# 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 (<u>for any reason</u>).

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

## ❑ PROPOSAL

**Stylized traits of financial time series:**

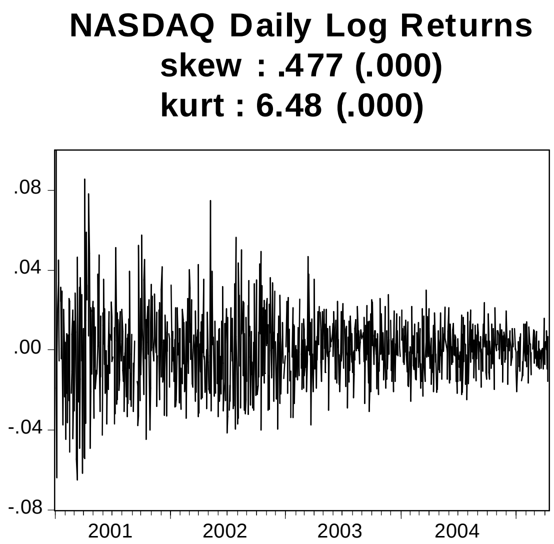
*Nonlinear* – asymmetries in returns.

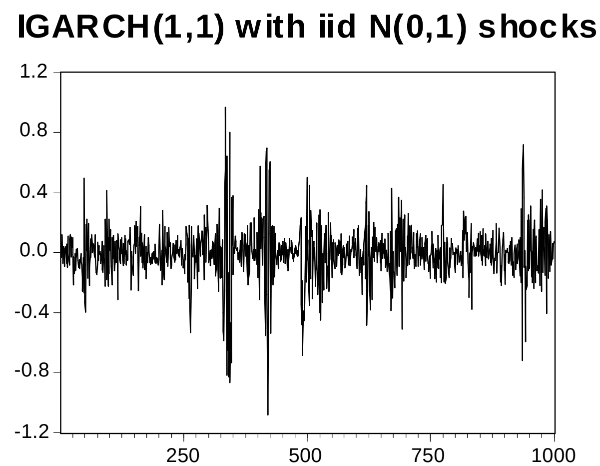
*Heavy-tailed* : unobserved shocks; and/or due to parametric structure.

($\varepsilon_t$) (IGARCH)

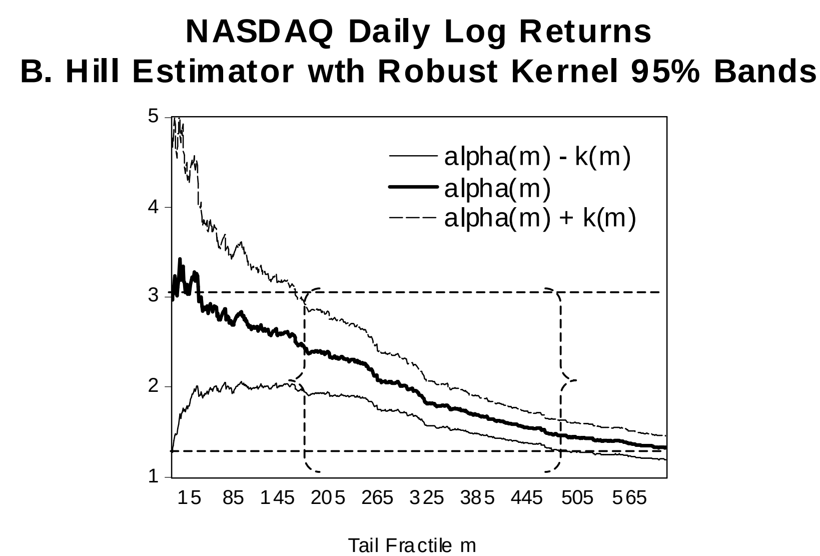
# ❑ PROPOSAL

*Clusters*: not as severe as IGARCH…




…but at least *infinite fourth moment.*



# ❑ PROPOSAL

## MODEL – Nonlinear ARX-Nonlinear GARCH

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

where $f$ and $h_t$ are *differentiable* in $\theta$: excludes TAR's, etc.

# ❑ PROPOSAL

## MODEL – Nonlinear ARX-Nonlinear GARCH

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

where $f$ and $h_t$ are *differentiable* in $\theta$: excludes TAR's, etc.

$x_t \in \Re^k$ contains $y_t, y_{t\text{-}1}, \ldots,$ and other variables and their lags.

$\Im_{\mathrm{t}} := \sigma(x_\tau : \tau \leq t)$.

$E[y_t \mid \Im_{t-1}] = f(x_{t-1}, \alpha)$ a.s. *if and only if* $\alpha = \alpha_0$.

$E[u_t^2 \mid \Im_{t-1}] = h_t(\theta)$ a.s. *if and only if* $\theta = \theta_0$.

## ❑ PROPOSAL

**MODEL – Nonlinear ARX-Nonlinear GARCH**

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

**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 $\{y_t, \varepsilon_t\}$.

Autoregressions : finite *variance* $\{\varepsilon_t\}$.

GARCH : at least finite *kurtosis* $\{y_t, \varepsilon_t\}$.

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 $\boldsymbol{E(\varepsilon_t^4) < \infty}$

(Francq and Zakoïan 2004).

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

(Hansen and Lee 1994, Lumsdaine 1996)

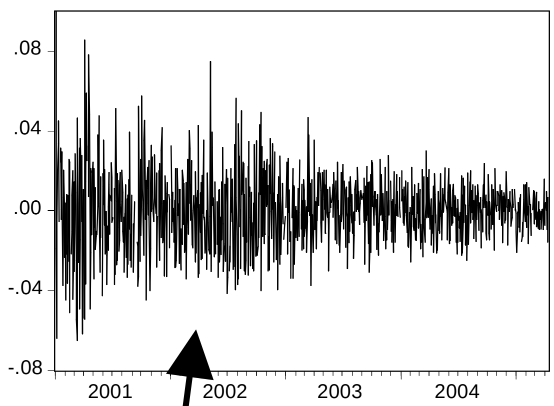
Non-stationary strong-ARCH (Jensen and Rahbek 2004)

**Shortcomings** : Only *linear* GARCH. Error *moments restrictive.*

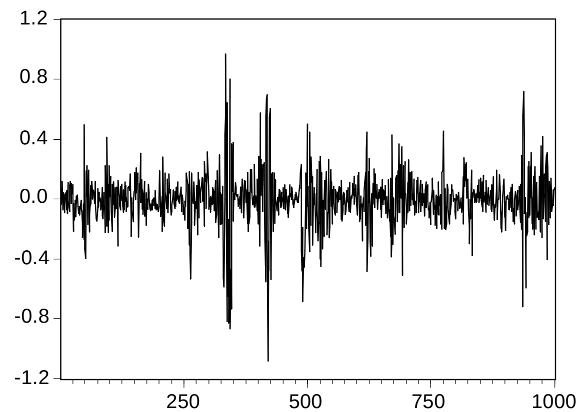
# ❑ WHAT WE KNOW…

*Asymmetric…*

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



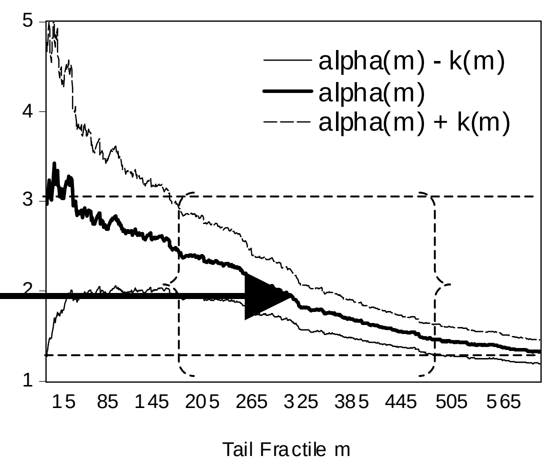

**IGARCH(1,1) with iid N(0,1) shocks**




…and *heavy-tailed.*

Doesn't look like IGARCH, but *could have infinite variance.*

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



## ❑ WHAT WE KNOW…

### LEAST ABSOLUTE DEVIATIONS

Asymptotic normality:

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

(Linton, O., J. Pan and H. Wang 2008; cf. Peng and Yao 2003)

**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**

<u>Trimming</u> : remove large values from estimation…

<u>Weighting</u> : diminish contribution of large values to criterion…

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

Čižek (2008) : Generalized Trimmed Estimators - **conditional mean**.

Rubert and Carroll (1980), Rousseeuw (1985), Welsh (1987),
Bassett (1991), Chen, Welsh and Chan (2001), Chan and Peng (2005),
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)

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

$$E \mid y_t \mid^p < \infty \text{ for some } p > 0$$

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

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

## ❑ 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

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

$$\Im_{\mathrm{t}} := \sigma(x_\tau : \tau \le t)$$

Estimating equations $m_t(\theta)$ map $m_t : \Re^k \rightarrow \Re^s, \; s \ge k$

Indentification : $E[m_t(\theta)] = 0$ *if and only if* $\theta = \theta_0$

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

where $\hat{\Omega}$ is positive semi-definite.

## ❑ TAIL-TRIMMED MDE – Set Up

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

Asymptotic problem :

$$\hat{\theta} = \arg\min\left\{Q_n(\theta)\right\}$$

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

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:

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

where $g_{t-1}^{(i)}$ are $\Im_{t\text{-}1}$ - measurable.

## ❑ TAIL-TRIMMED MDE – Estimation Equations

### GENERALIZED METHOD OF MOMENTS

**Example :** GMM estimating equations for pure ARCH(1)

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

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

$$\phi_t(\theta) = \frac{1}{h_t(\theta)}\exp\left\{-\frac{1}{2}\left(\frac{\left(y_t - f(x_{t-1},\alpha)\right)^2}{h_t^2(\theta)}\right)\right\}$$

## ❑ TAIL-TRIMMED MDE – Estimation Equations

We require $\{\varepsilon_t, \Im_t\}$ to form a *martingale difference sequence*.

We only consider pure Nonlinear GARCH for LAD and LLAD:

**LEAST ABSOLUTE DEVIATIONS**

$$m_t(\theta) = \varepsilon_t \times \mathrm{sgn}\left\{\varepsilon_t^2 - 1\right\} \times \frac{\partial}{\partial \theta} \ln h_t^2(\theta)$$

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

$$m_t(\theta) = \mathrm{sgn}\left\{\ln \varepsilon_t^2\right\} \times \frac{\partial}{\partial \theta} \ln h_t^2(\theta)$$

## ❑ TAIL-TRIMMED MDE – Tail Trimmed Equations

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

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

**# Trimmed *m*'s** : $k_{1,n} = left, \; k_{2,n} = right \; : \; k_{i,n} \to \infty, \; \dfrac{k_{i,n}}{n} \to 0.$

**Thresholds** : $\dfrac{n}{k_{1,n}} P\left(m_{i,t}^{(-)} < -l_{i,n}\right) \to 1$ : $l_{i,n}$ is lower $k_{1,n} / n^{th} \to 0$ - quantile.

$\dfrac{n}{k_{2,n}} P\left(m_{i,t}^{(+)} > u_{i,n}\right) \to 1$ : $u_{i,n}$ is lower $k_{2,n} / n^{th} \to 0$ - quantile.

## ❑ TAIL-TRIMMED MDE – Tail Trimmed Equations

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

**Thresholds** : $\dfrac{n}{k_{1,n}} P\left(m_{i,t}^{(-)} < -l_{i,n}\right) \to 1$ : $l_{i,n}$ is lower $k_{1,n}/n^{th} \to 0$ quantile.

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

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

Need to estimate $l_{i,n}$ and $c_{i,n}$ for practical applications.

Jonathan B. Hill, Dept. of Economics, University of North Carolina – Chapel Hill, jbhill@email.unc.edu 

## TAIL-TRIMMED MDE – Tail Trimmed Equations

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

**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)

## ❑ TAIL-TRIMMED MDE – Tail Trimmed MDE

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

**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),*

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

---

$v_n^2 := \| V_n \|$ 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),*

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

---

If $\| m_t(\theta_0) \|_2 < \infty$ then $v_n \sim K n^{1/2}$.

## ❑ ASYMPTOTIC THEORY FOR TTMDE

## MAIN RESULT

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

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

---

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

where $\Sigma_{\text{n}} = n \times E[m_{n,t}(\theta_0) m_{n,t}(\theta_0)']$ and $G_n = \frac{\partial}{\partial \theta} E[m_{n,t}(\theta)]|_{\theta=\theta_0}$ .

## ASYMPTOTIC THEORY FOR TTMDE

## MAIN RESULT

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

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

---

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 $G_n = \frac{\partial}{\partial \theta} E[m_{n,t}(\theta)]|_{\theta=\theta_0}$

- **ASSUMPTIONS –** A's ensure $Q_n$ is sufficiently smooth.

**A1.** $\| \hat{\Omega}_n - \Omega \| \xrightarrow{p} 0$, 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.** $\dfrac{\partial}{\partial \theta} E[m_{n,t}(\theta)]$ exists $\forall \theta \in U_{\theta_0}$, an arbitrary neighborhood of $\theta_0$.

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

Huber (1967), Pakes and Pollard (1989), Newey and McFadden (1994)

- **ASSUMPTIONS** – B's restrict dgp's { $m_t$, $m_{n,t}$, $x_t$ }

**B1.** $\{m_{n,t}(\theta), \mathfrak{I}_t\}$ is an adapted martingale difference array *iff* $\theta = \theta_0$.

**B2.** $\lim\limits_{z\to\infty} \dfrac{P(|m_{i,t}(\theta_0)| > z)}{z^{-\kappa_i} L(z)} \leq 1$ where $L$ is slowly varying, $\kappa_i > 0$.

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

*ii.* $\liminf\limits_{n\geq 1} \dfrac{n\left\|E[m_{n,t}(\theta_0) m_{n,t}(\theta_0)']\right\|}{\max\limits_{1\leq i\leq s}\left\{\max\{l_{i,n}, u_{i,n}\}\right\}^2} > 0$, $\max\limits_{1\leq i\leq s}\left\{\max\{l_{i,n}, u_{i,n}\}\right\}^2 = o(n)$.

**B4.** Regressorss $x_t$ are stationary and strong mixing with size 1.

- **EXAMPLES : each satisfy *martingale difference* B1 and *tail bound* B2**

1. Random Coefficient Autoregression

$$y_t = \alpha_t y_{t-1} + \varepsilon_t, \quad |\alpha_t| \leq \alpha \in (0,1), \quad \alpha_t \text{ is } \Im_{t-1} \text{- measurable.}$$

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

$$mds \text{ and } \lim_{\varepsilon\to\infty} P(\varepsilon_t > \varepsilon \mid \Im_{t-1})/\varepsilon^{\tilde{\kappa}} \leq K$$

## ❑ EXAMPLES : each satisfy *martingale difference* B1 and *tail bound* B2

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

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

AR - roots outside unit circle.

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

$E|\varepsilon_t|^{\iota} < 1$ for some infinitessimal $\iota > 0$ and $\beta_i \left\|\varepsilon_t\right\|_{\iota}^{1/\kappa} < 1,\ i = 1, 2.$

Covers nearly
all GARCH since
$\iota > 0$ is arbitrary.

- **EXAMPLES : each satisfy *martingale difference* B1 and *tail bound* B2**

3. Threshold GARCH

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

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

$E|\varepsilon_t|^{\iota} < 1$ for some infinitessimal $\iota > 0$ and $\beta_i \left\|\varepsilon_t\right\|_{\iota}^{1/\kappa} < 1, \; i = 1, 2.$

Covers nearly all GARCH since $\iota > 0$ 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.

Amazingly : **no simulation** study in Ling (2007) and Čižek (2008).

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

Hill and Renault (2008) : TTGMM - simulations; large sample distribution.

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

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

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

$$\hat{\theta}_n(\delta) = \arg\min_{\theta \in \Theta} \left\{\hat{Q}_n^{(\delta)}(\theta)\right\}$$

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

**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 : $\hat{\theta}_n(\hat{\delta}_n) \xrightarrow{p} \theta_0$.

---

Notice : select $\delta$ such that *untrimmed equation mean* is close to zero.

## ❑ SIMULATION STUDY

| MODEL | $\kappa$ |
|---|---|
| AR : $y_t = .9y_{t-1} + \varepsilon_t, \;\; \varepsilon_t \overset{iid}{\sim} P_{1.5}$ | 1.50 |
| GARCH : $y_t = h_t\varepsilon_t, \; h_t^2 = .3 + .3y_{t-1}^2 + .6h_{t-1}^2, \;\; \varepsilon_t \overset{iid}{\sim} P_{2.5}$ | 2.14 |
| IGARCH : $y_t = h_t\varepsilon_t, \; h_t^2 = .3 + .4y_{t-1}^2 + .6h_{t-1}^2, \; \varepsilon_t \overset{iid}{\sim} N_{0,1}$ | 2.00 |
| QARCH : $y_t = h_t\varepsilon_t, \; h_t^2 = (.3 + .8y_{t-1})^2, \;\; \varepsilon_t \overset{iid}{\sim} P_{2.5}$ | 2.88 |
| QARCH : $y_t = h_t\varepsilon_t, \; h_t^2 = (.3 + .8y_{t-1})^2, \;\; \varepsilon_t \overset{iid}{\sim} N_{0,1}$ | 3.53 |

## ❑ SIMULATION STUDY

- o Simulate 100 series of sample size ***n* = 1000**.
- o QML
- o GMM and TTGMM – weight matrix $\Omega = I$.
- o Inspect small sample distribution of $k^{th}$-parameter estimate :

Kolmogorov-Smirnov tests of normality over 100 estimates.

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

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

$$\text{min. KS over } \delta_i \in \{.01, ..., .99\}.$$

## ❑ SIMULATION STUDY

**QML, GMM, TTGMM**

| | **MEAN** | **STDEV** | **Z** | **KS** | **MEAN** | **STDEV** | **Z** | **KS** |
|---|---|---|---|---|---|---|---|---|
| | IID, *mean* = **1**, $P_{1.5}$ | | | | AR(1), *slope* = **.9**, $P_{1.5}$ | | | |
| QML | 1.01 | .151 | .07 | .272 | .889 | .013 | .04 | .253 |
| GMM | .939 | .212 | .08 | .172 | .890 | .012 | .07 | .187 |
| TTGMM | .992 | .139 | .06 | .052* | .901 | .007 | .04 | .075* |
| | GARCH, *slope* = **.6**, $P_{2.5}$ | | | | IGARCH, *slope* = **.6**, $N_{0,1}$ | | | |
| QML | .264 | .184 | .00 | .569 | .586 | .198 | .14 | .262 |
| GMM | .605 | .346 | .00 | .448 | .628 | .225 | .02 | .273 |
| TTGMM | .593 | .231 | .06 | .096* | .602 | .177 | .06 | .083* |
| | QARCH, *slope* = **.8**, $N_{0,1}$ | | | | QARCH, *slope* = **.8**, $P_{2.5}$ | | | |
| QML | .896 | .659 | .00 | .389 | .884 | .751 | .01 | .412 |
| GMM | .673 | .093 | .05 | .454 | .689 | .243 | .04 | .486 |
| TTGMM | .807 | .094 | .06 | .068* | .799 | .185 | .06 | .094* |

**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%.