Robust Minimum Distance Estimation for Nonlinear GARCH Processes

Robust Minimum Distance Estimation for Nonlinear Semi-Strong GARCH Models

Jonathan B. Hill
Dept. of Economics
University of North Carolina – Chapel Hill
PROPOSAL

1. Estimate parameters of Nonlinear ARX - Nonlinear GARCH robustly.
   Includes non-stationary GARCH (almost no parameter restrictions).

2. Minimum Distance Estimator is asymptotically normal under minimal restrictions on model parameters and errors, including arbitrarily heavy-tailed processes (for any reason).

3. Asymptotic normality is assured by tail trimming functions of the data.
PROPOSAL

Stylized traits of financial time series:

Nonlinear – asymmetries in returns.

Heavy-tailed: unobserved shocks; and/or due to parametric structure.

\((\varepsilon_t)\) (IGARCH)
PROPOSAL

Clusters: not as severe as IGARCH…

…but at least infinite fourth moment.
PROPOSAL

MODEL – Nonlinear ARX-Nonlinear GARCH

\[ y_t = f(x_{t-1}, \alpha) + u_t \quad \text{and} \quad u_t = h_t(\theta)\varepsilon_t, \quad \theta = [\alpha', \beta']' \in \mathbb{R}^k \]

where \( f \) and \( h_t \) are differentiable in \( \theta \): excludes TAR’s, etc.
PROPOSAL

MODEL – Nonlinear ARX-Nonlinear GARCH

\[ y_t = f(x_{t-1}, \alpha) + u_t \quad \text{and} \quad u_t = h_t(\theta) \varepsilon_t, \quad \theta = [\alpha', \beta']' \in \mathbb{R}^k \]

where \( f \) and \( h_t \) are differentiable in \( \theta \): excludes TAR’s, etc.

\[ x_t \in \mathbb{R}^k \] contains \( y_t, y_{t-1}, \ldots, \) and other variables and their lags.

\[ \mathcal{F}_t := \sigma(x_t : \tau \leq t). \]

\[ E[y_t \mid \mathcal{F}_{t-1}] = f(x_{t-1}, \alpha) \quad \text{a.s. if and only if} \quad \alpha = \alpha_0. \]

\[ E[u_t^2 \mid \mathcal{F}_{t-1}] = h_t(\theta) \quad \text{a.s. if and only if} \quad \theta = \theta_0. \]
PROPOSAL

MODEL – Nonlinear ARX-Nonlinear GARCH

\[ y_t = f(x_{t-1}, \alpha) + u_t \quad \text{and} \quad u_t = h_t(\theta) \epsilon_t, \quad \theta = [\alpha', \beta']' \in \mathbb{R}^k \]

EXAMPLES

ARX(p), Nonlinear AR(p), Random Coefficient Autoregression;

All GARCH(p,q) : IGARCH, GARCH with explosive roots, …;

Nonlinear GARCH : Smooth Transition GARCH, Quadratic GARCH,…;

AR-GARCH, STAR-STGARCH, etc.
WHAT WE KNOW...

GENERALIZED METHOD OF MOMENTS [GMM]

Severe moment restrictions on \( \{ y_t, \epsilon_t \} \).

- **Autoregressions**: finite variance \( \{ \epsilon_t \} \).
- **GARCH**: at least finite kurtosis \( \{ y_t, \epsilon_t \} \).

Hansen (1982), Newey and McFadden (1994)

**Shortcomings**: unsupportable *moment restrictions* for financial time series.
WHAT WE KNOW…

QUASI-MAXIMUM LIKELIHOOD [QML]

Asymptotic normality known for:

Strong GARCH and ARMA-GARCH under $E(\varepsilon_t^4) < \infty$

(Francq and Zakoïan 2004).

Strong-GARCH, semi-strong GARCH under at least $E(\varepsilon_t^4) < \infty$

(Hansen and Lee 1994, Lumsdaine 1996)

Non-stationary strong-ARCH (Jensen and Rahbek 2004)

Shortcomings: Only linear GARCH. Error moments restrictive.
Robust Minimum Distance Estimation for Nonlinear GARCH Processes

 WHAT WE KNOW…

NASDAQ Daily Log Returns
skew : .477 (.000)
kurt : 6.48 (.000)

IGARCH(1,1) with iid N(0,1) shocks

Asymmetric…

NASDAQ Daily Log Returns
B. Hill Estimator with Robust Kernel 95% Bands

…and heavy-tailed.

Doesn’t look like IGARCH, but could have infinite variance.
WHAT WE KNOW…

LEAST ABSOLUTE DEVIATIONS

Asymptotic normality:

Log-transformed LAD for non-stationary GARCH under $E(\varepsilon_t^2) < \infty$ if $mds$, or $E|\varepsilon|^p < \infty$ for some $p > 0$ if iid.


Shortcomings: Only linear GARCH. Moment restriction in $mds$ case.

LAD known to be robust to outliers. But so is tail-trimming…
WHAT WE KNOW…

TRIMMED and WEIGHTED OLS, QML and LAD

Trimming: remove large values from estimation…

Weighting: diminish contribution of large values to criterion…

Ling (2005, 2007): Weighted LAD, QML for ARMA-GARCH


Rubert and Carroll (1980), Rousseeuw (1985), Welsh (1987),
Agulló, Croux and Van Aelst (2008).
WHAT WE KNOW...

TRIMMED and WEIGHTED OLS, QML and LAD

Trim $y_t$ itself, or trim/weight criterion based on values of $y_t$.

Trimming/weighting *symmetrically* based on fixed quantile of $y_t: |y_t| > c$.

No theory for selecting $c$ (often: *no simulation work*)

*Non-time series and GARCH: linear models.*
WHAT WE KNOW…

EXAMPLE: Quasi-Maximum Trimmed Likelihood (Čižek 2008)

NLARX(p) : \( y_t = f(x_{t-1}, \alpha) + \varepsilon_t \)

\[ E \left| y_t \right|^p < \infty \text{ for some } p > 0 \]

QML : \( Q_n(\theta) = -\sum_{i=1}^{n} \ln \phi_t(\theta) \text{ where } \phi_t(\theta) := \frac{1}{\sigma} \phi \left( \frac{(y_t - f(x_{t-1}, \alpha))^2}{\sigma^2} \right) \)

QMTL : \( Q_n(\theta) = -\sum_{i=k+1}^{n} \ln \phi_{(i)}(\theta) \text{ where } \phi_{(1)} \geq \phi_{(2)} \geq \ldots, \text{ and } k ? \)
WHAT WE KNOW...

TAIL-TRIMMED GMM (Hill and Renault 2008)

Leading example of the following class of Tail-Trimmed MDE’s.

Contribution in present paper:

Extensive proofs for Nonlinear ARX-Nonlinear-GARCH

Main assumptions there are here turned into fundamental proofs.

(tail trimming impact, tail shape)

Optimally select the number of trimmed observations (unique?).
TAIL-TRIMMED MDE – Set Up

\[ y_t = f(x_{t-1}, \alpha) + u_t \quad \text{and} \quad u_t = h_t(\theta) \varepsilon_t, \quad \theta = [\alpha', \beta']' \in \mathbb{R}^k \]

\[ \mathcal{I}_t := \sigma(x_\tau : \tau \leq t) \]

Estimating equations \( m_t(\theta) \) map \( m_t : \mathbb{R}^k \rightarrow \mathbb{R}^s, \ s \geq k \)

Identification: \( E[m_t(\theta)] = 0 \ if \ and \ only \ if \ \theta = \theta_0 \)

MM criterion: \( Q_n(\theta) = \left( \frac{1}{n} \sum_{t=1}^{n} m_t(\theta) \right)' \hat{\Omega} \times \left( \frac{1}{n} \sum_{t=1}^{n} m_t(\theta) \right) \)

where \( \hat{\Omega} \) is positive semi-definite.
TAIL-TRIMMED MDE – Set Up

\[ y_t = f(x_{t-1}, \alpha) + u_t \quad \text{and} \quad u_t = h_t(\theta)\varepsilon_t, \quad \theta = [\alpha', \beta']' \in \mathbb{R}^k \]

Asymptotic problem:

\[ \hat{\theta} = \arg \min \{ Q_n(\theta) \} \]

\[ \sqrt{n} \left( \hat{\theta} - \theta_0 \right) = A_n \times \frac{1}{\sqrt{n}} \sum_{t=1}^{n} m_t(\theta_0) + o_p(1) \]

\[ \begin{array}{c}
  k \times s \\
  s \times 1
\end{array} \]

Estimating equations may be heavy-tailed even if \( y_t \) and \( \varepsilon_t \) are not.
TAIL-TRIMMED MDE – Estimation Equations

GENERALIZED METHOD OF MOMENTS

Estimating equations are at least:

\[
m_t(\theta) = \begin{bmatrix} 
\{y_t - f(x_{t-1}, \alpha)\} \times g_{t-1}^{(1)} \\
\{(y_t - f(x_{t-1}, \alpha))^2 - h_t(\theta)\} \times g_{t-1}^{(2)}
\end{bmatrix}
\]

where \( g_{t-1}^{(i)} \) are \( \mathcal{F}_{t-1} \) - measurable.
TAIL-TRIMMED MDE – Estimation Equations

GENERALIZED METHOD OF MOMENTS

Example: GMM estimating equations for pure ARCH(1)

\[
m_t(\theta) = \left[ \{ y_t^2 - h_t(\theta) \} \times g_{t-1}^{(2)} \right] = \begin{bmatrix}
y_t^2 - \beta_0 - \beta_1 y_{t-1}^2 \\
\{ y_t^2 - \beta_0 - \beta_1 y_{t-1}^2 \} \times y_{t-1}^2 \\
\{ y_t^2 - \beta_0 - \beta_1 y_{t-1}^2 \} \times y_{t-1}^4
\end{bmatrix}
\]

If \( E[y_t^8] = \infty \) then \( E[m_{3,t}^2(\theta_0)] = \infty \): Gaussian asymptotics fail.
TAIL-TRIMMED MDE – Estimation Equations

QUASI-MAXIMUM LIKELIHOOD

Estimating equations are:

\[ m_t(\theta) = \frac{\partial}{\partial \theta} \ln \phi_t(\theta) \]

where

\[ \phi_t(\theta) = \frac{1}{h_t(\theta)} \exp \left\{ -\frac{1}{2} \left( \frac{(y_t - f(x_{t-1}, \alpha))^2}{h_t^2(\theta)} \right) \right\} \]
TAIL-TRIMMED MDE – Estimation Equations

We require \( \{ \varepsilon_t, \mathcal{F}_t \} \) to form a martingale difference sequence.

We only consider pure Nonlinear GARCH for LAD and LLAD:

**LEAST ABSOLUTE DEVIATIONS**

\[
m_t(\theta) = \varepsilon_t \times \text{sgn}\{\varepsilon_t^2 - 1\} \times \frac{\partial}{\partial \theta} \ln h_t^2(\theta)
\]

**LOG-LEAST ABSOLUTE DEVIATIONS** (Peng and Yao 2003)

\[
m_t(\theta) = \text{sgn}\{\ln \varepsilon_t^2\} \times \frac{\partial}{\partial \theta} \ln h_t^2(\theta)
\]
TAIL-TRIMMED MDE – Tail Trimmed Equations

\[ y_t = f(x_{t-1}, \alpha) + u_t \quad \text{and} \quad u_t = h_t(\theta) \epsilon_t, \quad \theta = [\alpha', \beta']' \in \mathbb{R}^k \]

Tail - values: \[ m_{i,t}^{(-)} := m_{i,t} \times I(m_{i,t} < 0) \quad \text{and} \quad m_{i,t}^{(+)} := m_{i,t} \times I(m_{i,t} > 0) \]

# Trimmed m's: \[ k_{1,n} = \text{left}, \quad k_{2,n} = \text{right} : k_{i,n} \rightarrow \infty, \quad \frac{k_{i,n}}{n} \rightarrow 0. \]

Thresholds: \[ \frac{n}{k_{1,n}} P(m_{i,t}^{(-)} < -l_{i,n}) \rightarrow 1 : l_{i,n} \text{ is lower } k_{1,n} / n^{th} \rightarrow 0 \text{- quantile.} \]

\[ \frac{n}{k_{2,n}} P(m_{i,t}^{(+)} > u_{i,n}) \rightarrow 1 : u_{i,n} \text{ is lower } k_{2,n} / n^{th} \rightarrow 0 \text{- quantile.} \]
TAIL-TRIMMED MDE – Tail Trimmed Equations

\[ y_t = f(x_{t-1}, \alpha) + u_t \quad \text{and} \quad u_t = h_t(\theta) \varepsilon_t, \quad \theta = [\alpha', \beta']' \in \mathbb{R}^k \]

Thresholds : \[
\frac{n}{k_{1,n}} P\left( m_{i,t}^{(-)} < -l_{i,n} \right) \to 1 : \quad l_{i,n} \text{ is lower } k_{1,n} / n^{th} \to 0 \text{ quantile.}
\]

\[
\frac{n}{k_{2,n}} P\left( m_{i,t}^{(+)} > u_{i,n} \right) \to 1 : \quad u_{i,n} \text{ is upper } k_{2,n} / n^{th} \to 0 \text{ quantile.}
\]

Trimmed \( m \) : \[
m_{n,t}^{\prime} (\theta) = \left[ m_{i,t} (\theta) \times I(-l_{i,n} < m_{i,t} (\theta) < u_{i,n}) \right]_{i=1}^s
\]

Need to estimate \( l_{i,n} \) and \( c_{i,n} \) for practical applications.
TAIL-TRIMMED MDE – Tail Trimmed Equations

\[ y_t = f(x_{t-1}, \alpha) + u_t \quad \text{and} \quad u_t = h_t(\theta)e_t, \quad \theta = [\alpha', \beta']' \in \mathbb{R}^k \]

Tail values: \( m_{i,t}^{(-)} := m_{i,t} \times I(m_{i,t} < 0) \) and \( m_{i,t}^{(+) := m_{i,t} \times I(m_{i,t} > 0) \}

Order statistics: \( m_{i,(1)}^{(-)} \leq m_{i,(2)}^{(-)} \ldots \) and \( m_{i,(1)}^{(+) \geq m_{i,(2)}^{(+) \ldots}} \)

Trimmed \( m \): \( m_{n,t}(\theta) = \left[ m_{i,t}(\theta) \times I(-l_{i,n} < m_{i,t}(\theta) < u_{i,n}) \right]_{i=1}^s \)

Sample trimmed \( m \): \( \hat{m}_{n,t}(\theta) := \left[ m_{i,t} \times I(m_{i,(k_{1,n}+1)}^{(-)} < m_{i,t} < m_{i,(k_{2,n}+1)}^{(+)}) \right]_{i=1}^s \)

(removes \( k_{n,1} \) and \( k_{n,2} \) largest negative and positive \( m \)'s)

Jonathan B. Hill, Dept. of Economics, University of North Carolina – Chapel Hill, jbhill@email.unc.edu 24
TAIL-TRIMMED MDE – Tail Trimmed MDE

\[ y_t = f(x_{t-1}, \alpha) + u_t \quad \text{and} \quad u_t = h_t(\theta) \varepsilon_t, \quad \theta = [\alpha', \beta']' \in \mathbb{R}^k \]

Sample trimmed \( m : \hat{m}_{n,t}(\theta) := \left[ m_{i,t} \times I(m_{i,(k_{1,n}+1)} < m_{i,t} < m_{i,(k_{2,n}+1)}) \right]_{i=1}^s \)

TTMDE Criterion: \( \hat{Q}_n(\theta) = \left( \frac{1}{n} \sum_{t=1}^n \hat{m}_{n,t}(\theta) \right)' \times \hat{\Omega}_n \times \left( \frac{1}{n} \sum_{t=1}^n \hat{m}_{n,t}(\theta) \right) \)

TTMDE: \( \hat{\theta}_n = \arg \min_{\theta \in \Theta} \left\{ \hat{Q}_n(\theta) \right\} \)
ASYMPTOTIC THEORY FOR TTMDE

MAIN RESULT

Under A1-A4 and B1-B4 (below),

\[ V_n^{1/2} (\hat{\theta}_n - \theta_0) = A_n \times \sum_{t=1}^{n} m_{n,t}(\theta_0) + o_p(1) \xrightarrow{d} N(0, I_k) \]

for some matrix sequences \{V_n, A_n\}. 
ASYMPTOTIC THEORY FOR TTMDE

MAIN RESULT

Under A1-A4 and B1-B4 (below),

\[ V_{n}^{1/2} (\hat{\theta}_{n} - \theta_{0}) = A_{n} \times \sum_{t=1}^{n} m_{n,t} (\theta_{0}) + o_{p} (1) \overset{d}{\to} N(0, I_{k}) \]

\[ v_{n}^{2} := \| V_{n} \| \text{ reveals convergence rate of TTMDE vector.} \]

(Antoinne and Renault 2008 – GMM, linear combo’s, rate).
ASYMPTOTIC THEORY FOR TTMDE

MAIN RESULT

Under $A1$-$A4$ and $B1$-$B4$ (below),

$$V_n^{1/2} (\hat{\theta}_n - \theta_0) = A_n \times \sum_{t=1}^{n} m_{n,t} (\theta_0) + o_p(1) \overset{d}{\to} N(0, I_k)$$

If $\| m_t (\theta_0) \|_2 < \infty$ then $v_n \sim Kn^{1/2}$. 
ASYMPTOTIC THEORY FOR TTMDE

MAIN RESULT

Under A1-A4 and B1-B4 (below),

\[
V_n^{1/2} (\hat{\theta}_n - \theta_0) = A_n \times \sum_{t=1}^{n} m_{n,t} (\theta_0) + o_p(1) \overset{d}{\to} N(0, I_k)
\]

Asymptotically efficient \( \Omega_n = \Sigma_n^{-1} \Rightarrow V_n = n^2 \left( G_n \Sigma_n^{-1} G_n \right) \),

where \( \Sigma_n = n \times E[m_{n,t} (\theta_0) m_{n,t} (\theta_0)^\prime] \) and \( G_n = \frac{\partial}{\partial \theta} E[m_{n,t} (\theta)] \Big|_{\theta=\theta_0} \).
ASYMPTOTIC THEORY FOR TTMDE

MAIN RESULT

Under A1-A4 and B1-B4 (below),

\[ V_n^{1/2} (\hat{\theta}_n - \theta_0) = A_n \times \sum_{t=1}^{n} m_{n,t} (\theta_0) + o_p(1) \overset{d}{\rightarrow} N(0, I_k) \]

If \( k = s = 1 \) then \( v_n = \|V_n\|^{1/2} = \sqrt{n} \times \left( \frac{|G_n|}{\|m_{n,t}(\theta_0)\|_2} \right) \)

where \( G_n = \frac{\partial}{\partial \theta} E[m_{n,t}(\theta)] \bigg|_{\theta=\theta_0} \)
ASSUMPTIONS – A’s ensure $Q_n$ is sufficiently smooth.

A1. $\| \hat{\Omega}_n - \Omega \|^p \to 0$, where $\hat{\Omega}_n$ and $\Omega$ are p.s.d.

A2. i. $\hat{Q}_n (\hat{\theta}_n) \leq \inf_{\theta \in \Theta} \hat{Q}_n (\theta) + o_p (v_n^{-2})$ where $v_n^2 = \| V_n \| \to \infty$

   ii. $\sup_{\| \theta - \theta_0 \| > \delta} \left\{ \hat{Q}_n (\theta)^{-1} \right\} = O_p (1) \ \forall \delta > 0$. 

A3. $\frac{\partial}{\partial \theta} E[m_{n,t} (\theta)]$ exists $\forall \theta \in U_{\theta_0}$, an arbitrary neighborhood of $\theta_0$.

A4. $\sup_{\| \theta - \theta_0 \| \leq \delta_n} \left\{ v_n \left\| \frac{m_n (\Theta) - E[m_{n,t} (\theta)] - \overline{m}_n (\theta_0)}{1 + v_n \| \theta - \theta_0 \|} \right\|^p \right\} \to 0$ as $\delta_n \to 0$.

ASSUMPTIONS – B’s restrict dgp’s \{ m_t, m_{n,t}, x_t \}

B1. \{m_{n,t}(\theta), \mathcal{Z}_t\} is an adapted martingale difference array iff \( \theta = \theta_0 \).

\[ P(\|m_{i,t}(\theta_0)\| > z) \leq 1 \text{ where } L \text{ is slowly varying, } \kappa_i > 0. \]

B2. \[ \lim_{z \to \infty} \frac{P(\|m_{i,t}(\theta_0)\| > z)}{z^{-\kappa_i} L(z)} \leq 1 \text{ where } L \text{ is slowly varying, } \kappa_i > 0. \]

B3. i. \( E[m_{n,t}(\theta_0)m_{n,t}(\theta_0)'] \) is p.d. \( \forall n \geq N; \)

\[ \text{ii. } \liminf_{n \geq 1} \frac{n \|E[m_{n,t}(\theta_0)m_{n,t}(\theta_0)']\|}{\max \{ \max\{l_{i,n}, u_{i,n}\} \}^2} > 0, \max_{1 \leq i \leq s} \left\{ \max\{l_{i,n}, u_{i,n}\} \right\}^2 = o(n). \]

B4. Regressors \( x_t \) are stationary and strong mixing with size 1.
EXAMPLES: each satisfy martingale difference B1 and tail bound B2

1. Random Coefficient Autoregression

\[ y_t = \alpha_t y_{t-1} + \varepsilon_t, \quad |\alpha_t| \leq \alpha \in (0, 1), \quad \alpha_t \text{ is } \mathcal{F}_{t-1} - \text{measurable.} \]

\[ \varepsilon_t \text{ is mds and } \lim_{\varepsilon \to \infty} P(|\varepsilon_t| > \varepsilon) / \varepsilon^\kappa \leq K \text{ for some } \kappa > 0. \]

\[ \text{mds and } \lim_{\varepsilon \to \infty} P(\varepsilon_t > \varepsilon | \mathcal{F}_{t-1}) / \varepsilon^{\tilde{\kappa}} \leq K \]
EXAMPLES: each satisfy *martingale difference* B1 and *tail bound* B2

2. **AR(p)-GARCH(1,1)**

\[ y_t = \sum_{i=1}^{p} \alpha_i y_{t-i} + u_t \] where \( u_t = h_t \varepsilon_t \) and \( h_t^2 = \beta_0 + \beta_1 u_{t-1}^2 + \beta_2 h_{t-1}^2 \)

AR - roots outside unit circle.

\( \varepsilon_t \) is *mds* and \( \lim_{\varepsilon \to \infty} P(|\varepsilon_t| > \varepsilon)/\varepsilon^{\kappa_i} \leq K \) for some \( \kappa > 0 \).

\( E|\varepsilon_t|^\iota < 1 \) for some infinitessimal \( \iota > 0 \) and \( \beta_i \|\varepsilon_t\|^{1/\kappa}_t < 1, \ i = 1, 2. \)

Covers nearly all GARCH since \( \iota > 0 \) is arbitrary.
EXAMPLES: each satisfy *martingale difference* $B_1$ and *tail bound* $B_2$

3. **Threshold GARCH**

\[
y_t = h_t \varepsilon_t \quad \text{where} \quad h_t^2 = \beta_0 + \left( \beta_1 y_{t-1}^2 + \beta_2 h_{t-1}^2 \right) I(y_{t-1} < 0)
\]

$\varepsilon_t$ is *mds* and \( \lim_{\varepsilon \to \infty} P(|\varepsilon_t| > \varepsilon) / \varepsilon^{\kappa_i} \leq K \) for some $\kappa > 0$.

\[
E|\varepsilon_t|^t < 1 \text{ for some infinitessimal } t > 0 \text{ and } \beta_i \|\varepsilon_t\|_{t}^{1/\kappa} < 1, \ i = 1, 2.
\]

Covers nearly all GARCH since $t > 0$ is arbitrary.
OPTIMAL FRACTILE SELECTION

Literature

Ling (2005, 2007) : Least Absolute Weighted Deviations, QMWL

Čižek (2008) : Generalized Trimmed Estimators

Trim $k$ obs.’s or weight based on $|y_t| > c$ : $k$ and $c$ unknown.

No theory for selecting trimming or weighting criterion.


Aguilar, Hill and Renault (2008) : Truncated Simulated MM – $mse(c)$.

OPTIMAL FRACTILE SELECTION

**New theory**: select trimming fractiles \( \{ k_{1,n}, k_{2,n} \} \) by combination grid search and using untrimmed criterion.

**Assume**: \( k_{1,n} = \left[n^{\delta_1}\right] \) and \( k_{2,n} = \left[n^{\delta_2}\right] \) where \( \delta = [\delta_1, \delta_2] \in (0,1)^2 \)

If symmetric process (e.g. AR, GARCH, etc.) : \( \delta_1 = \delta_2 \).

Uncountably infinitely many \( \delta, \theta : E[m_{n,t}(\theta)] = 0!! \)

**Re-write**: \( \hat{m}_{n,t}^{(\delta)}(\theta) := \left[m_{i,t} \times I(m_{i,([n^{\delta_1}]+1)}^{(-)} < m_{i,t} < m_{i,([n^{\delta_2}]+1)}^{(+)}\right]_{i=1}^s \)
OPTIMAL FRACTILE SELECTION

Assume : $k_{1,n} = \lfloor n^{\delta_1} \rfloor$ and $k_{2,n} = \lfloor n^{\delta_2} \rfloor$ where $\delta = [\delta_1, \delta_2]' \in (0,1)^2$

Re-write : $\hat{m}_{n,t}^{(\delta)}(\theta) := \left[ m_{i,t} \times I(m_{i,([n^{\delta_1}] + 1)}^{(-)} < m_{i,t} < m_{i,([n^{\delta_2}] + 1)}^{(+)} \right]_{i=1}^s$

$\hat{Q}_n^{(\delta)}(\theta) = \left( \frac{1}{n} \sum_{t=1}^n \hat{m}_{n,t}^{(\delta)}(\theta) \right)' \hat{\Omega}^{(\delta)} \times \left( \frac{1}{n} \sum_{t=1}^n \hat{m}_{n,t}^{(\delta)}(\theta) \right)$

$\hat{\theta}_n(\delta) = \arg \min_{\theta \in \Theta} \left\{ \hat{Q}_n^{(\delta)}(\theta) \right\}$
Robust Minimum Distance Estimation for Nonlinear GARCH Processes

OPTIMAL FRACTILE SELECTION

Assume: \( k_{1,n} = \left[ n^{\delta_1} \right] \) and \( k_{2,n} = \left[ n^{\delta_2} \right] \) where \( \delta = [\delta_1, \delta_2]' \in (0,1)^2 \)

Untrimmed criterion: \( Q_n (\theta) = \left( \frac{1}{n} \sum_{t=1}^{n} m_t (\theta) \right)' \hat{\Omega} \times \left( \frac{1}{n} \sum_{t=1}^{n} m_t (\theta) \right) \)

Select fractile: \( \hat{\delta}_n = \arg \min_{\delta \in D} \left\{ Q_n (\hat{\theta}_n (\delta)) \right\} \) for compact \( D \subset (0,1)^2 \)

TTMDE: \( \hat{\theta}_n (\hat{\delta}_n) \)... Claim: \( \hat{\theta}_n (\hat{\delta}_n) \stackrel{p}{\rightarrow} \theta_0 \).

Notice: select \( \delta \) such that untrimmed equation mean is close to zero.
## SIMULATION STUDY

<table>
<thead>
<tr>
<th>MODEL</th>
<th>$\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AR</strong> : $y_t = 0.9 y_{t-1} + \varepsilon_t$, $\varepsilon_t \sim P_{1.5}$</td>
<td>1.50</td>
</tr>
<tr>
<td><strong>GARCH</strong> : $y_t = h_t \varepsilon_t$, $h_t^2 = 0.3 + 0.3 y_{t-1}^2 + 0.6 h_{t-1}^2$, $\varepsilon_t \sim P_{2.5}$</td>
<td>2.14</td>
</tr>
<tr>
<td><strong>IGARCH</strong> : $y_t = h_t \varepsilon_t$, $h_t^2 = 0.3 + 0.4 y_{t-1}^2 + 0.6 h_{t-1}^2$, $\varepsilon_t \sim N_{0,1}$</td>
<td>2.00</td>
</tr>
<tr>
<td><strong>QARCH</strong> : $y_t = h_t \varepsilon_t$, $h_t^2 = (0.3 + 0.8 y_{t-1})^2$, $\varepsilon_t \sim P_{2.5}$</td>
<td>2.88</td>
</tr>
<tr>
<td><strong>QARCH</strong> : $y_t = h_t \varepsilon_t$, $h_t^2 = (0.3 + 0.8 y_{t-1})^2$, $\varepsilon_t \sim N_{0,1}$</td>
<td>3.53</td>
</tr>
</tbody>
</table>
SIMULATION STUDY

- Simulate 100 series of sample size \( n = 1000 \).

- QML

- GMM and TTGMM – weight matrix \( \Omega = I \).

- Inspect small sample distribution of \( k^{th} \)-parameter estimate:

\[
\text{Kolmogorov-Smirnov tests of normality over 100 estimates.}
\]

\[
\text{TTGMM (AR & GARCH)}: \quad k_n = [n^\delta], \quad \text{min. KS over } \delta \in \{.01, ..., .99\}
\]

\[
\text{TTGMM (QARCH)}: \quad k_{1,n} = [n^{\delta_1}], \quad k_{2,n} = [n^{\delta_2}]
\]

\[
\text{min. KS over } \delta_i \in \{.01, ..., .99\}.
\]
## SIMULATION STUDY

<table>
<thead>
<tr>
<th></th>
<th>MEAN</th>
<th>STDEV</th>
<th>Z</th>
<th>KS</th>
<th>MEAN</th>
<th>STDEV</th>
<th>Z</th>
<th>KS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IID, mean = 1, ( P_{1.5} )</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QML</td>
<td>1.01</td>
<td>.151</td>
<td>.07</td>
<td>.272</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMM</td>
<td>.939</td>
<td>.212</td>
<td>.08</td>
<td>.172</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTGMM</td>
<td>.992</td>
<td>.139</td>
<td>.06</td>
<td>.052*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AR(1), slope = .9, ( P_{1.5} )</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QML</td>
<td>.889</td>
<td>.013</td>
<td>.04</td>
<td>.253</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMM</td>
<td>.939</td>
<td>.212</td>
<td>.08</td>
<td>.172</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTGMM</td>
<td>.992</td>
<td>.139</td>
<td>.06</td>
<td>.052*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>GARCH, slope = .6, ( P_{2.5} )</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QML</td>
<td>.264</td>
<td>.184</td>
<td>.00</td>
<td>.569</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMM</td>
<td>.605</td>
<td>.346</td>
<td>.00</td>
<td>.448</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTGMM</td>
<td>.593</td>
<td>.231</td>
<td>.06</td>
<td>.096*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IGARCH, slope = .6, ( N_{0,1} )</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QML</td>
<td>.586</td>
<td>.198</td>
<td>.14</td>
<td>.262</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMM</td>
<td>.628</td>
<td>.225</td>
<td>.02</td>
<td>.273</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTGMM</td>
<td>.602</td>
<td>.177</td>
<td>.06</td>
<td>.083*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>QARCH, slope = .8, ( N_{0,1} )</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QML</td>
<td>.896</td>
<td>.659</td>
<td>.00</td>
<td>.389</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMM</td>
<td>.673</td>
<td>.093</td>
<td>.05</td>
<td>.454</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTGMM</td>
<td>.807</td>
<td>.094</td>
<td>.06</td>
<td>.068*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>QARCH, slope = .8, ( P_{2.5} )</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QML</td>
<td>.884</td>
<td>.751</td>
<td>.01</td>
<td>.412</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMM</td>
<td>.689</td>
<td>.243</td>
<td>.04</td>
<td>.486</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTGMM</td>
<td>.799</td>
<td>.185</td>
<td>.06</td>
<td>.094*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( Z \): rejection frequency of \( Z \)-test at 5\% level.

\( KS \): 1\%, 5\%, 10\% c.v.’s = .136, .122, .107; \* = fail to reject \( H_0 \): \( N(0,1) \) at 5\%.