Heavy Tails and Financial Time Series Models

Richard A. Davis
Colorado State University
www.stat.colostate.edu/~rdavis

Thomas Mikosch
Bojan Basrak
Outline

Financial time series modeling
- General comments
- Characteristics of financial time series
- Examples (exchange rate, Amazon)
- Multiplicative models for log-returns (GARCH, SV)

Regular variation
- univariate case
- multivariate case

Applications of regular variation
- Stochastic recurrence equations (GARCH)
- Stochastic volatility
- Time-reversibility
- Point process convergence
- Extremes and extremal index
- Limit behavior of sample correlations

Wrap-up
One possible goal: Develop models that capture essential features of financial data.

Strategy: Formulate families of models that at least exhibit these key characteristics. (e.g., GARCH and SV)

Linkage with goal: Do fitted models actually capture the desired characteristics of the real data?

Answer wrt to GARCH and SV models: Yes and no. Answer may depend on the features.

Stărică’s paper: “Is GARCH(1,1) Model as Good a Model as the Nobel Accolades Would Imply?”

This paper discusses inadequacy of GARCH(1,1) model as a “data generating process” for the data.
Goal of this talk: compare and contrast some of the features of GARCH and SV models.

- Regular-variation of finite dimensional distributions
- Time-reversibility
- Point process convergence
- Extreme value behavior
- Sample ACF
Define $X_t = \ln (P_t) - \ln (P_{t-1})$ (log returns)

• heavy tailed

\[ P(|X_1| > x) \sim RV(-\alpha), \quad 0 < \alpha < 4. \]

• uncorrelated

\[ \hat{\rho}_X(h) \text{ near 0 for all lags } h > 0 \]

• $|X_t|$ and $X_t^2$ have slowly decaying autocorrelations

\[ \hat{\rho}_{|X|}(h) \text{ and } \hat{\rho}_{X^2}(h) \text{ converge to 0 slowly as } h \text{ increases.} \]

• process exhibits ‘volatility clustering’.
Example: Pound-Dollar Exchange Rates
Example: Pound-Dollar Exchange Rates
Hill’s estimate of alpha (Hill Horror plots-Resnick)
15 realizations from GARCH model fitted to exchange rates + exchange rate data. Which one is the real data?
Stărică Plots for Pound-Dollar Exchange Rates

ACF of the squares from the 15 realizations from the GARCH model on previous slide.
Stărică Plots for Pound-Dollar Exchange Rates

15 realizations from SV model fitted to exchange rates + real data. Which one is the real data?
Example: Amazon-returns
Hill’s estimate of alpha (Hill Horror plots-Resnick)
Stărică Plots for the Amazon Data

15 realizations from GARCH model fitted to Amazon + exchange rate data. Which one is the real data?
Stărică Plots for Amazon

ACF of the squares from the 15 realizations from the GARCH model on previous slide.
**Multiplicative models for log(returns)**

**Basic model**

\[ X_t = \ln (P_t) - \ln (P_{t-1}) \quad \text{(log returns)} \]

\[ = \sigma_t Z_t , \]

where

- \{Z_t\} is IID with mean 0, variance 1 (if exists). (e.g. N(0,1) or a \(t\)-distribution with \(\nu\) df.)
- \{\sigma_t\} is the volatility process
- \(\sigma_t\) and \(Z_t\) are independent.

**Properties:**

- \(E X_t = 0, \ \text{Cov}(X_t, X_{t+h}) = 0, \ h>0 \ (\text{uncorrelated if } \text{Var}(X_t) < \infty)\)
- conditional heteroscedastic (condition on \(\sigma_t\)).
Multiplicative models for log(returns)-cont

\[ X_t = \sigma_t Z_t \] (observation eqn in state-space formulation)

Two classes of models for volatility:

(i) GARCH(p,q) process (General AutoRegressive Conditional Heteroscedastic-observation-driven specification)

\[
\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \cdots + \alpha_p X_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_q \sigma_{t-q}^2.
\]

Special case: ARCH(1):

\[
X_t^2 = (\alpha_0 + \alpha_1 X_{t-1}^2)Z_t^2 = \alpha_1 Z_t^2 X_{t-1}^2 + \alpha_0 Z_t^2 = A_t X_{t-1}^2 + B_t
\] (stochastic recurrence eqn)

\[
\rho_{X^2}(h) = \alpha_1^h, \text{ if } \alpha_1^2 < 1/3.
\]
Multiplicative models for log(returns)-cont

GARCH(2,1): $X_t = \sigma_t Z_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \alpha_2 X_{t-2}^2 + \beta_1 \sigma_{t-1}^2$.

Then $Y_t = (\sigma_t^2, X_{t-1}^2)'$ follows the SRE given by

$$
\begin{bmatrix}
\sigma_t^2 \\
X_{t-1}^2
\end{bmatrix} =
\begin{bmatrix}
\alpha_1 Z_{t-1}^2 + \beta_1 \\
Z_{t-1}^2
\end{bmatrix}
\begin{bmatrix}
\sigma_{t-1}^2 \\
X_{t-2}^2
\end{bmatrix} +
\begin{bmatrix}
\alpha_0 \\
0
\end{bmatrix}
$$

Questions:

- Existence of a unique stationary solution to the SRE?
- Regular variation of the joint distributions?
Multiplicative models for log(returns)-cont

\[ X_t = \sigma_t Z_t \]  (observation eqn in state-space formulation)

(ii) stochastic volatility process (parameter-driven specification)

\[
\log \sigma_t^2 = \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j}, \quad \sum_{j=-\infty}^{\infty} \psi_j^2 < \infty, \{ \varepsilon_t \} \sim \text{IIDN}(0, \sigma^2) 
\]

\[
\rho_{X^2}(h) = \text{Cor}(\sigma_t^2, \sigma_{t+h}^2) / EZ_1^4
\]

Question:

• Joint distributions of process regularly varying if distr of \( Z_1 \) is regularly varying?
Two models for log(returns)-cont

GARCH(1,1):

\[ X_t = \sigma_t Z_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad \{Z_t\} \sim \text{IID}(0,1) \]

Stochastic Volatility:

\[ X_t = \sigma_t Z_t, \quad \log(\sigma_t^2) = \phi_0 + \phi_1 \log(\sigma_{t-1}^2) + \varepsilon_t, \quad \{\varepsilon_t\} \sim \text{IIDN}(0,\sigma^2) \]

Main question:

What intrinsic features in the data (if any) can be used to discriminate between these two models?
Regular variation — univariate case

**Def:** The random variable $X$ is *regularly varying with index* $\alpha$ if

$$P(|X|> t \cdot x)/P(|X|>t) \to x^{-\alpha} \text{ and } P(X> t)/P(|X|>t) \to p,$$

or, equivalently, if

$$P(X> t \cdot x)/P(|X|>t) \to p x^{-\alpha} \text{ and } P(X< -t \cdot x)/P(|X|>t) \to q x^{-\alpha},$$

where $0 \leq p \leq 1$ and $p+q=1$.

**Equivalence:**

$X$ is RV($\alpha$) *if and only if* $P(X \in t \cdot \cdot ) /P(|X|>t) \to_v \mu(\cdot \cdot )$

($\to_v$ vague convergence of measures on $\mathbb{R}\{0\}$). In this case,

$$\mu(dx) = \left(p \alpha x^{-\alpha-1} I(x>0) + q \alpha (-x)^{-\alpha-1} I(x<0)\right) dx$$

**Note:** $\mu(tA) = t^{-\alpha} \mu(A)$ for every $t$ and $A$ bounded away from 0.
Another formulation (polar coordinates):

Define the ± 1 valued rv $\theta$, $P(\theta = 1) = p$, $P(\theta = -1) = 1 - p = q$.

Then

$X$ is RV$(\alpha)$ if and only if

$$\frac{P(|X| > t x, X / |X| \in S)}{P(|X| > t)} \rightarrow x^{-\alpha} P(\theta \in S)$$

or

$$\frac{P(|X| > t x, X / |X| \in \bullet)}{P(|X| > t)} \rightarrow_{v} x^{-\alpha} P(\theta \in \bullet)$$

($\rightarrow_{v}$ vague convergence of measures on $S^0 = \{-1, 1\}$).
Multivariate regular variation of $X = (X_1, \ldots, X_m)$: There exists a random vector $\theta \in S^{m-1}$ such that

$$\frac{P(|X| > t x, X/|X| \in \cdot)}{P(|X| > t)} \rightarrow_{\nu} x^{-\alpha} P(\theta \in \cdot)$$

($\rightarrow_{\nu}$ vague convergence on $S^{m-1}$, unit sphere in $\mathbb{R}^m$).

- $P(\theta \in \cdot)$ is called the spectral measure
- $\alpha$ is the index of $X$.

Equivalence:

$$\frac{P(X \in t\cdot)}{P(|X| > t)} \rightarrow_{\nu} \mu(\cdot)$$

$\mu$ is a measure on $\mathbb{R}^m$ which satisfies for $x > 0$ and $A$ bounded away from 0,

$$\mu(xB) = x^{-\alpha} \mu(xA).$$
Regular variation — multivariate case (cont)

Examples:

1. If $X_1 > 0$ and $X_2 > 0$ are iid RV$(\alpha)$, then $X = (X_1, X_2)$ is multivariate regularly varying with index $\alpha$ and spectral distribution

$$P(\theta = (0,1)) = P(\theta = (1,0)) = .5 \text{ (mass on axes).}$$

Interpretation: Unlikely that $X_1$ and $X_2$ are very large at the same time.

Figure: plot of $(X_{t1}, X_{t2})$ for realization of 10,000.
2. If \( X_1 = X_2 > 0 \), then \( X = (X_1, X_2) \) is multivariate regularly varying with index \( \alpha \) and *spectral distribution*

\[
P( \theta = (1/\sqrt{2}, 1/\sqrt{2}) ) = 1.
\]

3. AR(1): \( X_t = 0.9 X_{t-1} + Z_t \), \( \{Z_t\} \sim \text{IID symmetric stable (1.8)} \)

Distr of \( \theta \):
- \( \pm(1.9)/\sqrt{1.81} \), W.P. 0.9898
- \( \pm(0,1) \), W.P. 0.0102

**Figure:** plot of \( (X_t, X_{t+1}) \) for realization of 10,000.
Applications of multivariate regular variation

• Domain of attraction for *sums of iid random vectors* (Rvaceva, 1962). That is, when does the partial sum

\[ a_n^{-1} \sum_{t=1}^{n} X_t \]

converge for some constants \(a_n\)?

• *Spectral measure* of multivariate stable vectors.

• *Domain of attraction* for componentwise maxima of iid random vectors (Resnick, 1987). Limit behavior of

\[ a_n^{-1} \bigvee_{t=1}^{n} X_t \]

• Weak convergence of *point processes* with iid points.

• Solution to *stochastic recurrence equations*, \(Y_t = A_t Y_{t-1} + B_t\)

• Weak convergence of *sample autocovariances*.
Linear combinations:

*\( X \sim \text{RV}(\alpha) \Rightarrow \) all linear combinations of *\( X \) are regularly varying.

i.e., there exist *\( \alpha \) and slowly varying fcn *\( L(.) \), s.t.

\[
P(c^T X > t)/(t^{\alpha}L(t)) \to w(c), \text{ exists for all real-valued } c,
\]

where

\[
w(tc) = t^{-\alpha}w(c).
\]

Use vague convergence with *\( A_c = \{y: c^T y > 1\} \), i.e.,

\[
\frac{P(X \in tA_c)}{t^{-\alpha}L(t)} = \frac{P(c^T X > t)}{P(|X| > t)} \to \mu(A_c) = w(c),
\]

where *\( t^{\alpha}L(t) = P(|X| > t) \).
Applications of multivariate regular variation (cont)

Converse?

\( \mathbf{X} \sim \text{RV}(\alpha) \iff \) all linear combinations of \( \mathbf{X} \) are regularly varying?

There exist \( \alpha \) and slowly varying fcn \( L(.) \), s.t.

\[
(\text{LC}) \quad P(\mathbf{c}^T \mathbf{X} > t)/(t^{\alpha L(t)}) \rightarrow w(\mathbf{c}), \text{ exists for all real-valued } \mathbf{c}.
\]

Theorem (Basrak, Davis, Mikosch, `02). Let \( \mathbf{X} \) be a random vector.

1. If \( \mathbf{X} \) satisfies (LC) with \( \alpha \) non-integer, then \( \mathbf{X} \) is \( \text{RV}(\alpha) \).

2. If \( \mathbf{X} > 0 \) satisfies (LC) for non-negative \( \mathbf{c} \) and \( \alpha \) is non-integer, then \( \mathbf{X} \) is \( \text{RV}(\alpha) \).

3. If \( \mathbf{X} > 0 \) satisfies (LC) with \( \alpha \) an odd integer, then \( \mathbf{X} \) is \( \text{RV}(\alpha) \).
Applications of multivariate regular variation (cont)

There exist $\alpha$ and slowly varyingfcn $L(.)$, s.t.

\[(LC) \quad P(c^T X > t)/(t^{\alpha}L(t)) \rightarrow w(c), \text{ exists for all real-valued } c.\]

1. If $X$ satisfies $(LC)$ with $\alpha$ non-integer, then $X$ is $RV(\alpha)$.
2. If $X > 0$ satisfies $(LC)$ for non-negative $c$ and $\alpha$ is non-integer, then $X$ is $RV(\alpha)$.
3. If $X > 0$ satisfies $(LC)$ with $\alpha$ an odd integer, then $X$ is $RV(\alpha)$.

Remark: Hult and Lindskog (2005) show that:

- 1 cannot be extended to integer $\alpha$.
- 2 cannot be extended to integer $\alpha$.
- It is unknown if 3 can be extended to even integers.
1. Kesten (1973). Under general conditions, (LC) holds with \( L(t) = 1 \) for stochastic recurrence equations of the form

\[
Y_t = A_t Y_{t-1} + B_t, \quad (A_t, B_t) \sim \text{IID},
\]

\[A_t \text{ } d \times d \text{ random matrices, } B_t \text{ random } d\text{-vectors.}\]

It follows that the distributions of \( Y_t \), and in fact all of the finite dim’l distrs of \( Y_t \) are regularly varying (if \( \alpha \) is non-even).

2. GARCH processes. Since squares of a GARCH process can be embedded in a SRE, the \textit{finite dimensional distributions} of a \textit{GARCH} are regularly varying.
Example of ARCH(1): \[ X_t = (\alpha_0 + \alpha_1 X_{t-1}^2)^{1/2}Z_t, \quad \{Z_t\} \sim \text{IID}. \]

\( \alpha \) found by solving \( E|\alpha_1 Z^2|^{\alpha/2} = 1. \)

\[
\begin{array}{c|cccc}
\alpha_1 & .312 & .577 & 1.00 & 1.57 \\
\alpha  & 8.00 & 4.00 & 2.00 & 1.00 \\
\end{array}
\]

Distr of \( \theta \):

\[
P(\theta \in \bullet) = E\{||B,Z||^{\alpha} I(\text{arg}((B,Z)) \in \bullet)\}/E||B,Z||^{\alpha}
\]

where

\[
P(B = 1) = P(B = -1) = .5
\]
**Examples (cont)**

Example of ARCH(1): \( \alpha_0=1, \alpha_1=1, \alpha=2, X_t=(\alpha_0+\alpha_1 X_{t-1}^2)^{1/2}Z_t, \) \( \{Z_t\}\sim\text{IID} \)

**Figures:** plots of \((X_t, X_{t+1})\) and estimated distribution of \(\theta\) for realization of 10,000.
Excursion to time-reversibility

Reversibility. A stationary sequence of random variables \( \{X_t\} \) is time-reversible if \((X_1, \ldots, X_n) =_d (X_n, \ldots, X_1)\) for all \( n > 1 \).

Results: i) IID sequences \( \{Z_t\} \) are time-reversible.

ii) Linear time series (with a couple obvious exceptions) are time-reversible iff Gaussian. (Breidt and Davis `91)

Application: If plot of time series does not look time-reversible, then it cannot be modeled as IID or a Gaussian process. Use the “flip and compare” inspection test!

![Series plots]
Reversibility.  *Does the following series look time-reversible?*
Examples (cont)

Example of ARCH(1): \( \alpha_0=1, \alpha_1=1, \alpha=2, X_t=(\alpha_0+\alpha_1 X^2_{t-1})^{1/2}Z_t, \{Z_t\}\sim\text{IID} \)

Is this process time-reversible?

**Figures:** plots of \((X_t, X_{t+1})\) and \((X_{t+1}, X_t)\) implies *non-reversible.*

---

**Figures:**

- Scatter plot of \((x_{t-1}, x_{t+1})\)
- Scatter plot of \((x_1, x_2)\)
Example: SV model $X_t = \sigma_t Z_t$

Suppose $Z_t \sim RV(\alpha)$ and

$$X_t = \sigma_t Z_t, \quad \log \sigma_t^2 = \phi_0 + \phi_1 \log \sigma_{t-1}^2 + \varepsilon_t, \quad \{\varepsilon_t\} \sim \text{IID N}(0, \sigma^2)$$

Then $Z_n = (Z_1, \ldots, Z_n)'$ is regular varying with index $\alpha$ and so is

$$X_n = (X_1, \ldots, X_n)' = \text{diag}(\sigma_1, \ldots, \sigma_n) Z_n$$

with spectral distribution concentrated on $(\pm 1, 0), (0, \pm 1)$.

**Figure:** plot of

$(X_t, X_{t+1})$ for realization of 10,000.
Example: SV model $X_t = \sigma_t Z_t$

- SV processes are time-reversible if log-volatility is Gaussian.
- Asymptotically time-reversible if log-volatility is non-Gaussian
**Point process convergence**

**Theorem** (Davis & Hsing `95, Davis & Mikosch `97). Let \( \{X_t\} \) be a stationary sequence of random \( m \)-vectors. Suppose

(i) finite dimensional distributions are jointly regularly varying (let \((\theta_{-k}, \ldots, \theta_k)\) be the vector in \( S^{(2k+1)m-1} \) in the definition).

(ii) mixing condition \( A(\alpha_n) \) or strong mixing.

(iii) \( \lim_{k \to \infty} \limsup_{n \to \infty} P( \bigvee_{k \leq |t| \leq n} |X_t| > a_n y \big| |X_0| > a_n y) = 0. \)

Then

\[
\gamma = \lim_{k \to \infty} E( |\theta_0^{(k)}| \alpha - \bigvee_{j=1}^k |\theta_j^{(k)}| ) / E |\theta_0^{(k)}| \alpha
\]

(extrimal index) exists. If \( \gamma > 0 \), then

\[
N_n := \sum_{t=1}^n \xi X_t / a_n \xrightarrow{d} N := \sum_{i=1}^\infty \sum_{j=1}^\infty \xi_{P_i Q_{ij}} ,
\]
Point process convergence (cont)

- \( (P_i) \) are points of a Poisson process on \((0, \infty)\) with intensity function
  \[ \nu(dy) = \gamma \alpha y^{-\alpha-1} dy. \]

- \( \sum_{j=1}^{\infty} \varepsilon_{Q_{ij}}, i \geq 1, \) are iid point process with distribution \(Q\), and \(Q\) is the weak limit of
  \[
  \lim_{k \to \infty} E\left( \sum_{j=1}^{k} \varepsilon_{\theta^{(k)}_j} \right) / E\left( \sum_{|l| \leq k} \varepsilon_{\theta^{(k)}_l} \right) \]

Remarks:

1. GARCH and SV processes satisfy the conditions of the theorem.

2. Limit distribution for sample extremes and sample ACF follows from this theorem.
Setup

- $X_t = \sigma_t Z_t$, \quad \{Z_t\} \sim \text{IID } (0,1)
- $X_t$ is RV ($\alpha$)
- Choose $\{b_n\}$ s.t. \quad nP(X_t > b_n) \to 1$

Then

$P^n(b_n^{-1}X_1 \leq x) \to \exp\{-x^{-\alpha}\}$.

Then, with $M_n = \max\{X_1, \ldots, X_n\}$,

(i) GARCH:

$P(b_n^{-1}M_n \leq x) \to \exp\{-\gamma x^{-\alpha}\}$,

\(\gamma\) is extremal index (0 < \(\gamma\) < 1).

(ii) SV model:

$P(b_n^{-1}M_n \leq x) \to \exp\{-x^{-\alpha}\}$,

extremal index $\gamma = 1$ no clustering.
Remarks about extremal index.

(i) $\gamma < 1$ implies clustering of exceedances

(ii) Numerical example. Suppose $c$ is a threshold such that

$$P^n(b_n^{-1}X_1 \leq c) \sim .95$$

Then, if $\gamma = .5$, $P(b_n^{-1}M_n \leq c) \sim (.95)^5 = .975$

(iii) $1/\gamma$ is the mean cluster size of exceedances.

(iv) Use $\gamma$ to discriminate between GARCH and SV models.

(v) Even for the light-tailed SV model (i.e., $\{Z_t\} \sim$IID $N(0,1)$, the extremal index is 1 (see Breidt and Davis `98)
Extremes for GARCH and SV processes (cont)

Absolute values of ARCH

![Graph showing absolute values of ARCH over time.](image-url)
Extremal Index Estimates for Amazon

\[ \gamma_1 = \text{block method} \]
\[ \gamma_2 = \frac{1}{\text{mean cluster size}} \]
\[ \gamma_3 = \text{interval method (Ferro and Segers)} \]
\[ \gamma_4 = \text{interval method (Ferro and Segers)} \]
Summary of results for ACF of GARCH(p,q) and SV models

\textbf{GARCH(p,q)}

$\alpha \in (0, 2)$:

$$\left( \hat{\rho}_X(h) \right)_{h=1, \ldots, m} \overset{d}{\longrightarrow} \left( V_h / V_0 \right)_{h=1, \ldots, m},$$

$\alpha \in (2, 4)$:

$$\left( n^{1-2/\alpha} \hat{\rho}_X(h) \right)_{h=1, \ldots, m} \overset{d}{\longrightarrow} \gamma_X^{-1}(0) \left( V_h \right)_{h=1, \ldots, m}. $$

$\alpha \in (4, \infty)$:

$$\left( n^{1/2} \hat{\rho}_X(h) \right)_{h=1, \ldots, m} \overset{d}{\longrightarrow} \gamma_X^{-1}(0) \left( G_h \right)_{h=1, \ldots, m}. $$

\textbf{Remark:} Similar results hold for the sample ACF based on $|X_t|$ and $X_t^2$. 
Summary of results for ACF of GARCH(p,q) and SV models (cont)

**SV Model**

\[ \alpha \in (0, 2): \]
\[ \left( \frac{n}{\ln n} \right)^{1/\alpha} \hat{\rho}_X(h) \xrightarrow{d} \frac{\| \sigma_1 \sigma_{h+1} \|^2_\alpha}{\| \sigma_1 \|_\alpha^2} \frac{S_h}{S_0}. \]

\[ \alpha \in (2, \infty): \]
\[ \left( n^{1/2} \hat{\rho}_X(h) \right)_{h=1, \ldots, m} \xrightarrow{d} \gamma_X^{-1}(0)(G_h)_{h=1, \ldots, m}. \]
Sample ACF for GARCH and SV Models (1000 reps)

(a) GARCH(1,1) Model, n=10000

(b) SV Model, n=10000
Sample ACF for Squares of GARCH (1000 reps)

(a) GARCH(1,1) Model, n=10000

(b) GARCH(1,1) Model, n=100000
Sample ACF for Squares of SV (1000 reps)

(c) SV Model, n=10000

(d) SV Model, n=100000
Amazon returns (GARCH model)

GARCH(1,1) model fit to Amazon returns:

\[ \alpha_0 = 0.0002493, \quad \alpha_1 = 0.0385, \quad \beta_1 = 0.957, \quad X_t = (\alpha_0 + \alpha_1 X_{t-1}^2)^{1/2} Z_t, \quad \{Z_t\} \sim \text{IID } t(3.672) \]

Simulation from GARCH(1,1) model
Amazon returns (SV model)

Stochastic volatility model fit to Amazon returns:

![ACF plots](image_url)
Wrap-up

- **Regular variation** is a flexible tool for modeling both **dependence** and **tail heaviness**.

- Useful for establishing **point process convergence** of heavy-tailed time series.

- **Extremal index** $\gamma < 1$ for GARCH and $\gamma = 1$ for SV.

- ACF has faster convergence for SV.