Supplement to "Inference for Parameters Defined by Moment Inequalities: A Recommended Moment Selection Procedure"

Donald W. K. Andrews*

Cowles Foundation for Research in Economics

Yale University

Panle Jia

Department of Economics

MIT

June 2008

Revised: September 2008

*Andrews gratefully acknowledges the research support of the National Science Foundation via grant numbers SES-0417911 and SES-0751517. The authors thank Steve Berry for numerous discussions and comments.

Abstract

This Supplement provides three Appendices to the paper Andrews and Jia (2008). Appendix A provides proofs of the asymptotic results of the paper. Appendix B provides supplemental numerical results to those reported in Section 6 of the paper. Appendix C contains details concerning the numerical results reported in Section 6 of the paper.

1 Appendix A

This is a theoretical Appendix that includes proofs of the results given in the paper Andrews and Jia (2008). The first subsection gives a more precise/detailed definition of Δ than appears in Section 4.4 of the paper. The second subsection gives an alternative parametrization of the moment inequality/equality model to that given in Section 2 of the paper. This parametrization is conducive to the calculation of the asymptotic properties of CS's and tests. It was first used in Andrews and Guggenberger (2009). This section also specifies the parameter space for the case of dependent observations and for the case where a preliminary estimator of a parameter τ appears. The third section provides proofs of the results stated in the paper.

We use the following notation throughout. Let $R_+ = \{x \in R : x \geq 0\}$, $R_{++} = \{x \in R : x > 0\}$, $R_{+,\infty} = R_+ \cup \{+\infty\}$, $R_{[+\infty]} = R \cup \{+\infty\}$, $R_{[\pm\infty]} = R \cup \{\pm\infty\}$, $K^p = K \times ... \times K$ (with p copies) for any set K, $\infty^p = (+\infty, ..., +\infty)'$ (with p copies). All limits are as $n \to \infty$ unless specified otherwise. Let "pd" abbreviate "positive definite," $cl(\Psi)$ denote the closure of a set Ψ, and 0_v denote a v-vector of zeros.

1.1 Definition of Δ

The set Δ, which appears in Section 4.4 of the paper, is defined as follows. Let the normalized mean vector and asymptotic correlation matrix of the sample moment functions be denoted by

$$\begin{aligned}\gamma_1(\theta, F) &= Diag^{-1/2}\left(AsyVar_F\left(n^{1/2}\overline{m}_n(\theta)\right)\right) E_F m(W_i, \theta) \geq 0_p \text{ and} \\ \Omega(\theta, F) &= AsyCorr_F\left(n^{1/2}\overline{m}_n(\theta)\right), \qquad (1.1)\end{aligned}$$

where $AsyVar_F(n^{1/2}\overline{m}_n(\theta))$ and $AsyCorr_F(n^{1/2}\overline{m}_n(\theta))$ denote the variance and correlation matrices, respectively, of the asymptotic distribution of $n^{1/2}\overline{m}_n(\theta)$ when the true parameter is θ and the true distribution is F.[1] Then, Δ is defined by

$$\begin{aligned}\Delta = \{&(h_1, \Omega) \in R^p_{+,\infty} \times cl(\Psi) : \exists \text{ a subsequence } \{w_n\} \text{ of } \{n\} \text{ and} \\ &\text{a sequence } \{(\theta_{w_n}, F_{w_n}) \in \mathcal{F} : n \geq 1\} \text{ with } \gamma_1(\theta_{w_n}, F_{w_n}) \geq 0_p \text{ and} \\ &\Omega(\theta_{w_n}, F_{w_n}) \in \Psi \text{ for which } w_n^{1/2}\gamma_1(\theta_{w_n}, F_{w_n}) \to h_1,\ \Omega(\theta_{w_n}, F_{w_n}) \to \Omega, \\ &\text{and } \theta_{w_n} \to \theta_* \text{ for some } \theta_* \text{ in } cl(\Theta)\}. \end{aligned} \tag{1.2}$$

1.2 Alternative Parametrization

In this section we specify a one-to-one mapping between the parameters (θ, F) with parameter space $\mathcal{F}$ and a new parameter $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ with corresponding parameter space Γ. The latter parametrization is amenable to establishing the asymptotic uniformity results of Theorem 1 of the paper.

As stated in the paper, the true value θ_0 $(\in \Theta \subset R^d)$ is assumed to satisfy the moment conditions:

$$\begin{aligned} &E_{F_0} m_j(W_i, \theta_0) \geq 0 \text{ for } j = 1, ..., p \text{ and} \\ &E_{F_0} m_j(W_i, \theta_0) = 0 \text{ for } j = p+1, ..., p+v, \end{aligned} \tag{1.3}$$

where $\{m_j(\cdot, \theta) : j = 1, ..., k\}$ are known real-valued moment functions, $k = p + v$, and $\{W_i : i \geq 1\}$ are i.i.d. or stationary random vectors with joint distribution F_0.

For the case where the sample moment functions depend on a preliminary estimator $\widehat{\tau}_n(\theta)$ of an identified parameter vector τ with true parameter τ_0, we define $m_j(W_i, \theta) = m_j(W_i, \theta, \tau_0)$, $m(W_i, \theta) = (m_1(W_i, \theta, \tau_0), ..., m_k(W_i, \theta, \tau_0))'$, $\overline{m}_{n,j}(\theta) = n^{-1}\sum_{i=1}^n m_j(W_i, \theta, \widehat{\tau}_n(\theta))$, and $\overline{m}_n(\theta) = (\overline{m}_{n,1}(\theta), ..., \overline{m}_{n,k}(\theta))'$. (Hence, in this case, $\overline{m}_n(\theta) \neq n^{-1}\sum_{i=1}^n m(W_i, \theta)$.)

We define $\gamma_1 = (\gamma_{1,1}, ..., \gamma_{1,p})' \in R^p_+$ by writing the moment inequalities in (1.3) as

[1] For dependent observations and when a preliminary estimator of a paramter τ appears, the parameter space $\mathcal{F}$ of (θ, F) is defined in Section 1.2 such that both $AsyVar_F(n^{1/2}\overline{m}_n(\theta))$ and $AsyCorr_F(n^{1/2}\overline{m}_n(\theta))$ exist. These limits equal $Var_F(m(W_i, \theta)))$ and $Corr_F(m(W_i, \theta)))$, respectively, in the case of i.i.d. observations with no preliminary estimator of a parameter τ.

moment equalities:

$$\sigma_{F,j}^{-1}(\theta)E_F m_j(W_i,\theta) - \gamma_{1,j} = 0 \text{ for } j = 1, ..., p, \tag{1.4}$$

where $\sigma_{F,j}^2(\theta)$ is the variance of the asymptotic distribution of $n^{1/2}\overline{m}_{n,j}(\theta)$ under (θ, F). Also, let $\Omega = \Omega(\theta, F) = AsyCorr_F(n^{1/2}\overline{m}_n(\theta))$ denote the correlation matrix of the asymptotic distribution of $n^{1/2}\overline{m}_n(\theta)$ under (θ, F). When no preliminary estimator of a parameter τ appears, $\sigma_{F,j}^2(\theta) = \lim_{n\to\infty} Var_F(n^{1/2}\overline{m}_{n,j}(\theta))$ and $\Omega(\theta, F) = \lim_{n\to\infty} Corr_F (n^{1/2}\overline{m}_n(\theta))$, where $Var_F(n^{1/2}\overline{m}_{n,j}(\theta))$ and $Corr_F(n^{1/2}\overline{m}_n(\theta))$ denote the finite-sample variance of $n^{1/2}\overline{m}_{n,j}(\theta)$ and correlation matrix of $n^{1/2}\overline{m}_n(\theta)$ under (θ, F), respectively. Let $\gamma_2 = (\gamma_{2,1}, \gamma_{2,2}) = (\theta, vech_*(\Omega(\theta, F))) \in R^q$, where $vech_*(\Omega)$ denotes the vector of elements of Ω that lie below the main diagonal, $q = d + k(k-1)/2$, and $\gamma_3 = F$.

For i.i.d. observations and no preliminary estimator of a parameter τ, the parameter space for γ is defined by $\Gamma = \{\gamma = (\gamma_1, \gamma_2, \gamma_3) :$ for some $(\theta, F) \in \mathcal{F}$, where $\mathcal{F}$ is defined in (2.2) of the paper, γ_1 satisfies (1.4), $\gamma_2 = (\theta, vech_*(\Omega(\theta, F)))$, and $\gamma_3 = F\}$.

For dependent observations and for sample moment functions that depend on a preliminary estimator $\widehat{\tau}_n(\theta)$, we specify the parameter space Γ for the moment inequality model using a set of high-level conditions. To verify the high-level conditions using primitive conditions one has to specify an estimator $\widehat{\Sigma}_n(\theta)$ of the asymptotic variance matrix $\Sigma(\theta)$ of $n^{1/2}\overline{m}_n(\theta)$. For brevity, we do not do so here. Since there is a one-to-one mapping from γ to (θ, F), Γ also defines the parameter space $\mathcal{F}$ of (θ, F). Let Ψ be a specified set of $k \times k$ correlation matrices. The parameter space Γ is defined to include parameters $\gamma = (\gamma_1, \gamma_2, \gamma_3) = (\gamma_1, (\theta, \gamma_{2,2}), F)$ that satisfy:

(i) $\theta \in \Theta$,
(ii) $\sigma_{F,j}^{-1}(\theta)E_F m_j(W_i,\theta) - \gamma_{1,j} = 0$ for $j = 1, ..., p$,
(iii) $E_F m_j(W_i,\theta) = 0$ for $j = p+1, ..., k$,
(iv) $\sigma_{F,j}^2(\theta) = AsyVar_F\left(n^{1/2}\overline{m}_{n,j}(\theta)\right)$ exists and lies in $(0, \infty)$ for $j = 1, ..., k$,
(v) $AsyCorr_F\left(n^{1/2}\overline{m}_n(\theta)\right)$ exists and equals $\Omega_{\gamma_{2,2}} \in \Psi$, and
(vi) $\{W_i : i \geq 1\}$ are stationary under F, (1.5)

where $\gamma_1 = (\gamma_{1,1}, ..., \gamma_{1,p})'$ and $\Omega_{\gamma_{2,2}}$ is the $k \times k$ correlation matrix determined by $\gamma_{2,2}$.[2] Furthermore, Γ must be restricted by enough additional conditions such that

[2]In Andrews and Guggenberger (2009), a strong mixing condition is imposed in condition (vi) of

under any sequence $\{\gamma_{n,h} = (\gamma_{n,h,1}, (\theta_{n,h}, vech_* \ (\Omega_{n,h})), \ F_{n,h}) : n \geq 1\}$ of parameters in Γ that satisfies $n^{1/2}\gamma_{n,h,1} \to h_1$ and $(\theta_{n,h}, vech_*(\Omega_{n,h})) \to h_2 = (h_{2,1}, h_{2,2})$ for some $h = (h_1, h_2) \in R^p_{+,\infty} \times R^q_{[\pm\infty]}$, we have

$$\begin{aligned}
&\text{(vii) } A_n = (A_{n,1}, ..., A_{n,k})' \to_d Z_{h_{2,2}} \sim N(0_k, \Omega_{h_{2,2}}) \text{ as } n \to \infty, \text{ where} \\
&A_{n,j} = n^{1/2} \left(\overline{m}_{n,j}(\theta_{n,h}) - E_{F_{n,h}} m_j(W_i, \theta_{n,h})\right) / \sigma_{F_{n,h},j}(\theta_{n,h}), \\
&\text{(viii) } \widehat{\sigma}_{n,j}(\theta_{n,h}) / \sigma_{F_{n,h},j}(\theta_{n,h}) \to_p 1 \text{ as } n \to \infty \text{ for } j = 1, ..., k, \\
&\text{(ix) } \widehat{D}_n^{-1/2}(\theta_{n,h}) \widehat{\Sigma}_n(\theta_{n,h}) \widehat{D}_n^{-1/2}(\theta_{n,h}) \to_p \Omega_{h_{2,2}} \text{ as } n \to \infty, \text{ and} \\
&\text{(x) conditions (vii)-(ix) hold for all subsequences } \{w_n\} \text{ in place of } \{n\},
\end{aligned} \tag{1.6}$$

where $\Omega_{h_{2,2}}$ is the $k \times k$ correlation matrix for which $vech_*(\Omega_{h_{2,2}}) = h_{2,2}$, $\widehat{\sigma}^2_{n,j}(\theta) = [\widehat{\Sigma}_n(\theta)]_{jj}$ for $1 \leq j \leq k$ and $\widehat{D}_n(\theta) = Diag\{\widehat{\sigma}^2_{n,1}(\theta), ..., \widehat{\sigma}^2_{n,k}(\theta)\}$ $(= Diag(\widehat{\Sigma}_n(\theta)))$.[3,4]

For example, for i.i.d. observations, conditions (i)-(vi) in (2.2) of the paper imply conditions (i)-(vi) in (1.5). Furthermore, conditions (i)-(vi) in (2.2) of the paper plus the definition of $\widehat{\Sigma}_n(\theta)$ in (2.5) of the paper and the additional condition (vii) in (2.2) of the paper imply conditions (vii)-(x) in (1.6). For a proof, see Lemma 2 of Andrews and Guggenberger (2009).

For dependent observations or when a preliminary estimator of a parameter τ appears, one needs to specify a particular variance estimator $\widehat{\Sigma}_n(\theta)$ before one can specify primitive "additional conditions" beyond conditions (i)-(vi) in (1.5) that ensure that Γ is such that any sequences $\{\gamma_{w_n,h} : n \geq 1\}$ in Γ satisfy (1.6). For brevity, we do not do so here.

We now specify the set Δ, defined in (1.2), in the parametrization introduced above. Define

$$H = \{h \in R^p_{[\pm\infty]} \times R^q_{[\pm\infty]} : \exists \text{ a subsequence } \{w_n\} \text{ of } \{n\} \text{ and a sequence}$$

(1.5). This condition is used to verify Assumption E0 in that paper and is not needed with RMS critical values.

[3]When a preliminary estimator $\widehat{\tau}_n(\theta)$ appears, $A_{n,j}$ can be written equivalently as $n^{1/2}\left(n^{-1}\sum_{i=1}^n m_j(W_i, \theta_{n,h}, \widehat{\tau}_n(\theta_{n,h})) - E_{F_{n,h}} m_j(W_i, \theta_{n,h}, \tau_0)\right)/\sigma_{F_{n,h},j}(\theta_{n,h})$, which typically is asymptotically normal with an asymptotic variance matrix $\Omega_{h_{2,2}}$ that reflects the fact that τ_0 has been estimated. When a preliminary estimator $\widehat{\tau}_n(\theta)$ appears, $\widehat{\Sigma}_n(\theta)$ needs to be defined to take account of the fact that τ_0 has been estimated. When no preliminary estimator $\widehat{\tau}_n(\theta)$ appears, $A_{n,j}$ can be written equivalently as $n^{1/2}(\overline{m}_{n,j}(\theta_{n,h}) - E_{F_{n,h}}\overline{m}_{n,j}(\theta_{n,h}))/\sigma_{F_{n,h},j}(\theta_{n,h})$.

[4]Condition (x) of (1.6) requires that conditions (vii)-(ix) must hold under any sequence of parameters $\{\gamma_{w_n,h} : n \geq 1\}$ that satisfies the conditions preceding (1.6) with n replaced by w_n.

$$\{\gamma_{w_n,h} \in \Gamma : n \geq 1\} \text{ for which } w_n^{1/2}\gamma_{w_n,h,1} \to h_1 \text{ and } \gamma_{w_n,h,2} \to h_2\}. \quad (1.7)$$

Then, Δ can be written equivalently as

$$\begin{aligned}\Delta = \{&(h_1, \Omega_{h_{2,2}}) \in R^p_{+,\infty} \times cl(\Psi) : h = (h_1, h_{2,1}, h_{2,2}) \in H \\ &\text{for some } h_{2,1} \in cl(\Theta), \text{ where } h_{2,2} = vech_*(\Omega_{h_{2,2}})\}. \end{aligned} \quad (1.8)$$

In words, Δ is the set of "slackness" parameters h_1 and correlation matrices Ω that correspond to some limit point h in H.

1.3 Proofs

The proof of Theorem 1 of the paper uses the following Lemmas. Let

$$CP_n(\gamma) = P_\gamma(T_n(\theta) \leq c_n(\theta)). \quad (1.9)$$

As above, for a sequence of constants $\{\zeta_n : n \geq 1\}$, $\zeta_n \to [\zeta_{1,\infty}, \zeta_{2,\infty}]$ denotes that $\zeta_{1,\infty} \leq \liminf_{n\to\infty} \zeta_n \leq \limsup_{n\to\infty} \zeta_n \leq \zeta_{2,\infty}$.

Lemma 1 *Suppose Assumptions* S, φ, κ, *and* η1 *hold. Let* $\{\gamma_{n,h} = (\gamma_{n,h,1}, \gamma_{n,h,2}, \gamma_{n,h,3}) : n \geq 1\}$ *be a sequence of points in* Γ *that satisfies* (i) $n^{1/2}\gamma_{n,h,1} \to h_1$ *for some* $h_1 \in R^p_{+,\infty}$ *and* (ii) $\gamma_{n,h,2} \to h_2$ *for some* $h_2 = (h_{2,1}, h_{2,2}) \in R^q_{[\pm\infty]}$. *Let* $h = (h_1, h_2)$ *and let* $\Omega_{h_{2,2}}$ *be the correlation matrix that corresponds to* $h_{2,2}$. *Then,*

(a) $CP_n(\gamma_{n,h}) \to [CP(h_1, \Omega_{h_{2,2}}, \eta(\Omega_{h_{2,2}})-), CP(h_1, \Omega_{h_{2,2}}, \eta(\Omega_{h_{2,2}}))]$ *and*

(b) *for any subsequence* $\{w_n : n \geq 1\}$ *of* $\{n\}$, *the result of part* (a) *holds with* w_n *in place of* n *provided conditions* (i) *and* (ii) *above hold with* w_n *in place of* n.

Lemma 2 *Suppose Assumptions* S(b)-(e) *hold. Then,* $q_S(\beta, \Omega)$ *is continuous on* $(R^p_{[+\infty]} \times R^v) \times \Psi$.

Proof of Theorem 1 of the Paper. First, we prove part (a). Let $\{\gamma_n^* = (\gamma_{n,1}^*, \gamma_{n,2}^*, \gamma_{n,3}^*) \in \Gamma : n \geq 1\}$ be a sequence such that $\liminf_{n\to\infty} CP_n(\gamma_n^*) = \liminf_{n\to\infty} \inf_{\gamma\in\Gamma} CP_n(\gamma)$ $(= AsyCS)$. Such a sequence always exists. Let $\{u_n : n \geq 1\}$ be a subsequence of $\{n\}$ such that $\lim_{n\to\infty} CP_{u_n}(\gamma_{u_n}^*)$ exists and equals $\liminf_{n\to\infty} CP_n(\gamma_n^*) = AsyCS$. Such a subsequence always exists.

Let $\gamma^*_{n,1,j}$ denote the jth component of $\gamma^*_{n,1}$ for $j = 1, ..., p$. Either (1) $\limsup_{n\to\infty} u_n^{1/2}\gamma^*_{u_n,1,j} < \infty$ or (2) $\limsup_{n\to\infty} u_n^{1/2}\gamma^*_{u_n,1,j} = \infty$. If (1) holds, then for some subsequence $\{w_n\}$ of $\{u_n\}$,

$$w_n^{1/2}\gamma^*_{w_n,1,j} \to h^*_{1,j} \text{ for some } h^*_{1,j} \in R_+. \tag{1.10}$$

If (2) holds, then for some subsequence $\{w_n\}$ of $\{u_n\}$,

$$w_n^{1/2}\gamma^*_{w_n,1,j} \to h^*_{1,j}, \text{ where } h^*_{1,j} = \infty. \tag{1.11}$$

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

$$\gamma^*_{w_n,2} \to h^*_2 \text{ for some } h^*_2 \in \mathrm{cl}(\Gamma_2). \tag{1.12}$$

By taking successive subsequences over the p components of $\gamma^*_{u_n,1}$ and $\gamma^*_{u_n,2}$, we find that there exists a subsequence $\{w_n\}$ of $\{u_n\}$ such that for each $j = 1, ..., p$ either (1.10) or (1.11) applies and (1.12) holds. In consequence, (i) $w_n^{1/2}\gamma_{w_n,h,1} \to h^*_1$ for some $h^*_1 \in R^p_{+,\infty}$, (ii) $\gamma_{w_n,h,2} \to h^*_2$ for some $h^*_2 \in R^q_{[\pm\infty]}$, (iii) $h^* = (h^*_1, h^*_2) \in H$ (for H defined in (1.7)), and (iv) $\lim_{n\to\infty} CP_{w_n}(\gamma^*_{w_n}) = AsyCS$. Hence, by Lemma 1(b),

$$\begin{aligned} AsyCS = \lim_{n\to\infty} CP_{w_n}(\gamma^*_{w_n}) &\geq CP(h^*_1, \Omega_{h^*_{2,2}}, \eta(\Omega_{h^*_{2,2}})-) \\ &\geq \inf_{(h_1,\Omega)\in\Delta} CP(h_1, \Omega, \eta(\Omega)-), \end{aligned} \tag{1.13}$$

where the second inequality holds because $(h^*_1, \Omega_{h^*_{2,2}}) \in \Delta$ by the definition of Δ in (1.8).

Next, by the definition of Δ in (1.8), for each $(h_1, \Omega_{h_{2,2}}) \in \Delta$, there exists a subsequence $\{t_n : n \geq 1\}$ of $\{n\}$ and a sequence of points $\{\gamma_{t_n,h} = (\gamma_{t_n,h,1}, \gamma_{t_n,h,2}, \gamma_{t_n,h,3}) \in \Gamma : n \geq 1\}$ such that conditions (i) and (ii) of Lemma 1 hold with t_n in place of n. Hence,

$$\begin{aligned} AsyCS &= \liminf_{n\to\infty} \inf_{(\theta,F)\in\mathcal{F}} P_F(T_n(\theta) \leq c_n(\theta)) \\ &\leq \liminf_{n\to\infty} CP_{t_n}(\gamma_{t_n,h}) \\ &\leq CP(h_1, \Omega_{h_{2,2}}, \eta(\Omega_{h_{2,2}})), \end{aligned} \tag{1.14}$$

where the second inequality holds by Lemma 1(b). Since (1.14) holds for all $(h_1, \Omega_{h_{2,2}}) \in \Delta$, we have

$$AsyCS \leq \inf_{(h_1,\Omega)\in\Delta} CP(h_1, \Omega, \eta(\Omega)). \tag{1.15}$$

Combining (1.13) and (1.15) establishes part (a) of the Theorem.

Part (b) of the Theorem follows from part (a) and Assumption $\eta 2$. Part (c) of the Theorem follows from part (a) and Assumption $\eta 3$. $\square$

Proof of Lemma 1. For notational simplicity, let Ω_0 denote $\Omega_{h_{2,2}}$. To establish part (a), we show below that

$$\begin{pmatrix} T_n(\theta_{n,h}) \\ c_n(\theta_{n,h}) \end{pmatrix} \to_d \begin{pmatrix} S\left(Z + (h_1, 0_v), \Omega_0\right) \\ q_S\left(\varphi\left(\kappa^{-1}(\Omega_0)[Z + (h_1, 0_v)], \Omega_0\right), \Omega_0\right) + \eta(\Omega_0) \end{pmatrix} \text{ as } n \to \infty \tag{1.16}$$

under $\{\gamma_{n,h} : n \geq 1\}$, where $Z \sim N(0_k, \Omega_0)$. Hence, by the definition of convergence in distribution, for every continuity point x of the asymptotic distribution of $T_n(\theta_{n,h}) - c_n(\theta_{n,h})$, we have

$$\begin{aligned} & P_{\gamma_{n,h}}(T_n(\theta_{n,h}) \leq c_n(\theta_{n,h}) + x) \\ \to\ & P\left(S\left(Z + (h_1, 0_v), \Omega_0\right) \leq q_S\left(\varphi\left(\kappa^{-1}(\Omega_0)[Z + (h_1, 0_v)], \Omega_0\right), \Omega_0\right) + \eta(\Omega_0) + x\right) \\ =\ & CP(h_1, \Omega_0, \eta(\Omega_0) + x). \end{aligned} \tag{1.17}$$

There exist continuity points $x > 0$ and $x < 0$ arbitrarily close to zero. Hence, we have

$$\begin{aligned} & \limsup_{n\to\infty} P_{\gamma_{n,h}}(T_n(\theta_{n,h}) \leq c_n(\theta_{n,h})) \\ \leq\ & \lim_{x\downarrow 0} \limsup_{n\to\infty} P_{\gamma_{n,h}}(T_n(\theta_{n,h}) \leq c_n(\theta_{n,h}) + x) \\ =\ & \lim_{x\downarrow 0} CP(h_1, \Omega_0, \eta(\Omega_0) + x) \\ =\ & CP(h_1, \Omega_0, \eta(\Omega_0)), \end{aligned} \tag{1.18}$$

where the first equality holds by (1.17) and the second equality holds because $CP(h_1, \Omega_0, \eta(\Omega_0) + x)$ is a df and hence is right-continuous. Analogously,

$$\begin{aligned} \liminf_{n\to\infty} P_{\gamma_{n,h}}(T_n(\theta_{n,h}) \leq c_n(\theta_{n,h})) &\geq \lim_{x\downarrow 0} CP(h_1, \Omega_0, \eta(\Omega_0) - x) \\ &= CP(h_1, \Omega_0, \eta(\Omega_0)-), \end{aligned} \tag{1.19}$$

where the equality holds by definition. Equations (1.18) and (1.19) combine to establish part (a).

Next, we prove (1.16). Using Assumption S(a), we have

$$T_n(\theta) = S\Big(\widehat{D}_n^{-1/2}(\theta)n^{1/2}\overline{m}_n(\theta),\ \widehat{D}_n^{-1/2}(\theta)\widehat{\Sigma}_n(\theta)\widehat{D}_n^{-1/2}(\theta)\Big). \tag{1.20}$$

For i.i.d. or dependent observations with or without preliminary estimators of identified parameters, (1.6) holds (using the fact that $\gamma \in \Gamma$ if and only if $(\theta, F) \in \mathcal{F}$ and using Lemma 2 of Andrews and Guggenberger (2009) to show that (1.6) holds for i.i.d. observations). By (1.6), the jth element of $\widehat{D}_n^{-1/2}(\theta_{n,h})n^{1/2}\overline{m}_n(\theta_{n,h})$ equals $(1 + o_p(1))(A_{n,j} + n^{1/2}\gamma_{n,h,1,j})$, where $\gamma_{n,h,1} = (\gamma_{n,h,1,1},\ ..., \gamma_{n,h,1,p})'$ and by definition $\gamma_{n,h,1,j} = 0$ for $j = p+1, ..., k$. If $h_{1,j} = \infty$ and $j \leq p$, where $h_1 = (h_{1,1}, ..., h_{1,p})'$, then $A_{n,j} + n^{1/2}\gamma_{n,h,1,j} \rightarrow_p \infty$ under $\{\gamma_{n,h} : n \geq 1\}$ by condition (vii) of (1.6) and the definition of $\{\gamma_{n,h} : n \geq 1\}$. Hence, if any element of h_1 equals ∞, $\widehat{D}_n^{-1/2}(\theta_{n,h})n^{1/2}\overline{m}_n(\theta_{n,h})$ does not converge in distribution (to a proper finite random vector) and the continuous mapping theorem cannot be applied to obtain the asymptotic distribution of the right-hand side of (1.20) or of the RMS critical value, which is defined by

$$c_n(\theta) = q_S\left(\varphi\Big(\xi_n(\theta), \widehat{\Omega}_n(\theta)\Big), \widehat{\Omega}_n(\theta)\right) + \eta(\widehat{\Omega}_n(\theta)). \tag{1.21}$$

To circumvent these problems, we consider k-vector-valued functions of $\widehat{D}_n^{-1/2}(\theta_{n,h}) \times n^{1/2}\overline{m}_n(\theta_{n,h})$ and $\xi_n(\theta_{n,h})$ that converge in distribution whether or not some elements of h_1 equal ∞. Then, we write the right-hand sides of (1.20) and (1.21) as continuous functions of these k-vectors and apply the continuous mapping theorem. Let $G(\cdot)$ be a strictly increasing continuous df on R, such as the standard normal df.

For $j \leq k$, we have

$$\begin{aligned} G_{\kappa,n,j} &= G(\xi_{n,j}(\theta_{n,h})) = G\left(\kappa^{-1}(\widehat{\Omega}_n(\theta_{n,h}))\widehat{\sigma}_{n,j}^{-1}(\theta_{n,h})n^{1/2}\overline{m}_{n,j}(\theta_{n,h})\right) \\ &= G\left(\kappa^{-1}(\widehat{\Omega}_n(\theta_{n,h}))\widehat{\sigma}_{n,j}^{-1}(\theta_{n,h})\sigma_{F_{n,h},j}(\theta_{n,h})\left[A_{n,j} + n^{1/2}\gamma_{n,h,1,j}\right]\right), \end{aligned} \tag{1.22}$$

where $A_{n,j}$ is defined in (1.6) and by definition $\gamma_{n,h,1,j} = 0$ for $j = p+1, ..., k$.

Let $Z = (Z_1, ..., Z_k)' \sim N(0_k, \Omega_0)$. Define $h_{1,j} = 0$ for $j = p+1, ..., k$. If $j \leq p$ and $h_{1,j} < \infty$ or if $j = p+1, ..., k$, then

$$G_{\kappa,n,j} \rightarrow_d G\left(\kappa^{-1}(\Omega_0)[Z_j + h_{1,j}]\right) \tag{1.23}$$

using (1.22), conditions (vii) and (viii) of (1.6) (which yield $A_{n,j} + n^{1/2}\gamma_{n,h,1,j} \to_d Z_j + h_{1,j}$), Assumption κ and condition (ix) of (1.6) (which yield $\kappa^{-1}(\widehat{\Omega}_n(\theta_{n,h})) \to_p \kappa^{-1}(\Omega_0)$), and the continuous mapping theorem.

If $j \leq p$ and $h_{1,j} = \infty$, then

$$G_{\kappa,n,j} \to_p 1 \tag{1.24}$$

using (1.22), $A_{n,j} = O_p(1)$, $\kappa^{-1}(\widehat{\Omega}_n(\theta_{n,h})) \to_p \kappa^{-1}(\Omega_0) > 0$, and $G(x) \to 1$ as $x \to \infty$. The results in (1.23)-(1.24) hold jointly and combine to give

$$\begin{aligned} G_{\kappa,n} &= (G_{\kappa,n,1}, ..., G_{\kappa,n,k})' \to_d G_{\kappa,\infty}, \text{ where} \\ G_{\kappa,\infty} &= (G(\kappa^{-1}(\Omega_0)[Z_1 + h_{1,1}]), ..., G(\kappa^{-1}(\Omega_0)[Z_k + h_{1,k}]))' \end{aligned} \tag{1.25}$$

and $G(Z_{h_{2,2,j}} + h_{1,j})$ denotes $G(\infty) = 1$ when $h_{1,j} = \infty$.

Let G^{-1} denote the inverse of G. For $x = (x_1, ..., x_k)' \in R^p_{[+\infty]} \times R^v$, let $G_{(k)}(x) = (G(x_1), ..., G(x_k))' \in (0,1]^p \times (0,1)^v$. For $z = (z_1, ..., z_k)' \in (0,1]^p \times (0,1)^v$, let $G^{-1}_{(k)}(z) = (G^{-1}(z_1), ..., G^{-1}(z_k))' \in R^p_{[+\infty]} \times R^v$. Define $\widetilde{q}_S(z, \Omega)$ as

$$\widetilde{q}_{S,\varphi}(z, \Omega) = q_S\left(\varphi(G^{-1}_{(k)}(z), \Omega), \Omega\right) \tag{1.26}$$

for $z \in (0,1]^p \times (0,1)^v$ and $\Omega \in \Psi$.

Assumption φ and Lemma 2 imply that $\widetilde{q}_{S,\varphi}(z, \Omega)$ is continuous at (z, Ω) for all $z \in \mathcal{Z}((h_1, 0_v), \Omega_0)$ and $\Omega = \Omega_0$, where

$$\begin{aligned} \mathcal{Z}((h_1, 0_v), \Omega_0) &= \left\{z \in (0,1]^p \times (0,1)^v : G^{-1}_{(k)}(z) \in \Xi((h_1, 0_v), \Omega)\right\} \text{ and} \\ P\left(G_{\kappa,\infty} \in \mathcal{Z}((h_1, 0_v), \Omega_0)\right) &= P\left(\kappa^{-1}(\Omega_0)[Z + (h_1, 0_v)] \in \Xi((h_1, 0_v), \Omega_0)\right) \\ &= 1, \end{aligned} \tag{1.27}$$

where $\Xi(\beta, \Omega)$ is defined in Assumption φ.

We now have

$$\begin{aligned} c_n(\theta_{n,h}) &= q_S\left(\varphi(\xi_n(\theta_{n,h}), \widehat{\Omega}_n(\theta_{n,h})), \widehat{\Omega}_n(\theta_{n,h})\right) + \eta(\widehat{\Omega}_n(\theta_{n,h})) \\ &= q_S\left(\varphi(G^{-1}_{(k)}(G_{\kappa,n}), \widehat{\Omega}_n(\theta_{n,h})), \widehat{\Omega}_n(\theta_{n,h})\right) + \eta(\widehat{\Omega}_n(\theta_{n,h})) \\ &= \widetilde{q}_{S,\varphi}\left(G_{\kappa,n}, \widehat{\Omega}_n(\theta_{n,h})\right) + \eta(\widehat{\Omega}_n(\theta_{n,h})) \\ &\to_d \widetilde{q}_{S,\varphi}\left(G_{\kappa,\infty}, \Omega_0\right) + \eta(\Omega_0) \end{aligned}$$

$$= q_S\left(\varphi(G_{(k)}^{-1}(G_{\kappa,\infty}),\Omega_0),\Omega_0\right)+\eta(\Omega_0)$$
$$= q_S\left(\varphi(\kappa^{-1}(\Omega_0)[Z+(h_1,0_v)],\Omega_0),\Omega_0\right)+\eta(\Omega_0), \quad (1.28)$$

where the first equality holds by the definition of $c_n(\theta_{n,h})$, the second equality holds by the definitions of $G_{\kappa,n}$ and $G_{(k)}^{-1}(\cdot)$, the third and fourth equalities hold by the definition of $\widetilde{q}_{S,\varphi}(\cdot,\cdot)$, the convergence holds by (1.25), condition (ix) of (1.6), Assumption $\eta 1$, and the continuous mapping theorem using (1.27), the last equality holds by the definitions of $G_{\kappa,\infty}$ and $G_{(k)}^{-1}(\cdot)$ and the definition that if $h_{1,j}=\infty$, then the corresponding element of $Z+(h_1,0_v)$ equals ∞.

We now use an analogous argument to that in (1.22)-(1.28) to show that

$$T_n(\theta_{n,h})\rightarrow_d S(Z+(h_1,0_v),\Omega_0). \quad (1.29)$$

The argument only differs from that given above in that (i) $\kappa(\cdot)$ is replaced by 1 throughout, (ii) the function $q_S(\varphi(m,\Omega),\Omega)$ is replaced by $S(m,\Omega)$, (iii) the function $\widetilde{q}_{S,\varphi}(z,\Omega)=q_S(\varphi(G_{(k)}^{-1}(z),\Omega),\Omega)$ is replaced by $\widetilde{S}(z,\Omega)=S(G_{(k)}^{-1}(z),\Omega)$, and (iv) the continuity argument in the paragraph containing (1.27) is replaced by the assertion that $\widetilde{S}(z,\Omega)$ is continuous at all $(z,\Omega)\in((0,1]^p\times(0,1)^v)\times\Psi$ by Assumption S(c).

The convergence in (1.28) and (1.29) is joint because the two results can be obtained by a single application of the continuous mapping theorem. Hence, the verification of (1.16) is complete and part (a) is proved.

Next, we prove part (b). By the same argument as above but using condition (x) of (1.6) in place of conditions (vii)-(ix), the results of (1.28) and 1.29 hold with $\{w_n\}$ in place of $\{n\}$ for any subsequence $\{w_n\}$. Hence, (1.16) and (1.17) hold with the same changes, which implies that part (b) holds. $\square$

Proof of Lemma 2. Given $(\beta_0,\Omega_0)\in(R^p_{[+\infty]}\times R^v)\times\Psi$, we consider three cases: (i) $q_S(\beta_0,\Omega_0)>0$, (ii) $q_S(\beta_0,\Omega_0)=0$ and either $v>0$ or both $v=0$ and $\beta_0\neq\infty^p$, and (iii) $q_S(\beta_0,\Omega_0)=0$, $v=0$, and $\beta_0=\infty^p$.

In case (i), given $\varepsilon>0$, we want to show that if (β,Ω) is sufficiently close to (β_0,Ω_0), then $|q_S(\beta,\Omega)-q_S(\beta_0,\Omega_0)|<\varepsilon$. Let $Z^*\sim N(0_k,I_k)$. By Assumption S(e), the df of $S(\Omega_0^{1/2}Z^*+\beta_0,\Omega_0)$ is strictly increasing at $x=q_S(\beta_0,\Omega_0)>0$. Hence, for some $\varepsilon_U>0$,

$$P\left(S(\Omega_0^{1/2}Z^*+\beta_0,\Omega_0)\leq q_S(\beta_0,\Omega_0)+\varepsilon\right)=1-\alpha+\varepsilon_U. \quad (1.30)$$

The df of $S(\Omega^{1/2}Z^*+\beta,\Omega)$ at $x>0$ is continuous in (β,Ω) at (β_0,Ω_0) by the bounded convergence theorem because

$$\begin{aligned}
&\text{(a) } S(\Omega^{1/2}Z^*+\beta,\Omega)\to S(\Omega_0^{1/2}Z^*+\beta_0,\Omega_0) \text{ a.s.,}\\
&\text{(b) } 1\left(S(\Omega^{1/2}Z^*+\beta,\Omega)\le x\right)\to 1\left(S(\Omega_0^{1/2}Z^*+\beta_0,\Omega_0)\le x\right) \text{ a.s.}\\
&\quad\text{except if } S(\Omega_0^{1/2}Z^*+\beta_0,\Omega_0)=x,\\
&\text{(c) } P\left(S(\Omega_0^{1/2}Z^*+\beta_0,\Omega_0)=x\right)=0, \text{ and}\\
&\text{(d) the indicator function is bounded,}
\end{aligned} \tag{1.31}$$

where (a) holds by Assumption S(c), (b) holds by (a), and (c) holds because the df of $S(\Omega_0^{1/2}Z^*+\beta_0,\Omega_0)$ is continuous at all $x>0$ by Assumption S(e).

In consequence, for all (β,Ω) sufficiently close to (β_0,Ω_0), we have

$$\begin{aligned}
&\big|P\left(S(\Omega^{1/2}Z^*+\beta,\Omega)\le q_S(\beta_0,\Omega_0)+\varepsilon\right)\\
&-P\left(S(\Omega_0^{1/2}Z^*+\beta_0,\Omega_0)\le q_S(\beta_0,\Omega_0)+\varepsilon\right)\Big|<\varepsilon_U/2.
\end{aligned} \tag{1.32}$$

Equations (1.30) and (1.32) imply that

$$P\left(S(\Omega^{1/2}Z^*+\beta,\Omega)\le q_S(\beta_0,\Omega_0)+\varepsilon\right)\ge 1-\alpha+\varepsilon_U/2. \tag{1.33}$$

The definition of a quantile and (1.33) imply that

$$q_S(\beta,\Omega)\le q_S(\beta_0,\Omega_0)+\varepsilon. \tag{1.34}$$

By a completely analogous argument, for (β,Ω) sufficiently close to (β_0,Ω_0), $q_S(\beta,\Omega)\ge q_S(\beta_0,\Omega_0)-\varepsilon$. Hence, $|q_S(\beta,\Omega)-q_S(\beta_0,\Omega_0)|<\varepsilon$ and the proof is complete for case (i).

In case (ii), $P(S(\Omega_0^{1/2}Z^*+\beta_0,\Omega_0)\le 0)\ge 1-\alpha$ because $q_S(\beta_0,\Omega_0)=0$. Also, in case (ii), $S(\Omega_0^{1/2}Z^*+\beta_0,\Omega_0)$ has a strictly increasing df for $x>0$ by Assumption S(e) (because $v=0$ and $\beta_0=\infty^p$ does not hold in case (ii)). These results imply that given $\varepsilon>0$, there exists $\varepsilon_1>0$ such that

$$P(S(\Omega_0^{1/2}Z^*+\beta_0,\Omega_0)\le\varepsilon)=1-\alpha+\varepsilon_1. \tag{1.35}$$

Because the df of $S(\Omega^{1/2}Z^* + \beta, \Omega)$ at $\varepsilon > 0$ is continuous in (β, Ω) by (1.31), for all (β, Ω) sufficiently close to (β_0, Ω_0), we have

$$\left| P\left(S(\Omega^{1/2}Z^* + \beta, \Omega) \leq \varepsilon\right) - P\left(S(\Omega_0^{1/2}Z^* + \beta_0, \Omega_0) \leq \varepsilon\right) \right| < \varepsilon_1/2. \tag{1.36}$$

Equations (1.35) and (1.36) imply

$$P\left(S(\Omega^{1/2}Z^* + \beta, \Omega) \leq \varepsilon\right) \geq 1 - \alpha. \tag{1.37}$$

This and the definition of a quantile imply that $q_S(\beta, \Omega) \leq \varepsilon$. Since $q_S(\beta, \Omega) \geq 0$ for all (β, Ω) by Assumption S(b), the proof for case (ii) is complete.

In case (iii), $S(\Omega_0^{1/2}Z^* + \beta_0, \Omega_0) = S(\infty^p, \Omega_0) = 0$ a.s. by Assumptions S(b) and S(d). This and the continuity in (β, Ω) at (β_0, Ω_0) of the df of $S(\Omega^{1/2}Z^* + \beta, \Omega)$ at $x > 0$, which holds by (1.31), give: for all $x > 0$,

$$\lim_{(\beta,\Omega)\to(\beta_0,\Omega_0)} P\left(S(\Omega^{1/2}Z^* + \beta, \Omega) \leq x\right) = P\left(S(\Omega_0^{1/2}Z^* + \beta_0, \Omega_0) \leq x\right) = 1. \tag{1.38}$$

Equation (1.38) implies that given any $x > 0$ for all (β, Ω) sufficiently close to (β_0, Ω_0), the df of $S(\Omega^{1/2}Z^* + \beta, \Omega)$ at $x > 0$ is greater than $1 - \alpha$ and hence $q_S(\beta, \Omega) \leq x$. Since $q_S(\beta, \Omega) \geq 0$ for all (β, Ω) and $x > 0$ is arbitrary, the proof for case (iii) is complete. $\square$

Proof of Lemma 1 of the Paper. Assumption LA3(a) holds by the Liapounov triangular array CLT for row-wise i.i.d. random variables with mean zero and variance one using Assumptions LA1(a), LA1(c), and LA3* and the Cramér-Wold device. Assumptions LA3(b) and LA3(c) hold by standard arguments using a weak law of large numbers for row-wise i.i.d. random variables with variance one using Assumptions LA1(a), LA1(c), and LA3*. Note that Assumption LA3 does not follow from (1.6) because in Assumption LA3 the functions are evaluated at θ_0, which is not the true value (unless $\lambda = 0$). $\square$

Proof of Theorem 2 of the Paper. The proof follows a similar line of argument to that of Lemma 1(a). We start by showing that under the given assumptions (1.16) holds with $(h_1, 0_v)$ replaced by $(h_1, 0_v) + \Pi_0\lambda$. By element-by-element mean-value expansions about $\theta = \theta_n$ and Assumptions LA1 and LA2, we obtain

$$D^{-1/2}(\theta_0, F_n)E_{F_n}m(W_i, \theta_0) \; = \; D^{-1/2}(\theta_n, F_n)E_{F_n}m(W_i, \theta_n)$$

$$+\Pi(\theta_n^*, F_n)(\theta_0 - \theta_n),$$
$$n^{1/2} D^{-1/2}(\theta_0, F_n) E_{F_n} m(W_i, \theta_0) \to (h_1, 0_v) + \Pi_0 \lambda, \quad (1.39)$$

where $D(\theta, F) = Diag\{\sigma_{F,1}^2(\theta), ..., \sigma_{F,k}^2(\theta)\}$, θ_n^* may differ across rows of $\Pi(\theta_n^*, F_n)$, θ_n^* lies between θ_0 and θ_n, $\theta_n^* \to \theta_0$, and $\Pi(\theta_n^*, F_n) \to \Pi_0$.

For the same reason as described above following (1.20), to obtain the asymptotic distribution of $T_n(\theta_0)$ we use the same type of argument as in the proof of Lemma 1(a). Let $G(\cdot)$ be a strictly increasing continuous df on R, such as the standard normal df. Using (1.39), Assumption LA3, and $\kappa^{-1}(\widehat{\Omega}_n(\theta_0)) \to_p \kappa^{-1}(\Omega(\theta_0))$ (which holds by Assumptions κ and LA3), for $j = 1, ..., k$, we have

$$\begin{aligned}
G_{\kappa,n,j}^0 &= G\left(\kappa^{-1}(\widehat{\Omega}_n(\theta_0))\widehat{\sigma}_{n,j}^{-1}(\theta_0) n^{1/2} \overline{m}_{n,j}(\theta_0)\right) \\
&= G\left(\kappa^{-1}(\widehat{\Omega}_n(\theta_0))\widehat{\sigma}_{n,j}^{-1}(\theta_0)\sigma_{F_n,j}(\theta_0)\left[A_{n,j}^0 + n^{1/2}\sigma_{F_n,j}^{-1}(\theta_0) E_{F_n} m_j(W_i, \theta_0)\right]\right), \\
G_{\kappa,n,j}^0 &\to_p 1 \text{ if } j \le p \text{ and } h_{1,j} = \infty, \qquad (1.40) \\
G_{\kappa,n,j}^0 &\to_d G\left(\kappa^{-1}(\Omega(\theta_0))[Z_j + h_{1,j} + \Pi_{0,j}'\lambda]\right) \text{ if } j \le p \text{ and } h_{1,j} < \infty, \\
G_{\kappa,n,j}^0 &\to_d G\left(\kappa^{-1}(\Omega(\theta_0))[Z_j + \Pi_{0,j}'\lambda]\right) \text{ if } j = p+1, ..., k, \\
G_{\kappa,n}^0 &= (G_{\kappa,n,1}^0, ..., G_{\kappa,n,k}^0) \to_d G_{\kappa,\infty}^0 = \\
&\quad (G(\kappa^{-1}(\Omega(\theta_0))[Z_1 + h_{1,1} + \Pi_{0,1}'\lambda]), ..., G(\kappa^{-1}(\Omega(\theta_0))[Z_k + \Pi_{0,k}'\lambda]))',
\end{aligned}$$

where $Z = (Z_1, ..., Z_k)'$ and $Z_j + h_{1,j} + \Pi_{0,j}'\lambda = \infty$ by definition if $h_{1,j} = \infty$. Now, the same argument as in (1.26)-(1.28) of the proof of Lemma 1(a) gives

$$c_n(\theta_0) \to_d q_S\left(\varphi(\kappa^{-1}(\Omega_0)[Z + (h_1, 0_v) + \Pi_0\lambda], \Omega_0), \Omega_0\right) + \eta(\Omega_0). \quad (1.41)$$

The only difference in the proof is that $\mathcal{Z}((h_1, 0_v), \Omega_0)$ and $\Xi((h_1, 0_v), \Omega)$ are replaced by $\mathcal{Z}((h_1, 0_v) + \Pi_0\lambda, \Omega_0)$ and $\Xi((h_1, 0_v) + \Pi_0\lambda, \Omega)$, respectively.

Next, by the same argument as in (1.29) in the proof of Lemma 1(a), we obtain

$$T_n(\theta_0) \to_d S([Z + (h_1, 0_v) + \Pi_0\lambda], \Omega_0). \quad (1.42)$$

Furthermore, the convergence in (1.41) and (1.42) is joint, which establishes that (1.16) holds with $(h_1, 0)$ replaced by $(h_1, 0_v) + \Pi_0\lambda$. Finally, given the latter result, the result of the Theorem holds by the same argument as in (1.17)-(1.19) in the proof of Lemma 1(a) with $(h_1, 0_v)$ replaced by $(h_1, 0_v) + \Pi_0\lambda$ and $CP(h_1, \Omega_0, \eta(\Omega_0))$ replaced by

$AsyPow(\mu, \Omega_0, S, \varphi, \kappa(\Omega_0), \eta(\Omega_0))$. $\square$

2 Appendix B

This Appendix gives supplemental numerical results to those given in the paper. Section 2.1 provides a table of the κ values that maximize average asymptotic power for various tests. These are the κ values that yield the asymptotic power reported in Table II of Section 6.2 of the paper. Section 2.1 also provides a table that is analogous to Table II of the paper but reports asymptotic sizes rather than asymptotic power.

Section 2.2 provides results that supplement those of Section 6.2 of the paper by comparing (S, φ) functions for a larger number of Ω matrices. These are results based on the best κ values in terms of average asymptotic power.

Section 2.4 provides additional asymptotic size and power results for some GMS and RMS tests that are not considered explicitly in the paper.

Section 2.5 provides comparative computation times for tests based on the QLR and MMM test statistics and the "asymptotic normal" and bootstrap versions of the t-test (i.e., $\varphi^{(1)}$) moment selection critical values.

2.1 κ Values That Maximize Average Asymptotic Power

The κ values that maximize average asymptotic power, i.e., the best κ values, which are used in the construction of Table II, are given in Table B-I.

Table B-II gives the asymptotic sizes of the RMS tests that appear in Table II and are based on the κ=Best tuning parameter and no size-correction factor, i.e., $\eta = 0$. The results show that the κ value that maximizes average asymptotic power also has quite good asymptotic size properties even with $\eta = 0$, with the exception of the SumMax/t-Test and QLR/$\varphi^{(3)}$ tests.

Table B-I. κ Values That Maximize (Size-Corrected) Asymptotic Power: MMM, Max, SumMax, & QLR Statistics; t-Test, $\varphi^{(3)}$, $\varphi^{(4)}$, & MMSC Critical Values[1]

	Crit.	$p = 10$			$p = 4$			$p = 2$		
Stat.	Val.	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}
MMM	t-Test	2.75	1.75	.25	2.50	1.50	.10	2.50	1.50	.50
Max	t-Test	2.50	1.25	.00	2.50	1.50	.50	2.50	1.50	.75
SumMax	t-Test	1.87	1.25	.25	2.25	1.50	.10	2.50	1.50	.50
QLR	t-Test	2.50	1.75	.00	2.75	1.50	.25	2.75	1.50	.75
QLR	$\varphi^{(3)}$	12.5†	3.00	1.25†	9.5*	2.25*	1.00*	8.00*	2.50*	.75*
QLR	$\varphi^{(4)}$	2.75†	1.75	.50†	2.75*	1.25*	.10*	2.75*	1.87*	.50*
QLR	MMSC	5.0	1.75	.10	7.5	1.50	.10	2.75	1.50	.75

[1] Results are based on (40000, 40000) size-correction and rejection probability repetitions for $p = 2, 4$ and (5000, 5000) repetitions for $p = 10$, unless noted otherwise.

*Results are based on (5000, 5000) repetitions.

†Results are based on (2000, 2000) repetitions.

Table B-II. Asymptotic Size Comparisons: Max, SumMax, & QLR Statistics; t-Test, $\varphi^{(3)}$, & $\varphi^{(4)}$ Critical Values with κ=Best[1] & $\eta = 0$

	Crit.	Tuning	$p = 10$			$p = 4$			$p = 2$		
Stat.	Val.	Par. κ	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}
MMM	t-Test	Best	.051	.055	.052	.053	.055	.052	.051	.053	.054
Max	t-Test	Best	.051	.056	.053	.051	.054	.050	.051	.053	.051
SumMax	t-Test	Best	.172	.153	.158	.109	.092	.123	.051	.053	.054
QLR	$\varphi^{(3)}$	Best	$.100^{\dagger}$	$.074^{*}$	$.052^{\dagger}$	$.101^{*}$	$.065^{*}$	$.051^{*}$	$.073^{*}$	$.059^{*}$	$.054^{*}$
QLR	$\varphi^{(4)}$	Best	$.054^{\dagger}$	$.054^{*}$	$.052^{\dagger}$	$.052^{*}$	$.058^{*}$	$.051^{*}$	$.051^{*}$	$.052^{*}$	$.053^{*}$
QLR	t-Test	Best	.057	.055	.054	.051	.055	.051	.051	.052	.052
QLR	MMSC	Best	.056	.055	.053	.055	.055	.052	.052	.052	.052

[1] κ=Best denotes the κ value that maximizes average asymptotic power. Except where stated otherwise, the results are based on (40000, 40000) critical value and rejection probability repetitions.

*Results are based on (5000, 5000) critical value and rejection probability repetitions.

†Results are based on (2000, 2000) critical value and rejection probability repetitions.

2.2 Comparison of (S, φ) Functions: 19 Ω Matrices

Here we compare the MMM/t-Test/κBest, QLR/t-Test/κBest, QLR/t-Test/κAuto, & QLR/MMSC/κBest tests. This section is quite similar to Section 6.2 of the paper except that 19 Ω matrices are considered here, rather than 3, and fewer tests are considered. The 19 Ω matrices are the same as those considered in Table III in Section 6.2.3 of the paper and defined in Appendix C below.

The qualitative results reported in Section 6.2 of the paper are found here to apply as well to the broader range of Ω matrices that are considered.

TABLE B-III. Asymptotic Power Comparisons (Size-Corrected) for 19 Ω Matrices: MMM & QLR Statistics; t-Test & MMSC Critical Values with κ=Best & κAuto[1]

(a) $p = 10$

Stat.	Crit. Val.	κ	$\delta(\Omega)$: -.99	-.975	-.95	-.9	-.8	-.7	-.6	-.5	-.4	-.2
MMM	t-Test	κBest	.19	.19	.19	.19	.21	.24	.29	.35	.43	.57
QLR	t-Test	κBest	.96	.94	.80	.58	.48	.48	.49	.51	.54	.61
QLR	t-Test	κAuto	.96	.94	.79	.58	.48	.47	.49	.51	.54	.61
QLR	MMSC	κBest	.96*	.96*	.83*	.65*	.52*	.50*	.52*	.54*	.56*	.61*
Power	Envelope	-	.98	.98	.94	.85	.74	.73	.74	.75	.77	.81

			$\delta(\Omega)$: 0.0	.2	.4	.6	.8	.9	.95	.975	.99
MMM	t-Test	κBest	.67	.36	.50	.85	.82	.80	.80	.79	.79
QLR	t-Test	κBest	.67	.37	.51	.85	.83	.82	.82	.81	.81
QLR	t-Test	κAuto	.67	.37	.51	.85	.83	.82	.82	.81	.81
QLR	MMSC	κBest	.66*	.36*	.50*	.84*	.83*	.82*	.81*	.80*	.81*
Power	Envelope	-	.85	.47	.59	.88	.85	.83	.82	.81	.81

[1] κ=Best denotes the κ value that maximizes average asymptotic power. Except where stated otherwise, the results are based on (40000, 40000) critical value and rejection probability repetitions.

*Results are based on (2000, 2000) critical value and rejection probability repetitions when determining the best κ value. Results reported in the table that use the best κ value are based on (5000, 5000) critical value and rejection probability repetitions.

TABLE B-III (Cont.)

(b) $p = 4$

Stat.	Crit. Val.	κ	$\delta(\Omega)$: -.99	-.975	-.95	-.9	-.8	-.7	-.6	-.5	-.4	-.2
MMM	t-Test	κBest	.31	.31	.31	.32	.34	.37	.42	.47	.52	.62
QLR	t-Test	κBest	.93	.89	.76	.62	.52	.52	.53	.56	.58	.63
QLR	t-Test	κAuto	.92	.88	.76	.62	.52	.52	.53	.55	.58	.63
QLR	MMSC	κBest	.94	.90	.78	.65	.56	.55	.56	.57	.59	.64
Power	Envelope	-	.95	.94	.87	.80	.70	.69	.70	.72	.73	.77
			$\delta(\Omega)$: 0.0	.2	.4	.6	.8	.9	.95	.975	.99	
MMM	t-Test	κBest	.68	.45	.59	.80	.79	.78	.78	.77	.77	
QLR	t-Test	κBest	.68	.46	.59	.80	.80	.80	.79	.78	.78	
QLR	t-Test	κAuto	.68	.45	.59	.80	.80	.79	.79	.78	.78	
QLR	MMSC	κBest	.68	.46	.59	.80	.80	.79	.79	.78	.78	
Power	Envelope	-	.80	.53	.65	.83	.80	.79	.79	.78	.78	

(c) $p = 2$

Stat.	Crit. Val.	κ	$\delta(\Omega)$: -.99	-.975	-.95	-.9	-.8	-.7	-.6	-.5	-.4	-.2
MMM	t-Test	κBest	.52	.52	.51	.51	.52	.54	.57	.59	.61	.65
QLR	t-Test	κBest	.86	.83	.76	.65	.60	.59	.60	.61	.62	.65
QLR	t-Test	κAuto	.84	.81	.75	.64	.60	.59	.60	.61	.62	.65
QLR	MMSC	κBest	.86	.83	.76	.65	.60	.59	.60	.61	.62	.65
Power	Envelope	-	.88	.86	.83	.75	.70	.69	.69	.70	.70	72
			$\delta(\Omega)$: 0.0	.2	.4	.6	.8	.9	.95	.975	.99	
MMM	t-Test	κBest	.69	.58	.65	.71	.72	.73	.73	.73	.73	
QLR	t-Test	κBest	.69	.58	.66	.72	.73	.73	.73	.73	.73	
QLR	t-Test	κAuto	.69	.58	.66	.72	.73	.73	.73	.73	.73	
QLR	MMSC	κBest	.69	.58	.66	.72	.73	.73	.73	.73	.73	
Power	Envelope	-	.75	.63	.70	.74	.74	.73	.73	.73	.73	

2.3 Comparison of RMS and GMS Procedures

In this section, we provide asymptotic size and power comparisons (based on fixed κ asymptotics) of several GMS tests and the recommended RMS test, which is the QLR/t-Test/κAuto test.

We consider GMS tests based on $(S, \varphi) =$ (MMM, t-Test), (QLR, t-Test), and (QLR, MMSC). The GMS tests depend on a tuning parameter κ $(= \kappa_n)$ that does not depend on Ω. We consider the values κ=2.35 and κ=1.87. The former corresponds to the BIC choice $\kappa_n = (\ln n)^{1/2}$ for $n = 250$ and the latter corresponds to the LIL choice $\kappa_n = (2 \ln \ln n)^{1/2}$ for $n = 300$. Note that the BIC choice yields $\kappa_n \in [2.15, 2.63]$ for $n \in [100, 1000]$ and the LIL choice yields $\kappa_n \in [1.75, 1.97]$ for $n \in [100, 1000]$.

Tables B-IV and B-V provide the asymptotic size and power results, respectively, for $p = 2, 4, 10$ and $\Omega = \Omega_{Neg}, \Omega_{Zero}, \Omega_{Pos}$. The critical values are obtained using $40,000$ simulation repetitions and both the size and power results are obtained using $40,000$ repetitions, which yields a simulation standard error of .0011. The power results are size-corrected.

Table B-IV shows that the GMS tests with κ=1.87 have asymptotic size that is close to .050 for Ω_{Pos}, is slightly above .050 for Ω_{Zero}, and is noticeably above .050 for Ω_{Neg}. For example, for Ω_{Neg}, the QLR/t-Test/κ=1.87 test has size .074, .080, and .078 for $p = 2, 4$, and 10, respectively. The amount of over-rejection is higher for the QLR/MMSC test than for the QLR/t-Test and MMM/t-Test tests.

The GMS tests with κ=2.35 have asymptotic size that is closer to .050 than when κ=1.87. There is still some over-rejection with Ω_{Neg}, especially for the QLR/MMSC/ κ=2.35 test. But it is noticeably smaller. For example, for Ω_{Neg}, the QLR/t-Test/κ=2.35 test has size .055, .059, and .059 for $p = 2, 4$, and 10, respectively.

The recommended RMS test has asymptotic size that is close to its nominal level .050. It is within three simulation standard errors of the nominal level for all cases considered. For Ω_{Neg}, it has size .046, .048, and .050 for $p = 2, 4$, and 10, respectively.

Based on Table B-IV, we conclude that some GMS tests have moderate to large problems of over-rejection asymptotically (under fixed κ) asymptotics for some Ω matrices. However, some GMS tests with κ=2.35 perform quite well and over-reject by a relatively small amount. The recommended RMS test performs well. It shows no sign of over-rejection and its asymptotic size is close to its nominal level.

Next, we discuss the asymptotic power results given in Table B-V. Table B-V shows that the GMS tests given by MMM/t-Test with κ=2.35 and κ=1.87 have quite low

Table B-IV. Asymptotic Size Comparisons for Nominal .05 Tests: MMM & QLR Statistics; t-Test & MMSC Critical Values with κ=2.35, κ=1.87, & κAuto

	Crit.	Tuning	$p=10$			$p=4$			$p=2$		
Stat.	Val.	Par. κ	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}
MMM	t-Test	2.35	.059	.051	.051	.053	.050	.050	.052	.050	.051
MMM	t-Test	1.87	.070	.054	.050	.068	.053	.050	.063	.052	.050
QLR	t-Test	2.35	.059	.051	.051	.059	.050	.049	.055	.050	.050
QLR	t-Test	1.87	.078	.054	.050	.080	.053	.050	.074	.052	.050
QLR	MMSC	2.35	.106	.051	.051	.093	.050	.049	.056	.050	.050
QLR	MMSC	1.87	.123	.054	.050	.115	.053	.050	.074	.052	.050
QLR	t-Test	Auto	.046	.049	.041	.048	.051	.047	.050	.050	.050

power compared to the recommended RMS test (i.e., QLR/t-Test/κAuto) for Ω_{Neg} and noticeably lower power for Ω_{Pos}. For Ω_{Neg}, the powers of the MMM/t-Test tests are decreasing in p rather quickly.

The GMS tests QLR/t-Test/κ=1.87 and QLR/MMSC/κ=1.87 have power that is the same as that of the RMS test for Ω_{Zero}. For Ω_{Pos}, these two GMS tests have power that is only slightly lower than that of the RMS test. On the other hand, for Ω_{Neg}, the power of these two GMS tests is noticeably less than that of the RMS test, especially for $p = 2, 4$. As discussed above, a drawback of these GMS tests is that they over-reject the null hypothesis with Ω_{Neg}.

The QLR/t-Test/κ=2.35 and QLR/MMSC/κ=2.35 tests have similar asymptotic power but the former has higher power for Ω_{Pos}, especially for $p = 10$. In fact, the QLR/t-Test/κ=2.35 is the best GMS test in terms of overall power. Its power is uniformly dominated by that of the recommended RMS test, but the differences in power are not large.

We conclude that (i) the best GMS test in terms of asymptotic size and power is the QLR/t-Test/κ=2.35, (ii) the recommended RMS test out-performs this GMS test in terms of asymptotic size and power in all cases considered, but the differences between the two are not large, and (iii) the recommended RMS test out-performs the other GMS tests considered by a noticeable margin in terms of asymptotic size and/or power.

Table B-V. Asymptotic Power Comparisons (Size-Corrected) for Nominal .05 Tests: MMM & QLR Statistics; PA, t-Test, & MMSC Critical Values with κ=2.35, κ=1.87, & κAuto

	Crit.	Tuning	$p = 10$			$p = 4$			$p = 2$		
Stat.	Val.	Par. κ	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}
MMM	t-Test	2.35	.19	.65	.68	.31	.68	.68	.51	.69	.68
MMM	t-Test	1.87	.16	.67	.72	.28	.69	.71	.48	.69	.69
QLR	t-Test	2.35	.58	.65	.79	.61	.68	.76	.65	.69	.70
QLR	t-Test	1.87	.55	.67	.81	.56	.69	.77	.60	.69	.71
QLR	MMSC	2.35	.58	.65	.75	.60	.68	.75	.64	.69	.70
QLR	MMSC	1.87	.56	.67	.78	.55	.69	.77	.60	.69	.71
QLR	t-Test	Auto	.58	.67	.82	.62	.69	.78	.65	.69	.72
Power	Envelope	-	.85	.85	.85	.80	.80	.80	.75	.75	.75

2.4 Additional Asymptotic Size & Power Results

Table B-VI reports asymptotic size results for some tests that are not considered in the text of the paper or Section 2.3 above. Table B-VII does likewise for asymptotic power.

Table B-VI. Asymptotic Size Comparisons of Nominal .05 Tests: MMM, Max, SumMax, & QLR Statistics; PA, t-Test, $\varphi^{(3)}$, $\varphi^{(4)}$, & MMSC Critical Values with κ=Best, κ=2.35, & κ=1.87; & $\eta = 0$[1]

	Crit.	Tuning	$p = 10$			$p = 4$			$p = 2$		
Stat.	Val.	Par. κ	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}
MMM	PA	-	.050	.050	.050	.050	.050	.050	.050	.050	.050
QLR	PA	-	.050	.050	.050	.050	.050	.050	.050	.050	.050
GEL	Const.	-	.021	.010	.000	.050	.025	.006	.047	.032	.026
MMM	t-Test	Best	.051	.055	.052	.053	.055	.052	.051	.053	.054
MMM	t-Test	2.35	.059	.051	.051	.053	.050	.050	.052	.050	.051
MMM	t-Test	1.87	.070	.054	.050	.068	.053	.050	.063	.052	.050
Max	PA	-	.050	.050	.050	.050	.050	.050	.050	.050	.050
Max	t-Test	Best	.051	.056	.053	.051	.054	.050	.051	.053	.051
Max	t-Test	2.35	.054	.051	.051	.051	.050	.050	.052	.050	.050
Max	t-Test	1.87	.063	.052	.050	.064	.052	.050	.063	.051	.050
SumMax	t-Test	Best	.172	.153	.158	.109	.092	.123	.051	.053	.054
SumMax	t-Test	2.35	.164	.149	.147	.103	.087	.118	.052	.062	.077
SumMax	t-Test	1.87	.172	.162	.153	.111	.090	.120	.063	.052	.050

Table B-VI. (Cont.)

Stat.	Crit. Val.	Tuning Par. κ	$p=10$ Ω_{Neg}	Ω_{Zero}	Ω_{Pos}	$p=4$ Ω_{Neg}	Ω_{Zero}	Ω_{Pos}	$p=2$ Ω_{Neg}	Ω_{Zero}	Ω_{Pos}
QLR	$\varphi^{(3)}$	Best	.100†	.074*	.052†	.101*	.065*	.051*	.073*	.059*	.054*
QLR	$\varphi^{(3)}$	2.35	.225†	.070†	046.†	.160*	.061*	.042*	.095*	.054*	.046*
QLR	$\varphi^{(3)}$	1.87	.280†	.085*	.051†	.184*	.069*	.051*	.113*	.062*	.052*
QLR	$\varphi^{(4)}$	Best	.054†	.054*	.052†	.052*	.058*	.051*	.051*	.052*	.053*
QLR	$\varphi^{(4)}$	2.35	.057†	.045†	.046†	.055*	.045*	.041*	.047*	.047*	.045*
QLR	$\varphi^{(4)}$	1.87	.079†	.053*	.050†	.080*	.052*	.051*	.076*	.052*	.050*
QLR	t-Test	Best	.057	.055	.054	.051	.055	.051	.051	.052	.052
QLR	t-Test	2.35	.059	.051	.051	.059	.050	.049	.055	.050	.050
QLR	t-Test	1.87	.078	.054	.050	.080	.053	.050	.074	.052	.050
QLR	t-Test	Auto	.046	.049	.041	.048	.051	.047	.050	.050	.050
QLR	MMSC	Best	.056	.055	.053	.055	.055	.052	.052	.052	.052
QLR	MMSC	2.35	.106	.051	.051	.093	.050	.049	.056	.050	.050
QLR	MMSC	1.87	.123	.054	.050	.115	.053	.050	.074	.052	.050

[1] κ=Best denotes the κ value that maximizes average asymptotic power. Unless stated otherwise, results are based on (40000, 40000) critical value and rejection probability repetitions.

*Results are based on (5000, 5000) critical value and rejection probability repetitions.

†Results are based on (2000, 2000) critical value and rejection probability repetitions.

Table B-VII. Asymptotic Power Comparisons (Size-Corrected) of Nominal .05 Tests: MMM, Max, SumMax, & QLR Statistics; t-Test, $\varphi^{(3)}$, $\varphi^{(4)}$, & MMSC Critical Values with κ=Best, κ=2.35, κ=1.87, & κAuto[1]

	Crit.	Tuning	$p = 10$			$p = 4$			$p = 2$		
Stat.	Val.	Par. κ	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}	Ω_{Neg}	Ω_{Zero}	Ω_{Pos}
MMM	PA	-	.04	.36	.36	.20	.52	.46	.48	.62	.59
QLR	PA	-	.28	.36	.70	.44	.52	.71	.58	.62	.65
GEL	Const.	-	.19	.18	.12	.44	.42	.39	.52	.54	.54
MMM	t-Test	Best	.19	.67	.79	.32	.69	.77	.51	.69	.71
MMM	t-Test	2.35	.19	.65	.68	.31	.68	.68	.51	.69	.68
MMM	t-Test	1.87	.16	.67	.72	.28	.69	.71	.48	.69	.69
Max	PA	-	.18	.44	.72	.30	.55	.71	.48	.63	.66
Max	t-Test	Best	.25	.59	.82	.35	.66	.79	.51	.69	.72
Max	t-Test	2.35	.25	.57	.79	.35	.65	.76	.51	.68	.71
Max	t-Test	1.87	.24	.59	.81	.34	.66	.77	.48	.69	.71
SumMax	t-Test	Best	.14	.55	.71	.24	.64	.65	.51	.69	.71
SumMax	t-Test	2.35	.14	.55	.69	.24	.62	.64	.51	.67	.63
SumMax	t-Test	1.87	.14	.55	.70	.24	.64	.64	.48	.69	.69

Table B-VII. (Cont.)

Stat.	Crit. Val.	Tuning Par. κ	$p=10$ Ω_{Neg}	$p=10$ Ω_{Zero}	$p=10$ Ω_{Pos}	$p=4$ Ω_{Neg}	$p=4$ Ω_{Zero}	$p=4$ Ω_{Pos}	$p=2$ Ω_{Neg}	$p=2$ Ω_{Zero}	$p=2$ Ω_{Pos}
QLR	$\varphi^{(3)}$	Best	.49†	.62*	.83†	.54*	.67*	.78*	.60*	.67*	.72*
QLR	$\varphi^{(3)}$	2.35	.38†	.64†	.83†	.50*	.67*	.78*	.58*	.68*	.72*
QLR	$\varphi^{(3)}$	1.87	.40†	.61*	.82†	.50*	.66*	.78*	.57*	.67*	.72*
QLR	$\varphi^{(4)}$	Best	.59†	.67*	.82†	.62*	.69*	.78*	.65*	.69*	.72*
QLR	$\varphi^{(4)}$	2.35	.57†	.63†	.79†	.60*	.66*	.75*	.64*	.67*	.69*
QLR	$\varphi^{(4)}$	1.87	.57†	.67*	.81†	.55*	.69*	.77*	.59*	.69*	.71*
QLR	t-Test	Best	.59	.67	.82	.62	.69	.78	.65	.69	.72
QLR	t-Test	2.35	.58	.65	.79	.61	.68	.76	.65	.69	.70
QLR	t-Test	1.87	.58	.67	.81	.56	.69	.77	.60	.69	.71
QLR	t-Test	Auto	.58	.67	.82	.62	.69	.78	.65	.69	.72
QLR	MMSC	Best	.65	.67	.78	.65	.69	.78	.65	.69	.72
QLR	MMSC	2.35	.58	.65	.75	.60	.68	.75	.64	.69	.70
QLR	MMSC	1.87	.56	.67	.82	.55	.69	.77	.60	.69	.71
Power	Envelope	-	.85	.85	.85	.80	.80	.80	.75	.75	.75

[1] κ=Best denotes the κ value that is best in terms of average asymptotic power. Unless stated otherwise, results are based on (40000, 40000) critical value and rejection probability repetitions.

*Results are based on (5000, 5000) critical value and rejection probability repetitions.

†Results are based on (2000, 2000) critical value and rejection probability repetitions.

2.5 Comparative Computation Times

As reported in the paper, to compute the recommended bootstrap RMS test, i.e., QLR/t-Test/κAuto/Boot, using 10,000 critical value simulation repetitions takes 1.3, 1.7, 3.2, 8.4, 17.2, and 52.0 seconds when $p = 2$, 4, 10, 20, 30, and 50, respectively, and $n = 250$ using a PC with a 3.4 GHz processor. For the asymptotic normal version of the recommended bootstrap RMS test, i.e., QLR/t-Test/κAuto/Norm, the times are .25, .31, .71, 2.4, 6.1, and 21.8 seconds, respectively.

In contrast, to compute the bootstrap version of the MMM/t-Test/κ=2.35 test using 10,000 critical value simulation repetitions takes .86, .98, 2.0, 5.9, 11.6, and 28.4 seconds when $p = 2$, 4, 10, 20, 30, and 50, respectively, and $n = 250$. For the asymptotic normal version of the MMM/t-Test/κ=2.35 test, the times are .008, .010, .029, .060, .090, and .18 seconds, respectively. Note that the computation times are not affected by whether κ is taken to be κAuto or κ=2.35. The difference between the results in the previous paragraph and this paragraph is due to the different statistics used: QLR and MMM.

The results indicate that the bootstrap version of the MMM-based test is between 1.4 and 1.8 times faster than the corresponding bootstrap version of the QLR-based test. On the other hand, the asymptotic normal version of the MMM-based test is very much faster (from 20 to 85 times) than asymptotic normal version of the QLR-based test. (This is because the generation of the bootstrap samples dominates the computation time for the bootstrap version of the MMM-based test.)

When constructing a CS, if the computation time is burdensome (because one needs to carry out many tests with different values of θ as the null value), then the results above suggest that a useful approach is to map out the general features of the CS using the asymptotic normal version of the MMM/t-Test/κ=2.35 test, which is very fast to compute, and then switch to the bootstrap version of the QLR/t-Test/κAuto test to find the boundaries of the CS more precisely.

2.6 Magnitude of RMS Critical Values

Table B-VIII provides information on the magnitude of the preferred RMS critical value when the size-correction factor $\widehat{\eta}$ is not included. (Recall that the RMS critical value equals $c_n(\theta, \widehat{\kappa}) + \widehat{\eta}$.) Specifically, the Table provides simulated values of the mean and standard deviation of the asymptotic distribution of the data-dependent quantile $c_n(\theta, \widehat{\kappa}) = q_{S_2}(\varphi^{(1)}(\xi_n(\theta), \widehat{\Omega}_n(\theta)), \widehat{\Omega}_n(\theta))$ in various scenarios. The mean values in Table

B-VIII can be compared with the values of the components $\eta_1(\delta)$ and $\eta_2(p)$ (given in Table I of the paper) of the size-correction factor $\widehat{\eta}$ $(= \eta_1(\widehat{\delta}_n(\theta)) + \eta_2(p))$ to see how large the quantile $c_n(\theta, \widehat{\kappa})$ is (on average) compared to the size-correction factor $\widehat{\eta}$.

The asymptotic distribution of $c_n(\theta, \widehat{\kappa})$ depends on h_1 and Ω. Table B-VIII considers the same three correlation matrices Ω_{Neg}, Ω_{Zero}, and Ω_{Pos} as considered elsewhere in the paper, see Section 6 of the paper for their definitions. Table B-VIII considers h_1 vectors that consist of $0's$ and $\infty's$. (Other h_1 vectors are of interest, but for brevity we do not consider them here.) When an element of h_1 equals ∞, the corresponding moment inequality is far from binding and the moment selection procedure detects this with probability one asymptotically and does not include this moment when computing $c_n(\theta, \widehat{\kappa})$. When an element of h_1 equals 0, the corresponding moment inequality is binding and the moment selection procedure includes this moment with high probability but not with probability one, even asymptotically. (It is for this reason that $c_n(\theta, \widehat{\kappa})$ is random asymptotically.) In consequence, the asymptotic distribution depends on h_1 through the "# of Zeros in h_1" and through the sub-matrix of Ω that corresponds to the "Zeros in h_1." The matrices Ω_{Neg}, Ω_{Zero}, and Ω_{Pos} are defined such that for any value of p the sub-matrix of Ω of dimension equal to the "# of Zeros in h_1" is the same (provided $p \geq$"# of Zeros in h_1"). In consequence, the results of Table B-VIII hold for any value of p. For example, if $p = 20$, $\Omega = \Omega_{Neg}$, and the "# of Zeros in h_1" is 5, one obtains the same mean and standard deviation of the asymptotic distribution of $c_n(\theta, \widehat{\kappa})$ as when $p = 15$, $\Omega = \Omega_{Neg}$, and the "# of Zeros in h_1" is 5.

The results of Table B-VIII, combined with the magnitudes of the size-correction factors given in Table I, show that the size-correction factor $\widehat{\eta} = \eta_1(\widehat{\delta}_n(\theta)) + \eta_2(p)$ typically is small compared to $c_n(\theta, \widehat{\kappa})$, but not negligible. For example, for $p = 10$, $\Omega = \Omega_{Zero} = I_{10}$, and $h_1 = (0, 0, 0, 0, 0, \infty, \infty, \infty, \infty, \infty)'$ (which corresponds to five moment inequalities being binding and five being very far from binding), the mean and standard deviation of the asymptotic distribution of $c_n(\theta, \widehat{\kappa})$ are 8.7 and .13, respectively, whereas the size-correction factor is .48.

Table B-VIII. Mean and Standard Deviation of the Asymptotic Distribution of the Data-Dependent RMS Critical Values Excluding the Size-Correction Factor $\widehat{\eta}$[1]

	Ω_{Neg}		Ω_{Zero}		Ω_{Pos}	
# of	Mean	SD	Mean	SD	Mean.	SD
Zero's in h_1	$c_n(\theta,\widehat{\kappa})$	$c_n(\theta,\widehat{\kappa})$	$c_n(\theta,\widehat{\kappa})$	$c_n(\theta,\widehat{\kappa})$	$c_n(\theta,\widehat{\kappa})$	$c_n(\theta,\widehat{\kappa})$
1	2.7	.00	2.7	.00	2.7	.00
2	5.0	.13	4.1	.53	3.5	.55
3	6.2	.11	5.2	.52	4.1	.68
4	7.5	.11	6.2	.54	4.5	.76
5	8.7	.13	7.2	.57	5.0	.82
6	9.8	.14	8.1	.59	5.3	.86
7	10.9	.16	8.9	.57	5.6	.89
8	11.9	.16	9.7	.63	5.9	.90
9	12.9	.17	10.6	.66	6.1	.92
10	13.8	.17	11.4	.68	6.3	.94
15	19.4	.24	15.0	.70	7.2	.98
20	24.5	.25	18.4	.78	7.9	1.0
25	29.9	.31	21.6	.85	8.4	1.0
30	35.2	.32	24.8	.93	8.8	1.0
35	40.5	.35	27.9	.99	9.1	1.0
40	45.8	.38	31.0	1.0	9.4	1.0
45	51.2	.42	34.0	1.1	9.7	1.0
50	56.4	.42	36.9	1.1	10.0	1.0

[1] Results are based on 40,000 simulation repetitions.

3 Appendix C

This Appendix contains the following: (i) the definition of the μ vectors used in Section 6 of the paper, (ii) a description of some details concerning the power assessment given in Section 6.3.2 of the paper concerning the recommended RMS test, (iii) a discussion of the determination and computation of the asymptotic power envelope, (iv) a discussion of the computation of the κ values that maximize average asymptotic power that are reported in Table II of the paper, and (v) a description of the numerical computation of $\eta_2(p)$, which is part of the recommended size-correction function $\eta(\cdot)$.

3.1 μ Vectors

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

$$\begin{aligned}\mathcal{M}_2(I_2) &= \{(-2.309, 0), (-2.309, 1), (-2.309, 2), (-2.309, 3),\\ &\qquad (-2.309, 4), (-2.309, 7), (-1.6263, -1.6263)\},\\ \mathcal{M}_2(\Omega_{Neg}) &= \{(-1.001, 0), (-1.804, 1), (-2.303, 2), (-2.309, 3),\\ &\qquad (-2.309, 4), (-2.309, 7), (-0.5165, -0.5165)\}, \qquad (3.1)\\ \mathcal{M}_2(\Omega_{Pos}) &= \mathcal{M}_k(I_p) \text{ except the last vector is } (-2.0040, -2.0040).\end{aligned}$$

The power envelope at each of these μ vectors is .750.

For $p = 4$, the μ vectors in $\mathcal{M}_4(I_4)$ are defined by

$$\begin{aligned}&\mathcal{M}_4(\Omega)\\ &= \{(-\mu_1, -\mu_1, 1, 1), (-\mu_2, -\mu_2, 2, 2), (-\mu_3, -\mu_3, 3, 3), (-\mu_4, -\mu_4, 4, 4), (-\mu_5, -\mu_5, 7, 7),\\ &\quad (-\mu_6, -\mu_6, 1, 7), (-\mu_7, -\mu_7, 2, 7), (-\mu_8, -\mu_8, 3, 7), (-\mu_9, -\mu_9, 4, 7),\\ &\quad (-\mu_{10}, 1, 1, 1), (-\mu_{11}, 2, 2, 2), (-\mu_{12}, 3, 3, 3), (-\mu_{13}, 4, 4, 4), (-\mu_{14}, 7, 7, 7),\\ &\quad (-\mu_{15}, 1, 1, 7), (-\mu_{16}, 2, 2, 7), (-\mu_{17}, 3, 3, 7), (-\mu_{18}, 4, 4, 7), (-\mu_{19}, -\mu_{19}, 0, 0),\\ &\quad (-\mu_{20}, 0, 0, 0), (-\mu_{21}, 25, 25, 25), (-\mu_{22}, -\mu_{22}, 25, 25), (-\mu_{23}, -\mu_{23}, -\mu_{23}, 25),\\ &\quad (-\mu_{24}, -\mu_{24}, -\mu_{24}, -\mu_{24})\}, \qquad (3.2)\end{aligned}$$

and the following: $\mu_j = 1.7388$ for $j = 1, ..., 9, 19, 22$; $\mu_j = -2.4705$ for $j = 10, ..., 18, 20, 21$; $\mu_{23} = 1.4242$; and $\mu_{24} = 1.2350$.

For $p = 4$, the μ vectors in $\mathcal{M}_4(\Omega_{Neg})$ are defined by (3.2) and the following: $\mu_1 =$

-0.5505, $\mu_j = -0.5526$ for $j = 2, ..., 5$, $\mu_6 = -0.5505$, $\mu_j = -0.5526$ for $j = 7, 8, 9$, $\mu_{10} = -1.8814$, $\mu_{11} = -2.4283$, $\mu_j = -2.4705$ for $j = 12, 13, 14, 17, 18, 21$, $\mu_{15} = -1.8814$, $\mu_{16} = -2.4283$, $\mu_{19} = -0.3176$, $\mu_{20} = -0.8624$, $\mu_{22} = -0.5526$, $\mu_{23} = -0.2607$, $\mu_{24} = -0.1756$.

For $p = 4$, the μ vectors in $\mathcal{M}_4(\Omega_{Pos})$ are defined by (3.2) and the following: $\mu_j = 2.4047$ for $j = 1, ..., 9, 19, 22$; $\mu_j = -2.4705$ for $j = 10, ..., 18, 20, 21$; $\mu_{23} = 2.2628$; and $\mu_{24} = -2.1293$.

For $p = 4$, the power envelope at each of the μ vectors is .800.

For $p = k = 10$, $\mathcal{M}_{10}(\Omega)$ includes 40 vectors:

$$
\begin{aligned}
&\mathcal{M}_{10}(\Omega) \\
&= \{(-\mu_1, -\mu_1, 1, ..., 1), (-\mu_2, -\mu_2, 2, ..., 2), (-\mu_3, -\mu_3, 3, ..., 3), (-\mu_4, -\mu_4, 4, ..., 4), \\
&\quad (-\mu_5, -\mu_5, 7, ..., 7), (-\mu_6, -\mu_6, 1, 1, 1, 7, ..., 7), (-\mu_7, -\mu_7, 2, 2, 2, 7, ..., 7), \\
&\quad (-\mu_8, -\mu_8, 3, 3, 3, 7, ..., 7), (-\mu_9, -\mu_9, 4, 4, 4, 7, ..., 7), (-\mu_{10}, -\mu_{10}, -\mu_{10}, -\mu_{10}, 1, ..., 1), \\
&\quad (-\mu_{11}, -\mu_{11}, -\mu_{11}, -\mu_{11}, 2, ..., 2), (-\mu_{12}, -\mu_{12}, -\mu_{12}, -\mu_{12}, 3, ..., 3), \\
&\quad (-\mu_{13}, -\mu_{13}, -\mu_{13}, -\mu_{13}, 4, ..., 4), (-\mu_{14}, -\mu_{14}, -\mu_{14}, -\mu_{14}, 7, ..., 7), \\
&\quad (-\mu_{15}, -\mu_{15}, -\mu_{15}, -\mu_{15}, 1, 1, 1, 7, 7, 7), (-\mu_{16}, -\mu_{16}, -\mu_{16}, -\mu_{16}, 2, 2, 2, 7, 7, 7), \\
&\quad (-\mu_{17}, -\mu_{17}, -\mu_{17}, -\mu_{17}, 3, 3, 3, 7, 7, 7), (-\mu_{18}, -\mu_{18}, -\mu_{18}, -\mu_{18}, 4, 4, 4, 7, 7, 7), \\
&\quad (-\mu_{19}, 1, ..., 1), (-\mu_{20}, 2, ..., 2), (-\mu_{21}, 3, ..., 3), (-\mu_{22}, 4, ..., 4), (-\mu_{23}, 7, ..., 7), \\
&\quad (-\mu_{24}, 1, 1, 1, 7, ..., 7), (-\mu_{25}, 2, 2, 2, 7, ..., 7), (-\mu_{26}, 3, 3, 3, 7, ..., 7), (-\mu_{27}, 4, 4, 4, 7, ..., 7), \\
&\quad (-\mu_{28}, -\mu_{28}, 0, ..., 0), (-\mu_{29}, -\mu_{29}, -\mu_{29}, -\mu_{29}, 0, ..., 0), (-\mu_{30}, 0, ..., 0), \\
&\quad (-\mu_{31}, 25, ..., 25), (-\mu_{32}, -\mu_{32}, 25, ..., 25), (-\mu_{33}, -\mu_{33}, -\mu_{33}, 25, ..., 25), \\
&\quad (-\mu_{34}, -\mu_{34}, -\mu_{34}, -\mu_{34}, 25, ..., 25), (-\mu_{35}, -\mu_{35}, -\mu_{35}, -\mu_{35}, -\mu_{35}, 25, ..., 25), \\
&\quad (-\mu_{36}, ..., -\mu_{36}, 25, 25, 25, 25), (-\mu_{37}, ..., -\mu_{37}, 25, 25, 25), (-\mu_{38}, ..., -\mu_{38}, 25, 25), \\
&\quad (-\mu_{39}, ..., -\mu_{39}, 25), (-\mu_{40}, ..., -\mu_{40})\}. \qquad (3.3)
\end{aligned}
$$

For $p = 10$, the μ vectors in $\mathcal{M}_{10}(I_{10})$ are defined by (3.3) and the following: $\mu_j = 1.8927$ for $j = 1, ..., 9, 28, 32$ $\mu_j = 1.3360$ for $j = 10, ..., 18, 29, 34$, $\mu_j = 2.6817$ for $j = 19, ..., 27, 30, 31$, $\mu_{33} = 1.5463$, $\mu_{35} = 1.1963$, $\mu_{36} = 1.0893$, $\mu_{37} = 1.0099$, $\mu_{38} = 0.9465$, $\mu_{39} = 0.8882$, and $\mu_{40} = 0.8440$.

For $p = 10$, the μ vectors in $\mathcal{M}_{10}(\Omega_{Neg})$ are defined by (3.3) and the following: $\mu_j = 0.6016$ for $j = 1, ..., 9$, $\mu_j = 0.3475$ for $j = 10, ..., 18$, $\mu_{19} = 1.9847$, $\mu_{20} = 2.5835$,

$\mu_j = 2.6817$ for $j = 21, 22, 23, 26, 27, 31$, $\mu_{24} = 1.9847$, $\mu_{25} = 2.5835$, $\mu_{28} = 0.5341$, $\mu_{29} = 0.3322$, $\mu_{30} = 1.1551$, $\mu_{32} = 0.6016$, $\mu_{33} = 0.4195$, $\mu_{34} = 0.3475$, $\mu_{35} = 0.2985$, $\mu_{36} = 0.2674$, $\mu_{37} = 0.2430$, $\mu_{38} = 0.2254$, $\mu_{39} = 0.2106$, and $\mu_{40} = 0.1993$.

For $p = 10$, the μ vectors in $\mathcal{M}_{10}(\Omega_{Pos})$ are defined by (3.3) and the following: $\mu_j = 2.6227$ for $j = 1, ..., 9$, $\mu_j = 2.4676$ for $j = 10, ..., 18$, $\mu_j = 2.6817$ for $j = 19, ..., 27$, $\mu_{29} = 2.6227$, $\mu_{30} = 2.6817$, $\mu_{31} = 2.6817$, $\mu_{32} = 2.6227$, $\mu_{33} = 2.5401$, $\mu_{34} = 2.4676$, $\mu_{35} = 2.4005$, $\mu_{36} = 2.3140$, $\mu_{37} = 2.2846$, $\mu_{38} = 2.2565$, $\mu_{39} = 2.2343$, and $\mu_{40} = 2.2066$.

For $p = 10$, the power envelope at each of the μ vectors is .850.

3.2 Automatic κ Power Assessment Details

The 19 matrices Ω that are considered in Table III in Section 6.3.2 of the paper are Toeplitz matrices with elements on the diagonals given by the $(p-1)$-vectors ρ defined as follows. For $p = 2$, ρ takes the values for δ specified in Table III. For $p = 4, 10$, if $\delta \geq 0$, $\rho = (\delta, ..., \delta)$. For $p = 4$, if $\delta = -.99$, $\rho = (-.99, .97, -.95)$; if $\delta = -.975$, $\rho = (-.975, .94, -.90)$; if $\delta = -.95$, $\rho = (-.95, .9, -.8)$; and if $-.9 \leq \delta < 0$, $\rho = (\delta/(-.9)) \times (-.9, .7, -.5)$. For $p = 10$, if $\delta = -.99$, $\rho = (-.99, .97, -.95, .93, -.91, .89, -.87, .85, -.83)$; if $\delta = -.975$, $\rho = (-.975, .94, -.90, .86, -.82, .78, -.76, .74, -.72)$; if $\delta = -.95$, $\rho = (-.95, .9, -.8, .7, -.6, .5, -.4, .3, -.2)$; and if $-.9 \leq \delta < 0$, $\rho = (\delta/(-.9)) \times (-.9, .8, -.7, .6, -.5, .4, -.3, .2, -.1)$.

The randomly generated Ω matrices discussed in Section 6.3.2 of the paper have the following distributions. For $p = 2, 4$, the 500 Ω matrices are i.i.d. with $\Omega = Diag^{-1/2}(BB')BB' \times Diag^{-1/2}(BB')$, where B is a p by p matrix with independent $N(\varsigma_p, 1)$ elements, $\varsigma_p = 0$ for $p = 2$ and $\varsigma_p = .65$ for $p = 4$. The mean ς_p for $p = 4$ is chosen so that there is a more balanced distribution of $\delta(\Omega)$ values than is obtained if one takes $\zeta_p = 0$. For $p = 10$, the 250 Ω matrices are i.i.d. Toeplitz matrices (because this makes computation of size-correction values very much faster) that are the correlation matrices for moving-average (MA) processes of order $p - 1$ whose MA parameters are randomly generated. Specifically, Ω is the correlation matrix of an MA process $Y = (Y_1, ..., Y_p)$, where $Y_i = \sum_{j=0}^{p-1} a_j \varepsilon_{i-j}$ and $\{\varepsilon_i : i \leq p\}$ are i.i.d. with mean zero and variance one. The 250 Ω matrices are obtained by taking $\{a_j : j = 0, ..., p-1\}$ to be i.i.d. with a mixture of uniform distributions. With probability .7, a_j has a uniform distribution with mean zero and variance one, and with probability .3, a_j has a uniform distribution with mean one and variance one. This distribution for a_j is chosen to yield a fairly balanced distribution of $\delta(\Omega)$ values across the 250 Ω matrices. We obtain 175 negative values of $\delta(\Omega)$, 75 positive values, and a range of [-.90, .20].

The set of alternative hypothesis mean vectors μ, denoted $\mathcal{M}_p(\Omega)$, used in Section 6.3.2 of the paper contains linear combinations of μ vectors in $\mathcal{M}_p(\Omega_{Neg})$, $\mathcal{M}_p(\Omega_{Zero})$, and $\mathcal{M}_p(\Omega_{Pos})$. Specifically, for a given matrix Ω, $\mathcal{M}_p(\Omega)$ is defined by: (i) $\mathcal{M}_p(\Omega) = \mathcal{M}_p(\Omega_{Neg})$ if $\delta(\Omega) \in [-1.0, -.90]$, (ii) if $\delta(\Omega) \in [-.9, 0]$, $\mathcal{M}_p(\Omega) = \{\mu : \mu = (1 + \delta/.9)\mu_{Zero,j} - (\delta/.9)\mu_{Neg,j}$ for $j = 1, ..., J_p\}$, where $\mu_{Zero,j}$ denotes the jth element of $\mathcal{M}_p(\Omega_{Zero})$ and analogously for $\mathcal{M}_p(\Omega_{Neg})$ and $\mathcal{M}_p(\Omega_{Pos})$ and J_p denotes the numbers of elements in $\mathcal{M}_p(\Omega_{Zero})$, (iii) if $\delta(\Omega) \in [0, .5]$, $\mathcal{M}_p(\Omega) = \{\mu : \mu = (1 - \delta/.5)\mu_{Zero,j} + (\delta/.5)\mu_{Pos,j}$ for $j = 1, ..., J_p\}$, and (iv) if $\delta(\Omega) \in [0.5, 1.0]$, $\mathcal{M}_p(\Omega) = \mathcal{M}_p(\Omega_{Pos})$.

3.3 Asymptotic Power Envelope

We obtain an upper bound on the asymptotic power envelope by considering the simple-versus-simple likelihood ratio (SSLR) test for the desired alternative distribution and some selected null distribution, with the critical value chosen so that the test has the desired asymptotic null rejection rate α at the specified null distribution. This method of obtaining an upper bound on a power envelope also has been exploited in different contexts by Müller and Watson (2008) and Andrews, Moreira, and Stock (2008). If the specified null distribution is such that the SSLR test has maximum rejection probability equal to α over all null distributions, then the specified null distribution is least favorable and the SSLR test actually provides the asymptotic power envelope at the alternative distribution considered.

We assume that one observes $(n^{1/2}\overline{m}_n(\theta_0), \Sigma)$ and H_0 is defined as

$$\begin{aligned} H_0 : E_F m_j(W_i, \theta_0) &\geq 0 \text{ for } j = 1, ..., p \text{ and} \\ E_F m_j(W_i, \theta_0) &= 0 \text{ for } j = p + 1, ..., k, \end{aligned} \tag{3.4}$$

where F denotes the true distribution of the data. (More precisely, by this we mean H_0: the true $(\theta, F) \in \mathcal{F}$ satisfies $\theta = \theta_0$.) The simple alternative is $H_1 : F = F_n$, where F_n is a $n^{1/2}$-local alternative with asymptotic mean vector μ_{Alt}. Asymptotically, the distribution of $n^{1/2}\overline{m}_n(\theta_0)$ under the alternative is $N(\mu_{Alt}, \Sigma)$. We take the specified asymptotic null distribution to be $N(\mu_{Null}, \Sigma)$, where μ_{Null} is defined to minimize $(\mu - \mu_{Alt})'\Sigma^{-1}(\mu - \mu_{Alt})$ over $\mu \in R^p_{[+\infty]}$. In the numerical results reported below, we find that this choice of null distribution is least favorable. Thus, the upper bound on the asymptotic power envelope, up to numerical accuracy (based on 40,000 simulation repetitions), is the asymptotic power envelope.

3.4 Computation of κ Values That Maximize Average Asymptotic Power

Here we discuss the computation of the κ values that maximize average asymptotic power. These best κ values are used in the asymptotic power comparisons given in Table II of the paper. For all of the RMS tests in Table II, the best κ values are determined by grid search to an accuracy of .25. On a subset of cases this is found to be sufficiently small that the average asymptotic power is within than .01 of the maximum based on a finer grid. The grids of κ values used for the t-Test critical values and each test statistic considered are: for Ω_{Neg} : $\{3.25, 3.0, 2.75, 2.5, 1.87, 1.0, .25\}$; for $\Omega = I_p$: $\{2.75, 2.5, 2.25, 2.0, 1.87, 1.75, 1.5, 1.25, 1.0, .25\}$, and for V_{Pos} : $\{2.75, 1.87, 1.25, 1.0, .75, .50, .25, .10, .00\}$. For all of the test statistics considered, the average power values are well-behaved as a function of κ, there is no difficulty in finding the best κ value, and the best κ value is within the interior of the range considered. To ensure the latter, for the QLR/MMSC test, the following κ values also are included in the grids $\{3.5, 3.75, 4.0, 4.25, 5.0, 6.0, 7.0, 7.25, 7.5, 7.75, 8.0, 10.0\}$. For the QLR/$\varphi^{(3)}$ test, the grid is extended to 16 for Ω_{Neg} and to 3.5 for Ω_{Zero}.

3.5 Numerical Computation of $\eta_2(p)$

The size-correction factor $\eta_2(p)$ is determined as follows. Let p and Ω be given. For given (h_1, Ω), we compute the .95 sample quantile of

$$\begin{aligned}&\left\{S_2\left(\Omega^{1/2}Z_r + (h_1, 0_v), \Omega\right) - q_{S_2}\left(\varphi^{(1)}\left(\kappa^{-1}(\Omega)[\Omega^{1/2}Z_r + (h_1, 0_v)], \Omega\right), \Omega\right)\right. \\ &\left.+\eta_1(\delta(\Omega)) : r = 1, ..., R\right\}, \qquad (3.5)\end{aligned}$$

where $Z_r \sim$ i.i.d. $N(0_k, I_k)$ for $r = 1, ..., R$, where $R = 40,000$. Call the sample quantile $\eta_{h_1,\Omega}$. Up to simulation error, $\eta_{h_1,\Omega}$ is the smallest value that satisfies

$$CP(h_1, \Omega, \eta_1(\delta(\Omega)) + \eta_{h_1,\Omega}) = 1 - \alpha. \qquad (3.6)$$

The same simulated random variables $\{Z_r : r = 1, ..., R\}$ are used for all (h_1, Ω) considered. The critical value $q_{S_2}\left(\varphi^{(1)}\left(\kappa^{-1}(\Omega)[\Omega^{1/2}Z_r + (h_1, 0_v)], \Omega\right), \Omega\right)$ in (3.5) is obtained by simulation for each r. (The number of simulation repetitions employed is R here too and the same random numbers are used for each r).

Let $\mathcal{H}_1$ denote the set of all p vectors whose elements are $0's$ and $\infty's$. By considering a variety of subcases, we find that size is attained for $\mu \in \mathcal{H}_1$. That is, it suffices to restrict attention to maximization of $\eta_{h_1,\Omega}$ over $\mathcal{H}_1$, rather than over $R^p_{+,\infty}$. In addition, we approximate the maximization of $\eta_{h_1,\Omega}$ over the parameter space Ψ for Ω to a maximization of a finite set $\Psi^* \subset \Psi$. Given this, $\eta_2(p) \in R$ is defined to be

$$\sup_{h_1 \in \mathcal{H}_1, \Omega \in \Psi^*} \eta_{h_1,\Omega}. \tag{3.7}$$

For $p \leq 10$, the set Ψ^* is a set of correlation matrices that includes: (i) 43 Toeplitz matrices Ω that are such that $\delta(\Omega)$ takes values in a grid between $-.99$ and $.99$, [5] and (ii) 500 randomly generated matrices Ω that are generated by $\Omega = Corr(V)$, where $V = BB'$ and B is a $p \times p$ matrix with i.i.d. N(0, 1) elements.[6] As the number of randomly generated matrices Ω goes to infinity, the maximum of $\eta_{h_1,\Omega}$ over Ψ^* approaches the maximum over $\eta_{h_1,\Omega}$ over Ψ. Since the same underlying random variables $\{Z_r : r = 1, ..., R\}$ are used for each (h_1, Ω) considered, an empirical process CLT guarantees that as R and the number of random matrices Ω considered go to infinity the calculated critical values converge to the desired value $\eta_2(p)$ that satisfies

$$\inf_{h_1 \in \mathcal{H}_1, \Omega \in \Psi} CP(h_1, \Omega, \eta_1(\delta(\Omega)) + \eta_2(p)) = 1 - \alpha. \tag{3.8}$$

[5] For any given value of $\delta = \delta(\Omega)$, these 43 matrices are defined just as the 19 Toeplitz matrices are defined in Section 3.2. The $\delta(\Omega)$ values considered are the 43 values specified by the endpoints for δ in Table I, but including $-.99$ and excluding -1.0 and 1.0.

[6] For Toeplitz matrices, the null rejection probabilities of all of the tests considered in this paper are invariant to permutations of the elements of the null mean vector μ. Hence, with Toeplitz matrices one does not need to consider all 2^p null mean vectors containing $0's$ and $\infty's$. It suffices to consider only p vectors, viz., $(0, \infty, ..., \infty), (0, 0, \infty, ..., \infty), ..., (0, ..., 0)$. For non-Toeplitz matrices, this invariance property does not hold. The 500 randomly generated Ω matrices typically are non-Toeplitz. For such matrices and $p \leq 7$, we consider all $2^p - 1$ μ vectors of $0's$ and $\infty's$ (excluding the $(\infty, ..., \infty)$ vector). For such matrices and $8 \leq p \leq 10$, it is not feasible to consider all $2^p - 1$ μ vectors. Instead we randomly select $2p - 1$ μ vectors out of the universe of $2^p - 1$ μ vectors. We select two distinct vectors with exactly 1 zero and $p - 1$ infinities, two distinct vectors with exactly 2 zeros and $p - 2$ infinities, etc.. Of course, there is only one vector with p zeros, which is the reason why only $2p - 1$ vectors are considered, not $2p$.

For $p \in \{15, 20, 25, ..., 50\}$, the set Ψ^* is a set of correlation matrices that includes (i) 43 Toeplitz matrices Ω that are such that $\delta(\Omega)$ takes values in a grid between $-.99$ and $.99$ as above, and (ii) 250 randomly generated Toeplitz matrices Ω. (Toeplitz matrices are considered because this makes computation of the size-correction values feasible). The randomly generated Toeplitz matrices are the correlation matrices of moving-average (MA) processes of random order q and random MA parameters. We take $q = p + [\chi^2_1]$, where χ^2_1 is a chi-squared random variable with one degree of freedom and $[\cdot]$ denotes the integer part. Given q, Ω is the $p \times p$ correlation matrix of a stationary MA process $Y = (Y_1, ..., Y_p)'$, where $Y_i = \sum_{j=0}^{q} a_j \varepsilon_{i-j}$ and $\{\varepsilon_i : i = ..., -1, 0, ...\}$ are i.i.d. with mean zero and variance one. The MA parameters $\{a_j : j = 0, ..., p-1\}$ are i.i.d. with a mixture of uniform distributions. With probability .7, a_j has a uniform distribution with mean zero and variance one, and with probability .3, a_j has a uniform distribution with mean one and variance one. This distribution for a_j is chosen to yield a balanced distribution of $\delta(\Omega)$ values across the 250 Ω matrices.[7]

To reduce the effects of simulation error and to generate $\eta_2(p)$ values for $p = 11, ..., 14, 16, ..., 19$, etc., we smooth the simulated $\eta_2(p)$ values across p by fitting a regression model to the computed values for $p = 2, 3, ..., 10, 15, 20, ..., 50$. We take the $\eta_2(p)$ values to be the predicted values from this regression. We consider regression models with linear, quadratic, and cubic terms with and without the restriction that $\eta_2(p) = 0$ for $p = 2$ (which just amounts to using an intercept or not in a shifted version of the regression function). The results from the different models quite similar. The values in Table I of the paper are based on the quadratic model with the restriction that $\eta_2(p) = 0$ for $p = 2$. It has an R^2 of .992.

[7]We also compute values of $\eta_2(p)$ for $p = \{2, 3, ..., 10\}$ using 250 randomly generated Toeplitz matrices Ω in place of the 500 randomly generated matrices Ω described in the paragraph above (3.8) (which are not necessarily Toeplitz). The former are not noticeably different from the latter.

References

Andrews, D. W. K. and P. Jia (2008): "Inference for Parameters Defined by Moment Inequalities: A Recommended Moment Selection Procedure," unpublished manuscript, Cowles Foundation, Yale University.

Andrews, D. W. K. and P. Guggenberger (2009): "Validity of Subsampling and "Plug-in Asymptotic" Inference for Parameters Defined by Moment Inequalities," *Econometric Theory,* 25, forthcoming.

Andrews, D. W. K., M. J. Moreira, and J. H. Stock (2008): "Efficient Two-Sided Nonsimilar Invariant Tests in IV Regression with Weak Instruments," *Journal of Econometrics,* forthcoming.

Müller, U. and M. Watson (2008): "Low-Frequency Robust Cointegration Testing," *Econometrica,* 76, 979-1016.