Moment Condition Tests for Heavy-Tailed Time Series: Supplementary Appendix C

Jonathan B. Hill[1] and Mike Aguilar[2]

Dept. of Economics, University of North Carolina

July 2011

In this appendix we present omitted simulation robustness checks (Section C.1), omitted proofs for Lemmas B.1-B.5 (Section C.2) and CLT Lemma B.6 (Section C.3), and supporting theory for the proofs (Section C.4). All citations are presented in a self-contained bibliography at the end.

C.1. OMITTED SIMULATION RESULTS We present robustness Check #1 by performing the omitted variables test with GMTTM and LAD plug-ins. Recall LAD satisfies fast convergence P1 when variance is infinite (Lemma 3.1), and when variance is finite fails to satisfy either P1 or P2 due to a $T^{1/2}$-rate and nonlinearity.

GMTTM is computed from exactly identified least squares equations $\mathcal{M}_t(\theta) = (y_t - \theta' x_t)x_t$ where $x_t = [y_{t-1}, y_{t-2}]'$. Under the null $\mathcal{M}_{i,t} = \epsilon_t y_{t-i}$ is symmetrically distributed hence we symmetrically trim $\tilde{m}_{i,T,t}(\theta) = \mathcal{M}_{i,t}(\theta) \times I(|\mathcal{M}_{i,t}(\theta)| \leq \mathcal{M}^{(a)}_{i,(\tilde{k}_T+1)}(\theta))$. The two-step estimator is then computed exactly as described above. The use of GMTTM and LAD leads to results qualitatively similar to the OLS case. See Tables D1 and D2.

Although Ling's (2005) Least Weighted Absolute Deviations is ruled out in all cases due to a comparatively slow rate of convergence, or nonlinearity, it also performs roughly the same (not shown).

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

[2] Dept. of Economics, University of North Carolina-Chapel Hill, maguilar@email.unc.edu.

Table D1 - Omitted Variables with GMTTM Plug-In

Heavy Tails (1.5)				Thin Tails (4.5)		
Test	H_0 : IID[a]	H_1 : AR		Test	H_0 : IID	H_1 : AR
T = 100						
TT-FIX[b]	.01, .06, .11[c]	.15, .32, .46		TT-FIX	.01, .04, .08	.17, .40, .54
TT-MCD	.01, .06, .11	.19, .38, .57		TT-MCD	01, .05, .9	.17, .40, .53
TT-OT	.01, .07, .14	.27, .50, .62		TT-OT	.02, .06, .12	.20, .40, .54
MC	.00, .02, .04	.09, .26, .41		MC	.00, .01, .03	.10, .32, .44
Wald	.01 ,.03, .05	.72, .86, .93		Wald	.01, .05, .09	.64, .80, .87
T = 500						
TT-FIX	.01, .06, .12	.87, .94, .99		TT-FIX	.01, .04, .08	.97, .99, 1.0
TT-MCD	..01, .07, .12	.76, .89, .92		TT-MCD	.01, .06, .12	.94, .95, .99
TT-OT	.01, .06, .12	.96, .99, .99		TT-OT	.01, .07, .13	.97, .97, .97
MC	.00, .01, .04	.19, .26, .50		MC	.00, .02, .03	.82, .91, .94
Wald	.01 ,.05, .11	.99, .99, .99		Wald	.01, .03, .06	1.0, .10, 1.0
T = 1000						
TT-FIX	.02, .07, .13	.97, .97, .97		TT-FIX	.01, .06, .11	.99, .99, .99
TT-MCD	.01, .05, .11	.99, .99, .99		TT-MCD	.01, .06, .10	.83, .90, .93
TT-OT	.01, .06, .12	.98, .98, .98		TT-OT	.01, .05, .10	.98, .98, .99
MC	.01, .03, .07	.33, .51, .63		MC	.01, .02, .04	.90, .93, .96
Wald	.01, .05, .10	.99, 1.0, 1.0		Wald	.02, .04, .06	1.0, 1.0, 1.0

a. The null is no omitted variables in an AR(2). The alternative is the second lag is omitted.
b. FIX indicates pre-chosen tail parameter $\lambda = .05$ in $k_T = [\lambda T/\ln(T)]$.
c. Values are rejection frequencies at the 1%, 5% and 10% levels.

Table D2 - Omitted Variables with LAD Plug-In

Heavy Tails (1.5)				Thin Tails (4.5)		
Test	H_0 : IID	H_1 : AR		Test	H_0 : IID	H_1 : AR
T = 100						
TT-FIX	.01, .05, .11	.16, .38, .51		TT-FIX	.01, .06, .09	.21, .47, .61
TT-MCD	.01, .06, .11	.19, .39, .54		TT-MCD	01, .06, .11	.22, .46, .58
TT-OT	.01, .07, .12	.33, .53, .66		TT-OT	.01, .07, .12	.26, .50, .60
MC	.00, .01, .04	.10, .27, .42		MC	.00, .01, .03	.16, .39, .51
Wald	.01 ,.03, .06	.74, .91, .94		Wald	.01, .05, .09	.70, .84, .93
T = 500						
TT-FIX	.02, .07, .13	.88, .97, .97		TT-FIX	.01, .04, .09	1.0, 1.0, 1.0
TT-MCD	.01, .06, .12	.82, .93, .98		TT-MCD	.01, .06, .10	.99, .99, .99
TT-OT	.01, .06, .12	.98, .98, .98		TT-OT	.01, .06, .11	.93, .97, .99
MC	.00, .02, .06	.23, .43, .57		MC	.00, .02, .02	91, .95, .98
Wald	.01 ,.05, .11	.99, .99, .99		Wald	.01, .04, .07	1.0, .10, 1.0
T = 1000						
TT-FIX	.01, .06, .12	.98, .98, .98		TT-FIX	.01, .05, .09	1.0, 1.0, 1.0
TT-MCD	.01, .06, .11	.99, .99, .99		TT-MCD	.01, .04, .09	.86, .94, .98
TT-OT	.01, .05, .11	.99, .99, .99		TT-OT	.01, .05, .11	.99, .99, .99
MC	.01, .03, .06	.36, .54, .64		MC	.00, .01, .01	.99, .99, .99
Wald	.01, .05, .10	1.0, 1.0, 1.0		Wald	.01, .04, .06	1.0, 1.0, 1.0

C.2. PROOFS OF LEMMAS B.1-B.5 The proofs of some of lemmas require additional results that were not required directly to prove Theorems 2.1 and 2.2. We therefore state in some cases an expanded version of a previously stated lemma.

Recall the fractile $\{k_{j,i,T}\}$ and threshold $\{l_{i,T}(\theta), u_{i,T}(\theta)\}$ construction:

$$k_{j,i,T} \to \infty,\ k_{j,i,T}/T \to 0,\ \ 1 \leq k_{1,i,T} + k_{2,i,T} < T$$

$$l_{i,T}(\theta) \to L_i(\theta) \ \text{ and } \ u_{i,T}(\theta) \to U_i(\theta) \ \text{ uniformly on compact } \Theta \subset \mathbb{R}^r,$$

and

$$\frac{T}{k_{1,i,T}} P\left(m_{i,t}(\theta) < -l_{i,T}(\theta)\right) = 1 \ \text{ and } \ \frac{T}{k_{2,i,T}} P\left(m_{i,t}\left(\theta\right) > u_{i,T}(\theta)\right) = 1. \tag{1}$$

Assume $L(T)$ is a slowly varying function, $L(T) \to \infty$, whose value and rate may change with the context. Write compactly throughout

$$c_{i,T}\left(\theta\right) := \max\left\{l_{i,T}\left(\theta\right), u_{i,T}\left(\theta\right)\right\}, \ \ c_T\left(\theta\right) = \max_{1 \leq i \leq q}\left\{c_{i,T}\left(\theta\right)\right\}$$

$$k_{i,T} = \max\left\{k_{1,i,T}, k_{2,i,T}\right\} \ \text{ and } \ k_T = \max_{1 \leq i \leq q}\left\{k_{i,T}\right\}$$

$$\hat{m}_T^*(\theta) := \frac{1}{T}\sum_{t=1}^T \hat{m}_{T,t}^*(\theta) \ \text{ and } \ m_T^*(\theta) := \frac{1}{T}\sum_{t=1}^T m_{T,t}^*(\theta)$$

Asymptotic arguments require the following constructions, some of which are already defined above. Estimating equation instantaneous and long run covariance matrices are

$$\Sigma_T\left(\theta\right) = E\left[\left\{m_{T,t}^*\left(\theta\right) - E[m_{T,t}^*(\theta)]\right\}\left\{m_{T,t}^*\left(\theta\right) - E[m_{T,t}^*(\theta)]\right\}'\right] \text{ and } \Sigma_T = \Sigma_T\left(\theta^0\right) \in \mathbb{R}^{q\times q}$$

$$S_T\left(\theta\right) := \sum_{s,t=1}^T E\left[\left\{m_{T,s}^*\left(\theta\right) - E[m_{T,s}^*(\theta)]\right\}\left\{m_{T,t}^*\left(\theta\right) - E[m_{T,t}^*(\theta)]\right\}'\right] \text{ and } S_T = S_T\left(\theta^0\right)$$

$$\tilde{\Sigma}_T\left(\theta\right) = E\left[\left\{\tilde{m}_{T,t}^*\left(\theta\right) - E[\tilde{m}_{T,t}^*(\theta)]\right\}\left\{\tilde{m}_{T,t}^*\left(\theta\right) - E[\tilde{m}_{T,t}^*(\theta)]\right\}'\right] \text{ and } \tilde{\Sigma}_T = \tilde{\Sigma}_T\left(\theta^0\right)$$

$$\tilde{S}_T\left(\theta\right) := \sum_{s,t=1}^T E\left[\left\{\tilde{m}_{T,s}^*\left(\theta\right) - E[\tilde{m}_{T,s}^*(\theta)]\right\}\left\{\tilde{m}_{T,t}^*\left(\theta\right) - E[\tilde{m}_{T,t}^*(\theta)]\right\}'\right] \text{ and } \tilde{S}_T = \tilde{S}_T\left(\theta^0\right),$$

and

$$\mathfrak{S}_T^*\left(\theta\right) := \sum_{s,t=1}^T E\left[\left\{\mathcal{M}_{T,s}^*(\theta) - E[\mathcal{M}_{T,s}^*(\theta)]\right\}\left\{\mathcal{M}_{T,t}^*(\theta) - E[\mathcal{M}_{T,t}^*(\theta)]\right\}'\right],$$

where $\mathcal{M}_{T,t}^*(\theta)$ contains all unique $m_{i,T,t}^*(\theta)$ and $\tilde{m}_{i,T,t}^*(\theta)$.

We abuse notation since $\tilde{\Sigma}_T(\theta)$, $\tilde{S}_T(\theta)$ and $\mathfrak{S}_T^*(\theta)$, which depict covariance in $\tilde{m}_{T,t}^*(\theta)$, may not exist for any θ. See conditions P1-P2 below. Population and sample Jacobia are

$$J_T(\theta) := \frac{\partial}{\partial\theta}E\left[m_{T,t}^*(\theta)\right] \in \mathbb{R}^{q\times r} \ \text{ and } \ J_T = J_T(\theta^0)$$

$$J_{T,t}^*(\theta) := \left[\frac{\partial}{\partial\theta}m_{i,t}(\theta) \times I_{i,T,t}\left(\theta\right)\right]_{i=1}^q \ \text{ and } \ J_T^*(\theta) := \frac{1}{T}\sum_{t=1}^T J_{T,t}^*(\theta),$$

and a scale matrix is

$$V_T(\theta) := T^2 J_T'(\theta) S_T^{-1}(\theta) J_T(\theta) \in \mathbb{R}^{r\times r} \text{ and } V_T := V_T(\theta^0).$$

Finally, recall the assumptions. The first set concerns $\hat{\theta}_T$.

P1 (*fast plug-in convergence*). *$\tilde{V}_T^{1/2}(\hat{\theta}_T - \theta^0) = O_p(1)$ and $||V_T\tilde{V}_T^{-1}|| \to 0$ where $\tilde{\Sigma}_T^2$ and $\tilde{S}_T$ may not exist.*

P2 (*slow plug-in convergence*).

a. $\tilde{V}_T \sim \mathcal{K}V_T$ for some positive definite $\mathcal{K} \in \mathbb{R}^{r\times r}$;

b. $\tilde{V}_T^{1/2}(\hat{\theta}_T - \theta^0) = \tilde{A}_T\sum_{t=1}^T\{\tilde{m}_{T,t} - E[\tilde{m}_{T,t}]\} \times (1 + o_p(1)) + o_p(1)$ for unique $\theta^0 \in \Theta$ where non-stochastic $\tilde{A}_T \in \mathbb{R}^{r\times p}$ has full column rank $\tilde{A}_T\tilde{S}_T^{-1}\tilde{A}_T' \to I_p$;

c. The limiting finite dimensional distributions for $\mathfrak{S}_T^{*-1/2}\{\mathcal{M}_{T,t}^* - E[\mathcal{M}_{T,t}^*]\}$ belong to the same class as those for $S_T^{-1}\{m_{T,t}^* - E[m_{T,t}^*]\}$.

The second set characterizes identification and related smoothness properties.

I1 (*integrability*). *m_t is integrable under the null (1).*

I2 (*identification*). *Under the null (1) the thresholds $\{l_{i,T}, u_{i,T}\}$ satisfy a sequence of fixed point bounds: $E[m_{T,t}^*] = o(||S_T||^{1/2}/T)$.*

I3 (*covariance*). $\sup_\theta ||A_T(\theta)|| < \infty$ *and* $\liminf_{T\geq N}\inf_\theta\{\lambda_{\min}(A_T(\theta))\} > 0$ *for each* $A_T(\theta) \in \{\Sigma_T(\theta), \tilde{\Sigma}_T(\theta), \tilde{S}_T(\theta), \mathfrak{S}_T^*(\theta)\}$ *if* $\tilde{\Sigma}_T(\theta)$ *and* $\mathfrak{S}_T^*(\theta)$ *exist.*

I4 (*moment smoothness*). $\liminf_{T\geq N}\sup_{||\theta-\theta^0||\leq\delta}\{||E[m_{T,t}^*(\theta)]||\} > ||E[m_{T,t}^*]||$ *for some* $N \geq 1$ *and any* $\delta > 0$.

D1 (*distribution*).

i. The finite dimensional distributions of $m_t(\theta)$ are strictly stationary and absolutely continuous with respect to Lebesgue measure on Θ.

ii. If $\sup_\theta E[m_{i,t}^2(\theta)] = \infty$ then $m_{i,t}(\theta)$ have for each t a common power-law tail $P(|m_{i,t}(\theta)| > m) = d_i(\theta)m^{-\kappa_i(\theta)}(1 + o(1))$ where $\inf_\theta \kappa_i(\theta) > 0$, $\kappa_i = \kappa_i(\theta^0) > 1$ under the null (1), $\inf_\theta d_i(\theta) > 0$ where $d_i(\theta)$ is for each θ a constant, and $\sup_\theta\{d_i^{-1}(\theta)m^{\kappa_i(\theta)}P(|m_{i,t}(\theta)| > m)\} \to 1$.

The third set concerns distribution properties.

D2 (*differentiability*). *$m_t(\theta)$ is continuous and differentiable on Θ-a.e.*

D3 (*mixing*). *$\mathcal{M}_{T,t}^*(\theta)$ is for each T strictly stationary over $1 \leq t \leq T$, and geometrically β-mixing: $\beta_l := \sup_{\mathcal{A}\subset\Im_{t+l}^{+\infty}} E|P(\mathcal{A}|\Im_{-\infty}^t) - P(\mathcal{A})| = o(\rho^l)$ for $\rho \in (0,1)$, where $\{\Im_t\}$ is some sequence of σ-fields adapted to $\{\mathcal{M}_{T,t}^*(\theta)\}$, and $\Im_t$ does not depend on T or θ.*

D4 (*moment envelopes*). $\sup_\theta |m_{i,t}(\theta)|$ *and* $\sup_\theta |(\partial/\partial\theta_j)m_{i,t}(\theta)|$ *are L_ι-bounded $\forall i, j$.*

D5 (*Jacobia rank and smoothness*).

i. $\sup_\theta ||A_T(\theta)|| < \infty$ *and $A_T(\theta)$ has full column rank for each $A_T(\theta) \in \{J_T(\theta), J_T^*(\theta), E[J_{T,t}^*(\theta)]\}$.*

ii. For all $\{\delta_T\}$, $\delta_T \to 0$, $\sup_{||\theta-\theta^0||\leq\delta_T}\{||J_T(\theta)||/||J_T||\} = 1 + o(1)$.

D6 (*indicator class*). *$\{I_{i,T,t}(\theta) : \theta \in \Theta\}$ satisfies metric entropy with L_2-bracketing $\mathcal{H}_{[\,]}(\varepsilon, \Theta, ||\cdot||_2) = O(\ln(\varepsilon))$, $\varepsilon \in (0,1)$.*

Finally, the HAC kernel function.

K1 (*kernel*). *$k(\cdot)$ is a member of the class $\{k : \mathbb{R} \to [-1,1] \mid k(0) = 1, k(x) = k(-x)\ \forall x \in \mathbb{R}, \int_{-\infty}^{\infty} |k(x)|dx$*

$< \infty$, $\int_{-\infty}^{\infty} |\varpi(\xi)|d\xi < \infty$, $k(\cdot)$ *is continuous at* 0 *and all but a finite number of points*$\}$, *where* $\varpi(\xi) := (2\pi)^{-1}\int_{-\infty}^{\infty} k(x)e^{i\xi x}dx < \infty$. *Further* $\sum_{s,t=1}^{T} |k((s-t)/\gamma_T)| = o(T^2)$, $\max_{1\leq s\leq T}\sum_{t=1}^{T} k((s-t)/\gamma_T) = o(T)$ *and bandwidth* $\gamma_T = o(T)$.

We require two preliminary results which we prove in Section C.4. The maximum thresholds $c_T(\theta)$ are uniformly bounded, a property used variously to bound tail trimmed covariances.

LEMMA C.1 (*threshold bound*). *Under D1 and I3* $\sup_\theta\{c_T(\theta)/||\Sigma_T(\theta)||^{1/2}\} = O((T/\min_{1\leq i\leq q}\{k_{i,T}\})^{1/2})$ *and* $\sup_\theta\{c_T(\theta)/||S_T(\theta)||^{1/2}\} = O(1/\min_{1\leq i\leq q}\{k_{i,T}\})$.

The equations $\hat{m}_T^*(\theta)$ satisfy a stochastic differentiability property which we use to prove sample Jacobian consistency.

LEMMA C.2 (*stochastic differentiability*). *Under D1-D6 and I3 for any* $\delta \geq 0$

$$\sup_{\theta\in U^0(\delta)}\left\{\frac{\left\|\{\hat{m}_T^*(\theta)-\hat{m}_T^*\}-\left\{E\left[m_{T,t}^*(\theta)\right]-E\left[m_{T,t}^*\right]\right\}\right\|}{1+\|J_T\|\times\left\|\theta-\theta^0\right\|}\right\} = \sup_{\theta\in U^0(\delta)}\left\{\frac{\|J_T^*(\theta)-J_T\|}{\|J_T\|}\right\}+o_p(1).$$

We now prove Lemmas B.1-B.6. First, an expanded version of Lemma B.1.

LEMMA B.1 (*covariance properties*). *Under D3 and* $\liminf_{T\geq N}\inf_\theta\{\lambda_{\min}(\Sigma_T(\theta))\} > 0$*:*

a. $\limsup_{T\geq N}\sup_\theta ||TS_T^{-1}(\theta)\Sigma_T(\theta)|| \leq K$;

b. $||\Sigma_T(\theta)|| = O((T/\min\{k_{i,T}\}) \times \max\{1, ||E[m_{T,t}^*((\theta))]||\})$ *and* $\sup_\theta ||\Sigma_T(\theta)|| = o(T \times \max\{1, \sup_\theta ||E[m_{T,t}^*((\theta))]||\})$;

c. If additionally D1 holds then $||S_T|| = O(T(T/\min\{k_{i,T}\}))$.

PROOF.

Claim (a): Assume $q = 1$ for ease of notation, the general case being similar. Trivially

$$S_T(\theta) = TE\left[m_{T,t}^{*2}(\theta)\right]\times\left(1+2\sum_{h=1}^{T-1}(1-h/T)\frac{E\left[m_{T,1}^*(\theta)\,m_{T,h+1}^*(\theta)\right]}{E\left[m_{T,t}^{*2}(\theta)\right]}\right) = TE\left[m_{T,t}^{*2}(\theta)\right]\times(1+2\mathcal{R}_T(\theta)).$$

Under by non-degeneracy I3 it must be the case $\mathcal{R} := \liminf_{T\to\infty}\inf_\theta \mathcal{R}_T(\theta) > -1/2$. Therefore

$$\limsup_{T\geq N}\sup_\theta\left\{TS_T^{-1}(\theta)\,E\left[m_{T,t}^{*2}(\theta)\right]\right\} \leq \frac{1}{1+2\mathcal{R}} = K < \infty.$$

Claim (b): If $||\Sigma_T(\theta)|| < \infty$ the claim is trivial, so assume at least one $E[m_{i,t}^2(\theta)] = \infty$, and assume without loss of generality $m_{i,t}(\theta)$ is symmetrically trimmed with two-tailed thresholds $c_{i,T}(\theta)$ and fractiles $k_{i,T}$: $(T/k_{i,T})P(|m_{i,t}(\theta)| > c_{i,T}(\theta)) = 1$. Power-law tail D1.ii implies $c_{i,T}(\theta) = d_i(\theta)^{1/\kappa_i(\theta)}(T/k_{i,T})^{1/\kappa_i(\theta)}$ for some $\kappa_i(\theta) \in (0,2]$ and $d_i(\theta) > 0$ that is for each θ a constant. Coupled with properties of trimmed variances for regularly varying tails if $\kappa_i(\theta) \in (1,2)$ then[3]

$$E\left[\left(m_{i,T,t}^*(\theta)\right)^2\right] \sim Kc_{i,T}^2(\theta)P\left(|m_{i,t}(\theta)| > c_{i,T}(\theta)\right) \sim Kc_{i,T}^2(\theta)(k_{i,T}/T) = K(T/k_{i,T})^{2/\kappa_i(\theta)-1}.$$

It is easy to show $(T/k_{i,T})^{2/\kappa_i(\theta)-1} = o(T)$ for all $\kappa_i(\theta) \geq 1$. Similarly if $\kappa_i(\theta) = 2$ then $E[(m_{i,T,t}^*(\theta))^2] \sim L(T) \to \infty$ a slowly varying function which is trivially $o(T)$. Now invoke the Cauchy-Schwartz inequality

[3] The trimmed variance property follows directly from Karamata's Theorem. See Resnick (1987: Theorem 0.6; Exercise .0.4.2.8).

to deduce $\Sigma_T(\theta) = o(T) = o(T \times \max\{1, ||E[m^*_{i,T,t}(\theta)]||\})$. If $\kappa_i(\theta) < 1$ then $|E[m^*_{i,T,t}(\theta)]| \sim c_{i,T}(\theta)(k_{i,T}/T)$ $= K(T/k_{i,T})^{1/\inf_{\theta\in\Theta_{2,i}}\kappa_i(\theta)-1}$, hence

$$\frac{E\left[\left(m^*_{i,T,t}(\theta)\right)^2\right]}{\left|E\left[m^*_{i,T,t}(\theta)\right]\right|^2} \sim K(T/k_{i,T}).$$

The uniform case is identical in lieu of uniform power law tail property D1.ii.

Claim (c): Simplify notation by assuming $q = 1$, $E[m^*_{T,t}] = 0$ and $\kappa \in (1,2]$. The case $E[m^*_{T,t}] \neq 0$ is identical, $\kappa > 2$ is trivial given geometric β-mixing, and $\kappa > 1$ is required for equation integrability. Note $S_T = TE[m^{*2}_{T,t}] + 2T\sum_{h=1}^{T-1}(1 - h/T)E[m^*_{T,1}m^*_{T,h+1}]$, so in lieu of claim (b) it suffices to prove $\sum_{i=1}^{T-1} E[m^*_{T,i}m^*_{T,i+1}] = O(T/k_T)$.

Define the quantile function $Q_T(u) = \inf\{m : P(|m^*_{T,i}| > m) \leq u\}$ 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:

$$\sum_{h=1}^{T-1}\left|E[m^*_{T,1}m^*_{T,h+1}]\right| \leq 2\sum_{h=1}^{T-1}\int_0^{\alpha_h} Q^2_T(u)du \leq 2\sum_{h=1}^{T-1}\int_0^{\rho^h} Q^2_T(u)du.$$

Under Paretian tail Assumption D1.i it is easy to verify

$$Q_T(u) = d^{1/\kappa}u^{-1/\kappa}\left(1 + o(1)\right),$$

and given tail-trimming $P(|m_t| \leq c_T) = k_T/T$ and $m^*_{T,t} = 0\ \forall |m_t| > c_T$ hence $\sup_{u\in[0,1]} Q_T(u) = c_T$ or $u \geq k_T/T$. Therefore

$$\begin{aligned}\sum_{h=1}^{T-1}\left|E[m^*_{T,1}m^*_{T,h+1}]\right| &\leq 2d^{2/\kappa}\sum_{h=1}^{T-1}\int_{k_T/T}^{\rho^h} u^{-2/\kappa}du = 2d^{1/\kappa}\sum_{h=1}^{T-1}\max\left\{0, (T/k_T)^{(2/\kappa-1)} - \rho^{-h(2/\kappa-1)}\right\} \\ &= 2d^{2/\kappa}\sum_{h=1}^{(\ln 1/\rho)^{-1}\ln(T/k_T)}\left\{(T/k_T)^{(2/\kappa-1)} - \rho^{-h(2/\kappa-1)}\right\}.\end{aligned}$$

Finally, $\sum_{h=1}^{(\ln 1/\rho)^{-1}\ln(T/k_T)}\{(T/k_T)^{2/\kappa-1} - \rho^{-h(2/\kappa-1)}\} \leq (\ln 1/\rho)^{-1}\ln(T/k_T) \times (T/k_T)^{2/\kappa-1} = O(T/k_T)$ for any $\kappa \in (1,2]$. This completes the proof. $\mathcal{QED}$.

Next, an expanded version of Lemma B.2.

LEMMA B.2 (*approximations*). *Under D1-D4, D6, I3 and P1 or P2:*

a. $\left\|\sum_{t=1}^T\left\{\hat{m}^*_{T,t}(\theta) - m^*_{T,t}(\theta)\right\}\right\| = o_p\left(\|S_T(\theta)\|^{1/2}\right)$ *for any* $\theta \in \Theta$

b. $\sup_\theta\left\{\left\|1/T\sum_{t=1}^T\{\hat{m}^*_{T,t}(\theta) - m^*_{T,t}(\theta)\}\right\|\right\} = o_p\left(\sup_\theta\left\|E[m^*_{T,t}(\theta)]\right\|\right)$.

Recall the kernel function $k_{T,s,t}$, *and define* $\hat{\mu}^*_{T,t}(\theta) := \hat{m}^*_{T,t}(\theta) - \hat{m}^*_T(\theta)$ *and* $\mu^*_{T,t}(\theta) := m^*_{T,t}(\theta) - m^*_T(\theta)$. *If additionally moment smoothness I4 and kernel property K1 hold then:*

c. $\sup_{\theta\in U^0(\delta_T)}\left\{\|\hat{m}^*_T(\theta) - m^*_T(\theta)\| / \left[1 + \|J_T\| \times \left\|\theta - \theta^0\right\|\right]\right\} = o_p(1)\ \forall\delta > 0.$

d. $\left\|S_T^{-1}\sum_{s,t=1}^T k_{T,s,t}\left\{\hat{\mu}^*_{T,s}(\hat{\theta}_T)\hat{\mu}^*_{T,t}(\hat{\theta}_T)' - \mu^*_{T,s}\mu^{*\prime}_{T,t}\right\}\right\| = o_p(1).$

PROOF. Assume θ and $m_t(\theta)$ are scalars and $m_t(\theta)$ is symmetrically trimmed for notational convenience, and write $\bar{I}_{T,t}(\theta) := 1 - I_{T,t}(\theta)$. Assume θ and $m_t(\theta)$ are scalars and $m_t(\theta)$ is symmetrically trimmed for notational convenience, and write $\bar{I}_{T,t}(\theta) := 1 - I_{T,t}(\theta)$.

Claim (a): Let $\theta \in \Theta$ be arbitrary, and write $m_t = m_t(\theta)$, $c_T = c_T(\theta)$, $\hat{m}^*_{T,t} = \hat{m}^*_{T,t}(\theta)$, $m^*_{T,t} = m^*_{T,t}(\theta)$, $\bar{I}_{T,t} = 1 - I_{T,t}(\theta)$, $\hat{I}_{T,t} = \hat{I}_{T,t}(\theta)$, and $S_T := S_T(\theta)$. First bound

$$\left\|\sum_{t=1}^{T}\left\{\hat{m}^*_{T,t} - m^*_{T,t}\right\}\right\| \leq \max_{1\leq t\leq T}\left\{\left\|m_t\left\{\hat{I}_{T,t} - I_{T,t}\right\}\right\|\right\} \times \sum_{t=1}^{T}\left\|\hat{I}_{T,t} - I_{T,t}\right\|.$$

By construction $||m_t\{\hat{I}_{T,t} - I_{T,t}\}|| \leq 2||m^{(a)}_{(k_T)} - c_T||$, where $m^{(a)}_{(k_T)}/c_T = 1 + O_p(k_T^{-1/2})$ follows under D1-D4 and D6 by Lemma D.2.1 of HR (2010a). Now use threshold bound Lemma C.1 to deduce

$$\max_{1\leq t\leq T}\left\{\left\|m_t\left\{\hat{I}_{T,t} - I_{T,t}\right\}\right\|\right\} \leq 2\left\|m^{(a)}_{(k_T)} - c_T\right\| = 2c_T\left\|m^{(a)}_{(k_T)}/c_T - 1\right\| = o_p\left(\|S_T\|^{1/2} k_T^{-1/2}\right).$$

Next, by construction and the triangle inequality

$$\sum_{t=1}^{T}\left\|\hat{I}_{T,t} - I_{T,t}\right\| \leq k_T^{1/2}\left\|\frac{1}{k_T^{1/2}}\sum_{t=1}^{T}\left\{\bar{I}_{T,t} - E\left[\bar{I}_{T,t}\right]\right\}\right\| + k_T^{1/2}\left\|k_T^{1/2}\left(\frac{T}{k_T}E\left[\bar{I}_{T,t}\right] - 1\right)\right\|$$

which is $O_p(k_T^{1/2})$ by the threshold construction (1) and an application of HR's (2010a: Lemma D.4) uniform indicator law. Therefore $\sum_{t=1}^{T}\{\hat{m}^*_{T,t} - m^*_{T,t}\} = o_p(||S_T||^{1/2})$.

Claim (b): Define

$$M^*_T := \max_{1\leq t\leq T}\left\{\sup_{\theta}\left\|m_t(\theta)\left\{\hat{I}_{T,t}(\theta) - I_{T,t}(\theta)\right\}\right\|\right\}.$$

and repeat the above argument to reach

$$\begin{aligned}\sup_{\theta}\left\|\frac{1}{T}\sum_{t=1}^{T}\left\{\hat{m}^*_{T,t}(\theta) - m^*_{T,t}(\theta)\right\}\right\| &\leq M^*_T \times \frac{k_T^{1/2}}{T}\sup_{\theta}\left\|\frac{1}{k_T^{1/2}}\sum_{t=1}^{T}\left\{\bar{I}_{T,t}(\theta) - E\left[\bar{I}_{T,t}(\theta)\right]\right\}\right\| \\ &\quad + M^*_T \times \frac{k_T^{1/2}}{T}\sup_{\theta}\left\|k_T^{1/2}\left(\frac{T}{k_T}E\left[\bar{I}_{T,t}(\theta)\right] - 1\right)\right\|.\end{aligned}$$

Uniform indicator law Lemma D.4 in HR (2010a) and the threshold construction (1) imply the right-hand-side is $O_p(M^*_T k_T^{1/2}/T)$.

We need only prove $M^*_T = o_p(\sup_\theta ||E[m^*_{T,t}(\theta)]||T/k_T^{1/2})$ to complete the proof. Since

$$\left|m_t(\theta)\left\{\hat{I}_{T,t}(\theta) - I_{T,t}(\theta)\right\}\right| \leq 2c_T(\theta)\left|m^{(a)}_{(k_T)}(\theta)/c_T(\theta) - 1\right|,$$

and $\sup_\theta |m^{(a)}_{(k_T)}(\theta)/c_T(\theta) - 1| = O_p(k_T^{-1/2})$ by Lemma D.2.1 of HR (2010a), use threshold bound Lemma C.1, and covariance bound Lemma B.1.b to deduce

$$\begin{aligned}M^*_T &\leq K\sup_{\theta}c_T(\theta)\sup_{\theta}\left|m^{(a)}_{(k_T)}(\theta)/c_T(\theta) - 1\right| \leq o_p\left(\sup_{\theta}\|\Sigma_T(\theta)\|^{1/2}T^{1/2}/k_T^{1/2}\right) \\ &= o_p\left(\sup_{\theta}\left\|E\left[m^*_{T,t}(\theta)\right]\right\|T/k_T^{1/2}\right).\end{aligned}$$

Claim (c): The claim follows from (b) and Jacobian smoothness $\sup_{\theta\in U^0(\delta)}||J_T(\theta)||/||J_T|| = O(1)$ under condition D5.ii, since by the definition of a derivative

$$\begin{aligned}\sup_{\theta\in U^0(\delta)}\left\{\frac{\left\|E\left[m^*_{T,t}(\theta)\right]\right\|}{1 + \|J_T\| \times \left\|\theta - \theta^0\right\|}\right\} &\leq \sup_{\theta\in U^0(\delta)}\left\{\frac{\left\|E\left[m^*_{T,t}\right]\right\| + \|J_T(\theta)\| \times \left\|\theta - \theta^0\right\|}{1 + \|J_T\| \times \left\|\theta - \theta^0\right\|}\right\} \\ &\leq \sup_{\theta\in U^0(\delta)}\left\{\frac{\left\|E\left[m^*_{T,t}\right]\right\|}{1 + \|J_T\| \times \left\|\theta - \theta^0\right\|}\right\} + K.\end{aligned}$$

Under the null $||E[m^*_{T,t}]|| \to 0$, while under the alternative use moment smoothness I4 to deduce $\sup_{\theta \in U^0(\delta)} ||E[m^*_{T,t}(\theta)]|| > ||E[m^*_{T,t}]||$. Under either hypothesis, therefore,

$$\sup_{\theta \in U^0(\delta)} \left\{ \frac{\left\| E\left[m^*_{T,t}(\theta)\right] \right\|}{1 + \|J_T\| \times \left\|\theta - \theta^0\right\|} \right\} + K.$$

Claim (d): We will prove the simpler result under the null

$$\left\| S_T^{-1} \sum_{s,t=1}^{T} k_{T,s,t} \left\{ \hat{m}^*_{T,s}(\hat{\theta}_T) \hat{m}^*_{T,t}(\hat{\theta}_T) - m^*_{T,s} m^{*\prime}_{T,t} \right\} \right\| = o_p(1).$$

A proof of the claim under either hypothesis

$$\left\| S_T^{-1} \sum_{s,t=1}^{T} k_{T,s,t} \left\{ \hat{\mu}^*_{T,s}(\hat{\theta}_T) \hat{\mu}^*_{T,t}(\hat{\theta}_T)' - \mu^*_{T,s} \mu^{*\prime}_{T,t} \right\} \right\| = o_p(1)$$

is similar, but with tedious added steps to handle stochastic centering $\hat{\mu}^*_{T,t}(\hat{\theta}_T) = \hat{m}^*_{T,t}(\hat{\theta}_T) - 1/T \sum_{t=1}^{T} \hat{m}^*_{T,t}(\hat{\theta}_T)$.

Write $m_t = m_t(\theta^0)$, $\hat{I}_{T,t} = \hat{I}_{T,t}(\theta^0)$, $I_{T,t} = I_{T,t}(\theta^0)$, $\bar{I}_{T,t} := 1 - I_{T,t}$, $\hat{m}^*_{T,t} = m_t \hat{I}_{T,t}$, and $m^*_{T,t} = m_t I_{T,t}$. We prove $||T^{-1} S_T^{-1} \sum_{s,t=1}^{T} k_{T,s,t} \{\hat{m}^*_{T,s} \hat{m}^*_{T,t} - m^*_{T,s} m^*_{T,t}\}|| = o_p(1)$ and $||T^{-1} S_T^{-1} \sum_{s,t=1}^{T} k_{T,s,t} \{\hat{m}^*_{T,s}(\hat{\theta}_T) \hat{m}^*_{T,t}(\hat{\theta}_T) - \hat{m}^*_{T,s} \hat{m}^*_{T,t}\}|| = o_p(1)$ in two steps. The claim then follows by the triangle inequality.

Step 1: Observe

$$\begin{aligned}
&\left\| S_T^{-1} \sum_{s,t=1}^{T} k_{T,s,t} \left\{ \hat{m}^*_{T,s} \hat{m}^*_{T,t} - m^*_{T,s} m^*_{T,t} \right\} \right\| \\
&\quad \le 2 \left\| S_T^{-1} \sum_{s,t=1}^{T} k_{T,s,t} m_s \left(\hat{I}_{T,s} - I_{T,s} \right) m^*_{T,t} \right\| \\
&\qquad + \left\| S_T^{-1} \sum_{s,t=1}^{T} k_{T,s,t} m_s \left(\hat{I}_{T,s} - I_{T,s} \right) m_t \left(\hat{I}_{T,t} - I_{T,t} \right) \right\| = \mathcal{A}_{1,T} + \mathcal{A}_{2,T}.
\end{aligned}$$

We only bound $\mathcal{A}_{1,T}$ since $\mathcal{A}_{2,T}$ is similar. Define for any $\delta > 0$

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

$$\begin{aligned}
\mathcal{A}_{1,T,\delta} := \sum_{t=-T+1}^{2T} &\left(\frac{1}{\gamma_T^{1/2}} \sum_{l=1-t}^{T-t} k(l/\gamma_T) S_T^{-1/2} m_{t+l} \left(\hat{I}^*_{T,t+l} - I^*_{T,t+l} \right) I\left(0 \le l \le [\gamma_T/\delta]\right) \right) \\
&\times \left(\frac{1}{\gamma_T^{1/2}} \sum_{j=1-t}^{T-t} \eta_{\delta,T,j} S_T^{-1/2} m^*_{T,t+j} I\left(0 \le j \le [\gamma_T/\delta]\right) \right) \times (1 + o_p(1)).
\end{aligned}$$

By CLT Lemma B.6

$$\left\| S_T^{-1/2} \sum_{t=1}^{T} m^*_{T,t} \right\|_2 = O(1).$$

Similarly, approximation Lemma B.2.a coupled with CLT Lemma B.6 and the Helly-Bray theorem imply

$$\left\| \frac{1}{T^{1/2}} S_T^{-1/2} \sum_{t=1}^{T} m^*_{T,t} \left(\hat{I}^*_{T,t} - I^*_{T,t} \right) \right\|_2 = O(1).$$

Now imitate de Jong and Davidson's (2000: Lemmas A.2-A.3) arguments to deduce[4]

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

Next, consider the components of $\mathcal{A}_{1,T,\delta}$. It is straightforward to generalize approximation Lemma B.2.a to a weighted version with $k(t/\gamma_T)$ under K1. Specifically, define $N_T(\delta):=\min\{T,[\gamma_T/\delta]+1\}$ and note by construction and variance non-degeneracy I3

$$\limsup_{T\geq N}\frac{N_T(\delta)}{\gamma_T}\leq K \text{ and } \sup_{\gamma\in\Gamma}\left\{S_{N_T(\delta)}/N_T^{1/2}(\delta)\right\}\left\{S_T/T^{1/2}\right\}^{-1}=O(1).$$

Now use stationarity to deduce for any δ

$$T^{1/2}\max_{-T+1\leq t\leq 2T}\left\|\frac{1}{\gamma_T^{1/2}}S_T^{-1/2}\sum_{l=1-t}^{T-t}k\left(l/\gamma_T\right)\left\{\hat{m}^*_{T,t+l}-m^*_{T,t+l}\right\}I\left(0\leq l\leq[\gamma_T/\delta]\right)\right\|_2$$

$$\leq\left\|\frac{N_T^{1/2}(\delta)}{\gamma_T^{1/2}}\left\{S_{N_T(\delta)}/N_T^{1/2}(\delta)\right\}\left\{S_T/T^{1/2}\right\}^{-1}\right\|^{1/2}\times\left\|S_{N_T(\delta)}^{-1/2}\sum_{t=1}^{N_T(\delta)}k\left(t/\gamma_T\right)\left\{\hat{m}^*_{T,t}-m^*_{T,t}\right\}\right\|_2$$

$$\to 0 \text{ as } T\to\infty.$$

Similarly, by a straightforward generalization of CLT Lemma B.6 for any δ

$$T^{1/2}\max_{-T+1\leq t\leq 2T}\left\|\frac{1}{\gamma_T^{1/2}}\sum_{j=1-t}^{T-t}\eta_{\delta,T,j}\frac{1}{T^{1/2}}S_T^{-1/2}m^*_{T,t+j}I\left(0\leq j\leq[\gamma_T/\delta]\right)\right\|_2$$

$$\leq\left\|\frac{N_T^{1/2}(\delta)}{\gamma_T^{1/2}}\left\{S_{N_T(\delta)}/N_T^{1/2}(\delta)\right\}\left\{S_T/T^{1/2}\right\}^{-1}\right\|^{1/2}\times\left\|S_{[\gamma_T/\delta]}^{-1/2}\sum_{t=1}^{N_T(\delta)}\eta_{\delta,T,j}m^*_{T,t}\right\|_2+o\left(1\right)$$

$$\to 0 \text{ as } T\to\infty.$$

Therefore

(3) $$\lim_{\delta\to 0}\limsup_{T\to\infty}\left\|\mathcal{A}_{1,T,\delta}\right\|_1=0.$$

Combine (2) and (3) to conclude $\mathcal{A}_{1,T}=o_p(1)$.

Step 2: Note

$$\left\|S_T^{-1}\sum_{s,t=1}^{T}k_{T,s,t}\left\{\hat{m}^*_{T,s}(\hat{\theta}_T)\hat{m}^*_{T,t}(\hat{\theta}_T)-\hat{m}^*_{T,s}\hat{m}^*_{T,t}\right\}\right\|$$

$$\leq 2\left\|S_T^{-1}\sum_{s,t=1}^{T}k_{T,s,t}\left\{\hat{m}^*_{T,s}(\hat{\theta}_T)-\hat{m}^*_{T,s}\right\}\hat{m}^*_{T,t}\right\|$$

$$+\left\|S_T^{-1}\sum_{s,t=1}^{T}k_{T,s,t}\left\{\hat{m}^*_{T,s}(\hat{\theta}_T)-\hat{m}^*_{T,s}\right\}\left\{\hat{m}^*_{T,t}(\hat{\theta}_T)-\hat{m}^*_{T,t}\right\}\right\|.$$

[4] Define $\mathcal{Z}_{T,t}:=S_T^{-1/2}m^*_{T,t}$. Recall an array $\{y_{T,t},\Im_t\}$ forms an L_p-mixingale with geometric decay if $||y_{T,t}-E[y_{T,t}|\Im_{t+d}]||_p\leq e_{T,t}\rho^d$ and $||E[y_{T,t}]-E[y_{T,t}|\Im_{t-d}]||_p\leq e_{T,t}\rho^d$ for some constants $\{e_{T,t}\}$ and $\rho\in(0,1)$. See Andrews (1988), cf. McLeish (1975). Under geometric β-mixing $\{m^*_{T,t},\Im_t\}$ is an L_2-mixingale with constants $e_{T,t}$ hence $1=E(\sum_{t=1}^T\mathcal{Z}_{T,t})^2\leq K\sum_{t=1}^T e^2_{T,t}/S_T$ (McLeish 1975: Theorem 1.6, Lemma 2.1). We can therefore always assume under stationarity $e_{T,t}=KS_T^{1/2}/T^{1/2}$ for each t. Thus $\{\mathcal{Z}_{T,t},\Im_t\}$ is an L_2-mixingale with constants $K/T^{1/2}$, so de Jong and Davidson's arguments apply to our setting. See also the proof of Lemma B.5, below.

We will bound the first term, the second is similar. Use the Taylor expansion argument in the proof of expansion Lemma B.3.a to deduce for some $||\theta_{T,*} - \theta^0|| \leq ||\hat{\theta}_T - \theta^0||$

$$
\begin{aligned}
&\left\| S_T^{-1} \sum_{s,t=1}^{T} k_{T,s,t} \left\{ \hat{m}^*_{T,s}(\hat{\theta}_T) - \hat{m}^*_{T,s} \right\} \hat{m}^*_{T,t} \right\| \\
&\qquad \leq \left\| S_T^{-1} \sum_{s,t=1}^{T} k_{T,s,t} \hat{J}_{T,s}(\theta_{T,*}) \hat{m}^*_{T,t} \right\| \times \left\| \hat{\theta}_T - \theta^0 \right\| \\
&\qquad\qquad + \left\| S_T^{-1} \sum_{s,t=1}^{T} k_{T,s,t} J_s(\theta_{T,*}) \left\{ \hat{I}_{T,s}\left(\theta_{T,*}\right) - \hat{I}_{T,s} \right\} \hat{m}^*_{T,t} \right\| \times \left\| \hat{\theta}_T - \theta^0 \right\| \\
&\qquad\qquad + \left\| S_T^{-1} \sum_{s,t=1}^{T} k_{T,s,t} J_s(\theta_{T,*}) \left\{ \hat{I}_{T,s}\left(\hat{\theta}_T\right) - \hat{I}_{T,s} \right\} \hat{m}^*_{T,t} \right\| \times \left\| \hat{\theta}_T - \theta^0 \right\| \\
&\qquad\qquad + \left\| S_T^{-1} \sum_{s,t=1}^{T} k_{T,s,t} m_s \left\{ \hat{I}_{T,s}\left(\hat{\theta}_T\right) - \hat{I}_{T,s} \right\} \hat{m}^*_{T,t} \right\| \\
&\qquad = \sum_{i=1}^{4} \mathcal{B}_{i,T}.
\end{aligned}
$$

The gist of de Jong and Davidson's (2000: p. 419-420) Fourier inversion argument applies. Extend their equation (A.51) to our environment to obtain

$$
\begin{aligned}
\mathcal{B}_{1,T} &\leq K \int_{-\infty}^{\infty} \left(\|J_T\|^{-1} \left\| \frac{1}{T} \sum_{s=1}^{T} e^{-i\xi s/\gamma_T} \hat{J}_{T,s}(\theta_{T,*}) \right\| \times \left\| T^{-1/2} S_T^{-1/2} \sum_{t=1}^{T} e^{i\xi t/\gamma_T} \hat{m}^*_{T,t} \right\| \right) \left|\varpi\left(\xi\right)\right| d\xi \\
&= K \int_{-\infty}^{\infty} \mathcal{C}_T\left(\xi\right) \mathcal{D}_T\left(\xi\right) \left|\varpi\left(\xi\right)\right| d\xi,
\end{aligned}
$$

where $\varpi(\xi)$ is defined under K1. Lemma B.2.a and Lemma B.6 render $\mathcal{D}_T\left(\xi\right) = O_p(1)$. Further, Jacobian consistency Lemma B.4 with $||\theta_{T,*} - \theta^0|| \leq ||\hat{\theta}_T - \theta^0|| = O_p(T^{-1/2}||S_T||^{1/2} \times ||J_T||^{-1})$ under P1 or P2, and K1 properties $\sum_{s,t=1}^{T} |k_{T,s,t}| = o(T^2)$, $\max_{1 \leq s \leq T} \sum_{t=1}^{T} |k_{T,s,t}| = o(T)$ and $\gamma_T = o(T)$ imply $\mathcal{C}_T\left(\xi\right) = o_p(1)$. Therefore $\int_{-\infty}^{\infty} \mathcal{C}_T\left(\xi\right) \mathcal{D}_T\left(\xi\right) \left|\varpi\left(\xi\right)\right| d\xi = o_p\left(1\right)$ by dominated convergence and K1. Similar arguments extend to the remaining terms. $\mathcal{QED}$.

LEMMA B.3 (*expansions*). *Under D1-D6:*

a. $m^*_T(\theta) = m^*_T(\tilde{\theta}) + J^*_T(\theta_*)(\theta - \tilde{\theta}) + r_T \times o_p(1)$ *and* $\hat{m}^*_T(\theta) = \hat{m}^*_T(\tilde{\theta}) + \hat{J}^*_T(\theta_*)(\theta - \tilde{\theta}) + r_T \times ||\theta - \tilde{\theta}||^{1/\iota} \times o_p(1)$ *for* $||\theta_* - \theta|| \leq ||\theta - \tilde{\theta}||$ *that may be different in different in each case, and tiny* $\iota > 0$.

b. $E[m^*_{T,t}(\theta)] - E[m^*_{T,t}(\tilde{\theta})] = J_T(\tilde{\theta})(\theta - \tilde{\theta}) + o(||J_T(\tilde{\theta})|| \times ||\theta - \tilde{\theta}||)$ *for any* $\theta, \tilde{\theta} \in \Theta$.

PROOF.

Claim (a): Assume θ and $m_t(\theta)$ are scalars and $m_t(\theta)$ is symmetrically trimmed to simplify notation.

We only expand $m^*_T(\theta)$ since $\hat{m}^*_T(\theta)$ is similar. Write $m^*_{T,t}(\theta) = m_t(\theta) \times I_{T,t}(\theta)$ where $I_{T,t}(\theta) = I(|m_t(\theta)| \leq c_T(\theta))$, and choose $||\theta - \tilde{\theta}|| \leq \delta$ for any $\delta > 0$. Use differentiability D2 to deduce by Taylor's theorem

$$m^*_{T,t}(\theta) = \left\{ m_t(\tilde{\theta}) + J_t(\theta_{T,\delta})(\theta - \tilde{\theta}) \right\} \times I_{T,t}\left(\theta\right)$$

where $||\theta_{T,\delta} - \tilde{\theta}|| \leq ||\theta - \tilde{\theta}||$, and $J_t(\theta) := (\partial/\partial\theta)m_t(\theta)$. Therefore

$$
\begin{aligned}
(4) \qquad m_T^*(\theta) - m_T^*(\tilde{\theta}) &= J_T^*(\theta_{T,\delta}) \times (\theta - \tilde{\theta}) + \frac{1}{T}\sum_{t=1}^{T} m_t(\theta) \times \left\{ I_{T,t}(\theta) - I_{T,t}(\tilde{\theta}) \right\} \\
&\quad + \frac{1}{T}\sum_{t=1}^{T} J_t(\theta_{T,\delta}) \times \left\{ I_{T,t}(\theta) - I_{T,t}(\theta_{T,\delta}) \right\} \times (\theta - \tilde{\theta}).
\end{aligned}
$$

We will show the second and third terms are $o_p(r_T \times ||\theta - \tilde{\theta}||^{1/\iota})$.

Consider the second term in (4) and use $I_{T,t}(\theta) - I_{T,t}(\tilde{\theta}) \in \{-1, 0, 1\}$ to bound

$$
\begin{aligned}
\left| \frac{1}{T}\sum_{t=1}^{T} m_t(\theta) \left\{ I_{T,t}(\theta) - I_{T,t}(\tilde{\theta}) \right\} \right| &\leq \frac{1}{T^{1/2}}\sum_{t=1}^{T} \left| m_t(\theta) \left\{ I_{T,t}(\theta) - I_{T,t}(\tilde{\theta}) \right\} \right| \\
&\quad \times \frac{1}{T^{1/2}}\sum_{t=1}^{T} \left| I_{T,t}(\theta) - I_{T,t}(\tilde{\theta}) \right| = A_T(\theta, \tilde{\theta}) \times B_T(\theta, \tilde{\theta}).
\end{aligned}
$$

The threshold construction (1), $I_{T,t}(\theta) \in \{0, 1\}$ and triangle inequality imply for any $p > 0$

$$
\sup_{\theta, \tilde{\theta} \in \Theta} E\left| I_{T,t}(\theta) - I_{T,t}(\tilde{\theta}) \right|^p = O(k_T/T)
$$

where $O(\cdot)$ is not a function of θ. Combined with D1.i continuity and boundedness of the finite dimensional distributions of $m_t(\theta)$ and the mean-value-theorem, it follows $E|I_{T,t}(\theta) - I_{T,t}(\tilde{\theta})|^p = O((k_T/T)) \times ||\theta - \tilde{\theta}||$. Now invoke stationarity D1.i, envelope bound D4 and the Cauchy-Schwartz inequality to deduce for tiny $\iota > 0$

$$
\left(E\left[A_T(\theta, \tilde{\theta})^\iota \right] \right)^{1/\iota} \leq T^{1/2} \left[E\left| m_t(\theta) \left\{ I_{T,t}(\theta) - I_{T,t}(\tilde{\theta}) \right\} \right|^\iota \right]^{1/\iota} = O\left(T^{1/2} (k_T/T)^{1/\iota} \right) \times \left\| \theta - \tilde{\theta} \right\|^{1/\iota}.
$$

Since $\iota > 0$ can be chosen arbitrarily small and $k_T/T \to 0$ by tail trimming, invoke Markov's inequality to conclude for some $r_T \to 0$ arbitrarily fast and $o_p(\cdot)$ not a function of θ

$$
A_T(\theta, \tilde{\theta}) = o_p\left(T^{1/2} \left(k_T^{1/2}/T \right)^{1/\iota} \left\| \theta - \tilde{\theta} \right\|^{1/\iota} \right) = o_p(r_T) \times \left\| \theta - \tilde{\theta} \right\|^{1/\iota}.
$$

Since $E|B_T(\theta, \tilde{\theta})| \leq T^{1/2}$ follows trivially from $|I_{T,t}(\theta) - I_{T,t}(\tilde{\theta})| \in \{0, 1\}$ we have shown for some $r_T \to 0$ arbitrarily fast

$$
\left| \frac{1}{T}\sum_{t=1}^{T} m_t(\theta) \left\{ I_{T,t}(\theta) - I_{T,t}(\tilde{\theta}) \right\} \right| \leq A_T(\theta, \tilde{\theta}) \times B_T(\theta, \tilde{\theta}) = o_p(r_T) \times \left\| \theta - \tilde{\theta} \right\|^{1/\iota}.
$$

Repeat the argument for the third term in (4) by invoking envelope bound D4 for $J_t(\theta)$.

Claim (b): The claim follows from the definition of a derivative. $\mathcal{QED}$.

LEMMA B.4 (*Jacobian*). *Under D1-D6, and P1 or P2* $\hat{J}_T^*(\hat{\theta}_T) = J_T(1 + o_p(1))$.

PROOF. Recall $J_T = J_T(\theta^0) = (\partial/\partial\theta)E[m_{T,t}^*(\theta)]|_{\theta^0}$ and write $\hat{m}_T^*(\theta) = 1/T\sum_{t=1}^{T} \hat{m}_{T,t}^*(\theta)$.

Denote by $e_i \in \mathbb{R}^r$ the unit vector (e.g. $e_2 = [0, 1, 0, ..., 0]'$), define a sequence of bounded positive numbers $\{\varepsilon_T\}$ that satisfies $\liminf_{T \geq 1} \varepsilon_T ||J_T|| > 0$ and $||\tilde{\theta}_T - \theta^0||/\varepsilon_T \xrightarrow{p} 0$. This is always possible in lieu of the plug-in rate and Lemma B.1.c: $||\tilde{\theta}_T - \theta^0||/\varepsilon_T = O_p(T^{-1}||S_T||^{1/2}) = o_p(1)$. Define

$$
\check{J}_{i,j,T}^*(\theta, \varepsilon_T) := \frac{1}{2\varepsilon_T} \times \frac{1}{T}\sum_{t=1}^{T} \left\{ \hat{m}_{j,T,t}^*(\theta + e_i\varepsilon_T) - \hat{m}_{j,T,t}^*(\theta - e_i\varepsilon_T) \right\}.
$$

Minkowski's inequality implies for arbitrary θ

$$\left\| \hat{J}_T^*(\tilde{\theta}_T) - J_T \right\| \leq \left\| \hat{J}_T^*(\tilde{\theta}_T) - \check{J}_T^*(\theta, \varepsilon_T) \right\| + \left\| \check{J}_T^*(\theta, \varepsilon_T) - J_T \right\|$$

Apply asymptotic expansion Lemma B.3.a to deduce for some $\tilde{\theta}_{T,*} \in \{\tilde{\theta}_T - e_i\varepsilon_T, \tilde{\theta}_T + e_i\varepsilon_T\}$

$$\hat{J}_T^*(\tilde{\theta}_T) = \check{J}_{i,j,T}^*(\tilde{\theta}_{T,*}, \varepsilon_T) + o_p\left(\|J_T\|\right), \quad \text{hence} \quad \left\| \hat{J}_T^*(\tilde{\theta}_T) - \check{J}_T^*(\theta, \varepsilon_T) \right\| = o_p\left(\|J_T\|\right).$$

Since $||\tilde{\theta}_{T,*} - \theta^0|| \leq ||\tilde{\theta}_T - \theta^0|| = o_p(1)$ it remains to show $||\check{J}_T^*(\tilde{\theta}_T, \varepsilon_T) - J_T|| = o_p(||J_T||)$ for any $||\tilde{\theta}_T - \theta^0|| \xrightarrow{p} 0$. Define

$$U^0\left(\delta_1, \delta_2\right) := \left\{ \theta \in \Theta : \delta_1 \leq \left\|\theta - \theta^0\right\| \leq \delta_2 \right\} \text{ for } 0 \leq \delta_1 \leq \delta_2$$

$$\mathcal{J}_T\left(\delta_1, \delta_2\right) := \sup_{\theta \in U^0(\delta_1, \delta_2)} \left\{ \frac{\|J_T^*(\theta) - J_T\|}{\|J_T\|} \right\}$$

Stochastic differentiability Lemma C.2 and the fact that $U^0(\delta_1, \delta_2) \subseteq U^0\left(0, \delta_2\right)$, and consistency $\tilde{\theta}_T \xrightarrow{p} \theta^0$ imply for large K and any non-zero constant vector $a \in \mathbb{R}^r/0$

$$\begin{aligned}
&\left\| \left\{ \hat{m}_T(\tilde{\theta}_T + a\varepsilon_T) - \hat{m}_T \right\} - \left\{ E\left[m_{T,t}^*\left(\tilde{\theta}_T + a\varepsilon_T \right) \right] - E\left[m_{T,t}^* \right] \right\} \right\| \\
&\qquad \leq K \left\{ 1 + \|J_T\| \times \left\| \tilde{\theta}_T + a\varepsilon_T - \theta^0 \right\| \right\} \times o_p\left(1\right) \times \left(\mathcal{J}_T\left(\delta_1, \delta_2\right) + o_p\left(1\right) \right) \\
&\qquad \leq K \left\{ 1 + \|J_T\| \times \left\| \tilde{\theta}_T - \theta^0 \right\| + \|J_T\| \times \|a\varepsilon_T\| \right\} \times \left(\mathcal{J}_T\left(\delta_1, \delta_2\right) + o_p\left(1\right) \right) \\
&\qquad = o_p\left(\varepsilon_T \|J_T\|\right) + O_p\left(\varepsilon_T \|J_T\| \times \mathcal{J}_T\left(\delta_1, \delta_2\right)\right).
\end{aligned}$$

Similarly, by differentiability of $E[m_{T,t}^*(\theta)]$,

$$\begin{aligned}
&\left\| \frac{E\left[m_{T,t}^*(\tilde{\theta}_T + a\varepsilon_T) \right] - E\left[m_{T,t}^* \right]}{\varepsilon_T} - aJ_T \right\| \\
&\qquad = \left\| J_T\varepsilon_T^{-1}\left(\tilde{\theta}_T + a\varepsilon_T - \theta^0 \right) - aJ_T + o_p\left(\|J_T\| \varepsilon_T^{-1} \left(\tilde{\theta}_T + \varepsilon_T - \theta^0 \right) \right) \right\| \\
&\qquad = \left\| J_T\varepsilon_T^{-1}\left(\tilde{\theta}_T - \theta^0 \right) \right\| + o_p\left(\|J_T\|\right) = o_p\left(\|J_T\|\right).
\end{aligned}$$

Replace $\tilde{\theta}_T + a\varepsilon_T$ with $\tilde{\theta}_T - a\varepsilon_T$ to deduce the same bounds. Therefore

$$\left\| \check{J}_T^*(\tilde{\theta}_T, \varepsilon_T) - J_T \right\| = \left\| \frac{\hat{m}_T^*(\tilde{\theta}_T + \varepsilon_T) - \hat{m}_T^*(\tilde{\theta}_T - \varepsilon_T)}{2\varepsilon_T} - J_T \right\| = o_p\left(\|J_T\|\right) + O_p\left(\|J_T\| \times \mathcal{J}_T(\delta_1, \delta_2)\right)$$

hence we have shown $\hat{J}_T^*(\tilde{\theta}_T) = J_T(1 + o_p(1)) + O_p(||J_T|| \times \mathcal{J}_T(\delta_1, \delta_2))$.

Since $0 \leq \delta_1 < \delta_2$ are arbitrary, the proof is complete if we show for some sequence of positive numbers $\{\delta_{1,T}\}$, $\delta_{1,T} \to 0$ and $\delta_{2,T} = 2\delta_{1,T}$:

$$\mathcal{J}_T(\delta_{1,T}, \delta_{2,T}) \xrightarrow{p} 0.$$

Define

$$\mathfrak{m}_T(\delta_1, \delta_2) = \sup_{\theta \in U^0(\delta_1, \delta_2)} \left\| E\left[m_{T,t}^*\left(\theta\right) \right] \right\|.$$

Now, for each $\theta \in U^0(\delta)$ we can always find a sequence $\{\theta_{T,\delta}\} \in U^0(\delta_1, \delta_2)$, $\theta_{T,\delta} \neq \theta^0$ for each finite $T \geq N$, such that

$$\begin{aligned}
\frac{E\left[m^*_{T,t}(\theta_{T,\delta})\right] - E\left[m^*_{T,t}\right]}{\left\|\theta_{T,\delta} - \theta^0\right\|} &= \frac{m^*_T(\theta_{T,\delta}) - m^*_T}{\left\|\theta_{T,\delta} - \theta^0\right\|} + o_p(1) \times \frac{\mathfrak{m}_T(\delta_1, \delta_2)}{\left\|\theta_{T,\delta} - \theta^0\right\|} \\
&= J^*_T(\theta) \times \frac{\left(\theta_{T,\delta} - \theta^0\right)}{\left\|\theta_{T,\delta} - \theta^0\right\|} \times (1 + o_p(1)) + o_p(1) \times \frac{\mathfrak{m}_T(\delta_1, \delta_2)}{\left\|\theta_{T,\delta} - \theta^0\right\|},
\end{aligned}$$

where each $o_p(1)$ term does not depend on θ. Moreover, by moment expansion Lemma B.3.b

$$\frac{E\left[m^*_{T,t}(\theta_{T,\delta})\right] - E\left[m^*_{T,t}\right]}{\left\|\theta_{T,\delta} - \theta^0\right\|} = J_T \times \frac{\left(\theta_{T,\delta} - \theta^0\right)}{\left\|\theta_{T,\delta} - \theta^0\right\|} \times (1 + o(1)).$$

Further, by construction $||\theta_{T,\delta} - \theta^0|| \geq \delta_{2,T}/2$. Together it follows

$$\sup_{\theta \in U^0(\delta)} \left\{\frac{\left\|J^*_T(\theta) - J_T\right\|}{\|J_T\|}\right\} = o_p(1) + o_p\left(\frac{\mathfrak{m}_T(\delta_1, \delta_2)}{\delta_{2,T}\|J_T\|}\right).$$

Therefore $\mathcal{J}_T(\delta_{1,T}, \delta_{2,T}) \xrightarrow{p} 0$ if $\mathfrak{m}_T(\delta_{1,T}, \delta_{2,T})/[\delta_{2,T}||J_T||] = O(1)$. By the definition of a derivative, the construction $U^0(\delta_{1,T}, \delta_{2,T}) \subseteq U^0(0, \delta_{2,T}) = U^0(\delta_{2,T})$ and moment smoothness I4

$$\mathfrak{m}_T(\delta_{1,T}, \delta_{2,T}) \leq K\delta_{2,T} \sup_{\theta \in U^0(\delta_{2,T})} \|J_T(\theta)\| \times (1 + o(1))$$

Now invoke Jacobian smoothness D5.ii to conclude

$$\frac{\mathfrak{m}_T(\delta_1, \delta_2)}{\delta_{2,T}\|J_T\|} \leq \frac{K\delta_{2,T}\|J_T\|(1 + o(1)) + o(1)}{\delta_{2,T}\|J_T\|} = O(1). \ \mathcal{QED}.$$

LEMMA B.5 (*HAC estimator*). *Under D1-D6, K1, I3, and P1 or P2* $\hat{S}_T(\hat{\theta}_T) = S_T(1 + o_p(1))$ *and* $\hat{S}_T(\hat{\theta}_T) = \check{S}_T(\hat{\theta}_T)(1 + o_p(1))$.

PROOF. We will only prove $\hat{S}_T(\hat{\theta}_T) = S_T(1 + o_p(1))$, the remaining claim being similar. Define $\hat{\mu}^*_{T,t}(\theta) := \hat{m}^*_{T,t}(\theta) - \hat{m}^*_T(\theta)$, $\mu^*_{T,t}(\theta) := m^*_{T,t}(\theta) - m^*_T(\theta)$ and

$$\mathcal{A}_{1,T} := S_T^{-1} \sum_{s,t=1}^{T} k_{T,s,t}\left\{\hat{\mu}^*_{T,s}(\hat{\theta}_T)\hat{\mu}^*_{T,t}(\hat{\theta}_T) - \mu^*_{T,s}\mu^*_{T,t}\right\} \quad \text{and} \quad \mathcal{A}_{2,T} := S_T^{-1} \sum_{s,t=1}^{T} k_{T,s,t}\mu^*_{T,s}\mu^*_{T,t} - I_q.$$

By the triangle inequality we must show each $\mathcal{A}_{i,T}(\gamma) \xrightarrow{p} 0$. Uniform cross-product approximation Lemma B.2.d states $\mathcal{A}_{1,T} \xrightarrow{p} 0$.

Next, we apply Theorem 2.1 of de Jong and Davidson (2000) to prove $\mathcal{A}_{2,T} \xrightarrow{p} 0$. It suffices to verify their Assumptions 1-3. de Jong and Davidson's Assumption 1 holds by K1. Now define $\mathcal{Z}_{T,t} := S_T^{-1}m^*_{T,t}$ and note under geometric β-mixing D3 $\{m^*_{T,t}, \Im_t\}$ forms a geometric L_2-mixingale with constants $e_{T,t}$ (cf. McLeish 1975: Theorem 2.1). Therefore $\{\mathcal{Z}_{T,t}, \Im_t\}$ forms a geometric L_2-mixingale with constants $\mathcal{E}_{T,t} := S_T^{-1/2}e_{T,t}$, hence $E(\sum_{t=1}^{T} \mathcal{Z}^2_{T,t}) \leq KS_T^{-1}\sum_{t=1}^{T} e^2_{T,t}$ by Theorem 1.6 in McLeish (1975). Since $E(\sum_{t=1}^{T} \mathcal{Z}^2_{T,t}) = 1$, under stationarity may can always assume $e_{T,t} = KS_T^{1/2}/T^{1/2}$ hence $\{\mathcal{Z}_{T,t}, \Im_t\}$ has constants $\mathcal{E}_{T,t} = K/T^{1/2}$. Therefore de Jong and Davidson's Assumption 2 holds. Finally, their Assumption 3 states $\gamma_T \times \max_{1 \leq t \leq T}\{\mathcal{E}^2_{T,t}\} = o(1)$ which holds given $\mathcal{E}_{T,t} = K/T^{1/2}$ and $\gamma_T = o(T)$ under K1. $\mathcal{QED}$.

C.3. PROOF OF LEMMA B.6 Recall the claim.

LEMMA B.6 (*clt*). *Under D1, D3 and I3* $r'S_T^{-1/2}\sum_{t=1}^T\{m_{T,t}^* - E[m_{T,t}^*]\} \xrightarrow{d} N(0,1)$ *for any conformable* $r'r = 1$. *If P2 also holds then* $r'\mathfrak{S}_T^{-1/2}\sum_{t=1}^T\{\mathcal{M}_{T,t}^* - E[\mathcal{M}_{T,t}^*]\} \xrightarrow{d} N(0,1)$.

Define for conformable $\lambda'\lambda = 1$

$$z_T(\lambda) = \sum_{t=1}^T z_{T,t}(\lambda) = \lambda' S_T^{-1/2} \sum_{t=1}^T \left\{m_{T,t}^* - E[m_{T,t}^*]\right\},$$

and note

$$|z_{T,t}(\lambda)| \leq \max_{1\leq i\leq q}\{c_{i,T}\} \times \|S_T\|^{-1/2}. \tag{5}$$

The first claim follows from the Cramér-Wold theorem if we prove $\sum_{t=1}^T z_{T,t}(\lambda) \xrightarrow{d} N(0,1)$. The second claim follows instantly from the first and P2.c.

Notice by (5) $z_T(\lambda)$ is bounded exactly like a scalar self-normalized tail-trimmed random variable. In particular, the following does not require λ and is not sensitive to symmetric or asymmetric trimming. Thus, to ease notation assume $q = 1$, $E[m_{T,t}^*] = 0$ and symmetric trimming, hence in compact notation

$$z_{T,t} = m_{T,t}^*/S_T^{1/2} = m_{T,t}/\sigma_T \text{ where } m_{T,t} := m_t I\left(|m_t| \leq c_T\right) \text{ and } \sigma_T := S_T^{1/2}.$$

If $E[m_t^2] < \infty$ the claim follows from extant results for mixing arrays (e.g. de Jong 1997, Hill 2009). Therefore assume $E[m_t^2] = \infty$, hence the tail index $\kappa \leq 2$ in Assumption D.1. The following proof borrows heavily from arguments presented in Hill (2010) based on a well known martingale difference approximation argument dating to Gordin (1969, 1973) and McLeish (1974, 1975).

Define sequences of positive real numbers $\{h_T,\, j_T,\, r_T\}$ used to form telescoping or Bernstein sums of $z_{T,t}$: assume $h_T \to \infty$, $j_T \to \infty$, and

$$r_T = [T/h_T],\ 1 \leq\ j_T \leq h_T \text{ and } j_T = o\left(h_T\right). \tag{6}$$

Now define an array of σ-fields $\{\mathcal{F}_{T,i}\}$: for any sequence of positive finite numbers $\{g_T\}$, $g_T \to \infty$,

$$\mathcal{F}_{T,i} := \sigma\left(m_t : -g_T \leq t \leq ih_T\right) = \Im_{-g_T}^{ih_T}, \text{ for } i = 1,...,r_T,$$

and define telescoping sums of $z_{T,t}$, and a martingale difference:

$$\mathcal{Z}_{T,i} := \sum_{t=(i-1)h_T+j_T+1}^{ih_T} z_{T,t} \text{ and } \mathcal{W}_{T,i} := E\left[\mathcal{Z}_{T,i}|\mathcal{F}_{T,i}\right] - E\left[\mathcal{Z}_{T,i}|\mathcal{F}_{T,i-1}\right].$$

Thus h_T is the block size, r_T the number of blocks, and j_T the buffer between blocks. Trivially we have the decomposition

$$\begin{aligned}\sum_{t=1}^T z_{T,t} &= \sum_{i=1}^{r_T}\mathcal{W}_{T,i} + \sum_{i=1}^{r_T}\left(\mathcal{Z}_{T,i} - E\left[\mathcal{Z}_{T,i}|\mathcal{F}_{T,i}\right]\right) + \sum_{i=1}^{r_T} E\left[\mathcal{Z}_{T,i}|\mathcal{F}_{T,i-1}\right] \\ &\quad + \sum_{i=1}^{r_T}\sum_{t=(i-1)h_T+1}^{(i-1)h_T+j_T} z_{T,t} + \sum_{t=r_T h_T+1}^{T} z_{T,t}.\end{aligned} \tag{7}$$

Since g_T is finite and m_t is geometrically β-mixing, $\mathcal{F}_{T,i}$-measurable $\mathcal{W}_{T,i}$ is both geometrically β-mixing and an adapted martingale difference.

Decomposition (7) with sequence properties (6) are well known devices for dependent process CLT's (Gordin 1969, 1973, McLeish 1974, 1975, de Jong 1997). Under (6), however, we can otherwise restrict the sequences $\{j_T, h_T\}$ anyway we choose without loss of generality. Under tail-trimming it helps to restrict the block size h_T in order to bound dispersion in $\mathcal{Z}_{T,i}$ and therefore $\mathcal{W}_{T,i}$:

$$h_T \to \infty \text{ such that } h_T = o\left(\min\left\{k_T^{1/2}, T^{\delta}\right\}\right) \text{ for tiny } \delta > 0. \tag{8}$$

Next, define functions of $\mathcal{W}_{T,i}$

$$\mathcal{U}_{T,i} := \sigma_T \mathcal{W}_{T,i}, \text{ and } \mathcal{K}_T \sim K_0 T^{\delta} c_T^2 \text{ for some } K_0 > 0 \text{ and tiny } \delta > 0,$$

a truncation function

$$\widetilde{\mathcal{U}}_{T,i}^2(K) := \mathcal{U}_{T,i}^2 I\left(\mathcal{U}_{T,i}^2 \leq K\right) = \sigma_T^2 \mathcal{W}_{T,i}^2 I\left(\sigma_T^2 \mathcal{W}_{T,i}^2 \leq K\right) \text{ and } \widetilde{\mathcal{U}}_{T,i}^2 = \widetilde{\mathcal{U}}_{T,i}^2(\mathcal{K}_T)$$

and index sets

$$\mathcal{I}_{T,i}^* := \{t : t \in [(i-1)h_T + j_T + 1, ..., ih_T]\} \text{ and } \mathcal{I}_T := \left\{t : t \in \bigcup\nolimits_{i=1}^{r_T} \mathcal{I}_{T,i}^*\right\}.$$

Since $\mathcal{W}_{T,i}$ is $\mathcal{F}_{T,i}$-measurable and $\mathcal{U}_{T,i}$ is a Borel function, $\mathcal{U}_{T,i}$ is β-mixing.

The required limit follows instantly from Lemmas C.5 and C.6 below: $\sum_{t=1}^T z_{T,t} = \sum_{i=1}^{r_T} \mathrm{W}_{T,i} + o_p(1) \xrightarrow{d} N(0,1)$. Lemmas C.5 and C.6 require two preliminary mixingale properties. The following assumption merely formalizes (6) and (8) which are imperative for proving the decomposition and *mds* CLT.

ASSUMPTION B1 (*Bernstein sequences*). *Let $1 \leq h_T \leq T$, $1 \leq j_T \leq h_T$, $j_T \to \infty$ and $j_T/h_T \to 0$. Further $h_T \to \infty$ such that $h_T = o(\min\{k_T^{1/2}, T^{\delta}\})$ for tiny $\delta > 0$.*

Lemma C.3 (*mixingale*). *Under D3 $\{m_{T,t}, \Im_t\}$ forms a geometric L_2-mixingale array with constants $e_{T,t} \leq K\sigma_T/T^{1/2}$ (McLeish 1975): $||E[m_{T,t}] - E[m_{T,t}|\Im_{t-d}]||_2 \leq e_{T,t}\zeta_d$ and $||m_{T,t} - Em[_{T,t}|\Im_{t+d}]||_2 \leq e_{T,t}\zeta_{d+1}$, where $\zeta_d = O(\rho^d)$, $\rho \in (0,1)$.*

Lemma C.4 (*mixingale*). *Under B1 and D3 $\{m_{T,t} - E[m_{T,t}|\mathcal{F}_{T,i}], \Im_t\}_{t \in \mathcal{I}_T}$ and $\{E[m_{T,t}|F_{T,i-1}], \Im_t\}_{t \in \mathcal{I}_T}$ form geometric L_2-mixingale arrays with constants $e_{T,t}\zeta_{j_T}^{\iota}$ and coefficients $\zeta_{j_T}^{1-\iota}$, where $e_{T,t}$ and ζ_{j_T} are defined by Lemma C.3.*

Lemma C.5 (*decomposition*). *Under B1, D1, D3 and I3 $\sum_{t=1}^T z_{T,t} = \sum_{i=1}^{r_T} \mathrm{W}_{T,i} + o_p(1)$. In particular a. $\sum_{t=r_T h_T+1}^T z_{T,t} \xrightarrow{p} 0$; b. $\sum_{i=1}^{r_T} \sum_{t=(i-1)h_T+1}^{(i-1)h_T+j_T} z_{T,t} \xrightarrow{p} 0$; c. $\sum_{i=1}^{r_T} E[\mathcal{Z}_{T,i}|\mathcal{F}_{T,i-1}] \xrightarrow{p} 0$; and d. $\sum_{i=1}^{r_T}(\mathcal{Z}_{T,i} - E[\mathcal{Z}_{T,i}|\mathcal{F}_{T,i}]) \xrightarrow{p} 0$.*

Lemma C.6 (*mds clt*). *Under B1, D1, D3 and I3 $\sum_{i=1}^{r_T} \mathcal{W}_{T,i} \xrightarrow{d} N(0,1)$.*

C.4. PROOFS OF LEMMAS C.1-.C.6 We repeatedly use an implication of Meng and Lin's (2009) array version of McLeish's (1975: Theorem 1.6) maximal inequality for geometric L_2-mixingale arrays $\{y_{T,t}, \Im_t\}$ with constants $e_{T,t}$:

$$E\left(\sum_{t=1}^T y_{T,t}\right)^2 \leq K \sum_{t=1}^T e_{T,t}^2. \tag{9}$$

PROOF OF LEMMA C.1. Under power law tail decay D1.ii Karamata's Theorem applies:

$$\sup_{\theta}\left\{\frac{\max_{1\leq i\leq q}\{c_{i,T}(\theta)\}}{\|\Sigma_T(\theta)\|^{1/2}}\right\} \leq K \times \sup_{\theta}\left\{\frac{\max_{1\leq i\leq q}\{c_{i,T}(\theta)\}}{\left(\sum_{i=1}^q c_{i,T}^2(\theta)(k_{i,T}/T)\right)^{1/2}}\right\} = O\left(\frac{T^{1/2}}{\min_{1\leq i\leq q}\{k_{i,T}\}^{1/2}}\right).$$

The second bound $\sup_\theta\{c_T(\theta)/||S_T(\theta)||^{1/2}\} = O(1/\min_{1\leq i\leq q}\{k_{i,T}\}^{1/2})$ follows from the first and Lemma B.1.a. $\mathcal{QED}$.

Recall

$$\check{S}_T(\theta) := \sum_{s,t=1}^{T} E\left[\left\{\hat{m}^*_{T,s}(\theta) - E[\hat{m}^*_{T,s}(\theta)]\right\}\left\{\hat{m}^*_{T,s}(\theta) - E[\hat{m}^*_{T,s}(\theta)]\right\}'\right].$$

PROOF OF LEMMA C.2. Apply Minkowski's inequality and the Lemma B.2.c uniform approximation to obtain

$$\begin{aligned}
&\sup_{\theta\in U^0(\delta)}\left\{\frac{\left\|\{\hat{m}^*_T(\theta) - \hat{m}^*_T\} - \left\{E\left[m^*_{T,t}(\theta)\right] - E\left[m^*_{T,t}\right]\right\}\right\|}{1 + \|J_T\| \times \left\|\theta - \theta^0\right\|}\right\} \\
&\quad\leq \sup_{\theta\in U^0(\delta)}\left\{\frac{\left\|\{m^*_T(\theta) - m^*_T\} - \left\{E\left[m^*_{T,t}(\theta)\right] - E\left[m^*_{T,t}\right]\right\}\right\|}{1 + \|J_T\| \times \left\|\theta - \theta^0\right\|}\right\} \\
&\qquad + 2\sup_{\theta\in U^0(\delta)}\left\{\frac{\|\hat{m}^*_T(\theta) - m^*_T(\theta)\|}{1 + \|J_T\| \times \left\|\theta - \theta^0\right\|}\right\} \\
&\quad= \sup_{\theta\in U^0(\delta)}\left\{\frac{\left\|\{m^*_T(\theta) - m^*_T\} - \left\{E\left[m^*_{T,t}(\theta)\right] - E\left[m^*_{T,t}\right]\right\}\right\|}{1 + \|J_T\| \times \left\|\theta - \theta^0\right\|}\right\} + o_p(1)
\end{aligned}$$

Equation and moment expansions Lemma B.3.a,b imply the last line is bounded by $\sup_{\theta\in U^0(\delta)}\{||J^*_T(\theta) - J_T||/||J_T||\} + o_p1)$. $\mathcal{QED}$.

PROOF OF LEMMA C.3. The mixingale claim follows from Lemma 2.1 in McLeish (1975). Use (9) to deduce trivially $1 = E(\sum_{t=1}^T m_{T,t}/\sigma_T)^2 \leq K\sum_{t=1}^T e^2_{T,t}/\sigma^2_T$, hence by stationarity we can always set $e_{T,t} = K\sigma_T/T^{1/2}$. $\mathcal{QED}$.

PROOF OF LEMMA C.4. Since $\Im_t$ is increasing it follows $\mathcal{F}_{T,i-1} = \Im^{(i-1)h_T}_{-g_T} \subset \Im^{t-j_T}_{-\infty}$ for each $1 \leq i \leq r_T$ and all $(i-1)h_T + j_T + 1 \leq t \leq ih_T$, and $\Im^{t-j_T}_{-\infty} \subset \Im^{t-d}_{-\infty}$ for sufficiently large T since $j_T \rightarrow \infty$. Apply Lemma C.3, Jensen's inequality, $j_T \geq d$ for sufficiently large T since $j_T \rightarrow \infty$ under Assumption B1, and geometric memory to deduce

$$\begin{aligned}
\left\|E\left(E\left[m_{T,t}|\mathcal{F}_{T,i-1}\right]\right)|\Im^{t-d}_{-\infty}\right\|_2 &\leq \left\|E\left(E\left[m_{T,t}|\Im^{(i-1)h_T}_{-g_T}\right]\right)|\Im^{t-d}_{-\infty}\right\|_2 \\
&= \left\|E\left(E\left[E\left(m_{T,t}|\Im^{t-j_T}_{-\infty}\right)|\Im^{(i-1)h_T}_{-g_T}\right]\right)|\Im^{t-d}_{-\infty}\right\|_2 \\
&\leq K\left\|E\left[m_{T,t}|\Im^{t-j_T}_{-\infty}\right]\right\|_2 \leq e_{T,t} \times \zeta_{j_T} = Ke_{T,t}\zeta^{\iota}_{j_T} \times \zeta^{1-\iota}_{j_T},
\end{aligned}$$

where $\zeta^{1-\iota}_{j_T} = O(\rho^{(1-\iota)j_T})$. An identical argument applies to $\{m_{T,t} - E[m_{T,t}|\mathcal{F}_{T,i}], \Im_t\}_{t\in\mathcal{I}_T}$. $\mathcal{QED}$.

PROOF OF LEMMA C.5. By (7) we need only prove (a)-(d).

Claims (a) and (b): By Lemma C.3 (9) applies. Use the construction $r_Th_T \sim T$ and $e_{T,t} = K\sigma_T/T^{1/2}$ to deduce

$$E\left(\sum_{t=r_Th_T+1}^{T} z_{T,t}\right)^2 = O\left(\sum_{t=r_Th_T+1}^{T} e^2_{T,t}/\sigma^2_T\right) = O\left((1 - r_Th_T/T)\right) = o(1).$$

Similarly, in conjunction with Assumption B1

$$E\left[\left(\sum_{i=1}^{r_T}\sum_{t=(i-1)h_T+1}^{(i-1)h_T+j_T} z_{T,t}\right)^2\right] = O\left(\sum_{i=1}^{r_T}\sum_{t=(i-1)h_T+1}^{(i-1)h_T+j_T} e_{T,t}^2/\sigma_T^2\right) = O\left(r_T j_T/T\right) = O(j_T/h_T) = o(1).$$

Claims (a) and (b) now follow from Chebyshev's inequality.

Claims (c) and (d): Invoke $r_T h_T \sim T$, Lemma C.4, (9) and $\zeta_{j_T} \to 0$ since $j_T \to \infty$:

$$E\left[\left(\sum_{i=1}^{r_T}\{\mathcal{Z}_{T,i} - E[\mathcal{Z}_{T,i}|\mathcal{F}_{T,i}]\}\right)^2\right] = O\left(\sum_{i=1}^{r_T}\sum_{t=(i-1)h_T+j_T}^{ih_T}\frac{e_{T,t}^2}{\sigma_T^2}\zeta_{j_T}^{2\iota}\right) = O\left(\frac{r_T h_T}{T}\zeta_{j_T}^{2\iota}\right) = O\left(\zeta_{j_T}^{2\iota}\right) = o(1).$$

Similarly $E(\sum_{i=1}^{r_T} E[[\mathcal{Z}_{T,i}|\mathcal{F}_{T,i-1}])^2 = O((r_T h_T/T)\zeta_{j_T}^{2\iota}) = o(1)$. $\mathcal{QED}$.

PROOF OF LEMMA C.6. We will show $\max_{1\leq i\leq r_T}|\mathrm{W}_{T,i}|$ is $o_p(1)$ and uniformly L_2-bounded, and $\sum_{i=1}^{r_T}\mathrm{W}_{T,i}^2 \xrightarrow{p} 1$. The CLT then follows from Theorem 2.3 of McLeish (1974).

By construction and threshold bound Lemma C.1 $\max_{1\leq i\leq r_T}|\mathcal{W}_{T,i}| \leq Kh_T c_T/\sigma_T \leq Kh_T/k_T^{1/2}$ *a.s.*, where by Assumption B1 (8) holds: $h_T/k_T^{1/2} \to 0$. Therefore $\max_{1\leq i\leq r_T}|\mathcal{W}_{T,i}|$ is both $o_p(1)$ and uniformly L_2-bounded by dominated convergence.

Next, by the triangle inequality $|\sum_{i=1}^{r_T}\mathrm{W}_{T,i}^2 - 1| \leq \sum_{i=1}^{5}\mathcal{E}_{T,i}$ where

$$\begin{aligned}
\mathcal{E}_{T,1} &= \left|\frac{1}{\sigma_T^2}\sum_{i=1}^{r_T}\left(\widetilde{\mathcal{U}}_{T,i}^2 - \mathcal{U}_{T,i}^2\right)\right|,\ \mathcal{E}_{T,2} = \left|\frac{1}{\sigma_T^2}\sum_{i=1}^{r_T}\left(\widetilde{\mathcal{U}}_{T,i}^2 - E\left[\widetilde{\mathcal{U}}_{T,i}^2\right]\right)\right| \\
\mathcal{E}_{T,3} &= \frac{1}{\sigma_T^2}\sum_{i=1}^{r_T}\left|E\left[\widetilde{\mathcal{U}}_{T,i}^2\right] - E\left[\mathcal{U}_{T,i}^2\right]\right|,\ \mathcal{E}_{T,4} = \left|\sum_{i=1}^{r_T}\left(E\left[\mathcal{W}_{T,i}^2\right] - E\left[\mathcal{Z}_{T,i}^2\right]\right)\right|,\ \mathcal{E}_{T,5} = \left|\sum_{i=1}^{r_T}E\left[\mathcal{Z}_{T,i}^2\right] - 1\right|.
\end{aligned}$$

Use Steps 1-3 and Markov's inequality to deduce $\mathcal{E}_{T,i} = o_p(1)$ for $i = 1-4$, and $\mathcal{E}_{T,5} = o_p(1)$ by Step 4. Therefore $\sum_{i=1}^{r_T}\mathrm{W}_{T,i}^2 \xrightarrow{p} 1$.

Step 1 ($\sigma_T^{-2}\sum_{i=1}^{r_T}E|\widetilde{\mathcal{U}}_{T,i}^2 - \mathcal{U}_{T,i}^2| = o(1)$): By construction $\sigma_T^2\mathrm{W}_{T,i}^2 \leq Kh_T^2 c_T^2$. Use $r_T h_T \sim T$, $h_T = o(T^\delta)$ for tiny $\delta > 0$ by Assumption B1, $T/\sigma_T^2 = O(1)$ by Lemma B.1, $\mathcal{K}_T \sim K_0 T^\delta c_T^2$ by construction, and stationarity to deduce for sufficiently large K_0

$$\frac{1}{\sigma_T^2}\sum_{i=1}^{r_T}E\left|\widetilde{\mathcal{U}}_{T,i}^2 - \mathcal{U}_{T,i}^2\right| = \frac{1}{\sigma_T^2}\sum_{i=1}^{r_T}E\left[\mathcal{U}_{T,i}^2 I\left(\mathcal{U}_{T,i}^2 > \mathcal{K}_T\right)\right] \leq K\frac{r_T}{\sigma_T^2}E\left[\mathcal{U}_{T,i}^2 I\left(KT^\delta c_T^2 > K_0 T^\delta c_T^2\right)\right] = o\left(1\right).$$

Step 2 ($\sigma_T^{-4}E(\sum_{i=1}^{r_T}\{\widetilde{\mathcal{U}}_{T,i}^2 - E[\widetilde{\mathcal{U}}_{T,i}^2]\})^2 = o(1)$): By Lemma C.3 $\{m_{T,t}, \Im_t\}$ forms a geometric L_2-mixingale array with constants $e_{T,t} \leq K\sigma_T/T^{1/2}$ and coefficients $\zeta_d = O(\rho^d)$ for $\rho \in (0,1)$. Similarly by Lemma C.4 $\{m_{T,t} - E[m_{T,t}|\mathcal{F}_{T,i}], \Im_t\}_{t\in\mathcal{I}_T}$ and $\{E[m_{T,t}|\mathcal{F}_{T,i-1}], \Im_t\}_{t\in\mathcal{I}_T}$ form L_2-mixingale arrays with constants $e_{T,t}\zeta_{j_T}^\iota$ and coefficients $\zeta_{j_T}^{1-\iota}$. Since sums and products of geometric β-mixing random variables $\mathcal{U}_{T,i}$ are geometric β-mixing, and

$$\frac{\mathcal{U}_{T,i}^2}{\sigma_T^2} = \left(\sum_{t=(i-1)h_T+j_T}^{ih_T} z_{T,t} - \sum_{t=(i-1)h_T+j_T}^{ih_T}\{z_{T,t} - E[z_{T,t}|\mathcal{F}_{T,i}]\} - \sum_{t=(i-1)h_T+j_T}^{ih_T} E[z_{T,t}|\mathcal{F}_{T,i-1}]\right)^2,$$

it is straightforward to show $\{\widetilde{\mathcal{U}}_{T,i}^2/\sigma_T^2, \mathcal{F}_{T,i}\}$ forms a geometric L_2-mixingale array with constants

$\sigma_T^{-2}(\sum_{t=(i-1)h_T+j_T}^{ih_T} e_{T,t})^2\zeta_{j_T}^{2\iota}$ by the argument of Lemma C.3. Apply (9) to conclude

$$\begin{aligned} E\left(\frac{1}{\sigma_T^2}\sum_{i=1}^{r_T}\{\widetilde{\mathcal{U}}_{T,i}^2 - E[\widetilde{\mathcal{U}}_{T,i}^2]\}\right)^2 &= O\left(\frac{1}{\sigma_T^4}\sum_{i=1}^{r_T}\left(\sum_{t=(i-1)h_T+j_T}^{ih_T} e_{T,t}\right)^4 \zeta_{j_T}^{4\iota}\right) \\ &= O\left(\frac{1}{\sigma_T^4}\frac{r_T h_T^4 \sigma_T^4}{T^2}\zeta_{j_T}^{4\iota}\right) = O\left(\frac{h_T^3}{T}\zeta_{j_T}^{4\iota}\right) = o(1), \end{aligned}$$

where the second equality uses $e_{T,t} = K\sigma_T/T^{1/2}$, the third $r_T h_T \sim T$, and the last $\zeta_{j_T}^{4\iota} \to 0$ and $h_T = o(T^\delta)$ by Assumption B1 for tiny $\delta > 0$.

Step 3 ($E|\sum_{i=1}^{r_T}(\mathrm{W}_{T,i}^2 - \mathcal{Z}_{T,i}^2)| = o(1)$): Invoke Lemma C.4 and an argument identical to the proof of Lemma C.5: for some tiny $\iota > 0$

$$\begin{aligned} &\|E[\mathcal{Z}_{T,i}|\mathcal{F}_{T,i-1}]\|_2 = O\left(h_T^{1/2}\frac{e_{T,t}}{\sigma_T}\zeta_{j_T}^{\iota}\right) = O\left(\frac{h_T^{1/2}}{T^{1/2}}\zeta_{j_T}^{\iota}\right) = O\left(\frac{h_T^{1/2}}{T^{1/2}}\rho^{\iota j_T}\right) \\ &\|\{\mathcal{Z}_{T,i} - E[\mathcal{Z}_{T,i}|\mathcal{F}_{T,i}]\}\|_2 = O\left(\frac{h_T^{1/2}}{T^{1/2}}\rho^{\iota j_T}\right). \end{aligned}$$

Further, by construction and the conditional Jensen's inequality $||\mathrm{W}_{T,i} + \mathcal{Z}_{T,i}||_2 \leq K||\mathcal{Z}_{T,i}||_2$, while by (9)

$$\max_{1\leq i\leq r_T} \|\mathcal{Z}_{T,i}\|_2 \leq K\left(\sum_{i=1}^{h_T} e_{T,t}^2/\sigma_T^2\right)^{1/2} \leq K\left(h_T/T\right)^{1/2}.$$

Now apply Minkowski and Cauchy-Schwarz inequalities, Lemma 3.1 and $r_T h_T \sim T$ to obtain

$$\begin{aligned} \left\|\sum_{i=1}^{r_T}\left\{\mathcal{W}_{T,i}^2 - \mathcal{Z}_{T,i}^2\right\}\right\|_1 &\leq \sum_{i=1}^{r_T}\|\mathcal{W}_{T,i} - \mathcal{Z}_{T,i}\|_2\,\|\mathcal{W}_{T,i} + \mathcal{Z}_{T,i}\|_2 \\ &\leq \sum_{i=1}^{r_T}\left\{\|E[\mathcal{Z}_{T,i}|\mathcal{F}_{T,i-1}]\|_2 + \|\mathcal{Z}_{T,i} - E[\mathcal{Z}_{T,i}|\mathcal{F}_{T,i}]\|_2\right\} \times \|\mathcal{Z}_{T,i}\|_2 \\ &= O\left(r_T \times \frac{h_T^{1/2}}{T^{1/2}}\rho^{\iota j_T} \times \frac{h_T^{1/2}}{T^{1/2}}\right) = O\left(\rho^{\iota j_T}\right) = o(1). \end{aligned}$$

Step 4 ($\sum_{i=1}^{r_T} E[\mathcal{Z}_{T,i}^2] \xrightarrow{p} 1$): The proof of Lemma C.5 reveals by the definitions of $\mathcal{Z}_{T,i}$ and $\mathrm{W}_{T,i}$

$$\begin{aligned} \sum_{t=1}^{T} z_{T,t} &= \sum_{i=1}^{r_T}\mathcal{W}_{T,i} + \sum_{i=1}^{r_T}\left(\mathcal{Z}_{T,i} - E[\mathcal{Z}_{T,i}|\mathcal{F}_{T,i}]\right) + \sum_{i=1}^{r_T}E[\mathcal{Z}_{T,i}|\mathcal{F}_{T,i-1}] \\ &\quad + \sum_{i=1}^{r_T}\sum_{t=(i-1)h_T+1}^{(i-1)h_T+j_T} z_{T,t} + \sum_{t=r_T h_T+1}^{T} z_{T,t} \\ &= \sum_{i=1}^{r_T}\mathcal{Z}_{T,i} + o_p(1) = \sum_{i=1}^{r_T}\sum_{t=(i-1)h_T+j_T+1}^{ih_T} z_{T,t} + o_p(1). \end{aligned}$$

Further $E(\sum_{t=1}^{T} z_{T,t})^2 = 1$ by the construction of $z_{T,t}$. Therefore (cf. de Jong 1997: Appendix)

$$\begin{aligned}
\left|\sum_{i=1}^{r_T} E\left[\mathcal{Z}_{T,i}^2\right] - 1\right| &= \left|\sum_{i=1}^{r_T} E\left[\mathcal{Z}_{T,i}^2\right] - E\left(\sum_{t=1}^{T} z_{T,t}\right)^2\right| \\
&\leq \left|\sum_{i=1}^{r_T} E\left[\mathcal{Z}_{T,i}^2\right] - E\left(\sum_{i=1}^{r_T} \sum_{t=(i-1)h_T+j_T+1}^{ih_T} z_{T,t}\right)^2\right| + o(1) \\
&= 2\left|\sum_{i=1}^{r_T} \sum_{j=i+1}^{r_T} \sum_{t=(i-1)h_T+j_T+1}^{ih_T} \sum_{s=(j-1)h_T+j_T+1}^{ih_T} E\left[z_{T,s} z_{T,t}\right]\right| + o(1) =: \mathcal{A}_T,
\end{aligned}$$

say. By Lemma C.3 $\{z_{T,t}, \Im_t\}$ forms a geometric L_2-mixingale array with constants bounded by $K/T^{1/2}$, hence de Jong's (1997) Lemma 4 applies: $\mathcal{A}_T = o(1)$. $\mathcal{QED}$.

REFERENCES

Andrews, D.W.K. (1988). Laws of Large Numbers for Dependent Non-Identically Distributed Random Variables, Econometric Theory 4, 458-467.

de Jong, R. (1997). Central Limit Theorems for Dependent Heterogeneous Random Variables, Econometric Theory 13, 353-367.

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

Gordin, M. I. (1969). The Central Limit Theorem for Stationary Processes, Soviet Math. Dokl. 10, 1174-1176.

Gordin, M. I. (1973). Abstracts of Communications, T.1: A-K, International Conference on Probability, Vilnius.

Hill, J.B. (2009). On Functional Central Limit Theorems for Dependent, Heterogeneous Arrays with Applications to Tail Index and Tail Dependence Estimation. Journal of Statistical Planning and Inference 139, 2091-2110.

Hill, J.B. (2010). Central Limit Theory for Tail-Trimmed Sums of Heavy-Tailed Dependent, Heterogeneous Data under revision for Stochastic Processes and Their Applications.

Hill, J.B. and E. Renault (2010a). Generalized Method of Moments with Tail Trimming, Dept. of Economics, University of North Carolina - Chapel Hill..

Hill, J.B. and E. Renault (2010b). Technical Appendix E : Generalized Method of Moments with Tail Trimming, Dept. of Economics, University of North Carolina - Chapel Hill.

Ling, S. (2005). Self-Weighted Least Absolute Deviation Estimation for Infinite Variance Autoregressive Models, Journal of the Royal Statistical Society Series B 67, 381-393.

McLeish, D.L. (1974). Dependent Central Limit Theorems and Invariance Principles, Annals of Probability 2, 620-628.

McLeish, D.L. (1975). Maximal Inequality and Dependent Strong Laws, Annals of Probability 3, 829-839.

Meng, Y. and Z. Lin (2009). Maximal inequalities and laws of large numbers for Lq-mixingale arrays, Statistics and Probability Letters 79, 1539-1547.

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

Rio, E. (1993). Covariance Inequalities for Strongly Mixing Processes, Annals of the Institute of Henri Poincaré - B 29, 29, 587-597.