# Robust M-Estimation for Heavy Tailed Nonlinear AR-GARCH with an Application to Financial Returns

Jonathan B. Hill

Dept. of Economics
University of North Carolina - Chapel Hill

Conference on New Developments in Econometrics and Time Series

Brussels, 2011

## ❑ **Heavy tail robust M-estimation**

**Nonlinear AR-GARCH Model**

$$y_t = f(x_t, \phi^0) + u_t \text{ where } u_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = g(u_{t-1}, \sigma_{t-1}^2, \theta^0), \ \theta^0 = [\phi^{0\prime}, \beta^{0\prime}]' \in \mathbb{R}^{p+s=q}, \text{ note } \phi^0 \in \mathbb{R}^p$$

The error $\epsilon_t$ is *iid*, $E[\epsilon_t] = 0$ and $E[\epsilon_t^2] = 1$.

The regressors $x_t \in \mathcal{R}^k$ are lags of $y_t$ (can allow others).

Response *f* and *g* are *twice differentiable* in all arguments.

Unique $\theta^0$ such that $y_t$ is stationary, ergodic.

## ❑ **Heavy tail robust M-estimation**

Methodology : negligible (i.e. tail) trimming on criterion

- *Easy to compute*
- *Standard asymptotics : Gaussian estimator limit*
- *Nearly the highest convergence rate possible*

## **Heavy tail robust M-estimation**

Methodology : negligible (i.e. tail) trimming on criterion

Tail-trimming method easily extends to *inference*

Moment Condition tests (Hill and Aguilar 10: JE)

Tests of volatility spillover (Aguilar and Hill 11)

Consistent tests of functional (Hill 2011: Springer)

Extends to *GMM* and *GEL* (Hill/Renault 10, Hill/Prokhorov 11)

## ❑ Quasi-Maximum Tail-Trimmed Likelihood

**Example: GARCH(1,1) - QMTTL**

$$\hat{\theta}_n = \underset{\theta \in \Theta}{\operatorname{arginf}} \sum_{t=1}^{n} \left\{ \ln \sigma_t^2(\theta) + \frac{y_t^2}{\sigma_t^2(\theta)} \right\} \times \hat{I}_{n,t}(\theta)$$

where $\hat{I}_{n,t}(\theta)$ negligibly trims : $\hat{I}_{n,t}(\theta) \in \{0,1\}$ and $\hat{I}_{n,t}(\theta) \xrightarrow{a.s.} 1$ *uniformly*.

## ❑ <u>**Quasi-Maximum Tail-Trimmed Likelihood**</u>

**Example: GARCH(1,1) - QMTTL**

$$\hat{\theta}_n = \underset{\theta \in \Theta}{\arg\inf} \sum_{t=1}^{n} \left\{ \ln \sigma_t^2(\theta) + \frac{y_t^2}{\sigma_t^2(\theta)} \right\} \times \hat{I}_{n,t}(\theta)$$

Negligibility eradicates *bias*, and ensures *normality*.

If $E[\varepsilon_t^4] = \infty$ : higher rate $n^{1/2}/L(n)$ than QML and QMWL (Ling 07)

Exceptional small sample properties.

Contrary to QML - Lumsdaine (95)

Straumann and Mikosch (06)

## Least Tail-Trimmed Squares

**Example: AR(*p*) - LTTS**

$$\hat{\theta}_n = \underset{\theta \in \Theta}{\text{arginf}} \sum_{t=1}^{n} (y_t - \phi' x_t)^2 \times \hat{I}_{n,t}(\phi)$$

If $u_t$ is iid and $E[u_t^2] = \infty$ then rate $n^{1/\kappa}/L(n)$ where $\kappa \in (0,2]$ is the tail index.

## Least Tail-Trimmed Squares

**Example: AR(*p*) - LTTS**

$$\hat{\theta}_n = \underset{\theta \in \Theta}{\operatorname{arginf}} \sum_{t=1}^{n} (y_t - \phi' x_t)^2 \times \hat{I}_{n,t}(\phi)$$

If $u_t$ is iid and $E[u_t^2] = \infty$ then rate $n^{1/\kappa}/L(n)$ where $\kappa \in (0,2]$ is the tail index.

Exceptional small sample properties.

Nearly the highest possible rate $n^{1/\kappa}$ (Davis, Knight, Liu 92)

## ❑ <u>**Asymptotic versus Small Sample (Other) Methods**</u>

Methods are *asymptotic* : <u>asymptotic normality</u> *and* <u>near maximum rate</u>

Log-LAD for GARCH (Peng, Yao 03, Linton et al 10)

Least Weighted Absolute Deviation for AR (Ling 05)

Quasi-Maximum Weighted Likelihood for ARMA-GARCH (Ling 07)

---

Small sample methods (e.g. Dufour, Hallin, Ibragimov : inference)

Bootstrap ($m$-out-on-$n$, $m = o(n)$): Lahiri 95, Arcones and Gine 89)

Non-Gaussian QML (e.g. Boudt and Croux 10)

## Nonlinear Least Tail-Trimmed Squares

Nonlinear AR-GARCH Model

$y_t = f(x_t, \phi^0) + u_t$ where $u_t$ may be GARCH

Intuition on trimming: **expansion around $\boldsymbol{\phi}^0$**

$$0 = \sum_{t=1}^{n} u_t f_t^{\phi} + \left\{ \sum_{t=1}^{n} u_t(\phi_*) f_t^{\phi,\phi}(\phi_*) - \sum_{t=1}^{n} f_t^{\phi}(\phi_*) f_t^{\phi}(\phi_*)' \right\} \times \left( \hat{\phi}_n - \phi^0 \right)$$

**Criterion**: $\sum_{t=1}^{n} (y_t - \phi' x_t)^2 \times \hat{I}_{n,t}(\phi)$: $\hat{I}_{n,t}(\phi)$ **reflects source of extremes**:

$u_t(\phi)$ and $f_t^{\phi}(\phi)$ and $f_t^{\phi,\phi}(\phi)$.

## Nonlinear Least Tail-Trimmed Squares

Tail-specific variables and *order statistics*:

$$w_t^{(-)}(\phi) := w_t(\phi)I(w_t(\phi) < 0) \;\; and \;\; w_{(1)}^{(-)}(\phi) \le w_{(2)}^{(-)}(\phi) \cdots$$

$$w_t^{(+)}(\phi) := w_t(\phi)I(w_t(\phi) \ge 0) \;\; and \;\; w_{(1)}^{(+)}(\phi) \ge w_{(2)}^{(+)}(\phi) \cdots$$

Number of trimmed variables: *intermediate order sequences*

$$k_{1,n}^{(w)}, k_{2,n}^{(w)} \to \infty \;\; and \;\; k_{1,n}^{(w)}, k_{2,n}^{(w)} = o(n)$$

Trimming $indicators$ : $\hat{I}_{n,t}^{(w)}(\phi) := I\left(w_{(k_{1,n}^{(w)})}^{(-)}(\phi) \le w_t(\phi) \le w_{(k_{2,n}^{(w)})}^{(+)}(\phi)\right) \in \{0,1\}$

**Trimmed variable** : $\hat{w}_{n,t}^{*}(\phi) := w_t(\phi)\hat{I}_{n,t}^{(w)}(\phi)$

- **Nonlinear Least Tail-Trimmed Squares**

**Symmetric** $w_t^{(a)}(\phi) := \lfloor w_t(\phi) \rfloor \ and \ w_{(1)}^{(a)}(\phi) \leq w_{(2)}^{(a)}(\phi) \cdots$

Number of trimmed variables: *intermediate order sequence*

$$k_n^{(w)} \to \infty \ and \ k_n^{(w)} = o(n)$$

Trimming $indicators$ : $\hat{I}_{n,t}^{(w)}(\phi) := I\left(\lfloor w_t(\phi) \rfloor \leq w_{(k_n^{(w)})}^{(a)}(\phi)\right) \in \{0,1\}$

**Trimmed variable :** $\hat{w}_{n,t}^*(\phi) := w_t(\phi)\hat{I}_{n,t}^{(w)}(\phi)$

---

**Trimming is negligible**: $\hat{I}_{n,t}^{(w)}(\phi) \to 1 \ a.s.$ hence $\hat{w}_{n,t}^*(\phi) \to w_t(\phi) \ a.s.$

## Nonlinear Least Tail-Trimmed Squares

NLLS indicators:

$$\hat{I}_{n,t}^{(u)}(\phi) = I\left(u_{(k_{1,n}^{(u)})}^{(-)}(\phi) \leq u_t(\phi) \leq u_{(k_{2,n}^{(u)})}^{(+)}(\phi)\right)$$

$$\hat{I}_{n,t}^{(f^{\phi})}(\phi) := \prod_{i=1}^{p} \hat{I}_{i,n,t}^{(f^{\phi})}(\phi) = \prod_{i=1}^{p} I\left({f^{\phi}}_{i,(k_{1,i,n}^{(f^{\phi})})}^{(-)}(\phi) \leq f_t^{\phi}(\phi) \leq {f^{\phi}}_{i,(k_{2,i,n}^{(f^{\phi})})}^{(+)}(\phi)\right)$$

$$\hat{I}_{n,t}^{(f^{\phi,\phi})}(\phi) := \prod_{1 \leq i \leq j \leq p} \hat{I}_{i,j,n,t}^{(f^{\phi,\phi})}(\phi)$$

**NLLTTS criterion** $\hat{Q}_n(\phi) := \sum_{t=1}^{n} \left(y_t - f(x_t, \phi)\right)^2 \hat{I}_{n,t}^{(u)}(\phi) \hat{I}_{n,t}^{(f^{\phi})}(\phi) \hat{I}_{n,t}^{(f^{\phi,\phi})}(\phi)$

**Estimator** $\hat{\phi}_n := \arg\inf_{\phi \in \Phi} \hat{Q}_n(\phi)$

## ❑ Nonlinear Least Tail-Trimmed Squares

Define $m_t(\phi) \coloneqq u_t(\phi) f_t^{\phi}(\phi)$ and $G_t(\phi) \coloneqq \frac{\partial}{\partial \phi} m_t(\phi)$

**Assumptions: A1-A5**

A1. $\{m_{i,t}(\phi), G_{i,j,t}(\phi)\}$ **absolutely continuous distributions**

Ensures $\hat{\phi}_n \coloneqq \arg\inf_{\phi \in \Phi} \hat{Q}_n(\phi)$ exists, and *first order condition*:

$$\sum_{t=1}^{n} u_t(\hat{\phi}_n) f_t^{\phi}(\hat{\phi}_n) \times \hat{I}_{n,t}^{(u)}(\hat{\phi}_n) \hat{I}_{n,t}^{(f^{\phi})}(\hat{\phi}_n) \hat{I}_{n,t}^{(f^{\phi,\phi})}(\hat{\phi}_n) = 0 \ \ a.s. \ \ (\textbf{only } \hat{\phi}_n)$$

Čižek (2008).

## ❑ **Nonlinear Least Tail-Trimmed Squares**

Recall $m_t(\phi) \coloneqq u_t(\phi) f_t^{\phi}(\phi)$ and $G_t(\phi) \coloneqq \frac{\partial}{\partial \phi} m_t(\phi)$

**Assumptions: A1-A5**

A1. $\{m_{i,t}(\phi), G_{i,j,t}(\phi)\}$ **absolutely continuous distributions**

A2. $\{m_{i,t}(\phi), G_{i,j,t}(\phi)\}$ **Paretian tails** $if\ E\left[m_{i,t}^2(\phi)\right] = \infty, E\left|G_{i,j,t}(\phi)\right| = \infty$

Simplifies trimmed moment representations: Karamata's Theorem

## ❑ **Nonlinear Least Tail-Trimmed Squares**

Recall $m_t(\phi) \coloneqq u_t(\phi) f_t^{\phi}(\phi)$ and $G_t(\phi) \coloneqq \frac{\partial}{\partial \phi} m_t(\phi)$

**Assumptions: A1-A5**

A1. $\{m_{i,t}(\phi), G_{i,j,t}(\phi)\}$ **absolutely continuous distributions**

A2. $\{m_{i,t}(\phi), G_{i,j,t}(\phi)\}$ **Paretian tails** $if\ E\left[m_{i,t}^2(\phi)\right] = \infty, E\left|G_{i,j,t}(\phi)\right| = \infty$

A3. $\{y_t, u_t\}$ are **geometrically $\boldsymbol{\beta}$-mixing**

Ensures CLT for tail-trimmed arrays (Hill 10: NED, strong mixing)

Ensures UCTL for indicators (Doukan et al 95)

## Nonlinear Least Tail-Trimmed Squares

Recall $m_t(\phi) \coloneqq u_t(\phi) f_t^{\phi}(\phi)$ and $G_t(\phi) \coloneqq \frac{\partial}{\partial \phi} m_t(\phi)$

**Assumptions: A1-A5**

A1. $\{m_{i,t}(\phi), G_{i,j,t}(\phi)\}$ **absolutely continuous distributions**

A2. $\{m_{i,t}(\phi), G_{i,j,t}(\phi)\}$ **Paretian tails** $if\ E\left[m_{i,t}^2(\phi)\right] = \infty, E\left|G_{i,j,t}(\phi)\right| = \infty$

A3. $\{y_t, u_t\}$ are **geometrically $\beta$-mixing**

A4. **Sufficient trimming**

If $E|m_t| = \infty$ then $k_{j,n}^{(u)}/n^{2(1-\kappa_{m_i})/(2-\kappa_{m_i})} \to \infty$ : $\kappa_{m_i}$ tail index $m_{i,t}$

Ensures sufficient trimming if $E|m_t| = \infty$ : as $\kappa_{m_i} \searrow 0$ then $k_{j,n}^{(u)} \nearrow n$.

## ❑ Nonlinear Least Tail-Trimmed Squares

Recall $m_t(\phi) \coloneqq u_t(\phi) f_t^{\phi}(\phi)$ and $G_t(\phi) \coloneqq \frac{\partial}{\partial \phi} m_t(\phi)$

**Assumptions: A1-A5**

A1. $\{m_{i,t}(\phi), G_{i,j,t}(\phi)\}$ **absolutely continuous distributions**

A2. $\{m_{i,t}(\phi), G_{i,j,t}(\phi)\}$ **Paretian tails** $if\ E\left[m_{i,t}^2(\phi)\right] = \infty, E\left|G_{i,j,t}(\phi)\right| = \infty$

A3. $\{y_t, u_t\}$ are **geometrically *β*-mixing**

A4. **Sufficient trimming**: $k_{j,n}^{(u)}/L(n) \to \infty$ for slowly varying $L(n) \to \infty$

A5. **Identification sufficiently fast**: $E[\widehat{m}_{n,t}^*] \to 0$

Trivial if $u_t$ is independent of $y_{t-i}$ and symmetric.

Otherwise: regularity condition.

## Nonlinear Least Tail-Trimmed Squares

Recall $m_t(\phi) := u_t(\phi) f_t^{\phi}(\phi)$ and $G_t(\phi) := \frac{\partial}{\partial \phi} m_t(\phi)$

**Scale construction for asymptotic normality**

Tail quantiles thresholds: for any random variable $w_t$

$$P\left(w_t(\phi) < \underline{-l_n^{(w)}}(\phi)\right) = \frac{k_{1,n}^{(w)}}{n} \quad and \quad P\left(w_t(\phi) > \underline{u_n^{(w)}}(\phi)\right) = \frac{k_{2,n}^{(w)}}{n}$$

Note $\sup_{\phi} \left| \frac{w^{(-)}_{\left(k_{1,n}^{(w)}\right)}(\phi)}{l_n^{(w)}(\phi)} - 1 \right| = O_p\left(\left(k_{1,n}^{(w)}\right)^{-1/2}\right)$ etc ...

## ❑ Nonlinear Least Tail-Trimmed Squares

Recall $m_t(\phi) \coloneqq u_t(\phi) f_t^{\phi}(\phi)$ and $G_t(\phi) \coloneqq \frac{\partial}{\partial \phi} m_t(\phi)$

### Scale construction for asymptotic normality

Trimming indicators : $I_{n,t}^{(w)}(\theta) \coloneqq I\left(-l_n^{(w)}(\theta) \leq w_t(\theta) \leq u_n^{(w)}(\theta)\right)$

Trimmed equations: $m_{n,t}^{*}(\phi) = u_t(\phi) I_{n,t}^{(u)}(\phi) \times f_t^{\phi}(\phi) I_{n,t}^{\left(f^{\phi}\right)}(\phi) I_{n,t}^{(f^{\phi,\phi})}(\phi)$

Longrun covariance : $S_n = \frac{1}{n} \sum_{s,t=1}^{n} E[m_{n,s}^{*} m_{n,t}^{*}{}']$ ($m_{n,t}^{*}$ is not *mds*).

Jacobian : $G_n = E\left[\frac{\partial}{\partial \phi} m_t \times I_{n,t}^{(u)} I_{n,t}^{(f^{\phi})} I_{n,t}^{(f^{\phi,\phi})}\right]$

Scale : $V_n = nG_n{}' S_n^{-1} G_n$

## Nonlinear Least Tail-Trimmed Squares

**Theorem:** *Under A1-A5* $V_n^{1/2}(\hat{\phi}_n - \phi^0) \xrightarrow{d} N(0, I_p)$ *where* $V_n = nG_n{'}S_n^{-1}G_n$

---

**Remark**: Rates $V_{n,i,i}^{1/2} \to \infty$ require response *f*, dependence *u*, tail parameters...

**Remark**: Mimics NLLS for thin-tailed data (and *mds* $u_t$):

$$V_n^{1/2}(\hat{\phi}_n - \phi^0) \sim n^{-1/2}(E[u_t^2 G_t])^{-1/2} \sum_{t=1}^{n} u_t G_t$$

where $V_n = n \times E[G_t]'(E[u_t^2 G_t])^{-1} E[G_t]$

NLLS rate of convergence is $n^{1/2}$ if tails thin.

## Least Tail-Trimmed Squares for AR

Linear AR model: $y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + u_t = \phi' x_t + u_t$ with *iid Paretian* error.

---

The error $u_t$ is iid, symmetric, continuous and Paretian:

$$P(|u_t| > u) = du^{-\kappa}(1 + o(1)) \quad where \;\; d, \kappa > 0$$

Trivially $f_t^{\phi} = x_t \;\; and \;\; f_t^{\phi,\phi} = 0$

## ❑ Least Tail-Trimmed Squares for AR

Linear AR model: $y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + u_t = \phi' x_t + u_t$ with *iid Paretian* error.

---

The error $u_t$ is iid, symmetric, continuous and Paretian:

$$P(|u_t| > u) = du^{-\kappa}(1 + o(1)) \;\; where \;\; d, \kappa > 0$$

LTTS criterion

$$\hat{Q}_n(\phi) := \sum_{t=1}^{n} (y_t - \phi' x_t)^2 \times \hat{I}_{n,t}^{(u)}(\phi) \hat{I}_{n,t}^{(x)}$$

where $\hat{I}\left(|u_t| < u_{(k_n^{(u)})}^{(a)}\right)$ and $\hat{I}_{n,t}^{(x)} := \prod_{i=1}^{p} I\left(|y_{t-i}| < y_{(k_n^{(y)})}^{(a)}\right)$

## **Least Tail-Trimmed Squares for AR**

Linear AR model: $y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + u_t = \phi' x_t + u_t$ with *iid Paretian* error.

---

Scale $V_n = nG_n{}' S_n^{-1} G_n$

Independence and neglig. : $G_n = -E\left[x_t x_t{}' I_{n,t}^{(u)} I_{n,t}^{(x)}\right] \sim -\left[E\left[y^*_{n,t-i} y^*_{n,t-j}\right]\right]_{i,j=1}^{p}$

Symmetry and independence: $S_n \sim E[u_{n,t}^{*2}] \times \left[E\left[y^*_{n,t-i} y^*_{n,t-j}\right]\right]_{i,j=1}^{p}$

LTTS scale mimics OLS: $V_n = n \times \left(E[u_{n,t}^{*2}]\right)^{-1} \times \left[E\left[y^*_{n,t-i} y^*_{n,t-j}\right]\right]_{i,j=1}^{p}$

- **Least Tail-Trimmed Squares for AR**

Linear AR model: $y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + u_t = \phi' x_t + u_t$ with *iid Paretian* error.

---

**Theorem:** *Under fractile property A4*

$$n^{1/2} \frac{\left(E\left[y_{n,t}^{*2}\right]\right)^{1/2}}{\left(E\left[u_{n,t}^{*2}\right]\right)^{1/2}} \left(\hat{\phi}_{i,n} - \phi_i^0\right) \xrightarrow{d} N(0, I_p)$$

---

Mixing and tail assumptions hold (An/Huang 96, Brockwell/Cline 85).

## Least Tail-Trimmed Squares for AR

Linear AR model: $y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + u_t = \phi' x_t + u_t$ with *iid Paretian* error.

---

Characterize $n^{1/2} \frac{\left(E\left[y_{n,t}^{*2}\right]\right)^{1/2}}{\left(E\left[u_{n,t}^{*2}\right]\right)^{1/2}} \rightarrow \infty$

by Karamata's Theorem: since $P\left(|y_t| > c_n^{(y)}\right) = k_n^{(y)}/n$

$$E\left[y_{n,t}^{*2}\right] \sim \left(c_n^{(y)}\right)^2 P\left(|y_t| > c_n^{(y)}\right) = K\left(n/k_n^{(y)}\right)^{2/\kappa - 1}$$

$$E\left[u_{n,t}^{*2}\right] \sim \left(c_n^{(u)}\right)^2 P\left(|u_t| > c_n^{(u)}\right) = K\left(n/k_n^{(u)}\right)^{2/\kappa - 1}$$

## ❑ Least Tail-Trimmed Squares for AR

Linear AR model: $y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + u_t = \phi' x_t + u_t$ with *iid Paretian* error.

---

$$V_{i,i,n}^{1/2} = n^{1/2} \frac{\left(E\left[y_{n,t}^{*2}\right]\right)^{1/2}}{\left(E\left[u_{n,t}^{*2}\right]\right)^{1/2}} \sim K n^{1/2} \; if \; \kappa > 2$$

$$V_{i,i,n}^{1/2} \sim K n^{1/2} \frac{\left(\left(n/k_n^{(y)}\right)^{2/\kappa-1}\right)^{1/2}}{\left(\left(n/k_n^{(u)}\right)^{2/\kappa-1}\right)^{1/2}} = K n^{1/2} \left(\frac{k_n^{(u)}}{k_n^{(y)}}\right)^{1/\kappa-1/2} \quad if \; \kappa < 2$$

Errors $u_t$ are adverse: *harshly trim*.

Regressors $y_{t-i}$ provide leverage: *lightly trim*.

## **Least Tail-Trimmed Squares for AR**

Linear AR model: $y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + u_t = \phi' x_t + u_t$ with *iid Paretian* error.

**If $\kappa$ < 2** : *maximize* $k_n^{(u)} \sim \lambda_u \frac{n}{\ln(n)} < \lambda_u n$ and *mininize* $k_n^{(y)} \sim \lambda_y \ln(n) > 0$

$$V_{i,i,n}^{1/2} \sim K n^{1/2} \left( \frac{\lambda_u n / \ln(n)}{\lambda_y \ln(n)} \right)^{1/\kappa - 1/2} = K \left( \frac{\lambda_u}{\lambda_y} \right)^{1/\kappa - 1/2} n^{1/\kappa} / L(n)$$

$$\frac{n^{1/\kappa}}{L(n)} \left( \hat{\phi}_{i,n} - \phi_i^0 \right) \xrightarrow{d} N\left( 0, \left( \frac{\lambda_y}{\lambda_u} \right)^{2/\kappa - 1} v^2 \right)$$ : effic.$\uparrow$ as $\lambda_y \searrow 0$ and $\lambda_u \uparrow$

Nearly highest possible rate $n^{1/\kappa}$ (Davis, Knight, Liu 92...)

## Quasi-Maximum Tail-Trimmed Likelihood

Stationary GARCH(1,1)

$y_t = \sigma_t \epsilon_t$ where $u_t = \sigma_t \epsilon_t$

$\sigma_t^2 = \omega^0 + \alpha^0 y_{t-1}^2 + \beta^0 \sigma_{t-1}^2$ $\omega^0 > 0,\ \alpha^0, \beta^0 \geq 0$ and $\theta = [\omega, \alpha, \beta]'$

$\epsilon_t$ is *iid* (0,1), with continuous distribution (**can allow *mds* error**)

$\epsilon_t$ has Paretian distribution if $E[\epsilon_t^4] = \infty$.

Volatility process: $h_0(\theta) = \omega^0$ *and* $h_t(\theta) = \omega^0 + \alpha^0 y_{t-1}^2 + \beta^0 h_{t-1}(\theta)$

## ❑ Quasi-Maximum Tail-Trimmed Likelihood

QML first order condition: $\frac{1}{n}\sum_{t=1}^{n}\left(\epsilon_t^2(\hat{\theta}_n) - 1\right)\frac{\partial}{\partial\theta}\ln\left(h_t(\hat{\theta}_n)\right) = 0$

If there are GARCH affects ($\alpha^0 + \beta^0 > 0$) then (Francq/Zakoian 06, 10)

$$\sup_{\theta\in\Theta} E\left|\frac{\partial}{\partial\theta_i}\ln(h_t(\theta))\right|^{2+\iota}, \sup_{\theta\in\Theta} E\left|\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln(h_t(\theta))\right|^{2+\iota} < \infty \; for \;\; tiny \;\; \iota > 0$$

Only error $\epsilon_t$ causes damaging extremes (*obviously*).

QML : need $E[\epsilon_t^4] < \infty$ for strong-GARCH (Francq/Zakoian 06)

## ❑ **Quasi-Maximum Tail-Trimmed Likelihood**

If GARCH affects then *heavy tails only due to error* $\epsilon_t$.

$$\hat{Q}_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}\left(\ln h_t(\theta) + \epsilon_t^2(\theta)\right) \times \hat{I}_{n,t}^{(\epsilon)}(\theta)$$

If no effects $\frac{\partial}{\partial\theta_i}\ln(h_t)$ and $\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln(h_t) \in \{K, y_{t-1}^2, y_{t-2}^2\}$

$$\hat{Q}_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}\left(\ln h_t(\theta) + \epsilon_t^2(\theta)\right) \times \hat{I}_{n,t}^{(\epsilon)}(\theta)\hat{I}_{n,t-1}^{(y)}\,\hat{I}_{n,t-2}^{(y)}$$

QMTTL estimator $\hat{\theta}_n = \text{arginf}_{\theta\in\Theta}\,\hat{Q}_n(\theta)$

## ❑ **Quasi-Maximum Tail-Trimmed Likelihood**

---

**Theorem:** *Under fractile A4 and identif. A5* $V_n^{1/2}\left(\hat{\theta}_n - \theta^0\right) \overset{d}{\rightarrow} N(0, I_3)$

If GARCH affects are known: $V_n = n\left(E\left[\epsilon_{n,t}^{*4} - 1\right]\right)^{-1} E[s_t s_t{}']$

Else $V_n = n\left(E\left[\epsilon_{n,t}^{*4} - 1\right]\right)^{-1} E\left[s_{n,t}^{*} s_{n,t}^{*}{}'\right]$ where $s_{i,n,t}^{*} := s_{i,t} I_{n,t-1}^{(y)} I_{n,t-2}^{(y)}$

where $s_t := \frac{\partial}{\partial \theta_i} \ln(h_t^*)$ and $h_t^*$ is the stationary solution.

---

Only need stationarity, $\epsilon_t \sim iid$, $E[\epsilon_t^2] = 1$, and fractile bound.

## **Quasi-Maximum Tail-Trimmed Likelihood**

If GARCH affects and $E[\epsilon_t^4] = \infty \;\; such\ that \;\; \kappa \in (2,4)$,

$$V_{i,i,n} \sim Kn \frac{1}{\left(n/k_n^{(\epsilon)}\right)^{4/\kappa - 1}} = n\left(k_n^{(\epsilon)}/n\right)^{4/\kappa - 1} = o(n)$$

No leverage effect: pure error effect, so maximal trimming (near $\lambda n$):

$$k_n^{(\epsilon)} \sim \frac{n}{L(n)} < \lambda n \;\text{ for any }\; \lambda \in (0,1) \;\; then \;\; V_{i,i,n}^{1/2} \sim \frac{n^{1/2}}{L(n)}$$

QML has rate (Hall and Yao 2003) $n^{1-2/\kappa}/L(n) < n^{1/2}/L(n), \;\; \kappa \in (2,4)$

**QMTTL rate is higher.**

## ❑ Simulation Study - LTTS for AR

10,000 samples, $n$ = 100 or 800

Model $y_t = .2 + .8y_{t-1} - .3y_{t-2} + u_t$

where $u_t$ is one of the following:

iid Pareto $:= P_\kappa$ with index $\kappa \in \{.75,\ 1.5,\ 2.5\}$

or

GARCH(1,1) $u_t = \sigma_t \epsilon_t,\ \sigma_t^2 = .3 + .6u_{t-1}^2 + .3\sigma_{t-1}^2$

$\epsilon_t$ is iid $P_{2.5}$ hence $\kappa_y = 1.5$

## **Simulation Study - LTTS for AR**

Fractiles: $k_n^{(u)} = [.05n/\ln(n)]$ $and$ $k_n^{(y)} = [\ln(n)]$

Least Weighted Absolute Deviations (Ling 05)

- weight: Huber (1977), down weights $u_t$ above $5^{th}$-percentile

- asymptotically normal, rate $n^{1/2}$

OLS - asymptotically normal, rate $n^{1/\kappa}$

(**LAD** - essentially the same as OLS)

## ❑ Simulation Study - LTTS for AR

Model $y_t = .2 + .8y_{t-1} - .3y_{t-2} + u_t$

**AR, iid $P_{.75}$ , $n$ = 100, $\phi_3 = -.3$**

| | MEAN | MSE | KS | Trim % |
|---|---|---|---|---|
| **LTTS** | -.302 | .014 | .097 | 8% |
| **LWAD** | -.301 | .032 | .112 | 5% |
| **OLS** | -.299 | .004 | .198 | - |

## Simulation Study - LTTS for AR

Model $y_t = .2 + .8y_{t-1} - .3y_{t-2} + u_t$

**AR, GARCH $P_{2.5}$, $n$ = 100, $\phi_3 = -.3$**

| | MEAN | MSE | KS | Trim % |
|---|---|---|---|---|
| **LTTS** | -.296 | .020 | .036 | 8% |
| **LWAD** | -.304 | .034 | .055 | 5% |
| **OLS** | -.310 | .029 | .084 | - |

## ❑ **Simulation Study - LTTS for AR**

Model $y_t = .2 + .8y_{t-1} - .3y_{t-2} + u_t$

**AR, iid $P_{.75}$ , $n$ = 800, $\phi_3 = -.3$**

| | MEAN | MSE | KS | Trim % |
|---|---|---|---|---|
| **LTTS** | -.298 | .0008 | .038 | 2% |
| **LWAD** | -.300 | .0006 | .043 | 5% |
| **OLS** | -.300 | .0006 | .258 | - |

## Simulation Study - LTTS for AR

Model $y_t = .2 + .8y_{t-1} - .3y_{t-2} + u_t$

**AR, GARCH $P_{2.5}$, $n$ = 800, $\phi_3 = -.3$**

| | MEAN | MSE | KS | Trim % |
|---|---|---|---|---|
| **LTTS** | -.297 | .002 | .036 | 5% |
| **LWAD** | -.301 | .029 | .048 | 5% |
| **OLS** | -.295 | .016 | .074 | - |

## ❑ Simulation Study - QMTTL for GARCH

Model $y_t = \sigma_t \epsilon_t$ and $\sigma_t^2 = .3 + .3y_{t-1}^2 + .6y_{t-1}^2$

The error $\epsilon_t$ is iid $P_{2.5}$ or $N(0,1)$

Tail index $\kappa_y \in \{1.5, 4.1\}$

Fractile: $k_n^{(\epsilon)} = [.05n/\ln(n)]$

---

Quasi-Maximum Weighted Likelihood (Ling 07) - $E[\epsilon_t^4] < \infty$ (*same as QML*)

Log-LAD (Peng and Yao 03, Linton et al 10)

## ❑ Simulation Study - QMTTL for GARCH

Model $y_t = \sigma_t \epsilon_t$ and $\sigma_t^2 = .3 + .3y_{t-1}^2 + .6y_{t-1}^2$

The error $\epsilon_t$ is iid $P_{2.5}$ or $N(0,1)$

**GARCH, iid $P_{2.5}$, $n$ = 100, $\theta_3 =.6$**

| | MEAN | MSE | KS | Trim % |
|---|---|---|---|---|
| **QMTTL** | .611 | .146 | .048 | 1% |
| **QMWL** | .685 | .147 | .236 | 5% |
| **QML** | .691 | .141 | .252 | - |
| **Log-LAD** | .526 | .176 | .170 | - |

## ❑ Simulation Study - QMTTL for GARCH

Model $y_t = \sigma_t \epsilon_t$ and $\sigma_t^2 = .3 + .3y_{t-1}^2 + .6y_{t-1}^2$

The error $\epsilon_t$ is iid $P_{2.5}$ or $N(0,1)$

**GARCH, iid $P_{2.5}$, $n$ = 800, $\theta_3 = .6$**

| | MEAN | MSE | KS | Trim % |
|---|---|---|---|---|
| **QMTTL** | .603 | .036 | .036 | 1% |
| **QMWL** | .674 | .082 | .284 | 5% |
| **QML** | .664 | .112 | .251 | - |
| **Log-LAD** | .603 | .070 | .076 | - |

## ❑ Simulation Study - QMTTL for GARCH

Model $y_t = \sigma_t \epsilon_t$ and $\sigma_t^2 = .3 + .3y_{t-1}^2 + .6y_{t-1}^2$

The error $\epsilon_t$ is iid $P_{2.5}$ or $N(0,1)$

**GARCH, iid $\boldsymbol{N(0,1)}$, $\boldsymbol{n}$ = 100, $\boldsymbol{\theta_3 = .6}$**

| | **MEAN** | **MSE** | **KS** | **Trim %** |
|---|---|---|---|---|
| **QMTTL** | .437 | .180 | .217 | 1% |
| **QMWL** | .509 | .196 | .191 | 5% |
| **QML** | .502 | .181 | .186 | - |
| **Log-LAD** | .416 | .228 | .288 | - |

## ❑ Simulation Study - QMTTL for GARCH

Model $y_t = \sigma_t \epsilon_t$ and $\sigma_t^2 = .3 + .3y_{t-1}^2 + .6y_{t-1}^2$

The error $\epsilon_t$ is iid $P_{2.5}$ or $N(0,1)$

**GARCH, iid $N(0,1)$, $n$ = 800, $\theta_3 = .6$**

| | MEAN | MSE | KS | Trim % |
|---|---|---|---|---|
| **QMTTL** | .563 | .065 | .144 | 1% |
| **QMWL** | .571 | .084 | .122 | 5% |
| **QML** | .576 | .072 | .112 | - |
| **Log-LAD** | .537 | .127 | .192 | - |