Robust Estimation and Inference for Heavy Tailed Nonlinear GARCH (Supplemental Material)

Jonathan B. Hill

Dept. of Economics, University of North Carolina

S.1 FRACTILE SELECTION

A variety of covariance methods have been proposed for determining the trimming portion for outlier robust estimation, including Minimum Covariance Determinant (e.g. Rousseeuw et al 2004, Agulló et al 2008), sign and rank methods, and bootstrap covariance for trimmed means (Léger and Romano 1990). We adopt a small sample bootstrap mean-squared-error method for our purposes, while space constraints force us to leave deeper theory issues aside.

In order to simplify notation we assume we only need to trim by the errors ϵ_t, covering GARCH models with a squared integrable scaled volatility derivative $\mathfrak{s}_t := h_t^{-1}(\partial/\partial\theta)h_t(\theta)|_{\theta^0}$, and we use $k_n^{(\epsilon)} = [\lambda n/\ln(n)]$. We want to approximate the small sample mean-squared-error $E[(\hat{\theta}_n - \theta^0)(\hat{\theta}_n - \theta^0)']$ and minimize it over λ. Let $\{\hat{\theta}_{r,n}(\lambda)\}_{r=1}^{R_n}$ be a sequence of bootstrap QMTTL, QMFTTL or MFTTM estimates computed with λ, use any consistent estimator $\tilde{\theta}_n$ of θ^0 that is not function of λ as a plug-in, and define

$$\hat{\mathcal{M}}_n(\lambda) := \frac{1}{R_n}\sum_{r=1}^{R_n}\left(\hat{\theta}_{r,n}(\lambda) - \tilde{\theta}_n\right) \times \left(\hat{\theta}_{r,n}(\lambda) - \tilde{\theta}_n\right)'.$$

In time series settings the prevalent method is a variant of the block bootstrap (cf. Künsch 1989). See Gonçalves and White (2005) for background theory. Any norm of $\hat{\mathcal{M}}_n(\lambda)$ may be minimized on Λ, including the matrix norm $||\hat{\mathcal{M}}_n(\lambda)||$.

As an experiment we compute $\hat{\mathcal{M}}_n(\lambda)$ for each of 1000 samples of size $n = 100$ for a linear GARCH(1,1) model $y_t = \sigma_t\epsilon_t$ and $\sigma_t^2 = .3 + .3y_{t-1}^2 + .6\sigma_{t-1}^2$, where $\epsilon_t \overset{i.i.d.}{\sim} (0,1)$ is either normal or Paretian with index 2.5. See Section 6 in the main paper for simulation details. We use $R_n = n$ bootstrap draws with replacement, sub-sample size $n/2$, and Log-LAD $\tilde{\theta}_n$.[1] In Figures S.1 and S.2 below we compute the simulation average $||\hat{\mathcal{M}}_n(\lambda)||$ over $\lambda \in [.01, 1.0]$. The trimming parameter λ is minimized at .06 for Paretian ϵ_t, and at .04 for Gaussian ϵ_t, and similar results arise for Smooth Transition GARCH, GJR-GARCH (Glosten et al 1993), and Asymmetric GARCH (Engle an Ng 1993). Hill (2011) shows in simulation experiments $\lambda_n^* := \arg\inf_{\lambda \in [0,1]} ||\hat{\mathcal{M}}_n(\lambda)||$ leads to a sharp, and approximately normal estimator $\hat{\theta}_n$. The fact that $||\hat{\mathcal{M}}_n(\lambda)||$ is minimized roughly at $\lambda = .05$ suggests we may simply pick a small value. We show in Section 6 of the main paper that this rule of thumb leads to superb results.

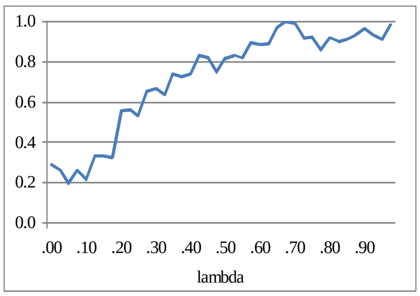

Figure S.1: $||\text{mse}(\lambda)||$, Paretian ϵ_t

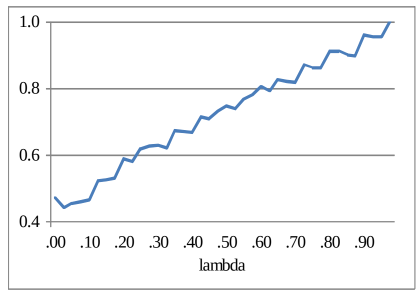

Figure S.2: $||\text{mse}(\lambda)||$, Normal ϵ_t

S.2 PROOFS OF LEMMAS A.4, A.7, and A.8

Let Assumptions 1-5 hold throughout. By Proposition A.2 it suffices to treat the infeasible estimator $\hat{\theta}_n^*$ and components $m_t^*(\theta)$ and $G_t^*(\theta)$. We therefore prove all lemmas for $m_t^*(\theta)$, $G_t^*(\theta)$, and so on. Since the notation "$*$" is repetitive we drop it everywhere and write $\hat{\theta}_n$ for $\hat{\theta}_n^*$, and

$$h_t(\theta) = h_t^*(\theta), \;\; m_t(\theta) = m_t^*(\theta), \;\; m_{n,t}(\theta) = m_{n,t}^*(\theta), \;\; G_t(\theta) = G_t^*(\theta), \;\; G_{n,t}(\theta) = G_{n,t}^*(\theta) = G_t^*(\theta) I_{n,t}^*(\theta).$$

Recall Jacobian and covariance matrices:

$$\mathcal{G}_n(\theta) := E\left[G_{n,t}(\theta)\right], \;\; \Sigma_n(\theta) := E\left[m_{n,s}(\theta) m_{n,t}(\theta)'\right] \;\; \text{and} \;\; \mathcal{S}_n(\theta) := \frac{1}{n} \sum_{s,t=1}^{n} E\left[m_{n,s}(\theta) m_{n,t}(\theta)'\right].$$

Assume *symmetric trimming* throughout to simplify notation. Let $w_t(\theta)$ denote any scalar $m_{i,t}(\theta)$ or $G_{i,j,t}(\theta)$ and let $\{k_n^{(w)}, \mathcal{C}_n^{(w)}(\theta)\}$ be the associated fractile and threshold sequences:

$$w_t(\theta) \in \{m_{i,t}(\theta), G_{i,j,t}(\theta)\}, \;\; P\left(|w_t(\theta)| > \mathcal{C}_n^{(w)}(\theta)\right) = \frac{k_n^{(w)}}{n}.$$

[1] The r^{th} bootstrapped sample is $\{y_{t_r^*}, y_{t_r^*+1}, ..., y_{t_r^*+[n/2]-1}\}$ where t_r^* is a uniform random draw from $\{1, ..., n - [n/2] + 1\}$.

We simply write $\mathcal{C}_n(\theta)$ and k_n whenever $w_t(\theta)$ is understood, and we drop θ^0.

We shorten arguments in lieu of Assumption 3.b by assuming $w_t(\theta)$ have power law tails for *any* θ:

$$\sup_{\theta\in\Theta}\left\{\left|c^{\kappa_w(\theta)}P\left(|w_t(\theta)|>c\right)-d_w(\theta)\right|\right\}\rightarrow 0 \textit{ as } c\rightarrow\infty,\ \inf_{\theta\in\Theta}\left\{\kappa_w(\theta),d_w(\theta)\right\}>0. \tag{1}$$

By construction of $\mathcal{C}_n(\theta)$ and (1) we have the following *fractile and threshold* properties:

$$\inf_{\theta\in\Theta}\frac{\mathcal{C}_n(\theta)}{(n/k_n)^{1/\kappa_w(\theta)}}\rightarrow(0,\infty) \quad\text{and}\quad \sup_{\theta\in\Theta}\frac{\mathcal{C}_n(\theta)}{(n/k_n)^{1/\kappa_w(\theta)}}\rightarrow(0,\infty)\,. \tag{FT}$$

Applications of Karamata's Theorem therefore gives the following *trimmed moments*:

$$\text{if }\ \kappa_w(\theta)<2:\ E\left[w_t^2(\theta)I\left(|w_t(\theta)|\leq\mathcal{C}_n(\theta)\right)\right]\sim K\left(\mathcal{C}_n(\theta)\right)^2\times(k_n/n)=K(n/k_n)^{2/\kappa_w(\theta)-1}$$

$$\text{if }\ \kappa_w(\theta)=2:\ E\left[w_t^2(\theta)I\left(|w_t(\theta)|\leq\mathcal{C}_n(\theta)\right)\right]\sim L\left(n\right).$$

Uniform bounds are similar in lieu of (FT): for example if $\kappa_w(\theta)<2$ then for finite $K>0$

$$\sup_{\theta\in\Theta}\left\{\frac{n}{k_n}\frac{\mathcal{C}_n^2(\theta)}{E\left[w_t^2(\theta)I\left(|w_t(\theta)|\leq\mathcal{C}_n(\theta)\right)\right]}\right\}\rightarrow K. \tag{TM}$$

LEMMA A.4 (*variance bounds*). *a.* $||\mathcal{S}_n(\theta)||\leq Kr_n(\theta)||\Sigma_n(\theta)||=o(n)$ *for some sequence of positive numbers* $\{r_n(\theta)\}$*, where in general* $\liminf_{n\rightarrow\infty}\inf_{\theta\in\Theta}r_n(\theta)>0$ *and* $\sup_{\theta\in\Theta}r_n(\theta)=O(\ln(n))$*, and* $r_n(\theta)\sim K$ *if* $m_t(\theta)$ *is finite dependent or each* $E[m_{i,t}^2(\theta)]<\infty$*; b.* $\Sigma_n(\theta)=o(n||E\left[m_{n,t}(\theta)\right]||^2)$ *and* $\sup_{\theta\in\Theta}||\Sigma_n(\theta)||$ $=o(n\sup_{\theta\in\Theta}||E[m_{n,t}(\theta)]||^2)$.

PROOF.

Claim (a): Assume $m_t(\theta)$ is a mean zero scalar to reduce notation. Drop θ everywhere since, by virtue of $\theta\in\Theta$ a bounded space, θ does not play any role.

Note $\mathcal{S}_n\sim E[m_{n,t}^2]+2\sum_{i=1}^{n-1}(1-i/n)E[m_{n,1}m_{n,i+1}]$. If $E[m_t^2]<\infty$ then $\mathcal{S}_n\sim K$ under geometric β-mixing (cf. Ibragimov 1962). If m_t is finite dependent then $\mathcal{S}_n=KE[m_{n,t}^2]$ has a finite limit if $E[m_t^2]<\infty$.

Finally, assume m_t is non-finite dependent and $E[m_t^2]=\infty$ and recall the Assumption 3 tail index $\kappa\in(1,2]$. By Assumption 4 $\liminf_{n\rightarrow\infty}\mathcal{S}_n/E[m_{n,t}^2]>0$ hence we need only prove $\sum_{i=1}^{n-1}\left|E\left[m_{n,1}m_{n,i+1}\right]\right|\leq KE[m_{n,t}^2]\ln(n)$ and $E[m_{n,t}^2]=o(n/\ln(n))$.

Define the quantile functions $Q_n(u) = \inf\{m : P(|m_{n,t}| > m) \leq u\}$ and $Q(u) = \inf\{m : P(|m_t| > m) \leq u\}$ for $u \in [0,1]$, recall geometric β-mixing implies α-mixing with coefficients $\alpha_h \leq K\rho^h$ for $\rho \in (0,1)$. Since we assume non-finite dependence, without loss of generality set $\alpha_h = \rho^h$ for notational simplicity. By Theorem 1.1 of Rio (1993)

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

By tail-trimming $Q_n(u) = 0 \; \forall u \leq k_n/n$ and $Q_n(u) = Q(u)$ otherwise. Further, under tail decay Assumption 3 $Q(u) = O(u^{-1/\kappa})$. Therefore

$$\begin{aligned}\sum_{i=1}^{n-1} |E[m_{n,1}m_{n,i+1}]| &\leq K\sum_{i=1}^{n-1}\int_{k_n/n}^{\rho^h} u^{-2/\kappa}du = K\sum_{i=1}^{n-1}\max\left\{0, (n/k_n)^{(2/\kappa-1)} - \rho^{-i(2/\kappa-1)}\right\} \\ &= K\sum_{i=1}^{\ln(n/k_n)}\left\{(n/k_n)^{(2/\kappa-1)} - \rho^{-i(2/\kappa-1)}\right\}.\end{aligned}$$

Further $\sum_{i=1}^{\ln(n/k_n)}\{(n/k_n)^{2/\kappa-1} - \rho^{-i(2/\kappa-1)}\} = K\ln(n/k_n) \times (n/k_n)^{2/\kappa-1}(1+O(1))$, hence

$$\sum_{i=1}^{n-1} |E[m_{n,1}m_{n,i+1}]| \leq K\ln(n/k_n) \times (n/k_n)^{2/\kappa-1}(1+O(1)) \leq K\ln(n/k_n) \times (n/k_n)^{2/\kappa-1}.$$

If $\kappa \in (1,2)$ then by Assumption 3 and therefore Karamata's Theorem (Resnick 1987: Theorem 0.6) $E[m_{n,t}^2] \sim Kc_n^2(k_n/n) \sim K(n/k_n)^{2/\kappa-1}$, hence

$$\sum_{i=1}^{n-1} |E[m_{n,1}m_{n,i+1}]| \leq K\ln(n/k_n) \times E[m_{n,t}^2] \leq K\ln(n) \times E[m_{n,t}^2].$$

If $\kappa = 2$ then $E[m_{n,t}^2] \sim L(n) \to \infty$ is slowly varying hence

$$\sum_{i=1}^{n-1} |E[m_{n,1}m_{n,i+1}]| \leq K\ln(n/k_n)\frac{1}{L(n)} \times E[m_{n,t}^2] \leq K\ln(n) \times E[m_{n,t}^2].$$

Finally, by Karamata's Theorem if $\kappa = 2$ then $E[m_{n,t}^2]$ is slowly varying in n hence $E[m_{n,t}^2] = o(n/\ln(n))$. If $\kappa \in (1,2)$ then $E[m_{n,t}^2] \sim Kc_n^2(k_n/n) = K(n/k_n)^{2/\kappa-1} = o(n^{2/\kappa-1}) = o(n/\ln(n))$.

Claims (b) and (c): Define $\underline{\kappa} := \inf_{\theta\in\Theta}\kappa_{m_i}(\theta)$ where $\kappa_{m_i}(\theta)$ is the moment supremum of $m_{i,t}(\theta)$. By (FT), (TM) and the claim (a) argument if $\underline{\kappa} > 1$ then $\sup_{\theta\in\Theta} E[m_{i,n,t}^2(\theta)] = o(n) = o(n\sup_{\theta\in\Theta}|E[m_{i,n,t}(\theta)]|)$. If $\underline{\kappa} \leq 1$ then apply (TM) to $m_{i,n,t}(\theta)$ to deduce

$$\frac{\sup_{\theta\in\Theta} E\left[m_{i,n,t}^2(\theta)\right]}{\sup_{\theta\in\Theta}|E[m_{i,n,t}(\theta)]|^2} \sim K\frac{(n/k_{i,n})^{2/\underline{\kappa}-1}}{(n/k_{i,n})^{2(1/\underline{\kappa}-1)}} = (n/k_{i,n}) = o(n).$$

A similar argument can be used to show $E[m_{i,n,t}^2(\theta)] = o(n|E[m_{i,n,t}(\theta)]|)$. $\mathcal{QED}$.

LEMMA A.7 (*CLT*). $n^{-1/2}\mathcal{S}_n^{-1/2}\sum_{t=1}^n m_{n,t} \xrightarrow{d} N(0, I_q)$.

PROOF. We prove the claim for the scalar case under symmetric trimming and $E[m_{n,t}] = 0$, the general case being similar. Write $k_n = k_n^{(m)}$, $\mathcal{C}_n = \mathcal{C}_n^{(m)}$ and $\mathfrak{S}_n^2 := E(\sum_{t=1}^n m_{n,t})^2$.

By the geometric β-mixing property and stationary, we need only verify (2.1) and (2.2) in Peligrad (1996: Theorem 2.1), stated below:

$$(2.1) \quad \sup_{n\geq 1} \frac{1}{\mathfrak{S}_n^2}\sum_{t=1}^n E\left[m_{n,t}^2\right] < \infty \qquad (2.2) \quad \frac{1}{\mathfrak{S}_n^2}\sum_{t=1}^n E\left[m_{n,t}^2 I\left(|m_{n,t}| > \varepsilon\mathfrak{S}_n\right)\right] \to 0.$$

By non-degeneracy Assumption 4.b and stationarity $nE[m_{n,t}^{*2}]/\mathfrak{S}_n^2 = E[m_{n,t}^{*2}]/\mathcal{S}_n = O(1)$, hence (2.1).

In view of stationarity, for (2.2) we must how $nE[m_{n,t}^2 I(|m_{n,t}| > \varepsilon\mathfrak{S}_n)]/\mathfrak{S}_n^2 \to 0$. Assume $E[m_{n,t}] = 0$ to reduce notation, and observe since $nE[m_{n,t}^2]/\mathfrak{S}_n^2 = O(1)$ it follows

$$\begin{aligned} E\left[m_{n,t}^2 I\left(|m_{n,t}| > \varepsilon\mathfrak{S}_n\right)\right] &= E\left[m_t^2 I\left(|m_t| \leq \mathcal{C}_n\right) I\left(|m_t| I\left(|m_t| \leq \mathcal{C}_n\right) > \varepsilon\mathfrak{S}_n\right)\right] \\ &= E\left[m_t^2 I\left(\varepsilon\mathfrak{S}_n \leq |m_t| \leq \mathcal{C}_n\right)\right] \\ &\leq E\left[m_t^2 I\left(\varepsilon n^{1/2}\left(E\left[m_{n,t}^2\right]\right)^{1/2} \leq |m_t| \leq \mathcal{C}_n\right)\right]. \end{aligned} \tag{2}$$

Write $(z)_+ := \max(0, z)$.

If $\kappa_m \in (0, 2)$ then by Karamata's Theorem $E[m_{n,t}^2] \sim \mathcal{C}_n^2 P(|m_t| > \mathcal{C}_n) \sim \mathcal{C}_n^2(k_n/n)$, hence

$$\frac{n^{1/2}\left(E\left[m_{n,t}^2\right]\right)^{1/2}}{\mathcal{C}_n} = \left(\frac{nE\left[m_{n,t}^2\right]}{\mathcal{C}_n^2}\right)^{1/2} \sim \left(\frac{n\mathcal{C}_n^2\left(k_n/n\right)}{\mathcal{C}_n^2}\right)^{1/2} = k_n^{1/2} \to \infty. \tag{3}$$

If $\kappa = 2$ then $E[m_{n,t}^2] = L(n) \to \infty$ is slowly varying and $\mathcal{C}_n = K(n/k_n)$, hence

$$\frac{n^{1/2}\left(E\left[m_{n,t}^2\right]\right)^{1/2}}{\mathcal{C}_n} \sim \left(\frac{nL(n)}{\mathcal{C}_n^2}\right)^{1/2} = \left(\frac{nL(n)}{n/k_n}\right)^{1/2} = \left(k_n L(n)\right)^{1/2} \to \infty. \tag{4}$$

In either case (3) or (4), together with (2) it follows there exists an $N \in \mathbb{N}$ sufficiently large that $\forall n \geq N$ we have $E[m_t^2 I(\varepsilon n^{1/2}(E[m_{n,t}^2])^{1/2} \leq |m_t| \leq \mathcal{C}_n)] = 0$, hence $nE[m_{n,t}^2 I(|m_{n,t}| > \varepsilon\mathfrak{S}_n)]/\mathfrak{S}_n^2 \to 0$ is immediate. $\mathcal{QED}$.

LEMMA A.8 (*Jacobian consistency*). *a.* $\widehat{\mathcal{G}}_n(\hat{\theta}_n) = E[G_t I_{n,t}] \times (1 + o_p(1))$; *b.* $(\partial/\partial\theta)E[m_{n,t}(\theta)]|_{\theta^0}$

$= E[G_t I_{n,t}] \times (1 + o(1))$.

PROOF.

Claim (a): Define $\tilde{\mathcal{G}}_n(\theta) := 1/n \sum_{t=1}^n G_{n,t}(\theta)$ and recall $\widehat{\mathcal{G}}_n(\theta) := 1/n \sum_{t=1}^n G_t(\theta)\hat{I}_{n,t}(\theta)$. We need only show $\widehat{\mathcal{G}}_n(\hat{\theta}_n) = \tilde{\mathcal{G}}_n(\hat{\theta}_n) \times (1+ o_p(1))$ and $||\tilde{\mathcal{G}}_n(\hat{\theta}_n) - \mathcal{G}_n|| = o_p(||\mathcal{G}_n||)$ where $o_p(1)$ are not functions of θ. The former holds by the same argument used to prove Lemma A.3.c. Now drop θ^0. We have for the latter

$$\left\| \frac{1}{n}\sum_{t=1}^n G_t(\hat{\theta}_n) I_{n,t}^{(G)}(\hat{\theta}_n) - E\left[G_t I_{n,t}^{(G)}\right]\right\|$$
$$\leq \left\| \frac{1}{n}\sum_{t=1}^n \left\{G_t I_{n,t}^{(G)} - E\left[G_t I_{n,t}^{(G)}\right]\right\}\right\| + \left\| \frac{1}{n}\sum_{t=1}^n \left\{G_t(\hat{\theta}_n) I_{n,t}^{(G)}(\hat{\theta}_n) - G_t I_{n,t}^{(G)}\right\}\right\| = \mathcal{A}_n + \mathcal{B}_n(\hat{\theta}_n).$$

If any $G_{i,j,t}$ is uniformly integrable then by geometric β-mixing Assumption 2 $\{G_{i,j,t} I_{n,t}^{(G)}\}$ satisfies Andrews' (1988) LLN: $\mathcal{A}_{i,j,n} := 1/n \sum_{t=1}^n \{G_{i,j,t} I_{n,t}^{(G)} - E[G_{i,j,t} I_{n,t}^{(G)}]\} \xrightarrow{p} 0$. Otherwise $\mathcal{A}_n = o_p(||E[G_t I_{n,t}^{(G)}]||)$ can be shown by exploiting geometric β-mixing and using the argument in the proof of ULLN Lemma A.5.b. In both cases $\mathcal{A}_n = o_p(||E[G_t I_{n,t}^{(G)}]||)$.

Finally, arguments from the proof of Lemma A.6 can be generalized to show $\mathcal{B}_n(\hat{\theta}_n) \leq K||\hat{\theta}_n - \theta^0||^{1/\iota}$ for any tiny $\iota > 0$. Therefore $\mathcal{B}_n(\hat{\theta}_n) = o_p(1)$ given consistency $\hat{\theta}_n \xrightarrow{p} \theta^0$ by Theorem 2.1, and since $\hat{\theta}_n$ and θ^0 are in compact Θ we can always assume $o_p(1)$ is not a function of θ. In lieu of non-degeneracy Assumption 4.b the proof is complete: $||\tilde{\mathcal{G}}_n(\hat{\theta}_n) - \mathcal{G}_n|| = o_p(||\mathcal{G}_n||) + o_p(1) = o_p(||\mathcal{G}_n||)$.

Claim (b): By expansion Lemma A.6.a we have for $\delta > 0$ arbitrarily large, and tiny $\iota > 0$,

$$\frac{1}{n}\sum_{t=1}^n \{m_{n,t}(\theta) - m_{n,t}\} = \frac{1}{n}\sum_{t=1}^n G_t I_{n,t} \times \left(\theta - \theta^0\right) + n^{-\delta} \times \left\|\theta - \tilde{\theta}\right\|^{1/\iota} \times o_p(1).$$

Invoke dominated convergence to deduce

$$\frac{E\left[m_{n,t}\left(\theta\right)\right] - E\left[m_{n,t}\right]}{\left\|\theta - \theta^0\right\|} = E\left[G_t I_{n,t}\right] \times \left(1 + o\left(\left\|\theta - \theta^0\right\|\right)\right) + o\left(n^{-\delta}\right) = \mathcal{G}_n \times \left(1 + o\left(\left\|\theta - \theta^0\right\|\right)\right) + o\left(n^{-\delta}\right).$$

Since $n^{-\delta} \to 0$ is arbitrary and $||\mathcal{G}_n||$ is non-degenerate under Assumption 4.b, it follows

$$\frac{E\left[m_{n,t}\left(\theta\right)\right] - E\left[m_{n,t}\right]}{\left\|\theta - \theta^0\right\|} = \mathcal{G}_n \times \left(1 + o\left(\left\|\theta - \theta^0\right\|\right)\right) + o\left(\left\|\mathcal{G}_n\right\|\right). \tag{5}$$

Identically, by the definition of a derivative

$$\frac{E\left[m_{n,t}\left(\theta\right)\right] - E\left[m_{n,t}\right]}{\left\|\theta - \theta^0\right\|} = \frac{\partial}{\partial\theta} E\left[m_{n,t}\left(\theta\right)\right] \times \left(1 + o\left(\left\|\theta - \theta^0\right\|\right)\right) + o\left(\left\|\mathcal{G}_n\right\|\right). \tag{6}$$

Equate (5) and (6) and take $||\theta - \theta^0|| \to 0$ to prove the claim. $\mathcal{QED}$.

S.3 PROOFS OF LEMMAS B.1-B.4.

LEMMA B.1 (*uniform indicator law*). *Define* $\mathcal{I}_{n,t}(\theta) := ((n/k_n)^{1/2})\{I(|w_t(\theta)| \leq \mathcal{C}_n(\theta)) - E[I(|w_t(\theta)| \leq \mathcal{C}_n(\theta))]\}$. *Then* $\{n^{-1/2}\sum_{t=1}^n \mathcal{I}_{n,t}(\theta) : \theta \in \Theta\} \Longrightarrow^* \{\mathcal{I}(\theta) : \theta \in \Theta\}$ *and* $E[\sup_{\theta\in\Theta} |n^{-1/2}\sum_{t=1}^n \mathcal{I}_{n,t}(\theta)|] = O(1)$ *where* $\{\mathcal{I}(\theta) : \theta \in \Theta\}$ *is a Gaussian process with uniformly bounded and uniformly continuous sample paths with respect to* L_2*-norm, and* $\Longrightarrow^*$ *denotes weak convergence on a Polish space.*[2]

PROOF. By construction $\mathcal{I}_{n,t}(\theta)$ is L_2-bounded uniformly on $1 \leq t \leq n$, $n \geq 1$, and Θ, and under Assumption 2 $\mathcal{I}_{n,t}(\theta)$ is geometrically β-mixing. Further, $\{\mathcal{I}_{n,t}(\theta) : \theta \in \Theta\}$ satisfies the metric entropy with L_2-bracketing bound[3] $\int_0^1 \mathcal{H}_{[\,]}(\varepsilon, \Theta, ||\cdot||_2)d\varepsilon < \infty$. This follows since $w_t(\theta)$ have absolutely continuous distributions under Assumption 3.a, hence the thresholds $\mathcal{C}_n(\theta)$ are continuous. Further, $w_t(\theta)$ have uniformly bounded densities uniformly on Θ by Assumption 3.a. Therefore $\mathcal{I}_{n,t}(\theta)$ is L_2-Lipschitz: $E[(\mathcal{I}_{n,t}(\theta) - \mathcal{I}_{n,t}(\tilde{\theta}))^2] \leq K||\theta - \tilde{\theta}||$. Proving the L_2-bracketing numbers satisfy $\int_0^1 \mathcal{H}_{[\,]}(\varepsilon, \Theta, ||\cdot||_2)d\varepsilon < \infty$ is then a classic exercise (Giné and Zinn 1984, Pollard 1984).

We may therefore apply Doukhan et al's (1995: Theorem 1; eq. (2.17), Application 4) uniform central limit theorem to deduce $\{1/n^{1/2}\sum_{t=1}^n \mathcal{I}_{n,t}(\theta) : \theta \in \Theta\} \Longrightarrow^* \{\mathcal{I}(\theta) : \theta \in \Theta\}$, a Gaussian process with a version that has uniformly bounded and uniformly continuous sample paths with respect to $||\cdot||_2$.

Finally, Doukhan et al's (1995: Theorem 2) uniform maximal inequality applies since their required bound (2.10) holds under their (2.17), which $\int_0^1 \mathcal{H}_{[\,]}(\varepsilon, \Theta, ||\cdot||_2)d\varepsilon < \infty$ ensures. Therefore $E[(\sup_{\theta\in\Theta} |n^{-1/2}\sum_{t=1}^n \mathcal{I}_{n,t}(\theta)|] = O(1)$. $\mathcal{QED}$.

LEMMA B.2 (*uniform order statistic*). *Define* $w_t^{(a)}(\theta) := |w_t(\theta)|$. *Then* $\sup_{\theta\in\Theta} |w_{(k_n)}^{(a)}(\theta)/\mathcal{C}_n(\theta) - 1| = O_p(k_n^{-1/2})$.

---

[2] See Dudley (1978), Giné and Zinn (1984), and Pollard (1984).

[3] The brackets $\{l, u\}$ of an index function class $\mathcal{F}$ satisfy $l \leq f \leq u$ for every member $f \in \mathcal{F}$, where $\{l, u\}$ may not be members of $\mathcal{F}$; an ε-L_2-bracket $\{l, u\}$ satisfies $||l - u|| \leq \varepsilon$; the L_2-bracketing numbers $\mathcal{N}_{[\,]}(\varepsilon, \Phi, ||\cdot||_2)$ are the number of ε-L_2-brackets required to cover $\mathcal{F}$, and metric entropy with L_2-bracketing is $\mathcal{H}_{[\,]}(\varepsilon, \Phi, ||\cdot||_2) = \ln(\mathcal{N}_{[\,]}(\varepsilon, \Phi, ||\cdot||_2))$. See Giné and Zinn (1984) and Pollard (1984). The property $\int_0^1 \mathcal{H}_{[\,]}^{1/2}(\varepsilon, \Phi, ||\cdot||_2)d\varepsilon < \infty$ ensures a required stochastic equicontinuity condition for weak convergence of a partial sum of $\mathcal{I}_{n,t}(\phi)$ (Dudley 1978).

PROOF. We first prove the pointwise limit, and then the uniform limit claim. Assume $\inf_{\theta\in\Theta} w_t(\theta) \geq 0$ hence $w_t^{(a)}(\theta) = w_t(\theta)$ for notational simplicity.

Step 1 (pointwise): Drop θ, and define $\mathcal{I}_n(u) := 1/k_n \sum_{t=1}^n I(w_t > \mathcal{C}_n e^{u/k_n^{1/2}})$ for arbitrary $u \in \mathbb{R}$. Under geometric β-mixing and power-tail decay Assumptions 2 and 3.b $\{k_n^{-1/2} I(w_t > \mathcal{C}_n e^u)\}$ satisfies the conditions of Hill's (2009: Theorem 2.1, Lemma 3.1) central limit theorem. Therefore point-wise

$$k_n^{1/2}\left\{\mathcal{I}_n(u) - E\{\mathcal{I}_n(u)\}\right\} \xrightarrow{d} N(0, \tilde{w}^2(u)), \text{ where } \tilde{w} : \mathbb{R}_+ \to \mathbb{R}_+, \text{ and } \sup_{u\geq 0} \tilde{w}^2(u) < \infty. \tag{7}$$

We need only show $k_n^{1/2}\ln(w_{(k_n)}/\mathcal{C}_n) \xrightarrow{d} N(0, v^2)$ follows from (7). By construction $k_n^{1/2}\ln(w_{(k_n)}/\mathcal{C}_n) \leq u$ for $u \in \mathbb{R}$ sufficiently if $\mathcal{I}_n(u) \leq \rho$ for any $\rho \in [0,1]$ to be chosen below, and $\mathcal{I}_n(u) \leq \rho$ if

$$\begin{aligned} k_n^{1/2}\left(\mathcal{I}_n(u) - E\left[\mathcal{I}_n(u)\right]\right) &\leq k_n^{1/2}\left(\rho - \frac{n}{k_n}P\left(w_t > \mathcal{C}_n e^{u/k_n^{1/2}}\right)\right) \\ &= k_n^{1/2}\left(\rho - \frac{n}{k_n}P\left(w_t > \mathcal{C}_n\right)\frac{P\left(w_t > \mathcal{C}_n e^{u/k_n^{1/2}}\right)}{P\left(w_t > \mathcal{C}_n\right)}\right). \end{aligned}$$

Let mapping $\mathfrak{b} : \mathbb{R}_+ \to \mathbb{R}_+$ satisfy $k_n^{1/2}\mathfrak{b}(\mathcal{C}_n) \to 0$, and let $\mathcal{O}(w) \in [0, w]$ be a contraction mapping that may be different in different places. By a generalization of Hsing's (1991: p. 1553) argument, power-law tail decay implies for some $\mathfrak{b}$

$$\frac{n}{k_n}P\left(w_t > \mathcal{C}_n\right) = \mathcal{O}(1) \times \left[1 + \mathcal{O}\left(\mathfrak{b}\left(\mathcal{C}_n\right)\right)\right] = \mathcal{O}(1) + o\left(1/k_n^{1/2}\right) \text{ as } n \to \infty$$

and

$$\frac{P\left(w_t > \mathcal{C}_n e^u\right)}{P\left(w_t > \mathcal{C}_n\right)} = \mathcal{O}(e^{-u\kappa}) \times \left(1 + \mathcal{O}\left(\mathfrak{b}\left(\mathcal{C}_n\right)\right)\right) = \mathcal{O}(e^{-u\kappa}) \times \left(1 + o\left(1/k_n^{1/2}\right)\right).$$

Now put $\rho = \mathcal{O}(1) \in [0,1]$ to deduce $k_n^{1/2}\ln(w_{(k_n)}/\mathcal{C}_n) \leq u$ sufficiently if

$$\begin{aligned} \kappa^{-1}k_n^{1/2}\left(\mathcal{I}_n(u) - E\left[\mathcal{I}_n(u)\right]\right) &\leq \kappa^{-1}k_n^{1/2}\left(\rho - \left(\mathcal{O}(1) + o\left(1/k_n^{1/2}\right)\right) \times \mathcal{O}(e^{-u\kappa}) \times \left(1 + o\left(1/k_n^{1/2}\right)\right)\right) \\ &\leq \kappa^{-1}k_n^{1/2}\left\{\rho - \mathcal{O}\left(e^{-u\kappa/k_n^{1/2}}\right) \times \left(1 + o\left(1/k_n^{1/2}\right)\right)\right\} \\ &\leq \kappa^{-1}k_n^{1/2}\left\{u\kappa/k_n^{1/2} + o\left(1/k_n^{1/2}\right)\right\} = u + o(1). \end{aligned}$$

Since $\kappa^{-1}k_n^{1/2}\{\mathcal{I}_n(u) - E[\mathcal{I}_n(u)]\} \xrightarrow{d} Z$ a mean-zero normal law with finite variance, it follows

$$\lim_{n\to\infty} P\left(k_n^{1/2}\ln\left(w_{(k_n)}/\mathcal{C}_n\right) \leq u\right) = \lim_{n\to\infty} P\left(\kappa^{-1}k_n^{1/2}\left(\mathcal{I}_n(u) - E\left[\mathcal{I}_n(u)\right]\right) \leq u + o(1)\right) = P(Z \leq u). \tag{8}$$

Therefore $k_n^{1/2}\ln(w_{(k_n)}/\mathcal{C}_n) \xrightarrow{d} N(0, v^2)$ where $v^2 < \infty$, hence $w_t/\mathcal{C}_n = 1 + O_p(k_n^{-1/2})$ by the mean-value-theorem.

Step 2 (uniform): Define $\mathcal{I}_n\,(u,\theta) := 1/k_n \sum_{t=1}^n I(w_t^{(a)}(\theta) > \mathcal{C}_n(\theta)e^{u/k_n^{1/2}})$ and $\mathcal{Z}_n(u,\theta) := k_n(n/k_n)^{1/2}\mathcal{I}_n(u,\theta)$. Invoke uniform tail properties (FT) and (TM) and repeat the argument leading to (8) to obtain for any $u \in \mathbb{R}$

$$P\left(\sup_{\theta\in\Theta}\left|k_n^{1/2}\ln\left(m_{(k_n)}^{(a)}(\theta)/\mathcal{C}_n(\theta)\right)\right| \leq u\right) = P\left(\kappa^{-1}\sup_{\theta\in\Theta}\left|n^{-1/2}\left(\mathcal{Z}_n\,(u,\theta) - E\left[\mathcal{Z}_n\,(u,\theta)\right]\right)\right| \leq u + o(1)\right).$$

Now invoke uniform indicator law Lemma B.1 for the right hand side and the mapping theorem to deduce $\sup_{\theta\in\Theta}|w_{(k_n)}^{(a)}(\theta)/\mathcal{C}_n(\theta) - 1| = O_p(k_n^{-1/2})$. $\mathcal{QED}$.

Finally, we require an approximation for a cross-product sum for HAC asymptotics.

LEMMA B.3 (*cross-product approximation*). *Under the kernel properties of Theorem 5.1* $\mathfrak{S}_n^{-1}\sum_{s,t=1}^n \mathcal{K}_{n,s,t}\{\hat{m}_{n,s}(\hat{\theta}_n)\hat{m}_{n,t}(\hat{\theta}_n) - m_{n,s}(\theta^0)m_{n,t}(\theta^0)\} = o_p(1)$.

PROOF. Assume $m_t(\theta)$ is a scalar to simplify notation. Write

$$\mathfrak{S}_n(\theta) := n\mathcal{S}_n(\theta) = \sum_{s,t=1}^n E\left[m_{n,s}(\theta)m_{n,t}(\theta)'\right].$$

It suffices to show

$$\mathcal{M}_{1,n} := \left|\frac{1}{\mathfrak{S}_n}\sum_{s,t=1}^n \mathcal{K}_{n,s,t}\left\{\hat{m}_{n,s}\hat{m}_{n,t} - m_{n,s}m_{n,t}\right\}\right| = o_p(1)$$
$$\mathcal{M}_{2,n} := \frac{1}{\mathfrak{S}_n}\sum_{s,t=1}^n \mathcal{K}_{n,s,t}\left\{\hat{m}_{n,s}(\hat{\theta}_n)\hat{m}_{n,t}(\hat{\theta}_n) - \hat{m}_{n,s}\hat{m}_{n,t}\right\} = o_p(1).$$

Step 1 ($\mathcal{M}_{1,n} = o_p(1)$): By the triangle inequality $\mathcal{M}_{1,n}$ is bounded by

$$K\left|\frac{1}{\mathfrak{S}_n}\sum_{s,t=1}^n \mathcal{K}_{n,s,t}m_s\left(\hat{I}_{n,s} - I_{n,s}\right)\left\{m_{n,t} - E\left[m_{n,t}\right]\right\}\right| + K\left|\frac{1}{\mathfrak{S}_n}\sum_{s,t=1}^n \mathcal{K}_{n,s,t}m_s\left(\hat{I}_{n,s} - I_{n,s}\right)E\left[m_{n,t}\right]\right|$$
$$+ K\left|\frac{1}{\mathfrak{S}_n}\sum_{s,t=1}^n \mathcal{K}_{n,s,t}m_s\left(\hat{I}_{n,s} - I_{n,s}\right)m_t\left(\hat{I}_{n,t} - I_{n,t}\right)\right| = \mathcal{A}_{1,n} + \mathcal{A}_{2,n} + \mathcal{A}_{3,n}.$$

Consider $\mathcal{A}_{1,n}$ and define for any $\delta > 0$

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

$$\mathcal{A}_{1,n,\delta} := \sum_{t=-n+1}^{2n} \left(\frac{1}{\gamma_n^{1/2}} \sum_{l=1-t}^{n-t} \mathcal{K}(l/\gamma_n) \frac{1}{\mathfrak{S}_n^{1/2}} m_{t+l} \left(\hat{I}_{n,t+l} - I_{n,t+l}\right) I\left(0 \le l \le [\gamma_n/\delta]\right) \right)$$
$$\times \left(\frac{1}{\gamma_n^{1/2}} \sum_{j=1-t}^{n-t} \eta_{\delta,n,j} \frac{1}{\mathfrak{S}_n^{1/2}} m_{n,t+j} I\left(0 \le j \le [\gamma_n/\delta]\right) \right) \times (1 + o_p(1)).$$

Trivially since $\mathfrak{S}_n = \sum_{s,t=1}^n E[m_{n,s} m_{n,t}]$ we have

$$E\left(\frac{1}{\mathfrak{S}_n^{1/2}} \sum_{t=1}^n \{m_{n,t} - E[m_{n,t}]\}\right)^2 = 1 = \sum_{t=1}^n \left(1/n^{1/2}\right)^2, \tag{9}$$

and by approximation Lemma A.3.a

$$E\left(\frac{1}{\mathfrak{S}_n^{1/2}} \sum_{t=1}^n m_t \left(\hat{I}_{n,t} - I_{n,t}\right)\right)^2 = o(1) \le K \sum_{t=1}^n \left(1/n^{1/2}\right)^2 = K. \tag{10}$$

In view of Lemmas A.1-A.3 in de Jong and Davidson (2000) it follows[4]

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \left\|\mathcal{A}_{1,n} - \mathcal{A}_{1,n,\delta} \times (1 + o_p(1))\right\|_1 = 0. \tag{11}$$

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

$$\limsup_{n\to\infty} \frac{N_n(\delta)}{\gamma_n} \le K \text{ and } \frac{S^2_{N_n(\delta)}/N_n(\delta)}{\mathfrak{S}_n/n} = O(1).$$

Approximation Lemma A.3.a and CLT Lemma A.7 generalize in straightforward ways to kernel-weighted versions under the suppositions of Theorem 5.1: for any δ

$$\frac{n^{1/2}}{\gamma_n^{1/2}} \times \max_{-n+1 \le t \le 2n} \left\| \frac{1}{\mathfrak{S}_n^{1/2}} \sum_{l=1-t}^{n-t} \mathcal{K}(l/\gamma_n) m_{t+l} \left(\hat{I}_{n,t+l} - I_{n,t+l}\right) I\left(0 \le l \le [\gamma_n/\delta]\right) \right\|_2$$
$$\le \frac{N_n^{1/2}(\delta)}{\gamma_n^{1/2}} \left\{ \frac{S_{N_n(\delta)}/N_n^{1/2}(\delta)}{\mathfrak{S}_n^{1/2}/n^{1/2}} \right\} \times \left\| \frac{1}{S_{N_n(\delta)}} \sum_{t=1}^{N_n(\delta)} \mathcal{K}(t/\gamma_n) \{\hat{m}_{n,t+l} - m_{n,t+l}\} \right\|_2 \to 0 \text{ as } n \to \infty,$$

[4] de Jong and Davidson (2000) invoke a mixingale maximal inequality due to McLeish (1975) solely to ensure partial sum variance bounds in their Lemma A.1. It suffices to replace their Lemma A.1 with (9) and (10) since these duplicate the same bound implied by Theorem 1.6 of McLeish (1975) with mixingale constants $1/n^{1/2}$.

and

$$\frac{n^{1/2}}{\gamma_n^{1/2}} \max_{-n+1\leq t\leq 2n} \left\| \sum_{j=1-t}^{n-t} \eta_{\delta,n,j} \frac{1}{\mathfrak{S}_n^{1/2}} m_{n,t+j} I\left(0 \leq j \leq [\gamma_n/\delta]\right) \right\|_2$$
$$\leq \frac{N_n^{1/2}(\delta)}{\gamma_n^{1/2}} \left\{ \frac{S_{N_n(\delta)}/N_n^{1/2}(\delta)}{\mathfrak{S}_n^{1/2}/n^{1/2}} \right\} \times \left\| \frac{1}{S_{N_n(\delta)}} \sum_{t=1}^{N_n(\delta)} \eta_{\delta,n,j} m_t \left(\hat{I}_{n,t} - I_{n,t}\right) \right\|_2 \to 0 \text{ as } n \to \infty.$$

Therefore

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \left\|\mathcal{A}_{1,n,\delta}\right\|_1 = 0. \tag{12}$$

Together (11) and (12) imply $\mathcal{A}_{1,n} = o_p(1)$. A similar argument applies to $\mathcal{A}_{3,n}$

Finally, $\mathcal{A}_{2,n}$ and $\mathcal{A}_{3,n}$ follow by the same argument for $m_s(\hat{I}_{n,s} - I_{n,s})$. For $E[m_{n,t}]$ in $\mathcal{A}_{2,n}$ use the Theorem 5.1 supposition $\max_{1\leq s\leq t} \sum_{t=1}^{n} |\mathcal{K}_{n,s,t}| = o(n)$, identification Assumption 5 and $\mathfrak{S}_n = n\mathcal{S}_n$, to deduce $o(||n\mathfrak{S}_n^{-1/2} E[m_{n,t}]||) = o(||n^{1/2}\mathcal{S}_n^{-1/2} E[m_{n,t}]||) = o(1)$.

Step 2 ($\mathcal{M}_{2,n} = o_p(1)$): Note $|\mathfrak{S}_n^{-1} \sum_{s,t=1}^{n} \mathcal{K}_{n,s,t}\{\hat{m}_{n,s}(\hat{\theta}_n)\hat{m}_{n,t}(\hat{\theta}_n) - \hat{m}_{n,s}\hat{m}_{n,t}\}|$ is bounded by

$$2\left|\frac{1}{\mathfrak{S}_n} \sum_{s,t=1}^{n} \mathcal{K}_{n,s,t} \left\{\hat{m}_{n,s}(\hat{\theta}_n) - \hat{m}_{n,s}\right\} \hat{m}_{n,t}\right| + \left|\frac{1}{\mathfrak{S}_n} \sum_{s,t=1}^{n} \mathcal{K}_{n,s,t} \left\{\hat{m}_{n,s}(\hat{\theta}_n) - \hat{m}_{n,s}\right\} \left\{\hat{m}_{n,t}(\hat{\theta}_n) - \hat{m}_{n,t}\right\}\right|.$$

Consider the first term, the second being similar. Similar to the proof of Lemma A.6, apply a Taylor expansion and multiple applications of Minkowski's inequality to deduce for some $||\theta_{n,*} - \theta^0|| \leq ||\hat{\theta}_n - \theta^0||$

$$\left|\frac{1}{\mathfrak{S}_n} \sum_{s,t=1}^{n} \mathcal{K}_{n,s,t} \left\{\hat{m}_{n,s}(\hat{\theta}_n) - \hat{m}_{n,s}\right\} \hat{m}_{n,t}\right|$$
$$\leq \left\|\frac{1}{\mathfrak{S}_n} \sum_{s,t=1}^{n} \mathcal{K}_{n,s,t} G_s(\theta_{n,*}) \hat{I}_{n,s}(\theta_{n,*}) \hat{m}_{n,t}\right\| \times \left\|\hat{\theta}_n - \theta^0\right\|$$
$$+ \left\|\frac{1}{\mathfrak{S}_n} \sum_{s,t=1}^{n} \mathcal{K}_{n,s,t} G_s(\theta_{n,*}) \left\{\hat{I}_{n,s}(\theta_{n,*}) - \hat{I}_{n,s}\right\} \hat{m}_{n,t}\right\| \times \left\|\hat{\theta}_n - \theta^0\right\|$$
$$+ \left\|\frac{1}{\mathfrak{S}_n} \sum_{s,t=1}^{n} \mathcal{K}_{n,s,t} G_s(\theta_{n,*}) \left\{\hat{I}_{n,s}(\hat{\theta}_n) - \hat{I}_{n,s}\right\} \hat{m}_{n,t}\right\| \times \left\|\hat{\theta}_n - \theta^0\right\|$$
$$+ \left\|\frac{1}{\mathfrak{S}_n} \sum_{s,t=1}^{n} \mathcal{K}_{n,s,t} m_s \left\{\hat{I}_{n,s}(\hat{\theta}_n) - \hat{I}_{n,s}\right\} \hat{m}_{n,t}\right\| = \sum_{j=1}^{4} \mathcal{B}_{j,n}.$$

By Theorem 2.2 $||\theta_{n,*} - \theta^0|| \leq ||\hat{\theta}_n - \theta^0|| = O_p(||\mathcal{V}_n||^{-1/2})$. Now define $\imath := (-1)^{1/2}$ and apply de Jong and Davidson's (2000: (A.51)) argument under the stated kernel properties to deduce

$$\begin{aligned}\mathcal{B}_{1,n} &\leq K\int_{-\infty}^{\infty}\left(\frac{1}{\|\mathcal{G}_n\|}\left\|\frac{1}{n}\sum_{s=1}^{n}e^{-\imath\xi s/\gamma_n}G_{n,s}(\theta_{n,*})\hat{I}_{n,s}(\theta_{n,*})\right\|\left\|\frac{1}{\mathfrak{S}_n^{1/2}}\sum_{t=1}^{n}e^{\imath\xi t/\gamma_n}\hat{m}_{n,t}\right\|\right)|\varpi(\xi)|\,d\xi \\ &= K\int_{-\infty}^{\infty}\mathcal{C}_n\mathcal{D}_n\,|\varpi(\xi)|\,d\xi,\end{aligned}$$

where $\varpi(\xi) = (2\pi)^{-1}\int_{-\infty}^{\infty}\mathcal{K}(x)e^{\imath\xi x}dx < \infty$. Then $\mathcal{C}_n = o_p(1)$ follows from Jacobian consistency Lemma A.8.a in view of $\theta_{n,*} \xrightarrow{p} \theta^0$, the Theorem 5.1 suppositions $\sum_{s,t=1}^{n}|\mathcal{K}_{n,s,t}| = o(n^2)$, $\max_{1\leq s\leq n}\sum_{t=1}^{n}|\mathcal{K}_{n,s,t}| = o(n)$, and $\gamma_n = o(n)$. Similarly $\mathcal{D}_n = O_p(1)$ follows from approximation Lemma A.3.a and CLT Lemma A.7. Therefore $\mathcal{B}_{1,n} = o_p(1)$ by dominated convergence. By exploiting arguments used in the proof of expansion Lemma A.6 it is straightforward to show the remaining $\mathcal{B}_{j,n} = o_p(1)$. $\mathcal{QED}$.

REFERENCES

[1] Agulló, J., C. Croux, and S. Van Aelst (2008). The Multivariate Least-Trimmed Squares Estimator, J. Mult. Anal. 99, 311-338.

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

[3] Engle, R. and V. Ng (1993). Measuring and Testing the Impact of News on Volatility, J. Fin. 48, 1749-1778.

[4] Doukhan, P., P. Massart, and E. Rio, 1995, Invariance Principles for Absolutely Regular Empirical Processes. Annal. I. H. P. 31, 393-427.

[5] Dudley, R. M. (1978). Central Limit Theorem for Empirical Processes, Ann. Prob. 6, 899-929.

[6] Glosten, L.R., R. Jagannathan, and D.E. Runkle (1993). On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, J. Fin. 48, 1779-1801.

[7] Giné, E. and J. Zinn (1984). Some Limit Theorems for Empirical Processes, Ann. Prob. 12, 929-989.

[8] Gonçalves, S. and H. White (2005). Bootstrap Standard Error Estimates for Linear Regression, J. Amer. Stat. Assoc. 100, 970-979.

[9] Hill, J.B. (2009). On Functional Central Limit Theorems for Dependent, Heterogeneous Arrays with Applications to Tail Index and Tail Dependence Estimation. J. Stat. Plan. Infer. 139, 2091-2110.

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

[11] Hsing, T. (1991). On Tail Index Estimation Using Dependent Data, Ann. Stat. 19, 1547-1569.

[12] Ibragimov, I.A. (1962). Some Limit Theorems for Stationary Processes, Theory Prob. Appl. 7, 349–382.

[13] Künsch, H. R. (1989), "The Jackknife and the Bootstrap for General Stationary Observations, Ann. Stat. 17, 1217–1241.

[14] Léger, C. and J.P. Romano (1990). Bootstrap Adaptive Estimation: The Trimmed-Mean Example, Can. J. Stat. 18, 297-314.

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

[16] Peligrad, M. (1996). On the Asymptotic Normality of Sequences of Weak Dependent Random Variables, J. Theor. Prob. 9, 703-715.

[17] Pollard, D. (1984). Convergence of Stochastic Processes, Springer-Verlang New York.

[18] Rio, E. (1993). Covariance Inequalities for Strongly Mixing Processes, Annal. I. H. P. 29, 587-597.

[19] Rousseeuw, P.J., S. Van Alest, K. Van Driessen, and J. Agulló (2004). Robust Multivariate Regression, Technometrics 46 293–305.