Tail and Non-Tail Memory with Applications to Random Volatility

Prepared for the 2nd International Workout on *Computational and Financial Econometrics* 2008

Jonathan B. Hill

Dept. of Economics

University of North Carolina – Chapel Hill

OBJECTIVES

❑ APPLICATIONS : ROBUST INFERENCE

Develop *tail dependence* properties for …

Tail shape inference (linear and nonlinear GARCH, stochastic vol.)

Non-parametric tail dependence inference

Develop *tail-trimmed* properties for…

Robust GMM, NLLS, QMLE for heavy-tailed random volatility models

(*robust = asymptotically Gaussian for persistent heavy-tailed data*)

OBJECTIVES

❑ THEORY

New notions of ***tail*** *dependence*, *persistence*, *decay*.

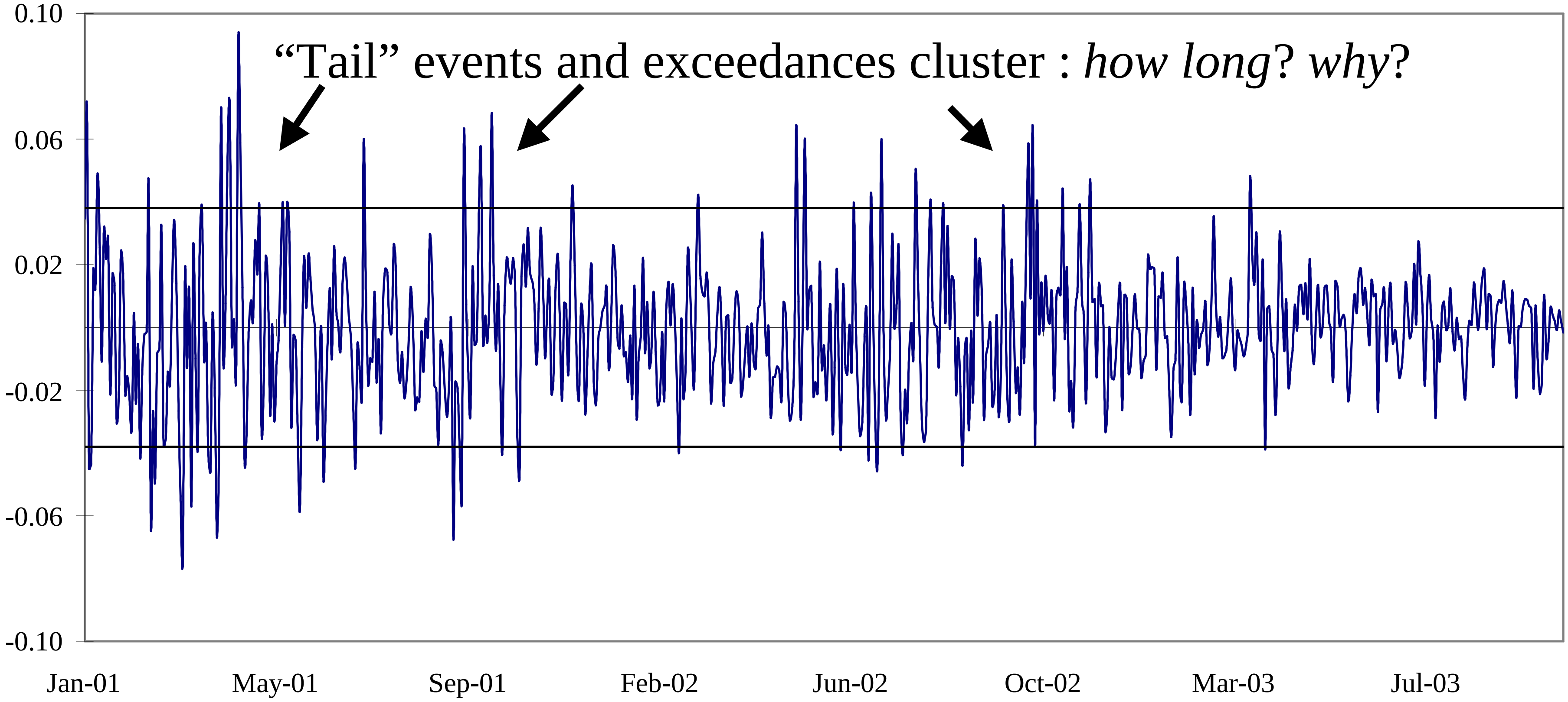
Tail Dependence $\{x_{t-h}, y_t\} : |x_{t-h}| > b_x \text{ } and \text{ } |y_t| > b_y$ as $\{b_x, b_y\} \to \infty$

Tail Events : $I(|x_t| > b_x)$ and $I(|y_t| > b_y)$ as $b_x, b_y \to \infty$

Tail Exceedances : $\max\{\ln|x_t| - \ln b_x, 0\}$ and $\max\{\ln|y_t| - \ln b_y, 0\}$

OBJECTIVES

NASDAQ Daily Log Returns and Upper 5th percentile : $n = 1044$



OBJECTIVES

❑ THEORY

New notions of ***tail*** *dependence*, *persistence*, *decay*.

Tail Dependence $\{x_{t-h}, y_t\} : |x_{t-h}| > b_x$ *and* $|y_t| > b_y$ as $\{b_x, b_y\} \to \infty$

New notions of ***tail-trimmed*** *dependence*, *persistence*, *decay*.

Non-Tail $\{x_{t-h}, y_t\} : |x_{t-h}| \leq b_x$ *and* $|y_t| \leq b_y$ as $\{b_x, b_y\} \to \infty$

LITERATURE

❑ TAIL DEPENDENCE

Joint Distribution Tail Exponents for Bivariate $\{X_t, Y_t\}$

$$P(W_{x,t-h} > \varepsilon, W_{y,t} > \varepsilon) = \varepsilon^{-1/\gamma_h} L_h(\varepsilon) \quad \text{where } \gamma \in [0,1]$$

$\{W_{x,t}, W_{y,t}\}$ are Unit Frechet transforms; $L_h(\varepsilon)$ is slowly varying

$\gamma_h < 1/2$ (neg. dep.), $\gamma_h = 1/2$ (indep.), $\gamma_h > 1/2$ (pos. dep.)

Ledford and Tawn (1996, 1997, 2003); Heffernan and Tawn (2004)

- Must know marginal distributions; **$\{X_t\}$ and $\{Y_t\}$ are iid**.
- ad hoc models of L_h (Weibull, Logistic,…).

LITERATURE

❑ TAIL DEPENDENCE

Joint Distribution Tail Exponents for Bivariate $\{X_t, Y_t\}$

$$P(W_{x,t-h} > \varepsilon, W_{y,t} > \varepsilon) = \varepsilon^{-1/\gamma_h} L_h(\varepsilon) \quad \text{where } \gamma \in [0,1]$$

$\{W_{x,t}, W_{y,t}\}$ are Unit Frechet transforms; $L_h(\varepsilon)$ is slowly varying

$\gamma_h < 1/2$ (neg. dep.), $\gamma_h = 1/2$ (indep.), $\gamma_h > 1/2$ (pos. dep.)

- o Assumes exponent γ_h denotes core expression of dependence.
- o Ignores bivariate dependence decay as $h \rightarrow \infty$.
- o Cannot capture degenerate cases (stochastic volatility).

LITERATURE

❑ TAIL DEPENDENCE

Tail Copula

$$\Lambda(h) = \lim_{b_x, b_y \to \infty} \frac{P(|X_{t-h}| > b_x, |Y_t| > b_y)}{\sqrt{P(|X_{t-h}| > b_x) \times P(|Y_t| > b_y)}} \quad \text{(two-tailed)}$$

Schmidt and Stadtmüller (2006); Klüppleberg, Kuhn, Peng (2007)

Ignores memory decay; **$\{X_t\}$ and $\{Y_t\}$ are iid**, and $h = 0$.

Cannot capture degenerate cases (stochastic volatility).

LITERATURE

- **TRIMMING** : $x_t \times I(|x_t| \leq b_x)$ **or** **TRUNCATION** : $x_t \times I(|x_t| \leq b_x) + b_x \times I(|x_t| > b_x)$

  Inference for mean, or least squares for *heavy tailed* $\{x_t\}$

  1. IID
  2. Weakly Dep. + finite variance (e.g. covariance stationary GARCH)
  3. Linear distributed lags and fixed quantile trimming
  4. Mean/location; Least Squares (*not* GMM; not *heavy-tail indifferent*)

Bickel (1965), Stigler (1973b), Csörgő, Horváhth and Mason (1986)
Griffin and Pruitt (1987), Hahn, Kuelbs, Weiner (1990)

Hahn and Weiner (1992)

Wu (2005), Agulló, Croux and Van Aelst (2008)

TAIL MEMORY

$\{x_t, y_t\}$ is a bivariate time series, x_t , $y_t > 0$, sample size = n.

Tail-specific data:

Two Tailed: $|x_t|$

Left Tailed: $-x_t \times I(x_t < 0)$

Right Tailed: $x_t \times I(x_t \geq 0)$

TAIL MEMORY

Extreme value within sample $\{x_t\}_{t=1}^{n}$: $x_t > b_{x,n} \rightarrow \infty$.

$\boldsymbol{k_n}$ = number of extremes : $k_n \rightarrow \infty$ and $k_n/n \rightarrow 0$ as $n \rightarrow \infty$.

$\boldsymbol{b_{x,n}} \rightarrow \infty$ is the $\boldsymbol{k_n/n^{th}}$ **upper quantile** of x_t :

$$P\left(x_t > b_{x,n}\right) \sim \frac{k_n}{n} \rightarrow 0$$

Leadbetter, Rootzen, Lindgren (1983), Resnick (1987),

Embrechts, Klüppleberg, Mikosch (2003)

TAIL MEMORY : Estimable

- **Coefficient of Tail Dependence:**

Joint / marginal tail probability discrepancy :

$$r_n(h) = \frac{n}{k_n}\left[P(x_{t-h} > b_{x,n}, y_t > b_{y,n}) - P(x_{t-h} > b_{x,n}) \times P(y_t > b_{y,n})\right]$$

$$= \frac{n}{k_n}\left[P_{x,y,h}(b_{x,n}, b_{y,n}) - P_x(b_{x,n}) \times P_y(b_{y,n})\right]$$

$$\lim_{n\to\infty} r_n(h) \in [0,1]$$

TAIL MEMORY : Estimable

- **Coefficient of Tail Dependence:** $r_n(h) = \frac{n}{k_n}[P_{x,y,h}(b_{x,n}, b_{y,n}) - P_x(b_{x,n})P_y(b_{y,n})]$

$$Local\ \ tail\ \ dep.:\ \frac{n}{k_n} r_n(h) \sim \frac{P(x_{t-h} > b_{x,n}, y_t > b_{y,n})}{P(x_{t-h} > b_{x,n}) \times P(y_t > b_{y,n})} - 1 \nrightarrow 0$$

$$Distant\ \ tail\ \ dep.:\ r_n(h) \sim \Lambda(h) = \frac{P(x_{t-h} > b_{x,n}, y_t > b_{y,n})}{\sqrt{P(x_{t-h} > b_{x,n}) \times P(y_t > b_{y,n})}} \nrightarrow 0$$

Tail index & copula: *<u>distant</u> tail dependence.*

Stochastic Volatility : *<u>local</u> tail dependent.*

TAIL MEMORY : Estimable

❑ Coefficient of Tail Dependence:

Easy to estimate :

$$\hat{r}_n(h) = \frac{1}{k_n}\sum_{t=1}^{n} I(x_{t-h} > x_{(k_n+1)}, y_t > y_{(k_n+1)}) - \frac{k_n}{n}$$

Order statistics $x_{(1)} \geq x_{(2)} \geq \ldots \geq x_{(n)}$

Asymptotic theory requires more *abstract* notions of *tail dependence*.

Do not want to restrict joint tail, non-extremes, extremal dependence.

TAIL MEMORY : Theoretical

❑ Tail Dependence Definitions

Extreme events $x_t > b_{x,n}$ can be predicted using information from "*near epoch*" of $\{\varepsilon_{t-l_n},...,\varepsilon_t\}$, where $l_n \rightarrow \infty$ as $n \rightarrow \infty$.

Example : x_t = log returns to NASDAQ

ε_t = vector of market specific news/innovations.

TAIL MEMORY : Theoretical

- **Extremal L_0-Approximability [L_0-E-APP]** ($\approx$ Pötscher and Prucha 1991)

$\{x_t\}$ is L_0-E-APP if there exists a function $h_{n,t}^{(l_n)}$ of $\{\varepsilon_{t-l_n},...,\varepsilon_t\}$:

$$\left| I\left(x_t > b_{x,n}\right) - h_{n,t}^{(l_n)} \right| \rightarrow 0 \text{ in probability as } n \rightarrow \infty.$$

As $n \rightarrow \infty$, *extremes* $|x_t| > b_n$ can be perfectly predicted from ε_t.

Says nothing about joint distribution, non-extremes, density smoothness.

TAIL MEMORY : Theoretical

❑ Coverage and Applications

***L_0*-E-APP** : L_0-APP, Near Epoch Dependence, mixingale, mixing.

: GARCH, IGARCH, Explosive GARCH, stochastic volatility

: linear and nonlinear GARCH, Threshold AR's

***L_0*-E-APP** implies : Asymptotically normal tail shape estimator.

Asymptotically normal tail dependence estimator.

TAIL MEMORY : Theoretical

GARCH(1,1)

$$x_t = h_t \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} (0,1), \quad E[\varepsilon_t] = 0, \quad h_t^2 = \alpha + \beta x_{t-1}^2 + \gamma h_{t-1}^2, \quad \alpha > 0, \quad \beta, \gamma \geq 0$$

Tails are regularly varying: $P(|x_t| > \varepsilon) = \varepsilon^{-\kappa} L(\varepsilon)$ (Basrak et al 2006)

$\{x_t\}$ is L_0 **-E-APP** for all $\alpha + \beta \leq 1$ and many $\alpha + \beta \geq 1$.

TAIL MEMORY : Theoretical

GARCH(1,1)

$$x_t = h_t \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} (0,1), \quad E[\varepsilon_t] = 0, \quad h_t^2 = \alpha + \beta x_{t-1}^2 + \gamma h_{t-1}^2, \quad \alpha > 0, \quad \beta, \gamma \geq 0$$

Tails are regularly varying: $P(|x_t| > \varepsilon) = \varepsilon^{-\kappa} L(\varepsilon)$ (Basrak et al 2006)

B. Hill's (1975) : $\sqrt{k_n}\{\hat{\kappa}_n - \kappa\} \overset{d}{\rightarrow} N(0, v_1^2)$

Coeff. of Tail Dep. : $\sqrt{k_n}\left\{\hat{r}_n(h) - r_n(h)\right\} \overset{d}{\rightarrow} N(0, v_2^2)$

Both cases : kernel variance estimators $\hat{v}_{i,n}^2 \overset{p}{\rightarrow} v_i^2$

TAIL MEMORY : Theoretical

L_0-E-APP in general (nonlinear GARCH; regime switching; …)

$$\sqrt{k_n}\{\hat{\kappa}_n - \kappa\} \xrightarrow{d} N(0, v_1^2)$$

$$\sqrt{k_n}\left\{\hat{r}_n(h) - r_n(h)\right\} \xrightarrow{d} N(0, v_2^2)$$

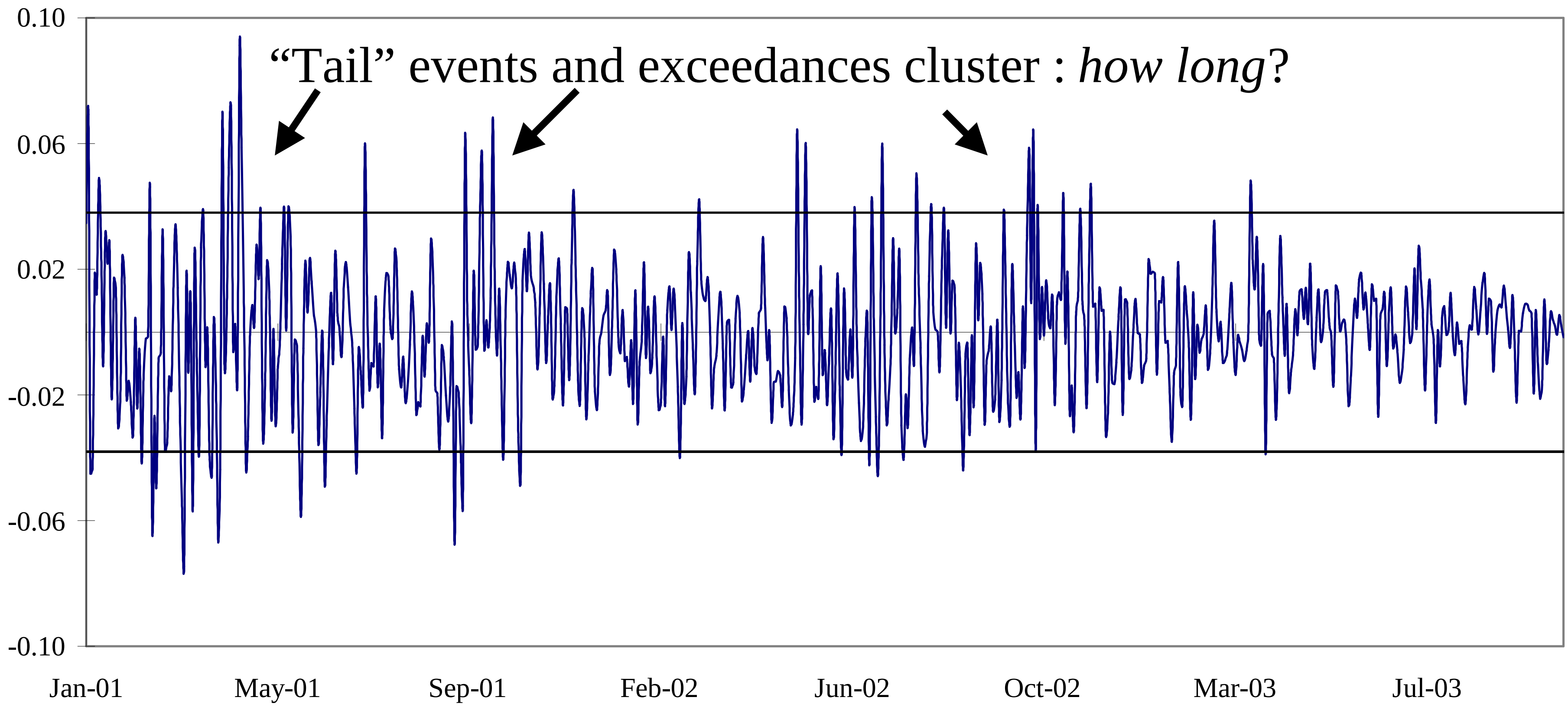
Do not need joint tail specification.

Do not need to restrict non-extremes.

Extremes allowed substantial dependence and heterogeneity.

TAIL MEMORY : Theoretical

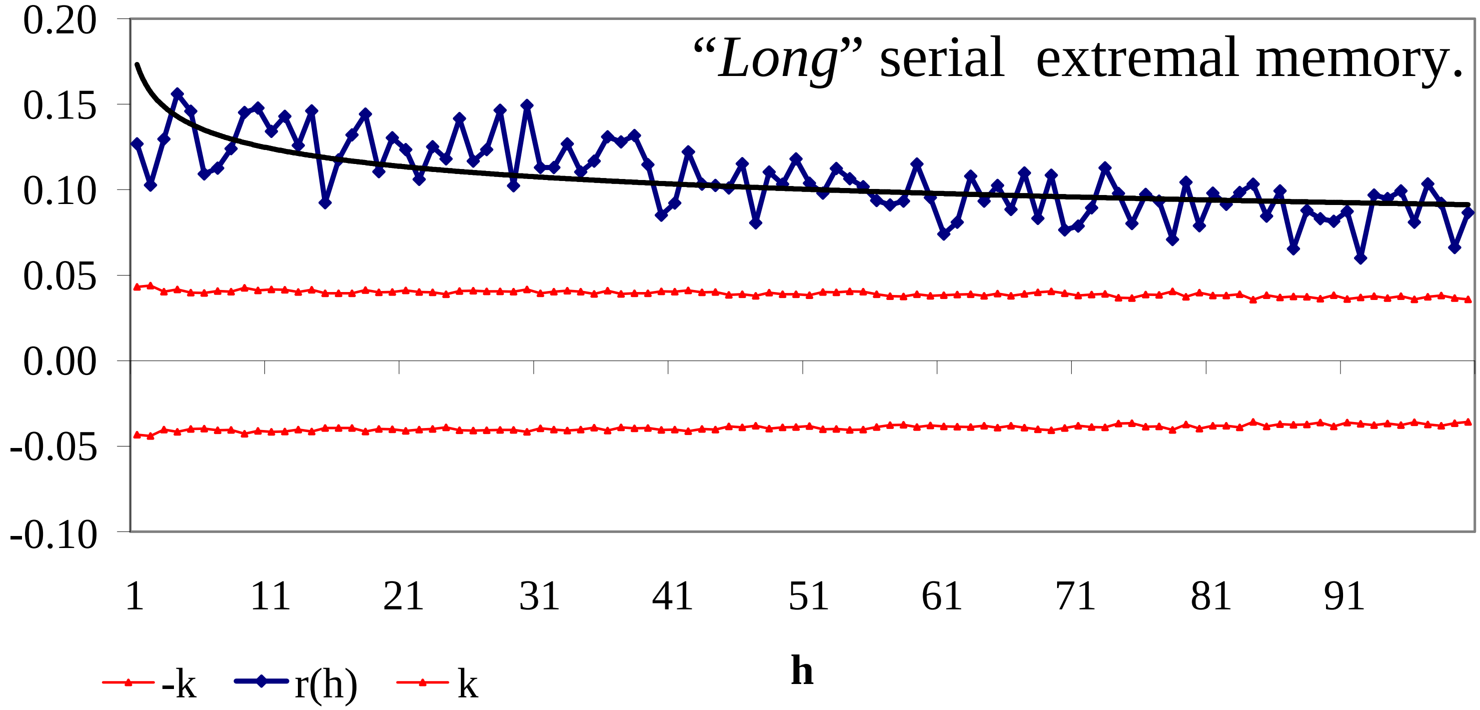
NASDAQ Daily Log Returns and Upper 5th percentile : n = 1044



TAIL MEMORY : Theoretical

$$r_n(h) = \frac{n}{k_n}[P_{x,y,h}(b_{x,n}, b_{y,n}) - P_x(b_{x,n})P_y(b_{y,n})]$$

NASDAQ Median Two Tailed Serial Tail Dependence Coefficient r(h) and 95% Band



NON-TAIL (TAIL-TRIMMED) MEMORY

❑ TAIL-TRIMMED PROCESS

$$X_{n,t}^{(tr)} := x_t \times I(|x_t| \leq b_n)$$

Inherently bounded with infinitely many moments:

$$\left|X_{n,t}^{(tr)}\right| \leq b_{x,n} \quad \text{and} \quad E\left|X_{n,t}^{(tr)}\right|^p < \infty \text{ for all } p > 0.$$

NON-TAIL (TAIL-TRIMMED) MEMORY

❑ POPULATION AND TAIL-TRIMMED MEMORY

If $\{x_t\}$ is $\boldsymbol{L_0}$ **– Approximable** (Pötscher and Prucha 1991)

$$\left|x_t - g_t^{(l)}\right| \rightarrow 0 \text{ in probability as } l \rightarrow \infty$$

for some function $g_t^{(l)}$ of near eoch $\{\varepsilon_{t-l},...,\varepsilon_t\}$

then **tail-trimmed** $X_{n,t}^{(tr)} := x_t \times I(|x_t| \leq b_n)$ satisfies LLN and CLT.

NON-TAIL (TAIL-TRIMMED) MEMORY

- ***L_0*-APP** : Near Epoch Dependence, mixingale, mixing.

  : GARCH, IGARCH, Explosive GARCH, stochastic volatility

  : linear and nonlinear GARCH, Threshold AR's

Tail - trimmed $X_{n,t}^{(tr)} := x_t \times I(|x_t| \leq b_n)$ satisfies **CLT**.

Tail-trimming ensures <u>standard asymptotics</u> for NLLS, <u>**GMM**</u>, QMLE.

NON-TAIL (TAIL-TRIMMED) MEMORY

- **EXAMPLE : Tail-Trimmed GMM (Hill and Renault)**

GARCH(1,1) $: \; x_t = \sigma_t \varepsilon_t, \;\; \varepsilon_t \sim \text{iid}, \quad \sigma_t^2 = \alpha + \beta x_{t-1}^2 + \gamma \sigma_{t-1}^2 \; : \; \theta = [\alpha, \beta, \gamma]$

$$m_t(\theta) = (x_t^2 - \alpha - \beta x_{t-1}^2 - \gamma \sigma_{t-1}^2) \times z_{t-1}^2 \; : \; z_{t-1}^2 = [1, x_{t-1}^2, \sigma_{t-1}^2, ...]'$$

Identification $: \; E\left[m_t(\theta)\right] = 0 \;\; iff \;\; \theta = \theta_0$

$$m_{n,t}^{(tr)}(\theta) \equiv m_t(\theta) \times I\left(m_t^{(a)}(\theta) \le m_{(k_n+1)}^{(a)}(\theta)\right) \; : \; \text{where } \; m_t^{(a)}(\theta) := |m_t(\theta)|$$

NON-TAIL (TAIL-TRIMMED) MEMORY

- **EXAMPLE : Tail-Trimmed GMM (Hill and Renault)**

GARCH(1,1) $: \; x_t = \sigma_t \varepsilon_t , \;\; \varepsilon_t \sim \text{iid}, \;\; \sigma_t^2 = \alpha + \beta x_{t-1}^2 + \gamma \sigma_{t-1}^2 \; : \; \theta = [\alpha, \beta, \gamma]$

$$m_{n,t}^{(tr)}(\theta) \equiv m_t(\theta) \times I\left(m_t^{(a)}(\theta) \le m_{(k_n+1)}^{(a)}(\theta)\right) \; : \; \text{where } m_t^{(a)}(\theta) := |m_t(\theta)|$$

$$\hat{\theta}^{tr} \;\text{ solves }\; \min\left\{\frac{1}{n}\sum_{t=1}^{n} m_{n,t}^{(tr)}(\theta)' \times \hat{\Omega} \times \frac{1}{n}\sum_{t=1}^{n} m_{n,t}^{(tr)}(\theta)\right\} \; : \; \hat{\Omega} \to \Omega \;\text{ (p.d.)}$$

NON-TAIL (TAIL-TRIMMED) MEMORY

❑ EXAMPLE : Tail-Trimmed GMM (Hill and Renault)

GARCH(1,1) $: \; x_t = \sigma_t \varepsilon_t , \;\; \varepsilon_t \sim \text{iid}, \quad \sigma_t^2 = \alpha + \beta x_{t-1}^2 + \gamma \sigma_{t-1}^2 \; : \; \theta = [\alpha, \beta, \gamma]$

$$\hat{\theta}^{tr} \;\; \text{solves} \;\; \min \left\{ \frac{1}{n} \sum_{t=1}^{n} m_{n,t}^{(tr)}(\theta)' \times \hat{\Omega} \times \frac{1}{n} \sum_{t=1}^{n} m_{n,t}^{(tr)}(\theta) \right\}$$

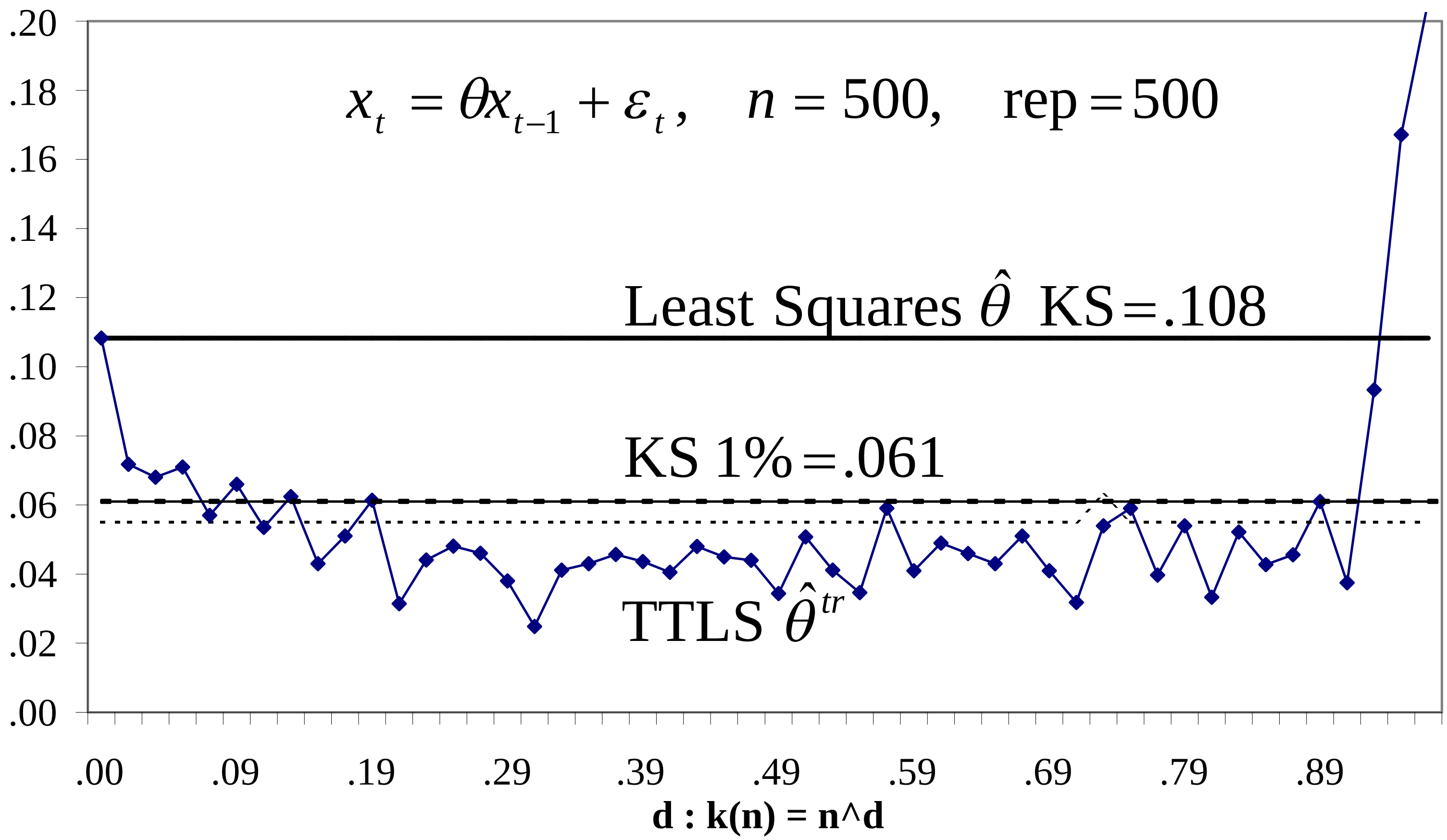
$$V_n^{1/2}(\hat{\theta}^{tr} - \theta_0) \overset{d}{\rightarrow} N(0, I_3) \;\; \text{for some matrix} \;\; V_n \in \Re^{3\times3}$$

Rate of convergence $\leq n^{1/2}$.

EXAMPLE : Tail-Trimmed Least Squares : AR(1)

$$x_t = .9x_{t-1} + \varepsilon_t \quad : \quad \varepsilon_t \overset{iid}{\sim} \text{Pareto } (\alpha = 1.75), \quad \text{trim } k_n \approx n^{\delta}, \ \delta \in (0,1)$$

Tail rimmed Least Squares :
Kolmogorov Smirnov Tests of Normality



SUMMARY

As long as x_t is predictable in some sense using near epoch $\{\varepsilon_{t-l_n},...,\varepsilon_t\}$:

- o **Extreme value estimators** are asymptotically normal :

$$\sqrt{k_n}\{\hat{\kappa}_n - \kappa\} \xrightarrow{d} N(0, v_1^2)$$

$$\sqrt{k_n}\left\{\hat{r}_n(h) - r_n(h)\right\} \xrightarrow{d} N(0, v_2^2)$$

Do need joint tail specification.

Non-extremes can be anything.

Extremes can be highly dependent.

- o **Tail Trimmed estimator(s)** are asymptotically normal (GMM)

- o Covers linear and nonlinear GARCH, linear and nonlinear ARFIMA…