Supplemental Material for
"Heavy Tail Robust Frequency Domain Estimation"

Jonathan B. Hill* and Adam McCloskey†
Dept. of Economics Dept. of Economics
University of North Carolina Brown University

September 12, 2014

A Introduction

This appendix presents the theory for minimum mean-squared-error selection of the trimming fractile $k_{T,h}$ in the special case where $y_t y_{t-h}$ has a symmetric distribution for $h \neq 0$ (Section B). It also contains details on the robust Whittle estimator (Section C), the omitted proofs of Theorems 2.3 and 3.3 (Section D), and omitted tables (Section E).

Let $\tilde{\gamma}_{T,h}$ be the quantity used in practice for centering:

$$\tilde{\gamma}_{T,h} \equiv \frac{1}{T}\sum_{t=h+1}^{T} y_t y_{t-h} \text{ if } P\left(y_t y_{t-h} > 0\right) < 1, \text{ else } \tilde{\gamma}_{T,h} \equiv 0,$$

and define its probability limit:

$$\tilde{\gamma}_h \equiv E\left[y_t y_{t-h}\right] \text{ if } P\left(y_t y_{t-h} > 0\right) < 1, \text{ else } \tilde{\gamma}_h \equiv 0.$$

Usable lags are

$$h = \{0, 1, ..., b_T\}$$

for a sequence of bandwidths $\{b_T\}$.

Define $\Im_s^t := \sigma(y_\tau : s \leq \tau \leq t)$ and mixing coefficients $\alpha_h := \sup_{\mathcal{A}\subset\Im_{-\infty}^t, \mathcal{B}\subset\Im_{t+h}^{\infty}} |P(\mathcal{A}\cap\mathcal{B}) - P(\mathcal{A})P(\mathcal{B})|$. Let $\mathcal{L}_2(\mathcal{F}) := \mathcal{L}_2(\Omega, \mathcal{F}, \mathcal{P})$ be the space of $\mathcal{F}$-measurable L_2-bounded random variables, define $\rho(\mathcal{A},\mathcal{B}) := \sup_{f\in\mathcal{L}_2(\mathcal{A}), g\in\mathcal{L}_2(\mathcal{B})} |corr(f,g)|$ and let $\mathfrak{S}_h$ and $\mathfrak{T}_h$ be non-empty subsets of $\mathbb{N}$ with

*First, and corresponding, author. Dept. of Economics, University of North Carolina, Chapel Hill; www.unc.edu/∼jbhill; jbhill@email.unc.edu.

†Dept. of Economics, Brown University, adam_mccloskey@brown.edu.

$\inf_{s\in\mathfrak{S}_h,t\in\mathfrak{T}_h}\{|s-t|\}\geq h$. Define the interlaced maximal correlation coefficient $\rho_h^*:=\sup_{\mathfrak{S}_h,\mathfrak{T}_h}\rho(\sigma(y_t : t\in\mathfrak{S}_h), \sigma(y_s : s\in\mathfrak{T}_h))$ where the supremum is taken over all $\mathfrak{S}_h$ and $\mathfrak{T}_h$.

We use the following assumptions.

Assumption A (data generating process).

1. *$\{y_t y_{t-h}\}$ is a stationary, L_p-bounded process, $p>1$, with an absolutely continuous non-degenerate distribution with unbounded support. Further, y_t has absolutely summable covariances, and is α-mixing $\alpha_h = O(h^{-p/(p-2)})$ with $\rho_1^* < 1$.*

2. *y_t has spectrum $f(\lambda,\theta_0)$ for unique θ_0 in the interior of compact $\Theta\subset\mathbb{R}^k$ with properties:*

(i) if $\theta_0\neq\theta$ then $f(\lambda;\theta)\neq f(\lambda;\theta_0)$;

(ii) $0<f(\lambda,\theta)\leq K<\infty$ for each $\lambda\in[-\pi,\pi]$ and $\theta\in\Theta$;

(iii) $f(\lambda,\theta)$ is twice continuously differentiable in θ, with derivatives $(\partial/\partial\theta)^i f(\lambda,\theta)$ for $i=1,2$ uniformly bounded on $[-\pi,\pi]\times\Theta$;

(iv) $h(\lambda,\theta)\in\{f(\lambda,\theta),(\partial/\partial\theta)f(\lambda,\theta)\}$ are uniformly Hölder continuous of degree $\alpha\in(1/2,1]$ in λ: $\sup_{\theta\in\Theta}||h(\lambda,\theta)-h(\omega,\theta)||\leq K|\lambda-\omega|^\alpha$ for all $\lambda,\omega\in[-\pi,\pi]$ and some $K>0$.

3. $\inf_{\theta\in\Theta}||\mathcal{H}(\theta)||>0$.

Assumption B (regularly varying tails). $P(y_t y_{t-h}-\tilde{\gamma}_h\leq -c)=\mathcal{L}_{h,1}(c)c^{-\kappa_{h,1}}$ *and* $P(y_t y_{t-h}-\tilde{\gamma}_h\geq c)=\mathcal{L}_{h,2}(c)c^{-\kappa_{h,2}}$ *where $\mathcal{L}_{h,i}(c)$ are slowly varying, and $\kappa_{h,i}>1$.*

Assumption B′ (second order power law and fractile rates). $P(y_t y_{t-h}-\tilde{\gamma}_h\leq -c) = d_{h,1}c^{-\kappa_{h,1}}(1+O(r_1(c))$ *and* $P(y_t y_{t-h}-\tilde{\gamma}_h\geq c)=d_{h,2}c^{-\kappa_{h,2}}(1+O(r_2(c)))$, *where $d_{h,i}>0$, $\kappa_{h,i}>1$, and r_i are measurable functions. Let $e_{h,i}>0$, $e_{h,0}\equiv\min\{e_{h,1},e_{h,2}\}$ and $\kappa_{h,0}\equiv\min\{\kappa_{h,1},\kappa_{h,2}\}$. Then $m_{T,h}\in\{1,...,T-h\}$ and $m_{T,h}\to\infty$, and either $r_i(c)=c^{-e_{h,i}}$ and $m_{T,h}=o((T-h)^{2e_{h,0}/(2e_{h,0}+\kappa_{h,0})})$, or $r_i(c)=\ln(c)^{-e_{h,i}}$ and $m_{T,h}=o(\ln(T-h)^{2e_{h,0}})$. Finally, $m_{T,h}/k_{T,h}\to\infty$.*

Assumption C (trimming and bandwidth rates).

1. *In general $k_{T,h}\to\infty$, $k_{T,h}/(T-h)\to 0$, and $k_{T,h}\sim K_{h,\tilde{h}}k_{T,\tilde{h}}$ for some $K_{h,\tilde{h}}>0$ and each $\tilde{h},h\in\{0,...,b_T\}$. If mean-centering at lag h is used then $k_{T,h}\to\infty$ at most at a slowly varying rate.*

2. *Let $b_T\leq T-1$, $b_T\to\infty$, and $T-b_T\to\infty$. Further $b_T/T^{1/(2\alpha)}\to\infty$ where $\alpha\in(1/2,1]$ is the Hölder continuity degree in Assumption A.2.iv.*

B Mean-Squared-Error Minimization

Write the bias and scale as $\mathcal{B}_T=[\mathcal{B}_{i,T}]_{i=1}^k$ and $\mathcal{V}_T=[\mathcal{V}_{i,j,T}]_{i,j=1}^k$. By Theorem 2.2 $T^{1/2}\mathcal{V}_T^{-1/2}(\hat{\theta}_T^*-\theta_0+\mathcal{B}_T)\xrightarrow{d}N(0,I_k)$, hence the *first order mean-squared-error* [*fmse*] $\mathcal{M}_{i,T}$ of $\hat{\theta}_{i,T}^*$ is

$$\mathcal{M}_{i,T}\equiv\mathcal{B}_{i,T}^2+\mathcal{V}_{i,i,T}/T.$$

The bias of $\hat{\theta}^*_T$ is complicated unless $y_t y_{t-h}$ is symmetrically distributed for $h > 1$, since then $\mathcal{B}_T = \Omega^{-1}(2\pi)^{-1}\varpi_0\{E[y_t^2 I(y_t^2 < c_{T,0})] - E[y_t^2]\}$ only depends on $E[y_t^2 I(y_t^2 < c_{T,0})]$. In this case, the *fmse* minimizing $k_{T,h}$ satisfies $k_{T,h} \to [0,\infty)$, a constant that depends on $\kappa \equiv \arg\sup\{\alpha > 0 : E|y_t|^\alpha < \infty\} > 2$.

Theorem B.1 *Let Assumptions A-C hold, and assume $y_t y_{t-h}$ for $h \neq 0$ have symmetric distributions. The* fmse *minimizing $k_{T,h} \sim K > 0$ if $\kappa \in (2,4) \cup (4,\infty)$, and $k_{T,h} = 0$ if $\kappa = 4$.*

Proof. By distribution symmetry of $y_t y_{t-h}$ for $h \neq 0$, Theorem 2.2 and Karamata's Theorem $\mathcal{B}_{i,T} = K\{E[y_t^2 I(y_t^2 < c_{T,0})] - E[y_t^2]\} \sim K(k_T/T)c_{T,0} \sim K(T/k_T)^{2/\kappa - 1}$, hence squared bias is $\mathcal{B}_{i,T}^2 \sim K(T/k_T)^{4/\kappa-2}$.

Next, by Theorem 2.2 $\mathcal{V}_{i,i,T} \sim KE[y_t^4 I(y_t^2 < c_{T,0})]$. If $\kappa > 4$ then $\mathcal{V}_{i,i,T} \sim K\{E[y_t^4] - E[y_t^4 I(y_t^2 \geq c_{T,0})]\}$, while by Karamata's Theorem $E[y_t^4 I(y_t^2 \geq c_{T,0})] \sim K(k_T/T)c_{T,0}^2 = K\left((k_T/T)\right)^{1-4/\kappa}$. Now use Theorem 2.2 to deduce by case that if $\kappa > 4$ then $\mathcal{V}_{i,i,T} \sim K\{E[y_t^4] - K(k_T/T)^{1-4/\kappa}\}$, if $\kappa = 4$ then $\mathcal{V}_{i,i,T} \sim \mathcal{L}(T) \to \infty$ for some slowly varying $\mathcal{L}(T)$, and if $\kappa < 4$ then $\mathcal{V}_{i,i,T} \sim K(T/k_T)^{4/\kappa-1}$. Therefore $\mathcal{M}_{i,T} \sim \tilde{\mathcal{M}}_{i,T}(k_T)$, where by case:

$$\kappa > 4 : \tilde{\mathcal{M}}_{i,T}(a) = K\left(\frac{a}{T}\right)^{2-4/\kappa} + K\frac{1}{T}\left\{E[y_t^4 - K\left(\frac{a}{T}\right)^{1-4/\kappa}\right\}$$

$$\kappa = 4 : \tilde{\mathcal{M}}_{i,T}(a) = K\left(\frac{a}{T}\right)^{2-4/\kappa} + K\frac{1}{T}\mathcal{L}(T) = K\frac{a}{T} + K\frac{1}{T}\mathcal{L}(T)$$

$$\kappa < 4 : \tilde{\mathcal{M}}_{i,T}(a) = K\left(\frac{a}{T}\right)^{2-4/\kappa} + K\frac{1}{T}\left(\frac{T}{a}\right)^{4/\kappa-1}.$$

Now let $a \in [0,\infty)$ be real valued and differentiate $\tilde{\mathcal{M}}_{i,T}(a)$ to deduce $\tilde{a} \equiv \arg\min_{a\geq 0}\tilde{\mathcal{M}}_{i,T}(a)$ satisfies:

$$\kappa > 4 : \frac{\partial}{\partial a}\tilde{\mathcal{M}}_{i,T}(\tilde{a}) = K\left(2 - \frac{4}{\kappa}\right)\frac{\tilde{a}^{1-4/\kappa}}{T^{2-4/\kappa}} - K\left(1 - \frac{4}{\kappa}\right)\frac{\tilde{a}^{-4/\kappa}}{T^{2-4/\kappa}} = 0 \text{ for some } \tilde{a} \sim K$$

$$\kappa = 4 : \frac{\partial}{\partial a}\tilde{\mathcal{M}}_{i,T}(\tilde{a}) = K\frac{1}{T} > 0 \text{ hence } \tilde{a} = 0$$

$$\kappa < 4 : \frac{\partial}{\partial a}\tilde{\mathcal{M}}_{i,T}(\tilde{a}) = K\left(2 - \frac{4}{\kappa}\right)\frac{\tilde{a}^{1-4/\kappa}}{T^{2-4/\kappa}} - K\left(\frac{4}{\kappa} - 1\right)\frac{T^{4/\kappa-2}}{\tilde{a}^{4/\kappa}} = 0 \text{ for some } \tilde{a} \sim K.$$

This completes the proof. $\mathcal{QED}$.

C Robust Whittle Estimator

Now suppose the spectral density has the form

$$f(\lambda,\theta)=\frac{\sigma^2}{2\pi}\times\frac{|\mathcal{B}(e^{-i\lambda},\beta)|^2}{|\mathcal{A}(e^{-i\lambda},\beta)|^2}=\frac{\sigma^2}{2\pi}\times\omega(\lambda,\beta)\text{ where }\theta=\left[\beta',\sigma^2\right]', \tag{C.1}$$

where $\mathcal{A}$ and $\mathcal{B}$ are continuously differentiable functions of $(\lambda,\beta)\in[-\pi,\pi]\times\mathfrak{B}$ and $\mathfrak{B}$ is a compact subset of $\mathbb{R}^k$. Autoregressions and squared GARCH processes, for example, satisfy (C.1). Notice Assumption A.2 imposes smoothness and boundedness properties on $f(\lambda,\theta)$ which instantly apply to $\omega(\lambda,\beta)$ under obvious restrictions on $\mathcal{A}$ and $\mathcal{B}$.

Define:

$$\widehat{\mathfrak{Y}}_{h,t}^{(0)}\equiv|y_ty_{t-h}-\tilde{\gamma}_{T,h}|\quad\text{and}\quad\widehat{\mathfrak{Y}}_{h,(1)}^{(0)}\geq\widehat{\mathfrak{Y}}_{h,(2)}^{(0)}\geq\cdots\widehat{\mathfrak{Y}}_{h,(T-h)}^{(0)}.$$

The negligibly transformed optimal bias-corrected Whittle estimator solves

$$\tilde{\beta}_T^{(obc)}=\underset{\beta\in\mathfrak{B}}{\arg\min}\sum_{j\in\mathcal{F}}\frac{\widehat{\mathcal{I}}_T^{(obc)}(\lambda_j)}{\omega(\lambda_j,\beta)}$$

where

$$\widehat{\mathcal{I}}_T^{(obc)}(\lambda)=\frac{1}{2\pi}\left(\widehat{\hat{\gamma}}_{T,0}^{(obc)}(\widehat{\mathfrak{Y}}_{0,(k_{T,h})}^{(0)})+2\sum_{h=1}^{b_T}\widehat{\hat{\gamma}}_{T,h}^{(obc)}(\widehat{\mathfrak{Y}}_{h,(k_{T,h})}^{(0)})\times\cos(\lambda h)\right).$$

See Section 3 in the main paper for details on $\widehat{\hat{\gamma}}_{T,h}^{(obc)}(\widehat{\mathfrak{Y}}_{h,(k_{T,h})}^{(0)})$ and the bandwidth b_T.

Define $\tilde{\mathcal{V}}_T:=\tilde{\Omega}^{-1}\tilde{\mathcal{S}}_T\tilde{\Omega}^{-1}$ where

$$\tilde{\Omega}:=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\partial\ln\omega(\lambda)}{\partial\theta}\frac{\partial\ln\omega(\lambda)}{\partial\theta'}d\lambda$$

$$\tilde{\mathcal{S}}_T:=T\times E\left[\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\mathcal{I}_T^*(\lambda)}{\omega(\lambda)}\frac{\partial\omega(\lambda)}{\partial\beta}d\lambda\right)\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\mathcal{I}_T^*(\lambda)}{\omega(\lambda)}\frac{\partial\omega(\lambda)}{\partial\beta'}d\lambda\right)\right].$$

The following theorem can be proved by exploiting our proof of Theorems 2.2 and 3.1 and following Hannan (1973)'s proof of his Theorem 2. We therefore omit the proof.

Theorem C.1 *Under Assumptions A, B′ and C we have* $T^{1/2}\tilde{\mathcal{V}}_T^{-1/2}(\tilde{\beta}_T^{(obc)}-\beta_0)\xrightarrow{d}N\left(0,I_k\right).$

The asymptotic variance has a classic structure if y_t is linear with a finite fourth moment. Consider

$$y_t=\sum_{i=0}^{\infty}\xi_i\left(\theta_0\right)\epsilon_{t-i}\text{ where }\sum_{i=0}^{\infty}\xi_i^2\left(\theta_0\right)<\infty,\ \xi_0\left(\theta_0\right)=1 \tag{C.2}$$

$$\sigma^2\left(\theta_0\right):=E[\epsilon_t^2]<\infty\text{ and }E[\epsilon_s\epsilon_t]=0\ \forall s\neq t,$$

where ϵ_t is a homoscedastic martingale difference. Define the σ-field $\Im_t := \sigma(y_\tau : \tau \leq t)$. Theorem C.1, negligibility, dominated convergence, and arguments in Dunsmuir (1979, proof of Theorem 2.1) imply the next result. Define the σ-field $\Im_t := \sigma(y_\tau : \tau \leq t)$.

Corollary C.2 *In addition to Assumption A, let y_t satisfy (C.2), and assume $E[\epsilon_t|\Im_{t-1}] = 0$ a.s., $E[\epsilon_t^2|\Im_{t-1}] = \sigma^2$ a.s., $E[\epsilon_t^3|\Im_{t-1}] = s$ a.s., and $E[\epsilon_t^4] = \mathcal{K} < \infty$. Let $k_{T,h} \to \infty$ and $k_{T,h}/(T-h) = o(1)$. Then $T^{1/2}(\tilde{\beta}_T^{(obc)} - \beta_0) \overset{d}{\to} N(0, \tilde{\Omega}^{-1})$.*

D Omitted Proofs

D.1 Proof of Theorem 2.3 (FD-QML Bias)

Theorem 2.3. *Under Assumptions A-D:*

i. $T^{1/2}\mathcal{V}_T^{-1/2}(\hat{\theta}_T^* - \theta_0) \overset{d}{\to} N(0, I_k)$ if either $\kappa > 4$; or $\kappa = 4$, $k_{T,h} = o(\ln(T))$ and $P(|y_t y_{t-h} - \tilde{\gamma}_h| \geq c) = d_{h,0}c^{-\kappa_{h,0}}(1 + o(1))$ where $\kappa_{0,0} = 2 \leq \kappa_{h,0}$;

ii. $T^{1/2}\mathcal{V}_T^{-1/2}(\hat{\theta}_T^* - \theta_0 + \mathcal{B}_T) \overset{d}{\to} N(0, I_k)$ if $\kappa \in (2, 4)$, where $T^{1/2}||\mathcal{V}_T^{-1/2}\mathcal{B}_T|| \sim K k_{T,h}^{1/2} \to \infty$ if $y_t y_{t-h}$ has a symmetric distribution, else $\lim_{T\to\infty} T^{1/2}||\mathcal{V}_T^{-1/2}\mathcal{B}_T||/k_{T,h}^{1/2} \in [0, \infty)$.

Proof.. We first characterize $\lim_{T\to\infty}(T/||\mathcal{V}_T||)^{1/2}\{E[y_t^2 I(y_t^2 < c_{T,0})] - E[y_t^2]\}$. By negligibility and Theorem 2.2, $\|\mathcal{V}_T\| \sim K[E(y_t^4 I(y_t^2 < c_{T,0})]$, and trivially

$$E[y_t^2 I(y_t^2 < c_{T,0})] - E[y_t^2] = E\left[\left(y_t^2 - E\left[y_t^2\right]\right) I\left(y_t^2 \geq c_{T,0}\right)\right],$$

hence

$$\left(\frac{T}{\|\mathcal{V}_T\|}\right)^{1/2} \left\{E[y_t^2 I(y_t^2 < c_{T,0})] - E\left[y_t^2\right]\right\} \sim -K\frac{T^{1/2}E\left[y_t^2 I\left(y_t^2 \geq c_{T,0}\right)\right]}{\left(E\left[y_t^4 I\left(y_t^2 < c_{T,0}\right)\right]\right)^{1/2}}. \tag{D.3}$$

Recall the tail components $\mathcal{L}_{h,0}(c) \equiv \mathcal{L}_{h,1}(c) + \mathcal{L}_{h,2}(c)$ and $\kappa_{h,0} \equiv \max\{\kappa_{h,1}, \kappa_{h,2}\}$, where $\kappa_{0,0} = \kappa/2 > 1$. By Karamata's Theorem (Resnick, 1987, Theorem 0.6):

$$E\left[y_t^2 I\left(y_t^2 \geq c_{T,0}\right)\right] \sim \frac{2}{\kappa - 2}\frac{k_{T,0}}{T}c_{T,0}.$$

The remaining $E[(y_t^4 I(y_t^2 < c_{T,0})]$ in (D.3) is $O(1)$ if $\kappa > 4$, and is evaluated for $\kappa \in (2, 4]$ by Karamata's Theorem: for some slowly varying $\tilde{\mathcal{L}}_4(c)$:

$$\kappa \in (2,4) : E\left[y_t^4 I\left(y_t^2 < c_{T,0}\right)\right] \sim \frac{4}{4-\kappa}\frac{k_{T,0}}{T}c_{T,0}^2 \tag{D.4}$$

$$\kappa = 4 : E\left[y_t^4 I\left(y_t^2 < c_{T,0}\right)\right] = \tilde{\mathcal{L}}_4(T).$$

We now deduce conditions for asymptotic bias $\lim_{T\to\infty}(T/||\mathcal{V}_T||)^{1/2}\{E[y_t^2 I(y_t^2 < c_{T,0})] - E[y_t^2]\} \neq 0$. Use $c_{T,0} \sim \mathcal{L}_{0,0}(T)^{2/\kappa}(T/k_{T,0})^{2/\kappa}$ to deduce if $\kappa > 4$ then $(T/||\mathcal{V}_T||)^{1/2}\{E[y_t^2 I(y_t^2 < c_{T,0})] -$

$E[y_t^2]\} \to 0$ only if $T^{1/2}\mathcal{L}_{0,0}(T)^{2/\kappa}(k_{T,0}/T)^{1-4/\kappa} \to 0$. The latter holds by slow variation for $\mathcal{L}_{0,0}(T)$, and slow variation for $k_{T,0}$ by Assumption D.

If $\kappa = 4$ then $c_{T,0} \sim \mathcal{L}_{0,0}(T)^{1/2}(T/k_{T,0})^{1/2}$ hence $(T/||\mathcal{V}_T||)^{1/2}\{E[y_t^2 I(y_t^2 < c_{T,0})] - E[y_t^2]\} \to 0$ only if

$$\frac{T^{1/2}\left|E\left[y_t^2 I\left(y_t^2 \geq c_{T,0}\right)\right]\right|}{\left(E\left[y_t^4 I\left(y_t^2 < c_{T,0}\right)\right]\right)^{1/2}} \sim K\frac{T^{1/2}\left(k_{T,0}/T\right)c_{T,0}}{\tilde{\mathcal{L}}_4(T)^{1/2}} \sim K\frac{k_{T,0}^{1/2}\mathcal{L}_{0,0}(T)^{1/2}}{\tilde{\mathcal{L}}_4(T)^{1/2}} \to 0. \tag{D.5}$$

If $P(|y_t y_{t-h} - E[y_t y_{t-h}]| \geq c) = d_{h,0}c^{-2}(1 + o(1))$ with $d_{h,0} > 0$, then $\mathcal{L}_{0,0}(T) \sim d$, and by direct integration $\tilde{\mathcal{L}}_4(T) = d_{0,0}\ln(T)$, hence $k_{T,h} \to \infty$ and $k_{T,h} = o(\ln(T))$ ensure (D.5).

If $\kappa \in (2,4)$ then we need:

$$\frac{T^{1/2}E\left[y_t^2 I\left(y_t^2 \geq c_{T,0}\right)\right]}{\left(E\left[y_t^4 I\left(y_t^2 < c_{T,0}\right)\right]\right)^{1/2}} \sim K\frac{T^{1/2}\left(k_{T,0}/T\right)c_{T,0}}{\left(\left(k_{T,0}/T\right)c_{T,0}^2\right)^{1/2}} = Kk_{T,0}^{1/2} \to 0,$$

which is ruled out by $k_{T,h} \to \infty$.

A similar set of derivations applies to any $\gamma_{T,h}^*(c_{T,h}) \equiv (T-h)(T-h-k_{T-h})^{-1}E[y_t y_{t-h}I(|y_t y_{t-h} - \tilde{\gamma}_h| < c_{T,h})]$ with $h \neq 0$, although there are two differences. First, when $y_t y_{t-h}$ has a symmetric distribution then $\gamma_{T,h}^*(c_{T,h}) = E[y_t y_{t-h}]$ for $h \neq 0$. Second, if y_t is independent then $y_t y_{t-h}$ for $h \neq 0$ has tail index $\kappa > 2$ and therefore $T^{1/2}\mathcal{V}_T^{-1/2}|\gamma_{T,h}^*(c_{T,h}) - E[y_t y_{t-h}]| = o(1)$ if $k_{T,h}/\ln(T) \to 0$ by repeating the above arguments. Otherwise since we trim symmetrically there is potential bias with each $\gamma_{T,h}^*(c_{T,h})$: $T^{1/2}\mathcal{V}_T^{-1/2}|\gamma_{T,h}^*(c_{T,h}) - E[y_t y_{t-h}]| = O(k_{T,h}^{1/2})$ when $\kappa \in (2,4)$, and $T^{1/2}\mathcal{V}_T^{-1/2}|\gamma_{T,h}^*(c_{T,h}) - E[y_t y_{t-h}]| = o(1)$ when $\kappa \geq 4$ as long as $k_{T,h}/\ln(T) \to 0$. Hence by covariance summability Assumption B.1 $\sum_{h=0}^{\infty} T^{1/2}\mathcal{V}_T^{-1/2}|\gamma_{T,h}^*(c_{T,h}) - E[y_t y_{t-h}]| = O(k_{T,h}^{1/2})$ when $\kappa \in (2,4)$ and $\sum_{h=0}^{\infty} T^{1/2}\mathcal{V}_T^{-1/2}|\gamma_{T,h}^*(c_{T,h}) - E[y_t y_{t-h}]| = o(1)$ when $\kappa \geq 4$ if $k_{T,h}/\ln(T) \to 0$. The claim now follows from the bias form $\mathcal{B}_T = \Omega^{-1}(2\pi)^{-1}\sum_{h=-\infty}^{\infty}\varpi_h\{\gamma_{T,h}^*(c_{T,h}) - E[y_t y_{t-h}]\}$ by Theorem 2.2. $\mathcal{QED}$.

D.2 Proof of Theorem 3.3 (Bias-Correction: Thin Tails)

Recall for Theorem 3.3 we assume tails decay exponentially fast for the sake of discussion:

$$P(|y_t y_{t-h} - \tilde{\gamma}_h| \ \geq \ c) = \vartheta_h \exp\{-\zeta_h c^{\delta_h}\} \text{ where } \vartheta_h, \zeta_h, \delta_h > 0. \tag{D.6}$$

Recall

$$\Omega \equiv \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\partial \ln f(\lambda;\theta_0)}{\partial\theta}\frac{\partial \ln f(\lambda;\theta_0)}{\partial\theta'}d\lambda \ \text{ and } \ \Pi \equiv \frac{\left(\mathcal{K} - 3\sigma^4\right)}{\sigma^8(\theta_0)}\frac{\partial\sigma^2(\theta_0)}{\partial\theta}\frac{\partial\sigma^2(\theta_0)}{\partial\theta'}. \tag{D.7}$$

Theorem 3.3. *Let Assumption A hold, assume (D.6) applies for each h, and let $k_{T,h} \to \infty$, $k_{T,h}/\ln(T) \to 0$, $m_{T,h} \to \infty$, and $m_{T,h}/T \to 0$. Then $T^{1/2}\hat{\mathcal{R}}_{T,h} \overset{p}{\to} 0$ hence $T^{1/2}\mathcal{V}_T^{-1/2}(\hat{\theta}_T^{(obc)} - \theta_0) \overset{d}{\to} N(0, I_k)$. Further, if the conditions of Corollary 2.5 hold then $T^{1/2}(\hat{\theta}_T^{(obc)} - \theta_0) \overset{d}{\to} N(0,\mathcal{V})$ where $\mathcal{V} = \Omega^{-1}(2\Omega + \Pi)\Omega^{-1}$ and Ω and Π are defined in (D.7).*

Proof. We will only show $T^{1/2}\hat{\mathcal{R}}_{T,0} \overset{p}{\to} 0$, the remaining $T^{1/2}\hat{\mathcal{R}}_{T,h} \overset{p}{\to} 0$ being similar. Drop the lag 0 subscript everywhere, e.g. $c_T = c_{T,0}$, $m_T = m_{T,0}$, $\hat{\kappa}_{m_T} = \hat{\kappa}_{0,0,m_T}$, etc. In order to ease the notational burden, we assume $\tilde{\gamma}_h = 0$. Write

$$\mathcal{Y}_t := y_t^2$$

and define order statistics $\mathcal{Y}_{(1)} \geq \mathcal{Y}_{(2)} \geq \cdots \geq \mathcal{Y}_{(T)}$.

For any tiny $\iota > 0$,

$$T^{1/2}\hat{\mathcal{R}}_{T,0} = \frac{1}{\hat{\kappa}_{m_T} - 1} \times \frac{k_T}{T^{1/2-\iota}} \times \frac{c_T}{T^{\iota}} \times \frac{\mathcal{Y}_{(k_T)}}{c_T}.$$

We will show $c_T/T^{\iota} \to 0$, $\mathcal{Y}_{(k_T)}/c_T \overset{p}{\to} 1$, and $1/\hat{\kappa}_{m_T} \overset{p}{\to} 0$ in three steps. In view of $k_T/T^{1/2-\iota} \to 0$ by the supposition $k_T/\ln(T) \to 0$ the proof is then complete.

Step 1 (c_T). By (D.6), and the definition of c_T:

$$c_T = \left(\frac{1}{\zeta}\ln\vartheta + \frac{1}{\zeta}\ln\frac{T}{k_T}\right)^{1/\delta} \sim \frac{1}{\zeta^{1/\delta}}\left(\ln(T)\right)^{1/\delta} = o(T^{\iota}). \tag{D.8}$$

Step 2 ($\mathcal{Y}_{(k_T)}$). We will prove $k_T^{1/2}(\ln\mathcal{Y}_{(k_T)} - \ln c_T) = O_p(1)$ hence $\mathcal{Y}_{(k_T)}/c_T = 1 + O_p(1/k_T^{1/2})$. Define, $\mathcal{I}_{T,t}(u) := (T/k_T)I(|y_t y_{t-h}| > c_T e^u)$ for any $u \in \mathbb{R}$, and $\mathcal{I}_T(u) := 1/T\sum_{t=1}^{T}\mathcal{I}_{T,t}(u)$. By construction $k_T^{1/2}(\ln\mathcal{Y}_{(k_T)} - \ln c_T) \leq u$ for $u \in \mathbb{R}$ *if and only if* $\mathcal{I}_T(u/k_T^{1/2}) \leq 1$, *if and only if*

$$\mathcal{I}_T(u/k_T^{1/2}) - E\left[\mathcal{I}_T(u/k_T^{1/2})\right] \leq 1 - \frac{T}{k_T}P\left(\mathcal{Y}_t > c_T e^{u/k_T^{1/2}}\right) = 1 - \frac{P\left(\mathcal{Y}_t > c_T e^{u/k_T^{1/2}}\right)}{P\left(\mathcal{Y}_t > c_T\right)}.$$

Now expand $P(|\mathcal{Y}_t| > c_T e^{u/k_T^{1/2}})$ around $u = 0$. Use $c_T \sim \zeta^{-1/\delta}(\ln(T))^{1/\delta}$ and $(\partial/\partial c)P(|\mathcal{Y}_t| > c) = -\zeta\vartheta\delta c^{\delta-1}\exp\{-\zeta c^{\delta}\}$, and the mean-value-theorem, to deduce $k_T^{1/2}(\ln\mathcal{Y}_{(k_T)} - \ln c_T) \leq u$ *if and only if* for some $|u^*| \leq |u|$:

$$\begin{aligned}
&k_T^{1/2}\left(\mathcal{I}_T(u/k_T^{1/2}) - E\left[\mathcal{I}_T(u/k_T^{1/2})\right]\right) \\
&\qquad \leq -\frac{1}{P\left(|\mathcal{Y}_t| > c_T\right)}\frac{\partial}{\partial c}P\left(\mathcal{Y}_t > c\right)|_{c=c_T e^{u_*/k_T^{1/2}}} \times c_T e^{u_*/k_T^{1/2}} \times u \\
&\qquad = \delta\vartheta\zeta\frac{T}{k_T}\exp\left\{-\zeta\left(c_T e^{u_*/k_T^{1/2}}\right)^{\delta}\right\}\left(c_T e^{u_*/k_T^{1/2}}\right)^{\delta-1} \times c_T e^{u_*/k_T^{1/2}}u \\
&\qquad = \delta\vartheta\frac{\ln(T)}{k_T}u \times (1 + o(1)).
\end{aligned}$$

Therefore $k_T^{1/2}(\ln\mathcal{Y}_{(k_T)} - \ln c_T)$ and $(k_T^{3/2}/(\delta\vartheta\ln(T))) \times (\mathcal{I}_T(u/k_T^{1/2}) - E[\mathcal{I}_T(u/k_T^{1/2})])$ have the same limit distribution. But in view of the Assumption A mixing properties and the zero mean

L_2-boundedness of $(T/k_T)^{1/2}\{I(\mathcal{Y}_t > c_T e^u) - E[I(\mathcal{Y}_t > c_T e^u)]\}$ we have by Theorem 1.6 and Lemma 2.1 in McLeish (1975):

$$k_T^{1/2}\left(\mathcal{I}_T(u/k_T^{1/2}) - E\left[\mathcal{I}_T(u/k_T^{1/2})\right]\right)$$

$$= \frac{1}{T^{1/2}}\sum_{t=1}^{T}\left(\frac{T}{k_T}\right)^{1/2}\left\{I\left(\mathcal{Y}_t > c_T e^{u/k_T^{1/2}}\right) - E\left[I\left(\mathcal{Y}_t > c_T e^{u/k_T^{1/2}}\right)\right]\right\} = O_p(1).$$

By the supposition $k_T = O(\ln(T))$ it therefore follows

$$\frac{k_T^{3/2}}{\ln(T)}\left(\mathcal{I}_T(u/k_T^{1/2}) - E\left[\mathcal{I}_T(u/k_T^{1/2})\right]\right) = \frac{k_T}{\ln(T)}k_T^{1/2}\left(\mathcal{I}_T(u/k_T^{1/2}) - E\left[\mathcal{I}_T(u/k_T^{1/2})\right]\right) = O_p(1).$$

Therefore $k_T^{1/2}(\ln \mathcal{Y}_{(k_T)} - \ln c_T) = O_p(1)$.

Step 3 ($\hat{\kappa}_{m_T}$). In view of distribution continuity we can define a sequence of positive numbers $\{\tilde{c}_T\}$ that satisfies

$$P\left(\left|y_t^2\right| > \tilde{c}_T\right) = \frac{m_T}{T}.$$

We have

$$\begin{aligned}\hat{\kappa}_{m_T}^{-1} &= \frac{1}{m_T}\sum_{j=1}^{m_T}\ln\left(\mathcal{Y}_{(j)}/\mathcal{Y}_{(m_T+1)}\right) \\ &= \frac{1}{m_T}\sum_{i=1}^{T}\ln\left(\mathcal{Y}_i/\mathcal{Y}_{(m_T+1)}\right)\times I\left(\mathcal{Y}_i > \mathcal{Y}_{(m_T+1)}\right) \\ &= \frac{1}{m_T}\sum_{i=1}^{T}\ln\left(\mathcal{Y}_i/\tilde{c}_T\right)\times I\left(\mathcal{Y}_i > \mathcal{Y}_{(m_T+1)}\right) + o_p(1) \\ &= \frac{1}{m_T}\sum_{i=1}^{T}\ln\left(\mathcal{Y}_i/\tilde{c}_T\right)\times I\left(\mathcal{Y}_i > \tilde{c}_T\right) + o_p(1) = \frac{1}{T}\sum_{i=1}^{T}\mathfrak{Y}_{T,i} + o_p(1),\end{aligned} \tag{D.9}$$

say. The third equality follows from $\mathcal{Y}_{(m_T+1)}/\tilde{c}_T = 1 + O_p(1/m_T^{1/2})$ by an application of Step 2 since

$$\frac{1}{m_T}\sum_{i=1}^{T}\ln\left(\mathcal{Y}_i/\mathcal{Y}_{(m_T+1)}\right)\times I\left(\mathcal{Y}_i > \mathcal{Y}_{(m_T+1)}\right) - \frac{1}{m_T}\sum_{i=1}^{T}\ln\left(\mathcal{Y}_i//\tilde{c}_T\right)\times I\left(\mathcal{Y}_i > \mathcal{Y}_{(m_T+1)}\right)$$

$$= \left(\ln\left(\mathcal{Y}_{(m_T+1)}\right) - \ln\left(\tilde{c}_T\right)\right)\times\frac{1}{m_T}\sum_{i=1}^{T}I\left(\mathcal{Y}_i > \mathcal{Y}_{(m_T+1)}\right)$$

$$= O_p\left(\frac{1}{m_T^{1/2}}\times\frac{1}{m_T}\sum_{i=1}^{T}I\left(\mathcal{Y}_i > \mathcal{Y}_{(m_T+1)}\right)\right) = O_p\left(1/m_T^{1/2}\right).$$

Notice this exploits the construction of $\mathcal{Y}_{(m_T+1)}$ and distribution continuity: $1/m_T \sum_{i=1}^T I(\mathcal{Y}_i > \mathcal{Y}_{(m_T+1)}) = 1$ $a.s.$ The fourth equality in (D.9) can be verified under very general conditions in view of the stationary mixing property and the fact that $\mathcal{Y}_{(m_T+1)}/\tilde{c}_T = 1 + O_p(1/m_T^{1/2})$. The method of proof is identical to Lemmas A.2 and A.3 in Hill (2012).

In order to prove $\hat{\kappa}_{m_T}^{-1} \xrightarrow{p} 0$ it remains to verify $1/T \sum_{i=1}^T \mathfrak{Y}_{T,i} \xrightarrow{p} 0$. The variable $\mathfrak{Y}_{T,i} = (T/m_T)\ln(\mathcal{Y}_i/\tilde{c}_T)I(\mathcal{Y}_i > \tilde{c}_T)$ has a positive finite mean that decays to zero. This follows by invoking exponential tail bound (D.6) and $P(y_t^2 > \tilde{c}_T) = m_T/T$ to obtain

$$\begin{aligned} E\left[\mathfrak{Y}_{T,i}\right] &= \frac{T}{m_T - 1}\int_0^\infty P\left(\ln\left(\mathcal{Y}_i/\tilde{c}_T\right) > u\right) du = \frac{T}{m_T - 1}\int_0^\infty P\left(\mathcal{Y}_i > \tilde{c}_T e^u\right) du \\ &= \frac{T}{m_T - 1} P\left(\mathcal{Y}_i > \tilde{c}_T\right) \int_0^\infty \frac{P\left(\mathcal{Y}_i > \tilde{c}_T e^u\right)}{P\left(\mathcal{Y}_i > \tilde{c}_T\right)} du \\ &= \frac{m_T}{m_T - 1}\int_0^\infty \exp\left\{-\zeta \tilde{c}_T^\delta \left(e^{\delta u} - 1\right)\right\} du > 0. \end{aligned}$$

Seeing that $\zeta > 0$, and (D.8) implies $\tilde{c}_T \to \infty$, it follows by dominated convergence $E[\mathfrak{Y}_{T,i}] \searrow 0$. Hence the Cesáro sum $1/T \sum_{i=1}^T E[\mathfrak{Y}_{T,i}] \searrow 0$. Now use Markov's inequality and $1/T \sum_{i=1}^T E[\mathfrak{Y}_{T,i}] \searrow 0$ to deduce $1/T \sum_{i=1}^T \mathfrak{Y}_{T,i} \xrightarrow{p} 0$. $\mathcal{QED}$.

E Omitted Tables

This section contains omitted simulation results for AR models with $\phi_0 \in \{0, .75\}$, and for GARCH models with $T = 2000$.

Table E.1 : FD-QML for AR(1) : $\phi_0 = .75$ and $\sigma_0^2 = 1$

$\phi_0 = .75$[a]

	T = 100					**T= 250**			
	Bias	Med[b]	RMSE	$KS^c_{.05}$		Bias	Med	RMSE	$KS_{.05}$
	Pareto Error: $\kappa = 2.25$					Pareto Error: $\kappa = 2.25$			
no trim[d]	-.0227	.7358	.0972	1.222	no trim	-.0072	.7451	.0427	1.435
trim-obc	-.0120	.7396	.0859	.9870	trim-obc	-.0032	.7460	.0575	1.231
	Pareto Error: $\kappa = 4.50$					Pareto Error: $\kappa = 4.50$			
no trim	-.0209	.7329	.0739	.9717	no trim	-.0058	.7474	.0449	1.047
trim-obc	-.0196	.7316	.0699	.9435	trim-obc	-.0035	.7464	.0430	.5494
	Normal Error					Normal Error			
no trim	-.0236	.7329	.0786	1.014	no trim	-.0081	.7454	.0447	1.136
trim-obc	-.0218	.7344	.0764	1.107	trim-obc	-.0080	.7444	.0440	.9054

	T = 500					**T = 1000**			
	Bias	Med	RMSE	$KS_{.05}$		Bias	Med	RMSE	$KS_{.05}$
	Pareto Error: $\kappa = 2.25$					Pareto Error: $\kappa = 2.25$			
no trim	-.0030	.7490	.0298	1.435	no trim	-.0013	.7494	.0331	1.386
trim-obc	-.0039	.7450	.0391	1.244	trim-obc	-.0024	.7468	.0319	1.174
	Pareto Error: $\kappa = 4.50$					Pareto Error: $\kappa = 4.50$			
no trim	-.0034	.7473	.0316	.9562	no trim	-.0023	.7488	.0210	.8675
trim-obc	-.0031	.7473	.0303	.3667	trim-obc	-.0017	.7495	.0209	.6421
	Normal Error					Normal Error			
no trim	-.0049	.7466	.0311	.6539	no trim	-.0007	.7502	.0212	.5285
trim-obc	-.0044	.7473	.0317	.7014	trim-obc	-.0010	.7485	.0206	.5841

$\sigma_0^2 = 1$

	T = 100					**T = 250**			
	Bias	Med[b]	RMSE	$KS^c_{.05}$		Bias	Med	RMSE	$KS_{.05}$
	Pareto Error: $\kappa = 2.25$					Pareto Error: $\kappa = 2.25$			
no trim[d]	.0187	1.011	.2632	3.131	no trim	-.0061	.9945	.2300	1.761
trim-obc	-.0029	.9896	.3973	1.846	trim-obc	-.0094	.9011	.3239	1.656
	Pareto Error: $\kappa = 4.50$					Pareto Error: $\kappa = 4.50$			
no trim	.0062	1.005	.0803	1.475	no trim	.0009	.9984	.0516	1.321
trim-obc	.0057	1.003	.0909	1.020	trim-obc	.0012	.9956	.1034	1.132
	Normal Error					Normal Error			
no trim	-.0020	.9914	.1413	.8621	no trim	.0035	1.000	.0932	.5672
trim-obc	.0035	1.000	.1475	.8798	trim-obc	.0019	1.001	.0899	.3610

	T = 500					**T = 1000**			
	Bias	Med	RMSE	$KS_{.05}$		Bias	Med	RMSE	$KS_{.05}$
	Pareto Error: $\kappa = 2.25$					Pareto Error: $\kappa = 2.25$			
no trim	.0016	.9975	.1927	2.817	no trim	.0059	1.002	.1756	1.931
trim-obc	.0029	.9905	.2887	1.369	trim-obc	.0070	1.000	.2954	1.314
	Pareto Error: $\kappa = 4.50$					Pareto Error: $\kappa = 4.50$			
no trim	.0008	1.000	.0036	1.3965	no trim	.0015	1.001	.0246	.9644
trim-obc	.0012	.9997	.0380	1.151	trim-obc	.0014	1.001	.0241	.8636
	Normal Error					Normal Error			
no trim	-.0044	.9976	.0599	.4404	no trim	-.0005	.9994	.0436	.7449
trim-obc	-.0049	.9968	.0601	.3193	trim-obc	-.0060	.9998	.0433	.5015

a. The model is $y_t = \phi_0 y_{t-1} + \epsilon_t$ where ϵ_t is iid Pareto or Normal, $E[\epsilon_t] = 0$ and $\sigma_0^2 \equiv E[\epsilon_t^2]$. b. "Med" is the median, and "RMSE" is the root-mean-squared-error.
c. "$KS_{.05}$" is the ratio of Kolmogorov-Smirnov statistic divided by its 5% critcal value. Values *greater* than one suggest non-normality at the 5% level.
d. "no-trim" is standard FD-QML. "trim-obc" is the optimal bias corrected tail-trimmed FD-QML

Table E.2 : FD-QML for AR(1) : $\phi_0 = .00$ and $\sigma_0^2 = 1$

$\phi_0 = .00^a$

	T = 100					**T = 250**			
	Bias	Med[b]	RMSE	$KS^c_{.05}$		Bias	Med	RMSE	$KS_{.05}$
	Pareto Error: $\kappa = 2.25$					Pareto Error: $\kappa = 2.25$			
no trim[d]	.0005	.0009	.0943	1.240	no trim	-.0021	-.0024	.0604	1.167
trim-obc	-.0016	-.0015	.1142	.9675	trim-obc	-.0023	-.0040	.0702	1.084
	Pareto Error: $\kappa = 4.50$					Pareto Error: $\kappa = 4.50$			
no trim	-0016	.0001	.0996	.9567	no trim	.0039	.0050	.0629	.8657
trim-obc	.0003	-.0007	.0947	.3225	trim-obc	.0018	.0033	.0603	.5423
	Normal Error					Normal Error			
no trim	.0022	-.0013	.0992	.6427	no trim	.0011	.0009	.0619	.3614
trim-obc	.0024	.0035	.1009	.4639	trim-obc	-.0008	.0001	.0625	.4343

	T = 500					**T = 1000**			
	Bias	Med	RMSE	$KS_{.05}$		Bias	Med	RMSE	$KS_{.05}$
	Pareto Error: $\kappa = 2.25$					Pareto Error: $\kappa = 2.25$			
no trim	-.0007	.0004	.0419	1.112	no trim	-.0006	-.0014	.0305	1.283
trim-obc	.0005	.0006	.0467	.7385	trim-obc	.0007	.0007	.0318	.4988
	Pareto Error: $\kappa = 4.50$					Pareto Error: $\kappa = 4.50$			
no trim	.0017	.0023	.0450	.7542	no trim	-.0012	-.0014	.0315	.6453
trim-obc	-.0012	-.0017	.0431	.4251	trim-obc	.0003	.0001	.0303	.3452
	Normal Error					Normal Error			
no trim	.0007	-.0008	.0442	.4949	no trim	.0025	.0022	.0303	.3592
trim-obc	.0007	.0009	.0455	.6124	trim-obc	.0008	.0010	.0317	.5007

$\sigma_0^2 = 1$

	T = 100					**T = 250**			
	Bias	Med[b]	RMSE	$KS^c_{.05}$		Bias	Med	RMSE	$KS_{.05}$
	Pareto Error: $\kappa = 2.25$					Pareto Error: $\kappa = 2.25$			
no trim[d]	.0069	1.006	.2622	1.167	no trim	-.0121	.9905	.2271	1.845
trim-obc	-.0094	.9942	.2944	.9967	trim-obc	-.0354	.9889	.2515	1.584
	Pareto Error: $\kappa = 4.50$					Pareto Error: $\kappa = 4.50$			
no trim	-.0047	.9961	.0732	1.2784	no trim	-.0048	.9955	.0622	.7787
trim-obc	-.0012	.9972	.0817	1.070	trim-obc	-.0032	.9964	.0494	.7846
	Normal Error					Normal Error			
no trim	-.0136	.9769	.1389	.8727	no trim	-.0025	.9937	.0904	.5458
trim-obc	.0133	.8717	.1400	.9227	trim-obc	-.0021	.9943	.0907	.6142

	T = 500					**T = 1000**			
	Bias	Med	RMSE	$KS_{.05}$		Bias	Med	RMSE	$KS_{.05}$
	Pareto Error: $\kappa = 2.25$					Pareto Error: $\kappa = 2.25$			
no trim	-.0008	.9953	.1954	1.986	no trim	.0062	1.004	.1766	1.931
trim-obc	-.0017	.9911	.2090	1.495	trim-obc	-.0053	.9983	.1935	1.733
	Pareto Error: $\kappa = 4.50$					Pareto Error: $\kappa = 4.50$			
no trim	-.0015	.9970	.0340	.7846	no trim	.0001	.9999	.0242	.9145
trim-obc	-.0011	.9979	.0347	.9595	trim-obc	.0002	1.000	.0246	.9075
	Normal Error					Normal Error			
no trim	-.0075	.9948	.0656	.4584	no trim	-.0017	.9978	.0433	.3097
trim-obc	-.0069	.9954	.0596	.4689	trim-obc	-.0016	.9976	.0433	.3422

a. The model is $y_t = \phi_0 y_{t-1} + \epsilon_t$ where ϵ_t is iid Pareto or Normal, $E[\epsilon_t] = 0$ and $\sigma_0^2 \equiv E[\epsilon_t^2]$. b. "Med" is the median, and "RMSE" is the root-mean-squared-error.
c. "$KS_{.05}$" is the ratio of Kolmogorov-Smirnov statistic divided by its 5% critcal value. Values *greater* than one suggest non-normality at the 5% level.
d. "no-trim" is standard FD-QML. "trim-obc" is the optimal bias corrected tail-trimmed FD-QML

Table E.3 : FD-QML for GARCH(1,1)[a], T = 2000

$[\alpha_0, \beta_0] = [.3, .6]$: $\kappa \approx 2.05$									
$\alpha_0 = .3$					$\beta_0 = .6$				
	Bias	Med[b]	RMSE	$KS^c_{.05}$		Bias	Med	RMSE	$KS_{.05}$
no trim[d]	.0367	.2929	.1988	2.593	no trim	.0028	.6201	.1791	2.307
trim-obc	.0041	.2912	.2195	1.687	trim-obc	.0014	.6372	.2092	1.671

$[\alpha_0, \beta_0] = [.2, .7]$: $\kappa \approx 3.02$									
$\alpha_0 = .2$					$\beta_0 = .7$				
	Bias	Med	RMSE	$KS_{.05}$		Bias	Med	RMSE	$KS_{.05}$
no trim	.0517	.2077	.1687	3.318	no trim	.0078	.7105	.1282	2.498
trim-obc	.0309	.2101	.1678	1.917	trim-obc	.0019	.7100	.1401	1.674

$[\alpha_0, \beta_0] = [.05, .9]$: $\kappa \approx 4.05$									
$\alpha_0 = .05$					$\beta_0 = .9$				
	Bias	Med	RMSE	$KS_{.05}$		Bias	Med	RMSE	$KS_{.05}$
no trim	.0345	.0570	.0850	4.657	no trim	-.0171	.8979	.0889	3.837
trim-obc	.0196	.0564	.0458	1.363	trim-obc	.0115	.9198	.0470	1.266

a. The model is $y_t = x_t^2 = \sigma_t^2 \epsilon_t^2$ where $\sigma_t^2 = 1 + \alpha_0 x_{t-1}^2 + \beta_0 \sigma_{t-2}^2$ where ϵ_t is iid standard normal.

b. "Med" is the median, and "RMSE" is the root-mean-squared-error.

c. "$KS_{.05}$" is the ratio of Kolmogorov-Smirnov statistic divided by its 5% critical value. Values *greater* than one suggest non-normality at the 5% level.

d. "no-trim" is standard FD-QML. "trim-obc" is the optimal bias corrected tail-trimmed FD-QML.

References

Dunsmuir, W., 1979. A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed with noise. Annals of Statistics 7, 490–506.

Hannan, E., 1973. The asymptotic theory of linear time-series models. Journal of Applied Probability 10, 130–145.

Hill, J. B., 2012. Least tail-trimmed squares for infinite variance autoregressions. Journal of Time Series Analysis 34, 168–186.

McLeish, D., 1975. A maximal inequality and dependent strong laws. Annals of Probability 3, 829–839.

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