Econometrica Supplementary Material

SUPPLEMENT TO "HYBRID AND SIZE-CORRECTED SUBSAMPLING METHODS"

(*Econometrica*, Vol. 77, No. 3, May 2009, 721–762)

BY DONALD W. K. ANDREWS AND PATRIK GUGGENBERGER[1]

This supplement contains 11 sections of results. Section S1 provides details concerning Tables II and III. Section S2 gives the proofs of the results in the paper. Section S3 introduces size-correction methods based on quantile adjustment. Section S4 provides results concerning power comparisons of size-corrected (SC) tests. Section S5 provides graphical illustrations of the critical value functions of fixed-critical-value (FCV), subsampling, and hybrid tests. Section S6 gives graphical illustrations of power comparisons of SC-FCV, SC-Sub, and SC-Hyb tests. Section S7 introduces and gives results for equal-tailed size-corrected tests. Section S8 defines a size-corrected combined (SC-Com) test that combines the SC-Sub and SC-Hyb tests. Section S9 gives asymptotic and finite-sample results for hybrid, SC, and PSC tests for the nuisance parameter near a boundary example of Andrews and Guggenberger (2009a), hereafter AG1. Section S10 provides a table of asymptotic and finite-sample results for upper and lower one-sided confidence intervals in the autoregressive parameter example considered in the paper. Section S10 also verifies the assumptions for that example. Section S11 verifies the assumptions for the conservative model selection example considered in the paper.

S1. DETAILS CONCERNING TABLES II AND III

TO IMPLEMENT the Kristensen and Linton (2006) estimator used in the results of Table II, we use two Newton–Raphson iterations (see their equation (17)), and to initialize the iteration we use their closed-form estimator (see p. 326), in particular, their equation (10), implemented with $w_1 = w_2 = w_3 = 1/3$ and with their $\widehat{\phi}$ winsorized to the interval [.001, .999]. In each iteration step, we initialize the $\widehat{\sigma}^2_{k,t}$ (p. 329, line 5 from the bottom) by setting it equal to the squared first data observation. For simplicity, this estimator has not been discretized and the GARCH$(1, 1)$ process has not been truncated to conform to the theoretical results given in the Section 3.4 of Andrews and Guggenberger (2008) for the asymptotic equivalence of feasible and infeasible quasi-GLS (QGLS) statistics. The subsample statistics use the full-sample estimator of the conditional heteroskedasticity $\{\widehat{\phi}_{n,i} : i \leq n\}$, which is justified because feasible and infeasible QGLS test statistics are asymptotically equivalent in the full sample and in subsamples.

In Table II, the parameter space for ρ is taken to be $[-.9, 1.0]$ to minimize the effect of the choice of the lower bound on the FS-Min values of the subsampling and hybrid CIs because in most practical applications in economics,

[1]Andrews gratefully acknowledges the research support of the National Science Foundation via Grants SES-0417911 and SES-0751517. Guggenberger gratefully acknowledges research support from a Sloan Fellowship, a faculty research Grant from UCLA in 2005, and from the National Science Foundation via Grant SES-0748922.

 DOI: 10.3982/ECTA7015

the parameter interval $(-1.0, -.9]$ is not of interest. The effects are small. For the parameter spaces $[-.999, 1.0]$ and $[-.9, 1.0]$, the respective FS-Min values of the symmetric subsampling CIs are 94.6 and 95.0 for case (i), 95.1 and 95.4 for case (ii), 92.8 and 94.6 for case (iv), and 95.6 and 95.8 for case (v). For the symmetric hybrid CIs, they are 95.9 and 96.0 for case (iii), 93.7 and 94.6 for case (iv), and 96.0 and 96.1 for case (v). For the equal-tailed hybrid CI, they are 93.1 and 93.5 for case (iv). No other results are affected by the choice of the lower bound of the parameter space.

The 119 subsamples used in Table III include 10 "wrap-around" subsamples that contain observations at the end and beginning of the sample, for example, observations indexed by $(110, \ldots, 120, 1)$. The choice of $q_n = 119$ subsamples is made because this reduces rounding errors when q_n is small when computing the sample quantiles of the subsample statistics. The values ν_α that solve $\nu_\alpha/(q_n + 1) = \alpha$ for $\alpha = .025, .95$, and $.975$ are the integers 3, 114, and 117. In consequence, the .025, .95, and .975 sample quantiles are given by the 3rd, 114th, and 117th largest subsample statistics. See Hall (1992, p. 307) for a discussion of this choice in the context of the bootstrap.

S2. PROOFS

For notational simplicity, throughout this section, we let c_g, c_h, c_∞, $c_{n,b}$, and cv abbreviate $c_g(1-\alpha)$, $c_h(1-\alpha)$, $c_\infty(1-\alpha)$, $c_{n,b}(1-\alpha)$, and $cv(1-\alpha)$, respectively.

S2.1. *Proof of Lemma 1*

LEMMA 1: *Suppose Assumptions* A–G, K, *and* T *hold. Then either* (i) *the addition of $c_\infty(1-\alpha)$ to the subsampling critical value is irrelevant asymptotically* (*i.e.*, $c_h(1-\alpha) \geq c_\infty(1-\alpha)$ *for all* $h \in H$ *and* $\mathrm{Max}_{\mathrm{Hyb}}(\alpha) = \mathrm{Max}_{\mathrm{Sub}}(\alpha)$) *or* (ii) *the nominal level α subsampling test over-rejects asymptotically* (*i.e.*, $\mathrm{AsySz}(\theta_0) > \alpha$) *and the hybrid test reduces the asymptotic over-rejection for at least one parameter value* $(g, h) \in GH$.

PROOF: If $c_h \geq c_\infty$ for all $h \in H$, then $\mathrm{Max}_{\mathrm{Hyb}}(\alpha) = \mathrm{Max}_{\mathrm{Sub}}(\alpha)$ and $\mathrm{Max}^-_{\mathrm{Hyb}}(\alpha) = \mathrm{Max}^-_{\mathrm{Sub}}(\alpha)$ follow immediately (where the latter three quantities are defined in Assumptions P and T). In addition, Assumption T implies that all of these quantities are equal. The latter, Theorem 1 of the paper, and Theorem 1(ii) of AG1 imply that the quantities equal $\mathrm{AsySz}(\theta_0)$ for the hybrid and subsampling tests.

On the other hand, suppose $c_h \geq c_\infty$ for all $h \in H$ does not hold. Then, for some $g \in H$, $c_g < c_\infty$. Given g, define $h_1 = (h_{1,1}, \ldots, h_{1,p})' \in H_1$ by $h_{1,m} = +\infty$ if $g_{1,m} > 0$, $h_{1,m} = -\infty$ if $g_{1,m} < 0$, $h_{1,m} = +\infty$ or $-\infty$ (chosen so that $(g, h) \in GH$) if $g_{1,m} = 0$ for $m = 1, \ldots, p$, and define $h_2 = g_2$. Let $h = (h_1, h_2)$. By construction, $(g, h) \in GH$. By Assumption K, $c_h = c_\infty$. Hence, we have

$$\mathrm{Max}_{\mathrm{Sub}}(\alpha) \geq 1 - J_h(c_g) > \alpha, \tag{S2.1}$$

where the second inequality holds because $c_g < c_\infty = c_h$ and c_h is the infimum of values x such that $J_h(x) \geq 1 - \alpha$ or, equivalently, $1 - J_h(x) \leq \alpha$. Equation (S2.1) and Theorem 1(ii) of AG1 imply that $\text{AsySz}(\theta_0) > \alpha$ for the subsampling test. The hybrid test reduces the asymptotic over-rejection of the subsampling test at (g, h) from being at least $1 - J_h(c_g) > \alpha$ to being at most $1 - J_h(c_\infty) = 1 - J_h(c_h) \leq \alpha$ (with equality if $J_h(\cdot)$ is continuous at c_h). *Q.E.D.*

S2.2. *Proof of Lemma 2*

LEMMA 2: *Suppose Assumptions* A–G, K, T, *and* Quant *hold. Then, the hybrid test based on* $T_n(\theta_0)$ *has* $\text{AsySz}(\theta_0) = \alpha$.

PROOF: Suppose Assumption Quant(i) holds. Then,

$$\text{(S2.2)} \qquad \begin{aligned} \text{Max}_{\text{Hyb}}(\alpha) &= \sup_{(g,h)\in GH} \bigl[1 - J_h(\max\{c_g, c_\infty\})\bigr] \\ &= \sup_{h\in H}[1 - J_h(c_\infty)] \\ &\leq \sup_{h\in H}[1 - J_h(c_h)] = \alpha, \end{aligned}$$

where the second equality and the inequality hold by Assumption Quant(i)(a) and the last equality holds because $1 - J_h(c_h) \leq \alpha$ by definition of c_h for all $h \in H$ and $1 - J_\infty(c_\infty) = \alpha$ by Assumption Quant(i)(b). By (S2.2) and Assumption Quant(i)(b), $\text{Max}_{\text{Hyb}}(\alpha) = \sup_{h\in H}[1 - J_h(c_\infty)] \geq 1 - J_\infty(c_\infty) = \alpha$.

Next, suppose Assumption Quant(ii) holds. By Assumption Quant(ii)(a), $p = 1$. Hence, given $(g, h) \in GH$, either (I) $|h_{1,1}| = \infty$ or (II) $|h_{1,1}| < \infty$. When (I) holds, $J_h = J_\infty$ by Assumption K and

$$\text{(S2.3)} \qquad 1 - J_h(\max\{c_g, c_\infty\}) \leq 1 - J_\infty(c_\infty) = \alpha.$$

When (II) holds, g must equal h^0 by the definition of GH. Hence,

$$\text{(S2.4)} \qquad 1 - J_h(\max\{c_g, c_\infty\}) \leq 1 - J_h(c_{h^0}) \leq \sup_{h\in H}[1 - J_h(c_h)] = \alpha,$$

where the second inequality holds because $c_{h^0} \geq c_h$ by Assumption Quant(ii)(b) and the equality holds by Assumption Quant(ii)(c). Hence, $\text{Max}_{\text{Hyb}}(\alpha) \leq \alpha$. In addition, $\text{Max}_{\text{Hyb}}(\alpha) \geq 1 - J_\infty(c_\infty) = \alpha$ by Assumption Quant(ii)(c). *Q.E.D.*

S2.3. *Proof of Theorem 2*

In this section, we prove Theorem 2 of the paper. For the reader's convenience, we repeat the definition of the size-corrected (SC) tests here. The size-corrected fixed-critical-value (SC-FCV), subsampling (SC-Sub), and hybrid (SC-Hyb) tests with nominal level α are defined to reject the null hypothesis $H_0: \theta = \theta_0$ when

$$\text{(S2.5)} \qquad \begin{aligned} &T_n(\theta_0) > cv(1-\alpha), \\ &T_n(\theta_0) > c_{n,b}(1-\alpha) + \kappa(\alpha), \\ &T_n(\theta_0) > \max\{c_{n,b}(1-\alpha), c_\infty(1-\alpha) + \kappa^*(\alpha)\}, \end{aligned}$$

respectively, where

$$\text{(S2.6)} \qquad \begin{aligned} &cv(1-\alpha) = \sup_{h \in H} c_h(1-\alpha), \\ &\kappa(\alpha) = \sup_{(g,h) \in GH} [c_h(1-\alpha) - c_g(1-\alpha)], \\ &\kappa^*(\alpha) = \sup_{h \in H^*} c_h(1-\alpha) - c_\infty(1-\alpha), \\ &H^* = \{h \in H : \text{for some } (g,h) \in GH, c_g(1-\alpha) < c_h(1-\alpha)\}. \end{aligned}$$

If H^* is empty, then $\kappa^*(\alpha) = -\infty$ by definition.

THEOREM 2: *Suppose Assumptions* A–G *and* K–M *hold*. *Then the* SC-FCV, SC-Sub, *and* SC-Hyb *tests satisfy* $\text{AsySz}(\theta_0) = \alpha$.

PROOF: First we note that Assumption L implies that cv, $\kappa(\alpha)$, and $\kappa^*(\alpha)$ are finite. Below we show that cv, $\kappa(\alpha)$, and $\kappa^*(\alpha)$ satisfy

$$\text{(S2.7)} \qquad \begin{aligned} &\sup_{h \in H} [1 - J_h(cv-)] \le \alpha, \\ &\sup_{(g,h) \in GH} \big(1 - J_h\big((c_g + \kappa(\alpha))-\big)\big) \le \alpha, \\ &\sup_{(g,h) \in GH} \big(1 - J_h\big(\max\{c_g, c_\infty + \kappa^*(\alpha)\}-\big)\big) \le \alpha, \end{aligned}$$

respectively. Given (S2.7), Theorem 1(i) of AG1 applied with $c_{\text{Fix}} = cv$ implies that the SC-FCV test satisfies $\text{AsySz}(\theta_0) \le \sup_{h \in H}[1 - J_h(cv-)] \le \alpha$, where the second inequality holds by (S2.7). Theorem 1(ii) of AG1 with $c_{n,b} + \kappa(\alpha)$ in place of $c_{n,b}$ implies that the SC-Sub test satisfies $\text{AsySz}(\theta_0) \le \sup_{(g,h) \in H}[1 - J_h((c_g + \kappa(\alpha))-)] \le \alpha$, where the second inequality holds by (S2.7). Theorem 1(ii) of AG1 with $\max\{c_{n,b}, c_\infty + \kappa^*(\alpha)\}$ in place of $c_{n,b}$ implies that the SC-Hyb test satisfies $\text{AsySz}(\theta_0) \le \sup_{(g,h) \in H}[1 - J_h(\max\{c_g, c_\infty + \kappa^*(\alpha)\}-)] \le \alpha$, where the second inequality holds by (S2.7). Hence, $\text{AsySz}(\theta_0) \le \alpha$ for SC-

FCV, SC-Sub, and SC-Hyb tests. Below we show that the reverse inequality also holds.

We now show that the first inequality in (S2.7) holds. For $h \in H$, if $c_h < \sup_{h^\dagger \in H} c_{h^\dagger}$, then

$$J_h\Big(\sup_{h^\dagger \in H} c_{h^\dagger}-\Big) \geq J_h(c_h) \geq 1-\alpha, \tag{S2.8}$$

where the first inequality holds because J_h is nondecreasing and the second inequality holds by the definition of c_h. For $h \in H$, if $c_h = \sup_{h^\dagger \in H} c_{h^\dagger}$, then

$$J_h\Big(\sup_{h^\dagger \in H} c_{h^\dagger}-\Big) = J_h(c_h-) = 1-\alpha, \tag{S2.9}$$

where the last equality holds by Assumption M(a)(ii). For cv defined in (S2.6), (S2.8), and (S2.9), combine to give

$$\sup_{h \in H}[1 - J_h(cv-)] = \sup_{h\in H}\Big[1 - J_h\Big(\sup_{h^\dagger \in H} c_{h^\dagger}-\Big)\Big] \leq \alpha. \tag{S2.10}$$

Hence, cv satisfies (S2.7).

Next, we prove that the second inequality in (S2.7) holds. For $(g, h) \in GH$, if $c_{h=} < c_g + \sup_{(g^\dagger,h^\dagger)\in GH}[c_{h^\dagger} - c_{g^\dagger}]$, then we have

$$J_h\big((c_g + \kappa(\alpha))-\big) = J_h\Big(\Big(c_g + \sup_{(g^\dagger,h^\dagger)\in GH}[c_{h^\dagger} - c_{g^\dagger}]\Big)-\Big) \geq J_h(c_h) \geq 1-\alpha, \tag{S2.11}$$

where the first inequality holds by the condition on (g, h) and the fact that J_h is nondecreasing.

For $(g, h) \in GH$, if $c_{h=} = c_g + \sup_{(g^\dagger,h^\dagger)\in GH}[c_{h^\dagger} - c_{g^\dagger}]$, then we have

$$\begin{aligned} J_h\big((c_g + \kappa(\alpha))-\big) &= J_h\Big(\Big(c_g + \sup_{(g^\dagger,h^\dagger)\in GH}[c_{h^\dagger} - c_{g^\dagger}]\Big)-\Big) \\ &= J_h(c_h-) = 1-\alpha, \end{aligned} \tag{S2.12}$$

where the second equality holds by the condition on (g, h) and the last equality holds by Assumption M(b)(ii). Combining (S2.11) and (S2.12) gives $\sup_{(g,h)\in GH}[1 - J_h((c_g + \kappa(\alpha))-)] \leq \alpha$, as desired.

The third inequality in (S2.7) holds by the following argument. Because $c_\infty + \kappa^*(\alpha) = \sup_{h^*\in H^*} c_{h^*}$, we need to show that $\sup_{(g,h)\in GH}[1 - J_h(\max\{c_g, \sup_{h^*\in H^*} c_{h^*}\}-)] \leq \alpha$. For all $(g, h) \in GH$, we have $\max\{c_g, \sup_{h^*\in H^*} c_{h^*}\} \geq c_h$ because $\max\{c_g, \sup_{h^*\in H^*} c_{h^*}\} < c_h$ implies that $c_g < c_h$, which implies that $h \in H^*$, which implies that $\sup_{h^*\in H^*} c_{h^*} \geq c_h$, which is a contradiction. Now, for any $(g, h) \in GH$ with $\max\{c_g, \sup_{h^*\in H^*} c_{h^*}\} > c_h$, we have $1 - J_h(\max\{c_g, \sup_{h^*\in H^*} c_{h^*}\}-) \leq 1 - J_h(c_h) \leq \alpha$, as desired. For any $(g, h) \in$

GH with $\max\{c_g, \sup_{h^*\in H^*} c_{h^*}\} = c_h$, Assumption M(c)(ii) implies that $J_h(x)$ is continuous at $x = c_h$. Hence, $1 - J_h(\max\{c_g, c_\infty + \kappa^*(\alpha)\}-) = 1 - J_h(c_h-) = 1 - J_h(c_h) = \alpha$, which completes the proof of the third inequality of (S2.7). This concludes the proof that $\text{AsySz}(\theta_0) \leq \alpha$ for the SC-FCV, SC-Sub, and SC-Hyb tests.

We now prove that these tests satisfy $\text{AsySz}(\theta_0) \geq \alpha$. By Theorem 1(i) of AG1 applied with $c_{\text{Fix}} = cv$, the SC-FCV test satisfies $\text{AsySz}(\theta_0) \geq \sup_{h\in H}[1 - J_h(cv)]$. Using (S2.6) and Assumption M(a)(i), $cv = \sup_{h\in H} c_h = c_{h^*}$ for some $h^* \in H$. Hence,

$$\text{(S2.13)} \quad \sup_{h\in H}[1 - J_h(cv)] = \sup_{h\in H}[1 - J_h(c_{h^*})] \geq 1 - J_{h^*}(c_{h^*}) = \alpha,$$

where the last equality holds by Assumption M(a)(ii). In consequence, for the SC-FCV test, $\text{AsySz}(\theta_0) \geq \alpha$.

Next, by Theorem 1(ii) of AG1 with $c_{n,b} + \kappa(\alpha)$ in place of $c_{n,b}$, the SC-Sub test satisfies $\text{AsySz}(\theta_0) \geq \sup_{(g,h)\in GH}[1 - J_h(c_g + \kappa(\alpha))]$. Using (S2.6) and Assumption M(b)(i), $\kappa(\alpha) = c_{h^*} - c_{g^*}$ for some $(g^*, h^*) \in GH$ as in Assumption M(b)(i). Hence,

$$\text{(S2.14)} \quad \sup_{(g,h)\in GH}\big[1 - J_h(c_g + \kappa(\alpha))\big] = \sup_{(g,h)\in GH}[1 - J_h(c_g + c_{h^*} - c_{g^*})] \geq 1 - J_{h^*}(c_{h^*}) = \alpha,$$

where the last equality holds by Assumption M(b)(ii). In consequence, for the SC-Sub test, $\text{AsySz}(\theta_0) \geq \alpha$.

Last, Theorem 1(ii) of AG1 with $\max\{c_{n,b}, c_\infty + \kappa^*(\alpha)\}$ in place of $c_{n,b}$ implies that the SC-Hyb test satisfies

$$\text{(S2.15)} \quad \text{AsySz}(\theta_0) \geq \sup_{(g,h)\in GH}\big[1 - J_h\big(\max\{c_g, c_\infty + \kappa^*(\alpha)\}\big)\big].$$

If H^* is not empty, then using (S2.6) and Assumption M(c)(i), $\kappa^*(\alpha) = c_{h^*} - c_\infty$ for some $h^* \in H^*$ as in Assumption M(c)(i). By the definition of H^*, there exists g^* such that $(g^*, h^*) \in GH$ and $c_{g^*} < c_{h^*}$. In consequence, the right-hand side of (S2.15) equals

$$\text{(S2.16)} \quad \sup_{(g,h)\in GH}\big[1 - J_h(\max\{c_g, c_{h^*}\})\big] \geq 1 - J_{h^*}(\max\{c_{g^*}, c_{h^*}\}) = 1 - J_{h^*}(c_{h^*}) = \alpha,$$

where the first equality uses $c_{g^*} < c_{h^*}$ and the last equality holds by Assumption M(c)(ii) because $(g^*, h^*) \in GH$ satisfies $c_{h^*} = \sup_{h\in H^*} c_h = \max\{c_{g^*}, \sup_{h\in H^*} c_h\}$. Combining (S2.15) and (S2.16) gives $\text{AsySz}(\theta_0) \geq \alpha$.

If H^* is empty, then $\kappa^*(\alpha) = -\infty$, $(h^0, h^0) \in GH$, where $h^0 = (0, h_2)$ for arbitrary $h_2 \in H_2$, and we have

$$\text{(S2.17)}\quad \sup_{(g,h)\in GH} \bigl[1 - J_h\bigl(\max\{c_g, c_\infty + \kappa^*(\alpha)\}\bigr)\bigr] = \sup_{(g,h)\in GH} [1 - J_h(c_g)] \geq 1 - J_{h^0}(c_{h^0}) = \alpha,$$

where the last equality holds by Assumption M(c)(ii) because $c_{h^0} = \max\{c_{h^0}, c_\infty + \kappa^*(\alpha)\}$. Combining (S2.15)–(S2.17) gives $\text{AsySz}(\theta_0) \geq \alpha$ for the SC-Hyb test. *Q.E.D.*

S2.4. *Proof of Theorem 3*

THEOREM 3: *Suppose Assumptions* A–G, K, L, N, *and* O *hold*. *Then* (a) $cv_{\widehat{\gamma}_{n,2}}(1-\alpha) - cv_{\gamma_{n,2}}(1-\alpha) \to_p 0$, $\kappa_{\widehat{\gamma}_{n,2}}(\alpha) - \kappa_{\gamma_{n,2}}(\alpha) \to_p 0$, *and* $\kappa^*_{\widehat{\gamma}_{n,2}}(\alpha) - \kappa^*_{\gamma_{n,2}}(\alpha) \to_p 0$ *under all sequences* $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) \in \Gamma : n \geq 1\}$ *and* (b) *the PSC-FCV*, *PSC-Sub*, *and PSC-Hyb tests satisfy* $\text{AsySz}(\theta_0) = \alpha$.

PROOF: The results of part (a) hold by an extension of Slutsky's theorem (to allow $\gamma_{n,2}$ to depend on n) using Assumption N and the uniform continuity of the functions in Assumption O(a)(i), (b)(i), and (c)(i). The proof of part (b) is split into two steps. In the first step, we consider the PSC tests with $\widehat{\gamma}_{n,2}$ replaced by the true value $\gamma_{n,2}$. In this case, using parts (ii) and (iii) of Assumption O(a), (b), and (c), the results of part (b) hold by a very similar argument to that given in the proof of Theorem 2 of the paper. In the second step, the results of parts (a) are combined with the results of the first step to obtain the desired results. This step holds because the results of parts (a) lead to the same limit distributions for the statistics in question whether they are based on $\widehat{\gamma}_{n,2}$ or the true value $\gamma_{n,2}$ by the argument used in the proof of Theorem 1(ii) of AG1. *Q.E.D.*

S2.5. *Proof of Theorem 4*

THEOREM 4(a): *Suppose Assumptions* A–G *and* P *hold*. *Then a subsampling test satisfies*

$$\lim_{n\to\infty} \text{AsySz}_n(\theta_0) = \text{AsySz}(\theta_0).$$

PROOF: Under Assumptions A–G, Theorem 1 of AG1 combined with Assumption P(ii) shows that $\text{AsySz}(\theta_0) = \sup_{(g,h)\in GH}(1 - J_h(c_g))$. First, we show that $\liminf_{n\to\infty} \text{AsySz}_n(\theta_0) \geq \text{AsySz}(\theta_0)$. Given $(g, h) = ((g_1, h_2), (h_1, h_2)) \in GH$, we construct a sequence $\{h_n = (h_{n,1}, h_{n,2}) \in H : n \geq 1\}$ such that $(g_n, h_n) \to (g, h)$ as $n \to \infty$, where $g_n = (g_{n,1}, g_{n,2}) = (\delta_n^r h_{n,1}, h_{n,2})$. Define $h_{n,2} =$

h_2 for all $n \geq 1$. We write $h_1 = (h_{1,1}, \ldots, h_{1,p})'$ and $h_{n,1} = (h_{n,1,1}, \ldots, h_{n,1,p})'$. For $m = 1, \ldots, p$, define

$$h_{n,1,m} = \begin{cases} h_{1,m}, & \text{if } g_{1,m} = 0 \text{ and } |h_{1,m}| < \infty, \\ (n/b_n)^{r/2}, & \text{if } g_{1,m} = 0 \text{ and } h_{1,m} = \infty, \\ -(n/b_n)^{r/2}, & \text{if } g_{1,m} = 0 \text{ and } h_{1,m} = -\infty, \\ (n/b_n)^r g_{1,m}, & \text{if } g_{1,m} \in (0, \infty) \text{ and } h_{1,m} = \infty, \\ (n/b_n)^r g_{1,m}, & \text{if } g_{1,m} \in (-\infty, 0) \text{ and } h_{1,m} = -\infty, \\ (n/b_n)^{2r}, & \text{if } g_{1,m} = \infty \text{ and } h_{1,m} = \infty, \\ -(n/b_n)^{2r}, & \text{if } g_{1,m} = -\infty \text{ and } h_{1,m} = -\infty. \end{cases} \tag{S2.18}$$

As defined, $(g_{n,1}, h_{n,1}) = (\delta_n^r h_{n,1}, h_{n,1}) \to (g_1, h_1)$ and $(g_n, h_n) \to (g, h)$.

We now have

$$\begin{aligned} \liminf_{n\to\infty} \text{AsySz}_n(\theta_0) &= \liminf_{n\to\infty} \sup_{h=(h_1,h_2)\in H} \big(1 - J_h\big(c_{(\delta_n^r h_1, h_2)}\big)\big) \\ &\geq \liminf_{n\to\infty} \big(1 - J_{h_n}\big(c_{(\delta_n^r h_{n,1}, h_{n,2})}\big)\big) \\ &= \liminf_{n\to\infty} \big(1 - J_{h_n}\big(c_{g_n}\big)\big) \\ &= 1 - J_h(c_g), \end{aligned} \tag{S2.19}$$

where the second equality holds by definition of g_n and the last equality holds by Assumption P because $(g_n, h_n) \to (g, h)$. Using the expression for $\text{AsySz}(\theta_0)$ given above, this establishes the desired result because (S2.19) holds for all $(g, h) \in GH$.

Next, we show that $\limsup_{n\to\infty} \text{AsySz}_n(\theta_0) \leq \text{AsySz}(\theta_0)$. For $h = (h_1, h_2) \in H$, let $\tau_n(h) = 1 - J_h(c_{(\delta_n^r h_1, h_2)})$. By definition, $\text{AsySz}_n(\theta_0) = \sup_{h\in H} \tau_n(h)$. There exists a sequence $\{h_n \in H : n \geq 1\}$ such that $\limsup_{n\to\infty} \sup_{h\in H} \tau_n(h) = \limsup_{n\to\infty} \tau_n(h_n)$. There exists a subsequence $\{u_n\}$ of $\{n\}$ such that $\limsup_{n\to\infty} \tau_n(h_n) = \lim_{n\to\infty} \tau_{u_n}(h_{u_n})$. There exists a subsequence $\{v_n\}$ of $\{u_n\}$ such that $(h_{v_n,1}, h_{v_n,2}, \delta_{v_n}^r h_{v_n,1}) \to (h_1^*, h_2^*, g_1^*)$ for some $h_1^* \in H_1$, $h_2^* \in H_2$, $g_1^* \in H_1$, where $(g^*, h^*) = ((g_1^*, h_2^*), (h_1^*, h_2^*)) \in GH$. Hence,

$$\begin{aligned} \limsup_{n\to\infty} \text{AsySz}_n(\theta_0) &= \lim_{n\to\infty} \tau_{u_n}\big(h_{u_n}\big) = \lim_{n\to\infty} \tau_{v_n}\big(h_{v_n}\big) \\ &= \lim_{n\to\infty} \big(1 - J_{h_{v_n}}\big(c_{(\delta_{v_n}^r h_{v_n,1}, h_{v_n,2})}\big)\big) = 1 - J_{h^*}(c_{g^*}) \\ &\leq \sup_{(g,h)\in GH} (1 - J_h(c_g)) = \text{AsySz}(\theta_0), \end{aligned} \tag{S2.20}$$

where the fourth equality holds by Assumption P and the results above. *Q.E.D.*

THEOREM 4(b): *Suppose Assumptions* A–G, K–M, Q, *and* R *hold*. *Then* (i) $\lim_{n\to\infty}\kappa(\delta_n,\alpha)=\kappa(\alpha)$ *and* $\lim_{n\to\infty}\kappa^*(\delta_n,\alpha)=\kappa^*(\alpha)$ *and* (ii) *the* ASC-Sub *and* ASC-Hyb *tests satisfy* $\text{AsySz}(\theta_0)=\alpha$.

PROOF: The first result of part (i) holds by the proof of Theorem 4(a) with $1-J_h(c_{(\delta_n^r h_1,h_2)})$ and $1-J_h(c_g)$ replaced by $c_{(h_1,h_2)}-c_{(\delta_n^r h_1,h_2)}$ and c_h-c_g, respectively, using Assumption Q in place of Assumption P. Next, we show the first result of part (ii). Using the first result of part (i), by the same argument as used to prove Theorem 1(ii) of AG1, $\text{AsySz}(\theta_0)$ for the ASC-Sub test equals $\text{AsySz}(\theta_0)$ for the SC-Sub test. By Theorem 2 of the paper, the latter equals α.

Now, we prove that the second result of part (i) holds with $\lim_{n\to\infty}$ and $=$ replaced by $\liminf_{n\to\infty}$ and $\geq$, respectively, even without imposing Assumption R. If H^* is empty, then $\liminf_{n\to\infty}\sup_{h\in H^*(\delta_n)}c_h\geq-\infty=\sup_{h\in H^*}c_h$. If H^* is nonempty, for any $(g,h)\in GH$ such that $h\in H^*$, define $(g_n,h_n)\in GH$ as in (S2.18). By $(g_n,h_n)\to(g,h)$, Assumption Q, and $c_g-c_h<0$, we obtain $c_{g_n}-c_{h_n}<0$ and $h_n\in H^*(\delta_n)$ for all n sufficiently large. Hence,

$$\liminf_{n\to\infty}\sup_{h\in H^*(\delta_n)}c_h\geq\liminf_{n\to\infty}c_{h_n}=c_h, \tag{S2.21}$$

where the equality uses $h_n\to h$ and Assumption Q. This inequality holds for all $h\in H^*$. Hence, $\liminf_{n\to\infty}\sup_{h\in H^*(\delta_n)}c_h\geq\sup_{h\in H^*}c_h$ and the proof is complete.

Next, we show the second result of part (ii) holds with $=$ replaced by $\leq$ even without imposing Assumption R. Using the second result of part (i) with $\lim_{n\to\infty}$ and $=$ replaced by $\liminf_{n\to\infty}$ and $\geq$, respectively, the $\limsup_{n\to\infty}$ of the rejection probability of the ASC-Hyb test is less than or equal to that of the SC-Hyb test and the latter equals α by Theorem 1.

To show that the second result of part (i) holds, it remains to show that it holds with $=$ replaced by $\leq$. First suppose that H^* is empty. Then $\kappa^*(\alpha)=-\infty$, $H^*(\delta)$ is empty for $\delta>0$ close to zero by Assumption R, and $\kappa^*(\delta_n,\alpha)=-\infty$ for n sufficiently large. Next, suppose that H^* is nonempty. Then, using Assumption R, it suffices to show that $\limsup_{n\to\infty}\sup_{h\in H^*(\delta_n)}c_h\leq\sup_{h\in H^\dagger}c_h$. As in the last paragraph of the proof of Theorem 4(a) (given above), there exists a sequence $\{h_n\in H^*(\delta_n):n\geq1\}$, a subsequence $\{u_n\}$ of $\{n\}$, and a subsequence $\{v_n\}$ of $\{u_n\}$ such that

$$\begin{aligned}&\limsup_{n\to\infty}\sup_{h\in H^*(\delta_n)}c_h=\lim_{n\to\infty}c_{h_{v_n}},\\&\left(h_{v_n,1},h_{v_n,2}\right)\to(h_1^*,h_2^*)=h^*,\\&\left(\delta_{v_n}^r h_{v_n,1},h_{v_n,2}\right)\to(g_1^*,h_2^*)=g^*\end{aligned} \tag{S2.22}$$

for some $(g^*,h^*)\in GH$. Since $h_{v_n}=(h_{v_n,1},h_{v_n,2})\in H^*(\delta_{v_n})$ for all n, we have $h^*\in H^\dagger$ by definition of $H^\dagger$. This, (S2.22), and Assumption Q yield

$$\limsup_{n\to\infty}\sup_{h\in H^*(\delta_n)}c_h=\lim_{n\to\infty}c_{h_{v_n}}=c_{h^*}\leq\sup_{h\in H^\dagger}c_h, \tag{S2.23}$$

which completes the proof of the second result of part (i). Given this, by the same argument as used to prove Theorem 1(ii) of AG1 with $c_{n,b}$ replaced by $\max\{c_{n,b}, c_\infty + \kappa^*(\delta_n, \alpha)\}$, $\text{AsySz}(\theta_0)$ for the ASC-Hyb test is equal to $\text{AsySz}(\theta_0)$ for the SC-Hyb test. By Theorem 2, the latter equals α. Hence, the second result of part (ii) holds. *Q.E.D.*

THEOREM 4(c): *Suppose Assumptions* A–G, K, L, N, O, Q, *and* S *hold*. *Then* (i) $\kappa_{\widehat{\gamma}_{n,2}}(\delta_n, \alpha) - \kappa_{\gamma_{n,2}}(\alpha) \to_p 0$ *and* $\kappa^*_{\widehat{\gamma}_{n,2}}(\delta_n, \alpha) - \kappa^*_{\gamma_{n,2}}(\alpha) \to_p 0$ *under all sequences* $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) \in \Gamma : n \geq 1\}$ *and* (ii) *the APSC-Sub and APSC-Hyb tests satisfy* $\text{AsySz}(\theta_0) = \alpha$.

PROOF: By Theorem 3, in part (a) it suffices to show that $\kappa_{\widehat{\gamma}_{n,2}}(\delta_n, \alpha) - \kappa_{\widehat{\gamma}_{n,2}}(\alpha) \to_p 0$ and $\kappa^*_{\widehat{\gamma}_{n,2}}(\delta_n, \alpha) - \kappa^*_{\widehat{\gamma}_{n,2}}(\alpha) \to_p 0$. To do so, we use the result that a sequence of random variables $\{X_n : n \geq 1\}$ satisfies $X_n \to_p 0$ if and only if for every subsequence $\{u_n\}$ of $\{n\}$ there is a subsequence $\{v_n\}$ of $\{u_n\}$ such that $X_{v_n} \to 0$ a.s. We apply this result with $X_n = \kappa_{\widehat{\gamma}_{n,2}}(\delta_n, \alpha) - \kappa_{\widehat{\gamma}_{n,2}}(\alpha)$. Hence, it suffices to show that given any $\{u_n\}$ there exists a subsequence $\{v_n\}$ of $\{u_n\}$ such that $X_{v_n} \to 0$ a.s. Given $\{u_n\}$, we apply the above subsequence result a second time with $X_n = \widehat{\gamma}_{n,2} - \gamma_{n,2}$ to guarantee that there is a subsequence $\{v_n\}$ of $\{u_n\}$ for which $\widehat{\gamma}_{v_n,2} - \gamma_{v_n,2} \to 0$ a.s. using Assumption N. The subsequence $\{v_n\}$ can be chosen such that $\gamma_{v_n,2} \to h_2$ for some $h_2 \in H_2$ because every sequence in H_2 has a convergent subsequence given that H_2 is closed with respect to R^q_∞. Now, the argument in the proof of Theorem 4(a) applied to the subsequence $\{v_n\}$ with $1 - J_h(c_{(\delta^r_{v_n} h_1, h_2)})$, $1 - J_h(c_g)$, and $h_{v_n,2} = h_2$ replaced by $c_{(h_1, \widehat{\gamma}_{v_n,2})} - c_{(\delta^r_{v_n} h_1, \widehat{\gamma}_{v_n,2})}$, $c_h - c_g$, and $h_{v_n,2} = \widehat{\gamma}_{v_n,2}$, respectively, and using Assumption Q in place of Assumption P gives the desired result.

The second result of part (i) holds using similar subsequence arguments to those above combined with variations of the proof of the second result of part (i) of Theorem 4(b) with H^*, $H^*(\delta_n)$, $H^\dagger$, and Assumption R replaced by $H^*_{h_2}$, $H^*_{\gamma_{n,2}}(\delta_n)$, $H^\dagger_{h_2}$, and Assumption S, respectively.

Given the results of part (i), part (ii) is proved using the same argument as used to prove part (ii) of Theorem 4(b). *Q.E.D.*

S3. SIZE CORRECTION BY QUANTILE ADJUSTMENT

We now briefly discuss SC methods based on quantile adjustment, as opposed to the method in Section 3 of the paper. Quantile-adjusted SC-Sub and SC-Hyb tests with nominal level α reject the null hypothesis $H_0 : \theta = \theta_0$ when

$$\text{(S3.1)} \qquad T_n(\theta_0) > c_{n,b}(1 - \xi(\alpha)),$$

$$T_n(\theta_0) > c^*_{n,b}(1 - \xi^*(\alpha)),$$

respectively, where $\xi(\alpha)$ ($\in (0,\alpha]$) and $\xi^*(\alpha)$ ($\in (0,\alpha]$) are the largest constants[2] that satisfy

$$\text{(S3.2)}\qquad \sup_{(g,h)\in GH}\big(1-J_h\big(c_g(1-\xi(\alpha))-\big)\big)\le\alpha,$$
$$\sup_{(g,h)\in GH}\big(1-J_h\big(\max\big\{c_g(1-\xi^*(\alpha)),c_\infty(1-\xi^*(\alpha))\big\}-\big)\big)\le\alpha.$$

In many cases, the quantile adjustment and the size-correction method of Section 3 give similar results. For many examples, we prefer the method based on (S2.5) and (S2.6) to that of (S3.1) and (S3.2) because the former are based on the explicit formulae for the adjustment factors $\kappa(\alpha)$ and $\kappa^*(\alpha)$ given in (S2.6).

S4. POWER COMPARISONS OF SIZE-CORRECTED TESTS

We now provide some results concerning power comparisons of SC tests that are referred to in Section 3.2. We consider three alternative assumptions concerning the shape of $c_h(1-\alpha)$. ("Quant" refers to quantile.)

ASSUMPTION Quant1: *$c_g(1-\alpha)\ge c_h(1-\alpha)$ for all $(g,h)\in GH$.*

ASSUMPTION Quant2: *$c_g(1-\alpha)\le c_h(1-\alpha)$ for all $(g,h)\in GH$ with strict inequality for some (g,h).*

ASSUMPTION Quant3: (i) *$H=H_1=R_{+,\infty}$,* (ii) *$c_h(1-\alpha)$ is uniquely maximized at $h^*\in(0,\infty)$, and* (iii) *$c_h(1-\alpha)$ is minimized at $h=0$ or $h=\infty$.*

THEOREM S1: *Suppose Assumptions* K *and* L *hold.*

(a) *Suppose Assumption Quant1 holds. Then* (i) $cv(1-\alpha)=\sup_{h_2\in H_2}c_{(0,h_2)}(1-\alpha)$, (ii) $\kappa(\alpha)=0$, (iii) $\kappa^*(\alpha)=-\infty$, (iv) $\max\{c_g(1-\alpha),c_\infty(1-\alpha)+\kappa^*(\alpha)\}=c_g(1-\alpha)+\kappa(\alpha)$, *and* (v) $c_g(1-\alpha)+\kappa(\alpha)\le cv(1-\alpha)$ *for all $g\in H$.*

(b) *Suppose Assumption Quant2 holds. Then* (i) $cv(1-\alpha)=c_\infty(1-\alpha)$, (ii) $\kappa^*(\alpha)=0$, (iii) $\max\{c_g(1-\alpha),c_\infty(1-\alpha)+\kappa^*(\alpha)\}=cv(1-\alpha)$, *and* (iv) $cv(1-\alpha)\le c_g(1-\alpha)+\kappa(\alpha)$ *for all $g\in H$.*

(c) *Suppose Assumption Quant3 holds. Then* (i) $cv(1-\alpha)=c_{h^*}(1-\alpha)$, (ii) $\kappa(\alpha)=c_{h^*}(1-\alpha)-c_0(1-\alpha)$, (iii) $\kappa^*(\alpha)=c_{h^*}(1-\alpha)-c_\infty(1-\alpha)$, (iv) $\max\{c_g(1-\alpha),c_\infty(1-\alpha)+\kappa^*(\alpha)\}=cv(1-\alpha)$ *for all $g\in H$,* (v) $cv(1-\alpha)\le c_g(1-\alpha)+\kappa(\alpha)$ *for all $g\in H$ such that $c_g(1-\alpha)\ge c_0(1-\alpha)$ (such as $g=h^*$), and likewise with strict inequalities, and* (vi) $cv(1-\alpha)>c_g(1-\alpha)+\kappa(\alpha)$ *for all $g\in H$ such that $c_g(1-\alpha)<c_0(1-\alpha)$ (there is no such $g\in H$ if $c_h(1-\alpha)$ is minimized at $h=0$).*

[2]If no such largest value exists, we take some value that is arbitrarily close to the supremum of the values that satisfy (S3.2).

COMMENTS: (i) In this comment and the next, we assume Assumption M holds, so that Theorem 2 holds. Theorem S1(a) shows that the subsampling and hybrid tests have correct asymptotic size under Assumption Quant1 and they have critical values less than or equal to that of the SC-FCV test. Theorem S1(b) shows that the FCV and hybrid tests have correct asymptotic size under Assumption Quant2 and they have critical values less than or equal to that of the SC-Sub test. If Assumption Quant1 (Quant2) holds with a strict inequality for $(g, h) = (h^0, h)$ for some $h = (h_1, h_2) \in H$, where $h^0 = (0, h_2) \in H$, then Theorem S1(a)(v) (respectively, (b)(iv)) holds with a strict inequality with g equal to this value of h.

(ii) Theorem S1(c)(iv) and (v) shows that under Assumption Quant3 the SC-Hyb and SC-FCV tests are asymptotically equivalent and are always more powerful than the SC-Sub test at some $(g, h) \in GH$. On the other hand, Theorem S1(c)(vi) shows that under Assumption Quant3 the SC-Sub test can be more powerful than the SC-Hyb and SC-FCV tests at some $(g, h) \in GH$ though not if $c_h(1 - \alpha)$ is minimized at $h = 0$.

The results above are relevant when the subsample statistics satisfy Assumption Sub1 (because then their asymptotic distribution typically is the same under the null and the alternative). On the other hand, if Assumption Sub2 holds, then the subsampling critical values typically diverge to infinity under fixed alternatives (at rate $b^{1/2} \ll n^{1/2}$ when $T_n(\theta_0)$ is a t statistic). For brevity, we do not investigate the relative magnitudes of the critical values of the SC-FCV, SC-Sub, and SC-Hyb tests for local alternatives when Assumption Sub2 holds.

In Section S8 below, we introduce a SC combined method that has power at least as good as that of the SC subsampling and hybrid tests, but it reduces to the SC hybrid test in most examples and, hence, may be of more interest theoretically than practically.

PROOF OF THEOREM S1: Assumption L guarantees that cv, $\kappa(\alpha)$, and $\kappa^*(\alpha)$ are well defined. Part (a)(i) follows from the definition of cv in (S2.6) and Assumption Quant1. Part (a)(ii) holds by definition of $\kappa(\alpha)$ in (S2.6) and the fact that $c_h - c_g \leq 0$ for all $(g, h) \in GH$ by Assumption Quant1 (with equality for some $(g, h) \in GH$). Part (a)(iii) holds by the definition of $\kappa^*(\alpha)$ in (S2.6) for the case where H^* is empty, because H^* is empty by Assumption Quant1. Part (a)(iv) follows from parts (a)(ii) and (a)(iii). Part (a)(v) follows from part (a)(ii) and the definition of cv in (S2.6).

Next, we prove part (b)(i). Given any $g = (g_1, g_2) = (g_{1,1}, \ldots, g_{1,p}, g_2) \in H$, let $g^\infty = (g_1^\infty, g_2) = (g_{1,1}^\infty, \ldots, g_{1,p}^\infty, g_2) \in H$ be such that $g_{1,m}^\infty = +\infty$ if $g_{1,m} > 0$, $g_{1,m}^\infty = -\infty$ if $g_{1,m} < 0$, and $g_{1,m}^\infty = +\infty$ or $-\infty$ (chosen so that $g^\infty \in H$) if $g_{1,m} = 0$ for $m = 1, \ldots, p$. By Assumption Quant2, $c_g \leq c_{g^\infty}$ because $(g, g^\infty) \in GH$. By Assumption K, $c_{g^\infty} = c_\infty$ for all $g \in H$. Hence, $cv = \sup_{h \in H} c_h = c_\infty$, which proves part (b)(i).

We now prove part (b)(ii). By Assumptions Quant2 and K, H^* is not empty and $\sup_{h\in H^*} c_h = c_\infty$. In consequence, $\kappa^*(\alpha) = 0$ by definition of $\kappa^*(\alpha)$ in (S2.6). Part (b)(iii) follows from parts (b)(i) and (b)(ii), and $c_g \le c_\infty$ follows by Assumptions Quant2 and K. We now prove part (b)(iv). By part (b)(i), it suffices to show that $c_g + \kappa(\alpha) \ge c_\infty$ for all $g \in H$. By the definition of $\kappa(\alpha)$ in (S2.6) and Assumptions Quant2 and K, $\kappa(\alpha) = c_\infty - \inf_{h_2\in H_2} c_{(0,h_2)}$. Hence, $c_g + \kappa(\alpha) = c_g + c_\infty - \inf_{h_2\in H_2} c_{(0,h_2)} \ge c_\infty$, where the inequality uses Assumption Quant2. This establishes part (b)(iv).

Part (c)(i) holds by Assumption Quant3(ii). Part (c)(ii) holds by definition of $\kappa(\alpha)$ in (S2.6) and Assumption Quant3(ii) and (iii). Part (c)(iii) holds by definition of $\kappa^*(\alpha)$ in (S2.6) and Assumption Quant3(ii). Part (c)(iv) holds because $\max\{c_g, c_\infty + \kappa^*(\alpha)\} = \max\{c_g, c_{h^*}\} = c_{h^*} = cv$ using parts (c)(i) and (c)(iii). Parts (c)(v) and (c)(vi) hold because $cv = c_{h^*}$ by part (c)(i) and $c_g + \kappa(\alpha) = c_{h^*} + c_g - c_0$ by part (c)(ii). *Q.E.D.*

S5. CRITICAL VALUE FUNCTIONS

In this section, we use graphs given in Figure S-1 to illustrate the asymptotic critical value (c.v.) functions of the hybrid, FCV, and subsampling tests for the case where $\gamma = \gamma_1 \in R_+$, (i.e., no subvectors γ_2 or γ_3 appear, $p = 1$, and $H = R_{+,\infty}$). The argument of the cv functions is $g \in H$. For example, the asymptotic subsampling c.v. function is $c_g(1-\alpha)$ for $g \in H$. In Figure S-1, the curved line is the subsampling c.v. function, the horizontal line is the FCV c.v. function (i.e., the constant $c_\infty(1-\alpha)$), and the hybrid c.v. function is the maximum of the two.

In Figure S-1(a), the subsampling and hybrid c.v. functions are the same and the corresponding tests have the desired asymptotic size α. (The latter holds because $c_\infty(1-\alpha)$ is less than or equal to the c.v. function at g for all $g \in R_+$, $c_0(1-\alpha)$ is greater than or equal to the c.v. function at g for all $g \in R_+$, and these two conditions are necessary and sufficient for a test to have asymptotic size α, assuming continuity of $J_h(\cdot)$ by Theorem 1 of AG1.) On the other hand, in Figure S-1(a), the FCV test has asymptotic size greater than α. In Figure S-1(b) and (d), the hybrid c.v. function equals the FCV c.v. function; both of these tests have asymptotic size α, whereas the subsampling test has asymptotic size greater than α. Figure S-1(a) and (b) illustrate the results of Lemma 1(i) and (ii), respectively.

Figure S-1(c) illustrates a case where the hybrid test has asymptotic size α, but both the FCV and subsampling tests have asymptotic size greater than α. In Figure S-1(a)–(d), Assumption Quant holds, so the hybrid test has correct asymptotic size, as established in Lemma 1.

Figure S-1(e) and (f) illustrate cases in which the function $c_g(1-\alpha)$ is maximized at an interior point $g \in (0, \infty)$. In these cases, the hybrid, FCV, and subsampling tests all have asymptotic size greater than α. Figure S-1(e) and (f) illustrate the results of Lemma 1(ii) and (i), respectively. In particular, in

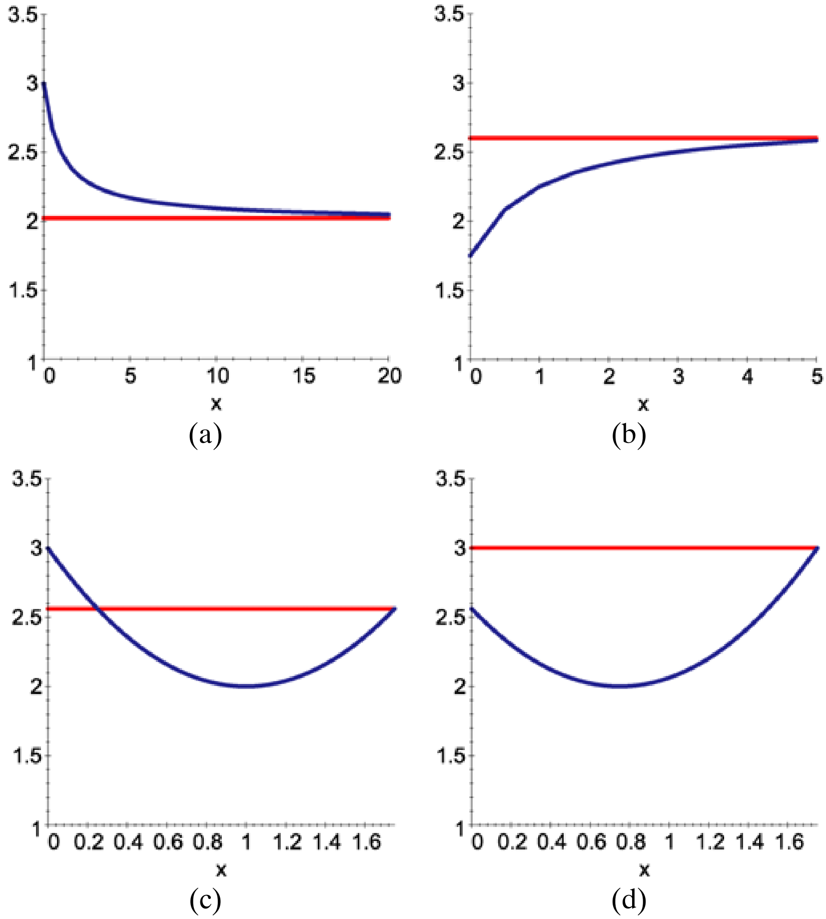


FIGURE S-1.—Hybrid, FCV, and subsample critical values as a function of $g \in H$: hybrid = max{curved line, horizontal line}; FCV = horizontal line; subsample = curved line.

Figure S-1(e), the over-rejection of the subsampling test for g close to zero is reduced for the hybrid test because its c.v. function is larger.

S6. GRAPHICAL POWER COMPARISONS

Next, we use graphs given in Figure S-2 to illustrate the power comparison between SC-FCV, SC-Sub, and SC-Hyb tests. Theorem S1 shows that (a) if $c_g(1-\alpha) \geq c_h(1-\alpha)$ for all $(g, h) \in GH$, then the SC-Sub, SC-Hyb, Sub, and Hyb tests are equivalent asymptotically and are more powerful than the SC-FCV test (see Figure S-2(a)); (b) if $c_g(1-\alpha) \leq c_h(1-\alpha)$ for all $(g, h) \in GH$, then the SC-FCV, SC-Hyb, FCV, and Hyb tests are equivalent asymptotically and are more powerful than the SC-Sub test (see Figure S-2(b)); and (c) if $H =$

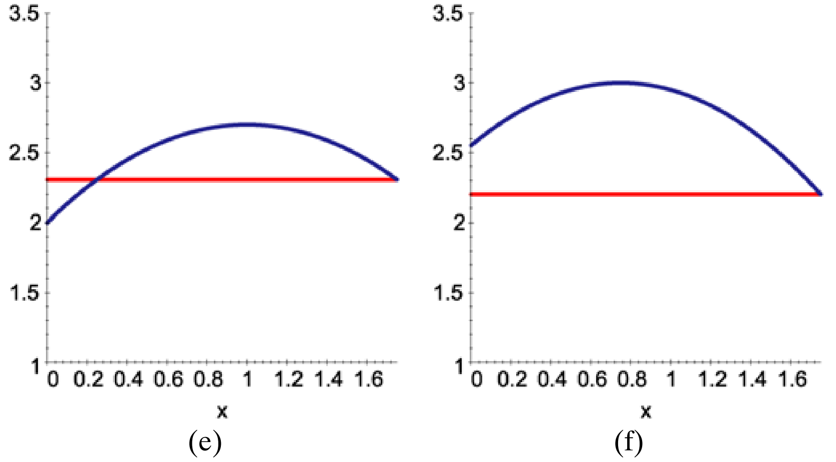


FIGURE S-1.—(*Continued*.)

$H_1 = R_{+,\infty}$ and $c_h(1-\alpha)$ is uniquely maximized at $h^* \in (0,\infty)$, then the SC-FCV and SC-Hyb tests are asymptotically equivalent and are either (i) more powerful than the SC-Sub test (see Figure S-2(e)) or (ii) more powerful than the SC-Sub test for some values of $(g,h) \in GH$ but less powerful for other values of $(g,h) \in GH$ (see Figure S-2(f)).

Figure S-2(c) illustrates the case where $c_g(1-\alpha)$ is not monotone but is maximized at $g=0$, the Hyb and SC-Hyb c.v. functions are the same, the Hyb c.v. function is lower than both the SC-Sub and SC-FCV c.v. functions, and so the Hyb test is more powerful than the SC-Sub and SC-FCV tests. Figure S-2(d) illustrates the case where $c_g(1-\alpha)$ is not monotone but is maximized at $g=\infty$, the Hyb, SC-Hyb, FCV, and SC-FCV c.v. functions are the same, the Hyb c.v. function is lower than the SC-Sub c.v. function, and so the Hyb test is more powerful than the SC-Sub test.

S7. EQUAL-TAILED SIZE-CORRECTED TESTS

This section introduces *equal-tailed* size-corrected FCV, subsampling, and hybrid t tests. It also introduces finite-sample adjusted asymptotics for equal-tailed tests. We suppose that $T_n(\theta_0) = \tau_n(\widehat{\theta}_n - \theta_0)/\widehat{\sigma}_n$.

Let $c_{\text{Fix}}(1-\alpha/2)$ and $c_{\text{Fix}}(\alpha/2)$ denote the critical values of the equal-tailed FCV test before size correction. Equal-tailed (i) SC-FCV, (ii) SC-Sub, and (iii) SC-Hyb tests are defined by (5.1) of the paper with the critical values $c^*_{n,b}(1-\alpha/2)$ and $c^{**}_{n,b}(\alpha/2)$ replaced by (i) $c_{\text{Fix}}(1-\alpha/2) + \kappa_{\text{ET,Fix}}(\alpha)$ and $c_{\text{Fix}}(\alpha/2) - \kappa_{\text{ET,Fix}}(\alpha)$, (ii) $c_{n,b}(1-\alpha/2) + \kappa_{\text{ET}}(\alpha)$ and $c_{n,b}(\alpha/2) - \kappa_{\text{ET}}(\alpha)$, and (iii) $\max\{c_{n,b}(1-\alpha/2), c_\infty(1-\alpha/2) + \kappa^*_{\text{ET}}(\alpha)\}$ and $\min\{c_{n,b}(\alpha/2), c_\infty(\alpha/2) - \kappa^*_{\text{ET}}(\alpha)\}$, respectively.

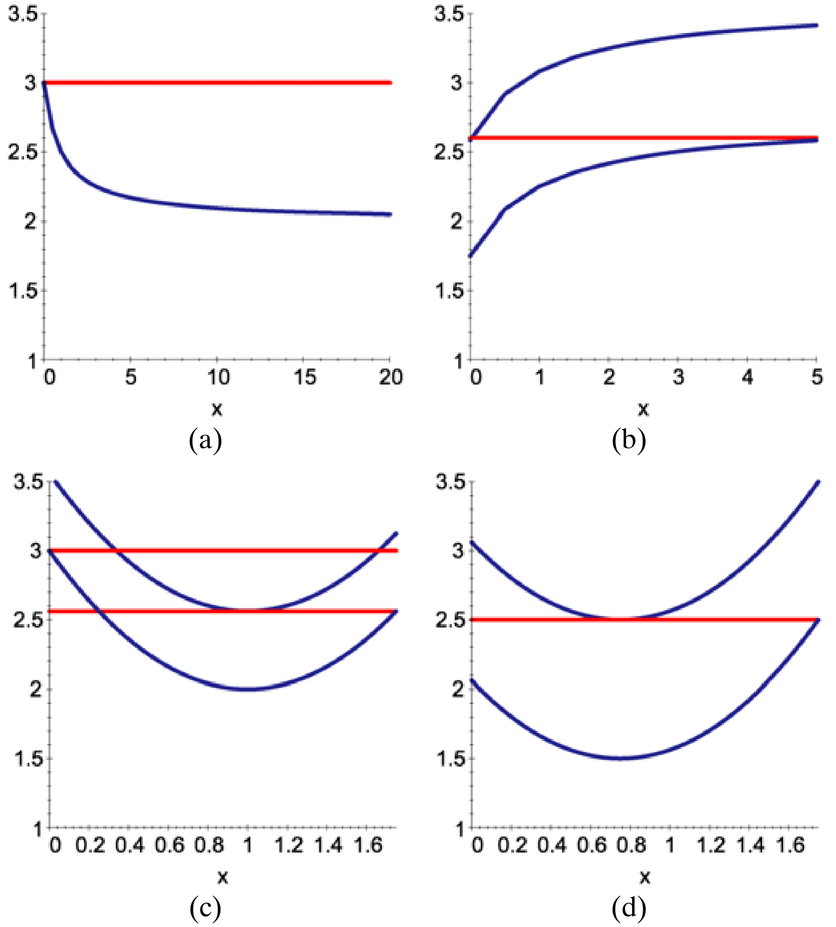
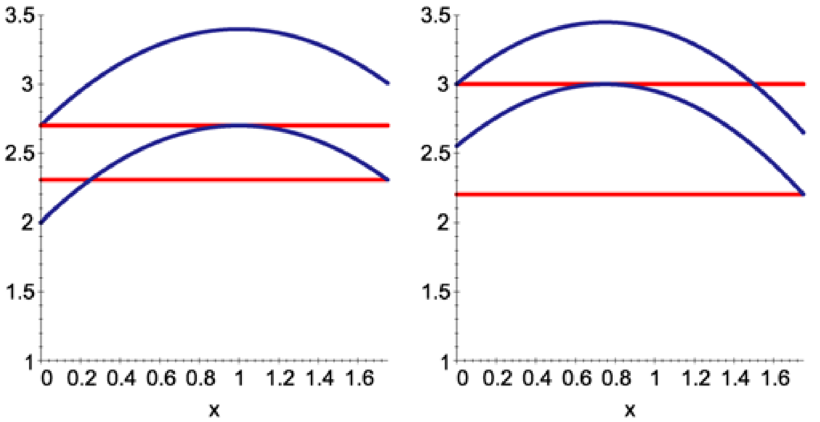


FIGURE S-2.—Critical values as a function of $g \in H$ for SC-Sub, SC-FCV, and SC-Hyb tests. In each panel the lower curve is $c_g(1-\alpha)$ and the lower horizontal is the FCV critical value. (a) Curve: SC-Sub and SC-Hyb, horizontal: SC-FCV. (b) Horizontal: SC-Hyb and SC-FCV, upper curve: SC-Sub. (c) Max{Lower Horizontal, Lower Curve}: SC-Hyb and SC-FCV, upper curve: SC-Sub, upper horizontal: SC-FCV. (d) Horizontal: SC-Hyb and SC-FCV, upper curve: SC-Sub. (e) Upper horizontal: SC-Hyb and SC-FCV, upper curve: SC-Sub. (f) Upper horizontal: SC-Hyb and SC-FCV, upper curve: SC-Sub.

By definition, the SC factors $\kappa_{\mathrm{ET,Fix}}(\alpha)$ $(\in [0,\infty))$, $\kappa_{\mathrm{ET}}(\alpha)$ $(\in [0,\infty))$, and $\kappa^*_{\mathrm{ET}}(\alpha)$ $(\in \{-\infty\} \cup [0,\infty))$, respectively, are the smallest values that satisfy

$$\text{(S7.1)} \qquad \sup_{h\in H}\Big[1 - J_h\big((c_{\mathrm{Fix}}(1-\alpha/2) + \kappa_{\mathrm{ET,Fix}}(\alpha))-\big) + J_h(c_{\mathrm{Fix}}(\alpha/2) - \kappa_{\mathrm{ET,Fix}}(\alpha))\Big] \le \alpha,$$



(e) (f)

FIGURE S-2.—(*Continued*.)

$$\sup_{(g,h)\in GH}\big[1-J_h\big((c_g(1-\alpha/2)+\kappa_{\mathrm{ET}}(\alpha))-\big)$$
$$+J_h(c_g(\alpha/2)-\kappa_{\mathrm{ET}}(\alpha))\big]\leq\alpha,$$
$$\sup_{(g,h)\in GH}\big[1-J_h\big(\max\{c_g(1-\alpha/2),c_\infty(1-\alpha/2)+\kappa^*_{\mathrm{ET}}(\alpha)\}-\big)$$
$$+J_h\big(\min\{c_g(\alpha/2),c_\infty(\alpha/2)-\kappa^*_{\mathrm{ET}}(\alpha)\}\big)\big]\leq\alpha.$$

(If no such smallest value exists, we take some value that is arbitrarily close to the infimum. If no value that satisfies (S7.1) exists, then size correction is not possible.)

For each test, the condition in (S7.1) guarantees that the "overall" asymptotic size of the test is less than or equal to α. It does not guarantee that the maximum (asymptotic) rejection probability in each tail is less than or equal to $\alpha/2$. If the latter is desired, then one should size-correct the lower and upper critical values of the equal-tailed test in the same way as one-sided t tests are size-corrected in the paper. (This can yield the overall size of the test to be strictly less than α if the (g,h) vector that maximizes the rejection probability is different for the lower and upper critical values.)

Given SC factors that satisfy (S7.1), the equal-tailed SC-FCV, SC-Sub, and SC-Hyb t tests have $\mathrm{AsySz}(\theta_0)\leq\alpha$ under Assumptions A–E, G, and J by the proofs of Corollary 2 of AG1 and Corollary 1 of the paper. (Only Assumptions A and B are needed for the SC-FCV tests.) Under continuity conditions on $J_h(x)$ at suitable (h,x) such that the inequalities in (S7.1) hold as equalities, these tests have $\mathrm{AsySz}(\theta_0)=\alpha$.

An alternative way to size-correct equal-tailed tests is the following method. This method has the advantage that if it is possible to produce an *equal-tailed* size-corrected test, then the procedure does so. Its disadvantage is that it is somewhat more complicated to implement.

First, let $\kappa_{\text{ET,Fix},1}(\alpha) \in [0,\infty)$, $\kappa_{\text{ET},1}(\alpha) \in [0,\infty)$, and $\kappa^*_{\text{ET},1}(\alpha) \in \{-\infty\} \cup [0,\infty)$ denote the smallest values such that

$$\text{(S7.2)} \quad \sup_{h\in H}\big[1 - J_h\big((c_{\text{Fix}}(1-\alpha/2) + \kappa_{\text{ET,Fix},1}(\alpha))-\big)\big] \le \alpha/2,$$
$$\sup_{(g,h)\in GH}\big(1 - J_h\big((c_g(1-\alpha/2) + \kappa_{\text{ET},1}(\alpha))-\big)\big) \le \alpha/2,$$
$$\sup_{(g,h)\in GH}\big(1 - J_h\big(\max\{c_g(1-\alpha/2), c_\infty(1-\alpha/2) + \kappa^*_{\text{ET},1}(\alpha)\}-\big)\big) \le \alpha/2.$$

Next, let $\kappa_{\text{ET,Fix},2}(\alpha) \in R$, $\kappa_{\text{ET},2}(\alpha) \in R$, and $\kappa^*_{\text{ET},2}(\alpha) \in \{-\infty\} \cup R$ denote the smallest values such that

$$\text{(S7.3)} \quad \sup_{h\in H}\big[1 - J_h\big((c_{\text{Fix}}(1-\alpha/2) + \kappa_{\text{ET,Fix},1}(\alpha))-\big) + J_h(c_{\text{Fix}}(\alpha/2) - \kappa_{\text{ET,Fix},2}(\alpha))\big] \le \alpha,$$
$$\sup_{(g,h)\in GH}\big[1 - J_h\big((c_g(1-\alpha/2) + \kappa_{\text{ET},1}(\alpha))-\big) + J_h(c_g(\alpha/2) - \kappa_{\text{ET},2}(\alpha))\big] \le \alpha,$$
$$\sup_{(g,h)\in GH}\big[1 - J_h\big(\max\{c_g(1-\alpha/2), c_\infty(1-\alpha/2) + \kappa^*_{\text{ET},1}(\alpha)\}-\big) + J_h\big(\min\{c_g(\alpha/2), c_\infty(\alpha/2) - \kappa^*_{\text{ET},2}(\alpha)\}\big)\big] \le \alpha.$$

The "alternative" SC equal-tailed FCV test rejects H_0 if $T_n(\theta_0) > c_{\text{Fix}}(1-\alpha/2) + \kappa_{\text{ET,Fix},1}(\alpha)$ or $T_n(\theta_0) < c_{\text{Fix}}(\alpha/2) - \kappa_{\text{ET,Fix},2}(\alpha)$. The alternative SC equal-tailed Sub and Hyb tests are defined analogously. We use alternative SC equal-tailed tests in the parameter of interest near a boundary example in Andrews and Guggenberger (2010). For all of the other examples, we use the SC equal-tailed tests defined in (S7.1).

If a parameter γ_2 appears in γ and γ_2 is consistently estimable, then PSC tests are more powerful asymptotically than SC tests (because they lead to smaller critical values under some distributions but still have the correct asymptotic size). Equal-tailed (i) PSC-FCV, (ii) PSC-Sub, and (iii) PSC-Hyb tests are defined as the SC versions are defined above, but with $\kappa_{\text{ET,Fix}}(\alpha)$, $\kappa_{\text{ET}}(\alpha)$, and $\kappa^*_{\text{ET}}(\alpha)$ replaced by $\kappa_{\text{ET,Fix},\widehat{\gamma}_{n,2}}(\alpha)$, $\kappa_{\text{ET},\widehat{\gamma}_{n,2}}(\alpha)$, and $\kappa^*_{\text{ET},\widehat{\gamma}_{n,2}}(\alpha)$, respectively. Here, the PSC factors $\kappa_{\text{ET,Fix},h_2}(\alpha)$ $(\in [0,\infty))$, $\kappa_{\text{ET},h_2}(\alpha)$ $(\in [0,\infty))$, and $\kappa^*_{\text{ET},h_2}(\alpha)$ $(\in \{-\infty\} \cup [0,\infty))$ are defined to be the smallest values that satisfy

$$(S7.4)\quad \sup_{h_1\in H_1}\big[1-J_{(h_1,h_2)}\big(\big(c_{\text{Fix}}(1-\alpha/2)+\kappa_{\text{ET,Fix},h_2}(\alpha)\big)-\big)$$
$$+J_{(h_1,h_2)}\big(c_{\text{Fix}}(\alpha/2)-\kappa_{\text{ET,Fix},h_2}(\alpha)\big)\big]\le\alpha,$$
$$\sup_{g_1,h_1\in H_1:((g_1,h_2),(h_1,h_2))\in GH}\big[1-J_{(h_1,h_2)}\big(\big(c_{(g_1,h_2)}(1-\alpha/2)+\kappa_{\text{ET},h_2}(\alpha)\big)-\big)$$
$$+J_{(h_1,h_2)}\big(c_{(g_1,h_2)}(\alpha/2)-\kappa_{\text{ET},h_2}(\alpha)\big)\big]\le\alpha,$$
$$\sup_{g_1,h_1\in H_1:((g_1,h_2),(h_1,h_2))\in GH}\big[1$$
$$-J_{(h_1,h_2)}\big(\max\big\{c_{(g_1,h_2)}(1-\alpha/2),c_\infty(1-\alpha/2)+\kappa^*_{\text{ET},h_2}(\alpha)\big\}-\big)$$
$$+J_{(h_1,h_2)}\big(\min\big\{c_{(g_1,h_2)}(\alpha/2),c_\infty(\alpha/2)-\kappa^*_{\text{ET},h_2}(\alpha)\big\}\big)\big]\le\alpha.$$

The (i) PSC-FCV, (ii) PSC-Sub, and (iii) PSC-Hyb equal-tailed tests all have $\text{AsySz}(\theta_0)\le\alpha$ under Assumptions N plus (i) Assumptions A and B, (ii) Assumptions A–E, G, and J, and (iii) Assumptions A–E, G, J, and K, respectively. (The proof is analogous to the proof of Theorem 3 combined with the proof of Theorem 2.) These tests have $\text{AsySz}(\theta_0)=\alpha$ provided the inequalities in (S7.4) hold as equalities.

The finite-sample adjustments introduced in Section 4 of the paper do not cover equal-tailed tests. For equal-tailed subsampling tests, we define the following finite-sample adjustment to $\text{AsySz}(\theta_0)$:

$$(S7.5)\quad \text{AsySz}_n(\theta_0)=\sup_{h\in H}\big[1-J_h\big(c_{(\delta_n^r h_1,h_2)}(1-\alpha/2)-\big)+J_h\big(c_{(\delta_n^r h_1,h_2)}(\alpha/2)\big)\big].$$

Define $\text{Max}^{r-}_{\text{ET,Sub}}(\alpha)$ as $\text{Max}_{\text{ET,Hyb}}(\alpha)$ is defined in (5.2) of the paper with $c_g^*(1-\alpha/2)$ and $c_g^{**}(\alpha/2)$ replaced by $c_g(1-\alpha/2)$ and $c_g(\alpha/2)-$, respectively, where "$-$" indicates the limit from the left. Define $\text{Max}^{\ell-}_{\text{ET,Sub}}(\alpha)$ analogously with $c_g^*(1-\alpha/2)$ and $c_g^{**}(\alpha/2)$ replaced by $c_g(1-\alpha/2)-$ and $c_g(\alpha/2)$. With the function that appears in Assumption P(i) altered to $(g,h)\to J_h(c_g(1-\alpha/2)-)-J_h(c_g(\alpha/2))$ and with $\text{Max}^{r-}_{\text{ET,Sub}}(\alpha)=\text{Max}^{\ell-}_{\text{ET,Sub}}(\alpha)$ in place of Assumption P(ii), the result of Theorem 4(a) (viz., $\text{AsySz}_n(\theta_0)\to\text{AsySz}(\theta_0)$) holds for equal-tailed subsampling tests. An analogous result holds for equal-tailed hybrid tests.

Based on (S7.5), we introduce finite-sample adjustments that can improve the asymptotic approximations upon which the equal-tailed SC and PSC subsampling and hybrid tests rely. Equal-tailed ASC and APSC subsampling and hybrid tests are defined just as SC and PSC subsampling and hybrid tests are defined, but using $\kappa_{\text{ET}}(\delta_n,\alpha)$, $\kappa^*_{\text{ET}}(\delta_n,\alpha)$, $\kappa_{\text{ET},\widehat{\gamma}_{n,2}}(\delta_n,\alpha)$, and $\kappa^*_{\text{ET},\widehat{\gamma}_{n,2}}(\delta_n,\alpha)$ in place of $\kappa_{\text{ET}}(\alpha)$ and $\kappa^*_{\text{ET}}(\alpha)$. The ASC factors $\kappa_{\text{ET}}(\delta,\alpha)$ ($\in[0,\infty)$) and

$\kappa^*_{\mathrm{ET}}(\delta,\alpha)$ ($\in \{-\infty\} \cup [0,\infty)$) are defined to be the smallest values that satisfy

$$\begin{aligned}
\text{(S7.6)} \quad & \sup_{(h_1,h_2)\in H} \bigl[1 - J_{(h_1,h_2)}\bigl(\bigl(c_{(\delta^r h_1,h_2)}(1-\alpha/2) + \kappa_{\mathrm{ET}}(\delta,\alpha)\bigr)-\bigr) \\
& \quad + J_{(h_1,h_2)}\bigl(c_{(\delta^r h_1,h_2)}(\alpha/2) - \kappa_{\mathrm{ET}}(\delta,\alpha)\bigr)\bigr] \le \alpha, \\
& \sup_{(h_1,h_2)\in H} \bigl[1 \\
& \quad - J_{(h_1,h_2)}\bigl(\max\bigl\{c_{(\delta^r h_1,h_2)}(1-\alpha/2), c_\infty(1-\alpha/2) + \kappa^*_{\mathrm{ET}}(\delta,\alpha)\bigr\}-\bigr) \\
& \quad + J_{(h_1,h_2)}\bigl(\min\bigl\{c_{(\delta^r h_1,h_2)}(\alpha/2), c_\infty(\alpha/2) - \kappa^*_{\mathrm{ET}}(\delta,\alpha)\bigr\}\bigr)\bigr] \le \alpha.
\end{aligned}$$

The APSC factors $\kappa_{\mathrm{ET},h_2}(\delta,\alpha)$ ($\in [0,\infty)$) and $\kappa^*_{\mathrm{ET},h_2}(\delta,\alpha)$ ($\in \{-\infty\} \cup [0,\infty)$) are defined to be the smallest values that satisfy

$$\begin{aligned}
\text{(S7.7)} \quad & \sup_{h_1\in H_1} \bigl[1 - J_{(h_1,h_2)}\bigl(\bigl(c_{(\delta^r h_1,h_2)}(1-\alpha/2) + \kappa_{\mathrm{ET},h_2}(\delta,\alpha)\bigr)-\bigr) \\
& \quad + J_{(h_1,h_2)}\bigl(c_{(\delta^r h_1,h_2)}(\alpha/2) - \kappa_{\mathrm{ET},h_2}(\delta,\alpha)\bigr)\bigr] \le \alpha, \\
& \sup_{h_1\in H_1} \bigl[1 \\
& \quad - J_{(h_1,h_2)}\bigl(\max\bigl\{c_{(\delta^r h_1,h_2)}(1-\alpha/2), c_\infty(1-\alpha/2) + \kappa^*_{\mathrm{ET},h_2}(\delta,\alpha)\bigr\}-\bigr) \\
& \quad + J_{(h_1,h_2)}\bigl(\min\bigl\{c_{(\delta^r h_1,h_2)}(\alpha/2), c_\infty(\alpha/2) - \kappa^*_{\mathrm{ET},h_2}(\delta,\alpha)\bigr\}\bigr)\bigr] \le \alpha.
\end{aligned}$$

The ASC and APSC tests have $\mathrm{AsySz}(\theta_0) = \alpha$ under conditions that are similar to those given in Section 4 of the paper. For brevity, we do not give details.

S8. SIZE-CORRECTED COMBINED TEST

Theorem S1(c)(iv)–(vi) and Figure S-2(f) show that in some contexts the SC-Hyb test can be more powerful than the SC-Sub test for some $(g,h) \in GH$ and vice versa for other $(g,h) \in GH$. This implies that a test that combines the SC-Hyb and SC-Sub tests can be more powerful than both. In this section, we introduce such a test. It is called the size-corrected combined (SC-Com) test. This test has power advantages over the SC-Hyb and SC-Sub tests in some cases. This is illustrated in Figure S-2(f) where the critical-value function of the SC-Com test is the minimum of the upper horizontal SC-Hyb critical-value function and the upper curved SC-Sub critical-value function. On the other hand, the SC-Com test has computational disadvantages because it requires computation of the critical values for both the SC-Sub and SC-Hyb tests, which requires calculation of $\kappa(\alpha)$ and $\kappa^*(\alpha)$ in cases where both the subsampling and hybrid tests need size correction. Furthermore, in most contexts, the SC-Hyb test is more powerful than the SC-Sub for all $(g,h) \in GH$, so the SC-Com test just reduces to the SC-Hyb test.

The size-corrected combined (SC-Com) test rejects $H_0 : \theta = \theta_0$ when

$$(S8.1)\qquad T_n(\theta_0) > c_{n,\mathrm{Com}}(1-\alpha), \quad \text{where}$$
$$c_{n,\mathrm{Com}}(1-\alpha) = \min\{c_{n,b}(1-\alpha) + \kappa(\alpha), \max\{c_{n,b}(1-\alpha), c_\infty(1-\alpha) + \kappa^*(\alpha)\}\},$$

where the constants $\kappa(\alpha)$ and $\kappa^*(\alpha)$ are defined in (3.2) of the paper.

The following result shows that the SC-Com test has $\mathrm{AsySz}(\theta_0) = \alpha$.

THEOREM S2: *Suppose Assumptions* A–G *and* K–M *hold*. *Then the* SC-Com *test satisfies* $\mathrm{AsySz}(\theta_0) = \alpha$.

COMMENTS: (i) By definition, the critical value, $c_{n,\mathrm{Com}}(1-\alpha)$, of the SC-Com test is less than or equal to those of the SC-Sub and SC-Hyb tests. By (3.2) of the paper, it is less than or equal to that of the SC-FCV test as well. Hence, the SC-Com test is at least as powerful as the SC-Sub, SC-Hyb, and SC-FCV tests.

(ii) A PSC-Com test can be defined as in (S8.1) with $\kappa(\alpha)$ and $\kappa^*(\alpha)$ replaced by $\kappa_{\widehat{\gamma}_{n,2}}(\alpha)$ and $\kappa^*_{\widehat{\gamma}_{n,2}}(\alpha)$.

(iii) An ASC-Com test can be defined as in (S8.1) with $\kappa(\alpha)$ and $\kappa^*(\alpha)$ replaced by $\kappa(\delta_n, \alpha)$ and $\kappa^*(\delta_n, \alpha)$, respectively. Suppose Assumptions A–G, K–M, and Q hold. Then the ASC-Com test satisfies $\mathrm{AsySz}(\theta_0) = \alpha$. This holds by the argument in the proof of Theorem 4(b) (see above) using the results of Theorem 4(b)(i).

(iv) An APSC-Com test can be defined as in (S8.1) with $\kappa(\alpha)$ and $\kappa^*(\alpha)$ replaced by $\kappa_{\widehat{\gamma}_{2,n}}(\delta_n, \alpha)$ and $\kappa^*_{\widehat{\gamma}_{2,n}}(\delta_n, \alpha)$, respectively. Suppose Assumptions A–G, K, L, N, O, R, and Q hold. Then the APSC-Com test satisfies $\mathrm{AsySz}(\theta_0) = \alpha$. This holds by the argument in the proof of Theorem 4(c) (see above) using the results of Theorem 4(c)(i).

PROOF OF THEOREM S2: By the same argument as in the proof of Theorem 1(ii) of AG1, the SC-Com test satisfies

$$(S8.2)\qquad \mathrm{AsySz}(\theta_0) \le \sup_{(g,h)\in GH} \big[1 - J_h\big(\min\{c_g(1-\alpha) + \kappa(\alpha), \max\{c_g(1-\alpha), c_\infty(1-\alpha) + \kappa^*(\alpha)\}\}-\big)\big].$$

By the proof of Theorem 2, the constants $\kappa(\alpha)$ and $\kappa^*(\alpha)$ defined in (3.2) of the paper are such that (S2.7) above holds and hence for all $(g, h) \in GH$,

$$(S8.3)\qquad 1 - J_h(c_g(1-\alpha) + \kappa(\alpha)-) \le \alpha,$$
$$1 - J_h\big(\max\{c_g(1-\alpha), c_\infty(1-\alpha) + \kappa^*(\alpha)\}-\big) \le \alpha.$$

Equations (S8.2) and (S8.3) combine to give $\text{AsySz}(\theta_0) \leq \alpha$.

The SC-Com test has $\text{AsySz}(\theta_0) \geq \alpha$ because its $\text{AsySz}(\theta_0)$ is greater than or equal to that of the SC-Sub test (because its critical value is no larger) and the latter equals α by Theorem 2. *Q.E.D.*

S9. TESTING WHEN A NUISANCE PARAMETER MAY BE NEAR A BOUNDARY

This example is a continuation of an example in AG1. It is a testing problem where a nuisance parameter may be near the boundary of the parameter space under the null hypothesis. The observations are $\{X_i \in R^2 : i \leq n\}$, which are i.i.d. with distribution F, $X_i = (X_{i1}, X_{i2})'$, $E_F X_i = (\theta, \mu)'$, and (X_{i1}, X_{i2}) have correlation ρ. The null hypothesis is $H_0 : \theta = 0$, that is, $\theta_0 = 0$. The parameter space for the nuisance parameter μ is $[0, \infty)$. The test statistic $T_n(\theta_0)$ equals $T_n^*(\theta_0)$, $-T_n^*(\theta_0)$, or $|T_n^*(\theta_0)|$, where $T_n^*(\theta_0)$ is a t statistic based on the Gaussian quasi-maximum likelihood estimator of θ that imposes the restriction that $\mu \in [0, \infty)$ and uses any consistent estimators of the standard deviations and correlation of X_{i1} and X_{i2} in the quasi-likelihood; see AG1 for details. In AG1, Assumptions A–G are verified.

Table S-I reports maximum (over $h_1 = \lim_{n\to\infty} n^{1/2}\mu_{n,h}/\sigma_{n,h,2}$) null rejection probabilities (as percentages) for several fixed values of h_2 $(= \lim_{n\to\infty} \rho_{n,h})$ for hybrid and several other nominal 5% tests.[3] Depending on the column, the probabilities are asymptotic or finite sample. The finite sample results are for the case of $n = 120$ and $b = 12$ with $\widehat{\sigma}_{n1}$, $\widehat{\sigma}_{n2}$, and $\widehat{\rho}_n$ being the sample standard deviations and correlation of X_{i1} and X_{i2}. To dramatically increase computational speed, here and in all of the tables below, finite-sample subsampling and hybrid results are based on $q_n = 119$ subsamples of consecutive observations.[4] Hence, only a small fraction of the "120 choose 12" available subsamples are used. In cases where subsampling and hybrid tests have correct asymptotic size, their finite-sample performance is expected to be better when all available subsamples are used than when only $q_n = 119$ are used. This should be taken into account when assessing the results of the tables. Panels (a), (b), and (c) of Table S-I give results for upper one-sided, symmetric two-sided, and equal-tailed

[3]The finite-sample results in Table S-I are based on 20,000 simulation repetitions. The asymptotic results are based on 50,000 simulation repetitions. For the asymptotic results, the search over h_1 is done with step size .05 on [0, 10] and also includes the value $h_1 = 9{,}999{,}999{,}999$. For the finite-sample results, the search over h_1 is done with step size .001 on [0, .5], with step size .05 on [.5, 1.0], and with step size 1.0 on [1.0, 10]. Calculations indicate that these step sizes are sufficiently small to yield accuracy to within $\pm .1$.

[4]This includes 10 wrap-around subsamples that contain observations at the end and beginning of the sample, for example, observations indexed by $(110, \ldots, 120, 1)$. The choice of $q_n = 119$ subsamples is made because this reduces rounding errors when q_n is small when computing the sample quantiles of the subsample statistics. The values ν_α that solve $\nu_\alpha/(q_n + 1) = \alpha$ for $\alpha =$ .025, .95, and .975 are the integers 3, 114, and 117. In consequence, the .025, .95, and .975 sample quantiles are given by the 3rd, 114th, and 117th largest subsample statistics. See Hall (1992, p. 307) for a discussion of this choice in the context of the bootstrap.

TABLE S-I

NUISANCE PARAMETER NEAR A BOUNDARY EXAMPLE: MAXIMUM (OVER h_1) NULL REJECTION PROBABILITIES ($\times 100$) FOR DIFFERENT VALUES OF THE CORRELATION h_2 FOR VARIOUS NOMINAL 5% TESTS FOR $n = 120$ AND $b = 12$, WHERE THE PROBABILITIES ARE ASYMPTOTIC, FINITE-SAMPLE ADJUSTED ASYMPTOTIC, AND EXACT

	Tests and Probabilities										
	Sub			PSC-Sub	APSC-Sub	FCV		PSC-FCV	Hyb		
h_2	Asy	Adj-Asy	Exact	Exact	Exact	Asy	Exact	Exact	Asy	Adj-Asy	Exact
				(a) Upper One-Sided Tests							
−1.0	50.2	49.5	49.8	4.9	13.5	5.0	5.2	5.1	5.0	5.0	5.2
−.95	33.8	22.9	25.6	5.1	9.0	5.0	5.2	5.1	5.0	5.0	5.2
−.80	20.2	12.1	13.1	3.1	6.2	5.0	5.1	4.9	5.0	5.0	4.7
−.40	8.3	6.5	5.9	4.8	4.6	5.0	4.9	4.8	5.0	5.0	3.7
.00	5.0	5.0	5.0	4.9	4.9	5.0	5.2	5.0	5.0	5.0	3.7
.20	5.0	5.0	4.9	5.2	4.9	5.6	5.7	5.2	5.0	5.0	3.8
.40	5.0	5.0	5.0	5.0	5.0	5.8	5.8	5.0	5.0	5.0	3.8
.60	5.0	5.0	5.3	5.3	5.3	5.6	5.7	5.1	5.0	5.0	3.9
.90	5.0	5.0	4.9	4.9	4.9	5.0	5.0	4.9	5.0	5.0	3.4
1.00	5.0	5.0	4.8	4.9	4.8	5.0	5.0	4.9	5.0	5.0	3.5
Max	50.2	49.5	49.8	5.3	13.5	5.8	5.8	5.2	5.0	5.0	5.2
				(b) Symmetric Two-Sided Tests							
.00	5.0	5.0	5.2	5.1	5.1	5.0	5.4	—	5.0	5.0	3.5
.20	5.2	5.2	5.2	4.9	5.0	5.0	5.3	—	5.0	5.0	3.5
.40	6.0	5.6	5.4	4.5	4.8	5.0	5.2	—	5.0	5.0	3.5
.60	7.5	6.5	6.0	4.0	4.6	5.0	5.3	—	5.0	5.0	3.7
.80	9.6	8.3	6.9	3.7	4.5	5.0	5.2	—	5.0	5.0	3.9
.95	10.1	10.0	8.3	4.2	4.2	5.0	5.7	—	5.0	5.0	4.5
1.00	10.1	10.1	8.4	4.1	4.1	5.0	5.1	—	5.0	5.0	4.2
Max	10.1	10.1	8.4	5.1	5.1	5.0	5.7	—	5.0	5.0	4.5
				(c) Equal-Tailed Two-Sided Tests							
.00	5.0	5.0	5.7	5.5	5.5	5.0	5.4	—	5.0	5.0	3.5
.20	5.4	5.2	5.9	5.4	5.6	5.0	5.3	—	5.0	5.0	3.6
.40	6.7	5.8	6.2	4.5	5.4	5.0	5.2	—	5.0	5.0	3.4
.60	9.9	7.0	7.8	3.9	5.7	5.0	5.3	—	5.0	5.0	3.8
.80	17.3	10.3	12.4	3.2	6.5	5.0	5.2	—	5.0	5.0	4.1
.95	32.4	21.0	24.3	3.5	9.1	5.0	5.7	—	5.0	5.0	4.7
1.00	52.7	51.8	52.7	4.6	13.5	5.0	5.1	—	5.0	5.0	4.2
Max	52.7	51.8	52.7	5.5	13.5	5.0	5.7	—	5.0	5.0	4.7

two-sided tests, respectively. The results for lower one-sided tests are the same as for the upper tests with the sign of h_2 changed (by symmetry) and, hence, are not given. The rows labeled Max give the size (asymptotic or $n = 120$) of the test considered. For brevity, we refer below to the numbers given in the tables as though they are precise, but of course they are subject to simulation error.

S9.1. *FCV, Subsampling, and Hybrid Tests*

Column 2 of Table S-I shows that subsampling tests have very large asymptotic size distortions for upper one-sided and equal-tailed two-sided tests (nominal 5% tests have asymptotic levels 50.2% and 52.7%, respectively), and moderate size distortions for symmetric two-sided tests (the nominal 5% test has asymptotic level 10.1%). Also, column 7 of Table S-I shows that the FCV tests have very small asymptotic size distortions for upper one-sided tests (the nominal 5% test has asymptotic level 5.8%), and no size distortions for symmetric and equal-tailed two-sided tests.

Column 10 of Table S-I shows that the nominal 5% hybrid test has asymptotic size of 5% for upper, symmetric, and equal-tailed tests. So, the hybrid test has correct asymptotic size for all three types of tests in this example.

Finite-sample results for the Sub, FCV, and Hyb tests are given in columns 4, 8, and 12 of Table S-I, respectively. For Hyb tests, the asymptotic approximations are fairly accurate, but tend to overestimate the finite-sample rejection rates somewhat for some values of h_2: finite-sample values vary between 3.4% and 5.2% compared to the asymptotic values of 5.0%. For FCV tests, the asymptotic approximations are found to be very accurate for upper tests and quite accurate for symmetric and equal-tailed tests.

The asymptotic approximations for the Sub test are found to be quite good for h_2 values where the (maximum) asymptotic probabilities equal 5.0%, but for h_2 values where they exceed 5.0%, they tend to overestimate the finite-sample values—sometimes significantly so (e.g., 33.8% versus 25.6% for $h_2 = -.95$ with upper Sub tests). Nevertheless, in the worst case scenarios (i.e., for h_2 values of 1.0 or -1.0, which yield the greatest asymptotic rejection probabilities), the asymptotic approximations are quite good. Hence, the asymptotic sizes and finite-sample sizes are close—50.2% versus 49.8%, 10.1% versus 8.4%, and 52.7% versus 52.7% for upper, symmetric, and equal-tailed tests, respectively.

The results in Table S-I for the columns headed Adj-Asy, PSC-Sub, APSC-Sub are discussed below.

S9.2. *Plug-In Tests*

The upper, symmetric, and equal-tailed subsampling tests and the upper FCV test need size correction in this example. Plug-in size correction is possible because estimation of the correlation parameter ρ is straightforward using the usual sample correlation estimator. Columns 5 and 9 of Table S-I provide the finite-sample (maximum) rejection probabilities of the nominal 5% PSC-Sub and PSC-FCV tests. Results for the symmetric and equal-tailed PSC-FCV tests are not given because the PSC-FCV and FCV tests are the same in these cases since the FCV test has correct asymptotic size. Results for the PSC-Hyb test are not given because it is the same as the Hyb test.

The results for the PSC-Sub tests are impressive. The finite-sample sizes of the upper, symmetric, and equal-tailed tests are 5.3%, 5.1%, and 5.5%, respectively, whereas the finite-sample sizes of the Sub tests are 49.8%, 8.4%, and 52.7%. The plug-in feature of the size-correction method yields (maximum) rejection probabilities across different h_2 values that are all reasonably close to 5.0%—ranging from 3.1% to 5.5%, with most being between 4.5% and 5.5%. Having these values all close to 5% is desirable from a power perspective.

The upper FCV test only requires minor size correction given that its asymptotic and finite-sample size is 5.8%. The PSC-FCV test provides improvement. Its finite-sample size is 5.2%.

S9.3. *Finite-Sample Adjusted Tests*

Column 3 of Table S-I gives the finite-sample adjusted-asymptotic rejection probabilities ($\times 100$) of the subsampling test. These values are noticeably closer to the finite-sample values given in column 4 than are the (unadjusted) asymptotic rejection probabilities given in column 2. For example, for the upper subsampling test and $h_2 = -.95$, the values for Adj-Asy, $n = 120$, and Asy are 22.9%, 25.6%, and 33.8%, respectively. Hence, the adjustment works pretty well for the subsampling test here. For the hybrid test, the adjusted-asymptotic and unadjusted-asymptotic rejection rates are all 5.0%, so the adjustment makes no difference for the hybrid test in this example.

Column 6 of Table S-I reports the finite-sample rejection probabilities of the APSC-Sub test. For upper and equal-tailed tests, the adjustment leads to overcorrection of the Sub test when the finite-sample correlation, denoted here by h_2, is close to -1 and 1, respectively, and appropriate size correction for other values of h_2. In consequence, for these cases the PSC-Sub test (see column 5) has better finite-sample size (viz., 5.3% and 5.5%) than the APSC-Sub test (13.5% and 13.5%). For symmetric tests, both of these size-corrected tests perform well.

In conclusion, in this example, the hybrid and PSC-Sub tests perform quite well in terms of finite-sample size for upper, symmetric, and equal-tailed tests. The APSC-Sub test performs well for the symmetric test, but not so well for upper and equal-tailed tests.

S10. AUTOREGRESSION EXAMPLE

This section provides results for the conditionally heteroskedastic autoregression example.

S10.1. *Upper and Lower One-Sided CIs*

First, we discuss Table S-II, which is analogous to Table II of the paper but provides results for upper and lower one-sided CIs rather than symmetric and equal-tailed two-sided CIs. (See the footnote to Table II of the paper and Section S1 above for details concerning the construction of Table II, which also apply to Table S-II.)

Due to the asymmetry of the asymptotic distribution J_h^* of the test statistic $T_n^*(\theta_0)$, the results for upper and lower one-sided CIs are quite different. Table S-II shows that upper one-sided FCV CIs have correct asymptotic size (up to simulation error). The same is true of upper one-sided hybrid CIs. Upper

TABLE S-II

AR EXAMPLE: CI COVERAGE PROBABILITIES ($\times$100) FOR NOMINAL 95% CIS

Case	Data Generating Process	$n = 131$ or Asy	Upper CIs FCV	Upper CIs Sub	Upper CIs Hyb	Lower CIs FCV	Lower CIs Sub	Lower CIs Hyb
(i)	GARCH	$-.90$	90.0	93.9	94.4	95.0	94.1	96.1
	MA = .15,	$-.50$	92.8	92.7	94.7	92.6	92.6	94.4
	AR = .80	.00	93.8	89.9	94.5	91.8	95.1	95.6
	$h_{27} = .86$	$\rho = .70$	95.9	83.6	95.9	88.4	97.7	97.7
		.80	96.7	83.2	96.7	86.7	97.8	97.8
		.90	97.7	83.9	97.7	84.0	97.9	97.9
		.97	98.9	89.2	98.9	74.7	97.5	97.5
		1.0	99.6	95.5	99.6	53.6	95.1	95.1
		FS-Min	90.0	83.2	93.9	53.6	92.4	94.4
		Asy	95.0	57.3	95.0	63.9	95.0	95.0
		Adj-Asy	—	82.6	95.3	—	95.0	95.2
(ii)	IGARCH	FS-Min	90.7	82.2	93.9	56.2	92.4	94.6
	MA = .20,							
	AR = .80							
(iii)	GARCH	FS-Min	90.4	86.2	94.1	60.6	92.9	95.0
	MA=.70,	Asy	95.0	77.2	95.0	79.2	95.0	95.0
	AR=.20	Adj-Asy	—	88.7	95.3	—	94.8	95.1
	$h_{27} = .54$							
(iv)	i.i.d.	FS-Min	90.5	82.8	94.0	50.4	92.7	94.3
	$h_{27} = 1$	Asy	95.2	47.4	95.0	53.4	95.0	95.0
		Adj-Asy	—	78.8	95.0	—	95.1	95.0
(v)	ARCH4	FS-Min	90.5	84.4	93.9	58.8	92.8	94.8
	(.3, .2, .2, .2)	Asy	95.0	77.2	95.0	79.2	95.0	95.0
	$h_{27} = .54$	Adj-Asy	—	88.7	95.3	—	94.8	95.1
(vi)	IARCH4	FS-Min	90.5	84.0	93.9	60.8	92.4	94.7
	(.3, .3, .2, .2)							
	Min over	Asy	94.8	47.5	94.8	54.5	94.9	94.9
	$h_{27} \in [0, 1]$	Adj-Asy	—	78.8	95.1	—	94.8	95.0

one-sided subsampling CIs, however, exhibit substantial asymptotic size distortion. The Adj-Asy size distortion of the subsampling CIs is noticeably smaller than the asymptotic size distortion and gives a better approximation to the finite-sample size distortion for sample size $n = 131$. The reason for the results just described for the upper one-sided FCV, hybrid, and subsampling CIs is that the upper tail of the asymptotic distribution J_h^* gets thinner as h_1 goes to zero. In consequence, the $1 - \alpha$ quantile of J_h^* is increasing in h_1, which leads to size distortion for the subsampling CI but not the FCV CI.

For lower one-sided CIs, the opposite is true. The lower tail of J_h^* gets thicker as h_1 goes to zero. In consequence, the lower one-sided FCV exhibits substantial asymptotic size distortion, whereas the subsampling and hybrid CIs have correct asymptotic size.

S10.2. *Verification of Assumptions for CI for an Autoregressive Parameter*

In this section, we verify the assumptions of Corollary 2 of the paper, namely Assumptions A–G, J, K–M, T, and TET, for the AR(1) example. We use Lemma 4 of AG1 to verify Assumption G. Lemma 4 of AG1 requires verification of Assumptions t1, Sub1, A, BB, C, DD, EE, and HH. Note that the latter assumptions imply Assumptions B and D. Corollary 2 of the paper establishes the desired results for the hybrid test. For the FCV and subsampling tests, the desired results hold under the same conditions by Corollary 1 in Appendix A2 of Andrews and Guggenberger (2009).

Assumptions BB(a) and (c) are verified by Proposition S1, stated below, that is proved in Andrews and Guggenberger (2008), hereafter AG-AR. Assumptions t1, Sub1, A, C, DD, F, J, T, M, TET, K, L, and BB(b) are verified in the next sub-section. Verifications of Assumptions E and EE are given in Sections S10.2.4 and S10.2.5 below for Model 1. For brevity, we do not verify these assumptions for Model 2. Finally, Assumption HH is verified in Section S10.2.6.

S10.2.1. *Verification of Assumptions t1, Sub1, A, C, DD, F, J, T, M, TET, K, L, and BB(b)*

Assumption t1 holds with $\tau_n = n^{1/2}$ by definition of $T_n^*(\theta_0)$. Assumptions Sub1 and A clearly hold. Assumption C holds by the choice of b_n. Assumption DD holds when the AR parameter is less than 1 by the assumption of a strictly stationary initial condition. In the unit root case, it holds by the i.i.d. assumption on the innovations for $i = 1, \dots, n$ and the fact that the test statistic $T_n^*(\theta_0)$ is invariant to the initial condition. Assumption F holds because J_h^* and $-J_h^*$ are strictly increasing on R for all $h \in H$, and $|J_h^*|$ is strictly increasing on R_+ and has support R_+ for all $h \in H$. For the same reason, Assumption J holds for $J_h = J_h^*$. Assumption T holds for $J_h = J_h^*$ and $J_h = -J_h^*$ because J_h^* is continuous on R and has support R for all $h \in H$. Assumption T holds for $J_h = |J_h^*|$ because $|J_h^*|$ is continuous on R_+ and has support R_+ for all $h \in H$. For the same reasons, Assumption M(a)(ii), (b)(ii), and (c)(ii) hold for $J_h = J_h^*, -J_h^*$, and $|J_h^*|$, and Assumption TET holds.

Assumption K holds because J_h^* is $N(0,1)$ for all $h = (h_1, h_2) \in H$ with $h_1 = \infty$.

Assumption L holds by properties of the Ornstein–Uhlenbeck process. Numerical calculations indicate that the supremum and infimum in this assumption are attained at $h_1 = 0$ or $h_1 = \infty$ (depending upon whether the supremum or infimum is being considered and whether $J_h = J_h^*, -J_h^*$, or $|J_h^*|$). This indicates that Assumption M(a)(i) holds. Numerical calculations also indicate that the supremum in Assumption M(b)(i) is attained at $h_1 = (0, \infty)$ or $(0, 0)$ for all $h_2 \in H_2$ depending upon whether $J_h = J_h^*, -J_h^*$, or $|J_h^*|$; hence, this assumption holds. Assumption M(c)(i) holds because $c_h(1-\alpha)$ is monotone in h_1 for each $h_2 \in H_2 = \Gamma_2$ (based on numerical calculations), which implies that either H^* is empty or $H^* = \{h \in H : h_1 > 0\}$, depending on whether $J_h = J_h^*, -J_h^*$, or $|J_h^*|$. When H^* is nonempty, $\sup_{h\in H^*} c_h(1-\alpha)$ is attained at $h_1 = \infty$.

Assumption BB(b) holds because $P_\gamma(\widehat{\sigma}_{n,b_n,j} > 0) = 1$ for all $n, b_n \geq 4$, $j = 1, \ldots, q_n$, and $\gamma \in \Gamma$.

S10.2.2. *Normalization Constants*

In this subsection, we specify the normalization constants a_n and d_n for which $a_n(\widehat{\rho}_n - \rho_n)$ and $d_n\widehat{\sigma}_n$ have nondegenerate asymptotic distributions under $\{\gamma_{n,h} : n \geq 1\}$. These constants appear in Assumptions BB, EE, and HH. The constants are rather complicated when the innovations exhibit conditional heteroskedasticity, so we show below how they simplify under conditional homoskedasticity, which should make them easier to interpret.

The normalization constants a_n and d_n depend on $\gamma_{n,h}$ and are denoted $a_n(\gamma_{n,h})$ and $d_n(\gamma_{n,h})$. They are defined as follows. Let $\{w_n : n \geq 1\}$ be any subsequence of $\{n\}$. Let $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) \in \Gamma : n \geq 1\}$ be a sequence for which $w_n\gamma_{n,1} \to \infty$ or $w_n\gamma_{n,1} \to h_1 < \infty$. Let $\rho_n = 1 - \gamma_{n,1}$. Define the 2-vectors

$$\text{(S10.1)} \quad X^1 = (Y_{n,0}^*/\phi_{n,1}, \phi_{n,1}^{-1})',$$
$$Z = \big(1, -E_{F_n}(Y_{n,0}^*/\phi_{n,1}^2)/E_{F_n}(\phi_{n,1}^{-2})\big)'.$$

Define

$$\text{(S10.2)} \quad a_{w_n}(\gamma_n) = w_n^{1/2} d_{w_n}(\gamma_n),$$
$$d_{w_n}(\gamma_n) = \begin{cases} \dfrac{E_{F_n}(Y_{n,0}^{*2}/\phi_{n,1}^2) - (E_{F_n}(Y_{n,0}^*/\phi_{n,1}^2))^2/E_{F_n}(\phi_{n,1}^{-2})}{(Z'E_{F_n}(X^1X^{1\prime}U_{n,1}^2/\phi_{n,1}^2)Z)^{1/2}}, \\ \quad \text{if } w_n\gamma_{n,1} \to \infty, \\ w_n^{1/2}, \\ \quad \text{if } w_n\gamma_{n,1} \to h_1 < \infty. \end{cases}$$

To simplify notation in this paragraph, we delete the subscript n in most expressions below and we omit the subscript F_n on expectations. In the case where $w_n\gamma_{n,1}\to\infty$ and $\rho\to 1$, the constants a_{w_n} and d_{w_n} in (S10.2) simplify to

$$\text{(S10.3)}\quad a_{w_n}=w_n^{1/2}\frac{E(Y_0^{*2}/\phi_1^2)}{(E(Y_0^{*2}U_1^2/\phi_1^4))^{1/2}}\quad\text{and}\quad d_{w_n}=\frac{E(Y_0^{*2}/\phi_1^2)}{(E(Y_0^{*2}U_1^2/\phi_1^4))^{1/2}}$$

up to lower order terms. This holds because by Lemma S1 below

$$\begin{aligned}\text{(S10.4)}\quad Z'E(X^1X^{1\prime}U_1^2/\phi_1^2)Z&=E(Y_0^{*2}U_1^2/\phi_1^4)\\&\quad-2E(Y_0^*U_1^2/\phi_1^4)E(Y_0^*/\phi_1^2)/E(\phi_1^{-2})\\&\quad+(E(Y_0^*/\phi_1^2))^2E(U_1^2/\phi_1^4)/(E(\phi_1^{-2}))^2\\&=E(Y_0^{*2}U_1^2/\phi_1^4)(1+O(1-\rho))\end{aligned}$$

and

$$\text{(S10.5)}\quad E(Y_0^{*2}/\phi_1^2)-(E(Y_0^*/\phi_1^2))^2/E(\phi_1^{-2})=E(Y_0^{*2}/\phi_1^2)(1+O(1-\rho)).$$

If, in addition, $\{U_i:i=\dots,0,1,\dots\}$ are i.i.d. with mean 0, variance $\sigma_U^2\in(0,\infty)$, and distribution F and $\phi_i=1$, then the constants a_{w_n} and d_{w_n} simplify to

$$\text{(S10.6)}\quad a_{w_n}=w_n^{1/2}(1-\rho_n^2)^{-1/2}\quad\text{and}\quad d_{w_n}=(1-\rho_n^2)^{-1/2}.$$

This follows because in the present case $\phi_i^2=1$, $EY_0^{*2}=\sum_{j=0}^{\infty}\rho^{2j}EU_{-j}^2=(1-\rho^2)^{-1}\sigma_U^2$, and $E(Y_0^{*2}U_1^2/\phi_1^2)=(1-\rho^2)^{-1}\sigma_U^4$.

Given the definitions of $a_n(\cdot)$ and $d_n(\cdot)$, $\tau_n=a_n(\gamma_{n,h})/d_n(\gamma_{n,h})=n^{1/2}$ does not depend on $\gamma_{n,h}$, as is required.

S10.2.3. *Preliminary Results From AG-AR*

In this subsection, we state the result proved in AG-AR that verifies Assumption B of the paper and Assumption BB(a) of AG1. We also state some other results proved in AG-AR because they are used below when verifying Assumptions E, EE, and HH.

We start by stating an assumption, Assumption INNOV, that specifies certain properties for the innovations $U_i=U_{n,i}$ and $\phi_i^2=\phi_{n,i}^2$. Assumption INNOV automatically holds when one is considering any sequence $\{\gamma_{n,h}:n\ge 1\}$. This follows from the definition of the parameter space $\mathcal{F}(\gamma_2)$ and the definition of a sequence $\{\gamma_{n,h}:n\ge 1\}$. Hence, when showing below that a property holds under a sequence $\{\gamma_{n,h}:n\ge 1\}$, it is sufficient to show that it holds under Assumption INNOV.

ASSUMPTION INNOV: (i) *For each* $n\ge 1$, $\{(U_{n,i},\phi_{n,i}^2):i=\dots,0,1,\dots\}$ *are stationary and strong mixing with* $E(U_{n,i}|\mathcal{G}_{n,i-1})=0$ *a.s.*, $E(U_{n,i}^2|\mathcal{G}_{n,i-1})=\sigma_{n,i}^2$

a.s. where $\mathcal{G}_{n,i}$ *is some nondecreasing sequence of* σ*-fields for* $i = \ldots, 1, 2, \ldots$ *for* $n \geq 1$ *for which* $(U_{n,j}, \phi^2_{n,j+1}) \in \mathcal{G}_{n,i}$ *for all* $j \leq i$, (ii) *the strong-mixing numbers* $\{\alpha_n(m) : m \geq 1\}$ *satisfy* $\alpha(m) = \sup_{n \geq 1} \alpha_n(m) = O(m^{-3\zeta/(\zeta-3)})$ *as* $m \to \infty$ *for some* $\zeta > 3$, (iv) $\sup_{n,i,s,t,u,v,A} E_{F_n} |\prod_{a \in A} a|^\zeta < \infty$, *where* $0 \leq i, s, t, u, v < \infty$, $n \geq 1$, *and* A *is any nonempty subset of* $\{U_{n,i-s}, U_{n,i-t}, U^2_{n,i+1}, U_{n,-u}, U_{n,-v}, U^2_{n,1}\}$, (v) $\phi^2_i \geq \delta > 0$ *a.s.*, (vi) $\lambda_{\min} E(X^1 X^{1\prime} U^2_{n,1}/\phi^2_{n,1}) \geq \delta > 0$, *where* $X^1 = (Y^*_{n,0}/\phi_{n,1}, \phi^{-1}_{n,1})'$, *and* (vii) *the following limits exist and are positive*: $h_{2,1} = \lim_{n\to\infty} EU^2_{n,i}$, $h_{2,2} = \lim_{n\to\infty} E(U^2_{n,i}/\phi^4_{n,i})$, $h_{2,3} = \lim_{n\to\infty} E(U^2_{n,i}/\phi^2_{n,i})$, $h_{2,4} = \lim_{n\to\infty} E\phi^{-1}_{n,i}$, $h_{2,5} = \lim_{n\to\infty} E\phi^{-2}_{n,i}$, *and* $h_{2,6} = \lim_{n\to\infty} E\phi^{-4}_{n,i}$.

Given that $\phi_{n,i}$ is bounded away from zero by Assumption INNOV(v), Assumption INNOV(iv) implies that $\sup_{n,i,s,t,u,v,A^*} E_{F_n} |\prod_{a \in A^*} a|^\zeta < \infty$, where $0 \leq i, s, t, u, v < \infty$, $n \geq 1$, and A^* is a nonempty subset of $\{U_{n,i-s}, U_{n,i-t}, U^2_{n,i+1}/\phi^4_{n,i+1}, U_{n,-u}, U_{n,-v}, U^2_{n,1}/\phi^4_{n,1}\}$ or a subset of $\{U_{n,i-s}, U_{n,i-t}, \phi^{-k}_{n,i+1}, U_{n,-u}, U_{n,-v}, \phi^{-k}_{n,1}\}$ for $k = 2, 3, 4$. The uniform bound on these expectations is needed in the application of the mixing inequality in (S10.15) used below in the verification of Assumption E.

Define $h_{n,1}$ by $\gamma_{n,h,1} = h_{n,1}/n$. Then $h_{n,1} \to h_1$ as $n \to \infty$ because $n\gamma_{n,h,1} \to h_1$. In this example, $h_{n,1} = 0$ corresponds to a unit root, that is, $\rho_n = 1 - \gamma_{n,h,1} = 1 - h_{n,1}/n = 1$. If $h_{n,1} = 0$, then the initial condition $Y^*_{n,0}$ is arbitrary. If $h_{n,1} > 0$, then under the assumptions in the paper the initial condition satisfies the following stationarity condition:

ASSUMPTION STAT: $Y^*_{n,0} = \sum_{j=0}^\infty \rho^j_n U_{n,-j}$, *where* $\rho_n = 1 - h_{n,1}/n$.

Let $W(\cdot)$ and $W_2(\cdot)$ be independent standard Brownian motions on $[0, 1]$ and let Z_1 be a standard normal random variable that is independent of $W(\cdot)$ and $W_2(\cdot)$. By definition,

$$
\text{(S10.7)} \quad I_h(r) = \int_0^r \exp(-(r-s)h_1)\, dW(s),
$$

$$
I^*_h(r) = \begin{cases} I_h(r) + \dfrac{1}{\sqrt{2h_1}} \exp(-h_1 r) Z_1 & \text{for } h_1 > 0, \\ W(r) & \text{for } h_1 = 0, \end{cases}
$$

$$
I^*_{D,h}(r) = I^*_h(r) - \int_0^1 I^*_h(s)\, ds,
$$

$$
Z_2 = \left(\int_0^1 I^*_{D,h}(r)^2\, dr \right)^{-1/2} \int_0^1 I^*_{D,h}(r)\, dW_2(r).
$$

Note that Z_2 has a $N(0, 1)$ distribution conditional on $(Z_1, W(\cdot))$. Hence, Z_2 has an unconditional $N(0, 1)$ distribution and is independent of $(Z_1, W(\cdot))$.

AG-AR proved the following proposition.

PROPOSITION S1: *Suppose* (i) *Assumption* INNOV *holds*, (ii) *Assumption* STAT *holds when* $\rho_n < 1$, (iii) $\rho_n \in [-1+\varepsilon, 1]$ *for some* $0 < \varepsilon < 2$, *and* (iv) $\rho_n = 1 - h_{n,1}/n$ *and* $h_{n,1} \to h_1 \in [0, \infty]$. *Then*

$$a_n(\widehat{\rho}_n - \rho_n) \to_d V_h, \quad d_n\widehat{\sigma}_n \to_d Q_h, \quad \frac{a_n(\widehat{\rho}_n - \rho_n)}{d_n\widehat{\sigma}_n} \to_d J_h,$$

where a_n, d_n, V_h, Q_h, *and* J_h *are defined as follows*.

(a) *In Model* 1, *for* $h_1 \in [0, \infty)$, $a_n = n$, $d_n = n^{1/2}$, V_h *is the distribution of*

$$\text{(S10.8)} \quad h_{2,7}\frac{\int_0^1 I^*_{D,h}(r)\,dW(r)}{h_{2,2}^{1/2}h_{2,1}^{1/2}\int_0^1 I^*_{D,h}(r)^2\,dr} + (1-h_{2,7}^2)^{1/2}\frac{\int_0^1 I^*_{D,h}(r)\,dW_2(r)}{h_{2,2}^{1/2}h_{2,1}^{1/2}\int_0^1 I^*_{D,h}(r)^2\,dr},$$

Q_h *is the distribution of*

$$\text{(S10.9)} \quad h_{2,2}^{-1/2}h_{2,1}^{-1/2}\left[\int_0^1 I^*_{D,h}(r)^2\,dr\right]^{-1/2},$$

and J_h *is the distribution of*

$$\text{(S10.10)} \quad h_{2,7}\frac{\int_0^1 I^*_{D,h}(r)\,dW(r)}{\left(\int_0^1 I^*_{D,h}(r)^2\,dr\right)^{1/2}} + (1-h_{2,7}^2)^{1/2}Z_2.$$

(b) *In Model* 1, *for* $h_1 = \infty$, a_n *and* d_n *are defined as in* (S10.2) *with* n *in place of* w_n, V_h *is a* $N(0,1)$ *distribution*, Q_h *is the distribution of the constant one*, *and* J_h *is a* $N(0,1)$ *distribution*.

In the remainder of this subsection, we state several other results that are proved in AG-AR and are used below when verifying Assumptions E, EE, and HH. In the proof of Proposition S1 for the case $n(1-\rho) \to \infty$, AG-AR showed the following results. If $\rho \to 1$,

$$\text{(S10.11)} \quad \frac{n^{-1/2}X_1' P_{X_2} U}{(E(Y_0^{*2}U_1^2/\phi_1^4))^{1/2}} \to_p 0 \quad \text{and} \quad \sum_{i=1}^n \zeta_i \to_d N(0,1), \quad \text{where}$$

$$\zeta_i = n^{-1/2}\frac{Y^*_{i-1}U_i/\phi_i^2}{(E(Y_0^{*2}U_1^2/\phi_1^4))^{1/2}}.$$

Furthermore,

$$\text{(S10.12)}\quad \frac{n^{-1}X_1'X_1}{E(Y_0^{*2}/\phi_1^2)} \to_p 1, \quad \frac{n^{-1}X_1'P_{X_2}X_1}{E(Y_0^{*2}/\phi_1^2)} \to_p 0,$$
$$\frac{n^{-1}X_1'M_{X_2}\Delta^2 M_{X_2}X_1}{E(Y_0^{*2}U_1^2/\phi_1^4)} \to_p 1.$$

If $\rho \to \rho^* < 1$, we have

$$\text{(S10.13)}\quad \frac{n^{-1}X_1'M_{X_2}X_1}{E(Y_0^{*2}/\phi_1^2) - (E(Y_0^*/\phi_1^2))^2/E(\phi_1^{-2})} \to_p 1.$$

AG-AR proved the following lemma which is helpful in determining the order of the normalization sequences $a_n(\gamma_{n,h})$ and $d_n(\gamma_{n,h})$ in the case where $h = \infty$.

LEMMA S1: *Suppose $n(1-\rho) \to \infty$, $\rho \to 1$, and Assumptions* INNOV *and* STAT *hold. Then we have*

$$E(Y_0^{*2}U_1^2/\phi_1^4) - (1-\rho^2)^{-1}(EU_1^2)^2/\phi_1^4 = O(1),$$
$$E(Y_0^{*2}/\phi_1^2) - (1-\rho^2)^{-1}EU_1^2 E\phi_1^{-2} = O(1),$$
$$E(Y_0^*/\phi_1^2) = O(1),$$
$$E(Y_0^*U_1^2/\phi_1^4) = O(1).$$

A more detailed version of the following lemma is proven in AG-AR as well.

LEMMA S2: *Suppose Assumptions* INNOV *and* STAT *hold, and $\rho_n \in (-1, 1]$, $\rho_n = 1 - h_{n,1}/n$, where $h_{n,1} \to h_1 \in (0, \infty)$. Then the following results hold jointly*:

(a) $n^{-1}\sum_{i=1}^n \phi_{n,i}^{-j} \to_p \lim_{n\to\infty} E_{F_n}\phi_{n,i}^{-j} = h_{2,(j+3)}$ *for* $j = 1, 2, 4$.

(b) $n^{-1}\sum_{i=1}^n U_{n,i}^2/\phi_{n,i}^4 \to_p \lim_{n\to\infty} E_{F_n}(U_{n,i}^2/\phi_{n,i}^4) = h_{2,2}$.

(c) $n^{-3/2}\sum_{i=1}^n Y_{n,i-1}^*/\phi_{n,i}^2 = O_p(1)$ *and* $n^{-3/2}\sum_{i=1}^n Y_{n,i-1}^* U_{n,i}^2/\phi_{n,i}^4 = O_p(1)$.

When $\rho_n = 1 - h_{n,1}/n$, where $h_{n,1} \to h_1 < \infty$, it is shown in AG-AR that

$$\text{(S10.14)}\quad n^{-2}\sum_{i=1}^n Y_{i-1}^{*2}\widehat{U}_i^2/\phi_i^4 = n^{-2}\sum_{i=1}^n Y_{i-1}^{*2}U_i^2/\phi_i^4 + o_p(1),$$
$$n^{-3/2}\sum_{i=1}^n Y_{i-1}^*\widehat{U}_i^2/\phi_i^4 = n^{-3/2}\sum_{i=1}^n Y_{i-1}^*U_i^2/\phi_i^4 + o_p(1),$$
$$n^{-1}\sum_{i=1}^n \widehat{U}_i^2/\phi_i^4 = n^{-1}\sum_{i=1}^n U_i^2/\phi_i^4 + o_p(1),$$

where $\widehat{U}_i/\phi_i$ is the ith residual from the LS regression of Y_i/ϕ_i on Y_{i-1}/ϕ_i and $1/\phi_i$.

S10.2.4. *Verification of Assumption E*

In this section, we verify Assumption E for Model 1. We make repeated use of the following well-known strong-mixing covariance inequality (see, e.g., Doukhan (1994, Theorem 3, p. 9)). Let X and Y be strong-mixing random variables with respect to sigma algebras $\mathcal{F}_i^j$ (for integers $i \le j$) such that $X \in \mathcal{F}_{-\infty}^n$ and $Y \in \mathcal{F}_{n+k}^{\infty}$ with strong-mixing numbers $\{\alpha(k): k \ge 1\}$. For $p, q > 0$ such that $1 - p^{-1} - q^{-1} > 0$, let $\|X\|_p = (E|X|^p)^{1/p}$ and $\|Y\|_q = (E|Y|^q)^{1/q}$. Then the following inequality holds:

$$\text{(S10.15)}\quad \operatorname{Cov}(X, Y) \le 8\|X\|_p\|Y\|_q\alpha(k)^{1-p^{-1}-q^{-1}}.$$

Below we apply the mixing inequality (S10.15) with $p = q = \zeta > 3$, where ζ appears in Assumption INNOV. Assumption INNOV(iv) will imply that the expression $\|X\|_p\|Y\|_q$ on the r.h.s. of the inequality is $O(1)$.

For verification of Assumption E, as argued in the next paragraph, it is enough to show that for all $x \in R$, $U_{n,b_n}(x) - E_{\gamma_n}U_{n,b_n}(x) \to_p 0$ under $\{\gamma_n : n \ge 1\}$ for all sequences $\{\gamma_n = (1-\rho_n, \gamma_{n,2}', \gamma_{n,3}')' \in \Gamma : n \ge 1\}$ that satisfy $n(1-\rho_n) \to h_1$, $b(1-\rho_n) \to g_1$, and $\gamma_{n,2} \to h_2 \in \Gamma_2$ for $(g, h) \in GH$, where $g = (g_1, h_2)$ and $h = (h_1, h_2)$.

To show $U_{n,b_n}(x) - E_{\gamma_n}U_{n,b_n}(x) \to_p 0$ under an arbitrary sequence $\{\gamma_n \in \Gamma : n \ge 1\}$ it is enough to show that for any subsequence $\{t_n\}$ there is a sub-subsequence $\{s_n\}$ such that $U_{s_n,b_{s_n}}(x) - E_{\gamma_{s_n}}U_{s_n,b_{s_n}}(x) \to_p 0$ under $\{\gamma_{s_n} \in \Gamma : n \ge 1\}$. Given any subsequence $\{t_n\}$, we can always construct a sub-subsequence $\{s_n\}$ such that $s_n(1-\rho_{s_n}) \to h_1$, $b_{s_n}(1-\rho_{s_n}) \to g_1$, and $\gamma_{s_n,2} \to h_2$ for $(g, h) \in GH$. Proceeding as in the proof of Lemma 6(iii) in AG1, we can define a sequence $\{\gamma_n^* : n \ge 1\}$ such that $n(1-\rho_n^*) \to h_1$, $b(1-\rho_n^*) \to g_1$, $\gamma_{n,2}^* \to h_2$, and $\gamma_{s_n}^* = \gamma_{s_n}$. It follows that $U_{n,b_n}(x) - E_{\gamma_n}U_{n,b_n}(x) \to_p 0$ holds under $\{\gamma_n^* : n \ge 1\}$ and therefore $U_{s_n,b_{s_n}}(x) - E_{\gamma_{s_n}}U_{s_n,b_{s_n}}(x) \to_p 0$ holds under $\{\gamma_{s_n} \in \Gamma : n \ge 1\}$.

For notational simplicity, in the rest of this section we let ρ denote ρ_n.

It is sufficient to show that for any given $x \in R$, $\operatorname{Var}(U_{n,b_n}(x)) \to 0$ under all sequences $\{\gamma_n \in \Gamma : n \ge 1\}$ that satisfy the conditions in the second paragraph of this subsection. Recall that $T_{n,b,k}(\rho)$ denotes the studentized t statistic based on the kth subsample and the full-sample version is defined as $T_n^*(\theta_n) = n^{1/2}(\widehat{\rho} - \rho_n)/\widehat{\sigma}$, where $\widehat{\rho} = (X_1'M_{X_2}X_1)^{-1}X_1'M_{X_2}Y$ and $\widehat{\sigma}^2 = (n^{-1}X_1'M_{X_2}X_1)^{-2}(n^{-1}X_1'M_{X_2}\Delta^2M_{X_2}X_1)$.[5] We write $T_{n,k}$ instead of $T_{n,b,k}(\rho)$ to simplify notation. Define

$$\text{(S10.16)}\quad I_{b,k} = 1\{T_{n,k} \le x\}.$$

[5]Here we deal with the upper one-sided case, so that $T_{n,b,k}(\rho) = T_{n,b,k}^*(\rho)$. The lower one-sided and symmetric two-sided cases can be dealt with using the same approach.

Stationarity of $I_{b,k}$ in k implies that

$$\text{(S10.17)}\quad \mathrm{Var}(U_{n,b_n}(x)) = q_n^{-1}\,\mathrm{Var}(I_{b,0}) + 2q_n^{-2}\sum_{k=1}^{q_n-1}(q_n - k)\,\mathrm{Cov}(I_{b,0}, I_{b,k}).$$

In this example, $q_n = n - b + 1$. Thus, it suffices to show $n^{-1}\sum_{k=0}^{n}|\,\mathrm{Cov}(I_{b,0}, I_{b,k})| \to 0$. This is implied by

$$\text{(S10.18)}\quad \sup_{k\geq k_n}|\,\mathrm{Cov}(I_{b,0}, I_{b,k})| \to 0$$

as $n\to\infty$ for some sequence $k_n\to\infty$ such that $k_n/n\to 0$.

By definition $T_{n,k} = b^{1/2}(\widehat{\rho}_{n,b,k} - \rho)/\widehat{\sigma}_{n,b,k}$. Below we show that for all $k \geq k_n$ we can write

$$\text{(S10.19)}\quad T_{n,k} = \widetilde{T}_{n,k} + \eta_{n,k}$$

for some random variables $\widetilde{T}_{n,k}$ and $\eta_{n,k}$ that are defined such that (i) $\widetilde{T}_{n,k}$ and $T_{n,0}$ are based on innovations, U_i, that are separated by at least b time periods and (ii) $\eta_{n,k} = o_p(1)$ uniformly in $k \geq k_n$ (by which we mean that $\forall\varepsilon > 0$, $\sup_{k\geq k_n}\Pr(|\eta_{n,k}| > \varepsilon) \to 0$). (Likewise, for an array $a_{n,k}$ of real numbers, we say that $a_{n,k}$ is $o(1)$ uniformly in $k \geq k_n$ if $\sup_{k\geq k_n}|a_{n,k}| \to 0$ as $n\to\infty$.) Note that the largest index of any U_i appearing in $T_{n,0}$ is $i = b$.

Using (S10.19), we show below that

$$\text{(S10.20)}\quad \sup_{k\geq k_n}|P(\widetilde{T}_{n,k}\leq x) - P(T_{n,k}\leq x)| \to 0,$$

$$\sup_{k\geq k_n}|P(T_{n,0}\leq x\ \&\ T_{n,k}\leq x) - P(T_{n,0}\leq x\ \&\ \widetilde{T}_{n,k}\leq x)| \to 0.$$

Using these results, we obtain

$$\begin{aligned}\text{(S10.21)}\quad \mathrm{Cov}(I_{b,0}, I_{b,k}) &= EI_{b,0}I_{b,k} - EI_{b,0}EI_{b,k}\\ &= P(T_{n,0}\leq x\ \&\ T_{n,k}\leq x) - P(T_{n,0}\leq x)P(T_{n,k}\leq x)\\ &= P(T_{n,0}\leq x\ \&\ \widetilde{T}_{n,k}\leq x) - P(T_{n,0}\leq x)P(\widetilde{T}_{n,k}\leq x)\\ &\quad + o(1)\\ &\leq \alpha(b) + o(1)\\ &= o(1),\end{aligned}$$

where the third equality holds by (S10.20), the fourth equality holds by the definition of the α-mixing numbers $\{\alpha_n(m) : m = 1, 2, \ldots\}$ of $\{U_{n,i} : i = \ldots, 0, 1, \ldots\}$, where $\alpha(m) = \sup_{n\geq 1}\alpha_n(m)$, and the fact that $\widetilde{T}_{n,k}$ and $T_{n,0}$ are separated by at least b time periods, the last equality holds by the strong-mixing assumption in the definition of $\mathcal{F}(\gamma_2)$, and the $o(1)$ expression is uniform in $k \geq k_n$ by (S10.20). Therefore (S10.18) holds and the proof is complete except for the verifications of (S10.19) and (S10.20).

Equation (S10.20) is established as follows. Equation (S10.19) and $P(T_{n,k} \leq x) \to J_h(x)$ as $n \to \infty$ where J_h is continuous (see Proposition S1) imply that for all $\varepsilon > 0$ there exist $\delta > 0$ and $n_0 \in N$ such that for $n \geq n_0$ we have

$$\text{(S10.22)}\quad \begin{aligned} &\sup_{k \geq k_n} P(|\widetilde{T}_{n,k} - T_{n,k}| > \delta) < \varepsilon/2, \\ &P(T_{n,k} \leq x + \delta) \leq P(T_{n,k} \leq x) + \varepsilon/2, \\ &P(T_{n,k} \leq x) \leq P(T_{n,k} \leq x - \delta) + \varepsilon/2. \end{aligned}$$

The latter two inequalities hold for all k because $T_{n,k}$ is identically distributed across k. These results lead to

$$\text{(S10.23)}\quad \begin{aligned} &\sup_{k \geq k_n} |P(\widetilde{T}_{n,k} \leq x) - P(T_{n,k} \leq x)| \\ &\quad = \sup_{k \geq k_n} \max\big\{P(\widetilde{T}_{n,k} \leq x) - P(T_{n,k} \leq x), \\ &\qquad - P(\widetilde{T}_{n,k} \leq x) + P(T_{n,k} \leq x)\big\} \\ &\quad \leq \sup_{k \geq k_n} \max\big\{P(\widetilde{T}_{n,k} \leq x) - P(T_{n,k} \leq x + \delta), \\ &\qquad - P(\widetilde{T}_{n,k} \leq x) + P(T_{n,k} \leq x - \delta)\big\} + \varepsilon/2 \\ &\quad \leq \sup_{k \geq k_n} P(|\widetilde{T}_{n,k} - T_{n,k}| > \delta) + \varepsilon/2 \leq \varepsilon, \end{aligned}$$

which proves the first result in (S10.20). The second result in (S10.20) can be proved in the same way. For example, the analogue of the third equation in (S10.22) holds because

$$\text{(S10.24)}\quad \begin{aligned} &P(T_{n,0} \leq x \;\&\; T_{n,k} \leq x) - P(T_{n,0} \leq x \;\&\; T_{n,k} \leq x - \delta) \\ &\quad \leq P(x - \delta < T_{n,k} \leq x) = P(x - \delta < T_{n,0} \leq x) < \varepsilon/2 \end{aligned}$$

for all k, for $\delta > 0$ small enough. This completes the proof of (S10.20).

It remains to establish (S10.19). We consider several cases: (i) $b(1-\rho) \to \infty$ with two subcases $\rho \to 1$ and $\rho \to \rho^* < 1$, (ii) $b(1-\rho) \to h_1 \in (0, \infty)$, (iii) $b(1-\rho) \to 0$ and $n(1-\rho) \to \infty$, and (iv) $n(1-\rho) \to h_1 \in [0, \infty)$.

PROOF OF (S10.19)—Case (i). $b(1-\rho)\to\infty$: By Proposition S1, we know that for

$$\text{(S10.25)}\quad d_b = \frac{E_{F_n}(Y_0^{*2}/\phi_1^2) - (E_{F_n}(Y_0^*/\phi_1^2))^2/E_{F_n}(\phi_1^{-2})}{(Z'E_{F_n}(X^1X^{1\prime}U_1^2/\phi_1^2)Z)^{1/2}} = \frac{d_{b1}}{d_{b2}^{1/2}} \quad \text{and}$$
$$a_b = b^{1/2}d_b,$$

we have $d_b\widehat{\sigma}_{n,b,k}\to_p 1$. Also, by (S10.12) and (S10.13) we have $d_{b1}^{-1}b^{-1}X_1'M_{X_2}\times X_1\to_p 1$, where here (with abuse of notation) X_1 and X_2 denote b-vectors containing data from the kth subsample. This implies that, uniformly in k,

$$\text{(S10.26)}\quad T_{n,k} = b^{-1/2}d_{b2}^{-1/2}\sum_{i=1}^{b}Y_{k+i-1}^*U_{k+i}/\phi_{k+i}^2 - \left(b^{-1}d_{b2}^{-1/2}\sum_{j=1}^{b}Y_{k+j-1}^*/\phi_{k+j}^2\right) \times\left(b^{-1}\sum_{j=1}^{b}\phi_{k+j}^{-2}\right)^{-1}b^{-1/2}\sum_{i=1}^{b}U_{k+i}/\phi_{k+i}^2 + o_p(1).$$

Consider first the subcase where $\rho\to 1$. In that case, (S10.11) implies further that, uniformly in k,

$$\text{(S10.27)}\quad T_{n,k} = b^{-1/2}d_{b2}^{-1/2}\sum_{i=1}^{b}Y_{k+i-1}^*U_{k+i}/\phi_{k+i}^2 + o_p(1).$$

Note that

$$\text{(S10.28)}\quad \sum_{i=1}^{b}Y_{k+i-1}^*U_{k+i}/\phi_{k+i}^2 = \sum_{i=1}^{b}\sum_{s=0}^{\infty}\rho^sU_{k+i-1-s}U_{k+i}/\phi_{k+i}^2.$$

Set

$$\text{(S10.29)}\quad \widetilde{T}_{n,k} = b^{-1/2}d_{b2}^{-1/2}\sum_{i=1}^{b}\sum_{s=0}^{k+i-2b-2}\rho^sU_{k+i-1-s}U_{k+i}/\phi_{k+i}^2$$

and note that the smallest index of any U_i appearing in $\widetilde{T}_{n,k}$ is $i=2b-1$. We are just left with showing that

$$\text{(S10.30)}\quad T_{n,k}-\widetilde{T}_{n,k} = b^{-1/2}d_{b2}^{-1/2}\sum_{i=1}^{b}\sum_{s=k+i-2b-1}^{\infty}\rho^sU_{k+i-1-s}U_{k+i}/\phi_{k+i}^2 + o_p(1) = o_p(1).$$

To show (S10.30), note that by Markov's inequality, we have

$$(\text{S10.31})\quad P\left(\left|b^{-1/2}d_{b2}^{-1/2}\sum_{i=1}^{b}\sum_{s=k+i-2b-1}^{\infty}\rho^s U_{k+i-1-s}U_{k+i}/\phi_{k+i}^2\right| > \varepsilon\right)$$

$$\leq \varepsilon^{-2}b^{-1}d_{b2}^{-1}\sum_{i,j=1}^{b}\sum_{\substack{s=k+i-2b-1\\ t=k+j-2b-1}}^{\infty}\rho^{s+t}EU_{k+i-1-s}$$

$$\times (U_{k+i}/\phi_{k+i}^2)U_{k+j-1-t}U_{k+j}/\phi_{k+j}^2$$

$$= O(b^{-1}(1-\rho))\sum_{i=1}^{b}\sum_{s,t=k+i-2b-1}^{\infty}\rho^{s+t}EU_{k+i-1-s}U_{k+i-1-t}U_{k+i}^2/\phi_{k+i}^4,$$

where the equality holds by Lemma S1 and the fact that $E(U_{k+i-1-s}(U_{k+i}/\phi_{k+i}^2)U_{k+j-1-t}U_{k+j}/\phi_{k+j}^2) = 0$ for $i \neq j$. The contribution of all terms with $s = t$ is $o(1)$ because

$$(\text{S10.32})\quad \sum_{s=k+i-2b-1}^{\infty}\rho^{2s} = \sum_{s=0}^{\infty}\rho^{2(s+k+i-2b-1)} \leq \rho^i\sum_{s=0}^{\infty}\rho^s = \rho^i(1-\rho)^{-1}$$

and $b^{-1}\sum_{i=1}^{b}\rho^i = o_p(1)$ since $b(1-\rho)\to\infty$. For the contributions with $s > t$, using (S10.15) and Assumption INNOV(iv), the r.h.s. of (S10.31) equals

$$(\text{S10.33})\quad O(b^{-1}(1-\rho))\sum_{i=1}^{b}\sum_{s>t=k+i-2b-1}^{\infty}\rho^{s+t}(s-t)^{-3-\varepsilon}$$

$$= O(b^{-1}(1-\rho))\sum_{i=1}^{b}\rho^i\sum_{s>t=0}^{\infty}\rho^{s+t}(s-t)^{-3-\varepsilon}$$

$$= O(b^{-1})\sum_{i=1}^{b}\rho^i$$

$$= o(1),$$

where the first equality holds by the change of variables $s \to s+k+i-2b-1$ and similarly for t, and the last equality uses $b(1-\rho)\to\infty$.

Next consider the subcase where $\rho \to \rho^* < 1$. In this case, define

$$(\text{S10.34})\quad \widetilde{T}_{n,k} = b^{-1/2}d_{b2}^{-1/2}\sum_{i=1}^{b}\sum_{s=0}^{k+i-2b-2}\rho^s U_{k+i-1-s}U_{k+i}/\phi_{k+i}^2$$

$$
\begin{aligned}
&-\left(b^{-1}d_{b2}^{-1/2}\sum_{j=1}^{b}\sum_{s=0}^{k+j-2b-2}\rho^{s}U_{k+j-1-s}/\phi_{k+j}^{2}\right)\\
&\times\left(b^{-1}\sum_{j=1}^{b}\phi_{k+j}^{-2}\right)^{-1}b^{-1/2}\sum_{i=1}^{b}U_{k+i}/\phi_{k+i}^{2}.
\end{aligned}
$$

Note that $(b^{-1}\sum_{j=1}^{b}\phi_{k+j}^{-2})^{-1}$, $b^{-1/2}\sum_{i=1}^{b}U_{k+i}/\phi_{k+i}^{2}$, and $d_{b2}^{-1/2}$ are all $O_p(1)$. (The first quantity is $O_p(1)$ by Lemma S2(a) and $h_{2,5}\geq\varepsilon_2>0$, the second quantity is $O_p(1)$ by a CLT, and the third quantity is $O_p(1)$ by Assumption INNOV(iv).) Therefore, it is enough to show that

$$
\begin{aligned}
&b^{-1/2}\sum_{i=1}^{b}\sum_{s=k+i-2b-1}^{\infty}\rho^{s}U_{k+i-1-s}U_{k+i}/\phi_{k+i}^{2}=o_p(1),\\
&b^{-1}\sum_{i=1}^{b}\sum_{s=k+i-2b-1}^{\infty}\rho^{s}U_{k+i-1-s}/\phi_{k+i}^{2}=o_p(1).
\end{aligned}
\tag{S10.35}
$$

Using Markov's inequality we have

$$
\begin{aligned}
&P\left(\left|b^{-1/2}\sum_{i=1}^{b}\sum_{s=k+i-2b-1}^{\infty}\rho^{s}U_{k+i-1-s}U_{k+i}/\phi_{k+i}^{2}\right|>\varepsilon\right)\\
&=O(b^{-1})\sum_{i,j=1}^{b}\sum_{\substack{s=k+i-2b-1\\ t=k+j-2b-1}}^{\infty}\rho^{s+t}EU_{k+i-1-s}U_{k+i}/\phi_{k+i}^{2}U_{k+j-1-t}U_{k+j}/\phi_{k+j}^{2}\\
&=O(b^{-1})\sum_{i=1}^{b}\sum_{s,t=k+i-2b-1}^{\infty}\rho^{s+t}\\
&=O(b^{-1})\sum_{i=1}^{b}\sum_{s,t=0}^{\infty}\rho^{s+t+2k+2i-4b-2}\\
&=O(b^{-1})\sum_{i=1}^{b}\rho^{2i}\sum_{s,t=0}^{\infty}\rho^{s+t}\\
&=o(1),
\end{aligned}
\tag{S10.36}
$$

where in the second equality we use Assumption INNOV(iv). The second term in (S10.35) can be handled analogously. This completes the proof for case (i).

For cases (ii)–(iv) we proceed as follows to establish (S10.19). Define

$$c_k = b^{-1}\sum_{j=1}^{b}\phi_{k+j}^{-2}, \tag{S10.37}$$

$$f_{k,i} = Y_{k+i-1}^{*} - c_k^{-1}b^{-1}\sum_{j=1}^{b}Y_{k+j-1}^{*}/\phi_{k+j}^{2}.$$

Note that

$$T_{n,k} = b^{1/2}(\widehat{\rho}_{n,b,k}-\rho)/\widehat{\sigma}_{n,b,k} = S_{1,k}/S_{2,k}^{1/2}, \quad \text{where} \tag{S10.38}$$

$$S_{1,k} = b^{-1}\sum_{i=1}^{b} f_{k,i}U_{k+i}/\phi_{k+i}^{2}, \quad S_{2,k} = b^{-2}\sum_{i=1}^{b} f_{k,i}^{2}\widehat{U}_{k+i}^{2}/\phi_{k+i}^{4},$$

and $\widehat{U} = M_X U$.[6] We show below that $S_{1,k}$ and $S_{2,k}$ can be written as

$$S_{1,k} = \widetilde{S}_{1,k} + \xi_{1,k} \quad \text{and} \quad S_{2,k} = \widetilde{S}_{2,k} + \xi_{2,k}, \tag{S10.39}$$

where $\widetilde{S}_{1,k}$ and $\widetilde{S}_{2,k}$ are separated from $S_{1,0}$ and $S_{2,0}$ by b time periods $\forall k \geq k_n$, $\xi_{1,k} = o_p(1)$, and $\xi_{2,k} = o_p(1)$.

Note that

$$\begin{aligned} f_{k,i} &= \sum_{s=0}^{\infty}\rho^{s}U_{k+i-1-s} - c_k^{-1}b^{-1}\sum_{j=1}^{b}\sum_{s=0}^{\infty}\rho^{s}U_{k+j-1-s}/\phi_{k+j}^{2} \\ &= \sum_{s=i-1}^{\infty}\rho^{s}U_{k+i-1-s} - c_k^{-1}b^{-1}\sum_{j=1}^{b}\sum_{s=j-1}^{\infty}\rho^{s}U_{k+j-1-s}/\phi_{k+j}^{2} \\ &\quad + \sum_{s=0}^{i-2}\rho^{s}U_{k+i-1-s} - c_k^{-1}b^{-1}\sum_{j=1}^{b}\sum_{s=0}^{j-2}\rho^{s}U_{k+j-1-s}/\phi_{k+j}^{2} \\ &= \rho^{i-1}\sum_{s=0}^{\infty}\rho^{s}U_{k-s} - c_k^{-1}b^{-1}\sum_{j=0}^{b-1}\rho^{j}\sum_{s=0}^{\infty}\rho^{s}U_{k-s}/\phi_{k+j+1}^{2} \\ &\quad + \sum_{s=1}^{i-1}\rho^{-s+i-1}U_{k+s} - c_k^{-1}b^{-1}\sum_{s=0}^{b-2}\sum_{j=1}^{b-(s+1)}\rho^{s}U_{k+j}/\phi_{k+j+s+1}^{2}, \end{aligned} \tag{S10.40}$$

[6]Strictly speaking, all sums over $i = 1, \ldots, b$ should be over $i = 1, \ldots, b-1$ because one observation from a block of length b is used as an initial observation given that lagged Y_i is a regressor. For notational simplicity, here and below, we sum to b rather than $b-1$.

where we used the transformation $s \mapsto -s + i - 1$ for the first sum of the last row, changed the sequence of summation over j and s, and applied the transformation $j \mapsto j + s$ in the second sum of the last row. Therefore, changing back the sequence of summation over j and s in the second sum, it follows that for $\rho < 1$,

$$\text{(S10.41)}\quad f_{k,i} = a_{k,i} \sum_{j=0}^{\infty} \rho^j U_{k-j} + \sum_{j=1}^{b-1} c_{k,i,j} U_{k+j}, \quad \text{where}$$

$$a_{k,i} = \rho^{i-1} - \frac{c_k^{-1}}{b} \sum_{l=0}^{b-1} \frac{\rho^l}{\phi_{k+l+1}^2},$$

$$c_{k,i,j} = 1(j \le i-1)\rho^{i-j-1} - \frac{c_k^{-1}}{b} \sum_{l=0}^{b-(j+1)} \frac{\rho^l}{\phi_{k+j+l+1}^2}.$$

Note that $a_{k,i}$ is random. When $\rho = 1$, (S10.41) simplifies to

$$\text{(S10.42)}\quad f_{k,i} = \sum_{j=1}^{b-1} \left(1(j \le i-1) - \frac{c_k^{-1}}{b} \sum_{l=0}^{b-(j+1)} \phi_{k+j+l+1}^{-2} \right) U_{k+j}.$$

By (S10.38) and (S10.45) below, (S10.42) implies that $T_{n,k}$ is separated from $T_{n,0}$ by at least b time periods when $k > 2b$. Thus, if $\rho = 1$ for all n, (S10.19) holds immediately. This leads us to only consider cases where $\rho < 1$ for all n. (Sequences in which $\rho = 1$ for some n and $\rho < 1$ for some n can be handled by analyzing subsequences.)

We now truncate the infinite sum in $f_{k,i}$ and for $k > 2b$ define

$$\text{(S10.43)}\quad f_{k,i}^t = a_{k,i} \sum_{j=0}^{k-2b-1} \rho^j U_{k-j} + \sum_{j=1}^{b-1} c_{k,i,j} U_{k+j}$$

$$= Y_{k+i-1}^{*t} - c_k^{-1} b^{-1} \sum_{j=1}^{b} Y_{k+j-1}^{*t} / \phi_{k+j}^2, \quad \text{where}$$

$$Y_{l-1}^{*t} = \sum_{s=0}^{l-2b-2} \rho^s U_{l-1-s}.$$

Note that $f_{k,i}^t$ is obtained from $f_{k,i}$ by deleting all U_p with subindices $p < 2b+1$. Define

$$\text{(S10.44)}\quad \widetilde{S}_{1,k} = b^{-1} \sum_{i=1}^{b} f_{k,i}^t U_{k+i} / \phi_{k+i}^2,$$

$$\xi_{1,k} = b^{-1}\sum_{i=1}^{b}(f_{k,i} - f^t_{k,i})U_{k+i}/\phi^2_{k+i}$$
$$= b^{-1}\sum_{i=1}^{b} a_{k,i}(U_{k+i}/\phi^2_{k+i}) \sum_{j=k-2b}^{\infty} \rho^j U_{k-j}.$$

For $k > 2b$, $\widetilde{S}_{1,k}$ depends only on innovations U_p for $p > 2b$, and $S_{1,0}$ and $S_{2,0}$ depend only on innovations U_p for $p \leq b$. Thus, for $k > 2b$, $\widetilde{S}_{1,k}$ is separated from $S_{1,0}$ and $S_{2,0}$ by at least b time periods.

Regarding $S_{2,k}$, note that by (S10.14) and the definition in (S10.37) we have

$$\text{(S10.45)}\quad S_{2,k} = b^{-2}\sum_{i=1}^{b} f^2_{k,i}U^2_{k+i}/\phi^4_{k+i} + o_p(1).$$

Set

$$\text{(S10.46)}\quad \widetilde{S}_{2,k} = b^{-2}\sum_{i=1}^{b}(f^t_{k,i})^2 U^2_{k+i}/\phi^4_{k+i}.$$

For $k > 2b$, $\widetilde{S}_{2,k}$ depends only on innovations U_p for $p > 2b$. By definition,

$$\text{(S10.47)}\quad \xi_{2,k} = b^{-2}\sum_{i=1}^{b}(f^2_{k,i} - (f^t_{k,i})^2)U^2_{k+i}/\phi^4_{k+i} + o_p(1).$$

We now show that $\xi_{1,k} = o_p(1)$ and $\xi_{2,k} = o_p(1)$ uniformly for $k \geq k_n$ for some sequence $k_n \to \infty$ such that $k_n/n \to 0$.

Case (ii). $b(1-\rho) \to h_1 \in (0, \infty)$: We first show that $\xi_{1,k} = o_p(1)$. Clearly, it is enough to show that

$$\text{(S10.48)}\quad b^{-1}\sum_{i=1}^{b}\rho^{i-1}U_{k+i}/\phi^2_{k+i}\sum_{j=k-2b}^{\infty}\rho^j U_{k-j} = o_p(1),$$
$$\left((c_k^{-1}/b)\sum_{l=0}^{b-1}\rho^l\phi^{-2}_{k+l+1}\right)\left(b^{-1}\sum_{i=1}^{b}U_{k+i}/\phi^2_{k+i}\right)\left(\sum_{j=k-2b}^{\infty}\rho^j U_{k-j}\right)$$
$$= o_p(1).$$

Note that by Lemma S2(a), c_k^{-1} and $b^{-1}\sum_{l=0}^{b-1}\rho^l\phi^{-2}_{k+l+1}$ are both $O_p(1)$. Applying Markov's inequality, it is therefore enough to show that the following quantity is $o_p(1)$:

$$\text{(S10.49)}\quad b^{-2}\sum_{i,j=1}^{b}\sum_{l,m=k-2b}^{\infty}\rho^{l+m}E(U_{k+i}/\phi_{k+i}^2)(U_{k+j}/\phi_{k+j}^2)U_{k-l}U_{k-m}$$
$$=O(1)b^{-2}\sum_{i=1}^{b}\sum_{l,m=k-2b}^{\infty}\rho^{l+m}EU_{k-l}U_{k-m}U_{k+i}^2/\phi_{k+i}^4,$$

where the equality uses $E(U_{k+i}/\phi_{k+i}^2)(U_{k+j}/\phi_{k+j}^2)U_{k-l}U_{k-m}=0$ for $k>2b$ unless $i=j$, by the martingale difference property of U_i.

Consider first the contribution of the summands in (S10.49) when $l=m$:

$$\text{(S10.50)}\quad b^{-2}\sum_{i=1}^{b}\sum_{l=k-2b}^{\infty}\rho^{2l}EU_{k-l}^2U_{k+i}^2/\phi_{k+i}^4=O(b^{-2})\sum_{i=1}^{b}\sum_{l=k-2b}^{\infty}\rho^{2l}$$
$$=O\big(\rho^{k-2b}(b(1-\rho))^{-1}\big),$$

where in the first equality we use Assumption INNOV(iv). Define $h_{n,1}^*$ by $\rho=\exp(-h_{n,1}^*/n)$. Because $b(1-\rho)\to h_1\in(0,\infty)$, we have $h_{n,1}^*\to\infty$. In consequence, there exists a sequence $\{k_n:n\geq 1\}$ such that $k_n/b\to\infty$, $k_n/n\to 0$, and $h_{n,1}^*k_n/n\to\infty$. For this sequence, $h_{n,1}^*(k_n-2b)/n\to\infty$, $\rho^{k_n-2b}=\exp(-h_{n,1}^*(k_n-2b)/n)\to 0$, and $\sup_{k\geq k_n}\rho^{2(k-2b)}\to 0$. This shows that the expression in (S10.50) is $o(1)$.

Therefore, we only need to consider the contributions in (S10.49) with $l>m$. We have

$$\text{(S10.51)}\quad b^{-2}\sum_{i=1}^{b}\sum_{l>m=k-2b}^{\infty}\rho^{l+m}EU_{k-l}U_{k-m}U_{k+i}^2/\phi_{k+i}^4$$
$$=O(1)b^{-2}\sum_{i=1}^{b}\sum_{l>m=k-2b}^{\infty}\rho^{l+m}(l-m)^{-3-\varepsilon}$$
$$=O(1)\rho^{k-2b}b^{-2}\sum_{i=1}^{b}\sum_{l>m=0}^{\infty}\rho^{m}(l-m)^{-3-\varepsilon}$$
$$=o(1),$$

where in the first equality we use (S10.15) and Assumption INNOV(iv).

Next we show $\xi_{2,k}=o_p(1)$. Note that up to a $o_p(1)$ term, $\xi_{2,k}=\xi_{21,k}-2\xi_{22,k}+\xi_{23,k}$, where

$$\text{(S10.52)}\quad \xi_{21,k}=b^{-2}\sum_{i=1}^{b}(Y_{k+i-1}^{*2}-Y_{k+i-1}^{*t2})U_{k+i}^2/\phi_{k+i}^4,$$

$$\xi_{22,k} = c_k^{-1} b^{-3/2} \sum_{i=1}^{b} \left(Y^*_{k+i-1} \left(b^{-3/2} \sum_{j=1}^{b} Y^*_{k+j-1}/\phi^2_{k+j} \right) - Y^{*t}_{k+i-1} \left(b^{-3/2} \sum_{j=1}^{b} Y^{*t}_{k+j-1}/\phi^2_{k+j} \right) \right) U^2_{k+i}/\phi^4_{k+i},$$

$$\xi_{23,k} = \left(\left(b^{-3/2} \sum_{j=1}^{b} Y^*_{k+j-1}/\phi^2_{k+j} \right)^2 - \left(b^{-3/2} \sum_{j=1}^{b} Y^{*t}_{k+j-1}/\phi^2_{k+j} \right)^2 \right) \times c_k^{-2} b^{-1} \sum_{i=1}^{b} U^2_{k+i}/\phi^4_{k+i}.$$

To show $\xi_{21,k} = o_p(1)$, note that

$$\xi_{21,k} = b^{-2} \sum_{i=1}^{b} \sum_{\substack{s,t=0, \\ s \text{ or } t \geq k+i-2b-1}}^{\infty} \rho^{s+t} U_{k+i-1-s} U_{k+i-1-t} U^2_{k+i}/\phi^4_{k+i} \tag{S10.53}$$

(where the second sum is over all $s, t = 0, \dots$ for which $s \geq k + i - 2b - 1$ or $t \geq k + i - 2b - 1$). By Markov's inequality, we have

$$\begin{aligned} P(|\xi_{21,k}| > \varepsilon) &\leq \varepsilon^{-2} b^{-4} \sum_{i,j=1}^{b} \sum_{\substack{s,t=0, \\ s \text{ or } t \geq k+i-2b-1}}^{\infty} \sum_{\substack{u,v=0, \\ u \text{ or } v \geq k+i-2b-1}}^{\infty} \rho^{s+t+u+v} \\ &\quad \times E(U^2_{k+i}/\phi^4_{k+i})(U^2_{k+j}/\phi^4_{k+j}) \\ &\quad \times U_{k+i-1-s} U_{k+i-1-t} U_{k+j-1-u} U_{k+j-1-v}. \end{aligned} \tag{S10.54}$$

Using (S10.15), Assumption INNOV(iv), and $\rho^{k-2b} \to 0$, one can show that the contribution of all summands for which at least two of the indices $k + i - 1 - s$, $k + i - 1 - t$, $k + j - 1 - u$, or $k + j - 1 - v$ coincide is $o(1)$. In what follows, we can therefore assume that these indices are all different. We can then assume $i \geq j$, $s > t$, and $u > v$. One has to separately investigate several cases regarding the order of the six indices $k + i - 1 - s < k + i - 1 - t < k + i$ and $k + j - 1 - u < k + j - 1 - v < k + j$. We will only deal with the case where, in the ordering of the indices $(k+i-1-s, k+i-1-t, k+j-1-u, k+j-1-v, k+j, k+i)$, the subindex $k+i-1-s$ is followed immediately by $k+i-1-t$ and the subindex $k+j-1-u$ is directly followed by $k+j-1-v$, $k+i-1-s \neq k + j$, and $k + j - 1 - u \neq k + i$. The other cases are dealt with analogously.

Equation (S10.15) and Assumption INNOV(iv) yield

$$\text{(S10.55)}\quad EU_{k+i-1-s}U_{k+i-1-t}U_{k+j-1-u}U_{k+j-1-v}(U_{k+j}^2/\phi_{k+j}^4)U_{k+i}^2/\phi_{k+i}^4 \leq O(\max\{s-t,u-v\})^{-3-\varepsilon}.$$

Therefore, the summands in (S10.54) equal

$$\text{(S10.56)}\quad O(b^{-4})\sum_{i\geq j=1}^{b}\sum_{\substack{s>t=0,\\ s\geq k+i-2b-1}}^{\infty}\sum_{\substack{u>v=0,\\ u\geq k+i-2b-1}}^{\infty}\rho^{s+t+u+v}(\max\{s-t,u-v\})^{-3-\varepsilon}$$

$$= O(b^{-4}\rho^{k-2b})\sum_{i\geq j=1}^{b}\sum_{\substack{s>t=0,\\ s\geq k+i-2b-1}}^{\infty}\rho^{t}(s-t)^{(-3-\varepsilon)/2} \times \sum_{\substack{u>v=0\\ u\geq k+i-2b-1}}^{\infty}\rho^{v}(u-v)^{(-3-\varepsilon)/2}$$

$$= o(b^{-2})\left(\sum_{t=0}^{\infty}\rho^{t}\sum_{s=t+1}^{\infty}(s-t)^{(-3-\varepsilon)/2}\right)^{2}.$$

By a change of variable $s \to s+t+1$, the r.h.s. of (S10.56) equals

$$\text{(S10.57)}\quad o(b^{-2})\left(\sum_{t=0}^{\infty}\rho^{t}\sum_{s=1}^{\infty}s^{(-3-\varepsilon)/2}\right)^{2} = o(b^{-2})\left(\sum_{t=0}^{\infty}\rho^{2t}\right)^{2} = o(b^{-2}(1-\rho)^{-2}) = o(1).$$

Next we deal with $\xi_{22,k}$. Note that by Lemma S2(a) and (c) we have $c_k^{-1} = O_p(1)$ and $b^{-3/2}\sum_{j=1}^{b}Y_{k+j-1}^{*}/\phi_{k+j}^2 = O_p(1)$. We add and subtract $Y_{k+i-1}^{*t}b^{-3/2} \times \sum_{j=1}^{b}Y_{k+j-1}^{*}/\phi_{k+j}^2$ which implies that it is enough to show that

$$\text{(S10.58)}\quad b^{-3/2}\sum_{i=1}^{b}(Y_{k+i-1}^{*}-Y_{k+i-1}^{*t})U_{k+i}^2/\phi_{k+i}^4 = o_p(1),$$

$$b^{-3/2}\sum_{j=1}^{b}(Y_{k+j-1}^{*}-Y_{k+j-1}^{*t})/\phi_{k+j}^2 = o_p(1),$$

$$b^{-3/2}\sum_{i=1}^{b}Y_{k+i-1}^{*t}U_{k+i}^2/\phi_{k+i}^4 = O_p(1).$$

The third statement holds by the first one and by Lemma S2(c). To prove the first two statements, note that

$$Y^*_{k+i-1} - Y^{*t}_{k+i-1} = \sum_{s=k+i-2b-1}^{\infty} \rho^s U_{k+i-1-s}. \tag{S10.59}$$

To show (S10.58), by Markov's inequality it is sufficient to show that

$$b^{-3} \sum_{i,j=1}^{b} \sum_{\substack{s=k+i-2b-1,\\ t=k+i-2b-1}}^{\infty} \rho^{s+t} EU_{k+i-1-s} U_{k+j-1-t} (U^2_{k+i}/\phi^4_{k+i}) U^2_{k+j}/\phi^4_{k+j} = o(1) \tag{S10.60}$$

$$b^{-3} \sum_{i,j=1}^{b} \sum_{\substack{s=k+i-2b-1,\\ t=k+i-2b-1}}^{\infty} \rho^{s+t} EU_{k+i-1-s} U_{k+j-1-t} \phi^{-2}_{k+i} \phi^{-2}_{k+j} = o(1).$$

These can be shown using the method employed above. Finally, $\xi_{23,k} = o_p(1)$ follows by similar steps to the ones above and Lemma S2(b). This completes the proof of case (ii).

Case (iii). $b(1-\rho) \to 0$ & $n(1-\rho) \to \infty$: Define $h^*_{n,1}$ and $h_{n,1}$ by $\rho = \exp(-h^*_{n,1}/n)$ and $\rho = 1 - h_{n,1}/n$. Let $t_n = bh^*_{n,1}/n$. For notational simplicity, we write $h^*_{n,1}$ and $h_{n,1}$ as h^*_n and h_n, respectively, in the remainder of the verification of Assumption E. Then we have

$$\begin{aligned} \rho^b &= \exp(-bh^*_n/n) = \exp(-t_n), \quad 1+\rho = 2 - h_n/n, \\ b(1-\rho) &= bh_n/n = t_n(h_n/h^*_n). \end{aligned} \tag{S10.61}$$

We have $b(1-\rho) \to 0 \Rightarrow \rho \to 1 \Rightarrow h^*_n/n \to 0 \Rightarrow h^*_n/h_n \to 1$, where the last implication follows from a mean-value expansion of $\exp(-h^*_n/n)$ about 0. In addition, $b(1-\rho) \to 0 \Rightarrow bh_n/n \to 0$. Combining these results gives $t_n = (bh_n/n)(h^*_n/h_n) \to 0$. Also, $n(1-\rho) \to \infty$ implies that $h_n \to \infty$ and $h^*_n \to \infty$.

Because $bh_n/n = b(1-\rho) \to 0$, it follows that $h_n = o(n/b)$. This and $h^*_n/h_n \to 1$ yield $h^*_n = o(n/b)$. By an expansion of $\exp(-h^*_n/n)$ about 0, we obtain

$$\begin{aligned} 0 = \rho - \rho &= \exp(-h^*_n/n) - (1 - h_n/n) \\ &= -h^*_n/n + 2^{-1}(h^*_n/n)^2 - 6^{-1}\exp(-h^{**}_n/n)(h^*_n/n)^3 + h_n/n, \end{aligned} \tag{S10.62}$$

where $h^{**}_n/n = o(1/b)$ because $h^*_n = o(n/b)$. Hence,

$$1 - h_n/h^*_n = 2^{-1}h^*_n/n - 6^{-1}(h^*_n/n)^2 \exp(-h^{**}_n/n). \tag{S10.63}$$

We first verify (S10.39) for $\xi_{1,k} = b^{-1}\sum_{i=1}^{b} a_{k,i}(U_{k+i}/\phi_{k+i}^2)\sum_{j=k-2b}^{\infty}\rho^j U_{k-j}$ defined in (S10.44). Note that by Markov's inequality we have

$$\text{(S10.64)}\quad P\left(\left|\sum_{s=k-2b}^{\infty}\rho^s U_{k-s}\right| > M(1-\rho)^{-1/2}\right) \le M^{-2}(1-\rho)\sum_{s=k-2b}^{\infty}\rho^{2s}EU_{k-s}^2 = O(M^{-2})$$

by Assumption INNOV(iv) and because U_i is a martingale difference sequence. Therefore,

$$\text{(S10.65)}\quad (1-\rho)^{1/2}\sum_{s=k-2b}^{\infty}\rho^s U_{k-s} = O_p(1).$$

To show $\xi_{1,k} = o_p(1)$, it is thus sufficient to show that

$$\text{(S10.66)}\quad \zeta_1 = (1-\rho)^{-1/2}b^{-1}\sum_{i=1}^{b} a_{k,i}(U_{k+i}/\phi_{k+i}^2) = o_p(1).$$

By adding and subtracting 1, we can write

$$\text{(S10.67)}\quad a_{k,i} = (\rho^{i-1}-1) - c_k^{-1}b^{-1}\sum_{l=0}^{b-1}(\rho^l - 1)\phi_{k+l+1}^{-2}.$$

Therefore,

$$\text{(S10.68)}\quad \zeta_1 = (1-\rho)^{-1/2}b^{-1} \times \sum_{i=1}^{b}\left((\rho^{i-1}-1) - c_k^{-1}b^{-1}\sum_{l=0}^{b-1}(\rho^l-1)\phi_{k+l+1}^{-2}\right)(U_{k+i}/\phi_{k+i}^2)$$

and it is enough to show that

$$\text{(S10.69)}\quad \zeta_{11} = (1-\rho)^{-1/2}b^{-1}\sum_{i=1}^{b}(\rho^{i-1}-1)(U_{k+i}/\phi_{k+i}^2) = o_p(1),$$

$$\zeta_{12} = (1-\rho)^{-1/2}b^{-1/2}\left(c_k^{-1}b^{-1}\sum_{l=0}^{b-1}(\rho^l-1)\phi_{k+l+1}^{-2}\right)b^{-1/2}\sum_{i=1}^{b}(U_{k+i}/\phi_{k+i}^2) = o_p(1).$$

To show $\zeta_{11} = o_p(1)$, by Markov's inequality, it is enough to show that

$$(1-\rho)^{-1}b^{-2}\sum_{i=1}^{b}(\rho^{i-1}-1)^2 = o(1), \tag{S10.70}$$

where we use the fact that U_{k+i}/ϕ^2_{k+i} is a martingale difference sequence and $E(U^2_{k+i}/\phi^4_{k+i})$ is uniformly bounded by Assumption INNOV(iv). Writing the sum in (S10.70) in closed form, it follows that it is enough to show that

$$\frac{1-\rho^{2b}-2(1-\rho^b)(1+\rho)+b(1-\rho)(1+\rho)}{b^2(1-\rho)^2} = o(1). \tag{S10.71}$$

Using (S10.61) and (S10.63) the l.h.s. of (S10.71) equals

$$\frac{1-\exp(-2t_n)-2(1-\exp(-t_n))(1+\rho)+t_n(h_n/h_n^*)(1+\rho)}{(t_n(h_n/h_n^*))^2}. \tag{S10.72}$$

We first show that replacing $(1+\rho)$ by 2 and (h_n/h_n^*) by 1 in (S10.72) is negligible in the sense that

$$t_n^{-2}\big[-2(1-\exp(-t_n))(1+\rho-2)+t_n((h_n/h_n^*)(1+\rho)-2)\big] = o(1). \tag{S10.73}$$

To show (S10.73), note that by (S10.63) $h_n/h_n^* = 1-2^{-1}h_n^*/n+6^{-1}(h_n^*/n)^2\times\exp(-h_n^{**}/n)$, where $h_n^{**}/n\to 0$. By a Taylor expansion for $\rho=\exp(-h_n^*/n)$ we have $\rho-1=-h_n^*/n+2^{-1}(h_n^*/n)^2\exp(-h_n^{++}/n)$ for some h_n^{++} such that $h_n^{++}/n\to 0$. By a Taylor expansion for $\exp(-t_n)$ we have $1-\exp(-t_n)=t_n-2^{-1}t_n^2\exp(t_n^*)$ for some t_n^* such that $t_n^*\to 0$. Multiplying out shows that the l.h.s. in (S10.73) is of order $t_n^{-2}(O(t_n(h_n^*/n)^2)+O(t_n^2h_n^*/n))$ which is $o(1)$.

By applying l'Hopital's rule twice, the limit of the expression in (S10.72) with $(1+\rho)$ replaced by 2 and (h_n/h_n^*) replaced by 1 equals 0 which completes the proof of $\zeta_{11}=o_p(1)$.

To show $\zeta_{12}=o_p(1)$, a central limit theorem for martingale difference sequences shows that $b^{-1/2}\sum_{i=1}^{b}(U_{k+i}/\phi^2_{k+i})=O_p(1)$. Furthermore, by Assumption INNOV(v) and (vii) we have $c_k^{-1}b^{-1}\sum_{l=0}^{b-1}(\rho^l-1)\phi^{-2}_{k+l+1}=O_p(1)b^{-1}\times\sum_{l=0}^{b-1}(\rho^l-1)$ and it is therefore enough to show that

$$(1-\rho)^{-1/2}b^{-1/2}b^{-1}\sum_{l=0}^{b-1}(\rho^l-1)=o(1) \quad \text{or} \tag{S10.74}$$

$$b^{-3/2}(1-\rho)^{-3/2}(1-\rho^b-b(1-\rho))=o(1).$$

Using analogous steps as in the proof for $\zeta_{11}=o_p(1)$ above then shows $\zeta_{12}=o_p(1)$. This completes the verification of (S10.39) for $\xi_{1,k}$.

We are left with showing that the component $b^{-2}\sum_{i=1}^{b}(f_{k,i}^2-(f_{k,i}^t)^2)U_{k+i}^2/\phi_{k+i}^4$ of $\xi_{2,k}$ in (S10.47) is $o_p(1)$. Using the definitions of $f_{k,i}$ and $f_{k,i}^t$ in (S10.41) and (S10.43), it follows that

$$
\begin{aligned}
\text{(S10.75)}\quad & f_{k,i}^2-(f_{k,i}^t)^2 \\
&= a_{k,i}^2\left(\sum_{j=0}^{\infty}\rho^j U_{k-j}\right)^2 - a_{k,i}^2\left(\sum_{j=0}^{k-2b-1}\rho^j U_{k-j}\right)^2 \\
&\quad + 2\sum_{j=1}^{b-1}c_{k,i,j}U_{k+j}\left(a_{k,i}\sum_{s=k-2b}^{\infty}\rho^s U_{k-s}\right) \\
&= a_{k,i}^2 \sum_{\substack{j,\ell=0,\\ j \text{ or } \ell=\geq k-2b}}^{\infty} \rho^{j+l}U_{k-j}U_{k-l} + 2\sum_{s=k-2b}^{\infty}\rho^s U_{k-s}\sum_{j=1}^{b-1}a_{k,i}c_{k,i,j}U_{k+j} \\
&= f_{1,k,i}+f_{2,k,i}.
\end{aligned}
$$

We first show that the contributions of $f_{1,k,i}$ to $\xi_{2,k}$ are $o_p(1)$. Note that

$$
\begin{aligned}
\text{(S10.76)}\quad & b^{-2}\sum_{i=1}^{b} f_{1,k,i}U_{k+i}^2/\phi_{k+i}^4 \\
&= \sum_{\substack{j,\ell=0,\\ j \text{ or } \ell=\geq k-2b}}^{\infty} \rho^{j+l}U_{k-j}U_{k-l}b^{-2}\sum_{i=1}^{b}a_{k,i}^2U_{k+i}^2/\phi_{k+i}^4 \\
&= O_p((1-\rho)^{-1})b^{-2}\sum_{i=1}^{b}a_{k,i}^2U_{k+i}^2/\phi_{k+i}^4.
\end{aligned}
$$

Using (S10.67), it is therefore enough to show that

$$
\begin{aligned}
\text{(S10.77)}\quad & (1-\rho)^{-1}b^{-2}\sum_{i=1}^{b}(\rho^{i-1}-1)^2U_{k+i}^2/\phi_{k+i}^4 = o_p(1), \\
& (1-\rho)^{-1}b^{-2}\left(c_k^{-1}b^{-1}\sum_{l=0}^{b-1}(\rho^l-1)\phi_{k+l+1}^{-2}\right)\sum_{i=1}^{b}(\rho^{i-1}-1)U_{k+i}^2/\phi_{k+i}^4 \\
&= o_p(1), \\
& (1-\rho)^{-1}b^{-2}\left(c_k^{-1}b^{-1}\sum_{l=0}^{b-1}(\rho^l-1)\phi_{k+l+1}^{-2}\right)^2\sum_{i=1}^{b}U_{k+i}^2/\phi_{k+i}^4 = o_p(1).
\end{aligned}
$$

To deal with the first term, it is enough to show that the law of large numbers (LLN) $b^{-1}\sum_{i=1}^{b} Z_{bi} = O_p(1)$ applies with $Z_{bi} = (b(1-\rho))^{-2}\ (\rho^{i-1} - 1)^2 U_{k+i}^2/\phi_{k+i}^4$. The LLN holds by White (1984, Theorem 3.47 with $r = \delta = 3/2$) because Z_{bi} is α-mixing of size 3, has finite mean by Assumption INNOV(iv) and because $(\rho^b - 1)(b(1-\rho))^{-1} = O(1)$, and because $\sum_{i=1}^{\infty}(i^{-3}E|Z_{bi} - EZ_{bi}|^3)^{2/3} < \infty$. The latter holds because by $(\rho^b - 1)(b(1-\rho))^{-1} = O(1)$ and Assumption INNOV(iv), $E|Z_{bi} - EZ_{bi}|^3$ is uniformly bounded.

The proofs for the second and third terms in (S10.77) are analogous. Just note that $c_k^{-1} = O_p(1)$ by Lemma S2(a) and that

$$b^{-1}\sum_{l=0}^{b-1} Z_{bl}^* = O_p(1) \tag{S10.78}$$

applies also with $Z_{bl}^* = (b(1-\rho))^{-1}(\rho^l - 1)\phi_{k+l+1}^{-2}$, $Z_{bl}^* = (b(1-\rho))^{-1}(\rho^l - 1)U_{k+i}^2/\phi_{k+i}^4$, and $Z_{bl}^* = U_{k+i}^2/\phi_{k+i}^4$ by White (1984, Theorem 3.47 with $r = \delta = 3/2$).

We next show that the contributions of $f_{2,k,i}$ to $\xi_{2,k}$ are $o_p(1)$. By (S10.65) and (S10.75) it is sufficient to show that

$$(1-\rho)^{-1/2}b^{-2}\sum_{j=1}^{b-1}\sum_{i=1}^{b} a_{k,i}c_{k,i,j}(U_{k+i}^2/\phi_{k+i}^4)U_{k+j} = o_p(1). \tag{S10.79}$$

By replacing $a_{k,i}$ and $c_{k,i,j}$ by their definitions we have

$$\begin{aligned}\sum_{j=1}^{b-1}\sum_{i=1}^{b} a_{k,i}c_{k,i,j}\left(\frac{U_{k+i}^2}{\phi_{k+i}^4}\right)U_{k+j} &= \sum_{j=1}^{b-1}\sum_{i=1}^{b}\left((\rho^{i-1} - 1) - c_k^{-1}b^{-1}\sum_{l=0}^{b-1}(\rho^l - 1)\phi_{k+l+1}^{-2}\right)\\ &\quad\times\left(1(j \le i-1)\rho^{i-j-1} - \frac{c_k^{-1}}{b}\sum_{l=0}^{b-(j+1)}\frac{\rho^l}{\phi_{k+j+l+1}^2}\right)\left(\frac{U_{k+i}^2}{\phi_{k+i}^4}\right)U_{k+j}.\end{aligned} \tag{S10.80}$$

Multiplying out in (S10.80), it is clear that to show (S10.79), it is sufficient to show that the following expressions multiplied by $(1-\rho)^{-1/2}b^{-2}$ are all $o_p(1)$:

$$\begin{aligned}&\sum_{j=1}^{b-1}\sum_{i=1}^{b}(\rho^{i-1} - 1)1(j \le i-1)\rho^{i-j-1}(U_{k+i}^2/\phi_{k+i}^4)U_{k+j},\\ &\sum_{j=1}^{b-1}\sum_{i=1}^{b} c_k^{-1}b^{-1}\sum_{l=0}^{b-1}(\rho^l - 1)\phi_{k+l+1}^{-2}1(j \le i-1)\rho^{i-j-1}\left(\frac{U_{k+i}^2}{\phi_{k+i}^4}\right)U_{k+j},\end{aligned} \tag{S10.81}$$

$$\left(\sum_{i=1}^{b}(\rho^{i-1}-1)(U_{k+i}^2/\phi_{k+i}^4)\right)\sum_{j=1}^{b-1}\left(\frac{c_k^{-1}}{b}\sum_{l=0}^{b-(j+1)}\frac{\rho^l}{\phi_{k+j+l+1}^2}\right)U_{k+j},$$

$$\left(c_k^{-1}b^{-1}\sum_{l=0}^{b-1}(\rho^l-1)\phi_{k+l+1}^{-2}\right)$$

$$\times\left(\sum_{j=1}^{b-1}\frac{c_k^{-1}}{b}\sum_{l=0}^{b-(j+1)}\frac{\rho^l}{\phi_{k+j+l+1}^2}U_{k+j}\right)\sum_{i=1}^{b}\left(\frac{U_{k+i}^2}{\phi_{k+i}^4}\right).$$

From the LLN in (S10.78) and from $c_k^{-1}=O_p(1)$ it follows that to show that the expressions in (S10.81) multiplied by $(1-\rho)^{-1/2}b^{-2}$ are $o_p(1)$, it is sufficient to show that

$$\text{(S10.82)}\quad (1-\rho)^{1/2}b^{-1}\sum_{j=1}^{b-1}\sum_{i=j+1}^{b}\frac{\rho^{i-1}-1}{b(1-\rho)}\rho^{i-j-1}\left(\frac{U_{k+i}^2}{\phi_{k+i}^4}\right)U_{k+j}=o_p(1),$$

$$(1-\rho)^{1/2}b^{-1}\sum_{j=1}^{b-1}\sum_{i=j+1}^{b}\rho^{i-j-1}\left(\frac{U_{k+i}^2}{\phi_{k+i}^4}\right)U_{k+j}=o_p(1),$$

$$(1-\rho)^{1/2}b^{-1}\sum_{j=1}^{b-1}\sum_{l=0}^{b-(j+1)}\rho^l\phi_{k+j+l+1}^{-2}U_{k+j}=o_p(1).$$

To see this, note that the first row in (S10.81) is clearly implied by the first row in (S10.82). The second row in (S10.81) is implied by the second row in (S10.82) because in (S10.81) we apply the LLN in (S10.78) to $b^{-1}\sum_{l=0}^{b-1}(\rho^l-1)\phi_{k+l+1}^{-2}$, which is thus $O_p(b(1-\rho))$. The same LLN argument applied to $b^{-1}\sum_{i=1}^{b}(\rho^{i-1}-1)(U_{k+i}^2/\phi_{k+i}^4)$ shows that the third row in (S10.81) is implied by the third row in (S10.82). The fourth row in (S10.81) is implied by the previous arguments and $b^{-1}\sum_{i=1}^{b}(U_{k+i}^2/\phi_{k+i}^4)=O_p(1)$.

For the third term in (S10.82), by Markov's inequality, it is enough to show that

$$\text{(S10.83)}\quad (1-\rho)b^{-2}\sum_{j=1}^{b-1}\sum_{l=0}^{b-(j+1)}\sum_{i=1}^{b-1}\sum_{m=0}^{b-(i+1)}\rho^{l+m}E\phi_{k+j+l+1}^{-2}U_{k+j}\phi_{k+i+m+1}^{-2}U_{k+i}$$
$$=o(1).$$

We can assume that $i\neq j$ because the contributions of all summands with $i=j$ can be bounded by $(1-\rho)b^{-2}\sum_{i=1}^{b-1}\sum_{l,m=0}^{b-(i+1)}E\phi_{k+j+l+1}^{-2}\phi_{k+i+m+1}^{-2}U_{k+i}^2$, which is $o(1)$ because $E\phi_{k+j+l+1}^{-2}\phi_{k+i+m+1}^{-2}U_{k+i}^2$ is uniformly bounded by Assumption INNOV(iv) and $(1-\rho)b^{-2}b^3=o(1)$. Using the same argument we can assume that all subindices $k+j+l+1, k+i+m+1, k+j$, and $k+i$ are different and also that $i>j$. We have to distinguish two subcases, namely $k+j+l+1>k+i$ and $k+j+l+1<k+i$. The contributions of all summands in the l.h.s. of (S10.83) satisfying $k+j+l+1>k+i$ can be bounded by

$$\text{(S10.84)}\quad O(1)(1-\rho)b^{-2}\sum_{i>j=1}^{b-1}\sum_{l=0}^{b-(j+1)}\sum_{\substack{m=0,\\ m\neq l}}^{b-(i+1)}(i-j)^{-3-\varepsilon}=O(1-\rho)\sum_{i>j=1}^{b-1}(i-j)^{-3-\varepsilon}$$
$$=o(1),$$

where the first expression uses the strong-mixing inequality (S10.15) and Assumption INNOV(iv) and the last equality uses $b(1-\rho)\to 0$. The contributions of all summands in the l.h.s. of (S10.83) satisfying $k+j+l+1\le k+i$ can be bounded by

$$\text{(S10.85)}\quad O(1)(1-\rho)b^{-2}\sum_{i>j=1}^{b-1}\sum_{l=0}^{b-(j+1)}\sum_{\substack{m=0,\\ m\neq l}}^{b-(i+1)}(m+1)^{-3-\varepsilon}$$
$$=O(1-\rho)b\sum_{m=0}^{b}(m+1)^{-3-\varepsilon}=o(1).$$

The first and second terms in (S10.82) are handled in exactly the same way. For the first term, recall that $(b(1-\rho))^{-1}(\rho^{i-1}-1)$ is $O(1)$ uniformly in i.

Case (iv). $n(1-\rho)\to h_1\in[0,\infty)$: Because $n(1-\rho)=h_n\to h_1<\infty$, it follows that $h_n=O(1)$ and $\rho_n\to 1$. Hence, $\exp(-h_n^*/n)=\rho_n\to 1$ and $h_n^*=o(n)$. By a mean-value expansion of $\exp(-h_n^*/n)$ about 0,

$$\text{(S10.86)}\quad 0=\rho_n-\rho_n=\exp(-h_n^*/n)-(1-h_n/n)=h_n/n-\exp(-h_n^{**}/n)h_n^*/n,$$

where $h_n^{**}=o(n)$ given that $h_n^*=o(n)$. Hence, $h_n-(1+o(1))h_n^*=0$ and thus $h_n^*/h_n\to 1$. Hence, $h_n^*=O(1)$ and $t_n=bh_n^*/n\to 0$. The proof for $\xi_{1,k}=o_p(1)$ and $\xi_{2,k}=o_p(1)$ used in case (iii) then goes through. *Q.E.D.*

S10.2.5. *Verification of Assumption EE*

In this section, we verify Assumption EE for Model 1. We verify Assumption EE using the same argument as for Assumption E given above, but with $T_{n,k}=S_{1,k}S_{2,k}^{-1/2}$ replaced by $d_{b_n}(\gamma_{n,h})\widehat{\sigma}_{n,b_n,k}$, where $d_{b_n}(\gamma_{n,h})$ is the normalization constant that appears in Assumption BB and is defined in (S10.2). In case (i) of the verification of Assumption E above, where $b(1-\rho)\to\infty$, we have $d_{b_n}(\gamma_{n,h})\widehat{\sigma}_{n,b_n,k}\to_p 1$ by Proposition S1(b). Thus, (S10.19) trivially holds in this case. In cases (ii)–(iv), we have $d_{b_n}(\gamma_{n,h})\widehat{\sigma}_{n,b_n,k}=S_{2,k}^{1/2}S_{3,k}^{-1}$ for

$S_{3,k} = b^{-2}X_1'M_{X_2}X_1$, where as above (with abuse of notation) X_1 and X_2 denote b-vectors containing data from the kth subsample. It is sufficient to show the equivalent of (S10.39) for $S_{3,k}$:

$$\text{(S10.87)}\quad S_{3,k} = \widetilde{S}_{3,k} + \xi_{3,k}$$
$$\text{for some } \widetilde{S}_{3,k} \text{ that is separated from } S_{3,0} \text{ by}$$
$$b \text{ time periods } \forall k \geq k_n \text{ and } \xi_{3,k} = o_p(1).$$

Easy calculations show that $S_{3,k} = b^{-2}\sum_{i=1}^{b} f_{k,i}^2/\phi_{k+i}^2$. Set $\widetilde{S}_{3,k} = b^{-2}\sum_{i=1}^{b}(f_{k,i}^t)^2/\phi_{k+i}^2$ and $\xi_{3,k} = b^{-2}\sum_{i=1}^{b}(f_{k,i}^2 - (f_{k,i}^t)^2)/\phi_{k+i}^2$. Then, proceeding exactly as in the verification of $S_{2,k} = \widetilde{S}_{2,k} + \xi_{2,k}$ in (S10.39) in the proof of Assumption E, (S10.87) follows.

S10.2.6. *Verification of Assumption HH*

Given the definitions in (S10.2), Assumption HH holds by the following calculations. For all sequences $\{\gamma_{n,h} = (\gamma_{n,h,1}, \gamma_{n,h,2}, \gamma_{n,h,3}) \in \Gamma : n \geq 1\}$ for which $b_n\gamma_{n,h,1} \to g_1$ for some $g_1 \in R_{+,\infty}$, if $b_n\gamma_{n,h,1} \to g_1 = \infty$, then $n\gamma_{n,h,1} \to \infty$ and

$$\text{(S10.88)}\quad \frac{a_{b_n}(\gamma_{n,h})}{a_n(\gamma_{n,h})} = \frac{b_n^{1/2}d_{b_n}(\gamma_{n,h})}{n^{1/2}d_n(\gamma_{n,h})} = \left(\frac{b_n}{n}\right)^{1/2} \to 0$$

using Assumption C(ii). If $n\gamma_{n,h,1} \to h_1 = \infty$ and $b_n\gamma_{n,h,1} \to g_1 < \infty$, let $h_{n,1} = n(1-\rho)$ and let $h_{n,1}^*$ be defined by $\rho = \exp(-h_{n,1}^*/n)$. By Lemma S1 and (S10.3), $b_n/(n^{1/2}d_n(\gamma_{n,h})) = O((1-\rho)^{1/2}b_n/n^{1/2}) = O((h_{n,1}/n)^{1/2}b_n/n^{1/2}) = O(h_{n,1}^{1/2}b_n/n)$. Given that $n\gamma_{n,h,1} \to h_1 = \infty$ and $b_n\gamma_{n,h,1} \to g_1 < \infty$ we are either in case (ii) or case (iii) of the proof of Assumption E. In case (iii), we showed above that $t_n = b_nh_{n,1}^*/n \to 0$, $h_{n,1}/h_{n,1}^* \to 1$, and $h_{n,1} \to \infty$. Therefore, $O(h_{n,1}^{1/2}b_n/n) = O(t_nh_{n,1}^{-1/2}) = o(1)$. In case (ii), $t_n = (b_nh_{n,1}/n)(h_{n,1}^*/h_{n,1}) \to g_1$ and thus $O(h_{n,1}^{1/2}b_n/n) = O(t_nh_{n,1}^{-1/2}) = o(1)$ because $h_{n,1} \to \infty$. Therefore,

$$\text{(S10.89)}\quad \frac{a_{b_n}(\gamma_{n,h})}{a_n(\gamma_{n,h})} = \frac{b_n}{n^{1/2}d_n(\gamma_{n,h})} \to 0.$$

If $n\gamma_{n,h,1} \to h_1 < \infty$, then

$$\text{(S10.90)}\quad \frac{a_{b_n}(\gamma_{n,h})}{a_n(\gamma_{n,h})} = \frac{b_n}{n} \to 0$$

using Assumption C(ii).

S11. CONSERVATIVE MODEL-SELECTION EXAMPLE

S11.1. *The Model*

Here we establish the asymptotic distribution of the test statistic $T_n^*(\theta_0)$ and verify Assumption G for this example.

The model is

$$\text{(S11.1)}\quad y_i = x_{1i}^*\theta + x_{2i}^*\beta_2 + x_{3i}^{*\prime}\beta_3 + \sigma\varepsilon_i \quad \text{for} \quad i = 1, \ldots, n, \quad \text{where}$$
$$x_i^* = (x_{1i}^*, x_{2i}^*, x_{3i}^{*\prime})' \in R^k, \quad \beta = (\theta, \beta_2, \beta_3')' \in R^k,$$

$x_{1i}^*, x_{2i}^*, \theta, \beta_2, \sigma, \varepsilon_i \in R$, and $x_{3i}^*, \beta_3 \in R^{k-2}$. The observations $\{(y_i, x_i^*) : i = 1, \ldots, n\}$ are i.i.d. The scaled error ε_i has mean 0 and variance 1 conditional on x_i^*. We consider testing $H_0 : \theta = \theta_0$ after carrying out a model-selection procedure to determine whether x_{2i}^* should enter the model. The model-selection procedure is based on a t test of $H_0^* : \beta_2 = 0$.

The inference problem described above covers the following (seemingly more general) inference problem. Consider the model

$$\text{(S11.2)}\quad y_i = z_i'\tau + \sigma\varepsilon_i \quad \text{for} \quad i = 1, \ldots, n, \quad \text{where}$$
$$z_i = (z_{1i}', z_{2i})' \in R^k, \quad \tau = (\tau_1', \tau_2)' \in R^k,$$

$z_{1i}, \tau_1 \in R^{k-1}$, and $z_{2i}, \tau_2 \in R$. We are interested in testing $\overline{H}_0 : a'\tau = \theta_0$ for a given vector $a \in R^k$ with $a \neq e_k$, where $e_k = (0, \ldots, 0, 1)'$, after using a (fixed-critical-value) t test to determine whether z_{2i} should enter the model. This testing problem can be transformed into the former one by writing

$$\text{(S11.3)}\quad \theta = a'\tau, \quad \beta_2 = \tau_2, \quad \beta_3 = B'\tau$$

for some matrix $B \in R^{k\times(k-2)}$ such that $D = [a : e_k : B] \in R^{k\times k}$ is nonsingular. As defined, $\beta = D'\tau$. Define $x_i^* = D^{-1}z_i$. Then $x_i^{*\prime}\beta = z_i'\tau$ and $H_0 : \theta = \theta_0$ is equivalent to $\overline{H}_0 : a'\tau = \theta_0$.

We now return to the model in (S11.1). To define the test statistic $T_n^*(\theta_0)$, we write the variables in matrix notation, and define the first and second regressors after projecting out the remaining regressors using finite-sample projections:

$$\text{(S11.4)}\quad Y = (y_1, \ldots, y_n)',$$
$$X_j^* = (x_{j1}^*, \ldots, x_{jn}^*)' \in R^n \quad \text{for} \quad j = 1, 2,$$
$$X_3^* = [x_{31}^* : \cdots : x_{3n}^*]' \in R^{n\times(k-2)},$$
$$X_j = M_{X_3^*}X_j^* \in R^n \quad \text{for} \quad j = 1, 2,$$
$$X = [X_1 : X_2] \in R^{n\times 2},$$

where $M_{X_3^*} = I_n - P_{X_3^*}$ and $P_{X_3^*} = X_3^*(X_3^{*\prime}X_3^*)^{-1}X_3^{*\prime}$. The n-vectors X_1 and X_2 correspond to the n-vectors X_1^* and X_2^*, respectively, with X_3^* projected out.

The restricted and unrestricted least squares (LS) estimators of θ and the unrestricted LS estimator of β_2 are

$$\text{(S11.5)}\quad \begin{aligned} \widetilde{\theta} &= (X_1'X_1)^{-1}X_1'Y, \\ \widehat{\theta} &= (X_1'M_{X_2}X_1)^{-1}X_1'M_{X_2}Y, \\ \widehat{\beta}_2 &= (X_2'M_{X_1}X_2)^{-1}X_2'M_{X_1}Y. \end{aligned}$$

The model-selection test rejects $H_0^*: \beta_2 = 0$ if

$$\text{(S11.6)}\quad |T_{n,2}| = \left| \frac{n^{1/2}\widehat{\beta}_2}{\widehat{\sigma}(n^{-1}X_2'M_{X_1}X_2)^{-1/2}} \right| > c, \quad \text{where}$$
$$\widehat{\sigma}^2 = (n-k)^{-1}Y'M_{[X_1^*:X_2^*:X_3^*]}Y$$

and $c > 0$ is a given critical value that does not depend on n. Typically, $c = z_{1-\alpha/2}$ for some $\alpha > 0$. The estimator $\widehat{\sigma}^2$ of σ^2 is the standard (unrestricted) unbiased estimator.

The test statistic, $T_n^*(\theta_0)$, for testing $H_0: \theta = \theta_0$ is a t statistic based on the restricted LS estimator of θ when the null hypothesis $H_0^*: \beta_2 = 0$ is not rejected and the unrestricted LS estimator when it is rejected:

$$\text{(S11.7)}\quad T_n^*(\theta_0) = \widetilde{T}_{n,1}(\theta_0)1(|T_{n,2}| \le c) + \widehat{T}_{n,1}(\theta_0)1(|T_{n,2}| > c), \quad \text{where}$$
$$\widetilde{T}_{n,1}(\theta_0) = \frac{n^{1/2}(\widetilde{\theta} - \theta_0)}{\widehat{\sigma}(n^{-1}X_1'X_1)^{-1/2}},$$
$$\widehat{T}_{n,1}(\theta_0) = \frac{n^{1/2}(\widehat{\theta} - \theta_0)}{\widehat{\sigma}(n^{-1}X_1'M_{X_2}X_1)^{-1/2}}.$$

Note that both $\widetilde{T}_{n,1}(\theta_0)$ and $\widehat{T}_{n,1}(\theta_0)$ are defined using the unrestricted estimator $\widehat{\sigma}$ of σ. One could define $\widetilde{T}_{n,1}(\theta_0)$ using the restricted LS estimator of σ, but this is not desirable because it leads to an inconsistent estimator of σ under sequences of parameters $\{\beta_2 = \beta_{2n} : n \ge 1\}$ that satisfy $\beta_{2n} \to 0$ and $n^{1/2}\beta_{2n} \nrightarrow 0$ as $n \to \infty$. For subsampling tests, one could define $\widetilde{T}_{n,1}(\theta_0)$ and $\widehat{T}_{n,1}(\theta_0)$ with $\widehat{\sigma}$ deleted because the scale of the subsample statistics offsets that of the original sample statistic. This does not work for hybrid tests because Assumption K fails if $\widehat{\sigma}$ is deleted.

The "model-selection" estimator, $\overline{\theta}$, of θ is

$$\text{(S11.8)}\quad \overline{\theta} = \widetilde{\theta}1(|T_{n,2}| \le c) + \widehat{\theta}1(|T_{n,2}| > c).$$

This estimator is used to recenter the subsample statistics. (One could also use the unrestricted estimator $\widehat{\theta}$ to recenter the subsample statistics.)

S11.2. *Proof of the Asymptotic Distributions of the Test Statistics*

In this section, we establish the asymptotic distribution J_h^* of $T_n^*(\theta_0)$ under a sequence of parameters $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) : n \geq 1\}$ (where $n^{1/2}\gamma_{n,1} \to h_1$, $\gamma_{n,2} \to h_2$, and $\gamma_{n,3} \in \Gamma_3(\gamma_{n,1}, \gamma_{n,2})$ for all n). Parts of the proof are closely related to calculations in Leeb (2006) and Leeb and Pötscher (2005). No papers in the literature that we are aware of consider subsampling methods for post-model-selection inference. For FCV tests, the main differences from Leeb (2006) are that here we consider (i) model selection among two models, (ii) errors that may be nonnormal, (iii) i.i.d. regressors, and (iv) t statistics, and (v) we prove the asymptotic results directly. In contrast, Leeb (2006) considered (i) multiple models, (ii) normal errors, (iii) fixed regressors, and (iv) normalized estimators, and (v) he proved the asymptotic results by establishing finite-sample results for the normal error case and taking their limits. The results in Leeb and Pötscher (2005) are a two-model special case of those given in Leeb (2006).

Using the definition of $T_n^*(\theta_0)$ in this example, we have

$$\text{(S11.9)}\quad P_{\theta_0,\gamma_n}(T_n^*(\theta_0) \leq x) = P_{\theta_0,\gamma_n}(\widetilde{T}_{n,1}(\theta_0) \leq x \;\&\; |T_{n,2}| \leq c) + P_{\theta_0,\gamma_n}(\widehat{T}_{n,1}(\theta_0) \leq x \;\&\; |T_{n,2}| > c).$$

Hence, it suffices to determine the limits of the two summands on the right-hand side. With this in mind, we show below that under $\{\gamma_n : n \geq 1\}$, when $|h_1| < \infty$,

$$\text{(S11.10)}\quad \begin{pmatrix} \widetilde{T}_{n,1}(\theta_0) \\ T_{n,2} \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \widetilde{Z}_{h,1} \\ Z_{h,2} \end{pmatrix} \sim N\left(\begin{pmatrix} -h_1 h_2 (1-h_2^2)^{-1/2} \\ h_1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right),$$
$$\begin{pmatrix} \widehat{T}_{n,1}(\theta_0) \\ T_{n,2} \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \widehat{Z}_{h,1} \\ Z_{h,2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ h_1 \end{pmatrix}, \begin{pmatrix} 1 & h_2 \\ h_2 & 1 \end{pmatrix}\right).$$

Given this, we have

$$\begin{aligned}\text{(S11.11)}\quad & P_{\theta_0,\gamma_n}(\widetilde{T}_{n,1}(\theta_0) \leq x \;\&\; |T_{n,2}| \leq c) \\ & \to P(\widetilde{Z}_{h,1} \leq x \;\&\; |Z_{h,2}| \leq c) \\ & = \Phi\big(x + h_1 h_2 (1-h_2^2)^{-1/2}\big)\Delta(h_1, c), \quad \text{where} \\ & \Delta(a,b) = \Phi(a+b) - \Phi(a-b),\end{aligned}$$

the equality uses the independence of $\widetilde{Z}_{h,1}$ and $Z_{h,2}$ and the normality of their distributions, and $\Delta(a,b) = \Delta(-a,b)$. In addition, we have

$$\text{(S11.12)}\quad P_{\theta_0,\gamma_n}(\widehat{T}_{n,1}(\theta_0) \leq x \;\&\; |T_{n,2}| > c) \to P(\widehat{Z}_{h,1} \leq x \;\&\; |Z_{h,2}| > c).$$

Next, we calculate the limiting probability in (S11.12). Let $f(z_2|z_1)$ denote the conditional density of $Z_{h,2}$ given $\widehat{Z}_{h,1}$. Let $\phi(z_1)$ denote the standard normal density. Given that

$$\text{(S11.13)}\quad \begin{pmatrix} \widehat{Z}_{h,1} \\ Z_{h,2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ h_1 \end{pmatrix}, \begin{pmatrix} 1 & h_2 \\ h_2 & 1 \end{pmatrix} \right),$$

the conditional distribution of $Z_{h,2}$ given $\widehat{Z}_{h,1} = z_1$ is $N(h_1 + h_2 z_1, 1 - h_2^2)$. We have

$$\begin{aligned}
\text{(S11.14)}\quad & P(\widehat{Z}_{h,1} \le x \;\&\; |Z_{h,2}| > c) \\
&= \int_{-\infty}^{x} \int_{|z_2|>c} f(z_2|z_1)\phi(z_1)\, dz_2\, dz_1 \\
&= \int_{-\infty}^{x} \left(1 - \int_{|z_2|\le c} (1-h_2^2)^{-1/2} \phi\left(\frac{z_2 - (h_1 + h_2 z_1)}{(1-h_2^2)^{1/2}}\right) dz_2\right) \phi(z_1)\, dz_1 \\
&= \int_{-\infty}^{x} \left(1 - \int_{|\overline{z}_2| \le c(1-h_2^2)^{-1/2}} \phi\left(\overline{z}_2 - \frac{h_1 + h_2 z_1}{(1-h_2^2)^{1/2}}\right) d\overline{z}_2\right) \phi(z_1)\, dz_1 \\
&= \int_{-\infty}^{x} \left(1 - \Delta\left(\frac{h_1 + h_2 z}{(1-h_2^2)^{1/2}}, \frac{c}{(1-h_2^2)^{1/2}}\right)\right) \phi(z)\, dz,
\end{aligned}$$

where the second equality holds by (S11.13), the third equality holds by change of variables with $\overline{z}_2 = z_2(1-h_2^2)^{-1/2}$, and the last equality holds by the definition of $\Delta(a, b)$.

Combining (S11.11), (S11.12), and (S11.14) gives the desired result,

$$\begin{aligned}
\text{(S11.15)}\quad J_h^*(x) &= \Phi\big(x + h_1 h_2 (1-h_2^2)^{-1/2}\big)\Delta(h_1, c) \\
&\quad + \int_{-\infty}^{x} \left(1 - \Delta\left(\frac{h_1 + h_2 t}{(1-h_2^2)^{1/2}}, \frac{c}{(1-h_2^2)^{1/2}}\right)\right) \phi(t)\, dt,
\end{aligned}$$

when $|h_1| < \infty$. When $|h_1| = \infty$, $J_h^*(x) = \Phi(x)$ (which equals the limit as $|h_1| \to \infty$ of $J_h^*(x)$ defined in (S11.15)). The proof of the latter result is given below in the paragraph containing (S11.29).

We now show that under $\{\gamma_n : n \ge 1\}$, when $|h_1| < \infty$, (S11.10) holds. Let $X_j^\perp = (x_{j1}^\perp, \ldots, x_{jn}^\perp)' \in R^n$ for $j = 1, 2$ and $X^\perp = (X_1^\perp, X_2^\perp)' \in R^{n\times 2}$. We use the following lemma.

LEMMA S3: *Given the assumptions stated in Section* 2.2 *of the paper, under a sequence of parameters* $\{\gamma_n = (\gamma_{n,1}, \gamma_{n,2}, \gamma_{n,3}) : n \ge 1\}$ (*where* $n^{1/2}\gamma_{n,1} \to h_1$, $\gamma_{n,2} \to h_2$, *and* $\gamma_{n,3} \in \Gamma_3(\gamma_1, \gamma_2)$ *for all* n), *and for* $Q = Q_n$ *as defined in* (2.7) *of the paper with the* (j, m) *element denoted* $Q_{n,jm}$, *we have* (a) $n^{-1}X'X - Q_n \to_p 0$, (b) $n^{-1}X_2'M_{X_1}X_2 - (Q_{n,22} - Q_{n,12}^2 Q_{n,11}^{-1}) \to_p 0$, (c) $n^{-1}X_1'M_{X_2}X_1 - (Q_{n,11} -$

$Q_{n,12}^2 Q_{n,22}^{-1}) \to_p 0$, (d) $\widehat{\sigma}/\sigma_n \to_p 1$, *and* (e) $n^{-1/2}X_j'\varepsilon = n^{-1/2}X_j^{\perp\prime}\varepsilon + o_p(1) = O_p(1)$ *for* $j = 1, 2$.

PROOF: The proofs of parts (a)–(d) are standard using a weak law of large numbers (WLLN) for $L^{1+\delta}$-bounded independent random variables for some $\delta > 0$ and taking into account the fact that $X_j = M_{X_3^*}X_j^*$ for $j = 1, 2$.

Next, we prove part (e). By definition of X_j, we have

$$\text{(S11.16)}\quad \begin{aligned} n^{-1/2}X_j'\varepsilon &= n^{-1/2}X_j^{*\prime}\varepsilon - n^{-1}X_j^{*\prime}X_3^*(n^{-1}X_3^{*\prime}X_3^*)^{-1}n^{-1/2}X_3^{*\prime}\varepsilon \\ &= n^{-1/2}X_j^{*\prime}\varepsilon - E_{G_n}x_{ji}^*x_{3i}^{*\prime}\big(E_{G_n}x_{3i}^*x_{3i}^{*\prime}\big)^{-1}n^{-1/2}X_3^{*\prime}\varepsilon + o_p(1) \\ &= n^{-1/2}X_j^{\perp\prime}\varepsilon + o_p(1), \end{aligned}$$

where G_n denotes the distribution of (ε_i, x_i^*) under γ_n, the second equality holds by the same WLLN as above combined with the Lindeberg triangular array central limit theorem (CLT) applied to $n^{-1/2}X_3^{*\prime}\varepsilon$, which yields $n^{-1/2}X_3^{*\prime}\varepsilon = O_p(1)$, and the third equality uses the definition that $x_{ji}^{\perp} = x_{ji}^* - E_{G_n}x_{ji}^*x_{3i}^{*\prime}(E_{G_n}x_{3i}^*x_{3i}^{*\prime})^{-1}x_{3i}^*$. The second equality of part (e) holds by the Lindeberg CLT. The Lindeberg condition is implied by a Liapounov condition, which holds by the moment bound in $\Gamma_3(\gamma_1, \gamma_2)$. *Q.E.D.*

We now prove the first result of (S11.10) (which assumes $|h_1| < \infty$). Using (S11.5) and (S11.6), we have

$$\text{(S11.17)}\quad \begin{aligned} T_{n,2} &= \frac{n^{1/2}\beta_2/\sigma_n + (n^{-1}X_2'M_{X_1}X_2)^{-1}n^{-1/2}X_2'M_{X_1}\varepsilon}{(\widehat{\sigma}/\sigma_n)(n^{-1}X_2'M_{X_1}X_2)^{-1/2}} \\ &= n^{1/2}\frac{\beta_2}{\sigma_n(Q_n^{22})^{1/2}}(1 + o_p(1)) \\ &\quad + (Q_n^{22})^{1/2}n^{-1/2}X_2'\big(I_n - P_{X_1}\big)\varepsilon(1 + o_p(1)) \\ &= n^{1/2}\gamma_{n,1}(1 + o_p(1)) \\ &\quad + (Q_n^{22})^{1/2}(e_2 - Q_{n,12}Q_{n,11}^{-1}e_1)'n^{-1/2}X'\varepsilon(1 + o_p(1)) \\ &= h_1 + (Q_n^{22})^{1/2}(e_2 - Q_{n,12}Q_{n,11}^{-1}e_1)'n^{-1/2}X^{\perp\prime}\varepsilon + o_p(1), \end{aligned}$$

where $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)'$, $e_1 = (1, 0)'$, $e_2 = (0, 1)'$, $Q_n = E_{G_n}x_i^{\perp}x_i^{\perp\prime}$, Q_n^{22} is the $(2, 2)$ element of Q_n^{-1}, the second equality uses Lemma S3(b) and (d), the fact that $Q_n^{22} = (Q_{n,22} - Q_{n,12}^2 Q_{n,11}^{-1})^{-1}$, and the fact that $\lambda_{\min}(Q_n) \geq \kappa > 0$ by definition of $\Gamma_3(\gamma_1, \gamma_2)$, the third equality uses the definition of $\gamma_{n,1}$ and Lemma S3(a), and the fourth equality holds by the assumption that $n^{1/2}\gamma_{n,1} \to h_1$ and Lemma S3(e).

Using (S11.5) and (S11.7), we have

$$\text{(S11.18)}\quad \widetilde{T}_{n,1}(\theta_0) = \frac{n^{1/2}(n^{-1}X_1'X_1)^{-1}n^{-1}X_1'X_2\beta_2/\sigma_n + (n^{-1}X_1'X_1)^{-1}n^{-1/2}X_1'\varepsilon}{(\widehat{\sigma}/\sigma_n)(n^{-1}X_1'X_1)^{-1/2}}$$
$$= n^{1/2}\frac{Q_{n,12}\beta_2}{\sigma_n Q_{n,11}^{1/2}}(1+o_p(1)) + Q_{n,11}^{-1/2}n^{-1/2}e_1'X'\varepsilon(1+o_p(1))$$
$$= h_1\frac{Q_{n,12}(Q_n^{22})^{1/2}}{Q_{n,11}^{1/2}} + Q_{n,11}^{-1/2}n^{-1/2}e_1'X^{\perp\prime}\varepsilon + o_p(1),$$

where the second equality uses Lemma S3(a) and (d), and the third equality uses the assumption that $n^{1/2}\gamma_{n,1} = n^{1/2}\beta_2/(\sigma_n^2 Q_n^{22})^{1/2} \to h_1$ and Lemma S3(e).

We have

$$\text{(S11.19)}\quad Q_n^{-1} = \frac{1}{Q_{n,11}Q_{n,22} - Q_{n,12}^2}\begin{bmatrix} Q_{n,22} & -Q_{n,12} \\ -Q_{n,12} & Q_{n,11}\end{bmatrix} \quad \text{and so}$$
$$\gamma_{n,2} = \frac{Q_n^{12}}{(Q_n^{11}Q_n^{22})^{1/2}} = \frac{-Q_{n,12}}{(Q_{n,11}Q_{n,22})^{1/2}},$$
$$Q_n^{22} = \frac{Q_{n,11}}{Q_{n,11}Q_{n,22} - Q_{n,12}^2} = (Q_{n,22})^{-1}(1-\gamma_{n,2}^2)^{-1},$$

where the first equality in the second line holds by the definition of $\gamma_{n,2}$ in (2.6) of the paper. This yields

$$\text{(S11.20)}\quad \frac{Q_{n,12}(Q_n^{22})^{1/2}}{Q_{n,11}^{1/2}} = \frac{Q_{n,12}(1-\gamma_{n,2}^2)^{-1/2}}{Q_{n,11}^{1/2}Q_{n,22}^{1/2}}$$
$$= -\gamma_{n,2}(1-\gamma_{n,2}^2)^{-1/2} = -h_2(1-h_2^2)^{-1/2} + o(1).$$

Combining (S11.17), (S11.18), and (S11.20) gives

$$\text{(S11.21)}\quad \begin{pmatrix} \widetilde{T}_{n,1}(\theta_0) \\ T_{n,2} \end{pmatrix} = \begin{pmatrix} -h_1h_2(1-h_2^2)^{-1/2} + Q_{n,11}^{-1/2}n^{-1/2}e_1'X^{\perp\prime}\varepsilon \\ h_1 + (Q_n^{22})^{1/2}(e_2 - Q_{n,12}Q_{n,11}^{-1}e_1)'n^{-1/2}X^{\perp\prime}\varepsilon \end{pmatrix} + o_p(1).$$

The first result of (S11.10) holds by (S11.21), the Lindeberg CLT, and the Cramér–Wold device. The Lindeberg condition is implied by a Liapounov condition, which holds by the moment bound in $\Gamma_3(\gamma_1, \gamma_2)$. The asymptotic covariance matrix is I_2 by the following calculations. The $(1,2)$ element of the asymptotic covariance matrix equals

$$\text{(S11.22)}\quad E_{G_n}Q_{n,11}^{-1/2}e_1'n^{-1}X^{\perp\prime}X^{\perp}(e_2 - Q_{n,12}Q_{n,11}^{-1}e_1)(Q_n^{22})^{1/2}$$
$$= Q_{n,11}^{-1/2}e_1'Q_n(e_2 - Q_{n,12}Q_{n,11}^{-1}e_1)(Q_n^{22})^{1/2} = 0,$$

where the first equality holds because $E_{G_n} x_i^{\perp} x_i^{\perp\prime} = Q_n$ and the second equality holds by algebra. The $(1,1)$ element equals

$$\text{(S11.23)}\quad E_{G_n} Q_{n,11}^{-1/2} e_1' n^{-1} X^{\perp\prime} X^{\perp} e_1 Q_{n,11}^{-1/2} = Q_{n,11}^{-1/2} e_1' Q_n e_1 Q_{n,11}^{-1/2} = 1.$$

The $(2,2)$ element equals

$$\begin{aligned}\text{(S11.24)}\quad & E_{G_n} (Q_n^{22})^{1/2} (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)' n^{-1} X^{\perp\prime} X^{\perp} (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)(Q_n^{22})^{1/2} \\ &= (Q_n^{22})^{1/2} (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)' Q_n (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)(Q_n^{22})^{1/2} \\ &= (Q_n^{22})^{1/2} (Q_{n,22}(1 - \gamma_{n,2}^2))(Q_n^{22})^{1/2} = 1,\end{aligned}$$

where the second equality holds by algebra and the definition of $\gamma_{n,2}$ and the third equality holds by the third result in (S11.19). This completes the proof of the first result in (S11.10).

Next, we prove the second result in (S11.10). Using (S11.7), we have

$$\begin{aligned}\text{(S11.25)}\quad \widehat{T}_{n,1}(\theta_0) &= \frac{(n^{-1} X_1' M_{X_2} X_1)^{-1} n^{-1/2} X_1' M_{X_2} \varepsilon}{(\widehat{\sigma}/\sigma_n)(n^{-1} X_1' M_{X_2} X_1)^{-1/2}} \\ &= (Q_n^{11})^{1/2} (e_1 - Q_{n,12} Q_{n,22}^{-1} e_2)' n^{-1/2} X^{\perp\prime} \varepsilon + o_p(1),\end{aligned}$$

where the second equality holds analogously to (S11.17). Combining (S11.17) and (S11.25) gives

$$\text{(S11.26)}\quad \begin{pmatrix} \widehat{T}_{n,1}(\theta_0) \\ T_{n,2} \end{pmatrix} = \begin{pmatrix} (Q_n^{11})^{1/2} (e_1 - Q_{n,12} Q_{n,22}^{-1} e_2)' n^{-1/2} X^{\perp\prime} \varepsilon \\ h_1 + (Q_n^{22})^{1/2} (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)' n^{-1/2} X^{\perp\prime} \varepsilon \end{pmatrix} + o_p(1).$$

The second result of (S11.10) holds by (S11.26), the Lindeberg CLT, and the Cramér–Wold device. The Lindeberg condition holds as above. The 2×2 asymptotic covariance matrix has off-diagonal element h_2 and diagonal elements equal to 1 by the following calculations. The $(1,2)$ element equals

$$\begin{aligned}\text{(S11.27)}\quad & E_{G_n} (Q_n^{11})^{1/2} (e_1 - Q_{n,12} Q_{n,22}^{-1} e_2)' n^{-1} X^{\perp\prime} X^{\perp} (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)(Q_n^{22})^{1/2} \\ &= (Q_n^{11})^{1/2} (e_1 - Q_{n,12} Q_{n,22}^{-1} e_2)' Q_n (e_2 - Q_{n,12} Q_{n,11}^{-1} e_1)(Q_n^{22})^{1/2} \\ &= (Q_n^{11})^{1/2} (-Q_{n,12}(1 - Q_{n,12}^2 Q_{n,11}^{-1} Q_{n,22}^{-1}))(Q_n^{22})^{1/2} \\ &= (Q_{n,11}(1 - \gamma_{n,2}^2))^{-1/2} (-Q_{n,12}(1 - \gamma_{n,2}^2))(Q_{n,22}(1 - \gamma_{n,2}^2))^{-1/2} \\ &= \frac{-Q_{n,12}}{(Q_{n,11} Q_{n,22})^{1/2}} = \frac{Q_n^{12}}{(Q_n^{11} Q_n^{22})^{1/2}} = \gamma_{n,2} = h_2 + o(1),\end{aligned}$$

where the second equality holds by algebra, the third equality holds by the second and third results of (S11.19) and the third result of (S11.19) with 22 and

11 interchanged, and the fifth and sixth equalities hold by the second result of (S11.19).

The $(1,1)$ element equals

$$\text{(S11.28)}\quad E_{G_n}(Q_n^{11})^{1/2}(e_1 - Q_{n,12}Q_{n,22}^{-1}e_2)'n^{-1}X^{\perp\prime}X^{\perp}(e_1 - Q_{n,12}Q_{n,22}^{-1}e_2)(Q_n^{11})^{1/2} = (Q_n^{11})^{1/2}(e_1 - Q_{n,12}Q_{n,22}^{-1}e_2)'Q_n(e_1 - Q_{n,12}Q_{n,22}^{-1}e_2)(Q_n^{11})^{1/2} = 1,$$

where the second equality holds by an analogous argument to that in (S11.24). The $(2,2)$ element equals 1 by (S11.24). This completes the proof of the second result in (S11.10).

Finally, we show that $J_h^*(x) = \Phi(x)$ when $|h_1| = \infty$. Equations (S11.25) and (S11.28) hold in this case, so $\widehat{T}_{n,1}(\theta_0) \xrightarrow{d} N(0,1)$ under $\{\gamma_n : n \geq 1\}$. The first three equalities of (S11.17) hold when $|h_1| = \infty$ and show that $|T_{n,2}| \xrightarrow{p} \infty$. These results combine to yield

$$\text{(S11.29)}\quad P_{\theta_0,\gamma_n}(\widetilde{T}_{n,1}(\theta_0) \leq x \ \&\ |T_{n,2}| \leq c) = o(1),$$
$$P_{\theta_0,\gamma_n}(\widehat{T}_{n,1}(\theta_0) \leq x \ \&\ |T_{n,2}| > c) = P_{\theta_0,\gamma_n}(\widehat{T}_{n,1}(\theta_0) \leq x) + o(1) \to \Phi(x)$$

for all $x \in R$. This and (S11.9) combine to give $P_{\theta_0,\gamma_n}(T_n^*(\theta_0) \leq x) \to \Phi(x)$ and $J_h^*(x) = \Phi(x)$ when $|h_1| = \infty$.

S11.3. *Verification of Assumption G*

Assumption G is verified in the conservative model-selection example by using a variant of the argument in the proof of Lemma 4 in AG1 with $\tau_n = a_n = n^{1/2}$ and $d_n = 1$. In the present case, (8.16) of AG1 holds with

$$\text{(S11.30)}\quad R_n(t) = q_n^{-1}\sum_{j=1}^{q_n} 1\big(\big|b_n^{1/2}(\overline{\theta} - \theta_0)/\widehat{\sigma}_{n,b,j}^{(1)}\big| \geq t\big) + q_n^{-1}\sum_{j=1}^{q_n} 1\big(\big|b_n^{1/2}(\overline{\theta} - \theta_0)/\widehat{\sigma}_{n,b,j}^{(2)}\big| \geq t\big), \quad \text{where}$$
$$\widehat{\sigma}_{n,b,j}^{(1)} = \widehat{\sigma}_{n,b,j}(b_n^{-1}X_{1,n,b,j}'X_{1,n,b,j})^{-1/2},$$
$$\widehat{\sigma}_{n,b,j}^{(2)} = \widehat{\sigma}_{n,b,j}(b_n^{-1}X_{1,n,b,j}'M_{X_{2,n,b,j}}X_{1,n,b,j})^{-1/2},$$

and $(X_{1,n,b,j}, X_{2,n,b,j}, \widehat{\sigma}_{n,b,j})$ denotes $(X_1, X_2, \widehat{\sigma})$ based on the jth subsample rather than the full sample. (Equation (8.16) of AG1 holds with $R_n(t)$ defined as in (S11.30) for all three versions of the tests: $T_n(\theta_0) = T_n^*(\theta_0)$, $-T_n^*(\theta_0)$, and $|T_n^*(\theta_0)|$.) As in the proof of Lemma 4 of AG1, it suffices to show that $R_n(t)$ converges in probability to zero under all sequences $\{\gamma_{n,h} : n \geq 1\}$ for all $t > 0$. The assumption that $b_n/n \to 0$ and the result established below that

$n^{1/2}(\overline{\theta}-\theta_0)/\sigma_n = O_p(1)$ under all sequences $\{\gamma_{n,h}: n \geq 1\}$ imply that for all $\delta > 0$, with probability approaching 1,

$$\text{(S11.31)}\quad R_n(t) \leq R_n^{(1)}(\delta, t) + R_n^{(2)}(\delta, t), \quad \text{where}$$
$$R_n^{(m)}(\delta, t) = q_n^{-1}\sum_{j=1}^{q_n} 1\big(\delta\sigma_n/\widehat{\sigma}_{n,b,j}^{(m)} \geq t\big)$$

for $m = 1, 2$. The variance of $R_n^{(m)}(\delta, t)$ goes to zero under $\{\gamma_{n,h}: n \geq 1\}$ by the same U-statistic argument for i.i.d. observations as used to establish Assumption E of AG1 in the i.i.d. case; see Section 3.3 of AG1. The expectation of $R_n^{(m)}(\delta, t)$ equals $P_{\theta_0,\gamma_{n,h}}(\widehat{\sigma}_{n,b,j}^{(m)}/\sigma_n \leq \delta/t)$. We have

$$\text{(S11.32)}\quad \widehat{\sigma}_{n,b,j}^{(1)}/\sigma_n = (\widehat{\sigma}_{n,b,j}/\sigma_n)\big[(b_n^{-1}X_{1,n,b,j}'X_{1,n,b,j})^{-1/2} - Q_{n,11}^{-1/2} + Q_{n,11}^{-1/2}\big]$$
$$= Q_{n,11}^{-1/2} + o_p(1),$$

where the second equality holds by Lemma S3 (or, more precisely, by the same argument as used to prove Lemma S3). In addition, $Q_{n,11}^{-1/2}$ is bounded away from zero as $n \to \infty$ by the definition of $\Gamma_3(\gamma_1, \gamma_2)$. In consequence, the expectation of $R_n^{(1)}(\delta, t)$ goes to zero for all δ sufficiently small. Since the mean and variance of $R_n^{(1)}(\delta, t)$ go to zero, $R_n^{(1)}(\delta, t) \to_p 0$ for $\delta > 0$ sufficient small. An analogous argument shows that $R_n^{(2)}(\delta, t) \to_p 0$ for $\delta > 0$ sufficient small. These results and (S11.31) yield $R_n(t) \to_p 0$ under all sequences $\{\gamma_{n,h}: n \geq 1\}$, as desired.

It remains to show that $n^{1/2}(\overline{\theta}-\theta_0)/\sigma_n = O_p(1)$ under all sequences $\{\gamma_{n,h}: n \geq 1\}$. We consider two cases: $|h_1| = \infty$ and $|h_1| < \infty$. First, suppose $|h_1| = \infty$. Then the first three equalities of (S11.17) hold and show that $|T_{n,2}| \to_p \infty$. In addition, $n^{1/2}(\widehat{\theta}-\theta_0)/\sigma_n = (\widehat{\sigma}/\sigma_n)(n^{-1}X_1'M_{X_2}X_1)^{-1/2} \times \widehat{T}_{n,1}(\theta_0) = O_p(1)$ by (S11.10), Lemma S3(c), and the definition of $\Gamma_3(\gamma_1, \gamma_2)$. Combining these results gives that when $|h_1| = \infty$,

$$\text{(S11.33)}\quad n^{1/2}(\overline{\theta}-\theta_0)/\sigma_n = \big[n^{1/2}(\widetilde{\theta}-\theta_0)/\sigma_n\big]1(|T_{n,2}| \leq c)$$
$$+ \big[n^{1/2}(\widehat{\theta}-\theta_0)/\sigma_n\big]1(|T_{n,2}| > c)$$
$$= o_p(1) + O_p(1).$$

Next, suppose $|h_1| < \infty$. Then $\widehat{T}_{n,1}(\theta_0) = O_p(1)$ and $\widetilde{T}_{n,1}(\theta_0) = O_p(1)$ by (S11.10). In addition, $\widehat{\sigma}/\sigma_n \to_p 1$, $(n^{-1}X_1'X_1)^{-1/2} = O_p(1)$, and $(n^{-1}X_1'M_{X_2} \times X_1)^{-1/2} = O_p(1)$ by Lemma S3 and the definition of $\Gamma_3(\gamma_1, \gamma_2)$. Combining these results gives that when $|h_1| < \infty$,

$$\text{(S11.34)}\quad n^{1/2}(\overline{\theta}-\theta_0)/\sigma_n = \big[n^{1/2}(\widetilde{\theta}-\theta_0)/\sigma_n\big]1(|T_{n,2}| \leq c)$$
$$+ \big[n^{1/2}(\widehat{\theta}-\theta_0)/\sigma_n\big]1(|T_{n,2}| > c)$$

$$
\begin{aligned}
&= (\widehat{\sigma}/\sigma_n)(n^{-1}X_1'X_1)^{-1/2}\widetilde{T}_{n,1}(\theta_0)1(|T_{n,2}| \le c) \\
&\quad + (\widehat{\sigma}/\sigma_n)(n^{-1}X_1'M_{X_2}X_1)^{-1/2}\widehat{T}_{n,1}(\theta_0)1(|T_{n,2}| > c) \\
&= O_p(1),
\end{aligned}
$$

which completes the verification of Assumption G.

REFERENCES

ANDREWS, D. W. K., AND P. GUGGENBERGER (2008): "Asymptotics for LS, GLS, and Feasible GLS Statistics in an AR(1) Model With Conditional Heteroskedaticity," Unpublished Manuscript, Cowles Foundation, Yale University. [1,27]

——— (2009): "Validity of Subsampling and 'Plug-In Asymptotic' Inference for Parameters Defined by Moment Inequalities," *Econometric Theory*, 25 (forthcoming). [27]

——— (2010): "Applications of Subsampling, Hybrid, and Size-Correction Methods," *Journal of Econometrics* (forthcoming). [18]

DOUKHAN, P. (1994): *Mixing Properties and Examples*. Lecture Notes in Statistics, Vol. 85. New York: Springer Verlag. [33]

HALL, P. (1992): *The Bootstrap and Edgeworth Expansion*. New York: Springer Verlag. [2,22]

KRISTENSEN, D., AND O. LINTON (2006): "A Closed-Form Estimator for the GARCH(1, 1) Model," *Econometric Theory*, 22, 323–337. [1]

LEEB, H. (2006): "The Distribution of a Linear Predictor After Model Selection: Unconditional Finite-Sample Distributions and Asymptotic Approximations," in *2nd Lehmann Symposium—Optimality*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, Vol. 49. Beachwood, OH: Institute of Mathematical Statistics, 291–311. [55]

LEEB, H., AND B. M. PÖTSCHER (2005): "Model Selection and Inference: Facts and Fiction," *Econometric Theory*, 21, 21–59. [55]

WHITE, H. (1984): *Asymptotic Theory for Econometricians*. New York: Academic Press. [49]

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

and

Dept. of Economics, University of California–Los Angeles, Los Angeles, CA 90095, U.S.A.; guggenbe@econ.ucla.edu.

Manuscript received March, 2007; final revision received May, 2008.