Heavy-Tail and Plug-In Robust Consistent Conditional Moment Tests of Functional Form: Supplemental Appendix

Jonathan B. Hill[1]
Dept. of Economics University of North Carolina

Feb. 2012

In Section C.1 we validate Hansen's (1996) bootstrapped p-value method for test functionals. In Section C.2 we complete the proof of Theorem 2.2 for the slow plug-in case P2 under H_0. In Section C.3 we prove Lemmas B.1-B.7. Finally, in Section C.4 we present all simulation results. All citations are presented in a bibliography at the end.

Recall the equation and covariance constructions

$$\mathcal{M}^*_{n,t}(\beta,\gamma) := [m^*_{n,t}(\beta,\gamma), \tilde{m}'_{n,t}(\beta)]' \in \mathbb{R}^{r+1}$$

$$\tilde{S}_n(\beta) := \sum_{s,t=1}^{n} E\left[\{\tilde{m}_{n,s}(\beta) - E[\tilde{m}_{n,s}(\beta)]\} \times \{\tilde{m}_{n,t}(\beta) - E[\tilde{m}_{n,t}(\beta)]\}'\right]$$

$$\mathfrak{S}^*_n(\beta,\gamma) := \sum_{s,t=1}^{n} E\left[\left\{\mathcal{M}^*_{n,s}(\beta,\gamma) - E[\mathcal{M}^*_{n,s}(\beta,\gamma)]\right\} \times \left\{\mathcal{M}^*_{n,t}(\beta,\gamma) - E[\mathcal{M}^*_{n,t}(\beta,\gamma)]\right\}'\right]$$

and threshold and fractile representations

$$c_{z,n}(\cdot) = \max\{l_{z,n}(\cdot), u_{z,n}(\cdot)\} \quad \text{and} \quad k_{j,n} = \max\{k_{j,\epsilon,n}, k_{j,1,n}, ..., k_{j,q,n}\}.$$

Recall if $\{A_n(\gamma,)B_n(\gamma)\}_{n\geq 1}$ are sequences of functions of γ we write $A_n(\gamma) \sim B_n(\gamma)$ *uniformly on* Γ when $\sup_{\gamma\in\Gamma}|A_n(\gamma)/B_n(\gamma)| \to 1$, and when $\sup_{\gamma\in\Gamma}|A_n(\gamma)/B_n(\gamma)| \stackrel{p}{\to} 1$ we write $A_n(\gamma) \stackrel{p}{\sim} B_n(\gamma)$ *uniformly on* Γ.

For ease of reference we first present all assumptions. Recall *f.d.d.* denotes *finite dimensional distributions.*

F1 (*fractiles*). *a.* $k_{j,\epsilon,n}/\ln(n) \to \infty$; *b. if* $\kappa_\epsilon \in (0,1)$ *then* $k_{j,\epsilon,n}/n^{2(1-\kappa_\epsilon)/(2-\kappa_\epsilon)} \to \infty$.

F2 (*non-degenerate trimmed variance*). $\liminf_{n\to\infty}\inf_{\beta\in\mathcal{B},\gamma\in\Gamma}\{S^2_n(\beta,\gamma)/n\} > 0$ *and* $\sup_{\beta\in\mathcal{B},\gamma\in\Gamma}\{n\sigma^2_n(\beta,\gamma)/S^2_n(\beta,\gamma)\} = O(1)$.

I1 (*identification by* $m^*_{n,t}(\gamma)$). *Under the null* $\sup_{\gamma\in\Gamma}|nS^{-1}_n(\gamma)E[m^*_{n,t}(\gamma)]| \to 0$.

K1 (*kernel and bandwidth*). $\omega(\cdot)$ *is integrable, and a member of the class* $\{\omega : \mathbb{R} \to [-1,1] \mid \omega(0) = 1, \omega(x) = \omega(-x)\ \forall x \in \mathbb{R}, \int_{-\infty}^{\infty}|\omega(x)|dx < \infty, \int_{-\infty}^{\infty}|\vartheta(\xi)|d\xi < \infty, \omega(\cdot)$ *is continuous at* 0 *and all but a finite number of points*$\}$, *where* $\vartheta(\xi) := (2\pi)^{-1}\int_{-\infty}^{\infty}\omega(x)e^{i\xi x}dx < \infty$. *Further* $\sum_{s,t=1}^{n}|\omega((s-t)/b_n)| = o(n^2)$, $\max_{1\leq s\leq n}|\sum_{t=1}^{n}\omega((s-t)/b_n)| = o(n)$ *and* $b_n = o(n)$.

[1] Dept. of Economics, University of North Carolina-Chapel Hill, www.unc.edu/∼jbhill, jbhill@email.unc.edu.

P1 (***fast (non)linear plug-ins***). *$\tilde{V}_n^{1/2}(\hat{\beta}_n - \beta^0) = O_p(1)$ and $\sup_{\gamma\in\Gamma} ||V_n(\gamma)\tilde{V}_n^{-1}|| \to 0$.*

P2 (***slow linear plug-ins***). *$\mathfrak{S}_n^*(\gamma)$ exists for each n, specifically $\sup_{\gamma\in\Gamma} ||\mathfrak{S}_n^*(\gamma)|| < \infty$ and $\liminf_{n\to\infty} \inf_{\gamma\in\Gamma} \lambda_{\min}(\mathfrak{S}_n^*(\gamma)) > 0$. Further:*

a. $\tilde{V}_n^{1/2}(\hat{\beta}_n - \beta^0) = O_p(1)$ and $\tilde{V}_n \sim \mathcal{K}(\gamma)V_n(\gamma)$, $\mathcal{K} : \Gamma \in \mathbb{R}^{q\times q}$, $\inf_{\gamma\in\Gamma} \lambda_{\min}(\mathcal{K}(\gamma)) > 0$.

b. $\tilde{V}_n^{1/2}(\hat{\beta}_n - \beta^0) = \tilde{A}_n \sum_{t=1}^n \{\tilde{m}_{n,t} - E[\tilde{m}_{n,t}]\} \times (1 + o_p(1)) + o_p(1)$ where non-stochastic $\tilde{A}_n \in \mathbb{R}^{q\times r}$ has full column rank and $\tilde{A}_n\tilde{S}_n^{-1}\tilde{A}_n' \to I_q$.
c. The f.d.d. of $\mathfrak{S}_n^*(\gamma)^{-1/2}\{\mathcal{M}_{n,t}^*(\gamma) - E[\mathcal{M}_{n,t}^*(\gamma)]\}$ belong to the same domain of attraction as the f.d.d. of $S_n^{-1}(\gamma)\{m_{n,t}^*(\gamma) - E[m_{n,t}^*(\gamma)]\}$.

P3 (***orthogonal equations and (non)linear plug-ins***). *$\tilde{V}_n^{1/2}(\hat{\beta}_n - \beta^0) = O_p(1)$ and $\limsup_{n\to\infty} \sup_{\gamma\in\Gamma} ||V_n^{\perp}(\gamma)\tilde{V}_n^{-1}|| < \infty$.*

R1 (***response***). *$f(\cdot,\beta)$ is for each $\beta \in \mathcal{B}$ a Borel measurable function, continuous and differentiable on $\mathcal{B}$ with Borel measurable gradient $g_t(\beta) := (\partial/\partial\beta)f(x_t,\beta)$.*

R2 (***moments***). *$E|y_t| < \infty$, and $E(\sup_{\beta\in\mathcal{B}} |f(x_t,\beta)|^{\iota}) < \infty$ and $E(\sup_{\beta\in\mathcal{B}} |(\partial/\partial\beta_i)f(x_t,\beta)|^{\iota}) < \infty$ for each i and some tiny $\iota > 0$.*

R3 (***distribution***).

a. The finite dimensional distributions of $\{y_t, x_t\}$ are strictly stationary, non-degenerate and absolutely continuous. The density function of $\epsilon_t(\beta)$ is uniformly bounded $\sup_{\beta\in\mathcal{B}} \sup_{a\in\mathbb{R}}\{(\partial/\partial a)P(\epsilon_t(\beta) \leq a)\} < \infty$.

b. Define $\kappa_\epsilon(\beta) := \text{argsup}_{\alpha > 0}\{E|\epsilon_t(\beta)|^\alpha < \infty\} \in (0,\infty]$, write $\kappa_\epsilon = \kappa_\epsilon(\beta^0)$, and let $\mathcal{B}_{2,\epsilon}$ denote the set of β such that variance is infinite $\kappa_\epsilon(\beta) \leq 2$. If $\kappa_\epsilon(\beta) \leq 2$ then $P(|\epsilon_t(\beta)| > c) = d(\beta)\epsilon^{-\kappa_\epsilon(\beta)}(1 + o(1))$ where $\inf_{\beta\in\mathcal{B}_{2,\epsilon}} d(\beta) > 0$ and $o(1)$ is not a function of β, hence $\lim_{c\to\infty} \sup_{\beta\in\mathcal{B}_{2,\epsilon}}\{d(\beta)^{-1}\epsilon^{\kappa_\epsilon(\beta)}P(|\epsilon_t(\beta)| > c)\} = 1$.

R4 (***mixing***). *$\{y_t, x_t\}$ are geometrically β-mixing: $\sup_{\mathcal{A}\subset\Im_{t+l}^{+\infty}} E|P(\mathcal{A}|\Im_{-\infty}^t) - P(\mathcal{A})| = o(\rho^l)$ for $\rho \in (0,1)$.*

W1 (***weight***).

a. $F : \mathbb{R} \to \mathbb{R}$ is Borel measurable, analytic and non-polynomial on some open interval $R_0 \subseteq \mathbb{R}$ containing 0.

b. $\sup_{u\in U} |F(u)| \leq K$ and $\inf_{u\in U} |F(u)| > 0$ on any compact subset $U \subset S_F$, with S_F the support of F.

Write the kernel function compactly as

$$\omega_{n,s,t} := \omega\left((s-t)/b_n\right).$$

C.1 TEST FUNCTIONAL AND BOOTSTRAPPED P-VALUE We only treat $\hat{T}_n(\gamma)$ under plug-in cases P1 or P2 for brevity since $\hat{T}_n^{\perp}(\gamma)$ under P3 follows similarly. Define $\Im_t := \sigma(y_\tau, x_{\tau+1} : \tau \leq 1)$. Let $h : \mathbb{R} \to \mathbb{R}$ be continuous on *R-a.e.* By Theorem 2.2 under the null $h(\hat{T}_n(\gamma)) \Longrightarrow h(z^2(\gamma))$, and if h is strictly monotonically increasing then $h(\hat{T}_n(\gamma)) \xrightarrow{p} \infty$ under H_1. The standard transforms are $\sup_{\gamma\in\Gamma} \hat{T}_n(\gamma)$ and $\int_\Gamma \hat{T}_n(\gamma)\mu(d\gamma)$ for any measure μ on Γ, absolutely continuous with respect to Lebesgue measure (cf. Davies 1977). In non-trivial cases $h(z^2(\gamma))$ is non-standard, hence we must approximate the p-value. This is all well known, but under trimming a wild bootstrap as in Hansen

(1996) is problematic since it involves an iid simulator and a martingale difference error under the null hypothesis.

Recall the construction of Hansen's (1996) approximate p-value. Let $h_n := h(\hat{T}_n(\gamma))$ be the test functional with the limit $h_n \Rightarrow h^0$ under H_0 and distribution $F^0(x) := P(h^0 \leq x)$, and operate on $p_n := 1 - F^0(h_n)$ the "asymptotic p-value". Under the null $p_n \overset{d}{\to} U[0,1]$ a uniform distribution, and under the global alternative $p_n \overset{p}{\to} 0$ as long as h is monotonic increasing, cf. Theorem 2.2. Now operate conditionally on the sample $\{x_t, y_t\}_{t=1}^n$, simulate a sequence $\{\varepsilon_t\}_{t=1}^n$ of iid $N(0,1)$ random variables ε_t, compute

$$\hat{h}_n := h\left(\frac{1}{\hat{S}_n^2(\hat{\beta}_n, \gamma)}\left(\sum_{t=1}^n \varepsilon_t \hat{m}_{n,t}^*(\hat{\beta}_n, \gamma)\right)^2\right),$$

and define $\hat{p}_n := 1 - \hat{F}_n(h_n)$ where $\hat{F}_n$ is the conditional distribution of $\hat{h}_n$.

There is a rather sharp hurdle concerning the wild bootstrap under tail-trimming. Notice $\hat{S}_n^2(\hat{\beta}_n, \gamma)$ is the original HAC estimator and $\sup_{\gamma\in\Gamma} |\hat{S}_n^2(\hat{\beta}_n, \gamma)/S_n^2(\gamma) - 1| \overset{p}{\to} 0$ by Lemma B.5, while $\mathcal{Z}_n(\gamma) := \sum_{t=1}^n \varepsilon_t m_{n,t}^*(\gamma)$ conditional on $\{x_t, y_t\}_{t=1}^n$ is a draw from a normal distribution with covariance kernel proportional to $\mathcal{K}_n(\gamma, \tilde{\gamma}) \overset{p}{\sim} \sum_{t=1}^n m_{n,t}^*(\gamma) m_{n,t}^*(\tilde{\gamma})$, cf. Lemma B.2 in Appendix B. Under the null $S_n^2(\gamma)/\mathcal{K}_n(\gamma, \gamma) \overset{p}{\to} 1$ can be verified only if

$$E\left(\sum_{t=1}^n m_{n,t}^*(\gamma)\right)^2 = nE\left[m_{n,t}^{*2}(\gamma)\right] \times (1 + o(1)).$$

The latter is trivial if ϵ_t is iid and symmetric under H_0 and symmetrically trimmed, and holds if ϵ_t is a martingale difference with respect to $\Im_t$ and $\epsilon_t I_{\epsilon,n,t}$ becomes a martingale sufficiently fast in the sense

$$\left(\frac{n}{E\left[m_{n,t}^{*2}(\gamma)\right]}\right)^{1/2} E\left[\epsilon_t I_{\epsilon,n,t} | \Im_{t-1}\right] \to 0. \tag{1}$$

The martingale difference property $E[\epsilon_t|\Im_{t-1}] = 0$ ensures by dominated convergence $E[\epsilon_t I_{\epsilon,n,t}|\Im_{t-1}] \to 0$, but the stronger condition (1) implies for each $i = 1, ..., n-1$

$$\frac{n}{E\left[m_{n,t}^{*2}(\gamma)\right]} \times E\left[\epsilon_1 I_{\epsilon,n,1} \epsilon_{i+1} I_{\epsilon,n,i+1} \times I_{g,n,1} F(\gamma'\psi_1) I_{g,n,i+1} F(\gamma'\psi_{i+1})\right] \to 0,$$

hence

$$\begin{aligned} E\left(\sum_{t=1}^n m_{n,t}^*(\gamma)\right)^2 &= nE\left[m_{n,t}^{*2}(\gamma)\right]\left(1 + \frac{1}{n}\sum_{i=1}^{n-1}(1 - i/n)\frac{n \times E\left[m_{n,1}^*(\gamma) m_{n,i+1}^*(\gamma)\right]}{E\left[m_{n,t}^{*2}(\gamma)\right]}\right) \\ &= nE\left[m_{n,t}^{*2}(\gamma)\right] \times (1 + o(1)). \end{aligned}$$

In general, as long as $E[\epsilon_t] = 0$ under either hypothesis then ϵ_t need only be iid (and possibly asymmetric) under H_0 and symmetrically trimmed with re-centering since under H_0

$$E\left[m_{n,t}^*(\gamma)|\Im_{t-1}\right] = I_{g,n,t} F(\gamma'\psi_t)\left(E\left[\epsilon_t I_{\epsilon,n,t}\right] - E\left[\epsilon_t I_{\epsilon,n,t}\right]\right) = 0.$$

See Section 3 in the main paper.

LEMMA C.1. *Let F1-F2, I1, K1, P1 or P2, R1-R4 and W1.b hold. Further, assume* $E(\sum_{t=1}^n m_{n,t}^*(\gamma))^2 = \sum_{t=1}^n E[m_{n,t}^{*2}(\gamma)] \times (1 + o(1))$. *Then* $|\hat{p}_n - p_n| \overset{p}{\to} 0$.

PROOF. Hansen (1996) requires β-mixing with at least a finite fourth moment, and a Lipschitz test weight all in order to prove his test statistic has a weak limit. We establish the required limit in the proof of Theorem 2.2 by exploiting weak limit theory developed in Section C.3, below, in particular UCLT Lemma B.7. All aspects of Hansen's (1996: p. 426-427) argument therefore goes through since it relies solely on the pre-established weak limit. $\mathcal{QED}$.

Now simulate R iid samples $\{\varepsilon_{r,t}\}_{t=1}^{n}$ of iid $N(0,1)$ random variables $\varepsilon_{r,t}$, and compute $\hat{h}_{r,n}$ using $\{\varepsilon_{r,t}\}_{t=1}^{n}$, and $\hat{p}_{n}^{R} := 1/R\sum_{r=1}^{R} I(\hat{h}_{r,n} > h_{n})$. Since $\hat{h}_{r,n}$ is, conditionally on z, iid and independent of h_n it follows $\lim_{R\rightarrow\infty}\hat{p}_{n}^{R} = 1 - \hat{F}_{n}(h_{n}) = \hat{p}_{n}$ by the Glivenko-Cantelli theorem. Hence $\hat{p}_{n}^{R}$ is a valid approximation of p_n for large R, provided $E(\sum_{t=1}^{n} m_{n,t}^{*}(\gamma))^{2} \sim \sum_{t=1}^{n} E[m_{n,t}^{*2}(\gamma)]$. In practice set $R = R_n \rightarrow \infty$ as $n \rightarrow \infty$.

C.2 PROOF OF THEOREM 2.2 UNDER PLUG-IN ASSUMPTION P2 Recall the claim.

THEOREM 2.2. *Let F1-F2, I1, K1, R1-R4 and W1 hold.*

i. Under H_0 and P1 or P2 there exists a Gaussian process $\{z(\gamma) : \gamma \in \Gamma\}$ on $\mathcal{C}[\Gamma]$ with zero mean, unit variance and covariance function $E[z(\gamma_1)z(\gamma_2)]$ such that $\{\hat{T}_n(\gamma) : \gamma \in \Gamma\} \Longrightarrow \{z(\gamma)^2 : \gamma \in \Gamma\}$.

ii. Under H_1 and P1 or P2 $\hat{T}_n(\gamma) \stackrel{p}{\rightarrow} \infty$ $\forall\gamma \in \Gamma/S$ where S has Lebesgue measure zero.

iii. Under P3 $\hat{T}_n^{\perp}(\gamma)$ satisfies cases (i) and (ii).

PROOF. It remains to prove claim (i) under slow plug-in Assumption P2. Define

$$M_{n,t}^{*}(\beta,\gamma) := m_{n,t}^{*}(\beta,\gamma) - E\left[m_{n,t}^{*}(\beta,\gamma)\right] \text{ and } \hat{M}_{n,t}^{*}(\beta,\gamma) := \hat{m}_{n,t}^{*}(\beta,\gamma) - E\left[\hat{m}_{n,t}^{*}(\beta,\gamma)\right]$$

$$\tilde{M}_{n,t}^{*} := \tilde{m}_{n,t}^{*}(\beta^{0}) - E\left[\tilde{m}_{n,t}^{*}(\beta^{0})\right] \text{ and } \check{S}_{n}^{2}(\beta,\gamma) = \sum_{s,t=1}^{n} E\left[\hat{M}_{n,s}^{*}(\beta,\gamma)\hat{M}_{n,t}^{*}(\beta,\gamma)\right]$$

We first state some required properties. Identification I1 imposes

$$\sup_{\gamma\in\Gamma}\left|S_{n}^{-1}(\gamma)E[m_{n,t}^{*}(\gamma)]\right| = o(1/n), \tag{2}$$

which implies the following long-run covariance relation uniformly on Γ:

$$E\left(\sum_{t=1}^{n} M_{n,t}^{*}(\gamma)\right)^{2} = S_{n}^{2}(\gamma) - n^{2}\left(E\left[m_{n,t}^{*}(\beta,\gamma)\right]\right)^{2} = S_{n}^{2}(\gamma)\left(1 + o\left(1\right)\right). \tag{3}$$

Uniform expansion Lemma B.3.a, coupled with Jacobian consistency Lemma B.4 and $\hat{\beta}_n \stackrel{p}{\rightarrow} \beta^0$ imply for any finite $\delta > 0$

$$\sup_{\gamma\in\Gamma}\left|\frac{1}{n}\sum_{t=1}^{n}\left\{\hat{m}_{n,t}^{*}(\hat{\beta}_{n},\gamma) - \hat{m}_{n,t}^{*}(\gamma)\right\} - J_{n}\left(\gamma\right)'\left(\hat{\beta}_{n} - \beta^{0}\right)\left(1 + o_{p}\left(1\right)\right)\right| = o_{p}\left(n^{-\delta}\right). \tag{4}$$

By approximation Lemma B.2.a

$$\sup_{\gamma\in\Gamma}\left|\frac{1}{S_{n}(\gamma)}\sum_{t=1}^{n}\left\{\hat{m}_{n,t}^{*}(\gamma) - m_{n,t}^{*}\left(\gamma\right)\right\}\right| = o_{p}\left(1\right). \tag{5}$$

By Lemma B.5 uniform HAC consistency

$$\sup_{\gamma\in\Gamma}\left|\frac{\hat{S}_n^2(\hat{\beta}_n,\gamma)}{S_n^2(\gamma)}-1\right|=o_p(1). \tag{6}$$

It can similarly be shown

$$\sup_{\gamma\in\Gamma}\left|\frac{\hat{S}_n^2(\hat{\beta}_n,\gamma)}{\check{S}_n^2(\hat{\beta}_n,\gamma)}-1\right|=o_p(1) \text{ where } \check{S}_n^2(\beta,\gamma):=E\left(\sum_{t=1}^n \hat{M}_{n,t}^*(\beta,\gamma)\right)^2. \tag{7}$$

Under P2 $\tilde{V}_n\sim\mathcal{K}(\gamma)V_n(\gamma)$ for some $\mathcal{K}:\Gamma\in\mathbb{R}^{q\times q}$ uniformly positive definite on Γ, and the plug-in is asymptotically linear

$$\hat{\beta}_n-\beta^0=\tilde{V}_n^{-1/2}\tilde{A}_n\sum_{t=1}^n\tilde{M}_{n,t}^*\times(1+o_p(1)) \;+\; o_p\left(\left\|\tilde{V}_n\right\|^{-1/2}\right) \tag{8}$$

for some non-stochastic sequence $\{\tilde{A}_n\}$, $\tilde{A}_n\in\mathbb{R}^{q\times r}$ that satisfies $\tilde{A}_n\tilde{S}_n^{-1}\tilde{A}_n'\to I_q$.

Substitute $\hat{\beta}_n-\beta^0$ from (8) into (4) and exploit the above uniform limits to obtain uniformly on Γ

$$\begin{aligned}\sum_{t=1}^n\hat{m}_{n,t}^*(\hat{\beta}_n,\gamma) &= \sum_{t=1}^n\hat{m}_{n,t}^*(\gamma)+nJ_n(\gamma)\left(\hat{\beta}_n-\beta^0\right)(1+o_p(1))+o_p\left(n^{-\delta}\right)\\ &= \sum_{t=1}^n m_{n,t}^*(\gamma)+nJ_n(\gamma)\tilde{V}_n^{-1/2}\tilde{A}_n\sum_{t=1}^n\tilde{M}_{n,t}^*\times(1+o_p(1))+o_p(S_n(\gamma))\\ &= \sum_{t=1}^n M_{n,t}^*(\gamma)+S_n(\gamma)\left\{nS_n^{-1}(\gamma)J_n(\gamma)\right\}\tilde{V}_n^{-1/2}\times\tilde{A}_n\sum_{t=1}^n\tilde{M}_{n,t}^*\times(1+o_p(1))+o_p(S_n(\gamma))\\ &= \sum_{t=1}^n M_{n,t}^*(\gamma)+\tilde{\mathcal{B}}_n(\gamma)\sum_{t=1}^n\tilde{M}_{n,t}^*\times(1+o_p(1))+o_p(S_n(\gamma)),\end{aligned}$$

say, where $\tilde{\mathcal{B}}_n(\gamma)\in\mathbb{R}^{1\times r}$. The second equality uses the facts that arbitrary $\delta>0$ and $\liminf_{n\to\infty}\{S_n^2(\gamma)\} > 0$ under F2 together imply $o_p\left(n^{-\delta}\right)=o_p(S_n(\gamma))$.

Now, define the selection matrix $\mathcal{R}_n(\gamma)=[1,\tilde{\mathcal{B}}_n(\gamma)]\in\mathbb{R}^{1\times(r+1)}$, and use $\mathcal{M}_{n,t}^*(\gamma)=[m_{n,t}^*(\gamma),\tilde{m}_{n,t}'(\gamma)]'$ to write

$$\sum_{t=1}^n\hat{m}_{n,t}^*(\hat{\beta}_n,\gamma)=\sum_{t=1}^n\mathcal{R}_n(\gamma)\left\{\mathcal{M}_{n,t}^*(\gamma)-E\left[\mathcal{M}_{n,t}^*(\gamma)\right]\right\}(1+o_p(1))+o_p(S_n(\gamma)). \tag{9}$$

Recall $\mathfrak{S}_n^*(\gamma)\in\mathbb{R}^{(r+1)\times(r+1)}$ is the covariance matrix for $\sum_{t=1}^n\mathcal{M}_{n,t}^*(\gamma)$, and define

$$\mathbb{S}_n^2(\gamma):=\mathcal{R}_n(\gamma)\mathfrak{S}_n^*(\gamma)\mathcal{R}_n'(\gamma)\in\mathbb{R}.$$

By construction $\sup_{\gamma\in\Gamma}\{S(\gamma)/\mathbb{S}_n(\gamma)\}=O(1)$. Therefore by (9) and UCLT Lemma B.7

$$\left\{\frac{1}{\mathbb{S}_n(\gamma)}\sum_{t=1}^n\hat{m}_{n,t}^*(\hat{\beta}_n,\gamma):\gamma\in\Gamma\right\}\Longrightarrow\{z(\gamma):\gamma\in\Gamma\}, \tag{10}$$

where $\{z(\gamma):\gamma\in\Gamma\}$ is a Gaussian process on C[Γ] with zero mean and unit variance, and covariance function $E[z(\gamma_1)z(\gamma_2)]$.

Equation (10) implies $E(\mathbb{S}_n^{-1}(\gamma)\sum_{t=1}^n \hat{m}^*_{n,t}(\hat{\beta}_n,\gamma)) \to 0$ by the Helly-Bray theorem, hence

$$\left\{\frac{1}{\mathbb{S}_n(\gamma)}\sum_{t=1}^n \left\{\hat{m}^*_{n,t}(\hat{\beta}_n,\gamma) - E\left[\hat{m}^*_{n,t}(\hat{\beta}_n,\gamma)\right]\right\} : \gamma \in \Gamma\right\} \tag{11}$$

$$= \left\{\frac{1}{\mathbb{S}_n(\gamma)}\sum_{t=1}^n \mathcal{R}_n(\gamma)\left\{\mathcal{M}^*_{n,t}(\gamma) - E\left[\mathcal{M}^*_{n,t}(\gamma)\right]\right\}(1+o_p(1)) + o_p(1) : \gamma \in \Gamma\right\}$$

$$\Longrightarrow \{z(\gamma) : \gamma \in \Gamma\}.$$

Therefore by HAC consistency (6) and (7) and the definition of $\check{S}_n(\hat{\beta}_n,\gamma)$ in (7)

$$\mathbb{S}_n(\gamma) = \check{S}_n(\hat{\beta}_n,\gamma) \times (1+o_p(1)) = \hat{S}_n(\hat{\beta}_n,\gamma) \times (1+o_p(1)). \tag{12}$$

Now combine (11) and (12) to deduce

$$\left\{\frac{1}{\hat{S}_n(\hat{\beta}_n,\gamma)}\sum_{t=1}^n \hat{m}^*_{n,t}(\hat{\beta}_n,\gamma) : \gamma \in \Gamma\right\}$$

$$= \left\{\frac{1}{\mathbb{S}_n(\gamma)}\sum_{t=1}^n \mathcal{R}_n(\gamma)\left\{\mathcal{M}^*_{n,t}(\gamma) - E\left[\mathcal{M}^*_{n,t}(\gamma)\right]\right\}(1+o_p(1)) + o_p(1) : \gamma \in \Gamma\right\}$$

$$\Longrightarrow \{z(\gamma) : \gamma \in \Gamma\}.$$

The claim now follows from the mapping theorem: $\{\hat{T}_n(\gamma) : \gamma \in \Gamma\} \Longrightarrow \{z^2(\gamma) : \gamma \in \Gamma\}$. $\mathcal{QED}$.

C.3 PROOFS OF LEMMAS B.1-B.7 We reduce notation by assuming symmetric trimming throughout since none of the following arguments rely on asymmetry. Let $\{c_{z,n}(\beta)\}$ be the sequence of two-tailed upper $k_{z,n}/n^{th}$ quantiles of $z_t(\beta)$:

$$P(|z_t(\beta)| > c_{z,n}(\beta)) = \frac{k_{z,n}}{n}. \tag{13}$$

Recall the absolute value notation:

$$z_t^{(a)}(\beta) := |z_t(\beta)|.$$

Write compactly

$$\sigma_n^2(\beta,\gamma) := E\left[m^{*2}_{n,t}(\beta,\gamma)\right] \quad \text{and} \quad S_n^2(\gamma) := E\left(\sum_{t=1}^n \left\{m^*_{n,t}(\gamma) - E[m^*_{n,t}(\gamma)]\right\}\right)^2,$$

and

$$\hat{m}^*_n(\beta,\gamma) := \frac{1}{n}\sum_{t=1}^n \hat{m}^*_{n,t}(\beta,\gamma) \;\text{ and }\; m^*_n(\beta,\gamma) := \frac{1}{n}\sum_{t=1}^n m^*_{n,t}(\beta,\gamma).$$

We repeatedly use the fact that $m_t = \epsilon_t F(\gamma'\psi_t)$ has the same decay rate as ϵ_t under Paretian tail R3.b because $F(u)$ is bounded by W1.b. This follows by properties of regularly varying tails: for some finite positive $K_u > K_l$

$$K_u c^{-\kappa_\epsilon}(1+o(1)) = P\left(\left|\epsilon_t \sup_{u\in\mathbb{R}} F(u)\right| > c\right) \geq P(|m_t| > c) \geq P\left(\left|\epsilon_t \inf_{u\in\mathbb{R}} F(u)\right| > c\right) = K_l c^{-\kappa_\epsilon}(1+o(1))$$

hence $P(|m_t| > c) = Kc^{-\kappa_\epsilon}(1 + o(1))$. Further, since $I_{n,t} = I_{\epsilon,n,t}I_{g,n,t}$ use dominated convergence and $I_{g,n,t} \overset{a.s.}{\to} 1$ to deduce for any $r > 0$

$$E\left[|m_t|^r I_{n,t}\right] = E\left[|m_t|^r I_{\epsilon,n,t}\right] \times \left(1 + \frac{E\left[|m_t|^r I_{\epsilon,n,t}\left(1 - I_{g,n,t}\right)\right]}{E\left[|m_t|^r I_{\epsilon,n,t}\right]}\right) = E\left[|m_t|^r I_{\epsilon,n,t}\right] \times \left(1 + o\left(1\right)\right).$$

Together, higher moments of $m_tI_{n,t}$ and $\epsilon_tI_{\epsilon,n,t}$ are identical up to a scale constant.

Impose F1-F2, K1, R1-R4, and W1.b throughout. Many results follow instantly from asymptotic theory and moment bounds developed in Hill (2011) for M-estimation under tail-trimming. We only provide proofs where significant changes are required.

In the following the sequence of positive non-random real numbers $\{r_n\}$, and terms $o_p(1)$, $O_p(1)$, $o(1)$ and $O(1)$, do not depend on β, γ and t, where $r_n \to 0$ is arbitrarily fast.[2] Let Λ be a compact subset of $(0, 1]$.

LEMMA B.1 (variance bounds).

a. $\sigma_n^2(\beta,\gamma) = o\left(n \max\left\{1, (E[m_{n,t}^*(\beta,\gamma)])^2\right\}\right)$, $\sup_{\gamma\in\Gamma}\left\{\frac{\sigma_n^2(\gamma)}{\max\left\{1, (E[m_{n,t}^*(\gamma)])^2\right\}}\right\} = o(n/\ln(n))$;

b. $S_n^2(\gamma) = \mathfrak{L}_n n\sigma_n^2(\gamma) = o(n^2)$ *for some sequence* $\{\mathfrak{L}_n\}$ *that satisfies* $\liminf_{n\to\infty}\mathfrak{L}_n > 0$, $\mathfrak{L}_n = K$ *if* ϵ_t *is finite dependent or* $E[\epsilon_t^2] < \infty$, *and otherwise* $\mathfrak{L}_n \leq K\ln(n/\min_{j\in\{1,2\}}\{k_{j,\epsilon,n}\}) \leq K\ln(n)$.

PROOF.

Claim (a): By the proof of Hill's (2011) Lemma B.2.a we have the first part of (a) $\sigma_n^2(\beta,\gamma) = o(n\max\{1, (E[m_{n,t}^*(\beta,\gamma)])^2\})$ hence

$$\sup_{\gamma\in\Gamma}\left\{\frac{\sigma_n^2(\gamma)}{\max\left\{1, (E[m_{n,t}^*(\gamma)])^2\right\}}\right\} = O(n/k_{\epsilon,n}).$$

Under fractile bound Assumption F1.a $n/k_{\epsilon,n} = o(n/\ln(n))$ hence the second part of (a) follows.

Claim (b): The following argument is used in Hill (2011: Lemma B.2.b) for the case $\kappa_\epsilon > 1$. The case $\kappa_\epsilon = 1$ is identical and $\kappa_\epsilon \in (0, 1]$ follows by invoking fractile bound F1.b. We present the proof below for completeness. Assume symmetric trimming to reduce notation, and asymmetric case being nearly identical.

If $E[\epsilon_t^2] < \infty$ the claim follows from geometric β-mixing. Under tail decay R3.b assume $\kappa_\epsilon < 2$, the case $\kappa_\epsilon = 2$ being similar. Since

$$S_n^2(\gamma) = nE\left[m_{n,t}^{*2}(\gamma)\right] + n\sum_{h=1}^{n-1}(1 - h/n)E\left[m_{n,1}^*(\gamma)m_{n,h+1}^*(\gamma)\right]$$

it suffices to prove the bounds

$$\left|\sum_{h=1}^{n-1} E\left[m_{n,1}^*(\gamma)m_{n,h+1}^*(\gamma)\right]\right| \leq \mathfrak{L}_n E[m_{n,t}^{*2}(\gamma)] = o(n).$$

Drop γ: $m_{n,t}^* = m_{n,t}^*(\gamma)$. Define the quantile function $Q_n(u) = \inf\{m : P(|m_{n,t}^*| > m) \leq u\}$ for $u \in [0, 1]$, and recall under geometric β-mixing $\alpha_h \leq \beta_h \leq K\rho^h$ for $\rho \in (0, 1)$, where α_h and β_h are α- and β-mixing coefficients. Assume $\beta_h = \rho^h$ without loss of generality. Theorem 1.1 of Rio (1993) applies:

$$\sum_{h=1}^{n-1} E\left|[m_{n,1}^* m_{n,h+1}^*]\right| \leq K\sum_{h=1}^{n-1}\int_0^{\rho^h} Q_n^2(u)du.$$

[2] For example $r_n = n^{-\delta}$ for any finite $\delta > 0$.

Under Paretian tail R3.b $Q_n(u) = O(u^{-1/\kappa_\epsilon})$, and by tail-trimming $\sup_{u\in[0,1]} Q_n(u) = c_{\epsilon,n}$ or $u \geq k_{\epsilon,n}/n$. Therefore

$$\sum_{h=1}^{n-1} \left|E[m^*_{n,1} m^*_{n,h+1}]\right| \leq K \sum_{h=1}^{n-1} \int_{k_{\epsilon,n}n}^{\rho^h} u^{-2/\kappa_\epsilon} du = K \sum_{h=1}^{\ln(n/k_{\epsilon,n})} \left\{(n/k_{\epsilon,n})^{(2/\kappa_\epsilon - 1)} - \rho^{-h(2/\kappa_\epsilon-1)}\right\}$$

which is trivially bounded by $K \ln(n/k_{\epsilon,n}) \times (n/k_{\epsilon,n})^{2/\kappa_\epsilon - 1} =: \mathcal{N}_n$.

Put $\mathfrak{L}_n = \ln(n/k_{\epsilon,n})$ and note $E[m^{*2}_{n,t}(\gamma)] \sim K\,(n/k_{\epsilon,n})^{2/\kappa_\epsilon-1}$ by Karamata's theorem and boundedness of $F(u)$. This proves the first bound $|\sum_{h=1}^{n-1} E[m^*_{n,1}(\gamma) m^*_{n,h+1}(\gamma)]| \leq \mathfrak{L}_n E[m^{*2}_{n,t}(\gamma)]$.

Finally, $(n/k_{\epsilon,n})^{2/\kappa_\epsilon-1} = o(n)$ is trivial for any $\kappa_\epsilon \in [1,2)$. If $\kappa_\epsilon \in (0,1)$, which occurs only under H_1, then $(n/k_{\epsilon,n})^{2/\kappa_\epsilon-1} = o(n)$ under fractile bound F1.b $k_{\epsilon,n}/n^{2(1-\kappa_\epsilon)/(2-\kappa_\epsilon)} \to \infty$. This proves the second bound. $\mathcal{QED}$.

LEMMA B.2 (approximations).

a. $\sup_{\gamma\in\Gamma} |S_n^{-1}(\gamma) \sum_{t=1}^n \{\hat{m}^*_{n,t}(\gamma) - m^*_{n,t}(\gamma)\}| = o_p(1)$.

b. Define $\hat{\mu}^*_{n,t}(\beta,\gamma) := \hat{m}^*_{n,t}(\beta,\gamma) - \hat{m}^*_n(\beta,\gamma)$ *and* $\mu^*_{n,t}(\beta,\gamma) := m^*_{n,t}(\beta,\gamma) - m^*_n(\beta,\gamma)$. *If additionally P1 or P2 holds then* $\sup_{\gamma\in\Gamma} |S_n^{-2}(\gamma) \sum_{s,t=1}^n \omega((s-t)/b_n)\{\hat{\mu}^*_{n,s}(\hat{\beta}_n,\gamma)\hat{\mu}^*_{n,t}(\hat{\beta}_n,\gamma) - \mu^*_{n,s}(\gamma)\mu^*_{n,t}(\gamma)\}| = o_p(1)$.

PROOF.

Claim (a): Use Lemma B.3 in Hill (2011) to deduce $S_n^{-1}(\gamma)\sum_{t=1}^n \{\hat{m}^*_{n,t}(\gamma) - m^*_{n,t}(\gamma)\} = o_p(1)$ for any $\gamma \in \Gamma$. By an almost identical argument the uniform claim $\sup_{\gamma\in\Gamma} |S_n^{-1}(\gamma)\sum_{t=1}^n \{\hat{m}^*_{n,t}(\gamma) - m^*_{n,t}(\gamma)\}| = o_p(1)$ can be shown by exploiting test weight lower and upper bounds W1.b and the fact that $F(u)$ does not enter the trimming indicators. Similar arguments are used to prove (b), below.

Claim (b): We will prove the notationally simpler claim under H_0

$$\sup_{\gamma\in\Gamma}\left|\frac{1}{S_n^2(\gamma)}\sum_{s,t=1}^n \omega_{n,s,t}\left\{\hat{m}^*_{n,s}(\hat{\beta}_n,\gamma)\hat{m}^*_{n,t}(\hat{\beta}_n,\gamma) - m^*_{n,s}(\gamma)m^*_{n,t}(\gamma)\right\}\right| = o_p(1).$$

The general claim under H_0 or H_1 is essentially identical but lengthy due to centering with $\hat{m}^*_n(\beta,\gamma)$. It suffices to show

$$\sup_{\gamma\in\Gamma}\left|\frac{1}{S_n^2(\gamma)}\sum_{s,t=1}^n \omega_{n,s,t}\left\{\hat{m}^*_{n,s}(\gamma)\hat{m}^*_{n,t}(\gamma) - m^*_{n,s}(\gamma)m^*_{n,t}(\gamma)\right\}\right| = o_p(1) \tag{14}$$

$$\sup_{\gamma\in\Gamma}\left|\frac{1}{S_n^2(\gamma)}\sum_{s,t=1}^n \omega_{n,s,t}\left\{\hat{m}^*_{n,s}(\hat{\beta}_n,\gamma)\hat{m}^*_{n,t}(\hat{\beta}_n,\gamma) - \hat{m}^*_{n,s}(\gamma)\hat{m}^*_{n,t}(\gamma)\right\}\right| = o_p(1). \tag{15}$$

Step 1 (bound (14)): Observe

$$\begin{aligned}
&\left|\frac{1}{S_n^2(\gamma)}\sum_{s,t=1}^n \omega_{n,s,t}\left\{\hat{m}^*_{n,s}(\gamma)\hat{m}^*_{n,t}(\gamma) - m^*_{n,s}(\gamma)m^*_{n,t}(\gamma)\right\}\right| \\
&\quad\leq K\left|\frac{1}{S_n^2(\gamma)}\sum_{s,t=1}^n \omega_{n,s,t} m_s(\gamma)\left(\hat{I}_{n,s} - I_{n,s}\right)m^*_{n,t}(\gamma)\right| \\
&\qquad + K\left|\frac{1}{S_n^2(\gamma)}\sum_{s,t=1}^n \omega_{n,s,t} m_s(\gamma)\left(\hat{I}_{n,s} - I_{n,s}\right)m^*_{n,t}(\gamma)\left(\hat{I}_{n,t} - I_{n,t}\right)\right| \\
&\quad= \mathcal{A}_{1,n}(\gamma) + \mathcal{A}_{2,n}(\gamma).
\end{aligned}$$

We only bound $\mathcal{A}_{1,n}(\gamma)$, $\mathcal{A}_{2,n}(\gamma)$ being similar. Define for any $\delta > 0$

$$\eta_\delta(x) := \frac{1}{(2\delta^2\pi)^{1/2}} \exp\left\{-x^2\delta^{-2}/2\right\} \quad \text{and} \quad \eta_{\delta,n,j} := \eta_\delta(j/b_n)$$

$$\mathcal{A}_{1,n,\delta}(\gamma) := \sum_{t=-n+1}^{2n} \left(\frac{1}{b_n^{1/2}} \sum_{l=1-t}^{n-t} k(l/b_n) \frac{1}{S_n(\gamma)} m_{t+l}(\gamma) \left(\hat{I}_{n,t+l} - I_{n,t+l}\right) I\left(0 \le l \le [b_n/\delta]\right) \right)$$
$$\times \left(\frac{1}{b_n^{1/2}} \sum_{j=1-t}^{n-t} \eta_{\delta,n,j} \frac{1}{S_n(\gamma)} m^*_{n,t+j}(\gamma) I\left(0 \le j \le [b_n/\delta]\right) \right) \times (1 + o_p(1)).$$

Trivially

$$\sup_{\gamma\in\Gamma} E\left(\frac{1}{S_n(\gamma)} \sum_{t=1}^n \left\{ m^*_{n,t}(\gamma) - E\left[m^*_{n,t}(\gamma)\right] \right\} \right)^2 = 1 = \sum_{t=1}^n \left(1/n^{1/2}\right)^2, \tag{16}$$

and by uniform approximation Claim (a) and UCLT Lemma B.7, for any chosen $K > 0$ there is an $N \in \mathbb{N}$ such that $\forall n \ge N$

$$\sup_{\gamma\in\Gamma} E\left(\frac{1}{S_n(\gamma)} \sum_{t=1}^n m^*_{n,t}(\gamma) \left(\hat{I}_{n,t} - I_{n,t}\right) \right)^2 = o(1) \le K \sum_{t=1}^n \left(1/n^{1/2}\right)^2 = K. \tag{17}$$

Therefore, by extending Lemmas A.2-A.3 of de Jong and Davidson (2000) to the uniform case we have[3]

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \sup_{\gamma\in\Gamma} \left\| \mathcal{A}_{1,n}(\gamma) - \mathcal{A}_{1,n,\delta}(\gamma) \times (1 + o_p(1)) \right\|_1 = 0. \tag{18}$$

Now consider $\mathcal{A}_{1,n,\delta}(\gamma)$, define $N_n(\delta) := \min\{n, [b_n/\delta] + 1\}$ and note by construction and variance non-degeneracy F2

$$\limsup_{n\to\infty} \frac{N_n(\delta)}{b_n} \le K \text{ and } \sup_{\gamma\in\Gamma} \left\{ \frac{S^2_{N_n(\delta)}(\gamma)/N_n(\delta)}{S^2_n(\gamma)/n} \right\} = O(1).$$

Uniform approximation Claim (a) and UCLT Lemma B.7 generalize to weighted versions under K1: for any δ

$$\frac{n^{1/2}}{b_n^{1/2}} \times \max_{-n+1\le t\le 2n} \sup_{\gamma\in\Gamma} \left\| \frac{1}{S_n(\gamma)} \sum_{l=1-t}^{n-t} k(l/b_n)\, m_{t+l}(\gamma) \left(\hat{I}_{n,t+l} - I_{n,t+l}\right) I\left(0 \le l \le [b_n/\delta]\right) \right\|_2$$
$$\le \frac{N_n^{1/2}(\delta)}{b_n^{1/2}} \sup_{\gamma\in\Gamma} \left\{ \frac{S_{N_n(\delta)}(\gamma)/N_n^{1/2}(\delta)}{S_n(\gamma)/n^{1/2}} \right\} \times \sup_{\gamma\in\Gamma} \left\| \frac{1}{S_{N_n(\delta)}(\gamma)} \sum_{t=1}^{N_n(\delta)} k(t/b_n) \left\{ \hat{m}^*_{n,t+l}(\gamma) - m^*_{n,t+l}(\gamma) \right\} \right\|_2$$
$$\to 0 \text{ as } n \to \infty,$$

[3] de Jong and Davidson (2000: Lemma A.1) invoke a mixingale maximal inequality due to McLeish (1975: Theorem 1.6) solely to ensure partial sum variance bounds. It suffices to replace their Lemma A.1 with (16) and (17) since these duplicate the same bound implied by McLeish (1975) with mixingale constants $1/n^{1/2}$.

and

$$\frac{n^{1/2}}{b_n^{1/2}} \max_{-n+1\leq t\leq 2n} \sup_{\gamma\in\Gamma} \left\| \sum_{j=1-t}^{n-t} \eta_{\delta,n,j} \frac{1}{S_n(\gamma)} m^*_{n,t+j}(\gamma) I\left(0 \leq j \leq [b_n/\delta]\right) \right\|_2$$

$$\leq \frac{N_n^{1/2}(\delta)}{b_n^{1/2}} \sup_{\gamma\in\Gamma} \left\{ \frac{S_{N_n(\delta)}(\gamma)/N_n^{1/2}(\delta)}{S_n(\gamma)/n^{1/2}} \right\} \times \sup_{\gamma\in\Gamma} \left\| \frac{1}{S_{N_n(\delta)}(\gamma)} \sum_{t=1}^{N_n(\delta)} \eta_{\delta,n,j} m^*_{n,t}(\gamma) \left(\hat{I}_{n,t} - I_{n,t}\right) \right\|_2$$

$$\rightarrow 0 \text{ as } n \rightarrow \infty.$$

Therefore

$$\lim_{\delta\rightarrow 0} \limsup_{n\rightarrow\infty} \sup_{\gamma\in\Gamma} \left\| \mathcal{A}_{1,n,\delta}(\gamma) \right\|_1 = 0. \tag{19}$$

Together (18) and (19) imply $\sup_{\gamma\in\Gamma}\{\mathcal{A}_{1,n}(\gamma)\} = o_p(1)$.

Step 2 (bound (15)): Note

$$\left| \frac{1}{S_n^2(\gamma)} \sum_{s,t=1}^{n} \omega_{n,s,t} \left\{ \hat{m}^*_{n,s}(\hat{\beta}_n,\gamma)\hat{m}^*_{n,t}(\hat{\beta}_n,\gamma) - \hat{m}^*_{n,s}(\gamma)\hat{m}^*_{n,t}(\gamma) \right\} \right|$$

$$\leq 2 \left| \frac{1}{S_n^2(\gamma)} \sum_{s,t=1}^{n} \omega_{n,s,t} \left\{ \hat{m}^*_{n,s}(\hat{\beta}_n,\gamma) - \hat{m}^*_{n,s}(\gamma) \right\} \hat{m}^*_{n,t}(\gamma) \right|$$

$$+ \left| \frac{1}{S_n^2(\gamma)} \sum_{s,t=1}^{n} \omega_{n,s,t} \left\{ \hat{m}^*_{n,s}(\hat{\beta}_n,\gamma) - \hat{m}^*_{n,s}(\gamma) \right\} \left\{ \hat{m}^*_{n,t}(\hat{\beta}_n,\gamma) - \hat{m}^*_{n,t}(\gamma) \right\} \right|.$$

Consider the first term, the second being similar. Apply a Taylor expansion to deduce for some $||\beta_{n,*} - \beta^0|| \leq ||\hat{\beta}_n - \beta^0||$,

$$\sup_{\gamma\in\Gamma} \left| \frac{1}{S_n^2(\gamma)} \sum_{s,t=1}^{n} \omega_{n,s,t} \left\{ \hat{m}^*_{n,s}(\hat{\beta}_n,\gamma) - \hat{m}^*_{n,s}(\gamma) \right\} \hat{m}^*_{n,t}(\gamma) \right|$$

$$\leq \sup_{\gamma\in\Gamma} \left\| \frac{1}{S_n^2(\gamma)} \sum_{s,t=1}^{n} \omega_{n,s,t} \hat{J}_{n,s}(\beta_{n,*},\gamma) \hat{m}^*_{n,t}(\gamma) \right\| \times \left\| \hat{\beta}_n - \beta^0 \right\|$$

$$+ \sup_{\gamma\in\Gamma} \left\| \frac{1}{S_n^2(\gamma)} \sum_{s,t=1}^{n} \omega_{n,s,t} J_s(\beta_{n,*},\gamma)) \left\{ \hat{I}_{n,s}(\beta_{n,*}) - \hat{I}_{n,s} \right\} \hat{m}^*_{n,t}(\gamma) \right\| \times \left\| \hat{\beta}_n - \beta^0 \right\|$$

$$+ \sup_{\gamma\in\Gamma} \left\| \frac{1}{S_n^2(\gamma)} \sum_{s,t=1}^{n} \omega_{n,s,t} J_s(\beta_{n,*},\gamma) \left\{ \hat{I}_{n,s}(\hat{\beta}_n) - \hat{I}_{n,s} \right\} \hat{m}^*_{n,t}(\gamma) \right\| \times \left\| \hat{\beta}_n - \beta^0 \right\|$$

$$+ \sup_{\gamma\in\Gamma} \left\| \frac{1}{S_n^2(\gamma)} \sum_{s,t=1}^{n} \omega_{n,s,t} m_s(\gamma) \left\{ \hat{I}_{n,s}(\hat{\beta}_n) - \hat{I}_{n,s} \right\} \hat{m}^*_{n,t}(\gamma) \right\|$$

$$= \sum_{i=1}^{4} \sup_{\gamma\in\Gamma} \left\{ \mathcal{B}_{i,n}(\gamma) \right\}.$$

Properties P1 or P2 ensure $||\beta_{n,*} - \beta^0|| \leq ||\hat{\beta}_n - \beta^0|| = O_p(||\tilde{V}_n||^{-1/2})$ and $\sup_{\gamma\in\Gamma} ||V_n(\gamma)\tilde{V}_n^{-1}|| = O(1)$. Define $\iota := (-1)^{1/2}$ and apply de Jong and Davidson's (2000: (A.51)) argument under K1 to deduce

$$\begin{aligned}\mathcal{B}_{1,n}(\gamma) &\leq K\int_{-\infty}^{\infty}\left(\frac{1}{\|J_n(\gamma)\|}\left\|\frac{1}{n}\sum_{s=1}^{n}e^{-\iota\xi s/b_n}\hat{J}_{n,s}(\beta_{n,*},\gamma)\right\|\left\|\frac{1}{S_n(\gamma)}\sum_{t=1}^{n}e^{\iota\xi t/b_n}\hat{m}^*_{n,t}(\gamma)\right\|\right)|\vartheta(\xi)|\,d\xi \\ &= K\int_{-\infty}^{\infty}\mathcal{C}_n(\gamma,\xi)\,\mathcal{D}_n(\gamma,\xi)\,|\vartheta(\xi)|\,d\xi,\end{aligned}$$

where the Fourier coefficients $\vartheta(\xi)$ is defined in K1. Uniform Jacobian consistency Lemma B.4 with K1 properties $\sum_{s,t=1}^{n}|\omega_{n,s,t}| = o(n^2)$, $\max_{1\leq s\leq n}\sum_{t=1}^{n}|\omega_{n,s,t}| = o(n)$, and $b_n = o(n)$ imply $\sup_{\gamma\in\Gamma}\{\mathrm{C}_n(\gamma,\xi)\} = o_p(1)$, and uniform approximation Lemma B.2.a and UCLT Lemma B.7 give $\sup_{\gamma\in\Gamma}\{\mathcal{D}_n(\gamma,\xi)\} = O_p(1)$. Therefore $\sup_{\gamma\in\Gamma}\{\mathcal{B}_{1,n}(\gamma)\} = o_p(1)$ by dominated convergence and K1. Similar arguments apply to the remaining terms by invoking uniform expansion Lemma B.3. $\mathcal{QED}$.

LEMMA B.3 (expansion). *Let* $\beta,\tilde{\beta}\in\mathcal{B}$. *For some sequence* $\{\beta_{n,*}\}$ *satisfying* $||\beta_{n,*}-\tilde{\beta}|| \leq ||\beta - \tilde{\beta}||$, *and for some tiny* $\iota > 0$ *and arbitrarily large finite* $\delta > 0$ *we have* $\sup_{\gamma\in\Gamma}|\hat{m}^*_n(\beta,\gamma) - \hat{m}^*_n(\tilde{\beta},\gamma) - \hat{J}^*_n(\beta_{n,*},\gamma)'(\beta-\tilde{\beta})| = n^{-\delta}\times||\beta-\tilde{\beta}||^{1/\iota}\times o_p(1)$.

PROOF. See Lemma B.5 of Hill (2011). $\mathcal{QED}$.

LEMMA B.4 (Jacobian). *Under P1 or P2* $\sup_{\gamma\in\Gamma}||J^*_n(\hat{\beta}_n,\gamma) - J_n(\gamma)(1+o_p(1))|| = o_p(1)$.

PROOF. Since F is continuous, differentiable and uniformly bounded, under the stated assumptions it is easy to verify the moment smoothness property $\liminf_{n\to\infty}\sup_{||\beta-\beta^0||\leq\delta}\{|E[m^*_{n,t}(\beta,\gamma)]|\} > |E[m^*_{n,t}(\gamma)]|$ for any $\gamma\in\Gamma$ and $\delta > 0$. Hence the claim can be proved exactly as in Hill (2011: Lemma B.7). $\mathcal{QED}$.

LEMMA B.5 (HAC). *Under P1 or P2* $\sup_{\gamma\in\Gamma}|\hat{S}^2_n(\hat{\beta}_n,\gamma)/S^2_n(\gamma) - 1| \xrightarrow{p} 0$.

PROOF. Define $\hat{\mu}^*_{n,t}(\hat{\beta}_n,\gamma) := \hat{m}^*_{n,t}(\hat{\beta}_n,\gamma) - \hat{m}^*_n(\hat{\beta}_n,\gamma)$, $\mu^*_{n,t}(\hat{\beta}_n,\gamma) := m^*_{n,t}(\hat{\beta}_n,\gamma) - m^*_n(\hat{\beta}_n,\gamma)$ and

$$\mathcal{A}_{1,n}(\gamma) := \frac{1}{S^2_n(\gamma)}\sum_{s,t=1}^{n}\omega_{n,s,t}\left\{\hat{\mu}^*_{n,s}(\hat{\beta}_n,\gamma)\hat{\mu}^*_{n,t}(\hat{\beta}_n,\gamma) - \mu^*_{n,s}(\gamma)\mu^*_{n,t}(\gamma)\right\}$$

$$\mathcal{A}_{2,n}(\gamma) := \sum_{s,t=1}^{n}\omega_{n,s,t}\frac{\mu^*_{n,s}(\gamma)\mu^*_{n,t}(\gamma)}{S^2_n(\gamma)} - 1.$$

By the triangle inequality we must show each $\sup_{\gamma\in\Gamma}|\mathcal{A}_{\iota,n}(\gamma)| \xrightarrow{p} 0$. The first term $\sup_{\gamma\in\Gamma}|\mathcal{A}_{1,n}(\gamma)| \xrightarrow{p} 0$ is identically Lemma B.2.b.

It remains to prove pointwise $\mathcal{A}_{2,n}(\gamma) \xrightarrow{p} 0$ and a stochastic equicontinuity property (e.g. Theorem 2.1 of Newey 1991)

$$\limsup_{n\to\infty} P\left(\sup_{\gamma,\tilde{\gamma}\in\Gamma:||\gamma-\tilde{\gamma}||\leq\delta}|\mathcal{A}_{2,n}(\gamma) - \mathcal{A}_{2,n}(\tilde{\gamma})| > \varepsilon\right) < \varepsilon.$$

Step 1 (pointwise): We will apply Theorem 2.1 of de Jong and Davidson (2000). It suffices to verify their Assumptions 1-3, where Assumption 1 holds by K1. Write $\mathcal{Z}_{n,t}(\gamma) : = S^{-1}_n(\gamma)m^*_{n,t}(\gamma)$. Under geometric β-mixing D3 $\{m^*_{n,t}(\beta^0\gamma,),\Im_t\}$ forms a geometric L_2-mixingale with constants $e_{n,t}(\gamma)$ (McLeish 1975: Theorem 2.1). Therefore $\{\mathcal{Z}_{n,t}(\gamma),\Im_t\}$ forms a geometric L_2-mixingale with constants $\mathcal{E}_{n,t}(\gamma) := S^{-1}_n(\gamma)e_{n,t}(\gamma)$, hence $E(\sum_{t=1}^{n}\mathcal{Z}_{n,t}(\gamma))^2 \leq KS^{-2}_n(\gamma)\sum_{t=1}^{n}e^2_{n,t}(\gamma)$ by Theorem 1.6 in McLeish

(1975). Since we assume stationarity, and trivially $E(\sum_{t=1}^{n}\mathcal{Z}_{n,t}(\gamma))^2 = 1$, we can always assume $e_{n,t}(\gamma) = KS_n(\gamma)/n^{1/2}$ hence $\{\mathcal{Z}_{n,t}(\gamma), \Im_t\}$ has constants $\mathcal{E}_{n,t}(\gamma) = K/n^{1/2}$. Coupled with bandwidth $b_n = o(n)$ under K1 their Assumptions 2 and 3 therefore hold..

Step 2 (stochastic equicontinuity): Since $F(\gamma'\psi_t)$ is differentiable and uniformly bounded the proof of stochastic equicontinuity is essentially identical to Step 2 of the proof of UCLT Lemma B.7. $\mathcal{QED}$.

LEMMA B.6 (ULLN). *Let* $\inf_{n\geq N}|E[m^*_{n,t}(\gamma)]| > 0$ *for some* $N \in \mathbb{N}$ *and all* $\gamma \in \Gamma/S$ *where* S *has measure zero. Then* $\sup_{\gamma\in\Gamma/S}\{1/n\sum_{t=1}^{n} m^*_{n,t}(\gamma)/E\left[m^*_{n,t}(\gamma)\right]\} \xrightarrow{p} 1$.

PROOF.

Step 1 (Pointwise LLN): By variance property Lemma B.1.b

$$E\left(\frac{1}{n}\sum_{t=1}^{n}\left\{\frac{m^*_{n,t}(\gamma) - E\left[m^*_{n,t}(\gamma)\right]}{\max\left\{1, \left|E\left[m^*_{n,t}(\gamma)\right]\right|\right\}}\right\}\right)^2 = \frac{S_n^2(\gamma)}{n^2 \times \max\left\{1, \left(E\left[m^*_{n,t}(\gamma)\right]\right)^2\right\}} = o(1).$$

Now invoke Chebyshev's inequality to deduce pointwise $1/n\sum_{t=1}^{n}\{m^*_{n,t}(\gamma) - E[m^*_{n,t}(\gamma)]\}/\max\{1, |E[m^*_{n,t}(\gamma)]|\} \xrightarrow{p} 0$.

Step 2 (ULLN): We first prove $\sup_{\gamma\in\Gamma/S}\{|1/n\sum_{t=1}^{n} m^*_{n,t}(\gamma)|/\sup_{\gamma\in\Gamma/S}|E[m^*_{n,t}(\gamma)]| \xrightarrow{p} 1$. Let $n \geq N$ such that $\sup_{\gamma\in\Gamma}|E[m^*_{n,t}(\gamma)]| > 0$, and define $h^*_{n,t}(\gamma) := m^*_{n,t}(\gamma)/\sup_{\gamma\in\Gamma}|E[m^*_{n,t}(\gamma)]|$. Since by construction $h^*_{n,t}(\gamma)$ is uniformly L_1-bounded on compact Γ it belongs to a separable Banach space. Therefore the L_1-bracketing numbers satisfy $N_{[\,]}(\varepsilon, \Gamma, ||\cdot||_1) < \infty$ (Dudley 1999: Proposition 7.1.7). Combined with the pointwise LLN $1/n\sum_{t=1}^{n}(h^*_{n,t}(\gamma) - E\left[h^*_{n,t}(\gamma)\right]) = o_p(1)$ we have a ULLN (cf. Theorem 7.1.5 of Dudley 1999)

$$\sup_{\gamma\in\Gamma}\left|\frac{1}{n}\sum_{t=1}^{n}\left\{h^*_{n,t}(\gamma) - E\left[h^*_{n,t}(\gamma)\right]\right\}\right| = o_p(1).$$

Now replace $m^*_{n,t}(\gamma)$ with $m^*_{n,t}(\gamma)/E[m^*_{n,t}(\gamma)]$ and repeat the argument to prove $\sup_{\gamma\in\Gamma/S}\{1/n\sum_{t=1}^{n} m^*_{n,t}(\gamma)/E[m^*_{n,t}(\gamma)]\} \xrightarrow{p} 1$. $\mathcal{QED}$.

LEMMA B.7 (UCLT). $\{S_n^{-1}(\gamma)\sum_{t=1}^{n}(m^*_{n,t}(\gamma) - E[m^*_{n,t}(\gamma)]) : \gamma \in \Gamma\} \Longrightarrow \{z(\gamma) : \gamma \in \Gamma\}$ *a scalar* $(0,1)$*-Gaussian process on* $C[\Gamma]$ *with covariance function* $E[z(\gamma_1)z(\gamma_2)]$ *and a.s. bounded sample paths. If P2 also holds then* $\{\mathfrak{S}_n^{-1/2}(\gamma)\sum_{t=1}^{n}\{\mathcal{M}^*_{n,t}(\gamma) - E[\mathcal{M}^*_{n,t}(\gamma)] : \gamma \in \Gamma\} \Longrightarrow \{\mathcal{Z}(\gamma) : \gamma \in \Gamma\}$ *an* $r + 1$ *dimensional Gaussian process on* $C[\Gamma]$ *with zero mean, covariance* I_{r+1}*, and covariance function* $E[\mathcal{Z}(\gamma_1)\mathcal{Z}(\gamma_2)']$.

PROOF. The metric space $\{\Gamma, ||\cdot||_2\}$ is totally bounded. Weak convergence on C[Γ] therefore follows if we demonstrate convergence in *f.d.d.* and uniform stochastic equicontinuity (*s.e.*) defined below (cf. Dudley 1978, Pollard 1984, 1990).

Step 1 (f.d.d.'s): Define

$$z_n(\gamma) := \sum_{t=1}^{n} z_{n,t}(\gamma) = \frac{1}{S_n(\gamma)}\sum_{t=1}^{n}\left\{m^*_{n,t}(\gamma) - E[m^*_{n,t}(\gamma)]\right\}, \tag{20}$$

and for any set of $h \in \mathbb{N}$ vectors $\{\gamma_1, ..., \gamma_h\}$, $\gamma_i \in \Gamma$, and conformable $\lambda'\lambda = 1$, $Z_n(\lambda, \gamma) := \sum_{i=1}^{h}\lambda_i z_n(\gamma_i)$. Test weight boundedness W1 and trimming imply $|z_{n,t}(\lambda, \gamma)| \leq Kc_{\epsilon,n}/S_n(\gamma)$. Therefore under non-degeneracy F2, the existence of a moment R2, distribution tails R3, and mixing R4, the tail-trimmed central limit theorem Lemma B.6 in Hill (2011) can be straightforwardly generalized to show $Z_n(\lambda,\gamma) \xrightarrow{d} N(0,1)$. Convergence in *f.d.d.* now follows from the Cramér-Wold theorem.

Similarly, under slow plug-in property P2 and the above argument, the limiting *f.d.d.* of

$$\mathcal{Z}_n(\gamma) := \mathfrak{S}_n^{-1/2}(\gamma) \sum_{t=1}^{n} \left\{ \mathcal{M}_{n,t}^*(\gamma) - E[\mathcal{M}_{n,t}^*(\gamma)] \right\} \tag{21}$$

are normal.

Step 2 (*s.e.*): Consider $z_n(\gamma)$ in (20) the argument for $\mathcal{Z}_n(\gamma)$ in (21) being identical. The required stochastic equicontinuity condition follows: $\forall \varepsilon > 0$ there exists $\delta > 0$ such that

$$\limsup_{n\to\infty} P\left(\sup_{\gamma,\tilde{\gamma}\in\Gamma:\|\gamma-\tilde{\gamma}\|\le\delta} |z_n(\gamma) - z_n(\tilde{\gamma})| > \varepsilon \right) < \varepsilon.$$

Apply differentiability, W1 boundedness of $F(\gamma'\psi_t)$ and the mean value theorem to obtain

$$\begin{aligned}
\left| \frac{m_{n,t}^*(\gamma)}{\sigma_n(\gamma)} - \frac{m_{n,t}^*(\tilde{\gamma})}{\sigma_n(\tilde{\gamma})} \right| &\le K \frac{|\epsilon_t I_{n,t}(\beta_0)|}{\|\epsilon_t I_{n,t}(\beta_0)\|_2} \times \sup_{\gamma,\tilde{\gamma}\in\Gamma} \left\| \frac{1}{F(\tilde{\gamma}'\psi_t)} \frac{\partial}{\partial\gamma} F(\gamma'\psi_t) \right\| \times \|\gamma - \tilde{\gamma}\| \\
&\quad + \frac{|\epsilon_t I_{n,t}(\beta_0)|}{\|\epsilon_t I_{n,t}(\beta_0)\|_2} \times E\left[\epsilon_t^2 I_{n,t} \sup_{\gamma\in\Gamma} \left(\frac{\partial}{\partial\gamma} F(\gamma'\psi_t) \right) \right] \times \|\gamma - \tilde{\gamma}\| \\
&\le K \left\{ \frac{|\epsilon_t I_{n,t}(\beta_0)|}{\|\epsilon_t I_{n,t}(\beta_0)\|_2} \right\} \times \|\gamma - \tilde{\gamma}\| = \mathcal{A}_{n,t} \times \|\gamma - \tilde{\gamma}\|.
\end{aligned}$$

By Lyapunov's inequality $E[\mathcal{A}_{n,t}] < \infty$. In lieu of boundedness and smoothness of the test weight F, an identical argument extends to partial sums:

$$\begin{aligned}
&\left\| \sup_{\gamma,\tilde{\gamma}\in\Gamma:\|\gamma-\tilde{\gamma}\|\le\delta} |z_n(\gamma) - z_n(\tilde{\gamma})| \right\|_2 \\
&\quad \le \frac{K}{n^{1/2}} \left\| \sup_{\gamma,\tilde{\gamma}\in\Gamma:\|\gamma-\tilde{\gamma}\|\le\delta} \left| \sum_{t=1}^{n} \{ z_{n,t}(\gamma) - z_{n,t}(\tilde{\gamma}) \} \right| \right\|_2 \\
&\quad \le K \left\| \sum_{t=1}^{n} \frac{\{ \epsilon_t I_{n,t}(\beta_0) - E[\epsilon_t I_{n,t}(\beta_0)] \}}{\| \sum_{t=1}^{n} \{ \epsilon_t I_{n,t}(\beta_0) - E[\epsilon_t I_{n,t}(\beta_0)] \} \|_2} \right\|_2 \times \sup_{\gamma,\tilde{\gamma}\in\Gamma:\|\gamma-\tilde{\gamma}\|\le\delta} \|\gamma - \tilde{\gamma}\| \le K\delta.
\end{aligned}$$

Now invoke Chebyshev's inequality to conclude

$$\begin{aligned}
\limsup_{n\to\infty} P\left(\sup_{\gamma,\tilde{\gamma}\in\Gamma:\|\gamma-\tilde{\gamma}\|\le\delta} |z_n(\gamma) - z_n(\tilde{\gamma})| > \varepsilon \right) &\le \varepsilon \times \varepsilon^{-3} \left\| \sup_{\gamma,\tilde{\gamma}\in\Gamma:\|\gamma-\tilde{\gamma}\|\le\delta} |z_n(\gamma) - z_n(\tilde{\gamma})| \right\|_2^2 \\
&\le \varepsilon \times \varepsilon^{-3} K^2 \delta^2 \le \varepsilon,
\end{aligned}$$

for any $0 < \delta \le \varepsilon^{3/2}/K$, which completes the proof. $\mathcal{QED}$.

C.4 COMPLETE SIMULATION RESULTS In the main paper we omitted simulation results for sample size $n = 5000$. The following are the complete results.

Table 1 - Empirical Size (Linear AR)

	iid ϵ_t ($\kappa = 1.5$)[a]			GARCH ϵ_t ($\kappa = 2$)			iid ϵ_t ($\kappa = \infty$)		
n	200	800	5000	200	800	5000	200	800	5000
	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%
TT-Orth-Fix[b,c]	.00,.01,.04[d]	.00,.02,.06	.01,.03,.09	.00,.01,.05	.00,.03,.09	.00,.03,.09	.00,.02,.05	.00,.03,.07	.01,.04,.09
TT-Fix	.00,.02,.07	.01,.04,.08	.01,.05,.10	.00,.01,.04	.00,.02,.05	.01,.06,.10	.00,.02,.06	.00,.02,.04	.00,.02,.05
TT-Orth-OT	.01,.04,.06	.02,.06,.12	.01,.06,.11	.01,.03,.04	.01,.02,.04	.02,.02,.04	.01,.03,.09	.02,.07,.12	.01,.05,.11
TT-OT	.23,.41,.52	.31,.46,.54	.42,.51,.58	.04,.17,.27	.15,.31,.40	.31,.45,.53	.04,.12,.20	.06,.14,.23	.06,.15,.23
CM-Orth[e]	.00,.01,.05	.00,.03,.07	.00,.02,.07	.00,.04,.11	.00,.04,.10	.00,.03,.08	.00,.04,.09	.00,.03,.09	.00,.04,.09
CM	.00,.01,.04	.01,.02,.08	.00,.04,.07	.00,.00,.02	.00,.01,.03	.00,.02,.06	.00,.01,.03	.00,.02,.04	.00,.02,.04
HW[f]	.17,.22,.25	.21,.24,.27	.20,.22,.24	.06,.15,.24	.80,.87,.89	.99,.99,.99	.00,.02,.05	.02,.05,.07	.01,.05,.09
RESET[g]	.00,.00,.02	.00,.01,.02	.00,.02,.02	.00,.03,.09	.01,.05,.11	.01,.05,.10	.00,.03,.08	.01,.05,.10	.01,.05,.10
McLeod-Li[g]	.02,.03,.03	.01,.02,.02	.01,.01,.02	.58,.70,.78	1.0,1.0,1.0	1.0,1.0,1.0	.01,.04,.07	.02,.05,.09	.01,.05,.10
Tsay[g]	.98,.99,1.0	1.0,1.0,1.0	1.0,1.0,1.0	.37,.47,.51	.72,.77,.80	.97,.98,1.0	.01,.05,.10	.01,.05,.10	.01,.05,.10

a. Moment supremum of the test error ϵ_t: $\kappa = \sup\{\alpha : E|\epsilon_t|^\alpha < \infty\}$

b. TT = Tail-Trimmed CM test with randomized nuisance parameter γ. FIX = fixed trimming parameter λ.

c. Orth = orthogonal equation transformation. OT = occupation time test over set of λ.

d. Rejection frequencies at 1%, 5% and 10% nominal levels.

e. Untrimmed randomized and sup-CM tests.

f. Hong and White's (1996) nonparametric test.

g. Ramsey's RESET test with 3 lags; McLeod and Li's test with 3 lags; Tsay's F-test.

Table 2 - Empirical Power[a] (Self-Exciting Threshold AR)

	iid ϵ_t ($\kappa = 1.5$)			GARCH ϵ_t ($\kappa = 2$)			iid ϵ_t ($\kappa = \infty$)		
n	200	800	5000	200	800	5000	200	800	5000
	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%
TT-Orth-Fix	.02,.12,.22	.08,.26,.38	.44,.66,.75	.01,.06,.11	.01,.05,.11	.02,.09,.20	.02,.05,.11	.02,.08,.15	.17,.35,.48
TT-Fix	.12,.18,.24	.21,.32,.39	.46,.57,.64	.01,.05,.11	.03,.12,.23	.35,.58,.71	.02,.07,.12	.09,.27,.43	.72,.87,.95
TT-Orth-OT	.19,.35,.46	.65,.83,.93	.94,.94,.94	.08,.17,.25	.12,.24,.39	.30,.43,.57	.06,.13,.24	.16,.28,.42	.53,.53,.58
TT-OT	.28,.30,.32	.38,.35,.37	.43,.44,.44	.11,.13,.21	.24,.25,.39	.41,.43,.44	.04,.13,.22	.39,.52,.57	.90,.90,.91
CM-Orth	.05,.21,.33	.11,.27,.42	.18,.41,.54	.02,.07,.13	.01,.05,.09	.01,.04,.10	.01,.04,.08	.02,.06,.10	.06,.21,.33
CM	.04,.11,.18	.12,.23,.29	.27,.39,.49	.01,.05,.10	.02,.12,.24	.30,.54,.71	.01,.07,.14	.08,.30,.44	.73,.87,.95
HW	.06,.10,.16	.17,.15,.30	.46,.51,.72	.04,.05,.09	.04,.07,.12	.02,.06,.11	.02,.06,.11	.16,.29,.40	.94,.99,.97
RESET	.03,.14,.28	.08,.28,.45	.24,.44,.63	.02,.12,.24	.15,.38,.53	.31,.58,.70	.20,.54,.73	1.0,1.0,1.0	1.0,1.0,1.0
McLeod-Li	.29,.45,.55	.71,.76,.83	.86,.93,1.0	.00,.00,.07	.01,.05,.10	.01,.05,.10	.07,.19,.27	.51,.69,.79	1.0,1.0,1.0
Tsay	.02,.02,.02	.00,.00,.00	.00,.00,.00	.13,.17,.19	.15,.17,.21	.04,.07,.10	.45,.65,.70	1.0,1.0,1.0	1.0,1.0,1.0

a. The rejection frequencies are adjusted for size distortions based on Table 1.

Table 3 - Empirical Power[a] (Bilinear AR)

	iid ϵ_t ($\kappa = 1.5$)			GARCH ϵ_t ($\kappa = 2$)			iid ϵ_t ($\kappa = \infty$)		
n	200	800	5000	200	800	5000	200	800	5000
	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%	1%,5%,10%
TT-Orth-Fix	.04,.16,.26	.22,.39,.49	.53,.71,.83	.01,.07,.11	.02,.08,.14	.18,.48,.62	.01,.05,.09	.01,.05,.11	.31,.56,.75
TT-Fix	.04,.13,.21	.13,.31,.42	.43,.60,.71	.02,.09,.17	.02,.09,.18	.23,.40,.52	.01,.04,.07	.01,.04,.10	.26,.36,.61
TT-Orth-OT	.08,.14,.18	.36,.38,.37	.51,.52,.57	.02,.05,.10	.04,.06,.09	.11,.15,.19	.01,.05,.05	.00,.01,.02	.04,.09,.14
TT-OT	.57,.61,.61	.68,.58,.60	.76,.78,.78	.27,.30,.36	.57,.52,.55	.67,.69,.70	.01,.07,.13	.06,.12,.16	.20,.28,.34
CM-Orth	.03,.14,.24	.21,.40,.51	.41,.53,.74	.02,.11,.21	.03,.13,.26	.14,.29,.43	.01,.04,.11	.01,.05,.11	.29,.57,.64
CM	.02,.08,.16	.08,.30,.41	.33,.40,.56	.01,.05,.10	.01,.05,.11	.04,.19,.28	.01,.04,.09	.01,.04,.09	.19,.35,.60
HW	.00,.00,.00	.00,.00,.00	.07,.18,.26	.02,.07,.07	.00,.00,.00	.00,.00,.00	.24,.38,.47	.87,.92,.97	.98,1.0,1.0
RESET	.02,.07,.14	.01,.06,.14	.03,.08,.15	.03,.07,.11	.01,.05,.09	.01,.06,.12	.02,.06,.12	.03,.17,.28	.15,.31,.52
McLeod-Li	.19,.26,.33	.35,.43,.51	.52,.63,.73	.00,.02,.04	.00,.00,.00	.01,.05,.10	.86,.93,.98	.99,1.0,1.0	1.0,1.0,1.0
Tsay	.03,.06,.10	.01,.05,.10	.01,.05,.10	.36,.36,.37	.20,.20,.23	.04,.07,.10	.76,.84,.88	.91,.95,.96	1.0,1.0,1.0

a. The rejection frequencies are adjusted for size distortions based on Table 1.

REFERENCES

Davies R. B. (1977). Hypothesis Testing when a Nuisance Parameter is Present Only under the Alternative, Biometrika 64, 247-254.

de Jong R.M., and Davidson J. (2000). Consistency of Kernel Estimators of Heteroscedastic and Autocorrelated Covariance Matrices, Econometrica 68, 407-423.

Dudley, R. M. (1999). Continuous Central Limit Theorems. Cambridge University Press: Cambridge.

Embrechts P., and Goldie C.M. (1980). On Closure and Factorization Properties of Subexponential Distributions, J. Aus. Math. Soc. Ser. A, 29, 243-256.

Hansen B.E. (1996). Inference When a Nuisance Parameter Is Not Identified Under the Null Hypothesis, Econometrica 64, 413-430.

Hill J.B. (2011). Robust M-Estimation for Heavy Tailed Nonlinear AR-GARCH, Working Paper, University of North Carolina - Chapel Hill.

McLeish, D. L. (1975). A Maximal Inequality and Dependent Strong Laws, Ann. Prob. 3, 829-839.

Newey, W.K. (1991). Uniform Convergence in Probability and Stochastic Equicontinuity, Econometrica 59, 1161-1167.

Pollard D. (1984). Convergence of Stochastic Processes. Springer-Verlang New York.

Pollard, D. (1990). Empirical Processes: Theory and Applications, NSF-CBMS Regional Conference Series in Probability and Statistics, Vol. 2.

Resnick, S.I. (1987). Extreme Values, Regular Variation and Point Processes. Springer-Verlag: New York.

Rio, E. (1993). Covariance Inequalities for Strongly Mixing Processes, Annales de l'Instritut Henri Poincaré 29, 587-597.