi$, and Ψ is a specified closed set of $k \times k$ correlation matrices. We assume $S(m, \Omega)$ satisfies Assumptions S.1–S.4. Examples of $S(m, \Omega)$ functions that satisfy these assumptions are given in (4.3).

Let $\widehat{c}_n(\theta, 1-\alpha)$ denote the GMS critical value defined in Andrews and Soares (2010) using a constant κ_n, such as $\kappa_n = (\ln n)^{1/2}$, where $\kappa_n \to \infty$ and $\kappa_n/n^{1/2} \to 0$.[25]

We consider a set $\mathcal{P}_n$ of distributions F for which one or more moment inequalities is violated by at least $r/n^{1/2}$, and the other moment inequalities are slack by at least $d_n/n^{1/2}$ for all $\theta \in \Theta$, where $d_n\kappa_n^{-1} \to \infty$. Let $\Omega_F(\theta) := Var_F(\widetilde{m}(W_i, \theta)) \in R^{k \times k}$ denote the variance/correlation matrix of $\widetilde{m}(W_i, \theta)$ under F. Let $\mathcal{J} := \{1, \ldots, k\}$.

Define

$$\begin{aligned}
\mathcal{P}_n := \{F : \forall\theta \in \Theta, \exists J(\theta) \subset \mathcal{J} \text{ with } J(\theta) \neq \varnothing \text{ such that} \\
E_F\widetilde{m}_j(W_i, \theta) \leq -r/n^{1/2} \text{ if } j \in J(\theta) \text{ and} \\
E_F\widetilde{m}_j(W_i, \theta) \geq d_n/n^{1/2} \text{ if } j \in \mathcal{J} \setminus J(\theta), \text{ and } \Omega_F(\theta) \in \Psi\}, \text{ and} \\
\mathcal{L}\Psi := \{(\ell, \Omega) \in R^k_{[\pm\infty]} \times \Psi : \text{for some subsequence } \{a_n\}_{n \geq 1} \text{ of } \{n\} \text{ with} \\
(\theta_{a_n}, F_{a_n}) \in \Theta \times \mathcal{P}_{a_n},\ a_n^{1/2} E_{F_{a_n}}\widetilde{m}(W, \theta_{a_n}) \to \ell \text{ and } \Omega_{F_{a_n}}(\theta_{a_n}) \to \Omega\}.
\end{aligned} \tag{17.2}$$

By the definition of $\mathcal{P}_n$, for $(\ell, \Omega) \in \mathcal{L}\Psi$, $\ell_j \leq -r$ or $\ell_j = \infty$ $\forall j \leq k$, where $\ell = (\ell_1, ..., \ell_k)'$.

For $\ell \in R^k_{[\pm\infty]}$, let $c_\ell(\Omega, 1-\alpha)$ denote the $1-\alpha$ quantile of $S(\Omega^{1/2}Z^* + \ell, \Omega)$, where $Z^* \sim N(0_k, I_k)$. For $\ell \in R^k_{[\pm\infty]}$, define $\pi(\ell) := (\pi_1(\ell), \ldots, \pi_k(\ell))'$ by $\pi_j(\ell) := \infty 1(\ell_j = \infty)$ for $j \leq k$, where $\infty \cdot 0 := 0$.

An upper bound on the maximum asymptotic coverage probability for any $\theta \in \Theta$ for GMS CS's

[25] We assume that the GMS function $\varphi(\xi, \Omega)$ satisfies Assumption A.4 of Bugni, Canay, and Guggenberger (2012) with $\xi = 0$ replaced with $\xi \leq 0$ in part (b).

under $\{\mathcal{P}_n\}_{n\geq 1}$ misspecification is given in the following lemma, which is proved using results in Bugni, Canay, and Guggenberger (2012). For example, if the upper bound for a nominal .95 CS is .70, then the asymptotic coverage probability for any potential pseudo-true value is at most .70, which indicates spurious precision of the CS.

Lemma 17.1 *Suppose the observations $\{W_i\}_{i\leq n}$ are i.i.d under each $F \in \mathcal{P}_n$ and $0 < \alpha < 1/2$. Under Assumptions* S.1*,* S.3*, and* S.4 (*stated in the Appendix*),

$$\limsup_{n\to\infty} \sup_{(\theta,F)\in\Theta\times\mathcal{P}_n} P_F\left(S_{n,Std}(\theta) \leq \widehat{c}_n(\theta, 1-\alpha)\right) \leq \sup_{(\ell,\Omega)\in\mathcal{L}\Psi} P\left(S(\Omega^{1/2}Z^* + \ell, \Omega) \leq c_{\pi(\ell)}(\Omega, 1-\alpha)\right).$$

Comments. (i). For the test functions $S(\cdot) = S_1(\cdot)$ and $S_4(\cdot)$, we show in the following subsection that the upper bound in Lemma 17.1 is strictly less than $1-\alpha$ for all $r > 0$.[26] Hence, these GMS CS's exhibit spurious precision under misspecification.

(ii). Under a mild condition, the inequality in Lemma 17.1 holds as an equality. Let $(\ell_\infty, \Omega_\infty) \in \mathcal{L}\Psi$ be a point that achieves the supremum on the right-hand side in Lemma 17.1. (Such a point always exists.) Let $J_\infty \subset \mathcal{J}$, denote the set of indices j for which $\ell_{\infty j} < \infty$, where $\ell_\infty = (\ell_{\infty 1}, ..., \ell_{\infty k})'$. Let $\ell(J_\infty, -r)$ denote the vector in $R^k_{[\pm\infty]}$ with jth element equal to $-r$ for $j \in J_\infty$ and all other elements equal to infinity. The inequality in Lemma 17.1 holds as an equality if $\ell(J_\infty, -r) \in \mathcal{L}_\infty := \{\ell \in R^k_{[\pm\infty]} : (\ell, \Omega_\infty) \in \mathcal{L}\Psi\}$.

The right-hand side in Lemma 17.1 equals $MaxCPM(r;\Omega_\infty, J_\infty) = P(S(\Omega_\infty^{1/2}Z^* + \ell(J_\infty, -r), \Omega_\infty) \leq c_{\ell(J_\infty,0)}(\Omega_\infty, 1-\alpha))$ for $\ell(J_\infty, -r)$ defined in Comment (ii), where $MaxCPM$ abbreviates "maximum coverage probability under misspecification".

17.3 Proof of Lemma 17.1 and Comment (i) to Lemma 17.1

Proof of Lemma 17.1. There always exists a subsequence $\{q_n\}$ of $\{n\}$ such that

$$\limsup_{n\to\infty} \sup_{(\theta,F)\in\Theta\times\mathcal{P}_n} P_F\left(S_{n,Std}(\theta) \leq \widehat{c}_n(\theta, 1-\alpha)\right) = \lim P_{F_{q_n}}\left(S_{q_n,Std}(\theta_{q_n}) \leq \widehat{c}_{q_n}(\theta_{q_n}, 1-\alpha)\right), \quad (17.3)$$

where $(\theta_{q_n}, F_{q_n}) \in \Theta \times \mathcal{P}_n$ $\forall n \geq 1$. We can take a further subsequence $\{w_n\}$ of $\{q_n\}$ such that

$$w_n^{1/2} E_{F_{w_n}} \widetilde{m}_j(W_i, \theta_{w_n}) \to \ell_\infty \text{ and } \Omega_{F_{w_n}}(\theta_{w_n}) \to \Omega_\infty \quad (17.4)$$

for some $(\ell_\infty, \Omega_\infty) \in \mathcal{L}\Psi$.

[26] For any test function $S(\cdot)$ satisfying the conditions, the upper bound in Lemma 17.1 is less than or equal to $1-\alpha$.

Let $J_F(\theta)$ denote the set $J(\theta)$ corresponding to F in the definition of $\mathcal{P}_n$ in (17.2). The first convergence result in (17.4) and the definition of $\mathcal{P}_n$ (including $d_n\kappa_n^{-1} \rightarrow \infty$) give: (i) $J_{F_{w_n}}(\theta_{w_n}) = J_\infty := \{j \leq k : \ell_{\infty j} < \infty\}$ for all n large, where $\ell_\infty = (\ell_{\infty 1}, ..., \ell_{\infty k})'$ and (ii) $\kappa_{w_n}^{-1} w_n^{1/2} E_{F_{w_n}} \widetilde{m}_j(W_i, \theta_{w_n}) \rightarrow \xi_\infty = (\xi_{\infty 1}, ..., \xi_{\infty k})'$, where by definition $\xi_{\infty j} = 0$ if $\ell_{\infty j} < \infty$ and $\xi_{\infty j} = \infty$ if $\ell_{\infty j} = \infty$ (i.e., $\xi_\infty = \pi(\ell_\infty)$ and $\pi(\xi_\infty) = \xi_\infty$).

We have $c_{\pi(\xi_\infty)}(\Omega_\infty, 1-\alpha) > 0$ by the discussion following Assumption A.7 of Bugni, Canay, and Guggenberger (2012) (using $\xi_\infty \neq \infty 1_k$). We have $S_{w_n,Std}(\theta_{w_n}) \rightarrow_d S(\Omega_\infty^{1/2} Z^* + \ell_\infty, \Omega_\infty)$ by Lemma S1.1 in the Supplemental Material of Bugni, Canay, and Guggenberger (2012) using (17.4), and $\widehat{c}_{w_n}(\theta_{w_n}, 1-\alpha) \rightarrow_p c_{\pi(\xi_\infty)}(\Omega_\infty, 1-\alpha)$ by a similar argument to that given in the proof of Lemma 2 of Andrews and Soares (2010) using (17.4) and the results in the previous paragraph. Finally, by applying Lemma 5 of Andrews and Guggenberger (2010) to the right-hand side (rhs) of (17.3), we obtain

$$\lim P_{F_{w_n}} \left(S_{w_n,Std}(\theta_{w_n}) \leq \widehat{c}_{w_n}(\theta_{w_n}, 1-\alpha)\right) = P\left(S(\Omega_\infty^{1/2} Z^* + \ell_\infty, \Omega_\infty) \leq c_{\pi(\ell_\infty)}(\Omega_\infty, 1-\alpha)\right). \tag{17.5}$$

The left-hand side of (17.3) equals the rhs of (17.5) because a subsequence has the same limit as the original sequence.

For any $(\ell, \Omega) \in \mathcal{L}\Psi$, we have $\ell_j \leq -r$ or $\ell_j = \infty$ $\forall j \leq k$, where $\ell = (\ell_1, ..., \ell_k)'$, by the definition of $\mathcal{P}_n$ in (17.2). Thus, $\ell_j \leq \pi(\ell_j)$ $(:= \infty 1(\ell_j = \infty))$ $\forall j \leq k$. In consequence, using Assumption S.1(i), we obtain $c_\ell(\Omega, 1-\alpha) \geq c_{\pi(\ell)}(\Omega, 1-\alpha)$ $\forall (\ell, \Omega) \in \mathcal{L}\Psi$.

We have

$$\begin{aligned} & P\left(S(\Omega_\infty^{1/2} Z^* + \ell_\infty, \Omega_\infty) \leq c_{\pi(\ell_\infty)}(\Omega_\infty, 1-\alpha)\right) \\ \leq & \sup_{(\ell,\Omega)\in\mathcal{L}\Psi} P\left(S(\Omega^{1/2} Z^* + \ell, \Omega) \leq c_{\pi(\ell)}(\Omega, 1-\alpha)\right) \\ \leq & \ 1-\alpha, \end{aligned} \tag{17.6}$$

where the first inequality holds because $(\ell_\infty, \Omega_\infty) \in \mathcal{L}\Psi$ by (17.4) and the second inequality holds because $c_\ell(\Omega, 1-\alpha)$ is the $1-\alpha$ quantile of $S(\Omega^{1/2} Z^* + \ell, \Omega)$ and $c_\ell(\Omega, 1-\alpha) \geq c_{\pi(\ell)}(\Omega, 1-\alpha)$ by the previous paragraph. Equations (17.3), (17.5), and (17.6) combine to prove the lemma. □

Now, we prove the result stated in Comment (i) to Lemma 17.1: "For the test functions $S(\cdot) = S_1(\cdot)$ and $S_4(\cdot)$, the upper bound in Lemma 17.1 is strictly less than $1-\alpha$ for all $r > 0$." The proof uses the following lemma.

Lemma 17.2 *Suppose X and Y are random variables with $1-\alpha$ quantiles c_X and c_Y, respectively,* *for some* $\alpha \in (0, 1/2)$, $X \leq Y$ *a.s., and* $P(X \leq c_X) = P(Y \leq c_Y) = 1 - \alpha$. *Then,* $c_X = c_Y$ *iff* $P(X \leq c_X, Y > c_X) = 0$.

Proof of Comment (i) to Lemma 17.1. If $c_{\ell_\infty}(\Omega_\infty, 1-\alpha) > c_{\pi(\ell_\infty)}(\Omega_\infty, 1-\alpha)$, then (17.3), (17.5), and Assumptions S.4(i) and (ii) establish the desired result:

$$\limsup_{n\to\infty} \sup_{(\theta,F)\in\Theta\times\mathcal{P}_n} P_F\left(S_{n,Std}(\theta) \leq \widehat{c}_n(\theta, 1-\alpha)\right) < 1-\alpha. \tag{17.7}$$

We have $\ell_\infty \leq \pi(\ell_\infty)$ with a strict inequality holding for one or more elements, because $\ell_{\infty j} \leq -r < 0 = \pi(\ell_\infty)_j$ for $j \in J_\infty := \{j \leq k : \ell_{\infty j} < \infty\}$, $\ell_{\infty j} = \pi(\ell_\infty)_j = \infty$ for $j \in \mathcal{J}\backslash J_\infty$, and $J_\infty \neq \varnothing$.

For notational simplicity, let

$$S_\ell := S(\Omega_\infty^{1/2} Z^* + \ell, \Omega_\infty) \text{ for } \ell \in R^k_{[\pm\infty]}. \tag{17.8}$$

By the discussion following Assumption A.7 of Bugni, Canay, and Guggenberger (2012), $c_\ell(\Omega_\infty, 1-\alpha) > 0$ for $\ell = \ell_\infty, \pi(\ell_\infty)$ and $\alpha \in (0, 1/2)$. Hence, by Assumption S.4, we have

$$P\left(S_\ell \leq c_\ell(\Omega_\infty, 1-\alpha)\right) = 1-\alpha \text{ for } \ell = \ell_\infty, \pi(\ell_\infty). \tag{17.9}$$

In addition, $S_{\pi(\ell_\infty)} \leq S_{\ell_\infty}$ a.s. by Assumption S.1(i) because $\ell_\infty \leq \pi(\ell_\infty)$. Using these results and Lemma 17.2, we have: $c_{\pi(\ell_\infty)}(\Omega_\infty, 1-\alpha) = c_{\ell_\infty}(\Omega_\infty, 1-\alpha)$ if and only if

$$P\left(S_{\pi(\ell_\infty)} \leq c_{\pi(\ell_\infty)}(\Omega_\infty, 1-\alpha), S_{\ell_\infty} > c_{\pi(\ell_\infty)}(\Omega_\infty, 1-\alpha)\right) = 0. \tag{17.10}$$

Thus, to prove the result of Comment (i), it suffices to show

$$P\left(S_{\pi(\ell_\infty)} \leq c, S_{\ell_\infty} > c\right) > 0 \tag{17.11}$$

for arbitrary $c > 0$, for $S(\cdot) = S_1(\cdot)$ and $S_4(\cdot)$. In the following, let $c > 0$ be an arbitrary positive number.

We consider the case where $S(\cdot) = S_1(\cdot)$ first. By the definition of $S_1(\cdot)$, $S_\ell = \sum_{j\in J_\infty}[\omega_j' Z^* + \ell_j]_-^2$ for $\ell = \ell_\infty, \pi(\ell_\infty)$, where ω_j denotes the jth column of $\Omega_\infty^{1/2}$ (because $\ell_{\infty j} = \pi(\ell_{\infty j}) = \infty$ for $j \in \mathcal{J}\backslash J_\infty$). Let v (> 0) denote the number of elements in J_∞. We have $\sum_{j\in J_\infty}[\omega_j' Z^* - \ell_{\infty j}]_-^2 > c$ if $\sum_{j\in J_\infty}[\omega_j' Z^* - r]_-^2 > c$ (because $\ell_{\infty j} \leq -r$ for $j \in J_\infty$), and the latter holds if $\omega_j' Z^* < r - \sqrt{c/v}$

for all $j \in J_\infty$. In addition, $\sum_{j \in J_\infty} [\omega_j' Z^*]_-^2 \leq c$ if $\omega_j' Z^* \geq -\sqrt{c/v}$ for all $j \in J_\infty$. It follows that

$$P\left(S_{\pi(\ell_\infty)} \leq c, S_{\ell_\infty} > c\right) \geq P\left(\bigcap_{j \in J_\infty} \{\omega_j' Z^* \in (-\sqrt{c/v}, r - \sqrt{c/v}]\}\right) > 0, \tag{17.12}$$

where the last inequality holds because the probability on its left-hand side is the probability that a multivariate normal v-vector with positive definite variance matrix lies in a set with positive Lebesgue measure (on R^v). This completes the proof of the result of Comment (i) for $S(\cdot) = S_1(\cdot)$.

Next, we consider the case where $S(\cdot) = S_4(\cdot)$. By the definition of $S_4(\cdot)$, $S_\ell = \max_{j \in J_\infty} [\omega_j' Z^* + \ell_j]_-^2$ for $\ell = \ell_\infty, \pi(\ell_\infty)$ (because $\ell_{\infty j} = \pi(\ell_{\infty j}) = \infty$ for $j \in \mathcal{J} \backslash J_\infty$). Thus, if $\omega_j' Z^* < r - \sqrt{c}$ for all $j \in J_\infty$, then $S_{\ell_\infty} > c$ (because $\ell_{\infty j} \leq -r$ for all $j \in J_\infty$). Also, if $\omega_j' Z^* \geq -\sqrt{c}$ for all $j \in J_\infty$, then $S_{\pi(\ell_\infty)} \leq c$. Hence, we obtain

$$P\left(S_{\pi(\ell_\infty)} \leq c, S_{\ell_\infty} > c\right) \geq P\left(\bigcap_{j \in J_\infty} \{\omega_j' Z^* \in (-\sqrt{c}, r - \sqrt{c}]\}\right) > 0, \tag{17.13}$$

where the last inequality holds for the same reason as given for the last inequality in (17.12). This completes the proof of the result of Comment (i) for $S(\cdot) = S_4(\cdot)$. □

Proof of Lemma 17.2. We have

$$\begin{aligned} 1 - \alpha = P(X \leq c_X) &= P(X \leq c_X, Y > c_X) + P(X \leq c_X, Y \leq c_X) \\ &= P(X \leq c_X, Y > c_X) + P(Y \leq c_X), \end{aligned} \tag{17.14}$$

where the first equality holds by assumption and the last equality holds because $X \leq Y$ a.s.

If $c_X = c_Y$, we have

$$1 - \alpha = P(X \leq c_X, Y > c_X) + P(Y \leq c_Y) = P(X \leq c_X, Y > c_X) + 1 - \alpha, \tag{17.15}$$

where the first equality holds by (17.14) and $c_X = c_Y$ and the second equality holds because $P(Y \leq c_Y) = 1 - \alpha$. Thus, the "only if" result of the lemma is proved.

If $P(X \leq c_X, Y > c_X) = 0$, then, by (17.14), $1 - \alpha = P(Y \leq c_X)$. Since $c_Y := \min\{y : P(Y \leq y) \geq 1 - \alpha\}$, this implies that $c_Y \leq c_X$. But, $X \leq Y$ a.s. implies $c_X \leq c_Y$. Hence, in this case, $c_X = c_Y$, which establishes the "if" result of the lemma. □

18 Additional Simulation Results for the Lower/Upper Bound Model

Here we provide some additional simulation results for the lower/upper bound model considered in Section 8. We give results for $k = 4$ and 8.

Figure 18.1 shows the rejection probabilities for the misspecified case with $k = 4$ under the "very slack" and "slack/almost binding" scenarios that were not reported in the main paper. Figure 18.2 does likewise for the correctly-specified case. The two figures confirm what we have already seen in the main paper: (i) when the model is misspecified, the SPUR1 and SPUR2 tests perform quite similarly, with their rejection probabilities reaching 1 fairly quickly as the distance between the null value and the MR identified set increases, and (ii) when the model is correctly specified, the SPUR2 test performs similarly to the GMS test when the length of identified set is .5 or larger, and likewise for the SPUR1 test when the length is 1. Again, we see that the SPUR2 test performs better than the SPUR1 test when the identified set is small, but not too small.

Next, we consider cases with $k = 8$. In this case, the moment inequalities are given as

$$\begin{aligned} &E_F W_{ij} \leq \theta \text{ for } 1 \leq j \leq 4 \text{ and} \\ &\theta \leq E_F W_{ij} \text{ for } 5 \leq j \leq 8. \end{aligned} \tag{18.1}$$

The definition of each scenario is analogous to the $k = 4$ cases, with each entry repeated twice. That is, if $\mu^4 = (\mu_1, \mu_2, \mu_3, \mu_4)' \in R^4$ is the mean vector used under some scenario for $k = 4$, then $\mu^8 = (\mu_1, \mu_1, \mu_2, \mu_2, \mu_3, \mu_3, \mu_4, \mu_4)' \in R^8$ is the mean vector used in the same scenario for $k = 8$. Figures 18.3 and Figure 18.4 give the simulation results for $k = 8$. These results show that the same qualitative results hold as for $k = 8$ as for $k = 4$.

19 Details for the Missing Data Model

In this section, we provide additional details for the missing data model considered in Section 8.2. Specifically, we provide derivations for (8.5), (8.6), and the line following (8.6), which gives an expression for the MR identified set.

Let $p_j := P(X_i = x_j) > 0$ for $j \leq 3$. In the simulations, we take $p_j = 1/3$ for $j \leq 3$. Some

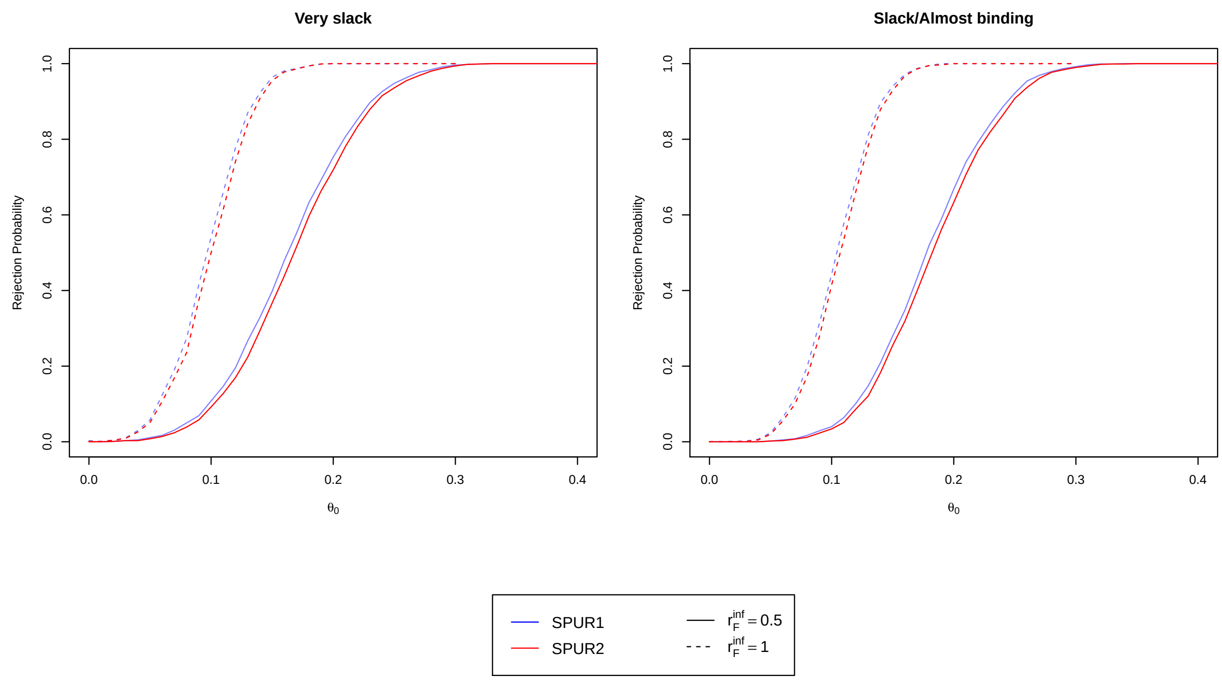


Figure 18.1: Rejection probabilities for (additional) misspecified cases for $k = 4$. Each plot shows, under different scenarios, the rejection probabilities of the SPUR1 and SPUR2 tests for the null hypothesis $H_0 : \theta = \theta_0$ for a range of θ_0 values, identified set $\Theta_I(F) = \{0\}$, and two different values of $r_F^{\inf}$.

calculations give

$$\begin{aligned} E_F m_1(W,\theta) &= p_1\theta_1, \\ E_F m_2(W,\theta) &= -p_2(\theta_1 + \widetilde{r}), \text{ and} \\ E_F m_3(W,\theta) &= p_3\theta_2. \end{aligned} \tag{19.1}$$

In consequence, the model is misspecified if and only if $\widetilde{r} > 0$, as stated in Section 8. If $\widetilde{r} \le 0$, $r_F^{\inf} = 0$.

Now, suppose $\widetilde{r} > 0$. Additional calculations give

$$\begin{aligned} Var_F\left(m_1(W,\theta)\right) &= (p_1 - p_1^2)\theta_1^2 + p_1 p_z, \\ Var_F\left(m_2(W,\theta)\right) &= (p_2 - p_2^2)(\theta_1 + \widetilde{r})^2 + p_2\left((1+\widetilde{r})^2(1/p_z - 1) + p_z\right), \text{ and} \\ Var_F\left(m_3(W,\theta)\right) &= (p_3 - p_3^2)\theta_2^2 + p_3 p_z. \end{aligned} \tag{19.2}$$

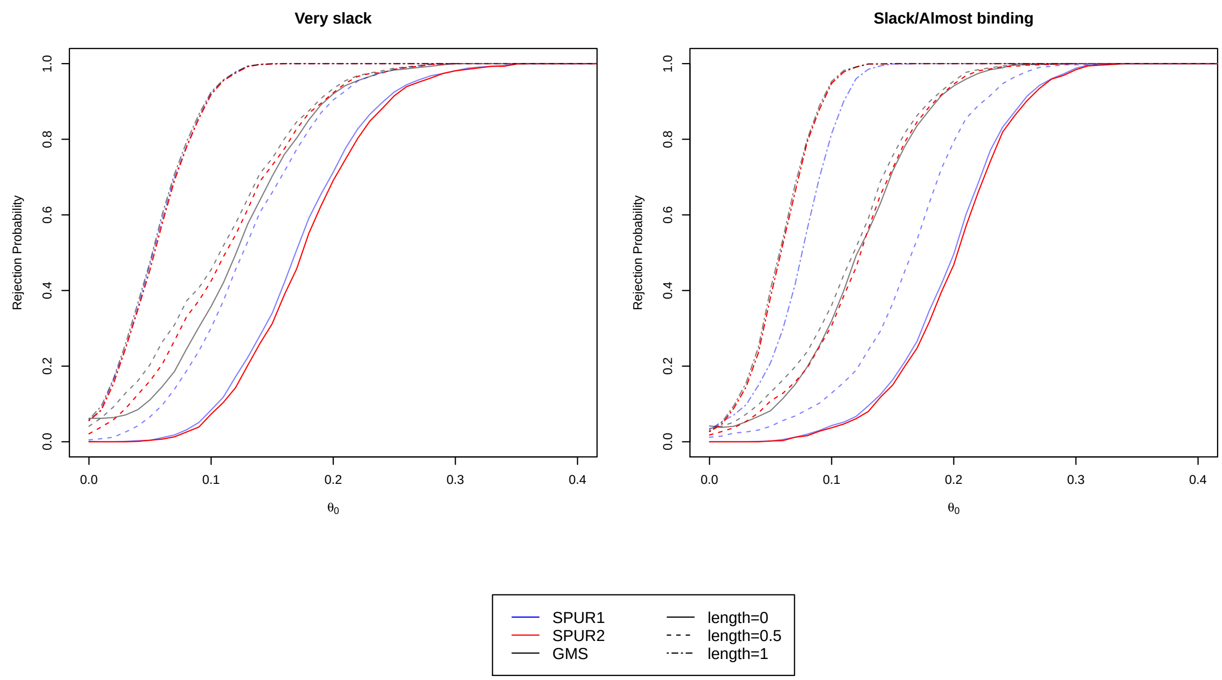


Figure 18.2: Rejection probabilities for (additional) correctly specified cases for $k = 4$. Each plot shows, under different scenarios, the rejection probabilities of the SPUR1, SPUR2 and standard GMS tests for the null hypothesis $H_0 : \theta = \theta_0$ for a range of θ_0 values and different lengths ℓ of the identified set $\Theta_I(F) = [-\ell, 0]$.

We relax the (standardized) inequalities by r. Then, by (19.1) and (19.2), the inequalities are

$$\begin{aligned} \frac{p_1\theta_1}{((p_1 - p_1^2)\theta_1^2 + p_1 p_z)^{1/2}} &\geq -r, \\ -\frac{p_2(\theta_1 + \widetilde{r})}{((p_2 - p_2^2)(\theta_1 + \widetilde{r})^2 + p_2((1+\widetilde{r})^2(1/p_z - 1) + p_z))^{1/2}} &\geq -r, \text{ and} \\ \frac{p_3\theta_2}{((p_3 - p_3^2)\theta_2^2 + p_3 p_z)^{1/2}} &\geq -r. \end{aligned} \tag{19.3}$$

By definition, $r_F^{\inf}$ is the smallest $r > 0$ such that there exists some $\theta \in \Theta$ that satisfies (19.3). The third inequality does not play a role in determining $r_F^{\inf}$. Hence, we focus on finding the smallest $r > 0$ such that there exists some θ_1 that satisfies the first two inequalities.

For arbitrary numbers a, b, and c with $a > 0$ and $b > 0$, consider the function

$$h(\theta_1) = \frac{\theta_1 + c}{(a(\theta_1 + c)^2 + b)^{1/2}}. \tag{19.4}$$

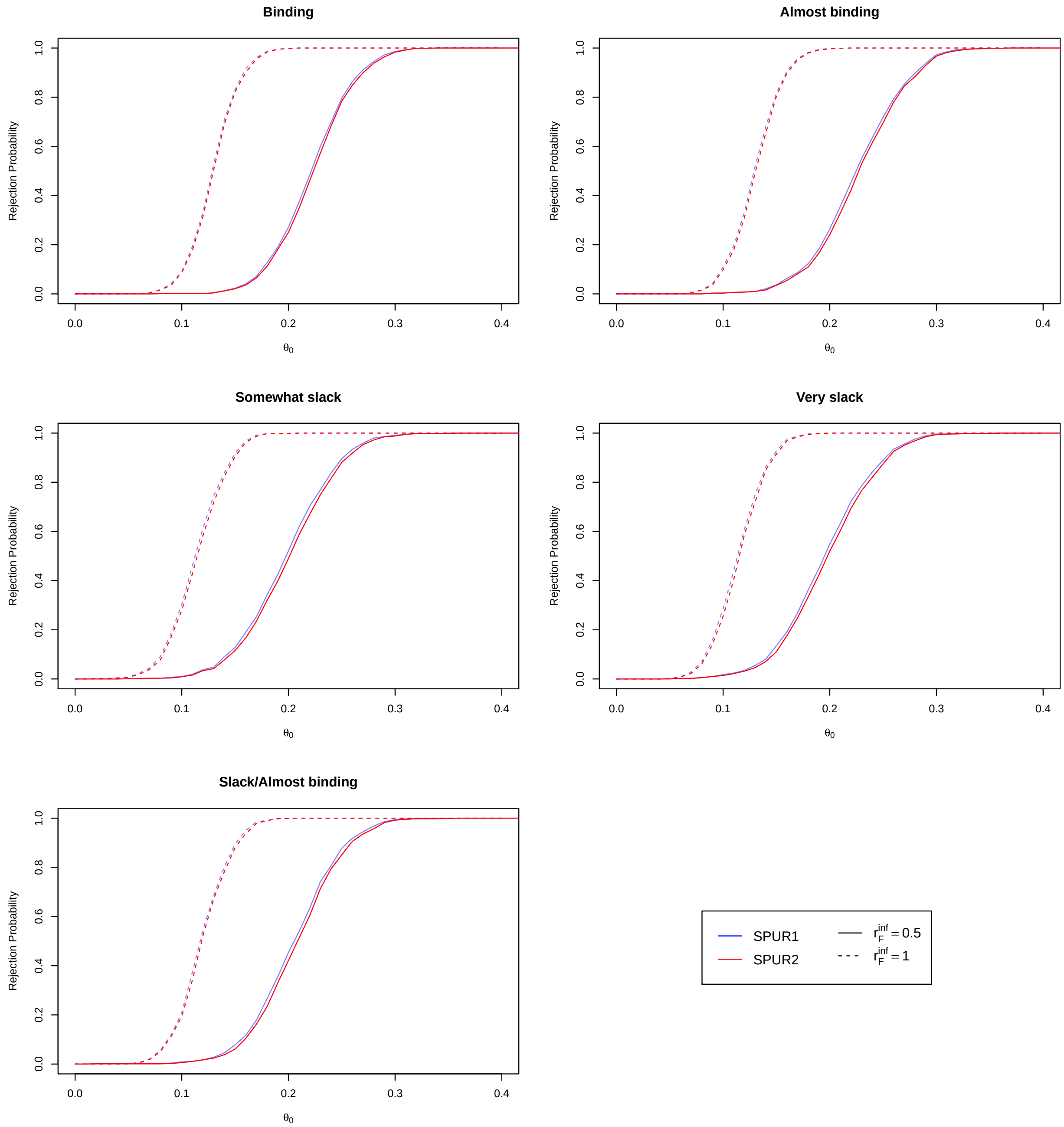


Figure 18.3: Rejection probabilities for misspecified cases for $k = 8$. Each plot shows, under different scenarios, the rejection probabilities of the SPUR1 and SPUR2 tests for the null hypothesis $H_0 : \theta = \theta_0$ for a range of θ_0 values, identified set $\Theta_I(F) = \{0\}$, and two different values of r_F^{inf}.

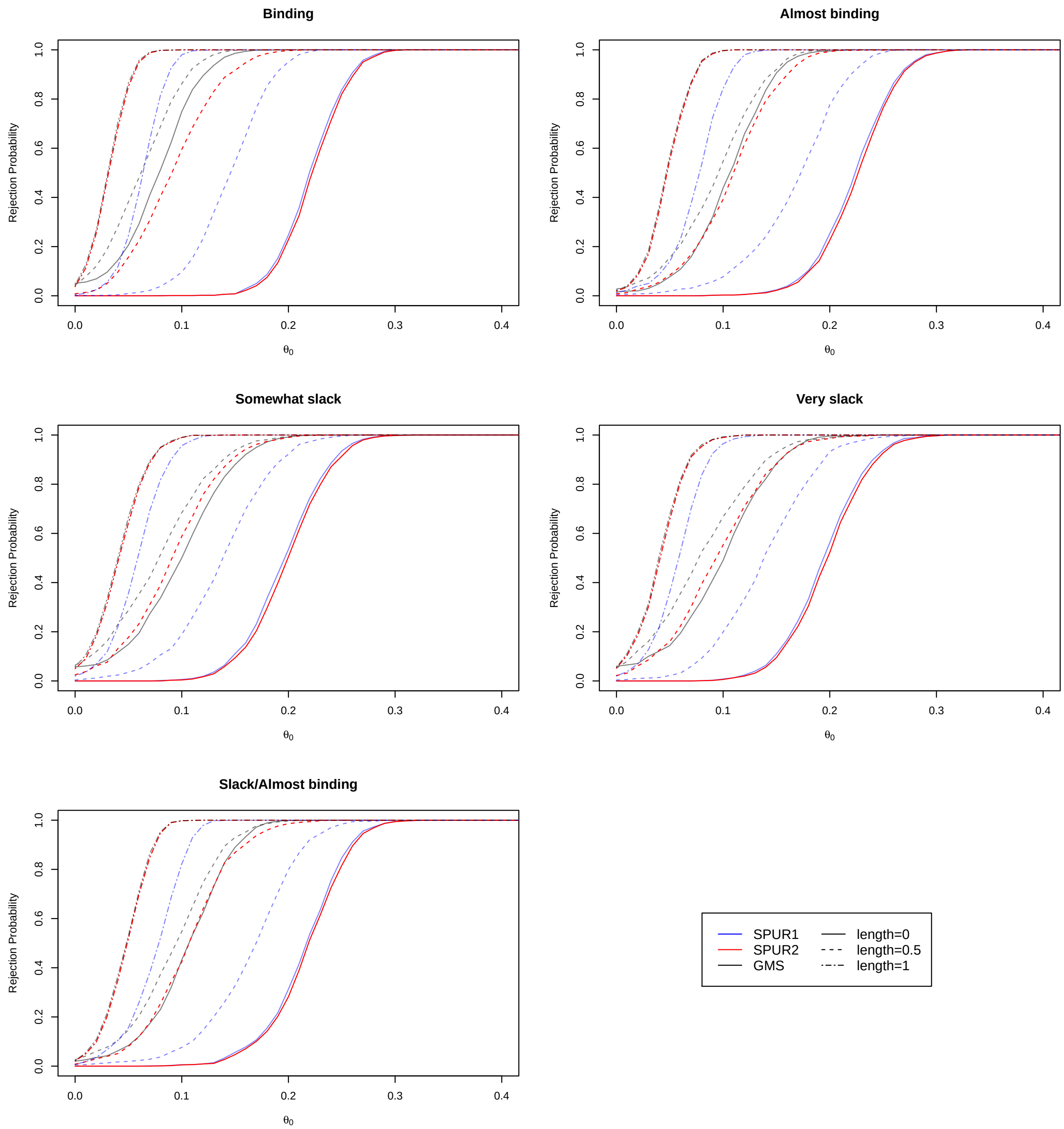


Figure 18.4: Rejection probabilities for correctly specified cases for $k = 8$. Each plot shows, under different scenarios, the rejection probabilities of the SPUR1, SPUR2 and standard GMS tests for the null hypothesis $H_0 : \theta = \theta_0$ for a range of θ_0 values and different lengths ℓ of the identified set $\Theta_I(F) = [-\ell, 0]$.

Calculation of the first derivative of $h(\cdot)$ shows that $h(\cdot)$ is strictly increasing. This implies that the left-hand sides of the first and second inequalities in (19.3) are strictly increasing and strictly decreasing functions of θ_1, respectively. Hence, if we let $\underline{\theta}_1(r)$ and $\overline{\theta}_1(r)$ denote the θ_1 values that solve the first and second inequalities as equalities, respectively, then θ_1 satisfies the two inequalities if and only if θ_1 lies in $[\underline{\theta}_1(r), \overline{\theta}_1(r)]$, where this interval is defined to be empty if $\underline{\theta}_1(r) > \overline{\theta}_1(r)$.

Some algebra gives

$$\begin{aligned} \underline{\theta}_1(r) &= -\left(\frac{p_z}{(p_1/r^2 + p_1 - 1))}\right)^{1/2} \text{ and} \\ \overline{\theta}_1(r) &= \left(\frac{(1+\widetilde{r})^2(1/p_z - 1) + p_z}{p_2/r^2 + p_2 - 1)}\right)^{1/2} - \widetilde{r}. \end{aligned} \tag{19.5}$$

Hence, if r is such that

$$\widetilde{r} \leq \left(\frac{(1+\widetilde{r})^2(1/p_z - 1) + p_z}{p_2/r^2 + p_2 - 1}\right)^{1/2} + \left(\frac{p_z}{p_1/r^2 + p_1 - 1)}\right)^{1/2}, \tag{19.6}$$

then the MR identified set under the relaxation r is non-empty. Since the rhs is increasing in r, $r_F^{\inf}$ must solve (19.6) as an equality. That is, $r_F^{\inf}$ is the value of r that makes $\underline{\theta}_1(r) = \overline{\theta}_1(r)$. Assuming $p_1 = p_2$, this gives

$$r_F^{\inf} = \left(\frac{p_1 \widetilde{r}^2}{\left(p_z^{1/2} + ((1+\widetilde{r})^2(1/p_z - 1) + p_z)^{1/2}\right)^2 + (1-p_1)\widetilde{r}^2}\right)^{1/2}. \tag{19.7}$$

Taking $p_1 = p_2 = 1/3$ gives (8.5).

Plugging the expression for $r_F^{\inf}$ in place of r in (19.5) gives

$$\underline{\theta}_1(r_F^{\inf}) = \overline{\theta}_1(r_F^{\inf}) = -\frac{p_z^{1/2}\widetilde{r}}{p_z^{1/2} + ((1+\widetilde{r})^2(1/p_z - 1) + p_z)^{1/2}} =: \theta_1^I(\widetilde{r}). \tag{19.8}$$

Thus, the only θ_1 value that satisfies (19.3) with $r = r_F^{\inf}$ is $\theta_1 = \theta_1^I(\widetilde{r})$. This gives (8.6).

Now, plugging in $r_F^{\inf}$ in place of r in the third inequality of (19.3) and taking $p_1 = p_2 = p_3 = 1/3$, one can see that any θ_2 such that $\theta_2 \geq \theta_1^I(\widetilde{r})$ satisfies (19.3) (with $r_F^{\inf}$ in place of r). This shows that $\Theta_I(F) = \{\theta_1^I(\widetilde{r})\} \times [\theta_1^I(\widetilde{r}), \infty)$.

20 Lemma 20.1 and Proofs of Lemmas 5.1, 5.2, and 20.1

The following is a sufficient condition for Assumption NLA, which first appears in Section 5.1.

Assumption LA. The null values $\{\theta_n\}_{n\geq 1}$ and distributions $\{F_n\}_{n\geq 1}$ satisfy: (i) $||\theta_n - \theta_{In}|| = O(n^{-1/2})$ for some sequence $\{\theta_{In} \in \Theta_I(F_n)\}_{n\geq 1}$, (ii) $n^{1/2}(E_{F_n}\widetilde{m}_j(W,\theta_{In}) + r_{F_n}^{\inf}) \to h_{Ij\infty}$ for some $h_{Ij\infty} \in R_{[\pm\infty]}$ $\forall j \leq k$, and (iii) $E_F\widetilde{m}(W,\theta)$ is Lipschitz on Θ uniformly over $\mathcal{P}$, i.e., there exists a constant $K < \infty$ such that $||E_F\widetilde{m}(W,\theta_1) - E_F\widetilde{m}(W,\theta_2)|| \leq K||\theta_1 - \theta_2||$ $\forall \theta_1, \theta_2 \in \Theta$, $\forall F \in \mathcal{P}$.

Under Assumption LA, $\{\theta_n\}_{n\geq 1}$ is a sequence of $n^{-1/2}$-local alternatives to the null hypothesis $\forall n \geq 1$. Assumption LA(ii) is the same as Assumption C.3 with $\{\theta_n\}_{n\geq 1}$ replaced by some sequence $\{\theta_{In}\}_{n\geq 1}$ in the MR identified set(s). Hence, by Lemma 5.1(a), $h_{Ij\infty} \geq 0$ $\forall j \leq k$.

A sufficient condition for Assumption CA is the following fixed alternative assumption.

Assumption FA. The null values $\{\theta_n\}_{n\geq 1}$ and distributions $\{F_n\}_{n\geq 1}$ satisfy: (i) The distributions $F_n = F_* \in \mathcal{P}$ and the null values $\theta_n = \theta_* \in \Theta$ do not depend on $n \geq 1$ and (ii) $E_{F_*}\widetilde{m}_j(W,\theta_*) + r_{F_*}^{\inf} < 0$ for some $j \leq k$.

Lemma 20.1 *Under Assumption* C.3, (a) *Assumption* N *implies Assumption* NLA*,* (b) *Assumption* LA *implies Assumption* NLA, *and* (c) *Assumption* FA *implies Assumption* CA*.*

Proof of Lemma 5.1. Part (a) holds because $r_{F_n}^{\inf} \geq 0$ by its definition in (2.5). The first result in part (b) holds because $n^{1/2} \geq 1$. The second result in part (b) holds because $|\ell_{j\infty}| < \infty$ implies $n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta_n) = O(1)$, which implies that $\widetilde{m}_{j\infty} := \widetilde{m}_j(\theta_\infty) = \lim_{n\to\infty} E_{F_n}\widetilde{m}_j(W,\theta_n) = 0$, using Assumptions C.1, C.2, and C.4.

Now, we prove part (c). If $\theta \in \Theta_I(F)$, then $r_F(\theta) = r_F^{\inf}$ (by the definition of $\Theta_I(F)$ in (2.7)), $r_{Fj}(\theta) \leq r_F^{\inf}$ $\forall j \leq k$ (by the definition of $r_{Fj}(\theta)$ in (2.6)), and $r_{Fj}(\theta) = r_F^{\inf}$ for some $j \leq k$. In consequence,

$$
\begin{aligned}
0 &= \max_{j\leq k}(r_{Fj}(\theta) - r_F^{\inf}) = \max_{j\leq k}(\max\{-E_F\widetilde{m}_j(W,\theta), 0\} - r_F^{\inf}) \\
&\geq \max_{j\leq k}(-E_F\widetilde{m}_j(W,\theta) - r_F^{\inf}) = -\min_{j\leq k}(E_F\widetilde{m}_j(W,\theta) + r_F^{\inf}), \qquad (20.1)
\end{aligned}
$$

where the second equality holds by the definition of $r_{Fj}(\theta)$ and the inequality is trivial.

Using (20.1), if $\theta_n \in \Theta_I(F_n)$ for n large, then

$$
\begin{aligned}
0 &\leq \liminf_{n\to\infty} \min_{j\leq k} n^{1/2}(E_{F_n}\widetilde{m}_j(W,\theta_n) + r_{F_n}^{\inf}) \\
&= \min_{j\leq k} \liminf_{n\to\infty} n^{1/2}(E_{F_n}\widetilde{m}_j(W,\theta_n) + r_{F_n}^{\inf}) = \min_{j\leq k} h_{j\infty}, \qquad (20.2)
\end{aligned}
$$

where the first equality holds by a subsequence argument and the second equality uses Assumption C.3. This establishes part (c).

Lastly, we prove part (d). If $\theta \in \Theta_I(F)$ and the model is correctly specified, then

$$r_F^{\inf} = \max_{j \le k} r_{Fj}(\theta) = \max_{j \le k} \max\{-E_F \widetilde{m}(W, \theta), 0\} = 0, \tag{20.3}$$

where the first two equalities hold by the definitions of $r_F^{\inf}$ and $r_{Fj}(\theta)$ in (2.5) and (2.6), respectively, and the last equality holds because $E_F \widetilde{m}(W, \theta) \ge 0_k \ \forall \theta \in \Theta_I(F)$ by correct model specification, see (2.4).

Equation (20.3) implies that under correct model specification, if $\theta_n \in \Theta_I(F_n)$ for all n large, then

$$h_{j\infty} = \lim n^{1/2}(E_{F_n} \widetilde{m}_j(W, \theta_n) + r_{F_n}^{\inf}) = \lim n^{1/2} E_{F_n} \widetilde{m}_j(W, \theta_n) = \ell_{j\infty} \ \forall j \le k. \tag{20.4}$$

We have $h_{j\infty}, \ell_{j\infty}, \widetilde{m}_{j\infty} \ge 0$ under correct model specification when $\theta_n \in \Theta_I(F_n)$ for all n large, because the moment inequalities all hold at $\theta_n \in \Theta_I(F_n)$, i.e., $E_{F_n} \widetilde{m}_j(W, \theta_n) \ge 0$, under correct model specification. This completes the proof of part (d). $\square$

Proof of Lemma 5.2. Because $r_F^{\inf} := \inf_{\theta \in \Theta} \max_{j \le k} r_{Fj}(\theta)$, see (2.5) and (2.6), for all F and $\theta \in \Theta$, we have

$$\max_{j \le k}(r_{Fj}(\theta) - r_F^{\inf}) \ge 0, \tag{20.5}$$

which establishes part (a).

Any $(\theta, b, \ell) \in \Lambda$ is the limit of some sequence $(\theta_n, b_n, \ell_n) \in \Lambda_{n,F_n}$ because $\Lambda_{n,F_n} \to_H \Lambda$ by Assumption C.7. That is, $b_n \to b$ and $\max_{j \le k} b_{nj} \to \max_{j \le k} b_j$. This and (20.5) applied with $(\theta, F) = (\theta_n, F_n)$ give

$$0 \le \max_{j \le k} n^{1/2}(r_{F_n j}(\theta_n) - r_{F_n}^{\inf}) = \max_{j \le k} b_{nj} \to \max_{j \le k} b_j, \tag{20.6}$$

which proves part (b) of the lemma.

Next, we prove part (c). The function $r_{F_n}(\theta) - r_{F_n}^{\inf}$ is lower semi-continuous on Θ (since $E_F \widetilde{m}_j(W, \theta)$ is upper semi-continuous on Θ by Assumption A.0(ii)) and $[x]_- := \max\{-x, 0\}$, Θ is compact by Assumption A.0(i), and a lower semi-continuous function on a compact set achieves its infimum. Hence, there exists $\widetilde{\theta}_n \in \Theta$ such that $r_F(\widetilde{\theta}_n) = r_F^{\inf} \ \forall n \ge 1$, which establishes part (c).

For part (d), let $(\widetilde{\theta}_n, \widetilde{b}_n, \widetilde{\ell}_n) \in \Lambda_{n,F_n}$ be such that $\widetilde{\theta}_n \in \Theta_I(F_n) \ \forall n \ge 1$. Such $(\widetilde{\theta}_n, \widetilde{b}_n, \widetilde{\ell}_n)$ exist because $\Theta_I(F_n)$ is non-empty $\forall n \ge 1$ by part (c). There exists a subsequence $\{q_n\}_{n \ge 1}$ of $\{n\}_{n \ge 1}$ and a $(\widetilde{\theta}, \widetilde{b}, \widetilde{\ell}) \in \Theta \times R_{[\pm\infty]}^{2k}$ such that $d((\widetilde{\theta}_{q_n}, \widetilde{b}_{q_n}, \widetilde{\ell}_{q_n}),\ (\widetilde{\theta}, \widetilde{b}, \widetilde{\ell})) \to 0$ because $(\Theta \times R_{[\pm\infty]}^{2k}, d)$ is a compact metric space under Assumption A.0(i). We have $(\widetilde{\theta}, \widetilde{b}, \widetilde{\ell}) \in \Lambda$ by the following argument:

$$
\begin{aligned}
0 \leq \inf_{(\theta,b,\ell)\in\Lambda} d((\theta,b,\ell),(\widetilde{\theta},\widetilde{b},\widetilde{\ell})) \;\leq\; & \inf_{(\theta,b,\ell)\in\Lambda} d((\theta,b,\ell),(\widetilde{\theta}_{q_n},\widetilde{b}_{q_n},\widetilde{\ell}_{q_n})) + d((\widetilde{\theta}_{q_n},\widetilde{b}_{q_n},\widetilde{\ell}_{q_n}),(\widetilde{\theta},\widetilde{b},\widetilde{\ell})) \\
& \to 0, \qquad (20.7)
\end{aligned}
$$

where the second inequality holds by the triangle inequality and the convergence holds using Assumption C.7 (i.e., $\Lambda_{n,F_n} \to_H \Lambda$). Thus, $\inf_{(\theta,b,\ell)\in\Lambda} d((\theta,b,\ell),(\widetilde{\theta},\widetilde{b},\widetilde{\ell})) = 0$. This implies that $(\widetilde{\theta},\widetilde{b},\widetilde{\ell}) \in \Lambda$, because Λ is a compact subset of $(\Theta \times R^{2k}_{[\pm\infty]}, d)$ by Assumption C.7, $d((\theta,b,\ell),(\widetilde{\theta},\widetilde{b},\widetilde{\ell}))$ is a continuous function of (θ,b,ℓ), and a continuous function on a compact set attains its infimum.

Since $\widetilde{\theta}_n \in \Theta_I(F_n)$, $r_{F_n}(\widetilde{\theta}_n) = r^{\inf}_{F_n}$ $\forall n \geq 1$. Hence, for all $n \geq 1$,

$$
\max_{j\leq k} \widetilde{b}_{nj} = \max_{j\leq k} n^{1/2}([E_{F_n}\widetilde{m}_j(W,\widetilde{\theta}_n)]_- - r^{\inf}_{F_n}) = n^{1/2}(r_{F_n}(\widetilde{\theta}_n) - r^{\inf}_{F_n}) = 0, \qquad (20.8)
$$

where the first equality holds by the definition of Λ_{n,F_n} in (5.3) and the second equality holds by the definition of $r_F(\theta)$ in (2.6). We obtain

$$
\max_{j\leq k} \widetilde{b}_j = \lim_{n\to\infty} \max_{j\leq k} \widetilde{b}_{nj} = 0, \qquad (20.9)
$$

which proves part (d) of the lemma since $(\widetilde{\theta},\widetilde{b},\widetilde{\ell}) \in \Lambda$.

Given any $(\theta^*, b^*, \ell^*) \in \Lambda$, there exists a sequence $\{(\theta^*_n, b^*_n, \ell^*_n) \in \Lambda_{n,F_n}\}_{n\geq 1}$ such that $(\theta^*_n, b^*_n, \ell^*_n) \to (\theta^*, b^*, \ell^*)$ because $\Lambda_{n,F_n} \to_H \Lambda$ by Assumption C.7. Hence, if $|\ell^*_j| < \infty$, we have

$$
|\widetilde{m}_j(\theta^*)| = \lim |E_{F_n}\widetilde{m}_j(W,\theta^*_n)| = \lim(n^{-1/2}(|\ell^*_j| + o(1))) = 0, \qquad (20.10)
$$

where the first equality uses Assumption C.4. This establishes part (e). $\square$

Proof of Lemma 20.1. Under Assumption N, Lemma 5.1(a) implies that $h_{j\infty} \geq 0$ $\forall j \leq k$, which establishes Assumption NLA and part (a).

Now, we establish part (b). Under Assumption LA, for all $j \leq k$, we have

$$
n^{1/2}|E_{F_n}\widetilde{m}_j(W,\theta_n) - E_{F_n}\widetilde{m}_j(W,\theta_{In})| \leq Kn^{1/2}||\theta_n - \theta_{In}|| = O(1), \qquad (20.11)
$$

where the inequality holds by Assumption LA(iii) and the equality holds by Assumption LA(i). In consequence, for all $j \leq k$, we have

$$h_{j\infty} = \lim_{n\to\infty} n^{1/2}(E_{F_n}\widetilde{m}_j(W,\theta_n) + r^{\inf}_{F_n})$$
$$= \lim_{n\to\infty} n^{1/2}(E_{F_n}\widetilde{m}_j(W,\theta_{In}) + r^{\inf}_{F_n}) + O(1) = h_{Ij\infty} + O(1) \geq O(1), \quad (20.12)$$

where the first equality holds by Assumption C.3, the second equality holds by (20.11), the third equality holds by Assumption LA(ii), and the inequality holds by Lemma 5.1(a) with θ_{In} in place of θ_n using Assumption LA(ii) in place of Assumption C.3. This completes the proof of part (b).

Under Assumption FA, we have

$$\min_{j\leq k} h_{j\infty} = \min_{j\leq k} \lim n^{1/2}(E_{F_*}\widetilde{m}_j(W,\theta_*) + r^{\inf}_{F_*}) = -\infty, \quad (20.13)$$

where the second equality holds because $E_{F_*}\widetilde{m}_j(W,\theta_*) + r^{\inf}_{F_*} < 0$ for some $j \leq k$ by Assumption FA(ii). Thus, Assumption CA holds, which establishes part (c). $\square$

21 Proof of Theorem 5.3

The proof of Theorem 5.3(b) uses the following lemma.

Lemma 21.1 *Suppose Assumptions* C.4 *and* C.5 *hold. Under* $\{F_n\}_{n\geq 1}$, *we have*

$$A^{\inf}_n = A^{\inf}_n(\Lambda_{n,F_n}) + o_p(1).$$

Proof of Lemma 21.1. For a given distribution F, define

$$\nu^{\sigma\dagger}_n(\theta) := n^{1/2}\left(\left(\frac{\widehat{\sigma}^2_{n1}(\theta)}{\sigma^2_{F1}(\theta)} - 1\right), ..., \left(\frac{\widehat{\sigma}^2_{nk}(\theta)}{\sigma^2_{Fk}(\theta)} - 1\right)\right)'. \quad (21.1)$$

Note that $\nu^{\sigma\dagger}_n(\theta)$ differs from $\nu^{\sigma}_n(\theta)$ (defined in (2.12)) because the former depends on $\widehat{\sigma}^2_{nj}(\theta)$, which is centered at the sample quantity $\overline{m}_{nj}(\theta)$, see (2.10), whereas the latter depends on $\widehat{\sigma}^2_{Fnj}(\theta)$, which is centered at the population quantity $E_F m_j(W_i,\theta)$. The following calculations show that

$\nu_{nj}^{\sigma\dagger}(\theta) = \nu_{nj}^{\sigma}(\theta) - n^{-1/2}(\nu_{nj}^{m}(\theta))^2$:

$$\begin{aligned}
\nu_{nj}^{\sigma\dagger}(\theta) &:= n^{1/2}\left(\frac{\widehat{\sigma}_{nj}^2(\theta)}{\sigma_{F_nj}^2(\theta)} - 1\right) = n^{-1/2}\sum_{i=1}^{n}\left[(\widetilde{m}_j(W_i,\theta) - \widetilde{m}_{nj}(\theta))^2 - 1\right] \\
&= n^{-1/2}\sum_{i=1}^{n}\left[(\widetilde{m}_j(W_i,\theta) - E_{F_n}\widetilde{m}_j(W,\theta))^2 - 1\right] - n^{1/2}(\widetilde{m}_{nj}(\theta) - E_{F_n}\widetilde{m}_j(W,\theta))^2 \\
&= \nu_{nj}^{\sigma}(\theta) - n^{-1/2}(\nu_{nj}^{m}(\theta))^2, \text{ and} \\
\nu_{nj}^{\sigma\dagger}(\theta) &= \nu_{nj}^{\sigma}(\theta) + o_p^{\Theta}(1) \qquad (21.2)
\end{aligned}$$

$\forall j \leq k$, where the last equality holds by Assumption C.5.

By (21.2), Assumption C.5, and the continuous mapping theorem, for all $j \leq k$,

$$\begin{aligned}
&\sup_{\theta\in\Theta}\left|\frac{\widehat{\sigma}_{nj}^2(\theta)}{\sigma_{F_nj}^2(\theta)} - 1\right| =: \sup_{\theta\in\Theta} n^{-1/2}\left|\nu_{nj}^{\sigma\dagger}(\theta)\right| = \sup_{\theta\in\Theta} n^{-1/2}\left|\nu_{nj}^{\sigma}(\theta)\right| + o_p^{\Theta}(n^{-1/2}) \rightarrow_p 0, \text{ and so,} \\
&\sup_{\theta\in\Theta}\left|\frac{\sigma_{F_nj}(\theta)}{\widehat{\sigma}_{nj}(\theta)} - 1\right| \rightarrow_p 0. \qquad (21.3)
\end{aligned}$$

We have

$$\begin{aligned}
n^{1/2}\left(\frac{\widehat{\sigma}_{nj}(\theta)}{\sigma_{F_nj}(\theta)} - 1\right) &= n^{1/2}\left(\left(1 + \left(\frac{\widehat{\sigma}_{nj}^2(\theta)}{\sigma_{F_nj}^2(\theta)} - 1\right)\right)^{1/2} - 1\right) \\
&= \frac{1}{2}(1 + o_p^{\Theta}(1))^{-1/2} n^{1/2}\left(\frac{\widehat{\sigma}_{nj}^2(\theta)}{\sigma_{F_nj}^2(\theta)} - 1\right) \\
&= \frac{1}{2}\nu_{nj}^{\sigma}(\theta) + o_p^{\Theta}(1), \qquad (21.4)
\end{aligned}$$

where the second equality holds by the following mean-value expansion, $(1+x)^{1/2} = 1 + (1/2)(1 + \widetilde{x})^{-1/2}x$, where $|\widetilde{x}| \leq |x|$, with $x := \widehat{\sigma}_{nj}^2(\theta)/\sigma_{F_nj}^2(\theta) - 1$ and $\sup_{\theta\in\Theta}|x| \leq \sup_{\theta\in\Theta}|\widehat{\sigma}_{nj}^2(\theta)/\sigma_{F_nj}^2(\theta) - 1| = o_p(1)$ by (21.3), and the last equality uses (21.2) and Assumption C.5.

For all $j \leq k$, we have

$$\begin{aligned}
n^{1/2}\left(\widehat{m}_{nj}(\theta) - E_{F_n}\widetilde{m}_j(W,\theta)\right) &= \frac{\sigma_{F_nj}(\theta)}{\widehat{\sigma}_{nj}(\theta)}\left(\nu_{nj}^{m}(\theta) - E_{F_n}\widetilde{m}_j(W,\theta)n^{1/2}\left(\frac{\widehat{\sigma}_{nj}(\theta)}{\sigma_{F_nj}(\theta)} - 1\right)\right) \\
&= (1 + o_p^{\Theta}(1))\left(\nu_{nj}^{m}(\theta) - \frac{1}{2}E_{F_n}\widetilde{m}_j(W,\theta)\nu_{nj}^{\sigma}(\theta) + o_p^{\Theta}(1)\right) \\
&= \nu_{nj}^{m\sigma}(\theta) + o_p^{\Theta}(1), \qquad (21.5)
\end{aligned}$$

where $\nu_{nj}^{m}(\theta) := n^{1/2}(\widetilde{m}_{nj}(\theta) - E_{F_n}\widetilde{m}_j(W,\theta))$, $\widetilde{m}_{nj}(\theta) = (\widehat{\sigma}_{nj}(\theta)/\sigma_{F_nj}(\theta))\widehat{m}_{nj}(\theta)$ is defined in (2.3), the second equality holds by (21.4), and the third equality holds by the definition of $\nu_{nj}^{m\sigma}(\theta)$ in (5.7) and Assumptions C.4 and C.5.

Next, we have

$$\sup_{\ell_j \in R} \left| \left[\nu_{nj}^{m\sigma}(\theta) + o_p^{\Theta}(1) + \ell_j \right]_- - \left[\nu_{nj}^{m\sigma}(\theta) + \ell_j \right]_- \right| = o_p^{\Theta}(1) \tag{21.6}$$

because the function $\chi(v,c) := [v+c]_- - [c]_-$ for $v, c \in R_{[\pm\infty]}$ satisfies

$$|\chi(v,c)| \leq |v|. \tag{21.7}$$

This holds because (i) if $c \leq 0$ and $\nu + c \leq 0$, then $\chi(\nu, c) = |\nu|$, (ii) if $c \leq 0$ and $\nu + c > 0$, then $\nu > -c$ and $\chi(\nu, c) = |c| \leq |\nu|$, and (iii) if $c > 0$, then $\chi(\nu, c) = [\nu + c]_- \leq [\nu]_- \leq |\nu|$.

We have

$$\begin{aligned} & n^{1/2}\left(\widehat{r}_{nj}(\theta) - r_{F_n}^{\inf}\right) \\ := {} & n^{1/2}\left([\widehat{m}_{nj}(\theta)]_- - r_{F_n}^{\inf}\right) \\ = {} & \left(\left[\nu_{nj}^{m\sigma}(\theta) + n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta)\right]_- - \left[n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta)\right]_- + s_{nj}(\theta, F_n)\right) + o_p^{\Theta}(1), \end{aligned} \tag{21.8}$$

where $s_{nj}(\theta, F) := n^{1/2}([E_F\widetilde{m}_j(W,\theta)]_- - r_F^{\inf})$, using (21.5) and (21.6).

For given $(\theta, b, \ell) \in \Lambda_{n,F_n}$, where Λ_{n,F_n} is defined in (5.3), we have

$$n^{1/2}E_{F_n}\widetilde{m}(W,\theta) = \ell_j \text{ and } s_{nj}(\theta, F_n) = b_j. \tag{21.9}$$

Using (21.8) and (21.9), we obtain

$$\begin{aligned} A_n^{\inf} := {} & \inf_{\theta \in \Theta} \max_{j \leq k} n^{1/2}\left(\widehat{r}_{nj}(\theta) - r_{F_n}^{\inf}\right) \\ = {} & \inf_{(\theta,b,\ell) \in \Lambda_{n,F_n}} \max_{j \leq k} \left([\nu_{nj}^{m\sigma}(\theta) + \ell_j]_- - [\ell_j]_- + b_j\right) + o_p(1) \\ =: {} & A_n^{\inf}(\Lambda_{n,F_n}) + o_p(1), \end{aligned} \tag{21.10}$$

where the first equality holds by the definitions in (4.1) and (5.2) and the last equality holds by the definition in (5.10). □

Proof of Theorem 5.3. First, we prove part (a). For $j \leq k$, we show that

$$n^{1/2}(\widehat{m}_{nj}(\theta_n) + r_{F_n}^{\inf}) \rightarrow_d T_{j\infty} \tag{21.11}$$

and the convergence holds jointly over $j \leq k$. Stacking these results for $j = 1, ..., k$ gives $T_n(\theta_n) \to_d T_\infty$ using the definitions of $T_n(\theta_n)$ and T_∞ in (5.2) and (5.8), respectively.

We have

$$\begin{aligned}
n^{1/2}\left(\widehat{m}_{nj}(\theta) + r_F^{\inf}\right) &= n^{1/2}\left(\frac{\overline{m}_{nj}(\theta)}{\widehat{\sigma}_{nj}(\theta)} + r_F^{\inf}\right) \\
&= \frac{\sigma_{Fj}(\theta)}{\widehat{\sigma}_{nj}(\theta)}\widehat{K}_{1nj}(\theta, F) + \frac{\sigma_{Fj}(\theta)}{\widehat{\sigma}_{nj}(\theta)}\widehat{K}_{2nj}(\theta, F) + K_{3nj}(\theta, F), \text{ where} \\
\widehat{K}_{1nj}(\theta, F) &:= n^{1/2}\left(\frac{\overline{m}_{nj}(\theta)}{\sigma_{Fj}(\theta)} - \frac{E_F m_j(W, \theta)}{\sigma_{Fj}(\theta)}\right), \\
\widehat{K}_{2nj}(\theta, F) &:= -n^{1/2}\left(\frac{\widehat{\sigma}_{nj}(\theta)}{\sigma_{Fj}(\theta)} - 1\right)\frac{E_F m_j(W, \theta)}{\sigma_{Fj}(\theta)}, \text{ and} \\
K_{3nj}(\theta, F) &:= n^{1/2}\left(\frac{E_F m_j(W, \theta)}{\sigma_{Fj}(\theta)} + r_F^{\inf}\right). \qquad (21.12)
\end{aligned}$$

By Assumption C.3,

$$K_{3nj}(\theta_n, F_n) \to h_{j\infty}. \qquad (21.13)$$

By (21.4) and Assumption C.5,

$$\frac{\sigma_{F_n j}(\theta_n)}{\widehat{\sigma}_{nj}(\theta_n)} \to_p 1. \qquad (21.14)$$

Given (21.14), to prove part (a), it remains to determine the asymptotic distributions of $\widehat{K}_{1nj}(\theta_n, F_n)$ and $\widehat{K}_{2nj}(\theta_n, F_n)$.

We have

$$n^{1/2}\left(\frac{\widehat{\sigma}^2_{nj}(\theta_n)}{\sigma^2_{F_n j}(\theta_n)} - 1\right) =: \nu^{\sigma\dagger}_{nj}(\theta_n) = \nu^{\sigma}_{nj}(\theta_n) + o_p^{\Theta}(1) \to_d G^{\sigma}_{j\infty}, \qquad (21.15)$$

where the two equalities hold by (21.2) and the convergence holds by Assumption C.5 (which implies stochastic equicontinuity of $\{\nu_n^{\sigma}(\cdot)\}_{n\geq 1}$) and Assumption C.1. Equation (21.15) and the δ-method applied with the function $g(x) = x^{1/2}$, for which $g'(x)|_{x=1} = 1/2$, give

$$n^{1/2}\left(\frac{\widehat{\sigma}_{nj}(\theta_n)}{\sigma_{F_n j}(\theta_n)} - 1\right) \to_d \frac{1}{2}G^{\sigma}_{j\infty}. \qquad (21.16)$$

By Assumptions C.1 and C.4, $E_{F_n}\widetilde{m}_j(W, \theta_n) = \widetilde{m}_j(\theta_n) + o(1) \to \widetilde{m}_j(\theta_\infty) := \widetilde{m}_{j\infty}$. This and (21.16) give

$$\widehat{K}_{2nj}(\theta_n, F_n) \to_d -\frac{\widetilde{m}_{j\infty}}{2}G^{\sigma}_{j\infty}. \qquad (21.17)$$

We have

$$\widehat{K}_{1nj}(\theta_n, F_n) := n^{1/2}\left(\widetilde{m}_{nj}(\theta_n) - E_{F_n}\widetilde{m}_{nj}(\theta_n)\right) = \nu^{m}_{nj}(\theta_n) \to_d G^{m}_{j\infty}, \qquad (21.18)$$

where $\nu_{nj}^m(\theta_n)$ denotes the jth element of $\nu_n^m(\theta_n)$ and the convergence holds by Assumption C.5.

Combining the results in (21.12)–(21.14), (21.17), (21.18) and, for the case where $h_{j\infty} = \pm\infty$, the fact that $G_{j\infty}^m - \widetilde{m}_{j\infty} G_{j\infty}^\sigma/2 = O_p(1)$ (by Assumptions C.4 and C.5), establishes (21.11). The results in (21.11) for $j \leq k$ hold jointly because they are all based on the convergence result in Assumption C.5. This completes the proof of part (a).

Next, we prove part (b). By Lemma 21.1, it suffices to show

$$A_n^{\inf}(\Lambda_{n,F_n}) \to_d A_\infty^{\inf}(\Lambda). \tag{21.19}$$

Let $\mathcal{D}$ be the space of functions from Θ to R^{2k}. Let $\mathcal{D}_0$ be the subset of uniformly continuous functions in $\mathcal{D}$. For a nonstochastic function $\nu(\cdot) \in \mathcal{D}$, let $\nu(\theta) = (\nu^m(\theta)', \nu^\sigma(\theta)')'$, and let $\nu_j^m(\theta)$ and $\nu_j^\sigma(\theta)$ denote the jth elements of $\nu^m(\theta)$ and $\nu^\sigma(\theta)$, respectively. Define

$$\begin{aligned}
g_n(\nu(\cdot)) &:= \inf_{(\theta,b,\ell)\in\Lambda_{n,F_n}} \max_{j\leq k} \left[\tau_j(\nu(\cdot),\theta,\ell) + b_j\right], \\
g(\nu(\cdot)) &:= \inf_{(\theta,b,\ell)\in\Lambda} \max_{j\leq k} \left[\tau_j(\nu(\cdot),\theta,\ell) + b_j\right], \text{ where} \\
\tau_j(\nu(\cdot),\theta,\ell) &:= [\nu_j^{m\sigma}(\theta) + \ell_j]_- - [\ell_j]_- \text{ and} \\
\nu_j^{m\sigma}(\theta) &:= \nu_j^m(\theta) - \frac{1}{2}\widetilde{m}_j(\theta)\nu_j^\sigma(\theta).
\end{aligned} \tag{21.20}$$

For the stochastic processes $\nu_n(\cdot)$ and $G(\cdot)$, we can write

$$A_n^{\inf}(\Lambda_{n,F_n}) = g_n(\nu_n(\cdot)) \text{ and } A_\infty^{\inf}(\Lambda) = g(G(\cdot)). \tag{21.21}$$

We want to show that $g_n(\nu_n(\cdot)) \to_d g(G(\cdot))$. By Assumption C.5, $\nu_n(\cdot) \Rightarrow G(\cdot)$ for $\nu_n(\cdot) \in \mathcal{D}$ a.s. and $G(\cdot) \in \mathcal{D}_0$ a.s. We use the extended CMT, see van der Vaart and Wellner (1996, Theorem 1.11.1), to establish the desired result, as in the proof of Theorem 3.1 in BCS. The extended CMT requires showing: for any deterministic sequence $\{\nu_n(\cdot) \in \mathcal{D}\}_{n\geq1}$ and deterministic $\nu(\cdot) \in \mathcal{D}_0$ such that $\sup_{\theta\in\Theta} ||\nu_n(\theta) - \nu(\theta)|| \to 0$, we have $g_n(\nu_n(\cdot)) \to g(\nu(\cdot))$. (For notational simplicity, we abuse notation here and consider a deterministic $\nu_n(\cdot)$ that differs from the random $\nu_n(\cdot)$ in Assumption C.5.) Once we have shown this, the proof of part (b) is complete.

Let $\{\nu_n(\cdot) \in \mathcal{D}\}_{n\geq1}$ and $\nu(\cdot) \in \mathcal{D}_0$ be deterministic and satisfy $\sup_{\theta\in\Theta} ||\nu_n(\theta) - \nu(\theta)|| \to 0$. We show

$$\text{(i) } \liminf_{n\to\infty} g_n(\nu_n(\cdot)) \geq g(\nu(\cdot)) \text{ and (ii) } \limsup_{n\to\infty} g_n(\nu_n(\cdot)) \leq g(\nu(\cdot)). \tag{21.22}$$

First, we establish (i) in (21.22). There exists a subsequence $\{a_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$ and there exists a sequence $\{(\overline{\theta}_{a_n}, \overline{b}_{a_n}, \overline{\ell}_{a_n}) \in \Lambda_{a_n,F_{a_n}}\}_{n\geq 1}$ such that

$$\begin{aligned}
\liminf_{n\to\infty} g_n(\nu_n(\cdot)) &= \lim_{n\to\infty} g_{a_n}(\nu_{a_n}(\cdot)) \text{ and} \\
\lim_{n\to\infty} g_{a_n}(\nu_{a_n}(\cdot)) &= \lim_{n\to\infty} \max_{j\leq k} \left[\tau_j(\nu_{a_n}(\cdot), \overline{\theta}_{a_n}, \overline{\ell}_{a_n}) + \overline{b}_{a_n j}\right],
\end{aligned} \tag{21.23}$$

where $\overline{b}_{a_n j}$ denotes the jth element of $\overline{b}_{a_n}$. Also, there exists a subsequence $\{e_n\}_{n\geq 1}$ of $\{a_n\}_{n\geq 1}$ and $(\overline{\theta}, \overline{b}, \overline{\ell}) \in \Theta \times R^{2k}_{[\pm\infty]}$ such that

$$d\left((\overline{\theta}_{e_n}, \overline{b}_{e_n}, \overline{\ell}_{e_n}), (\overline{\theta}, \overline{b}, \overline{\ell})\right) \to 0, \tag{21.24}$$

where d is defined following (5.2), by compactness of the metric space $(\Theta \times R^{2k}_{[\pm\infty]}, d)$ under Assumption A.0(i). We have $(\overline{\theta}, \overline{b}, \overline{\ell}) \in \Lambda$ by the same argument as used to show $(\widetilde{\theta}, \widetilde{b}, \widetilde{\ell}) \in \Lambda$ in (20.7) (but without the requirement that $\overline{\theta}_{a_n} \in \Theta_I(F_{a_n})\ \forall n \geq 1$) using (21.24) and Assumption C.7.

For all $j \leq k$,

$$\begin{aligned}
\lim_{n\to\infty} \tau_j(\nu_{e_n}(\cdot), \overline{\theta}_{e_n}, \overline{\ell}_{e_n}) &= \tau_{j\infty}(\nu(\cdot), \overline{\theta}, \overline{\ell}) \in R, \text{ where} \\
\tau_{j\infty}(\nu(\cdot), \overline{\theta}, \overline{\ell}) &:= \begin{cases} [\nu_j^{m\sigma}(\overline{\theta}) + \overline{\ell}_j]_- - [\overline{\ell}_j]_- & \text{if } |\overline{\ell}_j| < \infty \\ -\nu_j^{m\sigma}(\overline{\theta}) & \text{if } \overline{\ell}_j = -\infty \\ 0 & \text{if } \overline{\ell}_j = +\infty \end{cases} \\
&= [\nu_j^{m\sigma}(\overline{\theta}) + \overline{\ell}_j]_- - [\overline{\ell}_j]_- \\
&:= \tau_j(\nu(\cdot), \overline{\theta}, \overline{\ell}),
\end{aligned} \tag{21.25}$$

the equality on the first line holds by $\nu_{e_n}(\theta) \to \nu(\theta) = (\nu^m(\theta)', \nu^\sigma(\theta)')'$ uniformly over $\theta \in \Theta$ (by assumption), (21.24), $[\nu_n + c_n]_- - [c_n]_- \to -\nu$ as $(\nu_n, c_n) \to (\nu, -\infty)$ for $\nu \in R$, and $[\nu_n + c_n]_- - [c_n]_- \to 0$ as $(\nu_n, c_n) \to (\nu, +\infty)$ for $\nu \in R$, the equality on the third line holds using the notational convention in (5.6), the equality on the last line holds by the definition of $\tau_j(\nu(\cdot), \theta, \ell)$ in (21.20), and "$\in R$" in the first line holds using the rhs expression on the second line because $\nu_j^{m\sigma}(\overline{\theta})$ is finite since $\nu(\cdot)$ is assumed to be in $\mathcal{D}$, $\chi(\nu, c) := [\nu + c]_- - [c]_-$ for $\nu, c \in R$ satisfies $|\chi(\nu, c)| \leq |\nu|$ as shown in (21.7), and $\widetilde{m}_j(\overline{\theta})$ is finite by Assumption C.4.

Now, we have

$$\begin{aligned}
\liminf_{n\to\infty} g_n(\nu_n(\cdot)) &= \lim_{n\to\infty} \max_{j\le k} \left[\tau_j(\nu_{e_n}(\cdot), \overline{\theta}_{e_n}, \overline{\ell}_{e_n}) + \overline{b}_{e_n j}\right] \\
&= \max_{j\le k} \left[\tau_j(\nu(\cdot), \overline{\theta}, \overline{\ell}) + \overline{b}_j\right] \\
&\ge \inf_{(\theta,b,\ell)\in\Lambda} \max_{j\le k} \left[\tau_j(\nu(\cdot), \theta, \ell) + b_j\right] \\
&:= g(\nu(\cdot)), \qquad (21.26)
\end{aligned}$$

where the first equality holds by (21.23) and the fact that $\{e_n\}_{n\ge1}$ is a subsequence of $\{a_n\}_{n\ge1}$, the second equality holds by (21.25) (using the notational convention in (5.6) if $\overline{b}_j = \pm\infty$ for any $j \le k$), the inequality holds because $(\overline{\theta}, \overline{b}, \overline{\ell}) \in \Lambda$ by the paragraph containing (21.24), and the last equality holds by the definition of $g(\nu(\cdot))$ in (21.20). This establishes result (i) in (21.22).

Next, we establish result (ii) in (21.22). There exists $(\theta^\dagger, b^\dagger, \ell^\dagger) \in \Lambda$ such that

$$g(\nu(\cdot)) = \max_{j\le k} \left[\tau_j(\nu(\cdot), \theta^\dagger, \ell^\dagger) + b_j^\dagger\right] \qquad (21.27)$$

because Λ is compact under the metric d, defined following (5.2) (since it is assumed to be an element of $\mathcal{S}(\Theta \times R^{2k}_{[\pm\infty]})$) and $\tau_j(\nu(\cdot), \theta, \ell) + b_j$ is a continuous function of (θ, b, ℓ) under d that takes values in the extended real line. By Assumption C.7, $\Lambda_{n,F_n} \to_H \Lambda$. Hence, there is a sequence $\{(\theta_n^\dagger, b_n^\dagger, \ell_n^\dagger) \in \Lambda_{n,F_n}\}_{n\ge1}$ such that $d((\theta_n^\dagger, b_n^\dagger, \ell_n^\dagger), (\theta^\dagger, b^\dagger, \ell^\dagger)) \to 0$. We obtain

$$\begin{aligned}
\limsup_{n\to\infty} g_n(\nu_n(\cdot)) &:= \limsup_{n\to\infty} \inf_{(\theta,b,\ell)\in\Lambda_{n,F_n}} \max_{j\le k} \left[\tau_j(\nu_n(\cdot), \theta, \ell) + b_j\right] \\
&\le \limsup_{n\to\infty} \max_{j\le k} \left[\tau_j(\nu_n(\cdot), \theta_n^\dagger, \ell_n^\dagger) + b_{nj}^\dagger\right] \\
&= \max_{j\le k} \left[\tau_j(\nu(\cdot), \theta^\dagger, \ell^\dagger) + b_j^\dagger\right] \\
&= g(\nu(\cdot)), \qquad (21.28)
\end{aligned}$$

where the inequality holds because $(\theta_n^\dagger, b_n^\dagger, \ell_n^\dagger) \in \Lambda_{n,F_n}$ $\forall n \ge 1$, the second equality holds using $d((\theta_n^\dagger, b_n^\dagger, \ell_n^\dagger), (\theta^\dagger, b^\dagger, \ell^\dagger)) \to 0$ and (21.25) with $(\nu_n(\cdot), \theta_n^\dagger, \ell_n^\dagger)$ and $(\nu(\cdot), \theta^\dagger, \ell^\dagger)$ in place of $(\nu_{e_n}(\cdot), \overline{\theta}_{e_n}, \overline{\ell}_{e_n})$ and $(\nu(\cdot), \overline{\theta}, \overline{\ell})$, respectively, and the last equality holds by (21.27). This establishes result (ii) in (21.22) and completes the proof of part (b).

Now we prove part (c). We have

$$A_\infty^{\inf}(\Lambda) := \inf_{(\theta,b,\ell)\in\Lambda} \max_{j\le k} \left([G_j^{m\sigma}(\theta) + \ell_j]_- - [\ell_j]_- + b_j\right) > -\infty \text{ a.s.} \qquad (21.29)$$

because (I) $\max_{j\le k} b_j \ge 0\ \forall(\theta,b,\ell)\in\Lambda$ by Lemma 5.2(b) and (II) $\sup_{(\theta,b,\ell)\in\Lambda}|[G_j^{m\sigma}(\theta)+\ell_j]_- - [\ell_j]_-| \le \sup_{\theta\in\Theta}|G_j^{m\sigma}(\theta)| < \infty$ a.s. (because $\chi(\nu,c) := [\nu+c]_- - [c]_-$ satisfies $|\chi(\nu,c)|\le|\nu|$ as shown in (21.7), $|[\nu+c]_- - [c]_-| := 0$ if $\nu\in R$ and $c=+\infty$, $|[\nu+c]_- - [c]_-| := -\nu$ if $\nu\in R$ and $c=-\infty$ using (5.6), and $\sup_{\theta\in\Theta}|G_j^{m\sigma}(\theta)|<\infty$ a.s. since $G(\cdot)$ is bounded on Θ a.s. by Assumption C.5 and $\widetilde{m}_j(\cdot)$ is bounded on Θ by Assumption C.4).

To obtain the other half of part (c), i.e., $A_\infty^{\inf}(\Lambda)<\infty$ a.s., we use Lemma 5.2(d). We have

$$\begin{aligned} A_\infty^{\inf}(\Lambda) &:= \inf_{(\theta,b,\ell)\in\Lambda}\max_{j\le k}\left([G_j^{m\sigma}(\theta)+\ell_j]_- - [\ell_j]_- + b_j\right) \\ &\le \max_{j\le k}\left([G_j^{m\sigma}(\widetilde{\theta})+\widetilde{\ell}_j]_- - [\widetilde{\ell}_j]_- + \widetilde{b}_j\right) < \infty \text{ a.s.,} \end{aligned} \tag{21.30}$$

where $(\widetilde{\theta},\widetilde{b},\widetilde{\ell})\in\Lambda$ is as in Lemma 5.2(d), the first equality holds by the definition of $A_\infty^{\inf}(\Lambda)$ in (5.10), the first inequality holds because $(\widetilde{\theta},\widetilde{b},\widetilde{\ell})\in\Lambda$, and last inequality holds because (I) $\max_{j\le k}\widetilde{b}_j = 0$ by Lemma 5.2(d) and (II) $\sup_{(\theta,b,\ell)\in\Lambda}|[G_j^{m\sigma}(\theta)+\ell_j]_- - [\ell_j]_-|<\infty$ a.s. by (II) following (21.29). This completes the proof of part (c).

Now we prove part (d). Under Assumption NLA, for all $j\le k$, we have

$$T_{j\infty} := G_{j\infty}^{m\sigma} + h_{j\infty} > -\infty \text{ a.s.,} \tag{21.31}$$

where the first equality holds by (5.8) and the inequality holds because $|G_{j\infty}^{m\sigma}|<\infty$ a.s. by the definitions in (5.5) and (5.7) and Assumptions C.4 and C.5, and $h_{j\infty}>-\infty$ by Assumption NLA.

Part (e) follows from the convergence results for $T_n(\theta_n)$ and $A_n^{\inf}$ in parts (a) and (b), the convergence result for $\widehat{\Omega}_n(\theta_n)$ in Assumption C.6, the definition of $S_n := S_n(\theta_n)$ in (4.2) and (5.2), the continuity of $S(m,\Omega)$ at all $m\in R^k_{[+\infty]}$ and $\Omega\in\Psi$ by Assumption S.1(iii), and the fact that $T_{j\infty}>-\infty\ \forall j\le k$ and $A_\infty^{\inf}(\Lambda)\in R$ by parts (c) and (d).

Now, we establish part (f). If $\Lambda=\Lambda_I$, then part (f) holds immediately. So, we suppose that $\Lambda\backslash\Lambda_I$ is not empty. We show that for any $(\theta^*,b^*,\ell^*)\in\Lambda\backslash\Lambda_I$,

$$\max_{j\le k}\left[\tau_j(G(\cdot),\theta^*,\ell^*)+b_j^*\right] = \infty \text{ a.s.,} \tag{21.32}$$

where $\tau_j(\nu(\cdot),\theta,\ell)$ is defined in (21.20). Since $A_\infty^{\inf}(\Lambda)\in R$ a.s. by part (c), and $A_\infty^{\inf}(\Lambda) := \inf_{(\theta,b,\ell)\in\Lambda}\max_{j\le k}[\tau_j(G(\cdot),\theta,\ell)+b_j]$ by (5.10), (21.32) implies that $A_\infty^{\inf}(\Lambda)=A_\infty^{\inf}(\Lambda_I)$ a.s., which establishes the first result in part (f). The second result in part (f) follows from the first result provided the quantities θ_∞, T_∞, and Ω_∞ are well defined, which requires Assumptions C.1, C.3, and C.6.

For part (f), it remains to show (21.32). By Assumption C.8, Λ_I is compact. For any $(\theta^*, b^*, \ell^*) \in \Lambda \backslash \Lambda_I$, there is a neighborhood of (θ^*, b^*, ℓ^*) that lies in $\Lambda \backslash \Lambda_I$ and there exists a sequence $\{(\theta_n^*, b_n^*, \ell_n^*) \in \Lambda_{n,F_n}\}_{n \geq 1}$ such that $d((\theta_n^*, b_n^*, \ell_n^*), (\theta^*, b^*, \ell^*)) \to 0$ by Assumption C.7. In consequence, for n large, $(\theta_n^*, b_n^*, \ell_n^*) \notin \Lambda_{n,F_n}^{\eta_n}$. In turn, this implies that $\theta_n^* \notin \Theta_I^{\eta_n}(F_n)$ for n large using the definition of $\Lambda_{n,F_n}^{\eta_n}$ following (5.4).

Now, $\theta_n^* \notin \Theta_I^{\eta_n}(F_n)$ for all n large implies

$$
\begin{aligned}
&\max_{j \leq k} n^{1/2}[E_{F_n} \widetilde{m}_j(W, \theta_n^*) + r_{F_n}^{\inf}]_- > \eta_n \text{ for all } n \text{ large,} \\
&\max_{j \leq k} n^{1/2}(-E_{F_n} \widetilde{m}_j(W, \theta_n^*) - r_{F_n}^{\inf}) \to \infty, \text{ and} \\
&\max_{j \leq k} b_j^* = \lim \max_{j \leq k} b_{n,j}^* := \lim \max_{j \leq k} n^{1/2}([E_{F_n} \widetilde{m}_j(W, \theta_n^*)]_- - r_{F_n}^{\inf}) = \infty, \qquad (21.33)
\end{aligned}
$$

where the first line holds by the definition of $\Theta_I^{\eta}(F)$ in (5.4), the first line implies that $\min_{j \leq k} E_{F_n} \widetilde{m}_j(W, \theta_n^*) + r_{F_n}^{\inf} < 0$ for all n large, which is used to obtain the second line, the second line also uses $\eta_n \to \infty$ by Assumption C.8, the first equality in the third line holds by the convergence result for $\{(\theta_n^*, b_n^*, \ell_n^*)\}_{n \geq 1}$ in the previous paragraph, the second equality in the third line holds by $(\theta_n^*, b_n^*, \ell_n^*) \in \Lambda_{n,F_n}$ and the definition of $\Lambda_{n,F}$ in (5.3), and the third equality in the third line follows from the second line because $\min_{j \leq k} E_{F_n} \widetilde{m}_j(W, \theta_n^*) + r_{F_n}^{\inf} < 0$ for n large implies $\min_{j \leq k} E_{F_n} \widetilde{m}_j(W, \theta_n^*) < 0$ for n large, since $r_{F_n}^{\inf} \geq 0$ by (2.5).

The result $\max_{j \leq k} b_j^* = \infty$ in (21.33) implies that (21.32) holds because $|\tau_j(G(\cdot), \theta^*, \ell^*)| < \infty$ a.s. (using Assumptions C.4 and C.5, the definition of $\tau_j(\nu(\cdot), \theta, \ell)$ in (21.20), and explanation (II) following (21.29)). This completes the proof of part (f).

Part (g) holds because $T_{j\infty} := G_{j\infty}^{m\sigma} + h_{j\infty} = -\infty$ for some $j \leq k$ by (5.8), Assumption CA, and the notational convention in (5.6).

Next, we prove part (h). We have $T_{nj}(\theta_n) \to_p h_{j\infty} = -\infty$ for some $j \leq k$ by parts (a) and (g) and $A_n^{\inf} \to_d A_\infty^{\inf}(\Lambda) \in R$ by parts (b) and (c). Thus,

$$
\varphi_n := \min_{j \leq k}(T_{nj}(\theta_n) + A_n^{\inf}) \to_p -\infty. \qquad (21.34)
$$

Using this, we obtain

$$
\begin{aligned}
S_n := S_n(\theta_n) &= S\left(T_n(\theta_n) + A_n^{\inf} 1_k, \widehat{\Omega}_n(\theta_n)\right) = |\varphi_n|^\chi S\left([T_n(\theta_n) + A_n^{\inf} 1_k]/|\varphi_n|, \widehat{\Omega}_n(\theta_n)\right) \\
&\geq |\varphi_n|^\chi \min_{j \leq k} S\left(c_j, \widehat{\Omega}_n(\theta_n)\right) = |\varphi_n|^\chi \left(\min_{j \leq k} S\left(c_j, \Omega_\infty\right) + o_p(1)\right) \to_p \infty, \qquad (21.35)
\end{aligned}
$$

where c_j is a k-vector of ∞'s but with -1 as its jth element, the second equality holds by (5.2), the third equality holds with $\chi > 0$ by Assumption S.3, the inequality holds with probability that goes to one as $n \to \infty$ (wp$\to$1) because $(T_{nj}(\theta_n) + A_n^{\inf})/|\varphi_n| = -1$ for some $j \leq k$ wp$\to$1 by the definition of φ_n and $\varphi_n \to_p -\infty$, $S(m, \Omega)$ is nonincreasing in m for all $\Omega \in \Psi$ by Assumption S.1(i), and $[T_n(\theta_n) + A_n^{\inf} 1_k]/|\varphi_n| < \infty$ $\forall j \leq k$, the last equality holds by Assumptions C.6 and S.1(iii), and the convergence holds because $\min_{j \leq k} S(c_j, \Omega_\infty) > 0$ by Assumption S.2 and the fact that c_j has a negative element for all $j \leq k$, $|\varphi_n| \to_p \infty$ and $\chi > 0$.

Lastly, the results in parts (a)–(e) hold jointly because they are all based on the convergence result in Assumption C.5, which establishes part (i). $\square$

22 Asymptotic Rejection Probabilities of SPUR1 Tests

The first subsection of this section provides a theorem, Theorem 22.1, that is the key ingredient to the proofs of Theorems 7.1 and 12.1. It provides asymptotic NRP bounds, asymptotic $n^{-1/2}$-local power bounds, and consistency results for the nominal level α SPUR1 test $\phi_{n,SPUR1}(\theta_n)$, defined in (4.4), under drifting subsequences of distributions and parameter values. The second subsection states several lemmas that are used in the proof of Theorem 22.1. The third subsection provides the proof of Theorem 22.1 using these lemmas.

To establish the asymptotic properties of bootstrap critical values for a given sequence of distributions $\{F_n\}_{n \geq 1}$, it is convenient to have a single probability space $(\Omega, \mathcal{F}, P_\bigtriangledown)$ on which all of the random vectors $\{W_i\}_{i \leq n}$ for $n \geq 1$ and the bootstrap random variables (or vectors) $\{\zeta_i\}_{i \leq n}$ for all $n \geq 1$ are defined. Since F_n changes with n, this requires that we consider triangular arrays of random vectors, not sequences. Let $\{W_{ni}\}_{i \leq n, n \geq 1} := \{W_{ni} : i \leq n, n \geq 1\}$ be a triangular array of random vectors on $(\Omega, \mathcal{F}, P_\bigtriangledown)$ such that, for each $n \geq 1$, $\{W_{ni}\}_{i \leq n}$ has the same distribution as $\{W_i\}_{i \leq n} \sim F_n$. Analogously, let $\{\zeta_{ni}\}_{i \leq n, n \geq 1}$ be a triangular array of bootstrap random variables (or vectors) on $(\Omega, \mathcal{F}, P_\bigtriangledown)$ such that for each $n \geq 1$, $\{\zeta_{ni}\}_{i \leq n}$ has the same distribution as $\{\zeta_i\}_{i \leq n}$ and $\{\zeta_{ni}\}_{i \leq n, n \geq 1}$ is independent of $\{W_{ni}\}_{i \leq n, n \geq 1}$.

For notational simplicity, but with some abuse of notation, we let all of the statistics being considered, including S_n, $S_n^*(\theta_n)$, and $\widehat{c}_n(\theta_n, 1-\alpha)$, which are defined as functions of $\{W_i\}_{i \leq n} \sim F_n$ and $\{\zeta_i\}_{i \leq n}$, also denote the corresponding statistics defined when using the triangular arrays $\{W_{ni}\}_{i \leq n, n \geq 1}$ and $\{\zeta_{ni}\}_{i \leq n, n \geq 1}$. For events that only depend on n random vectors for a single n, such as $S_n^*(\theta_n) \in B_n$ for some fixed set $B_n \subset R$, we have $P_\bigtriangledown(S_n^*(\theta_n) \in B_n) = P_{F_n}(S_n^*(\theta_n) \in B_n)$. But, for events that depend on statistics for multiple values of n, such as $\{S_n^*(\theta_n)\}_{n \geq 1}$, we use the probability space $(\Omega, \mathcal{F}, P_\bigtriangledown)$. In particular, when we condition on the entire triangular array

$\{W_{ni}\}_{i\leq n,n\geq 1}$, we need to use $(\Omega, \mathcal{F}, P_{\bigtriangledown})$.

22.1 Statement of Theorem 22.1

Let $\{\nu_n^*(\theta) \in R^{2k} : \theta \in \Theta\}$ be a bootstrap version of the empirical process $(\nu_n^m(\cdot)', \nu_n^{\sigma\dagger}(\theta)')'$ defined in (2.12) and (21.1). It is defined as follows:

$$\begin{aligned}
&\nu_n^*(\theta) := (\nu_n^{m*}(\theta)', \nu_n^{\sigma*}(\theta)')', \text{ where} \\
&\nu_{nj}^{m*}(\theta) := n^{1/2}\left(\widetilde{m}_{nj}^*(\theta) - \widehat{m}_{nj}(\theta)\right), \ \widetilde{m}_{nj}^*(\theta) := \frac{\overline{m}_{nj}^*(\theta)}{\widehat{\sigma}_{nj}(\theta)}, \ \overline{m}_{nj}^*(\theta) := n^{-1}\sum_{i=1}^n m_j(W_i^*, \theta), \\
&\nu_{nj}^{\sigma*}(\theta) := n^{1/2}\left(\frac{\widehat{\sigma}_{nj}^{*2}(\theta)}{\widehat{\sigma}_{nj}^2(\theta)} - 1\right), \ \widehat{\sigma}_{nj}^{*2}(\theta) := n^{-1}\sum_{i=1}^n (m_j(W_i^*, \theta) - \overline{m}_{nj}^*(\theta))^2 \ \forall j \leq k, \\
&\nu_n^{m*}(\theta) = (\nu_{n1}^{m*}(\theta), ..., \nu_{nk}^{m*}(\theta))', \text{ and } \nu_n^{\sigma*}(\theta) = (\nu_{n1}^{\sigma*}(\theta), ..., \nu_{nk}^{\sigma*}(\theta))'.
\end{aligned} \tag{22.1}$$

We employ the following bootstrap convergence (BC) assumption.

Assumption BC.6. $\{\nu_n^*(\cdot)|\{W_{ni}\}_{i\leq n,n\geq 1}\} \Rightarrow G(\cdot)$ a.s.$[P_{\bigtriangledown}]$, where $G(\cdot)$ is as in Assumption C.5.

Assumption BC.6 is verified below for i.i.d. observations using Lemma D.2(8) of BCS under Assumptions A.1–A.4. To allow the general results to apply to non-i.i.d. observations, including time series observations, we employ Assumption BC.6 here, rather than impose Assumptions A.1–A.4.

The following theorem uses S_∞, which is defined in (5.11). The distribution of S_∞ is the asymptotic distribution of the SPUR test statistic, see Theorem 5.3. The theorem also uses $c_{L\infty,EGMS}(1-\alpha)$ and $c_{U\infty,EGMS}(1-\alpha)$, which are defined just below (12.7) and are the $1-\alpha$ quantiles of the asymptotic distributions of the lower and upper bounds on the EGMS bootstrap statistic $S_{n,EGMS}^*(\theta)$ defined in (12.4)–(12.7).

Theorem 22.1 *For* $\alpha \in (0,1)$ *and for sequences* $\{F_n\}_{n\geq 1}$ *and* $\{\theta_n\}_{n\geq 1}$ *that satisfy Assumptions* A.0, A.6, BC.1, BC.2, BC.6, C.1–C.8, *and* S.1 *for a subsequence* $\{p_n\}_{n\geq 1}$ *in place of* $\{n\}_{n\geq 1}$, *there exists a subsequence* $\{a_n\}_{n\geq 1}$ *of* $\{p_n\}_{n\geq 1}$ *for which the nominal level* α SPUR1 *test* $\phi_{n,SPUR1}(\theta_n)$ *for testing* $H_0 : \theta_n \in \Theta_I(F_n)$ *satisfies*

(a) $\limsup_{n\to\infty} P_{F_{a_n}}(\phi_{a_n,SPUR1}(\theta_{a_n}) = 1) \leq P(S_\infty > c_{L\infty,EGMS}(1-\alpha))$ *provided Assumptions* A.5, BC.3, *and* NLA *hold for the subsequence* $\{p_n\}_{n\geq 1}$ *in place of* $\{n\}_{n\geq 1}$,

(b) $\liminf_{n\to\infty} P_{F_{a_n}}(\phi_{a_n,SPUR1}(\theta_{a_n}) = 1) \geq P(S_\infty > c_{U\infty,EGMS}(1-\alpha))$ *provided Assumptions* A.10, BC.4, BC.5, *and* NLA *hold for the subsequence* $\{p_n\}_{n\geq 1}$ *in place of* $\{n\}_{n\geq 1}$,

(c) $\limsup_{n\to\infty} P_{F_{a_n}}(\phi_{a_n,SPUR1}(\theta_{a_n}) = 1) \leq \alpha$ *provided Assumptions* A.5, A.7, BC.3, *and* N *hold for the subsequence* $\{p_n\}_{n\geq 1}$ *in place of* $\{n\}_{n\geq 1}$, *and*

(d) $\liminf_{n\to\infty} P_{F_{a_n}}(\phi_{a_n,SPUR1}(\theta_{a_n}) = 1) = 1$ *provided Assumptions* A.10, BC.4, CA, S.2, *and* S.3 *hold for the subsequence* $\{p_n\}_{n\geq 1}$ *in place of* $\{n\}_{n\geq 1}$.

Comments. (i). Theorem 22.1(a) and (b) provide upper and lower bounds, respectively, on the asymptotic power of the SPUR1 test under null and $n^{-1/2}$-local-alternative distributions for certain subsequences.

(ii). Theorem 22.1(c) shows that the nominal level α SPUR1 test has asymptotic NRP's equal to α or less for certain subsequences. Theorem 22.1(c) also holds without imposing Assumption BC.1 and with $sd_{1j\infty}(\theta) := 1$ in Assumption BC.2. The proof of this is given following the proof of Theorem 22.1.

(iii). Theorem 22.1(d) establishes that the SPUR1 test $\phi_{n,SPUR1}(\theta_n)$ is consistent for certain subsequences under Assumption CA, which includes all fixed alternatives, as well as (drifting) local alternatives that deviate from the null by more than $n^{-1/2}$-local alternatives.

(iv). When Theorem 22.1 is used below to prove Theorems 7.1 and 12.1, the subsequences that are employed are ones in which the $\limsup_{n\to\infty}$ and $\liminf_{n\to\infty}$ in Theorem 22.1 are actually limits as $n \to \infty$.

22.2 Lemmas Used in the Proof of Theorem 22.1

Lemma 22.2 below provides upper and lower bounds on the asymptotic rejection probabilities of a test based on the SPUR test statistic and a generic bootstrap critical value under drifting sequences of distributions and parameters values under high-level conditions, namely, Assumptions CV.1–CV.3. The method employed is somewhat similar to that of Theorem 4.1 of BCS. Next, in Lemmas 22.3–22.5 below, we verify these high-level conditions for the EGMS bootstrap critical value, which is defined in Section 4.1.

Let $S_n^*(\theta)$ denote a nonnegative generic bootstrap (or some other) statistic that is used to calculate a critical value, such as $S_n^*(\theta) := S_{n,EGMS}^*(\theta)$ in (6.3). The bootstrap statistic $S_n^*(\theta)$ depends on $\{W_i\}_{i\leq n}$ and on some other independent random variables $\{\zeta_i\}_{i\leq n}$ that are used to construct the bootstrap sample. Let $\widehat{c}_n(\theta, 1-\alpha)$ be the $1-\alpha$ conditional quantile of $S_n^*(\theta)$ given $\{W_i\}_{i\leq n}$ for $\alpha \in (0,1)$. Let $\phi_n(\theta_n)$ denote the nominal level α test that rejects $H_0 : \theta_n \in \Theta_I(F_n)$ if

$$S_n(\theta_n) > \widehat{c}_n(\theta_n, 1-\alpha). \tag{22.2}$$

Let $X \geq_{ST} Y$ denote that X is stochastically greater than or equal to Y. That is, $P(Y > x) \leq P(X > x)$ for all $x \in R$.

To establish the asymptotic rejection probability results, we assume the existence of sequences of bounding random variables $\{S^*_{Ln}(\theta_n)\}_{n\geq 1}$ for which $S^*_{Ln}(\theta_n) \leq S^*_n(\theta_n)$ for almost all realizations of the bootstrap random variables wp→1 with respect to the randomness in the sample and likewise for some upper-bound random variables $\{S^*_{Un}(\theta_n)\}_{n\geq 1}$.

Assumption CV.1. There exist nonnegative random variables $\{S^*_{Ln}(\theta_n)\}_{n\geq 1}$ such that (i) $P_\nabla(S^*_{Ln}(\theta_n) \leq S^*_n(\theta_n)|\{W_{ni}\}_{i\leq n, n\geq 1}) = 1$ wp→1 and (ii) $\{S^*_{Ln}(\theta_n)|\{W_{ni}\}_{i\leq n, n\geq 1}\} \to_d S^*_{L\infty}$ a.s.$[P_\bigtriangledown]$ for some $S^*_{L\infty} \in R$ a.s. that does not depend on the conditioning value of $\{W_{ni}\}_{i\leq n, n\geq 1}$.[27]

Assumption CV.2. $S^*_{L\infty}$ satisfies $S^*_{L\infty} \geq_{ST} S_\infty$.

Assumption CV.3. There exist nonnegative random variables $\{S^*_{Un}(\theta_n)\}_{n\geq 1}$ such that (i) $P_\nabla(S^*_{Un}(\theta_n) \geq S^*_n(\theta_n)|\{W_{ni}\}_{i\leq n, n\geq 1}) = 1$ wp→1 and (ii) $\{S^*_{Un}(\theta_n)|\{W_{ni}\}_{i\leq n, n\geq 1}\} \to_d S^*_{U\infty}$ a.s.$[P_\bigtriangledown]$ for some $S^*_{U\infty} \in R$ a.s. that does not depend on the conditioning value of $\{W_{ni}\}_{i\leq n, n\geq 1}$.

Assumptions CV.1 and CV.3 are used to obtain upper and lower bounds, respectively, on asymptotic rejection probabilities under null and $n^{-1/2}$-local alternative distributions. For example, when Assumption CV.1 is employed with $S^*_n(\theta_n) = S^*_{n,EGMS}(\theta_n)$, we define the statistic $S^*_{Ln}(\theta_n)$ to equal $S^*_{Ln,EGMS}(\theta_n)$ in (12.4), which is defined using $\inf_{\theta\in\Theta_I^{\eta_n}(F_n)}$ (where $\Theta_I^{\eta_n}(F_n)$ is nonrandom), whereas $S^*_{n,EGMS}(\theta_n)$ is defined using $\inf_{\theta\in\widehat{\Theta}_n}$ (where $\widehat{\Theta}_n$ is random) and several other simplifications. These changes lead to simpler asymptotic behavior of $S^*_{Ln}(\theta_n)$ than $S^*_n(\theta_n)$. The same is true when Assumption CV.3 is employed with $S^*_n(\theta_n) = S^*_{n,EGMS}(\theta_n)$ and $S^*_{Un}(\theta_n) = S^*_{Un,EGMS}(\theta_n)$ (defined in (12.6)).

Assumption CV.2 is only employed in conjunction with Assumption N, i.e., when S_∞ is an asymptotic null distribution of S_n. Under Assumption LA, the distribution of S_∞ is larger than under Assumption N and $S^*_{L\infty} \geq_{ST} S_\infty$ typically fails.

Let $c_{L\infty}(1-\alpha)$ and $c_{U\infty}(1-\alpha)$ denote the $1-\alpha$ quantiles of $S^*_{L\infty}$ and $S^*_{U\infty}$, respectively.

Lemma 22.2 *Suppose that under* $\{F_n\}_{n\geq 1}$ *and* $\{\theta_n\}_{n\geq 1}$, *Assumptions* A.0, C.1–C.7, *and* S.1(iii) *hold. For* $\alpha \in (0,1)$, *let* $\phi_n(\theta_n)$ *be the nominal level* α *test defined in* (22.2). *Then,*

(a) $\limsup_{n\to\infty} P_{F_n}(\phi_n(\theta_n) = 1) \leq P(S_\infty > c_{L\infty}(1-\alpha))$ *provided Assumptions* CV.1 *and* NLA *hold,*

(b) $\limsup_{n\to\infty} P_{F_n}(\phi_n(\theta_n) = 1) \geq P(S_\infty > c_{U\infty}(1-\alpha))$ *provided Assumptions* CV.3, BC.5, *and* NLA *hold,*

(c) $\limsup_{n\to\infty} P_{F_n}(\phi_n(\theta_n) = 1) \leq \alpha$ *provided Assumptions* A.7, CV.1, CV.2, N, *and* S.1(ii) *hold, and*

(d) $\limsup_{n\to\infty} P_{F_n}(\phi_n(\theta_n) = 1) = 1$ *provided Assumptions* CA, CV.3, S.2, *and* S.3 *hold.*

[27] In Assumption CV.1(ii), $\{S^*_{Ln}(\theta_n)|\{W_{ni}\}_{i\leq n, n\geq 1}\} \to_d S^*_{L\infty}$ a.s.$[P_\bigtriangledown]$ means $P_\bigtriangledown(S^*_{Ln}(\theta_n) \to_d S^*_{L\infty}|\{W_{ni}\}_{i\leq n, n\geq 1}) := P_\bigtriangledown(\{\omega : S^*_{Ln}(\theta_n) \to_d S^*_{L\infty}|\{W_{ni}(\omega)\}_{i\leq n, n\geq 1}\}) = 1$.

Comment. For any subsequence $\{a_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$, Lemma 22.2 holds with a_n in place of n throughout, including the assumptions. (The proof just needs to be changed by replacing n by a_n throughout.)

The next three lemmas verify Assumptions CV.1–CV.3 for the EGMS critical values employed by the SPUR1 tests $\phi_{n,SPUR1}(\theta_n)$. More precisely, given any subsequence $\{p_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$, the lemmas verify Assumptions CV.1–CV.3 when these assumptions are defined for some subsequence $\{a_n\}_{n\geq 1}$ of $\{p_n\}_{n\geq 1}$, rather than for $\{n\}_{n\geq 1}$.

The EGMS critical values are based on the bootstrap random variables $S^*_{n,EGMS}(\theta_n)$. In the following lemmas, the "lower bound" random variables $S^*_{Ln,EGMS}(\theta)$, $T^*_{Lnj,EGMS}(\theta)$, and $A^{*\inf}_{Ln,EGMS}$ are defined in (12.4); the asymptotic distributions of these random variables $S^*_{L\infty,EGMS}$, $T^*_{Lj\infty,EGMS}$, and $A^{*\inf}_{L\infty,EGMS}$ are defined in (12.5); the "upper bound" random variables $S^*_{Un,EGMS}(\theta)$, $T^*_{Unj,EGMS}(\theta)$, and $A^{*\inf}_{Un,EGMS}$ are defined in (12.6); and the asymptotic distributions of the latter random variables $S^*_{U\infty,EGMS}$, $T^*_{Uj\infty,EGMS}$, and $A^{*\inf}_{U\infty,EGMS}$ are defined in (12.7). As above, we assume that all of the statistics are functions of the triangular arrays $\{W_{ni}\}_{i\leq n,n\geq 1}$ and $\{\zeta_{ni}\}_{i\leq n,n\geq 1}$ that are defined on a single probability space $(\Omega,\mathcal{F},P_{\bigtriangledown})$.

The following lemma provides the asymptotic distributions of $S^*_{Ln,EGMS}(\theta_n)$ and $S^*_{Un,EGMS}(\theta_n)$.

Lemma 22.3 *For sequences* $\{F_n\}_{n\geq 1}$ *and* $\{\theta_n\}_{n\geq 1}$ *that satisfy Assumptions* A.0, A.5, A.6, BC.1–BC.3, BC.6, C.1, C.2, C.4–C.7, *and* S.1 *for a subsequence* $\{p_n\}_{n\geq 1}$ *in place of* $\{n\}_{n\geq 1}$, *there exists a subsequence* $\{a_n\}_{n\geq 1}$ *of* $\{p_n\}_{n\geq 1}$ *for which* (a) $\{T^*_{La_nj,EGMS}(\theta_{a_n})|\{W_{ni}\}_{i\leq n,n\geq 1}\}\rightarrow_d T^*_{Lj\infty,EGMS}$ *a.s.*$[P_{\bigtriangledown}]$ $\forall j\leq k$, (b) $\{A^{*\inf}_{La_n,EGMS}|\{W_{ni}\}_{i\leq n,n\geq 1}\}\rightarrow_d A^{*\inf}_{L\infty,EGMS}$ *a.s.*$[P_{\bigtriangledown}]$, (c) $\{S^*_{La_n,EGMS}(\theta_{a_n})|\{W_{ni}\}_{i\leq n,n\geq 1}\}\rightarrow_d S^*_{L\infty,EGMS}$ *a.s.*$[P_{\bigtriangledown}]$ *and* $S^*_{L\infty,EGMS}\in[0,\infty)$ *a.s., and* (d) *parts* (a)–(c) *hold with* U *in place of* L *throughout and Assumptions* A.10 *and* BC.4 *in place of Assumptions* A.5 *and* BC.3.

Comment. Lemma 22.3(c) and (d) verify the convergence results in Assumptions CV.1(ii) and CV.3(ii) for the subsequences $\{S^*_{La_n,EGMS}(\theta_{a_n})\}_{n\geq 1}$ and $\{S^*_{Ua_n,EGMS}(\theta_{a_n})\}_{n\geq 1}$, respectively.

The following lemma verifies Assumptions CV.1(i) and CV.3(i) for a subsequence $\{p_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$.

Lemma 22.4 *For sequences* $\{F_n\}_{n\geq 1}$ *and* $\{\theta_n\}_{n\geq 1}$ *that satisfy Assumptions* A.0, A.5, A.6, BC.1, BC.3, C.4, C.5, C.7, *and* S.1(i) *for a subsequence* $\{p_n\}_{n\geq 1}$ *in place of* $\{n\}_{n\geq 1}$, (a) $P_{F_{p_n}}(T^*_{Lp_nj,EGMS}(\theta_{p_n})\geq T^*_{p_nj,EGMS}(\theta_{p_n})|\{W_{ni}\}_{i\leq n,n\geq 1})=1\ \forall j\leq k$ *wp→1,* (b) $P_{F_{p_n}}(A^{*\inf}_{Lp_n,EGMS}\geq A^{*\inf}_{p_n,EGMS}|\{W_{ni}\}_{i\leq n,n\geq 1})=1$ *wp→1,* (c) $P_{F_{p_n}}(S^*_{Lp_n,EGMS}(\theta_n)\leq S^*_{p_n,EGMS}(\theta_n)|\{W_{ni}\}_{i\leq n,n\geq 1})=1$ *wp→1, and* (d) *parts* (a)–(c) *hold with* U *in place of* L *throughout, the inequalities reversed throughout, and Assumptions* A.10 *and* BC.4 *in place of Assumptions* A.5 *and* BC.3.

The following lemma verifies Assumption CV.2 with $S^*_{L\infty} = S^*_{L\infty,EGMS}$ for sequences $\{\theta_n\}_{n\geq1}$ of null parameter values (i.e., under Assumption N).

Lemma 22.5 *For sequences* $\{F_n\}_{n\geq1}$ *and* $\{\theta_n\}_{n\geq1}$ *that satisfy Assumptions* A.5, A.6, BC.1–BC.3, C.1, C.3–C.5, C.8, N, *and* S.1(i) *for a subsequence* $\{p_n\}_{n\geq1}$ *in place of* $\{n\}_{n\geq1}$, *we have* $S^*_{L\infty,EGMS} \geq S_{I\infty}$ *for all sample realizations.*

22.3 Proof of Theorem 22.1

Proof of Theorem 22.1. Given any subsequence $\{p_n\}_{n\geq1}$ of $\{n\}_{n\geq1}$, we take the subsequence $\{a_n\}_{n\geq1}$ of $\{p_n\}_{n\geq1}$ as in Lemma 22.3. We apply Lemma 22.2 with $S^*_n(\theta_n)$, $S^*_{Ln}(\theta_n)$, and $S^*_{Un}(\theta_n)$ in Lemma 22.2 and Assumptions CV.1 and CV.3 equal to $S^*_{n,EGMS}(\theta_n)$, $S^*_{Ln,EGMS}(\theta_n)$, and $S^*_{Un,EGMS}(\theta_n)$, respectively, and with the subsequence $\{a_n\}_{n\geq1}$ in place of $\{n\}_{n\geq1}$ (see the Comment following Lemma 22.2), which establishes all of the results of the theorem. All of the assumptions in parts (a)–(d) of Lemma 22.2, which need to hold with $\{a_n\}_{n\geq1}$ in place of $\{n\}_{n\geq1}$, are imposed in the corresponding parts (a)–(d) of the theorem based on $\{n\}_{n\geq1}$, except Assumptions CV.1–CV.3. The assumptions based on $\{n\}_{n\geq1}$ imply those based on $\{a_n\}_{n\geq1}$. Thus, it remains to verify Assumption CV.1 (defined using $\{a_n\}_{n\geq1}$ in place of $\{n\}_{n\geq1}$) in parts (a) and (c) of Theorem 22.1, Assumption CV.2 in part (c), and Assumption CV.3 (defined using $\{a_n\}_{n\geq1}$ in place of $\{n\}_{n\geq1}$) in parts (b) and (d).

As required by Assumptions CV.1 and CV.3, $S^*_{Ln,EGMS}(\theta_n) \geq 0$ and $S^*_{Un,EGMS}(\theta_n) \geq 0$ by Assumption S.1(ii).

The assumptions of parts (a) and (c) of the theorem include all of the assumptions imposed in Lemmas 22.3(c) and 22.4(c). Lemma 22.3(c) verifies the convergence result of Assumption CV.1(ii) for the subsequence $\{a_n\}_{n\geq1}$ with $S^*_{L\infty} = S^*_{L\infty,EGMS}$ and the requirement of Assumption CV.1(ii) that $S^*_{L\infty} = S^*_{L\infty,EGMS} \in [0,\infty)$ a.s. Lemma 22.4(c) verifies Assumption CV.1(i) for the subsequence $\{p_n\}_{n\geq1}$, and hence, also for its subsequence $\{a_n\}_{n\geq1}$. The requirement of Assumption CV.1(ii) that "$S^*_{L\infty} = S^*_{L\infty,EGMS}$ does not depend on the conditioning value of $\{W_{ni}\}_{i\leq n,n\geq1}$" holds by the definition of $S^*_{L\infty,EGMS}$ in (12.5). Hence, Assumption CV.1 (defined using $\{a_n\}_{n\geq1}$ in place of $\{n\}_{n\geq1}$) holds in parts (a) and (c) of the theorem.

Assumption CV.2 holds in part (c) of Theorem 22.1 with $S^*_{L\infty} = S^*_{L\infty,EGMS}$ by Lemma 22.5, because part (c) imposes all of the assumptions of Lemma 22.5.

The assumptions of parts (b) and (d) of the theorem include all of the assumptions imposed in Lemmas 22.3(d) and 22.4(d). Lemma 22.3(d) verifies the convergence result of Assumption CV.3(ii) for the subsequence $\{a_n\}_{n\geq1}$ with $S^*_{U\infty} = S^*_{U\infty,EGMS}$ and the requirement of Assumption CV.3(ii) that $S^*_{U\infty} = S^*_{U\infty,EGMS} \in [0,\infty)$ a.s. Lemma 22.4(d) verifies Assumption CV.3(i). The requirement of Assumption CV.3(ii) that "$S^*_{U\infty} = S^*_{U\infty,EGMS}$ does not depend on the conditioning value of $\{W_{ni}\}_{i\le n, n\ge 1}$" holds by the definition of $S^*_{U\infty,EGMS}$ in (12.7). Hence, Assumption CV.3 (defined using $\{a_n\}_{n\ge 1}$ in place of $\{n\}_{n\ge 1}$) holds in parts (b) and (d) of the theorem. This completes the proof. □

Next, we show that Theorem 22.1(c) also holds without imposing Assumption BC.1 and with $sd_{1j\infty}(\theta) := 1$ in Assumption BC.2, as stated in Comment (ii) to Theorem 22.1. Consider the bootstrap statistic $S^*_{n,EGMS}(\theta)$ defined using $sd^*_{anj}(\theta) := 1$ for $a = 1, 3$ and using φ^*_j rather than φ_j (where φ^*_j and φ_j are defined in Assumption A.5) for $j \le k$. We claim that this adjusted statistic is stochastically less than or equal to the original statistic $S^*_{n,EGMS}(\theta)$ defined in Section 4.1. This implies that the bootstrap critical value based on the adjusted $S^*_{n,EGMS}(\theta)$ statistic is less than or equal to that based on the original $S^*_{n,EGMS}(\theta)$ statistic. In turn this implies that if the test based on the adjusted $S^*_{n,EGMS}(\theta)$ statistic satisfies the result of Theorem 22.1(c), then the test based on the original $S^*_{n,EGMS}(\theta)$ statistic also satisfies the result of Theorem 22.1(c), which is the desired result.

The test based on the adjusted statistic $S^*_{n,EGMS}(\theta)$ satisfies the assumptions of Theorem 22.1(c) if the original test does with $sd_{1j\infty}(\theta) := 1$ in Assumption BC.2 and with the exception that the adjusted test does not require Assumption BC.1 because no statistics $sd^*_{anj}(\theta)$ for $a = 1, 3$ and $j \le k$ appear in its definition. Hence, under the assumptions of Theorem 22.1(c), but without imposing Assumption BC.1 and with $sd_{1j\infty}(\theta) := 1$ in Assumption BC.2, the adjusted test satisfies the result of Theorem 22.1(c).

To complete the argument above, it remains to show that the adjusted statistic $S^*_{n,EGMS}(\theta)$ is stochastically less than or equal to the original statistic $S^*_{n,EGMS}(\theta)$. This holds if the adjusted versions of $T^*_{nj,EGMS}(\theta)$ and $A^*_{n,EGMS}(\theta)$ are greater than or equal to the original statistics $T^*_{nj,EGMS}(\theta)$ and $A^*_{n,EGMS}(\theta)$ statistics, respectively, defined in Section 4.1, with probability one. The adjusted version of $T^*_{nj,EGMS}(\theta)$ depends on

$$\varphi^*_j(\overline{\xi}_{nj}(\theta)), \text{ where } \overline{\xi}_{nj}(\theta) := \kappa_n^{-1} n^{1/2}(\widehat{m}_{nj}(\theta) + \widehat{r}_n(\theta)), \tag{22.3}$$

whereas the original version of $T^*_{nj,EGMS}(\theta)$ depends on $\varphi_j(\xi_n(\theta), \widehat{\Omega}_n(\theta))$. We have

$$\varphi^*_j(\overline{\xi}_{nj}(\theta)) \ge \varphi^*_j(\xi_{nj}(\theta)) \ge \varphi_j(\xi_n(\theta), \widehat{\Omega}_n(\theta)), \tag{22.4}$$

where the first inequality holds because (i) if $\xi_{nj}(\theta) < 0$, then $\varphi^*_j(\xi_{nj}(\theta)) = 0$ by Assumption A.5(i) and (ii) and $\varphi^*_j(\overline{\xi}_{nj}(\theta)) \ge 0$ by Assumptions A.5(ii) and (iii), and (ii) if $\xi_{nj}(\theta) \ge 0$, then $\overline{\xi}_{nj}(\theta) \ge$

$\xi_{nj}(\theta) := (sd^*_{1nj}(\theta)\kappa_n)^{-1}n^{1/2}(\widehat{m}_{nj}(\theta) + \widehat{r}_n(\theta))$ (since $sd^*_{1nj}(\theta) \geq 1$ by its definition following (6.4)) and $\varphi^*_j(\cdot)$ is nondecreasing by Assumption A.5(ii), and the second inequality holds by Assumption A.5(i). Equation (22.4) gives the desired "greater than or equal to" result for the adjusted versus original $T^*_{nj,EGMS}(\theta)$ statistics. A completely analogous argument gives the desired "greater than or equal to" result for the adjusted versus original $A^*_{n,EGMS}(\theta)$ statistics.

23 Proofs of Lemmas 22.2–22.5

23.1 Proof of Lemma 22.2

Proof of Lemma 22.2. For notational simplicity, let $S^*_n := S^*_n(\theta_n)$, $S^*_{Ln} := S^*_{Ln}(\theta_n)$, $S^*_{Un} := S^*_{Un}(\theta_n)$, $c_{L\infty} := c_{L\infty}(1-\alpha)$, $c_{U\infty} := c_{U\infty}(1-\alpha)$, $\widehat{c}_n := \widehat{c}_n(\theta_n, 1-\alpha)$, and $c_\infty := c_\infty(1-\alpha)$. Let $\widehat{c}_{Ln}$ and $\widehat{c}_{Un}$ denote the $1-\alpha$ conditional quantiles of $S^*_{Ln}(\theta_n)$ and $S^*_{Un}(\theta_n)$, respectively, given $\{W_{ni}\}_{i\leq n,n\geq 1}$. Note that $\widehat{c}_{Ln}$ and $\widehat{c}_{Un}$ are random and depend on the conditioning value of $\{W_{ni}\}_{i\leq n,n\geq 1}$, whereas $c_{L\infty}$ and $c_{U\infty}$ denote the $1-\alpha$ conditional (or unconditional) quantiles of $S^*_{L\infty}$ and $S^*_{U\infty}$, respectively, which are nonrandom and do not depend on $\{W_{ni}\}_{i\leq n,n\geq 1}$ by Assumptions CV.1(ii) and CV.3(ii), respectively.

First, we prove part (a). If $S^*_{Ln}(\theta_n) \leq S^*_n(\theta_n)$ with probability one (with respect to the bootstrap randomness) conditional on $\{W_{ni}\}_{i\leq n,n\geq 1}$, then the $1-\alpha$ conditional quantile of $S^*_{Ln}(\theta_n)$ given $\{W_{ni}\}_{i\leq n,n\geq 1}$, which is $\widehat{c}_{Ln}$, is less than or equal to the $1-\alpha$ conditional quantile of $S^*_n(\theta_n)$ given $\{W_{ni}\}_{i\leq n,n\geq 1}$, which is $\widehat{c}_n$, as a consequence of the definition of a quantile. By Assumption CV.1(i), the "if" condition in the previous sentence holds wp→1 (with respect to the randomness in the sample, i.e., $\{W_{ni}\}_{i\leq n,n\geq 1}$). Hence, Assumption CV.1(i) implies that $\widehat{c}_{Ln} \leq \widehat{c}_n$ wp→1, which implies that $\widehat{c}_{Ln} \leq \widehat{c}_n + o_p(1)$, where the $o_p(1)$ term refers to randomness in the sample. This gives

$$\begin{aligned}\limsup_{n\to\infty} P_{F_n}(\phi_n(\theta_n) = 1) &= \limsup_{n\to\infty} P_{F_n}(S_n > \widehat{c}_n) \\ &\leq \limsup_{n\to\infty} P_{F_n}(S_n + o_p(1) > \widehat{c}_{Ln}). \end{aligned} \tag{23.1}$$

Now, take an arbitrary $\varepsilon > 0$. Then, there exists $\varepsilon^* \in (0, \varepsilon)$ such that $c_{L\infty} - \varepsilon^*$ is a continuity point of $S^*_{L\infty}$. We have

$$\begin{aligned}&\limsup_{n\to\infty} P_\triangledown(S^*_{Ln} \leq c_{L\infty} - \varepsilon | \{W_{ni}\}_{i\leq n,n\geq 1}) \\ \leq &\limsup_{n\to\infty} P_\triangledown(S^*_{Ln} \leq c_{L\infty} - \varepsilon^* | \{W_{ni}\}_{i\leq n,n\geq 1}) = P(S^*_{L\infty} \leq c_{L\infty} - \varepsilon^*) < 1-\alpha \end{aligned} \tag{23.2}$$

a.s.$[P_\triangledown]$, where the equality holds by Assumption CV.1(ii) and the second inequality holds by the definition of the quantile $c_{L\infty}$.

Note that

$$\left\{\limsup_{n\to\infty} P_{\bigtriangledown}(S^*_{Ln} \le c_{L\infty} - \varepsilon | \{W_{ni}\}_{i\le n, n\ge 1}) < 1-\alpha\right\} \subset \liminf_{n\to\infty} \{c_{L\infty} - \varepsilon \le \widehat{c}_{Ln}\}, \tag{23.3}$$

because for a sample path $\omega \in \Omega$ included in the left-hand side event, we must have that $c_{L\infty} - \varepsilon \le \widehat{c}_{Ln}$ for large n (where $\liminf_{n\to\infty} B_n := \cup_{k\ge 1} \cap_{n\ge k} B_n$ for $B_n \subset \Omega$). Taking expectations, we obtain

$$\begin{aligned} & P_{\bigtriangledown}\left(\limsup_{n\to\infty} P_{\bigtriangledown}(S^*_{Ln} \le c_{L\infty} - \varepsilon | \{W_{ni}\}_{i\le n, n\ge 1}) < 1-\alpha\right) \\ \le\ & P_{\bigtriangledown}\left(\liminf_{n\to\infty} \{c_{L\infty} - \varepsilon \le \widehat{c}_{Ln}\}\right) \\ \le\ & \liminf_{n\to\infty} P_{\bigtriangledown}\left(c_{L\infty} - \varepsilon \le \widehat{c}_{Ln}\right) \\ =\ & \liminf_{n\to\infty} P_{F_n}\left(c_{L\infty} - \varepsilon \le \widehat{c}_{Ln}\right), \end{aligned} \tag{23.4}$$

where the second inequality follows from $P_{\bigtriangledown}(\liminf_{n\to\infty} B_n) \le \liminf_{n\to\infty} P_{\bigtriangledown}(B_n)$ for $B_n \in \mathcal{F}$ (which holds because $P_{\bigtriangledown}(\liminf_{n\to\infty} B_n) = \lim_{k\to\infty} P_{\bigtriangledown}(\cap_{n\ge k} B_n) \le \lim_{k\to\infty} \inf_{n\ge k} P_{\bigtriangledown}(B_n)$) and the equality holds because $\widehat{c}_{Ln}$ depends only on $\{W_{ni}\}_{i\le n}$ or $\{W_i\}_{i\le n}$, which have the same distribution. Since the probability on the first line of (23.4) equals one by (23.2), we have shown that

$$\liminf_{n\to\infty} P_{F_n}\left(c_{L\infty} - \varepsilon \le \widehat{c}_{Ln}\right) = 1 \ \forall \varepsilon > 0. \tag{23.5}$$

Next, we have

$$\begin{aligned} & \limsup_{n\to\infty} P_{F_n}(S_n + o_p(1) > \widehat{c}_{Ln}) \\ =\ & \limsup_{n\to\infty} P_{F_n}(S_n + o_p(1) > \widehat{c}_{Ln} \ \&\ c_{L\infty} - \varepsilon \le \widehat{c}_{Ln}) \\ \le\ & \limsup_{n\to\infty} P_{F_n}(S_n + o_p(1) > c_{L\infty} - \varepsilon \ \&\ c_{L\infty} - \varepsilon \le \widehat{c}_{Ln}) \\ =\ & \limsup_{n\to\infty} P_{F_n}(S_n + o_p(1) > c_{L\infty} - \varepsilon), \end{aligned} \tag{23.6}$$

where the two equalities hold using (23.5) and the inequality is straightforward.

By Theorem 5.3(e), we have

$$S_n \to_d S_\infty \tag{23.7}$$

using Assumptions A.0, C.1–C.7, S.1(iii), and NLA. Consider a sequence $\{\varepsilon_m\}_{m\ge 1}$ such that $c_\infty - \varepsilon_m$

is a continuity point of S_∞ for all $m \geq 1$ and $\varepsilon_m \downarrow 0$. Then, we have

$$\begin{aligned}\limsup_{n\to\infty} P_{F_n}(S_n + o_p(1) > \widehat{c}_{Ln}) &\leq \lim_{m\to\infty} \limsup_{n\to\infty} P_{F_n}(S_n + o_p(1) > c_{L\infty} - \varepsilon_m) \\ &= \lim_{m\to\infty} P(S_\infty > c_{L\infty} - \varepsilon_m) \\ &= P(S_\infty > c_{L\infty}), \end{aligned} \tag{23.8}$$

where the inequality holds by (23.6), the first equality holds by (23.7) and the definition of $\{\varepsilon_m\}_{m\geq 1}$, and the second equality holds by the monotone convergence theorem. This and (23.1) complete the proof of part (a).

Next, we prove part (c). By Assumption S.1(ii), there are two possible cases: (i) $c_\infty = 0$ and (ii) $c_\infty > 0$. First, if $c_\infty = 0$, the result follows immediately because

$$\limsup_{n\to\infty} P_{F_n}(S_n > \widehat{c}_n) \leq \limsup_{n\to\infty} P_{F_n}(S_n > 0) \leq \alpha, \tag{23.9}$$

where the first inequality holds because $\widehat{c}_n \geq 0$ (since S_n^* is nonnegative by assumption) and the second holds by Assumption A.7(ii).

Second, we consider the case where $c_\infty > 0$. By (23.1), it suffices to show

$$\limsup_{n\to\infty} P_{F_n}(S_n + o_p(1) > \widehat{c}_{Ln}) \leq \alpha. \tag{23.10}$$

By Lemma 20.1, under Assumption C.3, Assumption N implies NLA. Hence, the assumptions of part (c) imply those of part (a) and (23.8) holds under the assumptions of part (c). Using (23.8), we have

$$\limsup_{n\to\infty} P_{F_n}(S_n + o_p(1) > \widehat{c}_{Ln}) \leq P(S_\infty > c_{L\infty}) \leq P(S_\infty > c_\infty) = \alpha, \tag{23.11}$$

where the second inequality holds by Assumption CV.2 because $S_{L\infty}^* \geq_{ST} S_\infty$ implies that $c_{L\infty} \geq c_\infty$, and the equality holds by Assumption A.7(i).

Now, we prove part (b). The proof is similar to that of part (a), but there are some differences, such as the need for Assumption BC.5, and we use parts of the proof of part (c) in the proof of part (d), so we provide the details. By the same argument as in the paragraph containing (23.1), but with Assumption CV.3 in place of Assumption CV.1, we obtain $\widehat{c}_{Un} \geq \widehat{c}_n$ wp→1, which implies that

$$
\begin{aligned}
\widehat{c}_{Un} &\geq \widehat{c}_n + o_p(1), \text{ and} \\
\liminf_{n\to\infty} P_{F_n}(\phi_n(\theta_n) = 1) &= \liminf_{n\to\infty} P_{F_n}(S_n > \widehat{c}_n) \\
&\geq \liminf_{n\to\infty} P_{F_n}(S_n + o_p(1) > \widehat{c}_{Un}),
\end{aligned} \tag{23.12}
$$

where the $o_p(1)$ terms refer to randomness in the sample, not bootstrap randomness.

Consider an arbitrary $\varepsilon > 0$. There exists an $\varepsilon^* \in (0, \varepsilon)$ such that $c_{U\infty} + \varepsilon^*$ is a continuity point of $S^*_{U\infty}$. We have

$$
\begin{aligned}
&\liminf_{n\to\infty} P_{\bigtriangledown}(S^*_{Un} \leq c_{U\infty} + \varepsilon | \{W_{ni}\}_{i\leq n, n\geq 1}) \\
&\geq \liminf_{n\to\infty} P_{\bigtriangledown}(S^*_{Un} \leq c_{U\infty} + \varepsilon^* | \{W_{ni}\}_{i\leq n, n\geq 1}) \\
&= P(S^*_{U\infty} \leq c_{U\infty} + \varepsilon^*) \\
&> 1 - \alpha
\end{aligned} \tag{23.13}
$$

a.s.$[P_{\bigtriangledown}]$, where the equality holds by Assumption CV.3(ii) and the second inequality holds by Assumption BC.5 and the definition of the quantile $c_{U\infty}$.

Note that

$$
\left\{\liminf_{n\to\infty} P_{\bigtriangledown}(S^*_{Un} \leq c_{U\infty} + \varepsilon | \{W_{ni}\}_{i\leq n, n\geq 1}) > 1 - \alpha\right\} \subset \liminf_{n\to\infty} \{c_{U\infty} + \varepsilon \geq \widehat{c}_{Un}\}, \tag{23.14}
$$

because for a sample path $\omega \in \Omega$ included in the left-hand side event, we must have that $c_{U\infty} + \varepsilon \geq \widehat{c}_{Un}$ for large n. Taking expectations, we obtain

$$
\begin{aligned}
&P_{\bigtriangledown}\left(\liminf_{n\to\infty} P_{\bigtriangledown}(S^*_{Un} \leq c_{U\infty} + \varepsilon | \{W_{ni}\}_{i\leq n, n\geq 1}) > 1 - \alpha\right) \\
&\leq P_{\bigtriangledown}\left(\liminf_{n\to\infty} \{c_{U\infty} + \varepsilon \geq \widehat{c}_{Un}\}\right) \\
&\leq \liminf_{n\to\infty} P_{\bigtriangledown}(c_{U\infty} + \varepsilon \geq \widehat{c}_{Un}) \\
&= \liminf_{n\to\infty} P_{F_n}(c_{U\infty} + \varepsilon \geq \widehat{c}_{Un})
\end{aligned} \tag{23.15}
$$

for the same reasons as in (23.4). Since the probability on the first line of (23.15) equals one by (23.13), we have shown that

$$
\liminf_{n\to\infty} P_{F_n}(c_{U\infty} + \varepsilon \geq \widehat{c}_{Un}) = 1. \tag{23.16}
$$

Next, we have

$$\begin{aligned}
&\liminf_{n\to\infty} P_{F_n}(S_n + o_p(1) > \widehat{c}_{Un}) \\
&= \liminf_{n\to\infty} P_{F_n}(S_n + o_p(1) > \widehat{c}_{Un} \ \&\ c_{U\infty} + \varepsilon \geq \widehat{c}_{Un}) \\
&\geq \liminf_{n\to\infty} P_{F_n}(S_n + o_p(1) > c_{U\infty} + \varepsilon \ \&\ c_{U\infty} + \varepsilon \geq \widehat{c}_{Un}) \\
&= \liminf_{n\to\infty} P_{F_n}(S_n + o_p(1) > c_{U\infty} + \varepsilon), \qquad (23.17)
\end{aligned}$$

where the two equalities hold using (23.16) and the inequality is straightforward.

Consider a sequence $\{\varepsilon_m\}_{m\geq 1}$ such that $c_\infty + \varepsilon_m$ is a continuity point of S_∞ for all $m \geq 1$ and $\varepsilon_m \downarrow 0$. Then, we have

$$\begin{aligned}
\liminf_{n\to\infty} P_{F_n}(S_n + o_p(1) > \widehat{c}_{Un}) &\geq \lim_{m\to\infty} \liminf_{n\to\infty} P_{F_n}(S_n + o_p(1) > c_{U\infty} + \varepsilon_m) \\
&= \lim_{m\to\infty} P(S_\infty > c_{U\infty} + \varepsilon_m) \\
&= P(S_\infty > c_{U\infty}), \qquad (23.18)
\end{aligned}$$

where the inequality holds by (23.17), the first equality holds by (23.7) and the definition of $\{\varepsilon_m\}_{m\geq 1}$, and the second equality holds by the monotone convergence theorem. This and (23.12) complete the proof of part (b).

Lastly, we establish part (d). By Theorem 5.3(h), $S_n \to_p \infty$ (using Assumptions A.0, C.1–C.7, CA, S.1(iii), S.2, and S.3). Hence, it suffices to show that $\widehat{c}_n = O_p(1)$. First, suppose the support of $S^*_{U\infty}$ is bounded above. Then, there exists $c < \infty$ such that $P(S^*_{U\infty} \leq c) = 1$. For any $\varepsilon > 0$, we have

$$P_\bigtriangledown(S^*_{Un} \leq c + \varepsilon | \{W_{ni}\}_{i\leq n, n\geq 1}) \to P(S^*_{U\infty} \leq c + \varepsilon) = 1 \text{ a.s.}[P_\bigtriangledown] \qquad (23.19)$$

by Assumption CV.3(ii). We obtain

$$\begin{aligned}
\liminf_{n\to\infty} P_{F_n}(S^*_n \leq c + \varepsilon) &\geq \liminf_{n\to\infty} P_{F_n}(S^*_{Un} \leq c + \varepsilon) \\
&= \liminf_{n\to\infty} E_\bigtriangledown P_\bigtriangledown(S^*_{Un} \leq c + \varepsilon | \{W_{ni}\}_{i\leq n, n\geq 1}) = 1, \qquad (23.20)
\end{aligned}$$

where the inequality holds by Assumption CV.3(i), the first equality holds by the law of iterated expectations, and the second equality holds by the dominated convergence theorem using (23.19). Since $\widehat{c}_n$ is the $1-\alpha$ quantile of S^*_n, (23.20) implies that $\widehat{c}_n \leq c + 2\varepsilon$ wp→1, which implies that $\widehat{c}_n = O_p(1)$, as desired.

Next, we consider the case where the support of $S^*_{U\infty}$ is not bounded above. Then, there exists $\alpha_1 < \alpha$ such that the $1-\alpha_1$ quantile of $S^*_{U\infty}$ exceeds the $1-\alpha$ quantile of $S^*_{U\infty}$ (and is finite because $S^*_{U\infty} \in R$ a.s. by Assumption CV.3(ii)). By (23.13), but with $c_{U\infty}$ defined to equal the $1 - \alpha_1$ quantile of $S^*_{U\infty}$, rather than the $1 - \alpha$ quantile, we obtain the result of (23.13) (with the rhs being $1 - \alpha$ and without imposing Assumption BC.5). In consequence, (23.14)–(23.16) give

$$\liminf_{n\to\infty} P_{F_n}\left(c_{U\infty} + \varepsilon \geq \widehat{c}_{Un}\right) = 1, \tag{23.21}$$

where $c_{U\infty}$ is the $1 - \alpha_1$ quantile of $S^*_{U\infty}$. Since $c_{U\infty} < \infty$, this yields $\widehat{c}_{Un} = O_p(1)$. By (23.12), we have

$$\liminf_{n\to\infty} P_{F_n}(\phi_n(\theta_n) = 1) \geq \liminf_{n\to\infty} P_{F_n}(S_n + o_p(1) > \widehat{c}_{Un}) = 1, \tag{23.22}$$

where the equality holds because $S_n \to_p \infty$ by Theorem 5.3(h) and $\widehat{c}_{Un} = O_p(1)$. This completes the proof of part (d). $\square$

23.2 Proof of Lemma 22.3

Proof of Lemma 22.3. First, we prove part (a). For all $j \leq k$, we have

$$n^{1/2}\left(\widehat{m}_{nj}(\theta) - E_{F_n}\widetilde{m}_j(W,\theta)\right) = O^{\Theta}_p(1), \tag{23.23}$$

by (21.5) and Assumption C.5. Hence, we obtain

$$\sup_{\theta\in\Theta} |\widehat{m}_{nj}(\theta) - \widetilde{m}_j(\theta)| = o_p(1) \tag{23.24}$$

using Assumption C.4. Now, we use the result that for any sequence of random variables $\{X_n\}_{n\geq1}$ on $(\Omega, \mathcal{F}, P_{\triangledown})$ for which $X_n \to_p 0$, there exists a subsequence $\{c_n\}_{n\geq1}$ of $\{n\}_{n\geq1}$ such that $X_{c_n} \to 0$ a.s.$[P_{\triangledown}]$, e.g., see Theorem 9.2.1 of Dudley (1989). We apply this result with the original sequence $\{n\}_{n\geq1}$ replaced by some subsequence $\{p_n\}_{n\geq1}$. Using this and (23.24), given any subsequence $\{p_n\}_{n\geq1}$ of $\{n\}_{n\geq1}$, there exists a subsequence $\{c_n\}_{n\geq1}$ of $\{p_n\}_{n\geq1}$ such that

$$\sup_{\theta\in\Theta} |\widehat{m}_{c_nj}(\theta) - \widetilde{m}_j(\theta)| = o(1) \text{ a.s.}[P_{\triangledown}]. \tag{23.25}$$

By the continuity of $\widetilde{m}_j(\theta)$ (Assumption C.4) and $\theta_n \to \theta_\infty$ (Assumption C.1), (23.25) gives

$$\widehat{m}_{c_nj}(\theta_{c_n}) \to \widetilde{m}_j(\theta_\infty) \text{ a.s.}[P_{\triangledown}]. \tag{23.26}$$

Conditional on $\{W_{ni}\}_{i\leq n,n\geq 1}$, for the subsequence $\{c_n\}_{n\geq 1}$, we have

$$c_n^{1/2}\left(\frac{\widehat{\sigma}^*_{c_nj}(\theta_{c_n})}{\widehat{\sigma}_{c_nj}(\theta_{c_n})}-1\right)\to_d \frac{1}{2}G^{\sigma}_{j\infty} \text{ a.s.}[P_{\bigtriangledown}]\ \forall j\leq k. \tag{23.27}$$

This holds by the delta method, as in (21.16) with $\widehat{\sigma}^{*2}_{nj}(\theta_n)$ and $\widehat{\sigma}^2_{nj}(\theta_n)$ in place of $\widehat{\sigma}^2_{nj}(\theta_n)$ and $\sigma^2_{F_nj}(\theta_n)$, respectively, and using Assumption BC.6 in place of (21.15).

Next, suppressing the dependence of various quantities on θ_{c_n} for notational simplicity, we have: conditional on $\{W_{ni}\}_{i\leq n,n\geq 1}$,

$$\begin{aligned}
T^{**}_{c_nj} :=&\ c_n^{1/2}\left(\frac{\overline{m}^*_{c_nj}}{\widehat{\sigma}^*_{c_nj}}-\frac{\overline{m}_{c_nj}}{\widehat{\sigma}_{c_nj}}\right)\\
=&\ \left(\frac{\widehat{\sigma}_{c_nj}}{\widehat{\sigma}^*_{c_nj}}\right)\left(c_n^{1/2}\left(\frac{\overline{m}^*_{c_nj}}{\widehat{\sigma}_{c_nj}}-\frac{\overline{m}_{c_nj}}{\widehat{\sigma}_{c_nj}}\right)-\frac{\overline{m}_{c_nj}}{\widehat{\sigma}_{c_nj}}c_n^{1/2}\frac{\widehat{\sigma}^*_{c_nj}-\widehat{\sigma}_{c_nj}}{\widehat{\sigma}_{c_nj}}\right)\\
=&\ \left(\frac{\widehat{\sigma}_{c_nj}}{\widehat{\sigma}^*_{c_nj}}\right)\left(\nu^{m*}_{c_nj}-\widehat{m}_{c_nj}c_n^{1/2}\frac{\widehat{\sigma}^*_{c_nj}-\widehat{\sigma}_{c_nj}}{\widehat{\sigma}_{c_nj}}\right)\\
\to_d&\ G^m_{j\infty}-\frac{1}{2}\widetilde{m}_{j\infty}G^{\sigma}_{j\infty} =: G^{m\sigma}_{j\infty} \text{ a.s.}[P_{\bigtriangledown}]
\end{aligned}\tag{23.28}$$

$\forall j\leq k$, where $\widetilde{m}_{j\infty}=\widetilde{m}_j(\theta_\infty)$ by (5.5), $G^m_{j\infty}:=G^m_j(\theta_\infty)$ and $G^{\sigma}_{j\infty}:=G^{\sigma}_j(\theta_\infty)$ by (5.7), the second equality holds by algebra, the third equality uses the definition of $\nu^{m*}_{c_nj}(\theta_{c_n})$ in (6.1), the convergence holds by (23.26), (23.27), and Assumptions BC.6 and C.1, and the last equality holds by (5.7).

We have $T^*_{Lnj,EGMS}(\theta_n)=T^{**}_{nj}+\varphi^*_j(\xi_{nj}(\theta_n))$ by (6.1), (12.4), and (23.28), and $T^*_{Lj\infty,EGMS}=G^{m\sigma}_{j\infty}+\varphi^*_j(h^*_{j\infty})$ by (12.5) for all $j\leq k$. By (23.28), there exists a subsequence $\{c_n\}_{n\geq 1}$ of $\{p_n\}_{n\geq 1}$ for which $\{T^{**}_{c_nj}(\theta_{c_n})|\{W_{ni}\}_{i\leq n,n\geq 1}\}\to_d G^{m\sigma}_{j\infty}$ a.s.$[P_{\bigtriangledown}]$. Hence, part (a) holds if there exists a subsequence $\{a_n\}_{n\geq 1}$ of $\{c_n\}_{n\geq 1}$ for which

$$\{\varphi^*_j(\xi_{a_nj}(\theta_{a_n}))|\{W_{ni}\}_{i\leq n,n\geq 1}\}\to\varphi^*_j(h^*_{j\infty}) \text{ a.s.}[P_{\bigtriangledown}]\ \forall j\leq k. \tag{23.29}$$

We have

$$
\begin{aligned}
& \xi_{nj}(\theta_n) \\
:= \ & (sd^*_{1nj}(\theta_n)\kappa_n)^{-1} n^{1/2} \left(\widehat{m}_{nj}(\theta_n) + \widehat{r}_n(\theta_n) \right) \\
= \ & (sd^*_{1nj}(\theta_n)\kappa_n)^{-1} n^{1/2} \left(\widehat{m}_{nj}(\theta_n) - E_{F_n} \widetilde{m}_j(W,\theta_n) \right) \\
& +(sd^*_{1nj}(\theta_n)\kappa_n)^{-1} \Big([n^{1/2}(\widehat{m}_{nj}(\theta_n) - E_{F_n} \widetilde{m}_j(W,\theta_n)) + n^{1/2} E_{F_n} \widetilde{m}_j(W,\theta_n)]_- \Big. \\
& \qquad \Big. -[n^{1/2} E_{F_n} \widetilde{m}_j(W,\theta_n)]_- \Big) \\
& +(sd^*_{1nj}(\theta_n)\kappa_n)^{-1} n^{1/2} (E_{F_n} \widetilde{m}_j(W,\theta_n) + r_{F_n}(\theta_n)) \\
\to_p \ & h^*_{j\infty}, \qquad (23.30)
\end{aligned}
$$

where the first equality holds by definition, see (12.4), the second equality holds using (2.6) and (4.1), and the convergence holds using $sd^*_{1nj}(\theta_n)/sd_{1j\infty} \to_p 1$ by Assumption BC.1 and $sd_{1j\infty} := sd_{1j\infty}(\theta_\infty)$, $n^{1/2}(\widehat{m}_{nj}(\theta_n) - E_{F_n}\widetilde{m}_j(W,\theta_n)) = O_p(1)$ (by (23.23)), $|\chi(\nu,c)| := |[\nu+c]_- - [c]_-| \leq |\nu|$ for $\nu, c \in R$ (by (21.7)), $\kappa_n \to \infty$ (by Assumption A.6(i)), and Assumption BC.2 (which relies on Assumption BC.1 for the definition of $sd_{1j\infty}(\theta)$).

Equation (23.30) and the continuity of $\varphi^*_j(\xi_j)$ at all $\xi_j \in R_{[+\infty]}$ (by Assumption A.5(ii)) give $d(\varphi^*_j(\xi_{nj}(\theta_n)), \varphi^*_j(h^*_{j\infty})) \to_p 0$ for all $j \leq k$. Now, we use the result that for any sequence of random variables $\{X_n\}_{n\geq 1}$ on $(\Omega, \mathcal{F}, P_\bigtriangledown)$ for which $X_n \to_p 0$, there exists a subsequence $\{a_n\}_{n\geq 1}$ of $\{c_n\}_{n\geq 1}$ such that $X_{a_n} \to 0$ a.s.$[P_\bigtriangledown]$. Thus, there exists a subsequence $\{a_n\}_{n\geq 1}$ of $\{p_n\}_{n\geq 1}$ such that (23.29) holds, which completes the proof of part (a).

Now, we prove part (b). Define

$$
\nu^{m\sigma*}_{nj}(\theta) := \nu^{m*}_{nj}(\theta) - \frac{1}{2}\widetilde{m}_j(\theta)\nu^{\sigma*}_{nj}(\theta) \ \forall j \leq k. \qquad (23.31)
$$

We show that under $\{F_n\}_{n\geq 1}$, conditional on $\{W_{ni}\}_{i\leq n, n\geq 1}$, for the subsequence $\{c_n\}_{n\geq 1}$ of $\{p_n\}_{n\geq 1}$ defined above,

$$
\sup_{\theta\in\Theta} |\widehat{\nu}^*_{c_n j}(\theta) - \nu^{m\sigma*}_{c_n j}(\theta)| = o_p(1) \text{ a.s.}[P_\bigtriangledown]. \qquad (23.32)
$$

This, Assumption BC.6, (23.25), and the continuous mapping theorem give: under $\{F_n\}_{n\geq 1}$, conditional on $\{W_{ni}\}_{i\leq n, n\geq 1}$, for the subsequence $\{c_n\}_{n\geq 1}$ of $\{p_n\}_{n\geq 1}$,

$$
\widehat{\nu}^*_{c_n j}(\cdot) = \nu^{m\sigma*}_{c_n j}(\cdot) + o^\Theta_p(1) \Rightarrow G^{m\sigma}_j(\cdot) \text{ a.s.}[P_\bigtriangledown]. \qquad (23.33)
$$

The proof of (23.32) is quite similar to (21.4) and (21.5), but with bootstrap quantities in place of original sample quantities. By the same argument as in (21.4) with $\widehat{\sigma}^*_{nj}(\theta)$ and $\widehat{\sigma}_{nj}(\theta)$ in place of $\widehat{\sigma}_{nj}(\theta)$ and $\sigma_{F_nj}(\theta)$, respectively, we obtain

$$n^{1/2}\left(\frac{\widehat{\sigma}^*_{nj}(\theta)}{\widehat{\sigma}_{nj}(\theta)}-1\right)=\frac{1}{2}\nu^{\sigma*}_{nj}(\theta)+o^\Theta_p(1)\text{ a.s.}[P_\bigtriangledown], \tag{23.34}$$

using Assumption BC.6 in place of Assumption C.5 and (21.2). Next, we have: conditional on $\{W_{ni}\}_{i\le n,n\ge1}$, for the subsequence $\{c_n\}_{n\ge1}$,

$$\begin{aligned}\widehat{\nu}^*_{c_nj}(\theta) &:= c_n^{1/2}\left(\frac{\overline{m}^*_{c_nj}(\theta)}{\widehat{\sigma}^*_{c_nj}(\theta)}-\widehat{m}_{c_nj}(\theta)\right)=\frac{\widehat{\sigma}_{c_nj}(\theta)}{\widehat{\sigma}^*_{c_nj}(\theta)}\left(\nu^{m*}_{c_nj}(\theta)-\widehat{m}_{c_nj}(\theta)c_n^{1/2}\left(\frac{\widehat{\sigma}^*_{c_nj}(\theta)}{\widehat{\sigma}_{c_nj}(\theta)}-1\right)\right)\\ &= (1+o^\Theta_p(1))\left(\nu^{m*}_{c_nj}(\theta)-\frac{1}{2}\widetilde{m}_j(\theta)\nu^{\sigma*}_{c_nj}(\theta)+o^\Theta_p(1)\right)=\nu^{m\sigma*}_{c_nj}(\theta)+o^\Theta_p(1)\text{ a.s.}[P_\bigtriangledown],\end{aligned} \tag{23.35}$$

where the third equality holds by (23.25) and (23.34), and the fourth equality holds by the definition of $\nu^{m\sigma*}_{nj}(\theta)$ in (23.31) and Assumption BC.6. This proves (23.32).

Next, we have

$$\begin{aligned} n^{1/2}\widehat{m}_{nj}(\theta) &= \frac{\sigma_{F_nj}(\theta)}{\widehat{\sigma}_{nj}(\theta)}\left(\nu^m_{nj}(\theta)+n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta)\right)\\ &= \widehat{\omega}_{nj}(\theta)+n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta),\text{ where}\\ \widehat{\omega}_{nj}(\theta) &:= \frac{\sigma_{F_nj}(\theta)}{\widehat{\sigma}_{nj}(\theta)}\nu^m_{nj}(\theta)-n^{1/2}\left(\frac{\widehat{\sigma}_{nj}(\theta)}{\sigma_{F_nj}(\theta)}-1\right)\frac{\sigma_{F_nj}(\theta)}{\widehat{\sigma}_{nj}(\theta)}E_{F_n}\widetilde{m}_j(W,\theta)=O^\Theta_p(1),\end{aligned} \tag{23.36}$$

where $\nu^m_{nj}(\theta)$ denotes the jth element of $\nu^m_n(\theta)$ defined in (2.12), and the second equality on the last line holds by Assumptions C.4 and C.5 and (21.4). Now, we have

$$\begin{aligned} n^{1/2}\left([\widehat{m}_{nj}(\theta)]_- - \widehat{r}^{\inf}_n\right) &= n^{1/2}([\widehat{m}_{nj}(\theta)]_- - [E_F\widetilde{m}_j(W,\theta)]_-)-n^{1/2}(\widehat{r}^{\inf}_n-r^{\inf}_{F_n})+b_{nj}(\theta)\\ &= \widehat{d}_{nj}(\theta)+b_{nj}(\theta),\text{ where}\\ \widehat{d}_{nj}(\theta) &:= \chi(\widehat{\omega}_{nj}(\theta),n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta))-n^{1/2}(\widehat{r}^{\inf}_n-r^{\inf}_{F_n})=O^\Theta_p(1),\end{aligned} \tag{23.37}$$

the first equality uses the definition $b_{nj}(\theta):=n^{1/2}([E_{F_n}\widetilde{m}_j(W,\theta)]_- - r^{\inf}_{F_n})$ in (12.1), the second equality uses $\chi(\nu,c):=[\nu+c]_- - [c]_-$, and the second equality on the last line holds because $|\chi(\nu,c)|\le|v|\ \forall\nu,c\in R$ by (21.7), $\widehat{\omega}_{nj}(\theta)=O^\Theta_p(1)$ by (23.36), and $n^{1/2}(\widehat{r}^{\inf}_n-r^{\inf}_{F_n}):=A^{\inf}_n=O_p(1)$ by (5.2) and Theorem 5.3(b) (which uses Assumptions A.0, C.4, C.5, and C.7).

For $b^*_j=(sd_{3j\infty}(\theta)\kappa_n)^{-1}n^{1/2}([E_{F_n}\widetilde{m}_j(W,\theta)]_- - r^{\inf}_{F_n})$ as in $\Lambda^{*\eta_n}_{n,F_n}$ (defined in (12.3)), we obtain

$$\xi^b_{nj}(\theta):=(sd^*_{3nj}(\theta)\kappa_n)^{-1}n^{1/2}\left([\widehat{m}_{nj}(\theta)]_- - \widehat{r}^{\inf}_n\right)=(sd^*_{3nj}(\theta)\kappa_n)^{-1}\widehat{d}_{nj}(\theta)+\frac{sd_{3j\infty}(\theta)}{sd^*_{3nj}(\theta)}b^*_j, \tag{23.38}$$

where the first equality holds by definition, see (12.4), and the second equality holds by (23.37).

Using (23.36), (23.38), and the definition of $\Lambda^{*\eta_n}_{n,F_n}$, we can write $A^{*\inf}_{Ln,EGMS}$ in (12.4) as

$$A^{*\inf}_{Ln,EGMS} = \inf_{(\theta,b,b^*,\ell,j^*)\in\Lambda^{*\eta_n}_{n,F_n}} \max_{j\le k}\Bigg(\chi(\widehat{\nu}^*_{nj}(\theta),\ell_j)+1(j\neq j^*)b_j \tag{23.39}$$
$$+1(j=j^*)\varphi^*_j\left((sd^*_{3nj}(\theta)\kappa_n)^{-1}\widehat{d}_{nj}(\theta)+\frac{sd_{3j\infty}(\theta)}{sd^*_{3nj}(\theta)}b^*_j\right)\Bigg),$$

where $(\theta,b_j,b^*_j,\ell_j,j^*)\in\Lambda^{*\eta_n}_{n,F_n}$ implies that $b_j := b_{nj}(\theta)$, $b^*_j := (sd_{3j\infty}(\theta)\kappa_n)^{-1}b_{nj}(\theta)$, $\ell_j := n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta)$, and $j^* := j_n(\theta)$ and $\chi(\nu,c) := [\nu+c]_- - [c]_-$.

We have $(sd^*_{3nj}(\theta)\kappa_n)^{-1}\widehat{d}_{nj}(\theta) = o^\Theta_p(1)$ by (23.36), (23.37), Assumption A.6(i), and $sd^*_{3nj}(\theta)\ge 1$ (by its definition following (6.8)). Also, by Assumption BC.1 and $sd^*_{3nj}(\theta)\ge 1$, we have $\sup_{\theta\in\Theta}|sd_{3j\infty}(\theta)/sd^*_{3nj}(\theta)-1|\to_p 0$. Hence, by the same argument as used to establish (23.25), there exists a subsequence $\{a_n\}_{n\ge1}$ (different from that in the proof of part (a)) of $\{c_n\}_{n\ge1}$ for which

$$\sup_{\theta\in\Theta}|(sd^*_{3a_nj}(\theta)\kappa_{a_n})^{-1}\widehat{d}_{a_nj}(\theta)|\to 0 \text{ a.s.}[P_\bigtriangledown] \text{ and } \sup_{\theta\in\Theta}\left|\frac{sd_{3j\infty}(\theta)}{sd^*_{3a_nj}(\theta)}-1\right|\to 0 \text{ a.s.}[P_\bigtriangledown]. \tag{23.40}$$

In addition, by (23.33), under $\{F_n\}_{n\ge1}$, conditional on $\{W_{ni}\}_{i\le n,n\ge1}$, the subsequence $\{a_n\}_{n\ge1}$ of $\{p_n\}_{n\ge1}$ is such that

$$\widehat{\nu}^*_{a_nj}(\cdot)=\nu^{m\sigma*}_{a_nj}(\cdot)+o^\Theta_p(1)\Rightarrow G^{m\sigma}_j(\cdot) \text{ a.s.}[P_\bigtriangledown]. \tag{23.41}$$

Define

$$\overrightarrow{A}^{*\inf}_{Ln,EGMS} := \inf_{(\theta,b,b^*,\ell,j^*)\in\Lambda^{*\eta_n}_{n,F_n}}\max_{j\le k}\Big(\chi(\nu^{m\sigma*}_{nj}(\theta),\ell_j)+1(j\neq j^*)b_j$$
$$+1(j=j^*)\varphi^*_j(\mu_{1nj}(\theta)+\mu_{2nj}(\theta)b^*_j)\Big),\text{ where}$$
$$\mu_{1nj}(\theta) := (sd^*_{3nj}(\theta)\kappa_n)^{-1}\widehat{d}_{nj}(\theta) \text{ and } \mu_{2nj}(\theta) := \frac{sd_{3j\infty}(\theta)}{sd^*_{3a_nj}(\theta)}, \tag{23.42}$$

$\mu_{sn}(\theta)=(\mu_{sn1}(\theta),...,\mu_{snk}(\theta))'$ for $s=1,2$, and $\mu_n(\theta)=(\mu_{1n}(\theta)',\mu_{2n}(\theta)')'$.

By (23.39), (23.41), and (23.42), we obtain:

$$A^{*\inf}_{La_n,EGMS}=\overrightarrow{A}^{*\inf}_{La_n,EGMS}+o_p(1) \text{ a.s.}[P_\bigtriangledown], \tag{23.43}$$

using the continuity of $\varphi^*_j(\xi_j)$ on $R^k_{[\pm\infty]}$ by Assumption A.5(ii) and the continuity of $\chi(\nu,c)$ on

$R \times R_{[\pm\infty]}$ under d. Hence, to establish part (b), it suffices to show: conditional on $\{W_{ni}\}_{i \leq n, n \geq 1}$, for the subsequence $\{a_n\}_{n \geq 1}$,

$$\left\{ \overrightarrow{A}^{*\inf}_{La_n,EGMS} | \{W_{ni}\}_{i \leq n, n \geq 1} \right\} \rightarrow_d A^{*\inf}_{L\infty,EGMS} \text{ a.s.}[P_{\bigtriangledown}]. \tag{23.44}$$

To prove (23.44), we use a similar (but more complicated) argument to that used to prove Theorem 5.3(b) based on the extended continuous mapping theorem. As above, let $\mathcal{D}$ be the space of functions from Θ to R^{2k}. Let $\mathcal{D}_0$ be the subset of uniformly continuous functions in $\mathcal{D}$. For non-stochastic functions $\nu(\cdot) \in \mathcal{D}$ and $\mu(\cdot) : \Theta \rightarrow R^{2k}$ with $\mu(\theta) = (\mu_{11}(\theta), ..., \mu_{1k}(\theta), \mu_{21}(\theta), ..., \mu_{2k}(\theta))'$, define

$$\begin{aligned} \widetilde{g}_n(\nu(\cdot), \mu(\cdot)) :=& \inf_{(\theta,b,b^*,\ell,j^*) \in \Lambda^{*\eta_n}_{n,F_n}} \max_{j \leq k} \Big(\tau_j(\nu(\cdot), \theta, \ell) + 1(j \neq j^*) b_j \\ & \qquad +1(j = j^*) \varphi^*_{j^*}(\mu_{1j^*}(\theta) + \mu_{2j^*}(\theta) b^*_{j^*}) \Big), \\ \widetilde{g}(\nu(\cdot), \mu(\cdot)) :=& \inf_{(\theta,b,b^*,\ell,j^*) \in \Lambda^*_I} \max_{j \leq k} \Big(\tau_j(\nu(\cdot), \theta, \ell) + 1(j \neq j^*) b_j \\ & \qquad +1(j = j^*) \varphi^*_{j^*}(\mu_{1j^*}(\theta) + \mu_{2j^*}(\theta) b^*_{j^*}) \Big), \end{aligned} \tag{23.45}$$

where $\nu(\theta) = (\nu^m(\theta)', \nu^\sigma(\theta)')'$, $\nu^m_j(\theta)$ and $\nu^\sigma_j(\theta)$ denote the jth elements of $\nu^m(\theta)$ and $\nu^\sigma(\theta)$, respectively, and $\tau_j(\nu(\cdot), \theta, \ell)$ is defined in (21.20). Note that

$$\overrightarrow{A}^{*\inf}_{Ln,EGMS} = \widetilde{g}_n(\nu^*_n(\cdot), \mu_n(\cdot)) \text{ and } A^{*\inf}_{L\infty,EGMS} = \widetilde{g}(G(\cdot), \mu_\infty(\cdot)), \tag{23.46}$$

where $\mu_\infty(\cdot)$ is the constant function that equals $(0'_k, 1'_k)'$ for all $\theta \in \Theta$.

We want to show $\{\widetilde{g}_{a_n}(\nu^*_{a_n}(\cdot), \mu_{a_n}(\cdot)) | \{W_{ni}\}_{i \leq n, n \geq 1}\} \rightarrow_d \widetilde{g}(G(\cdot), \mu_\infty(\cdot))$ a.s.$[P_{\bigtriangledown}]$, where $\{\nu^*_{a_n}(\cdot) | \{W_{ni}\}_{i \leq n, n \geq 1}\} \Rightarrow G(\cdot)$ a.s.$[P_{\bigtriangledown}]$ by Assumption BC.6 and $\sup_{\theta \in \Theta} ||\mu_{a_n}(\theta) - \mu_\infty(\theta)|| = o(1)$ a.s.$[P_{\bigtriangledown}]$ by (23.40) and the definition of $\mu_n(\theta)$ following (23.42). We use the extended CMT to establish this result. For notational simplicity, we employ n, rather than a_n, in the proof of this result. The extended CMT requires showing that for any deterministic sequences $\{\nu_n(\cdot) \in \mathcal{D}\}_{n \geq 1}$ and $\{\mu_n(\cdot) : \Theta \rightarrow R^{2k}\}_{n \geq 1}$ and deterministic $\nu(\cdot) \in \mathcal{D}_0$ such that $\sup_{\theta \in \Theta} ||\nu_n(\theta) - \nu(\theta)|| \rightarrow 0$ and $\sup_{\theta \in \Theta} ||\mu_n(\theta) - \mu_\infty(\theta)|| \rightarrow 0$, we have $\widetilde{g}_n(\nu_n(\cdot), \mu_n(\cdot)) \rightarrow \widetilde{g}(\nu(\cdot), \mu_\infty(\cdot))$. (For notational simplicity, we abuse notation here and consider a deterministic $\nu_n(\cdot)$ that differs from the random $\nu_n(\cdot)$ in Assumption C.5.) Once we have shown this, the proof of part (b) is complete.

The proof of $\widetilde{g}_n(\nu_n(\cdot), \mu_n(\cdot)) \rightarrow \widetilde{g}(\nu(\cdot), \mu_\infty(\cdot))$ is an extension of the proof of $g_n(\nu_n(\cdot)) \rightarrow g(\nu(\cdot))$

in (21.22)–(21.28) in the proof of Theorem 5.3(b). We show

$$\begin{aligned}&\text{(i) } \liminf_{n\to\infty} \widetilde{g}_n(\nu_n(\cdot),\mu_n(\cdot)) \ge g(\nu(\cdot),\mu_\infty(\cdot)) \text{ and}\\ &\text{(ii) } \limsup_{n\to\infty} \widetilde{g}_n(\nu_n(\cdot),\mu_n(\cdot)) \le g(\nu(\cdot),\mu_\infty(\cdot)). \qquad (23.47)\end{aligned}$$

First, we establish (i) in (23.47). There exists a subsequence $\{c_n\}_{n\ge1}$ of $\{n\}_{n\ge1}$ and a sequence $\{(\overline{\theta}_{c_n},\overline{b}_{c_n},\overline{b}^*_{c_n},\overline{\ell}_{c_n},\overline{j}^*_{c_n})\in\Lambda^{*\eta_{c_n}}_{c_n,F_{c_n}}\}_{n\ge1}$ such that

$$\begin{aligned}\liminf_{n\to\infty}\widetilde{g}_n(\nu_n(\cdot),\mu_n(\cdot)) &= \lim_{n\to\infty}\widetilde{g}_{c_n}(\nu_{c_n}(\cdot),\mu_{c_n}(\cdot)) \text{ and}\\ \lim_{n\to\infty}\widetilde{g}_{c_n}(\nu_{c_n}(\cdot),\mu_{c_n}(\cdot)) &= \lim_{n\to\infty}\max_{j\le k}\Big(\tau_j(\nu_{c_n}(\cdot),\overline{\theta}_{c_n},\overline{\ell}_{c_n}) + 1(j\ne\overline{j}^*_{c_n})\overline{b}_{c_nj} \qquad (23.48)\\ &\quad + 1(j=\overline{j}^*_{c_n})\varphi^*_{\overline{j}^*_{c_n}}(\mu_{1\overline{j}^*_{c_n}}(\overline{\theta}_{c_n})+\mu_{2\overline{j}^*_{c_n}}(\overline{\theta}_{c_n})\overline{b}^*_{c_n\overline{j}^*_{c_n}})\Big),\end{aligned}$$

where $\overline{b}_{c_nj}$, $\overline{b}^*_{c_nj}$, and $\overline{\ell}_{c_nj}$ denote the jth elements of $\overline{b}_{c_n}$, $\overline{b}^*_{c_n}$, and $\overline{\ell}_{c_n}$, respectively. Also, there exists a subsequence $\{q_n\}_{n\ge1}$ of $\{c_n\}_{n\ge1}$ and $(\overline{\theta},\overline{b},\overline{b}^*,\overline{\ell},\overline{j}^*)\in\Theta\times R^{3k}_{[\pm\infty]}\times\{1,...,k\}$ such that

$$d\left((\overline{\theta}_{q_n},\overline{b}_{q_n},\overline{b}^*_{q_n},\overline{\ell}_{q_n},\overline{j}^*_{q_n}),(\overline{\theta},\overline{b},\overline{b}^*,\overline{\ell},\overline{j}^*)\right)\to0, \qquad (23.49)$$

where d is defined following (5.2), by compactness of the metric space $(\Theta\times R^{3k}_{[\pm\infty]}\times\{1,...,k\},d)$ under Assumption A.0(i). We have $(\overline{\theta},\overline{b},\overline{b}^*,\overline{\ell},\overline{j}^*)\in\Lambda^*_I$ by the same argument as used to show $(\widetilde{\theta},\widetilde{b},\widetilde{\ell})\in\Lambda$ in (20.7) (but without the requirement that $\overline{\theta}_{q_n}\in\Theta_I(F_{q_n})\ \forall n\ge1$) using (23.49) and Assumption BC.3.

For all $j\le k$,

$$\lim_{n\to\infty}\tau_j(\nu_{q_n}(\cdot),\overline{\theta}_{q_n},\overline{\ell}_{q_n})=\tau_j(\nu(\cdot),\overline{\theta},\overline{\ell})\in R \qquad (23.50)$$

by (21.25) using $\nu_{q_n}(\theta)\to\nu(\theta)$ uniformly over $\theta\in\Theta$ (by assumption) and (23.49).

In addition, we have, for all $j\le k$,

$$\begin{aligned}&1(j\ne\overline{j}^*_{q_n})\overline{b}_{q_nj}\to1(j\ne\overline{j}^*)\overline{b}_j \text{ and}\\ &1(j=\overline{j}^*_{q_n})\varphi^*_{\overline{j}^*_{q_n}}(\mu_{1\overline{j}^*_{q_n}}(\overline{\theta}_{q_n})+\mu_{2\overline{j}^*_{q_n}}(\overline{\theta}_{q_n})\overline{b}^*_{q_n\overline{j}^*_{q_n}})\to1(j=\overline{j}^*)\varphi^*_{\overline{j}^*}(\overline{b}^*_{\overline{j}^*}), \qquad (23.51)\end{aligned}$$

where the first line holds by (23.49) and the second line holds by (23.49), $\sup_{\theta\in\Theta}||\mu_{q_n}(\theta)-\mu_\infty(\theta)||\to0$, and the continuity of $\varphi^*_{\overline{j}}(\cdot)$ on $R^k_{[\pm\infty]}$ under d by Assumption A.5(ii), and the fact that $d(\varphi^*_{\overline{j}^*_{q_n}}(\mu_{1\overline{j}^*_{q_n}}(\overline{\theta}_{q_n})+\mu_{2\overline{j}^*_{q_n}}(\overline{\theta}_{q_n})\overline{b}^*_{q_n\overline{j}^*_{q_n}}),\varphi^*_{\overline{j}^*}(\overline{b}^*_{\overline{j}^*}))\to0$ implies that $\varphi^*_{\overline{j}^*_{q_n}}(\mu_{1\overline{j}^*_{q_n}}(\overline{\theta}_{q_n})+\mu_{2\overline{j}^*_{q_n}}(\overline{\theta}_{q_n})\overline{b}^*_{q_n\overline{j}^*_{q_n}})\to\varphi^*_{\overline{j}^*}(\overline{b}^*_{\overline{j}^*})$ (as a sequence of numbers in $R_{[+\infty]}$) even if $\varphi^*_{\overline{j}^*}(\overline{b}^*_{\overline{j}^*})=+\infty$.

Now, we have

$$
\begin{aligned}
&\liminf_{n\to\infty} \widetilde{g}_n(\nu_n(\cdot),\mu_n(\cdot))\\
=& \lim_{n\to\infty}\max_{j\le k}\left(\tau_j(\nu_{q_n}(\cdot),\overline{\theta}_{q_n},\overline{\ell}_{q_n}) + 1(j\neq \overline{j}^*_{q_n})\overline{b}_{q_nj} + 1(j=\overline{j}^*_{q_n})\varphi^*_{\overline{j}^*_{q_n}}(\mu_{1\overline{j}^*_{q_n}}(\overline{\theta}_{q_n}) + \mu_{2\overline{j}^*_{q_n}}(\overline{\theta}_{q_n})\overline{b}^*_{q_n\overline{j}^*_{q_n}})\right)\\
=& \max_{j\le k}\left(\tau_j(\nu(\cdot),\overline{\theta},\overline{\ell}) + 1(j\neq\overline{j}^*)\overline{b}_j + 1(j=\overline{j}^*)\varphi^*_{\overline{j}^*}(\overline{b}^*_{\overline{j}^*})\right)\\
\ge& \inf_{(\theta,b,b^*,\ell,j^*)\in\Lambda^*_I}\max_{j\le k}\left(\tau_j(\nu(\cdot),\theta,\ell)+1(j\neq j^*)b_j + 1(j=j^*)\varphi^*_{j^*}(b^*_{j^*})\right)\\
:=& \widetilde{g}(\nu(\cdot),\mu_\infty(\cdot)),
\end{aligned}\tag{23.52}
$$

where the first equality holds by (23.48) and the fact that $\{q_n\}_{n\ge1}$ is a subsequence of $\{c_n\}_{n\ge1}$, the second equality holds by (23.50) (using the notational convention that $\nu + c = c$ when $\nu\in R$ and $c = \pm\infty$ if $\overline{b}_j = \pm\infty$ for any $j\le k$) and (23.51), the inequality holds because $(\overline{\theta},\overline{b},\overline{b}^*,\overline{\ell},\overline{j}^*)\in\Lambda^*_I$ by the paragraph containing (23.49), and the last equality holds by the definition of $\widetilde{g}(\nu(\cdot),\mu(\cdot))$ in (23.45) with $\mu(\cdot)=\mu_\infty(\cdot)$. This establishes result (i) in (23.47).

Next, we establish result (ii) in (23.47). There exists $(\theta^\dagger, b^\dagger, b^{\dagger *}, \ell^\dagger, j^{\dagger *})\in\Lambda^*_I$ such that

$$
\widetilde{g}(\nu(\cdot),\mu_\infty(\cdot)) = \max_{j\le k}\left(\tau_j(\nu(\cdot),\theta^\dagger,\ell^\dagger) + 1(j\neq j^{\dagger *})b^\dagger_j + 1(j=j^{\dagger *})\varphi^*_{j^{\dagger *}}(b^{\dagger *}_{j^{\dagger *}})\right)\tag{23.53}
$$

because Λ^*_I is compact under the metric d defined following (5.2) with $a_* = d_\theta + 3k + 1$ (since it is assumed to be an element of $\mathcal{S}(\Theta\times R^{3k}_{[\pm\infty]}\times\{1,...,k\})$) and $\tau_j(\nu(\cdot),\theta,\ell) + 1(j\neq j^*)b_j + 1(j = j^*)\varphi^*_{j^*}(b^*_{j^*})$ is a continuous function of (θ,b,b^*,ℓ,j^*) under d that takes values in the extended real line using Assumption A.5(ii). By Assumption BC.3, $\Lambda^{*\eta_n}_{n,F_n}\to_H \Lambda^*_I$. Hence, there is a sequence $\{(\theta^\dagger_n, b^\dagger_n, b^{\dagger *}_n, \ell^\dagger_n, j^{\dagger *}_n)\in\Lambda^{*\eta_n}_{n,F_n}\}_{n\ge1}$ such that $d((\theta^\dagger_n, b^\dagger_n, b^{\dagger *}_n, \ell^\dagger_n, j^{\dagger *}_n),(\theta^\dagger, b^\dagger, b^{\dagger *}, \ell^\dagger, j^{\dagger *}))\to 0$. We obtain

$$
\begin{aligned}
&\limsup_{n\to\infty}\widetilde{g}_n(\nu_n(\cdot),\mu_n(\cdot))\\
:=& \limsup_{n\to\infty}\inf_{(\theta,b,b^*,\ell,j^*)\in\Lambda^{*\eta_n}_{n,F_n}}\max_{j\le k}(\tau_j(\nu_n(\cdot),\theta,\ell) + 1(j\neq j^*)b_j\\
&\qquad +1(j=j^*)\varphi^*_{j^*}(\mu_{1nj^*}(\theta) + \mu_{2nj^*}(\theta)b^*_{j^*}))\\
\le& \limsup_{n\to\infty}\max_{j\le k}\left(\tau_j(\nu_n(\cdot),\theta^\dagger_n,\ell^\dagger_n) + 1(j\neq j^{\dagger *}_n)b^\dagger_{nj} + 1(j=j^{\dagger *}_n)\varphi^*_{j^{\dagger *}_n}(\mu_{1nj^{\dagger *}_n}(\theta) + \mu_{2nj^{\dagger *}_n}(\theta)b^{\dagger *}_{j^{\dagger *}_n})\right)\\
=& \max_{j\le k}\left(\tau_j(\nu(\cdot),\theta^\dagger,\ell^\dagger) + 1(j\neq j^{\dagger *})b^\dagger_j + 1(j=j^{\dagger *})\varphi^*_{j^{\dagger *}}(b^{\dagger *}_{j^{\dagger *}})\right)\\
=& \widetilde{g}(\nu(\cdot),\mu_\infty(\cdot)),
\end{aligned}\tag{23.54}
$$

where the inequality holds because $(\theta^\dagger_n, b^\dagger_n, b^{\dagger *}_n, \ell^\dagger_n, j^{\dagger *}_n)\in\Lambda^{*\eta_n}_{n,F_n}$ $\forall n\ge1$, the second equality holds using $d((\theta^\dagger_n, b^\dagger_n, b^{\dagger *}_n, \ell^\dagger_n, j^{\dagger *}_n),(\theta^\dagger, b^\dagger, b^{\dagger *}, \ell^\dagger, j^{\dagger *}))\to 0$, (23.50) with $(\nu_n(\cdot),\theta^\dagger_n,\ell^\dagger_n)$ and $(\nu(\cdot),\theta^\dagger,\ell^\dagger)$ in place of $(\nu_{q_n}(\cdot),\overline{\theta}_{q_n},\overline{\ell}_{q_n})$ and $(\nu(\cdot),\overline{\theta},\overline{\ell})$, respectively, and (23.51) with $(\theta^\dagger_{nj}, b^\dagger_{nj}, b^{\dagger *}_{nj}, \ell^\dagger_{nj}, j^{\dagger *}_n)$ and

$(\theta^\dagger_j, b^\dagger_j, b^{\dagger *}_j, \ell^\dagger_j, j^{\dagger *})$ in place of $(\overline{\theta}_{q_nj}, \overline{b}_{q_nj}, \overline{b}^*_{q_nj}, \overline{\ell}_{q_nj}, \overline{j}^*_{q_n})$ and $(\overline{\theta}_j, \overline{b}_j, \overline{b}^*_j, \overline{\ell}_j, \overline{j}^*)$, respectively, and the last equality holds by (23.53). This establishes result (ii) in (23.47) and completes the proof of part (b).

For notational simplicity, we let the subsequence $\{a_n\}_{n\geq 1}$ of $\{p_n\}_{n\geq 1}$ differ in the proofs of parts (a) and (b). However, by taking successive subsequences across the proofs of parts (a) and (b), we can obtain a single subsequence $\{a_n\}_{n\geq 1}$ of $\{p_n\}_{n\geq 1}$ for which both parts (a) and (b) (and part (d)) hold, as stated in the theorem.

The convergence result of part (c) follows from parts (a) and (b), $\widehat{\Omega}_n(\theta_n) \to_p \Omega_\infty$ (by Assumption C.6), the continuity of $S(m, \Omega)$ by Assumption S.1(iii), and the continuous mapping theorem. We have $S^*_{L\infty,EGMS} \geq 0$ a.s. by Assumption S.1(ii). The function $S(m, \Omega)$ can be arbitrarily large only if m_j is arbitrarily small (i.e., m_j is negative and arbitrarily large in absolute value) for some $j \leq k$, by Assumption S.1(i). We have $T^*_{Lj\infty,EGMS}$ and $A^{*\inf}_{L\infty,EGMS}$ (defined in (12.5)) are in R a.s. by Assumptions A.5, C.4, and C.5, and $\chi(G^{m\sigma}_j(\theta), \ell_j) \geq -|G^{m\sigma}_j(\theta)|$ (because $\chi(\nu, c) \geq -|\nu|$ by (21.7). This yields $S^*_{L\infty,EGMS} < \infty$ a.s., which completes the proof of part (c).

Lastly, we prove part (d) of the theorem. The random variables $T^*_{Unj,EGMS}(\theta)$ and $T^*_{Uj\infty,EGMS}$ (defined in (12.6) and (12.7)) are the same as $T^*_{Lnj,EGMS}(\theta)$ and $T^*_{Lj\infty,EGMS}$ (defined in (12.4) and (12.5)), respectively, except the former are defined using φ^{**}_j, which satisfies Assumption A.10, whereas the latter are defined using φ^*_j, which satisfies Assumption A.5. In consequence, the proof of part (a) also applies with U in place of L.

Next, we consider the U version of part (b) that is stated in part (d). By definition, see (12.4) and (12.6), we have

$$\begin{aligned} A^{*\inf}_{Un,EGMS} &:= \inf_{\theta \in \Theta^{\eta_{Un}}_I(F_n)} \min_{j \leq k} \left(-[\widehat{\nu}^*_{nj}(\theta)]_+ + \varphi^{**}_j(\xi^b_{nj}(\theta)) \right), \text{ whereas} \\ A^{*\inf}_{Ln,EGMS} &:= \inf_{\theta \in \Theta^{\eta_n}_I(F_n)} \max_{j \leq k} \Big(\chi(\widehat{\nu}^*_{nj}(\theta), n^{1/2} E_{F_n} \widetilde{m}_j(W, \theta)) + 1(j \neq j_n(\theta)) b_{nj}(\theta) \\ &\qquad\qquad + 1(j = j_n(\theta)) \varphi^*_j(\xi^b_{nj}(\theta)) \Big). \end{aligned} \tag{23.55}$$

In consequence, analogously to (23.39), we can write $A^{*\inf}_{Un,EGMS}$ as

$$A^{*\inf}_{Un,EGMS} = \inf_{(\theta,b,b^*) \in \Lambda^{*\eta_{Un}}_{Un,F_n}} \min_{j \leq k} \left(-[\widehat{\nu}^*_{nj}(\theta)]_+ + \varphi^{**}_j \left((sd^*_{3nj}(\theta)\kappa_n)^{-1} \widehat{d}_{nj}(\theta) + \frac{sd_{3j\infty}(\theta)}{sd^*_{3nj}(\theta)} b^*_j \right) \right), \tag{23.56}$$

where $\Lambda^{*\eta_{Un}}_{Un,F_n}$ is defined just below (12.3) with $\eta = \eta_{Un}$.

By the same arguments as in (23.39)–(23.43), using Assumption A.10 in place of A.5, we obtain

$$\begin{aligned} A^{*\inf}_{Ua_n,EGMS} &= \overrightarrow{A}^{*\inf}_{Ua_n,EGMS} + o(1) \text{ a.s.}[P_{\bigtriangledown}], \text{ where} \\ \overrightarrow{A}^{*\inf}_{Un,EGMS} &:= \inf_{(\theta,b,b^*)\in\Lambda^{*\eta_{Un}}_{Un,F_n}} \min_{j\le k} \left(-[\nu^{m\sigma*}_{nj}(\theta)]_+ + \varphi^{**}_j(\mu_{1nj}(\theta)+\mu_{2nj}(\theta)b^*_j)\right). \end{aligned} \tag{23.57}$$

In place of the definitions in (23.45), for nonstochastic functions $\nu(\cdot)\in\mathcal{D}$ and $\mu(\cdot):\Theta\to R^{2k}$ with $\mu(\theta)=(\mu_{11}(\theta),...,\mu_{1k}(\theta),\mu_{21}(\theta),...,\mu_{2k}(\theta))'$, we now define

$$\begin{aligned} g_{Un}(\nu(\cdot),\mu(\cdot)) &:= \inf_{(\theta,b,b^*)\in\Lambda^{*\eta_{Un}}_{Un,F_n}} \min_{j\le k} \left(-[\nu^{m\sigma}_j(\theta)]_+ + \varphi^{**}_j(\mu_{1j}(\theta)+\mu_{2j}(\theta)b^*_j)\right) \text{ and} \\ g_U(\nu(\cdot),\mu(\cdot)) &:= \inf_{(\theta,b,b^*)\in\Lambda^{*}_{U,I}} \min_{j\le k} \left(-[\nu^{m\sigma}_j(\theta)]_+ + \varphi^{**}_j(\mu_{1j}(\theta)+\mu_{2j}(\theta)b^*_j)\right), \text{ where} \\ \nu^{m\sigma}_j(\theta) &:= \nu^m_j(\theta) - \frac{1}{2}\widetilde{m}_j(\theta)\nu^\sigma_j(\theta). \end{aligned} \tag{23.58}$$

The remainder of the proof of the U version of part (b) goes through as in the proof of the L version given above, using Assumptions A.10 and BC.4 in place of A.5 and BC.3.

The U version of part (c) that is stated in part (d) goes through as in the proof of the L version above, using Assumption A.10 in place of A.5. This completes the proof of part (d). □

23.3 Proof of Lemma 22.4

The proof of Lemma 22.4 uses the following lemma. The set $\Theta^\eta_I(F)$ for a positive constant η is defined in (5.4) by $\Theta^\eta_I(F) := \{\theta\in\Theta : \max_{j\le k}[E_F\widetilde{m}_j(W,\theta) + r^{\inf}_F]_- \le \eta/n^{1/2}\}$. The set $\widehat{\Theta}_n$ is defined in (BC.5) by $\widehat{\Theta}_n := \{\theta\in\Theta : \max_{j\le k}[\widehat{m}_{nj}(\theta) + \widehat{r}^{\inf}_n]_- \le \tau_n/n^{1/2}\}$.

Lemma 23.1 *Suppose that under* $\{F_n\}_{n\ge1}$ *and* $\{\theta_n\}_{n\ge1}$*, Assumptions* A.0, C.4, C.5, *and* C.7 *are satisfied.*

(a) *Let* $\{\eta_n\}_{n\ge1}$ *and* $\{\tau_n\}_{n\ge1}$ *be any sequences of positive constants that satisfy* $\tau_n\to\infty$ *and* $\eta_n/\tau_n\to0$*. Then,*

$$P_{F_n}(\widehat{\Theta}_n \supseteq \Theta^{\eta_n}_I(F_n)) \to 1.$$

(b) *Let* $\{\eta_{Un}\}_{n\ge1}$ *and* $\{\tau_n\}_{n\ge1}$ *be any sequences of positive constants that satisfy* $\eta_{Un}\to\infty$ *and* $\tau_n/\eta_{Un}\to0$*. Then,*

$$P_{F_n}(\Theta^{\eta_{Un}}_I(F_n) \supseteq \widehat{\Theta}_n) \to 1.$$

Proof of Lemma 22.4. For notational simplicity, we replace $\{p_n\}_{n\ge1}$ by $\{n\}_{n\ge1}$ throughout the proof of this lemma. Part (c) follows from parts (a) and (b) using the definitions of $S^*_{Ln,EGMS}(\theta_n)$

and $S^*_{n,EGMS}(\theta_n)$ in (12.4) and (6.3), respectively, and Assumption S.1(i), which requires that $S(m,\Omega)$ is nonincreasing in $m \in R^k$ $\forall (m,\Omega) \in R^k_{[+\infty]} \times \Psi$.

To prove part (a), note that $T^*_{Lnj,EGMS}(\theta)$ and $T^*_{nj,EGMS}(\theta)$ only differ because the former depends on $\varphi^*_j(\xi_{nj}(\theta))$, whereas the latter depends on $\varphi_j(\xi_n(\theta),\widehat{\Omega}_n(\theta))$. By Assumption A.5(i), $\varphi^*_j(\xi_j) \geq \varphi_j(\xi,\Omega)$ $\forall j \leq k$, $\forall (\xi,\Omega) \in R^k_{[+\infty]} \times \Psi$. This gives $\varphi^*_j(\xi_{nj}(\theta)) \geq \varphi_j(\xi_n(\theta),\widehat{\Omega}_n(\theta))$ for all sample and bootstrap realizations. Hence, $T^*_{Lnj,EGMS}(\theta_n) \geq T^*_{nj,EGMS}(\theta_n)$ for all sample and bootstrap realizations, $\forall j \leq k$, $\forall n \geq 1$, and part (a) holds.

Next, we prove part (b). By definition, see (12.4) and (6.11), we have

$$\begin{aligned} A^{*\inf}_{Ln,EGMS} &:= \inf_{\theta \in \Theta_I^{\eta_n}(F_n)} \max_{j \leq k} \Big(\chi(\widehat{\nu}^*_{nj}(\theta), n^{1/2} E_{F_n} \widetilde{m}_j(W,\theta)) + 1(j \neq j_n(\theta)) b_{nj}(\theta) \\ &\qquad\qquad + 1(j = j_n(\theta)) \varphi^*_j(\xi^b_{nj}(\theta)) \Big) \text{ and} \\ A^{*\inf}_{n,EGMS} &:= \inf_{\theta \in \widehat{\Theta}_n} \min_{j_1 \in \widehat{J}_n(\theta)} \max_{j \leq k} \Big(\widehat{\chi}^*_{nj,EGMS}(\theta) + 1(j \neq j_1) \widehat{b}_{nj,EGMS}(\theta) \\ &\qquad\qquad + 1(j = j_1) \varphi_j(\xi^b_n(\theta), \widehat{\Omega}_n(\theta)) \Big). \end{aligned} \tag{23.59}$$

The bootstrap random variables $A^{*\inf}_{Ln,EGMS}$ and $A^{*\inf}_{n,EGMS}$ differ in five ways. Specifically, $A^{*\inf}_{Ln,EGMS}$ versus (vs.) $A^{*\inf}_{n,EGMS}$ are defined with (i) $\inf_{\theta \in \Theta_I^{\eta_n}(F_n)}$ vs. $\inf_{\theta \in \widehat{\Theta}_n}$, (ii) $\varphi^*_j(\xi^b_{nj}(\theta))$ vs. $\varphi_j(\xi^b_n(\theta), \widehat{\Omega}_n(\theta))$, (iii) $b_{nj}(\theta)$ vs. $\widehat{b}_{nj,EGMS}(\theta)$, (iv) $\chi(\widehat{\nu}^*_{nj}(\theta), n^{1/2} E_{F_n} \widetilde{m}_j(W,\theta))$ vs. $\widehat{\chi}^*_{nj,EGMS}(\theta)$, and (v) $j = j_n(\theta)$ or $j \neq j_n(\theta)$ vs. $\min_{j_1 \in \widehat{J}_n(\theta)}$ with $j = j_1$ or $j \neq j_1$.

Lemma 23.1(a) applies because Lemma 22.4 imposes Assumptions A.0, C.4, C.5, and C.7, $\tau_n \to \infty$ by Assumptions A.6(ii), and $\eta_n/\tau_n \to 0$ by Assumption BC.3. By Lemma 23.1(a), for any bootstrap random function $K^*_n(\theta)$,

$$P_\nabla \left(\inf_{\theta \in \Theta_I^{\eta_n}(F_n)} K^*_n(\theta) \geq \inf_{\theta \in \widehat{\Theta}_n} K^*_n(\theta) \,\middle|\, \{W_{ni}\}_{i \leq n, n \geq 1} \right) = 1 \text{ wp} \to 1. \tag{23.60}$$

By Assumption A.5(i), we have

$$\varphi^*_j(\xi^b_n(\theta)) \geq \varphi_j(\xi^b_n(\theta), \widehat{\Omega}_n(\theta)) \ \forall \theta \in \Theta \tag{23.61}$$

for all sample and bootstrap realizations.

We have

$$\begin{aligned} &\widehat{b}_{nj,EGMS}(\theta) := n^{1/2} \left([\widehat{m}_{nj}(\theta)]_- - \widehat{r}^{\inf}_n \right) - s d^*_{3nj}(\theta) \kappa_n = \widehat{d}_{nj}(\theta) + b_{nj}(\theta) - s d^*_{3nj}(\theta) \kappa_n, \text{ and so,} \\ &\sup_{\theta \in \Theta} \left(\widehat{b}_{nj,EGMS}(\theta) - b_{nj}(\theta) \right) \leq \sup_{\theta \in \Theta} \left(\widehat{d}_{nj}(\theta) - \kappa_n \right) \to_p -\infty, \end{aligned} \tag{23.62}$$

where the first equality in the first line holds by definition, see (6.8), the second equality holds by (23.37), and the second line follows from the first line, the last line of (23.37), $sd^*_{3nj}(\theta) \geq 1$ by definition, and $\kappa_n \to \infty$ (by Assumption A.6(i)) and the inequality on the second line holds for all bootstrap realizations because $\widehat{d}_{nj}(\theta)$ does not depend on any bootstrap quantities. Equation (23.62) implies that

$$\sup_{\theta\in\Theta}(\widehat{b}_{nj,EGMS}(\theta) - b_{nj}(\theta)) \leq 0 \ \forall j \leq k, \text{ for all bootstrap realizations, wp} \to 1. \tag{23.63}$$

Now, we show

$$P_\nabla\left(\chi(\widehat{\nu}^*_{nj}(\theta), n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta)) \geq \widehat{\chi}^*_{nj,EGMS}(\theta) \ \forall\theta\in\Theta|\{W_{ni}\}_{i\leq n,n\geq 1}\right) = 1 \text{ wp} \to 1. \tag{23.64}$$

By the footnote following (6.7), $\chi(\nu, c)$ is nondecreasing in c for $\nu > 0$ and nonincreasing in c for $\nu < 0$. Using this and the definition of $\chi(\nu, c_1, c_2)$ in (6.6), we obtain: for $\nu \geq 0$, $\chi(\nu, c_1, c_2) = \chi(\nu, c_1) \leq \chi(\nu, c)$ $\forall c \geq c_1$. And, for $\nu < 0$, $\chi(\nu, c_1, c_2) = \chi(\nu, c_2) \leq \chi(\nu, c)$ $\forall c \leq c_2$. These results yield: for all $\widehat{\nu}^*_{nj}(\theta) \geq 0$,

$$\begin{aligned}
\widehat{\chi}^*_{nj,EGMS}(\theta) &:= \chi\left(\widehat{\nu}^*_{nj}(\theta),\ n^{1/2}\widehat{m}_{nj}(\theta) - sd^*_{2nj}(\theta)\kappa_n,\ n^{1/2}\widehat{m}_{nj}(\theta) + sd^*_{2nj}(\theta)\kappa_n\right) \\
&= \chi\left(\widehat{\nu}^*_{nj}(\theta),\ n^{1/2}\widehat{m}_{nj}(\theta) - sd^*_{2nj}(\theta)\kappa_n\right) \\
&\leq \chi\left(\widehat{\nu}^*_{nj}(\theta),\ n^{1/2}\widehat{m}_{nj}(\theta) - \kappa_n\right) \\
&\leq \chi\left(\widehat{\nu}^*_{nj}(\theta),\ n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta)\right)
\end{aligned} \tag{23.65}$$

provided $n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta) \geq n^{1/2}\widehat{m}_{nj}(\theta) - \kappa_n$, where the first inequality holds because $sd^*_{2nj}(\theta) \geq 1$ and $\chi(v, c)$ is nondecreasing in c for $v \geq 0$, as stated above. Similarly, for $\widehat{\nu}^*_{nj}(\theta) < 0$,

$$\begin{aligned}
\widehat{\chi}^*_{nj,EGMS}(\theta) &= \chi\left(\widehat{\nu}^*_{nj}(\theta),\ n^{1/2}\widehat{m}_{nj}(\theta) + sd^*_{2nj}(\theta)\kappa_n\right) \\
&\leq \chi\left(\widehat{\nu}^*_{nj}(\theta),\ n^{1/2}\widehat{m}_{nj}(\theta)\kappa_n\right) \\
&\leq \chi\left(\widehat{\nu}^*_{nj}(\theta),\ n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta)\right)
\end{aligned} \tag{23.66}$$

provided $n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta) \leq n^{1/2}\widehat{m}_{nj}(\theta) + \kappa_n$.

By (23.36), which uses Assumptions C.4 and C.5, $n^{1/2}\widehat{m}_{nj}(\theta) = n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta) + O^\Theta_p(1)$. Hence,

$$\liminf_{n\to\infty} P_{F_n}\left(n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta) \in \left[n^{1/2}\widehat{m}_{nj}(\theta) - \kappa_n, n^{1/2}\widehat{m}_{nj}(\theta) + \kappa_n\right] \ \forall\theta\in\Theta\right) = 1 \tag{23.67}$$

using $\kappa_n \to \infty$ by Assumption A.6(i). The combination of (23.65)–(23.67) establishes (23.64).

Define

$$\overline{A}^{*\inf}_{Ln,EGMS} := \inf_{\theta \in \Theta_I^{\eta_n}(F_n)} \min_{j_1 \in \widehat{J}_n(\theta)} \max_{j \leq k} \Big(\chi(\widehat{\nu}^*_{nj}(\theta), n^{1/2} E_{F_n} \widetilde{m}_j(W,\theta)) + 1(j \neq j_1) b_{nj}(\theta) + 1(j = j_1) \varphi^*_j(\xi^b_{nj}(\theta)) \Big). \tag{23.68}$$

Combining (23.60)–(23.64) and (23.68) gives

$$P_\nabla(\overline{A}^{*\inf}_{Ln,EGMS} \geq A^{*\inf}_{n,EGMS} | \{W_{ni}\}_{i \leq n, n \geq 1}) = 1 \text{ wp} \to 1. \tag{23.69}$$

Next, we show that

$$P_\nabla(j_n(\theta) \in \widehat{J}_n(\theta)\ \forall \theta \in \Theta | \{W_{ni}\}_{i \leq n, n \geq 1}) = 1 \text{ wp} \to 1, \tag{23.70}$$

where $j_n(\theta) := \arg\max_{j \leq k} b_{nj}(\theta)$ is defined in (12.1) and $\widehat{J}_n(\theta) := \{j \in \{1,...,k\} : \widehat{r}_{nj}(\theta) \geq \widehat{r}_n(\theta) - sd^*_{4nj}(\theta) n^{-1/2} \kappa_n\}$ is defined in (6.10). We have $j_n(\theta) \in \widehat{J}_n(\theta)$ iff $\widehat{r}_{nj_n(\theta)}(\theta) \geq \widehat{r}_n(\theta) - sd^*_{4nj_n(\theta)}(\theta) n^{-1/2} \kappa_n$ if $n^{1/2}(\widehat{r}_{nj_n(\theta)}(\theta) - \widehat{r}^{\inf}_n) - n^{1/2}(\widehat{r}_n(\theta) - \widehat{r}^{\inf}_n) \geq -\kappa_n$ because $sd^*_{4nj_n(\theta)}(\theta) \geq 1$ by definition. By (23.37), $n^{1/2}(\widehat{r}_{nj}(\theta) - \widehat{r}^{\inf}_n) = b_{nj}(\theta) + O^\Theta_p(1)\ \forall j \leq k$ (since $\widehat{r}_{nj}(\theta) = [\widehat{m}_{nj}(\theta)]_-$ by (4.1)). Hence, $n^{1/2}(\max_{j \leq k} \widehat{r}_{nj}(\theta) - \widehat{r}^{\inf}_n) = \max_{j \leq k} b_{nj}(\theta) + O^\Theta_p(1)$. Taking $j = j_n(\theta)$, these results combine to give $n^{1/2}(\widehat{r}_{nj_n(\theta)}(\theta) - \widehat{r}^{\inf}_n) - n^{1/2}(\widehat{r}_n(\theta) - \widehat{r}^{\inf}_n) = b_{nj_n(\theta)}(\theta) - \max_{j \leq k} b_{nj}(\theta) + O^\Theta_p(1) = O^\Theta_p(1)$ using the definition of $j_n(\theta)$, where the $O^\Theta_p(1)$ term does not depend on any bootstrap quantities. Since $O^\Theta_p(1) \geq -\kappa_n$ holds wp→1 using Assumption A.6(i) (i.e., $\kappa_n \to \infty$), (23.70) is proved.

For a suitably defined random function $w(j_1,\theta)$ on $\{1,...,k\} \times \Theta$, $A^{*\inf}_{Ln,EGMS}$ and $\overline{A}^{*\inf}_{Ln,EGMS}$ can be written as $\inf_{\theta \in \Theta_I^{\eta_n}(F_n)} w(j_n(\theta),\theta)$ and $\inf_{\theta \in \Theta_I^{\eta_n}(F_n)} \min_{j_1 \in \widehat{J}_n(\theta)} w(j_1,\theta)$, respectively. Since $w(j_n(\theta),\theta) \geq \min_{j_1 \in \widehat{J}_n(\theta)} w(j_1,\theta)$ when $j_n(\theta) \in \widehat{J}_n(\theta)$ and the latter event satisfies (23.70), we obtain

$$P_\nabla(A^{*\inf}_{Ln,EGMS} \geq \overline{A}^{*\inf}_{Ln,EGMS} | \{W_{ni}\}_{i \leq n, n \geq 1}) = 1 \text{ wp} \to 1. \tag{23.71}$$

This and (23.69) establish the result of part (b) of the lemma.

Now, we prove part (d) of the lemma. The proofs of the U versions of parts (c) and (a) stated in part (d) are the same as the L version proofs given above with the inequalities reversed using Assumption A.10(i) in place of A.5(i).

The proof of the U version of part (b) stated in part (d) is as follows. By definition, see (12.6),

$$\begin{aligned} A^{*\inf}_{Un,EGMS} &:= \inf_{\theta\in\Theta_I^{\eta_{Un}}(F_n)} \min_{j\le k} \left(-[\widehat{\nu}^*_{nj}(\theta)]_+ + \varphi_j^{**}(\xi^b_{nj}(\theta))\right) \\ &= \inf_{\theta\in\Theta_I^{\eta_{Un}}(F_n)} \min_{j_1\le k} \max_{j\le k} \left(-[\widehat{\nu}^*_{nj}(\theta)]_+ + 1(j\neq j_1)(-\infty) + 1(j=j_1)\varphi_j^{**}(\xi^b_{nj}(\theta))\right), \end{aligned} \tag{23.72}$$

where the second equality holds because on the rhs the max over $j \le k$ is attained for $j = j_1$ (since for $j \neq j_1$ the term in parentheses equals $-\infty$).

In contrast, consider $A^{*\inf}_{n,EGMS}$, which is defined in (23.59). The bootstrap random variables $A^{*\inf}_{Un,EGMS}$ and $A^{*\inf}_{n,EGMS}$ differ in five ways. Specifically, $A^{*\inf}_{Un,EGMS}$ vs. $A^{*\inf}_{n,EGMS}$ are defined with (i) $\inf_{\theta\in\Theta_I^{\eta_{Un}}(F_n)}$ vs. $\inf_{\theta\in\widehat{\Theta}_n}$, (ii) $\varphi_j^{**}(\xi^b_{nj}(\theta))$ vs. $\varphi_j(\xi^b_n(\theta),\widehat{\Omega}_n(\theta))$, (iii) $-\infty$ vs. $\widehat{b}_{nj,EGMS}(\theta)$, (iv) $-[\widehat{\nu}^*_{nj}(\theta)]_+$ vs. $\widehat{\chi}^*_{nj,EGMS}(\theta)$, and (v) $\min_{j_1\le k}$ vs. $\min_{j_1\in\widehat{J}_n(\theta)}$. Lemma 23.1(b) applies because Lemma 22.4(d) imposes Assumptions A.0, C.4, C.5, and C.7, $\tau_n \to \infty$ by Assumptions A.6(ii), and $\eta_{Un}/\tau_n \to 0$ by Assumption BC.4. Hence, by Lemma 23.1(b), for any bootstrap random function $K^*_n(\theta)$,

$$P_\nabla\left(\inf_{\theta\in\Theta_I^{\eta_{Un}}(F_n)} K^*_n(\theta) \le \inf_{\theta\in\widehat{\Theta}_n} K^*_n(\theta) \,\middle|\, \{W_{ni}\}_{i\le n,n\ge1}\right) = 1 \text{ wp} \to 1. \tag{23.73}$$

By Assumption A.10(i),

$$\varphi_j^{**}(\xi^b_n(\theta)) \le \varphi_j(\xi^b_n(\theta),\widehat{\Omega}_n(\theta)) \;\forall\theta\in\Theta \tag{23.74}$$

for all sample and bootstrap realizations. Since $\widehat{b}_{nj,EGMS}(\theta) \in R_{[\pm\infty]}$, we have $-\infty \le \widehat{b}_{nj,EGMS}(\theta)$ $\forall\theta\in\Theta$ for all sample and bootstrap realizations.

For all $\widehat{\nu}^*_{nj}(\theta) \ge 0$,

$$\widehat{\chi}^*_{nj,EGMS}(\theta) = \chi(\widehat{\nu}^*_{nj}(\theta), n^{1/2}\widehat{m}_{nj}(\theta) - sd^*_{2nj}(\theta)\kappa_n) \ge \chi(\widehat{\nu}^*_{nj}(\theta),-\infty) = -\widehat{\nu}^*_{nj}(\theta), \tag{23.75}$$

where the first equality holds by the definition of $\widehat{\chi}^*_{nj,EGMS}(\theta)$ in (6.7), the inequality holds because $\chi(\nu,c)$ is nondecreasing in c for $\nu \ge 0$ by the footnote following (6.7), and the last equality uses $\chi(\nu,-\infty) = -\nu$ by (5.6). Similarly, for all $\widehat{\nu}^*_{nj}(\theta) < 0$,

$$\widehat{\chi}^*_{nj,EGMS}(\theta) = \chi(\widehat{\nu}^*_{nj}(\theta), n^{1/2}\widehat{m}_{nj}(\theta) + sd^*_{2nj}(\theta)\kappa_n) \ge \chi(\widehat{\nu}^*_{nj}(\theta),+\infty) = 0, \tag{23.76}$$

where the first equality holds by the definition of $\widehat{\chi}^*_{nj,EGMS}(\theta)$, the inequality holds because $\chi(\nu,c)$ is nonincreasing in c for $\nu < 0$ by the footnote following (6.7), and the last equality uses $\chi(\nu,+\infty) = 0$ by (5.6). Hence, for all sample and bootstrap realizations, $\widehat{\chi}^*_{nj,EGMS}(\theta) \ge -[\widehat{\nu}^*_{nj}(\theta)]_+$.

Because $\widehat{J}_n(\theta) \subset \{1, ..., k\}$, the $\min_{j_1 \leq k}$ is less than or equal to the $\min_{j_1 \in \widehat{J}_n(\theta)}$. In consequence of the results above, we obtain $P_{\nabla}(A^{*\inf}_{Un,EGMS} \leq A^{*\inf}_{n,EGMS} | \{W_{ni}\}_{i \leq n, n \geq 1}) = 1$ wp→1, which establishes the U version of part (b) stated in part (d) of the lemma. □

Proof of Lemma 23.1. First, we prove part (a). We have

$$
\begin{aligned}
P_{F_n}(\widehat{\Theta}_n \supseteq \Theta_I^{\eta_n}(F_n)) &\geq P_{F_n}\left(\sup_{\theta \in \Theta_I^{\eta_n}(F_n)} \max_{j \leq k} n^{1/2}[\widehat{m}_{nj}(\theta) + \widehat{r}_n^{\inf}]_- \leq \tau_n \right) \\
&= P_{F_n}\left(\sup_{\theta \in \Theta_I^{\eta_n}(F_n)} \max_{j \leq k} n^{1/2}([\widehat{m}_{nj}(\theta)]_- - \widehat{r}_n^{\inf}) \leq \tau_n \right), \quad (23.77)
\end{aligned}
$$

where the inequality holds by the definition of $\widehat{\Theta}_n$ and the equality holds because for $b, c \geq 0$, $[a+b]_- \leq c$ if and only if $[a]_- - b \leq c$. To see this, first note that $[a+b]_- \leq c$ and $[a]_- - b \leq c$ are equivalent to $\max\{-a-b-c, -c\} \leq 0$ and $\max\{-a-b-c, -b-c\} \leq 0$, respectively. The "only if" part follows by observing that $\max\{-a-b-c, -c\} \geq \max\{-a-b-c, -b-c\}$. Now, suppose $[a]_- - b \leq c$ so that either (i) $a \geq 0$ or (ii) $a < 0$ and $-a-b \leq c$. If (i) is the case, $[a+b]_- = 0 \leq c$, and if (ii) is the case, $[a+b]_- = \max\{-a-b, 0\} \leq \max\{c, 0\} \leq c$.

We have

$$
\begin{aligned}
& \sup_{\theta \in \Theta_I^{\eta_n}(F_n)} \max_{j \leq k} n^{1/2}([\widehat{m}_{nj}(\theta)]_- - \widehat{r}_n^{\inf}) \\
= & \sup_{\theta \in \Theta_I^{\eta_n}(F_n)} \max_{j \leq k} n^{1/2}([\widehat{m}_{nj}(\theta)]_- - r_{F_n}^{\inf}) + n^{1/2}(r_{F_n}^{\inf} - \widehat{r}_n^{\inf}) \\
= & \sup_{\theta \in \Theta_I^{\eta_n}(F_n)} \max_{j \leq k} n^{1/2}([\widehat{m}_{nj}(\theta)]_- - r_{F_n}^{\inf}) + O_p(1) \\
= & \sup_{\theta \in \Theta_I^{\eta_n}(F_n)} \max_{j \leq k} \Big([\nu_{nj}^{m\sigma}(\theta) + n^{1/2} E_{F_n} \widetilde{m}_j(W, \theta)]_- - [n^{1/2} E_{F_n} \widetilde{m}_j(W, \theta)]_- \\
& \quad + n^{1/2}([E_{F_n} \widetilde{m}_j(W, \theta)]_- - r_{F_n}^{\inf}) \Big) + O_p(1) \\
\leq & \sup_{\theta \in \Theta_I^{\eta_n}(F_n)} \max_{j \leq k} |\nu_{nj}^{m\sigma}(\theta)| + \eta_n + O_p(1) \\
= & \; O_p(1) + \eta_n, \quad (23.78)
\end{aligned}
$$

where the second equality holds by Theorem 5.3(b) (which requires Assumptions A.0, C.4, C.5, and C.7), the third equality holds by (21.5) and (21.6), the inequality holds by the definition of $\Theta_I^{\eta_n}(F_n)$, the same reasoning as given following (23.77), and (21.7), and the last equality holds by Assumption C.5.

It follows that

$$\begin{aligned}
& P_{F_n}\left(\sup_{\theta\in\Theta_I^{\eta_n}(F_n)} \max_{j\leq k} n^{1/2}([\widehat{m}_{nj}(\theta)]_- - \widehat{r}_n^{\inf}) \leq \tau_n\right) \\
\geq\ & P_{F_n}(O_p(1)+\eta_n \leq \tau_n) \\
=\ & P_{F_n}(O_p(1/\tau_n)+\eta_n/\tau_n \leq 1) \\
\to\ & 1, \qquad (23.79)
\end{aligned}$$

where the convergence holds because $\tau_n \to \infty$ and $\eta_n/\tau_n \to 0$. Combining this with (23.77) gives the result of part (a).

Next, we prove part (b). Note that

$$P_{F_n}(\Theta_I^{\eta_{Un}}(F_n) \supseteq \widehat{\Theta}_n) \geq P_{F_n}\left(\sup_{\theta\in\widehat{\Theta}_n} \max_{j\leq k} n^{1/2}([E_{F_n}\widetilde{m}_j(W,\theta)]_- - r_{F_n}^{\inf}) \leq \eta_{Un}\right) \qquad (23.80)$$

by the definition of $\Theta_I^{\eta_{Un}}(F_n)$ and the same reasoning as given following (23.77).

We have

$$\begin{aligned}
& \sup_{\theta\in\widehat{\Theta}_n} \max_{j\leq k} n^{1/2}([E_{F_n}\widetilde{m}_j(W,\theta)]_- - r_{F_n}^{\inf}) \\
=\ & \sup_{\theta\in\widehat{\Theta}_n} \max_{j\leq k} n^{1/2}([E_{F_n}\widetilde{m}_j(W,\theta)]_- - [\widehat{m}_{nj}(\theta)]_- + [\widehat{m}_{nj}(\theta)]_- - \widehat{r}_n^{\inf} + \widehat{r}_n^{\inf} - r_{F_n}^{\inf}) \\
=\ & \sup_{\theta\in\widehat{\Theta}_n} \max_{j\leq k} n^{1/2}([E_{F_n}\widetilde{m}_j(W,\theta)]_- - [\widehat{m}_{nj}(\theta)]_- + [\widehat{m}_{nj}(\theta)]_- - \widehat{r}_n^{\inf}) + O_p(1) \\
\leq\ & \sup_{\theta\in\widehat{\Theta}_n} \max_{j\leq k} n^{1/2}([E_{F_n}\widetilde{m}_j(W,\theta)]_- - [\widehat{m}_{nj}(\theta)]_-) + \tau_n + O_p(1) \\
=\ & \sup_{\theta\in\widehat{\Theta}_n} \max_{j\leq k} \left([n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta)]_- - [\nu_{nj}^{m\sigma}(\theta) + n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta)]_-\right) + \tau_n + O_p(1) \\
\leq\ & \sup_{\theta\in\widehat{\Theta}_n} \max_{j\leq k} \left|\nu_{nj}^{m\sigma}(\theta)\right| + \tau_n + O_p(1) \\
=\ & O_p(1) + \tau_n, \qquad (23.81)
\end{aligned}$$

where the second equality holds by Theorem 5.3(b) (which requires Assumptions A.0, C.4, C.5, and C.7), the first inequality holds by the definition of $\widehat{\Theta}_n$ and the same reasoning as given following (23.77), the third equality holds by (21.5) and (21.6), the second inequality holds by (21.7), and the last equality holds by Assumption C.5.

It follows that

$$
\begin{aligned}
& P_{F_n}\left(\sup_{\theta\in\widehat{\Theta}_n}\max_{j\leq k} n^{1/2}([E_{F_n}\widetilde{m}_j(W,\theta)]_- - r_{F_n}^{\inf}) \leq \eta_{Un}\right) \\
\geq\ & P_{F_n}(O_p(1)+\tau_n \leq \eta_{Un}) \\
=\ & P_{F_n}(O_p(1/\eta_{Un})+\tau_n/\eta_{Un} \leq 1) \\
\to\ & 1, \qquad (23.82)
\end{aligned}
$$

where the convergence holds because $\eta_{Un}\to\infty$ and $\tau_n/\eta_{Un}\to 0$. Combining this with (23.80) gives the result of part (b). $\square$

23.4 Proof of Lemma 22.5

Proof of Lemma 22.5. Given the definitions of $S^*_{L\infty,EGMS}$ and $S_{I\infty}$ in (12.5) and (5.11), respectively, and Assumption S.1(i), it suffices to show that $T^*_{Lj\infty,EGMS}\leq T_{j\infty}$ and $A^{*\inf}_{L\infty,EGMS}\leq A^{\inf}_{\infty}(\Lambda_I)$ for all sample realizations, where $T^*_{Lj\infty,EGMS}$, $T_{j\infty}$, $A^{*\inf}_{L\infty,EGMS}$, $A^{\inf}_{\infty}(\Lambda)$, and Λ_I are defined in (12.5), (5.8), (12.5), (5.10), and Assumption C.8, respectively, using quantities that are defined in Assumptions C.1 and C.3–C.5. We have

$$
T^*_{Lj\infty,EGMS} := G^{m\sigma}_{j\infty}+\varphi^*_j(h^*_{j\infty})\leq G^{m\sigma}_{j\infty}+h_{j\infty} := T_{j\infty} \qquad (23.83)
$$

for all sample realizations, where the inequality holds because (i) $h_{j\infty}\geq 0$ by Lemma 5.1(a) (which imposes Assumptions C.3 and N), (ii) $\varphi^*_j(h^*_{j\infty})\leq h_{j\infty}$ holds immediately if $h_{j\infty}=\infty$, and (iii) if $0\leq h_{j\infty}<\infty$, then $h^*_{j\infty}=0$ (since $n^{1/2}(E_{F_n}\widetilde{m}_j(W,\theta_n)+r^{\inf}_{F_n})\to h_{j\infty}$ and $(sd_{1nj}(\theta_n)\kappa_n)^{-1}$ $\times n^{1/2}(E_{F_n}\widetilde{m}_j(W,\theta_n)+r^{\inf}_{F_n})\to h^*_{j\infty}$ by Assumptions C.3 and BC.2, $sd_{1nj}(\theta_n)\geq 1/2$ for n large, which holds by Assumption BC.1 and $sd^*_{1nj}(\theta_n)\geq 1$, and $\kappa_n\to\infty$), $h^*_{j\infty}=0$ implies $\varphi^*_j(h^*_{j\infty})=0$ by Assumption A.5(iii), and hence, $\varphi^*_j(h^*_{j\infty})\leq h_{j\infty}$.

Now, we show $A^{*\inf}_{L\infty,EGMS}\leq A^{\inf}_{\infty}(\Lambda_I)$. We can write $A^{*\inf}_{L\infty,EGMS}=\inf_{(\theta,b,b^*,\ell,j^*)\in\Lambda^*_I}K_L(\theta,b,b^*,\ell,$ $j^*)$ and $A^{\inf}_{\infty}(\Lambda_I)=\inf_{(\theta,b,\ell)\in\Lambda_I}K(\theta,b,\ell)$ for random functions $K_L(\cdot)$ and $K(\cdot)$ defined in (23.85) below. To show $A^{*\inf}_{L\infty,EGMS}\leq A^{\inf}_{\infty}(\Lambda_I)$, it suffices to show that for any $(\theta,b,\ell)\in\Lambda_I$ there exists $(\theta,b,b^*,\ell,j^*)\in\Lambda^*_I$ for which $K_L(\theta,b,b^*,\ell,j^*)\leq K(\theta,b,\ell)$ for all sample realizations.

To this end, we claim: Given any $(\theta,b,\ell)\in\Lambda_I$, there exists an element $(\theta,b,b^*,\ell,j^*)\in\Lambda^*_I$.

This claim is proved as follows. By Assumption C.8, given any $(\theta,b,\ell)\in\Lambda_I$, there exists a sequence $\{(\overline{\theta}_n,\overline{b}_n,\overline{\ell}_n)\in\Lambda^{\eta_n}_{n,F_n}\}_{n\geq 1}$ such that $d((\overline{\theta}_n,\overline{b}_n,\overline{\ell}_n),(\theta,b,\ell))\to 0$, where $\overline{\theta}_n\in\Theta^{\eta_n}_I(F_n)$ for all $n\geq 1$ by the definition of $\Lambda^{\eta_n}_{n,F_n}$ following (5.4). Given $\{\overline{\theta}_n\}_{n\geq 1}$, consider the corresponding sequence $\{(\overline{\theta}_n,\overline{b}_n,b^*_n,\ell_n,j^*_n)\in\Lambda^{*\eta_n}_{n,F_n}\}_{n\geq 1}$ for $\Lambda^{*\eta_n}_{n,F_n}$ defined in (12.3), where $b^*_{nj}:=(sd_{3j\infty}(\overline{\theta}_n)\kappa_n)^{-1}\overline{b}_{nj}$, $j^*_n:=$

$\arg\max_{j\leq k}\overline{b}_{nj}$, and j_n^* is the smallest arg max value if the arg max is not unique. By Assumption BC.3, $\Lambda_{n,F_n}^{*\eta_n} \rightarrow_H \Lambda_I^*$ for Λ_I^* compact (under d). In consequence, there exist a subsequence $\{u_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$ and an element $(\overline{\theta}, \overline{b}, b^*, \ell, j^*)$ of Λ_I^* for which

$$d((\overline{\theta}_{u_n}, \overline{b}_{u_n}, b_{u_n}^*, \ell_{u_n}, j_{u_n}^*), (\overline{\theta}, \overline{b}, b^*, \ell, j^*)) \rightarrow 0 \text{ and } (\overline{\theta}, \overline{b}) = (\theta, b), \tag{23.84}$$

where the equality holds because $d((\overline{\theta}_n, \overline{b}_n, \overline{\ell}_n), (\theta, b, \ell)) \rightarrow 0$, which completes the proof of the claim.

Given any $(\theta, b, \ell) \in \Lambda_I$, take $(\theta, b, b^*, \ell, j^*) \in \Lambda_I^*$ as in the previous paragraph. Then, we have

$$\begin{aligned} K_L(\theta, b, b^*, \ell, j^*) &:= \max_{j\leq k}\left(\chi(G_j^{m\sigma}(\theta), \ell_j) + 1(j \neq j^*)b_j + 1(j = j^*)\varphi_{j^*}^*(b_{j^*}^*)\right) \\ &\leq \max_{j\leq k}\left[\chi(G_j^{m\alpha}(\theta), \ell_j) + b_j\right] := K(\theta, b, \ell) \end{aligned} \tag{23.85}$$

for all sample realizations, where the first and last equalities hold by the definitions of $A_{L\infty,EGMS}^{*\inf}$ and $A_\infty^{\inf}(\Lambda_I)$ and the inequality holds because, as we show below, $\varphi_{j^*}^*(b_{j^*}^*) \leq b_{j^*}$. As argued above, (23.85) implies that $A_{L\infty,EGMS}^{*\inf} \leq A_\infty^{\inf}(\Lambda_I)$, which we set out to prove.

Now, we show $\varphi_{j^*}^*(b_{j^*}^*) \leq b_{j^*}$. For notational simplicity, suppose (23.84) holds with n in place of u_n. We have $j_n^* \rightarrow j^*$ by (23.84), and hence, $j_n^* = j^*$ for n large (because $j_n^* \in \{1, ..., k\}$), where $j_n^* := j_n(\overline{\theta}_n)$ by the definition of $\Lambda_{n,F_n}^{*\eta_n}$ in (12.3) for $j_n(\overline{\theta}_n)$ defined in (12.1). We have $\overline{b}_{nj} \rightarrow b_j$ and $b_{nj}^* \rightarrow b_j^*$ by (23.84), where $\overline{b}_{nj} = b_{nj}(\overline{\theta}_n)$ and $b_{nj}^* = (sd_{3j\infty}(\overline{\theta}_n)\kappa_n)^{-1}\overline{b}_{nj}$ by the definition of $\Lambda_{n,F_n}^{*\eta_n}$ for $b_{nj}(\theta)$ defined in (12.1). Hence, we have $\overline{b}_{nj_n^*} \rightarrow b_{j^*}$ and $b_{nj_n^*}^* \rightarrow b_{j^*}^*$, where $b_{nj_n^*}^* = (sd_{3j_n^*\infty}(\overline{\theta}_n)\kappa_n)^{-1}\overline{b}_{nj_n^*} = (sd_{3j_n(\overline{\theta}_n)\infty}(\overline{\theta}_n)\kappa_n)^{-1}b_{nj_n(\overline{\theta}_n)}(\overline{\theta}_n) \geq 0$ for all $n \geq 1$ by (12.2). This, $sd_{3j_n(\overline{\theta}_n)\infty}(\overline{\theta}_n) \geq 1/2$ for n large (by Assumption BC.1 and $sd_{3nj_n(\overline{\theta}_n)}^*(\theta) \geq 1$, which holds by its definition following (6.8)), and $\kappa_n \rightarrow \infty$ (by Assumption A.6(i)) imply that $b_{j^*} \geq b_{j^*}^* \geq 0$. In addition, it implies that if $0 \leq b_{j^*} < \infty$, then $b_{j^*}^* = 0$ (since $\kappa_n \rightarrow \infty$). Hence, we obtain: if $0 \leq b_{j^*} < \infty$, then $\varphi_{j^*}^*(b_{j^*}^*) = 0 \leq b_{j^*}$ because $\varphi_{j^*}^*(0) = 0$ by Assumption A.5(iii). On the other hand, if $b_{j^*} = \infty$, then $\varphi_{j^*}^*(b_{j^*}^*) \leq \infty = b_{j^*}^*$ by the definition of $\varphi_j^*(\cdot)$, which completes the proof of the lemma. □

24 Proof of Theorem 7.1

The proof of Theorem 7.1 uses the following lemma, which provides sufficient conditions for Assumptions C.5 and C.6 to hold for the case of i.i.d. observations.

Let $\rightarrow_u$ denote uniform convergence over Θ^2.

We assume the covariance kernel converges uniformly.

Assumption C.9. $\Omega_{F_n}(\cdot,\cdot) \to_u \Omega_\infty(\cdot,\cdot)$ for some continuous $R^{2k\times 2k}$-valued function $\Omega_\infty(\cdot,\cdot)$ on Θ^2.

The following Lemma is based on Lemma D.2 of BCS.

Lemma 24.1 *Assumptions* A.0–A.4, C.1, *and* C.9 *imply Assumptions* C.5 *and* C.6 *with the covariance kernel of* $G(\cdot)$ *in Assumption* C.5 *equal to* $\Omega_\infty(\cdot,\cdot)$ *and with* Ω_∞ *in Assumption* C.6 *equal to the upper left* $k \times k$ *submatrix of* $\Omega_\infty(\theta_\infty,\theta_\infty)$.

Comment. For any subsequence $\{q_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$, Lemma 24.1 holds with q_n in place of n throughout, including the assumptions. (The proof just needs to be changed by replacing n by q_n throughout.)

Proof of Theorem 7.1. First, we prove the result of part (b) for the $CS_{n,SPUR1}$ CS. Let $\phi_n(\theta)$ abbreviate $\phi_{n,SPUR1}(\theta)$. There always exist sequences $\{F_n\}_{n\geq 1}$ and $\{\theta_n \in \Theta_I(F_n)\}_{n\geq 1}$ and a subsequence $\{q_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$ such that

$$\liminf_{n\to\infty} \inf_{F\in\mathcal{P}} \inf_{\theta\in\Theta_I(F)} P_F(\phi_n(\theta)=0) = \liminf_{n\to\infty} P_{F_n}(\phi_n(\theta_n)=0) = \lim P_{F_{q_n}}(\phi_{q_n}(\theta_{q_n})=0). \quad (24.1)$$

The left-hand side expression equals the uniform coverage probability in Theorem 7.1(b) using the definition of the SPUR1 CS in (4.5). By (24.1), it suffices to show that the rhs of (24.1) is $1-\alpha$ or greater with $\{q_n\}_{n\geq 1}$ replaced by some subsequence $\{a_n\}_{n\geq 1}$ of $\{q_n\}_{n\geq 1}$ (because the limit under the subsequence $\{a_n\}_{n\geq 1}$ is the same as the limit under the original subsequence $\{q_n\}_{n\geq 1}$). The rhs of (24.1) defined with $\{a_n\}_{n\geq 1}$ is $1-\alpha$ or greater by Theorem 22.1(c) provided the assumptions of Theorem 22.1(c) hold for some subsequence $\{p_n\}_{n\geq 1}$ of $\{q_n\}_{n\geq 1}$. Note that Theorem 22.1(c) holds without imposing Assumption BC.1 and with $sd_{1j\infty}(\theta) := 1$ in Assumption BC.2 by Comment (ii) following Theorem 22.1(c). Hence, it remains to verify that Assumptions BC.2 (with $sd_{1j\infty}(\theta) := 1$), BC.3, BC.6, and C.1–C.8 hold for some subsequence $\{p_n\}_{n\geq 1}$ (of $\{q_n\}_{n\geq 1}$) in place of $\{n\}_{n\geq 1}$ (because Assumptions A.0, A.5–A.7, and S.1, which are imposed in Theorem 22.1(c), are also imposed in the present theorem, and Assumption N, which is imposed in Theorem 22.1(c), holds because $\theta_{a_n} \in \Theta_I(F_{a_n})$ $\forall n \geq 1$ in (24.1) by construction).

Under Assumptions A.4 and A.8, by Lemma D.7 of BCS, given $\{q_n\}_{n\geq 1}$, there exists a subsequence $\{u_n\}_{n\geq 1}$ of $\{q_n\}_{n\geq 1}$, a continuous $R^{k\times k}$-valued function Ω_∞ on Θ^2, and a continuous R^k-valued function $\widetilde{m}$ on Θ for which (i) $\Omega_{F_{u_n}} \to_u \Omega_\infty$, where $\to_u$ denotes uniform convergence (over Θ^2 in this case), (ii) $E_{F_{u_n}}\widetilde{m}(W,\cdot) \to_u \widetilde{m}(\cdot)$, and hence, Assumption C.4 holds for the subsequence $\{u_n\}_{n\geq 1}$, and (iii) Assumptions C.7, C.8, and BC.3 hold for the subsequence $\{u_n\}_{n\geq 1}$. Strictly speaking, Lemma D.7 of BCS only establishes $\Omega_{F_{u_n}} \to_u \Omega_\infty$ and the subsequence versions of Assumptions C.7 and C.8, but $E_{F_{u_n}} \widetilde{m}(W, \cdot) \to_u \widetilde{m}(\cdot)$ and the subsequence version of Assumption BC.3 are established in the same ways as $\Omega_{F_{u_n}} \to_u \Omega_\infty$ (but using Assumption A.8 in place of Assumption A.4) and the subsequence versions of Assumptions C.7 and C.8, respectively.

Assumption C.1 holds for a subsequence $\{\overline{u}_n\}_{n\geq 1}$ of $\{u_n\}_{n\geq 1}$ because $\{\theta_{u_n}\}_{n\geq 1}$ is a sequence in the compact set Θ (by Assumption A.0(i)).

Assumptions C.5 and C.6 hold for the subsequence $\{u_n\}_{n\geq 1}$ by applying a subsequence version of Lemma 24.1, which imposes Assumptions A.0–A.4, C.1, and C.9. Assumptions A.0–A.4 are imposed in the present theorem and the subsequence version of Assumption C.9 holds by (i) above.

Assumptions C.2, C.3, and BC.2 hold for a subsequence $\{p_n\}_{n\geq 1}$ of $\{\overline{u}_n\}_{n\geq 1}$ because $\{\overline{u}_n^{1/2} E_{F_{\overline{u}_n}} \widetilde{m}(W, \theta_{\overline{u}_n})\}_{n\geq 1}$, $\{\overline{u}_n^{1/2}(E_{F_{\overline{u}_n}} \widetilde{m}(W, \theta_{\overline{u}_n}) + r^{\inf}_{F_{\overline{u}_n}})\}_{n\geq 1}$, and $\{\kappa_{\overline{u}_n}^{-1} \overline{u}_n^{1/2}(E_{F_{\overline{u}_n}} \widetilde{m}(W, \theta_{\overline{u}_n}) + r^{\inf}_{F_{\overline{u}_n}})\}_{n\geq 1}$ are sequences taking values in $R^k_{[\pm\infty]}$, which is compact under d (defined following (5.2) with $a_* = k$).

Assumption BC.6 holds for the subsequence $\{p_n\}_{n\geq 1}$ by Lemma D.2(8) of BCS because Assumptions A.1–A.4 of this paper imply Assumptions A.1–A.4 of BCS and $\Omega_{F_{u_n}} \to_u \Omega_\infty$ implies $\Omega_{F_{p_n}} \to_u \Omega_\infty$ (because $\{p_n\}_{n\geq 1}$ is a subsequence of $\{u_n\}_{n\geq 1}$).

This concludes the proof that the assumptions employed in Theorem 22.1(c) hold for the subsequence $\{p_n\}_{n\geq 1}$ of $\{q_n\}_{n\geq 1}$, which completes the proof of part (b) for $CS_{n,SPUR1}$.

The proof of part (a) for the SPUR1 test is essentially the same as that of part (b) for the SPUR1 CS, but with θ_0 in place of θ_n $\forall n \geq 1$.

Next, we prove part (b) for the SPUR2 CS. Let $\{F_n\}_{n\geq 1}$ and $\{\theta_n\}_{n\geq 1}$ denote sequences of distributions in $\mathcal{P}$ for which

$$\limsup_{n\to\infty} \sup_{F\in\mathcal{P}} \sup_{\theta\in\Theta_I(F)} P_F(\phi_{n,SPUR2}(\theta) = 1) = \limsup_{n\to\infty} P_{F_n}(\phi_{n,SPUR2}(\theta_n) = 1). \tag{24.2}$$

Such sequences always exists. The left-hand side expression in (24.2) equals one minus the uniform coverage probability in Theorem 7.1(b) using the definition of the SPUR2 CS in (4.5).

We use the following Bonferroni argument. Define

$$\phi_{n,SPUR2}(\theta, r) := \begin{cases} \phi_{n,GMS}(\theta, \alpha_2) & \text{if } r = 0 \\ \phi_{n,SPUR1}(\theta, \alpha_2) & \text{if } r > 0. \end{cases} \tag{24.3}$$

Then, $\phi_{n,SPUR2}(\theta) = \min_{0 \le r \le \widehat{r}_{n,UP}(\alpha_1)} \phi_{n,SPUR2}(\theta, r)$. We have

$$\begin{aligned}
&\limsup_{n\to\infty} P_{F_n}(\phi_{n,SPUR2}(\theta_n) = 1) \\
\le\ &\limsup_{n\to\infty} P_{F_n}(\phi_{n,SPUR2}(\theta_n) = 1 \ \&\ r_{F_n}^{\inf} \le \widehat{r}_{n,UP}(\alpha_1)) \\
&+\limsup_{n\to\infty} P_{F_n}(\phi_{n,SPUR2}(\theta_n) = 1 \ \&\ r_{F_n}^{\inf} > \widehat{r}_{n,UP}(\alpha_1)) \\
\le\ &\limsup_{n\to\infty} P_{F_n}(\phi_{n,SPUR2}(\theta_n) = 1 \ \&\ r_{F_n}^{\inf} \le \widehat{r}_{n,UP}(\alpha_1)) + \alpha_1 \\
=\ &\limsup_{n\to\infty} P_{F_n}(\min_{0 \le r \le \widehat{r}_{n,UP}} \phi_{n,SPUR2}(\theta_n, r) = 1 \ \&\ r_{F_n}^{\inf} \le \widehat{r}_{n,UP}(\alpha_1)) + \alpha_1, \qquad (24.4)
\end{aligned}$$

where the inequality holds because $\liminf_{n\to\infty} P_{F_n}(r_{F_n}^{\inf} \in [0, \widehat{r}_{n,UP}(\alpha_1)]) \ge 1 - \alpha_1$.

First, consider the case where $r_{F_n}^{\inf} > 0$ for all n large. Under the null hypothesis, the rhs of (24.4) is less than or equal to

$$\limsup_{n\to\infty} P_{F_n}(\phi_{n,SPUR1}(\theta_n, \alpha_2) = 1) + \alpha_1 \le \alpha_2 + \alpha_1 = \alpha, \qquad (24.5)$$

where the inequality holds because the nominal level α_2 test $\phi_{n,SPUR1}(\theta_n, \alpha_2)$ has asymptotic size α_2 or less by Theorem 7.1(b) for the SPUR1 CS (which allows for drifting sequences of null values θ_n).

Next, consider the case where $r_{F_n}^{\inf} = 0$ for all n large. Under the null hypothesis, the rhs of (24.4) is less than or equal to

$$\limsup_{n\to\infty} P_{F_n}(\phi_{n,GMS}(\theta_n, \alpha_2) = 1) + \alpha_1 \le \alpha_2 + \alpha_1 = \alpha, \qquad (24.6)$$

where the inequality holds because the model is correctly specified (i.e., $r_{F_n}^{\inf} = 0$) for n large and the $\phi_{n,GMS}(\theta_n, \alpha_2)$ test has asymptotic size α_2 or less in this case. The latter holds by the same argument as used to prove Theorem 7.1(b) for the SPUR1 CS (which allows for drifting sequences of null values θ_n), but with the test statistic $S_n(\theta)$ defined in (4.2) with $\widehat{r}_n^{\inf}$ replaced by the true value $r_{F_n}^{\inf} = 0$ and with the EGMS bootstrap statistic replaced by the GMS bootstrap statistic $S_{n,GMS}^*(\theta)$ defined just above (4.7), which is suitable because $r_{F_n}^{\inf} = 0$.

The result of part (b) for the SPUR2 CS holds because the rhs of (24.4) for the sequence $\{F_n\}_{n\ge1}$ is α or less by considering subsequences of $\{n\}$ where either (24.5) or (24.6) applies.

The proof of part (a) for the SPUR2 test is analogous to that of part (b) for the SPUR2 CS with θ_0 in place of θ_n $\forall n \ge 1$. $\square$

Proof of Lemma 24.1. Now we verify Assumption C.5 using Lemma D.2(1) of BCS, which imposes their Assumptions A.1–A.4 and M.2 and $\Omega_{F_n} \to_u \Omega_\infty$ for some Ω_∞. Assumptions A.1–A.4

in this paper imply A.1–A.4 in BCS, Assumption A.0(i) is the same as BCS's M.2, and Assumption C.9 implies $\Omega_{F_n} \to_u \Omega_\infty$. Lemma D.2(1) of BCS gives $\nu_n^m(\cdot) \Rightarrow G^m(\cdot)$, whereas Assumption C.5 concerns $\nu_n(\cdot) := (\nu_n^m(\cdot)', \nu_n^\sigma(\cdot)')'$. However, by the same argument as in the proof of Lemma D.2(1) applied to $\nu_n(\cdot)$, rather than $\nu_n^m(\cdot)$, we obtain

$$\nu_n(\cdot) \Rightarrow G(\cdot), \tag{24.7}$$

where $G(\cdot)$ is as in Assumption C.5, using equicontinuity of $\nu_n(\cdot)$ in our Assumption A.2, rather than of $\nu_n^m(\cdot)$ in BCS's Assumption A.2, and using $4+a$ finite moments in our Assumption A.3, rather than $2+a$ finite moments in BCS's Assumption A.3. Hence, Assumption C.5 holds.

Next, we verify Assumption C.6. Lemma D.2(5) of BCS gives $\sup_{\theta\in\Theta} ||\widehat{\Omega}_n(\theta) - \Omega_{\infty 11}(\theta,\theta)|| \to_p 0$, where $\Omega_{\infty 11}(\theta,\theta)$ denotes the upper left $k \times k$ submatrix of $\Omega_\infty(\theta,\theta)$, because Assumptions A.1–A.4 in this paper imply Assumptions A.1–A.4 of BCS and $\Omega_{F_n} \to_u \Omega_\infty$ by Assumption C.9. By Assumption C.1, $\theta_n \to \theta_\infty$, and by Assumption C.9, $\Omega_\infty(\theta,\theta')$ is continuous on Θ^2. These results combine to yield $\widehat{\Omega}_n(\theta_n) \to_p \Omega_{\infty 11}(\theta_\infty,\theta_\infty) := \Omega_\infty$, which verifies Assumption C.6. $\square$

25 Proof of Theorem 12.1

Proof of Theorem 12.1. We prove part (a) first. There always exists a subsequence $\{p_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$ such that

$$\limsup_{n\to\infty} \; P_{F_n}(\phi_{n,EGMS}(\theta_n) = 1) = \lim P_{F_{p_n}}(\phi_{p_n,EGMS}(\theta_{q_n}) = 1). \tag{25.1}$$

By Theorem 22.1(a) applied with $\{p_n\}_{n\geq 1}$ defined in (25.1), there exists a subsequence $\{a_n\}_{n\geq 1}$ of $\{p_n\}_{n\geq 1}$ such that

$$\lim P_{F_{p_n}}(\phi_{p_n,EGMS}(\theta_{q_n}) = 1) = \lim P_{F_{a_n}}(\phi_{a_n,EGMS}(\theta_{a_n}) = 1) \leq P(S_\infty > c_{\infty,EGMS}(1-\alpha)), \tag{25.2}$$

where the equality holds because a subsequence has the same limit as the original sequence and the inequality holds by Theorem 22.1(a) with $\{p_n\}_{n\geq 1}$ defined in (25.2), which imposes Assumptions A.0, A.5, A.6, BC.1–BC.3, BC.6, C.1–C.8, NLA, and S.1 defined using the subsequence $\{p_n\}_{n\geq 1}$ in place of $\{n\}_{n\geq 1}$. Assumptions A.0, A.5, A.6, BC.1–BC.3, C.1, C.4, C.7, C.8, NLA, and S.1 (among others) defined using $\{n\}_{n\geq 1}$ are imposed in Theorem 12.1(a), which implies that the subsequence $\{p_n\}_{n\geq 1}$ versions of them also hold. Hence, it remains to verify Assumptions BC.6, C.5, and C.6 (defined using $\{p_n\}_{n\geq 1}$ in place of $\{n\}_{n\geq 1}$). Assumptions C.5 and C.6 hold for the subsequence

$\{p_n\}_{n\geq 1}$ by applying a subsequence version of Lemma 24.1, which imposes Assumptions A.0–A.4, C.1, and C.9. These assumptions are also imposed in the theorem. Assumption BC.6 holds for the subsequence $\{p_n\}_{n\geq 1}$ by Lemma D.2(8) of BCS because Assumptions A.1–A.4 of this paper imply Assumptions A.1–A.4 of BCS and $\Omega_{F_n} \to_u \Omega_\infty$ (by Assumption C.9) implies $\Omega_{F_{p_n}} \to_u \Omega_\infty$. This completes the proof of part (a).

Next, we prove part (b). There always exists a subsequence $\{p_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$ and, by Theorem 22.1(b), there exists a subsequence $\{a_n\}_{n\geq 1}$ of $\{p_n\}_{n\geq 1}$ such that

$$\begin{aligned}\liminf_{n\to\infty} P_{F_n}(\phi_{n,EGMS}(\theta_n) = 1) &= \lim P_{F_{p_n}}(\phi_{p_n,EGMS}(\theta_{p_n}) = 1) \\ &= \lim P_{F_{a_n}}(\phi_{a_n,EGMS}(\theta_{a_n}) = 1) \\ &\geq P(S_\infty > c_{\infty,EGMS}(1-\alpha)), \end{aligned} \tag{25.3}$$

where the second equality holds because a subsequence has the same limit as the original sequence and the inequality holds by Theorem 22.1(b) (with $\{p_n\}_{n\geq 1}$ defined in (25.3)), which employs the same assumptions as Theorem 22.1(a) except with Assumptions A.10, BC.4, and BC.5 in place of A.5 and BC.3. Given the assumptions imposed in part (b) of the theorem (which include Assumptions A.10, BC.4, and BC.5), it remains to verify Assumptions BC.6, C.5, and C.6 defined using the subsequence $\{p_n\}_{n\geq 1}$ in place of $\{n\}_{n\geq 1}$. These assumptions are verified by the same argument as in the proof of part (a) above.

Now, we prove part (c). There always exists a subsequence $\{p_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$ and, by Theorem 22.1(d), there exists a subsequence $\{a_n\}_{n\geq 1}$ of $\{p_n\}_{n\geq 1}$ such that

$$\begin{aligned}\liminf_{n\to\infty} P_{F_n}(\phi_{n,EGMS}(\theta_n) = 1) &= \lim P_{F_{p_n}}(\phi_{p_n,EGMS}(\theta_{p_n}) = 1) \\ &= \lim P_{F_{a_n}}(\phi_{a_n,EGMS}(\theta_{a_n}) = 1) = 1, \end{aligned} \tag{25.4}$$

where the second equality holds because a subsequence has the same limit as the original sequence and the third equality holds by Theorem 22.1(d) provided the assumptions of Theorem 22.1(d) hold for the subsequence $\{p_n\}_{n\geq 1}$ defined in (25.4) in place of $\{n\}_{n\geq 1}$. All of the latter assumptions hold by the assumptions imposed in Theorem 12.1(c) for the sequence $\{n\}_{n\geq 1}$, except Assumptions BC.6, C.5, and C.6 defined using the subsequence $\{p_n\}_{n\geq 1}$ in place of $\{n\}_{n\geq 1}$. These assumptions are verified by the same argument as given in the proof of part (b), which completes the proof of part (c). $\square$

26 Proof of Theorem 9.1 and Rate of Convergence of $\widehat{\Theta}_n$

This section proves Theorem 9.1 (i.e., it shows that $\widehat{\Theta}_n$, defined in (6.5), is uniformly consistent for $\Theta_I(F)$) and it establishes the rate of convergence of $d_H(\widehat{\Theta}_n, \Theta_I(F_n))$ to zero under suitable conditions. These results are similar to results in Theorem 3.1 of Chernozhukov, Hong, and Tamer (2007).

26.1 Consistency and Rate of Convergence of $\widehat{\Theta}_n$ under $\{F_n\}_{n\geq 1}$

Here we establish consistency and rate of convergence results for $\widehat{\Theta}_n$ under a drifting sequence of distributions $\{F_n\}_{n\geq 1}$.

The set $\Theta_{I,\varepsilon}(F_n)$, which is an ε-expansion of $\Theta_I(F_n)$, is defined in Section 9. The following assumption ensures that $\inf_{\theta\in\Theta\backslash\Theta_{I,\varepsilon}(F_n)} \max_{j\leq k}\ [E_{F_n}\widetilde{m}_j(W_i,\theta)]_- - r_{F_n}^{\inf}$ is bounded away from zero under $\{F_n\}_{n\geq 1}$.

Assumption C.10. For all $\varepsilon > 0$,

$$\liminf_{n\to\infty}\left(\inf_{\theta\in\Theta\backslash\Theta_{I,\varepsilon}(F_n)} \max_{j\leq k}\ [E_{F_n}\widetilde{m}_j(W_i,\theta)]_- - r_{F_n}^{\inf}\right) > 0.$$

The following minorant condition for the population moments is similar to (4.1) of Chernozhukov, Hong, and Tamer (2007). It is used to determine the rate of convergence of $d_H(\widehat{\Theta}_n, \Theta_I(F_n))$ to zero.

Assumption C.11. There exist positive constants C, ε, and γ such that for all $\theta\in\Theta$ and $n\geq 1$,

$$\max_{j\leq k}[E_{F_n}\widetilde{m}_j(W_i,\theta)]_- - r_{F_n}^{\inf} \geq C\cdot(\min\{d(\theta,\Theta_I(F_n)),\varepsilon\})^{\gamma}.$$

Typically, Assumption C.11 holds with $\gamma = 1$.

Part (a) of the following lemma is used in the proof of Theorem 9.1 given below. Part (b) provides a rate of convergence result for $\widehat{\Theta}_n$.

Lemma 26.1 *Suppose Assumptions* A.0, C.4, C.5, C.7, *and* C.10 *hold under* $\{F_n\}_{n\geq 1}$. *Suppose the positive constants* $\{\tau_n\}_{n\geq 1}$ *that appear in* (6.5) *satisfy* $\tau_n\to\infty$ *and* $\tau_n/n^{1/2} = o(1)$. *Then,*

(a) $d_H(\widehat{\Theta}_n, \Theta_I(F_n)) = o_p(1)$ *and*

(b) $d_H(\widehat{\Theta}_n, \Theta_I(F_n)) = O_p((\tau_n/n^{1/2})^{1/\gamma})$ *provided Assumption* C.11 *also holds.*

Comment. When $F_n = F$ for all $n\geq 1$ for some $F\in\mathcal{P}$, Assumption C.10 holds by the definitions of $r_F^{\inf}$ and $\Theta_{I,\varepsilon}(F)$ under Assumption A.0. In consequence, Lemma 26.1(a) establishes the result of Theorem 9.1 with $\sup_{F\in\mathcal{P}}$ deleted and without imposing Assumption A.9.

26.2 Proofs of Lemma 26.1 and Theorem 9.1

The proof of Lemma 26.1(b) uses the following lemma, which shows that Assumption C.11 implies a similar minorant condition on the sample analogue of the left-hand side of Assumption C.11.

Lemma 26.2 *Suppose Assumptions* A.0, C.4, C.5, C.7, *and* C.11 *hold under* $\{F_n\}_{n\geq 1}$. *Then, there exist positive constants* κ, ε, *and* γ *such that for any* $\delta \in (0,1)$ *there exists positive constants* κ_δ *and* N_δ *such that*

$$\max_{j\leq k}([\widehat{m}_{nj}(\theta)]_- - \widehat{r}_n^{\inf}) \geq \kappa \cdot (\min\{d(\theta, \Theta_I(F_n)), \varepsilon\})^\gamma$$

for all $\theta \in \{\theta \in \Theta : d(\theta, \Theta_I(F_n)) \geq (\kappa_\delta/n^{1/2})^{1/\gamma}\}$ *with probability at least* $1-\delta$ *for all* $n \geq N_\delta$.

Proof of Lemma 26.1. The proof is similar to that of Theorem 3.1 of Chernozhukov, Hong, and Tamer (2007). For part (a), we have

$$\sup_{\theta\in\Theta_I(F_n)} d(\theta, \widehat{\Theta}_n) = 0 \text{ wp} \to 1 \tag{26.1}$$

because $\Theta_I(F_n) \subset \widehat{\Theta}_n$ wp $\to 1$ by Lemma 23.1(a) (which requires Assumptions A.0, C.4, C.5, and C.7). For part (a), it remains to show $\sup_{\theta\in\widehat{\Theta}_n} d(\theta, \Theta_I(F_n)) = o_p(1)$.

By Assumption C.10, for arbitrary $\varepsilon > 0$, we have

$$\zeta_\varepsilon := \liminf_{n\to\infty} \inf_{\theta\in\Theta\backslash\Theta_{I,\varepsilon}(F_n)} \max_{j\leq k} [E_{F_n}\widetilde{m}_j(W_i,\theta)]_- - r_{F_n}^{\inf} > 0. \tag{26.2}$$

By (23.81) (which requires Assumptions A.0, C.4, C.5, and C.7), we have

$$\sup_{\theta\in\widehat{\Theta}_n} \max_{j\leq k}[E_{F_n}\widetilde{m}_j(W_i,\theta)]_- - r_{F_n}^{\inf} \leq O_p(1/n^{1/2}) + \tau_n/n^{1/2} = o_p(1), \tag{26.3}$$

where the equality holds because $\tau_n/n^{1/2} = o(1)$. Combining (26.2) and (26.3), it follows that

$$\begin{aligned} &\lim P_{F_n}\left(\inf_{\theta\in\Theta\backslash\Theta_{I,\varepsilon}(F_n)} \max_{j\leq k} [E_{F_n}\widetilde{m}_j(W_i,\theta)]_- > \sup_{\theta\in\widehat{\Theta}_n} \max_{j\leq k} [E_{F_n}\widetilde{m}_j(W_i,\theta)]_-\right) \\ &\geq \lim P_{F_n}(\zeta_\varepsilon/2 > o_p(1)) \\ &= 1. \end{aligned} \tag{26.4}$$

Thus, $\lim P_{F_n}(\widehat{\Theta}_n \subset \Theta_{I,\varepsilon}(F_n)) = 1$ and $\sup_{\theta\in\widehat{\Theta}_n} d(\theta, \Theta_I(F_n)) \leq \varepsilon$ wp$\to 1$. Since $\varepsilon > 0$ is arbitrary, we have $\sup_{\theta\in\widehat{\Theta}_n} d(\theta, \Theta_I(F_n)) = o_p(1)$, which completes the proof of part (a).

For part (b), take the positive constants $(\kappa, \varepsilon, \gamma, \delta, N_\delta, \kappa_\delta)$ as in Lemma 26.2. We can take $N'_\delta \geq N_\delta$ such that $2\tau_n > \kappa \cdot \kappa_\delta$ and $\varepsilon_n := (2\tau_n/(n^{1/2}\kappa))^{1/\gamma} < \varepsilon$ for $n \geq N'_\delta$, because $\tau_n \to \infty$ and $\tau_n/n^{1/2} = o(1)$. As defined, $\varepsilon_n > (\kappa_\delta/n^{1/2})^{1/\gamma}$ for $n \geq N'_\delta$. Hence,

$$\Theta \backslash \Theta_{I,\varepsilon_n} \subset \{\theta \in \Theta : d(\theta, \Theta_I(F_n)) \geq (\kappa_\delta/n^{1/2})^{1/\gamma}\} \tag{26.5}$$

for $n \geq N'_\delta$. In consequence, with probability at least $1-\delta$ for $n \geq N'_\delta$, we have

$$\begin{aligned}
\inf_{\theta \in \Theta \backslash \Theta_{I,\varepsilon_n}(F_n)} \max_{j \leq k}([\widehat{m}_{nj}(\theta)]_- - \widehat{r}_n^{\inf}) &\geq \kappa \cdot \inf_{\theta \in \Theta \backslash \Theta_{I,\varepsilon_n}(F_n)} (\min\{d(\theta, \Theta_I(F_n)), \varepsilon\})^\gamma \\
&\geq \kappa \cdot (\min\{\varepsilon_n, \varepsilon\})^\gamma \\
&= \kappa \cdot \varepsilon_n^\gamma \\
&:= 2\tau_n/n^{1/2} \\
&> \tau_n/n^{1/2} \\
&\geq \sup_{\theta \in \widehat{\Theta}_n} \max_{j \leq k}([\widehat{m}_{nj}(\theta)]_- - \widehat{r}_n^{\inf}),
\end{aligned} \tag{26.6}$$

where the first inequality holds by Lemma 26.2 and (26.5), the second inequality holds by the definition of $\Theta_{I,\varepsilon_n}(F_n)$, the first equality holds by the definition of N'_δ, the second equality holds by the definition of ε_n, and the last holds inequality by the definition of $\widehat{\Theta}_n$.

Equation (26.6) implies $\widehat{\Theta}_n \subset \Theta_{I,\varepsilon_n}(F_n)$, and hence, $\sup_{\theta \in \widehat{\Theta}_n} d(\theta, \Theta_I(F_n)) \leq \varepsilon_n$ with probability at least $1-\delta$ for $n \geq N'_\delta$. Combining this with (26.1) gives

$$d_H(\widehat{\Theta}_n, \Theta_I(F_n)) = O_p(\varepsilon_n) = O_p((\tau_n/n^{1/2})^{1/\gamma}), \tag{26.7}$$

which completes the proof of part (b). □

Proof of Lemma 26.2. By (23.78) with Θ in place of $\Theta_I^{\eta_n}(F_n)$ throughout and with $[E_{F_n}\widetilde{m}_j(W,\theta)]_- - r_{F_n}^{\inf}$ in place of η_n in the last two lines (which makes the inequality into an equality), we have

$$\max_{j \leq k}([\widehat{m}_{nj}(\theta)]_- - \widehat{r}_n^{\inf}) = \max_{j \leq k}[E_{F_n}\widetilde{m}_j(\theta)]_- - r_{F_n}^{\inf} + O_p^\Theta(1/n^{1/2}) \tag{26.8}$$

using Assumptions A.0, C.4, C.5, and C.7. Hence, for any $\delta \in (0,1)$, there exist positive constants κ_δ and N_δ such that with probability at least $1-\delta$, we have

$$\begin{aligned}
\max_{j \leq k}([\widehat{m}_{nj}(\theta)]_- - \widehat{r}_n^{\inf}) &\geq C \cdot (\min\{d(\theta, \Theta_I(F_n)), \varepsilon\})^\gamma + O_p^\Theta(1/n^{1/2}) \\
&\geq C \cdot (\min\{d(\theta, \Theta_I(F_n)), \varepsilon\})^\gamma - (C/2)\kappa_\delta/n^{1/2}
\end{aligned} \tag{26.9}$$

for all $\theta \in \Theta$ and $n \geq N_\delta$, where C, ε, and γ are as in Assumption C.11 and the first inequality uses (26.8) and Assumption C.11. Without loss in generality, we can take $N_\delta \geq (\kappa_\delta/\varepsilon^\gamma)^2$. Hence, $\kappa_\delta/N_\delta^{1/2} \leq \varepsilon^\gamma$.

For all $n \geq N_\delta$, we have

$$\kappa_\delta/n^{1/2} \leq (\min\{d(\theta, \Theta_I(F_n)), \varepsilon\})^\gamma \tag{26.10}$$

for all $\theta \in \{\theta \in \Theta : d(\theta, \Theta_I(F_n)) \geq (\kappa_\delta/n^{1/2})^{1/\gamma}\}$. Combining (26.9) and (26.10) establishes the lemma with $\kappa = C/2$. □

Proof of Theorem 9.1. Let an arbitrary $\varepsilon > 0$ be given. There always exists a sequence $\{F_n \in \mathcal{P}\}_{n\geq 1}$ (that may depend on ε) such that

$$\limsup_{n\to\infty} \sup_{F\in\mathcal{P}} P_F(d_H(\widehat{\Theta}_n, \Theta_I(F)) > \varepsilon) = \limsup_{n\to\infty} P_{F_n}(d_H(\widehat{\Theta}_n, \Theta_I(F_n)) > \varepsilon). \tag{26.11}$$

There always exists a subsequence $\{w_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$ such that

$$\limsup_{n\to\infty} P_{F_n}(d_H(\widehat{\Theta}_n, \Theta_I(F_n)) > \varepsilon) = \lim P_{F_{w_n}}(d_H(\widehat{\Theta}_{w_n}, \Theta_I(F_{w_n})) > \varepsilon). \tag{26.12}$$

Given any subsequence $\{a_n\}_{n\geq 1}$ of $\{w_n\}_{n\geq 1}$, there exists a subsequence $\{u_n\}_{n\geq 1}$ of $\{a_n\}_{n\geq 1}$ such that Assumptions C.4, C.7, and C.9 hold for the subsequence $\{u_n\}_{n\geq 1}$ by the proof of Theorem 7.1, which uses Lemma D.7 of BCS and relies on Assumptions A.4 and A.8. Given Assumption A.9, Assumption C.10 also holds for the subsequence $\{u_n\}_{n\geq 1}$. By Lemma 24.1, Assumptions A.0–A.4 and C.9 imply Assumption C.5. Hence, Assumptions C.4, C.5, C.7, and C.10 hold for the subsequence $\{u_n\}_{n\geq 1}$. In consequence, by Lemma 26.1(a) applied with n replaced by u_n, which utilizes Assumptions A.0, C.4, C.5, C.7, and C.10, we have

$$\lim P_{F_{u_n}} d_H(\widehat{\Theta}_{u_n}, \Theta_I(F_{u_n})) > \varepsilon) = 0. \tag{26.13}$$

This implies that the same result holds for the subsequence $\{w_n\}_{n\geq 1}$, which completes the proof using (26.11) and (26.12) because $\varepsilon > 0$ is arbitrary. □

27 Assumptions

For ease of reference, we state all of the assumptions used in the paper and Supplemental Material here.

Assumption A.0. (i) Θ is compact and non-empty and (ii) $E_F\widetilde{m}_j(W, \theta)$ is upper semi-continuous on Θ $\forall j \leq k$, $\forall F \in \mathcal{P}$.

Assumption A.1. The observations $W_1, ..., W_n$ are i.i.d. under F and $\{\widetilde{m}_j(\cdot,\theta) : \mathcal{W} \to R\}$ and $\{\widetilde{m}_j^2(\cdot,\theta) : \mathcal{W} \to R\}$ are measurable classes of functions indexed by $\theta \in \Theta$ $\forall j \leq k$, $\forall F \in \mathcal{P}$.

Assumption A.2. The empirical process $\nu_n(\cdot)$ is asymptotically ρ_F-equicontinuous on Θ uniformly in $F \in \mathcal{P}$.

Assumption A.3. For some $a > 0$, $\sup_{F\in\mathcal{P}} E_F \sup_{\theta\in\Theta} ||\widetilde{m}(W,\theta)||^{4+a} < \infty$.

Assumption A.4. The covariance kernel $\Omega_F(\theta,\theta')$ satisfies: for all $F \in \mathcal{P}$,

$\lim_{\delta\to 0} \sup_{||(\theta_1,\theta_1')-(\theta_2,\theta_2')||<\delta} ||\Omega_F(\theta_1,\theta_1') - \Omega_F(\theta_2,\theta_2')|| = 0$.

Assumption A.5. Given the function $\varphi : R^k_{[+\infty]} \times \Psi \to R^k_{[+\infty]}$, which appears in (6.4) and (6.11), there is a function $\varphi^* : R^k_{[+\infty]} \to R^k_{[+\infty]}$ that takes the form $\varphi^*(\xi) = (\varphi_1^*(\xi_1), ..., \varphi_k^*(\xi_k))'$ and $\forall j \leq k$, (i) $\varphi_j^*(\xi_j) \geq \varphi_j(\xi,\Omega) \geq 0$ $\forall(\xi,\Omega) \in R^k_{[+\infty]} \times \Psi$, (ii) φ_j^* is nondecreasing and continuous under the metric d, and (iii) $\varphi_j^*(\xi_j) = 0$ $\forall \xi_j \leq 0$ and $\varphi_j^*(\infty) = \infty$.

Assumption A.6. (i) $\kappa_n \to \infty$. (ii) $\tau_n \to \infty$.

Assumption A.7. Under $\{F_{q_n}\}_{n\geq 1}$ and $\{\theta_{q_n}\}_{n\geq 1}$, (i) if $c_\infty(1-\alpha) > 0$, then $P(S_\infty = c_\infty(1-\alpha)) = 0$, and (ii) if $c_\infty(1-\alpha) = 0$, then $\limsup_{n\to\infty} P_{F_{q_n}}(S_{q_n} > 0) \leq \alpha$.

Assumption A.8. $E_F\widetilde{m}(W,\theta)$ is equicontinuous on Θ over $F \in \mathcal{P}$. That is, $\lim_{\delta\downarrow 0} \sup_{F\in\mathcal{P}}$ $\sup_{||\theta-\theta'||<\delta} ||E_F\widetilde{m}(W,\theta) - E_F\widetilde{m}(W,\theta')|| = 0$.

Assumption A.9. For all $\varepsilon > 0$, $\inf_{F\in\mathcal{P}} \inf_{\theta\in\Theta\backslash\Theta_{I,\varepsilon}(F)} \max_{j\leq k} [E_F\widetilde{m}_j(W_i,\theta)]_- - r_F^{\inf} > 0$.

Assumption A.10. Given the function $\varphi : R^k_{[+\infty]} \times \Psi \to R^k_{[+\infty]}$, which appears in (6.4) and (6.11), there is a function $\varphi^{**} : R^k_{[+\infty]} \to R^k_{[+\infty]}$ that takes the form $\varphi^{**}(\xi) = (\varphi_1^{**}(\xi_1), ..., \varphi_k^{**}(\xi_k))'$ and $\forall j \leq k$, (i) $\varphi_j^{**}(\xi_j) \leq \varphi_j(\xi,\Omega)$ $\forall(\xi,\Omega) \in R^k_{[+\infty]} \times \Psi$, (ii) φ_j^{**} is continuous, and (iii) $\varphi_j^{**}(\xi_j) = 0$ $\forall \xi_j \leq 0$ and $\varphi_j^{**}(\infty) = \infty$.

Assumption S.1. (i) $S(m,\Omega)$ is nonincreasing in $m \in R^k_{[+\infty]}$ $\forall \Omega \in \Psi$.

(ii) $S(m,\Omega) \geq 0$ $\forall m \in R^k$, $\forall \Omega \in \Psi$.

(iii) $S(m,\Omega)$ is continuous at all $m \in R^k_{[+\infty]}$ and $\Omega \in \Psi$.

Assumption S.2. $S(m,\Omega) > 0$ iff $m_j < 0$ for some $j \leq k$, $\forall \Omega \in \Psi$.

Assumption S.3. For some $\chi > 0$, $S(am,\Omega) = a^\chi S(m,\Omega)$ $\forall a > 0$, $\forall m \in R^k$, $\forall \Omega \in \Psi$.

Assumption S.4. For all $h \in (-\infty,\infty]^k$, all $\Omega \in \Psi$, and $Z \sim N(0_k,\Omega)$, the distribution function of $S(Z+h,\Omega)$ at $x \in R$ is (i) continuous for $x > 0$, (ii) strictly increasing for $x > 0$ unless $h = (\infty,\ldots,\infty)' \in R^k_{[\pm\infty]}$, and (iii) less than 1/2 for $x = 0$ if $h_j = 0$ for some $j \leq k$.

The following assumptions apply to a drifting sequence of null values $\{\theta_n\}_{n\geq 1}$ and distributions $\{F_n\}_{n\geq 1}$.

Assumption C.1. $\theta_n \to \theta_\infty$ for some $\theta_\infty \in \Theta$.

Assumption C.2. $n^{1/2}E_{F_n}\widetilde{m}_j(W,\theta_n) \to \ell_{j\infty}$ for some $\ell_{j\infty} \in R_{[\pm\infty]}$ $\forall j \leq k$.

Assumption C.3. $n^{1/2}(E_{F_n}\widetilde{m}_j(W,\theta_n) + r_{F_n}^{\inf}) \to h_{j\infty}$ for some $h_{j\infty} \in R_{[\pm\infty]}$ $\forall j \leq k$.

Assumption C.4. $\sup_{\theta\in\Theta} ||E_{F_n}\widetilde{m}(W,\theta) - \widetilde{m}(\theta)|| \to 0$ for some nonrandom bounded continuous R^k-valued function $\widetilde{m}(\cdot)$ on Θ.

Assumption C.5. $\nu_n(\cdot) := (\nu_n^m(\cdot)', \nu_n^\sigma(\cdot)')' \Rightarrow G(\cdot) := (G^m(\cdot)', G^\sigma(\cdot)')'$ as $n \to \infty$, where $\{G(\theta) : \theta \in \Theta\}$ is a mean zero R^{2k}-valued Gaussian process with bounded continuous sample paths a.s. and $G^m(\theta), G^\sigma(\theta) \in R^k$.

Assumption C.6. $\widehat{\Omega}_n(\theta_n) \to_p \Omega_\infty$ for some $\Omega_\infty \in \Psi$.

Assumption C.7. $\Lambda_{n,F_n} \to_H \Lambda$ for some non-empty set $\Lambda \in \mathcal{S}(\Theta \times R^{2k}_{[\pm\infty]})$.

Assumption C.8. $\Lambda^{\eta_n}_{n,F_n} \to_H \Lambda_I$ for some non-empty set $\Lambda_I \in \mathcal{S}(\Theta \times R^{2k}_{[\pm\infty]})$, where $\{\eta_n\}_{n\geq 1}$ is a sequence of positive constants for which $\eta_n \to \infty$.

Assumption C.9. $\Omega_{F_n}(\cdot,\cdot) \to_u \Omega_\infty(\cdot,\cdot)$ for some continuous $R^{2k\times 2k}$-valued function $\Omega_\infty(\cdot,\cdot)$ on Θ^2.

Assumption C.10. For all $\varepsilon > 0$,

$$\liminf_{n\to\infty} \left(\inf_{\theta\in\Theta\backslash\Theta_{I,\varepsilon}(F_n)} \max_{j\leq k} \; [E_{F_n}\widetilde{m}_j(W_i,\theta)]_- - r_{F_n}^{\inf} \right) > 0.$$

Assumption C.11. There exist positive constants C, ε, and γ such that for all $\theta \in \Theta$ and $n \geq 1$,

$$\max_{j\leq k}[E_{F_n}\widetilde{m}_j(W_i,\theta)]_- - r_{F_n}^{\inf} \geq C \cdot (\min\{d(\theta,\Theta_I(F_n)),\varepsilon\})^\gamma.$$

The following assumptions apply to a drifting sequence of null values $\{\theta_n\}_{n\geq 1}$ and distributions $\{F_n\}_{n\geq 1}$.

Assumption BC.1. $\sup_{\theta\in\Theta} |sd^*_{anj}(\theta) - sd_{aj\infty}(\theta)| \to_p 0$ as $n \to \infty$ for some nonrandom continuous real-valued functions $sd_{aj\infty}(\theta)$ on Θ for $j \leq k$ and $a = 1, 3$.

Assumption BC.2. $(sd_{1j\infty}\kappa_n)^{-1} n^{1/2}(E_{F_n}\widetilde{m}_j(W,\theta_n) + r_{F_n}^{\inf}) \to h^*_{j\infty}$ for some $h^*_{j\infty} \in R_{[\pm\infty]}$ $\forall j \leq k$.

Assumption BC.3. $\Lambda^{*\eta_n}_{n,F_n} \to_H \Lambda^*_I$ for some non-empty set $\Lambda^*_I \in \mathcal{S}(\Theta \times R^{3k}_{[\pm\infty]} \times \{1,...,k\})$ for some constants $\{\eta_n\}_{n\geq 1}$ that satisfy $\eta_n \to \infty$ and $\eta_n/\tau_n \to 0$ for $\{\tau_n\}_{n\geq 1}$ as in Assumption A.6(ii).

Assumption BC.4. $\Lambda^{*\eta_{Un}}_{Un,F_n} \to_H \Lambda^*_{U,I}$ for some non-empty set $\Lambda^*_{U,I} \in \mathcal{S}(\Theta \times R^{2k}_{[\pm\infty]})$ for constants $\{\eta_{Un}\}_{n\geq 1}$ that satisfy $\eta_{Un} \to \infty$ and $\tau_n/\eta_{Un} \to 0$ for $\{\tau_n\}_{n\geq 1}$ as in Assumption A.6(ii).

Assumption BC.5. The distribution of $S^*_{U\infty,EGMS}$ is continuous at $c_{U\infty,EGMS}(1-\alpha)$.

Assumption BC.6. $\{\nu^*_n(\cdot)|\{W_{ni}\}_{i\leq n,n\geq 1}\} \Rightarrow G(\cdot)$ a.s.$[P_\bigtriangledown]$, where $G(\cdot)$ is as in Assumption C.5.

Assumption NLA. $\min_{j\leq k} h_{j\infty} > -\infty$.

Assumption CA. $\min_{j\leq k} h_{j\infty} = -\infty$.

Assumption N. $\theta_n \in \Theta_I(F_n)$ $\forall n \geq 1$.

Assumption LA. The null values $\{\theta_n\}_{n\geq 1}$ and distributions $\{F_n\}_{n\geq 1}$ satisfy: (i) $||\theta_n - \theta_{In}|| = O(n^{-1/2})$ for some sequence $\{\theta_{In} \in \Theta_I(F_n)\}_{n\geq 1}$, (ii) $n^{1/2}(E_{F_n}\widetilde{m}_j(W, \theta_{In}) + r_{F_n}^{\inf}) \to h_{Ij\infty}$ for some $h_{Ij\infty} \in R_{[\pm\infty]}$ $\forall j \leq k$, and (iii) $E_F\widetilde{m}(W, \theta)$ is Lipschitz on Θ uniformly over $\mathcal{P}$, i.e., there exists a constant $K < \infty$ such that $||E_F\widetilde{m}(W, \theta_1) - E_F\widetilde{m}(W, \theta_2)|| \leq K||\theta_1 - \theta_2||$ $\forall \theta_1, \theta_2 \in \Theta$, $\forall F \in \mathcal{P}$.

Assumption FA. The null values $\{\theta_n\}_{n\geq 1}$ and distributions $\{F_n\}_{n\geq 1}$ satisfy: (i) $F_n = F_* \in \mathcal{P}$ and $\theta_n = \theta_* \in \Theta$ do not depend on $n \geq 1$ and (ii) $E_{F_*}\widetilde{m}_j(W, \theta_*) + r_{F_*}^{\inf} < 0$ for some $j \leq k$.

Assumption A.7$_\Delta$. $P(A_{\infty,\Delta}^{\inf}(\Lambda_\Delta) = c_{\infty,\Delta}(\alpha)) = 0$.

Assumption IS. The sequence $\{F_n\}_{n\geq 1}$ is such that there exists a sequence $\{\theta_n^I \in \Theta_I(F_n)\}_{n\geq 1}$ for which $n^{1/2}E_{F_n}\widetilde{m}_j(W, \theta_n^I) \to \infty$ $\forall j \leq k$.

Assumption MM. The sequence $\{F_n\}_{n\geq 1}$ is such that $n^{1/2}r_{F_n}^{\inf} \to \infty$.

Assumption CV.1. There exist nonnegative random variables $\{S_{Ln}^*(\theta_n)\}_{n\geq 1}$ such that (i) $P_\bigtriangledown(S_{Ln}^*(\theta_n) \leq S_n^*(\theta_n)|\{W_{ni}\}_{i\leq n, n\geq 1}) = 1$ wp$\to$1 and (ii) $\{S_{Ln}^*(\theta_n)|\{W_{ni}\}_{i\leq n, n\geq 1}\} \to_d S_{L\infty}^*$ a.s.$[P_\bigtriangledown]$ for some $S_{L\infty}^* \in R$ a.s. that does not depend on the conditioning value of $\{W_{ni}\}_{i\leq n, n\geq 1}$.

Assumption CV.2. $S_{L\infty}^*$ satisfies $S_{L\infty}^* \geq_{ST} S_\infty$.

Assumption CV.3. There exist nonnegative random variables$\{S_{Un}^*(\theta_n)\}_{n\geq 1}$ such that (i) $P_\bigtriangledown(S_{Un}^*(\theta_n) \geq S_n^*(\theta_n)|\{W_{ni}\}_{i\leq n, n\geq 1}) = 1$ wp$\to$1 and (ii) $\{S_{Un}^*(\theta_n)|\{W_{ni}\}_{i\leq n, n\geq 1}\} \to_d S_{U\infty}^*$ a.s.$[P_\bigtriangledown]$ for some $S_{U\infty}^* \in R$ a.s. that does not depend on the conditioning value of $\{W_{ni}\}_{i\leq n, n\geq 1}$.

References

Andrews, D. W. K., and P. Guggenberger (2009), "Validity of Subsampling and 'Plug-in Asymptotic' Inference for Parameters Defined by Moment Inequalities," *Econometric Theory,* 25, 669–709.

——— (2010), "Asymptotic Size and a Problem with Subsampling and with the m Out of n Bootstrap," *Econometric Theory,* 26, 426–468.

Andrews, D. W. K., and S. Kwon (2019), "Confidence Intervals for Measures of Misspecification in Moment Inequality and/or Equality Models," unpublished manuscript, Cowles Foundation, Yale University.

Andrews, D. W. K., and G. Soares (2010), "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica,* 78, 119–157.

Bugni, F. A., I. A. Canay, and P. Guggenberger (2012), "Distortions of Asymptotic Confidence Size in Locally Misspecified Moment Inequality Models," *Econometrica,* 80, 1741–1768.

Bugni, F. A., I. A. Canay, and X. Shi (2015), "Specification Tests for Partially Identified Models Defined by Moment Inequalities," *Journal of Econometrics,* 185, 259–282.

Chernozhukov, V., H. Hong, and E. Tamer (2007), "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica,* 75, 1243–284.

Dudley, R. M. (1989), *Real Analysis and Probability.* Pacific Grove, CA: Wadsworth and Brook/Cole.

van der Vaart, A., and J. Wellner (1996), *Weak Convergence and Empirical Processes.* New York: Springer.