Dependence and Tail Modeling II

The Extremes and Tails of Financial Time Series[1]

Thomas Mikosch

University of Copenhagen

www.math.ku.dk/∼mikosch/Tehran10

Joint work with Richard A. Davis (Columbia)

Mikosch (2003), Davis and Mikosch (2009a,b,c)

[1]Tehran, May 19, 2010

Contents

1. "Stylized facts"

- Consider log-returns

$$\begin{aligned} X_t &= \log(P_t/P_{t-1}) = \log P_t - \log P_{t-1} \\ &= \log\left(1 + \frac{P_t - P_{t-1}}{P_{t-1}}\right) \\ &\approx \frac{P_t - P_{t-1}}{P_{t-1}}, \qquad t = 0, 1, 2 \ldots, \end{aligned}$$

  where (P_t) is a speculative price series (share price, stock index, foreign exchange (FX) rate,...) with time scale: days, weeks, hours, ..., high frequency data.

- Why log-returns? (X_t) is unit free.

  Common belief: prices P_t "increase" exponentially on average, (X_t) is "stationary".

1.1. Marginal distribution.

- sample mean close to zero
- sample variance of order $10^{-5}, 10^{-6}, \ldots$
- distribution is roughly symmetric in the center
- density is sharply peaked at zero (leptokurtic)
- data are non-Gaussian

- heavy tails on both sides,

$$P(X_t > x) \approx x^{-\alpha} \quad \text{as } x \to \infty,\ \alpha \in (3,5)$$

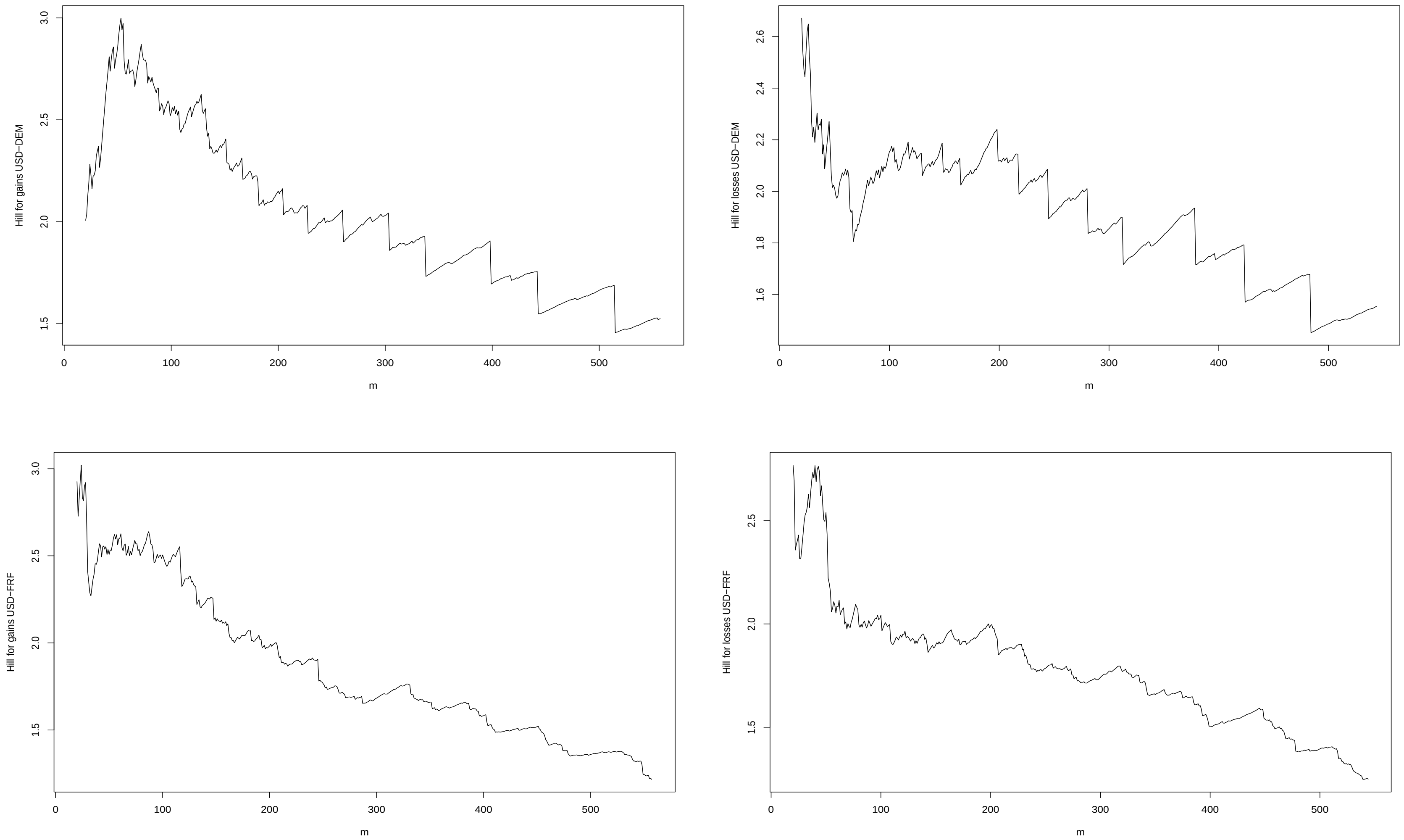

FIGURE 1. Hill estimation (based on up to 50% of the order statistics) for 5 minute foreign exchange rate log-returns, USD-DEM (top) and USD-FRF (bottom). Left: gains. Right: losses.

1.2. Dependence: autocorrelations, clustering of extremes.

- Classical time series analysis: main goal is second order structure of (Gaussian) stationary time series (X_t)
- This structure is determined by

  autocovariance function (ACVF)

$$\gamma_X(h) = \mathrm{cov}(X_0, X_h)\,, \quad h \in \mathbb{Z}\,.$$

  autocorrelation function (ACF)

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}\,, \quad h \in \mathbb{Z}\,.$$

- ACF determines dependence structure of stationary Gaussian (X_t).

- **ACF used for parameter estimation, model testing, prediction of Gaussian/non-Gaussian time series (ARMA, FARIMA,...)**

  Brockwell and Davis (1991,1996)

- **Since one does not know the ACF/ACVF of real-life data one needs to estimate them: sample ACVF and sample ACF**

$$\gamma_{n,X}(h) = \frac{1}{n}\sum_{t=1}^{n-|h|}(X_t - \overline{X}_n)(X_{t+h} - \overline{X}_n)$$

$$\rho_{n,X}(h) = \frac{\gamma_{n,X}(h)}{\gamma_{n,X}(0)}, \quad h \in \mathbb{Z}.$$

- **If (X_t) is *stationary ergodic*, var$(X_t) < \infty$,**

$$\gamma_{n,X}(h) \stackrel{\text{a.s.}}{\to} \gamma_X(h), \quad \rho_{n,X}(h) \stackrel{\text{a.s.}}{\to} \rho_X(h).$$

The ACF stylized fact

- **Sample ACF $\rho_{n,X}$ of returns are negligible** **(possible exception: 1st lag)**

- **Sample ACFs $\rho_{n,|X|}$, ρ_{n,X^2} are positive and decay very slowly** **(typical for "long" time series)**

- **This is often interpreted as long memory or long-range dependence (LRD),** see Samorodnitsky and Taqqu (1994), Doukhan et al. (2003)

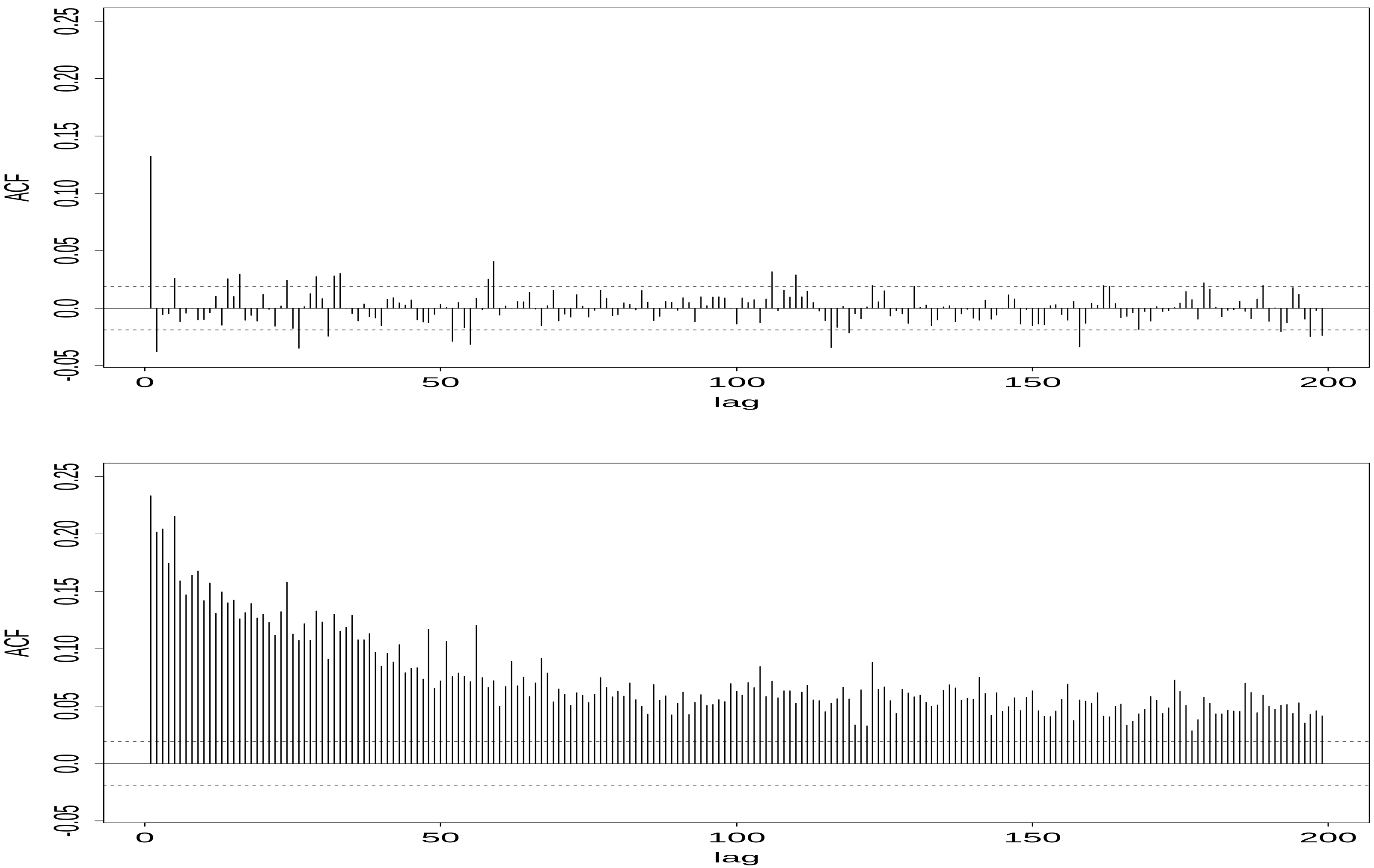

Figure 2. Sample ACFs for the log-returns *(top)* and absolute log-returns *(bottom)* of the *S&P500*. Here and in what follows, the horizontal lines in graphs displaying sample ACFs are set as the **95%** confidence bands $(\pm 1.96/\sqrt{n})$ corresponding to the ACF of iid Gaussian noise.

The extremal stylized fact

- **Covariances and correlations are not good tools for describing the dependence of extremes.**
- **High/low level exceedances of returns tend to appear in clusters.**
- **Extremal serial dependence can be measured by the extremal index or the extremogram; see Talk III.**

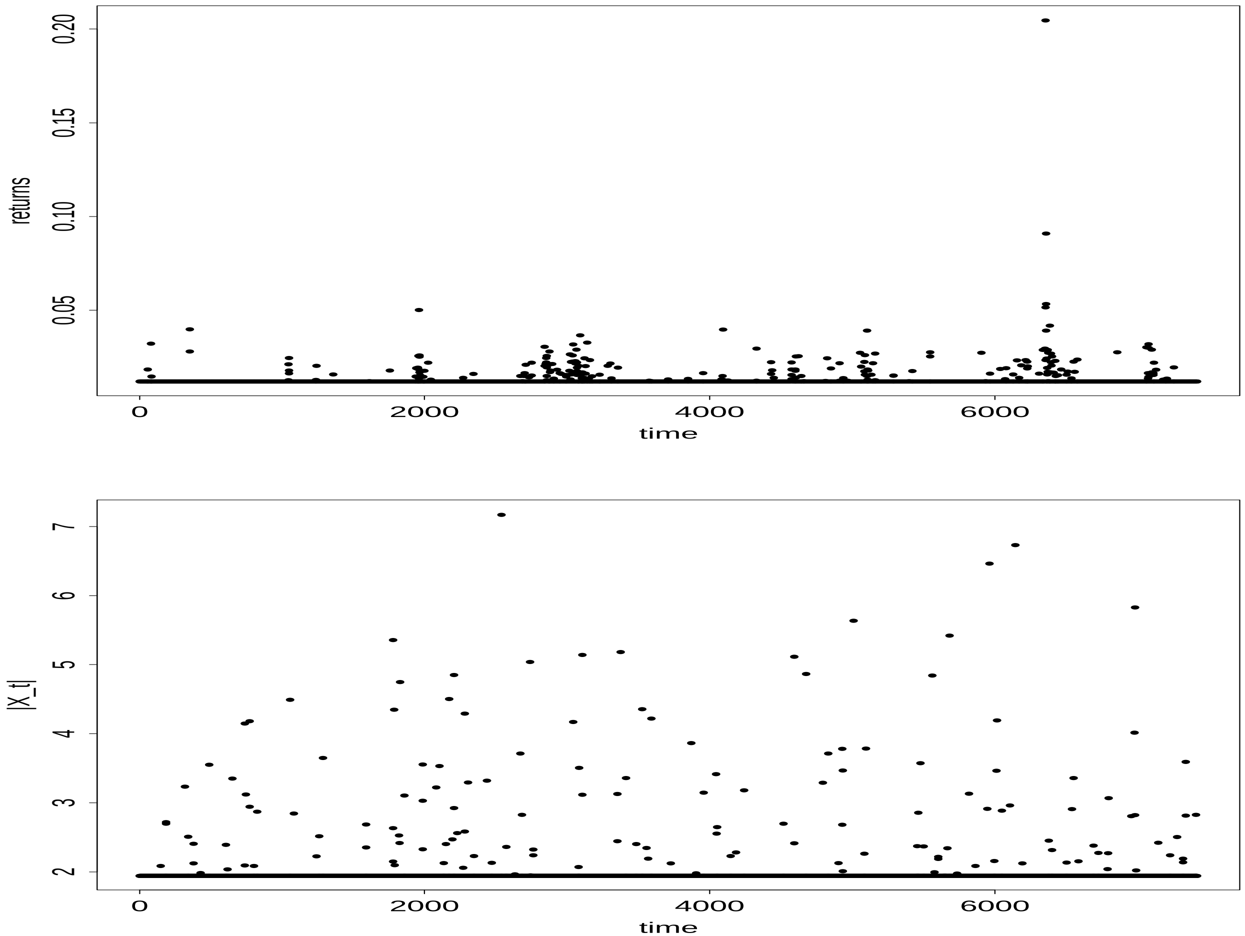

FIGURE 3. *Top:* Absolute returns $|\boldsymbol{X}_t|$ of the *S&P500* series for which both $|\boldsymbol{X}_t|$ and $|\boldsymbol{X}_{t+1}|$ exceed the $\mathbf{87\%}$ quantile of the data. The latter is indicated by the bottom line. *Bottom:* The same kind of plot for an iid sequence from a student distribution with $\mathbf{4}$ degrees of freedom. In the former case pairwise exceedances occur in clusters, in the latter case exceedances appear uniformly scattered through time.

Stylized fact: aggregational Gaussianity

- **The center of the distribution of**

$$\left(\boldsymbol{X}_{t+1} + \cdots + \boldsymbol{X}_{t+h} - \overline{\boldsymbol{X}_{t,t+h}}\right) / \sqrt{\boldsymbol{h}}$$

  becomes close to the normal distribution.

- **This is an indication of CLT behavior and points at the fact that** $\mathrm{var}(\boldsymbol{X}_t) < \infty$ **(no infinite variance stable distributions).**

- **This is difficult to establish by statistical means since aggregation means that one uses less data.**

2. Can classical time series analysis model returns?

- Classical time series analysis is about linear processes[2]

$$X_t = \sum_{j=0}^{\infty} \psi_j \, Z_{t-j} \,, \quad (Z_t) \text{ iid}$$

in particular ARMA(p, q) processes:

$$\varphi(B) X_t = X_t - \varphi_1 X_{t-1} - \cdots - \varphi_p X_{t-p} =$$

$$\theta(B) Z_t = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \,,$$

where

$$\varphi(z) = 1 - \sum_{i=1}^{p} \varphi_i \, z^i \,, \quad \theta(z) = 1 + \sum_{j=1}^{q} \theta_j \, z^j \,,$$

and $B^k A_t = A_{t-k}$ is the backshift operator.

[2] Brockwell, Davis (1991,1996)

Examples. AR(2) process

$$(1 - \varphi_1 B - \varphi_2 B^2) X_t = X_t - \varphi_1 X_{t-1} - \varphi_2 X_{t-2} = Z_t \,.$$

MA(2) process

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} = (1 + \theta_1 B + \theta_2 B^2) Z_t \,.$$

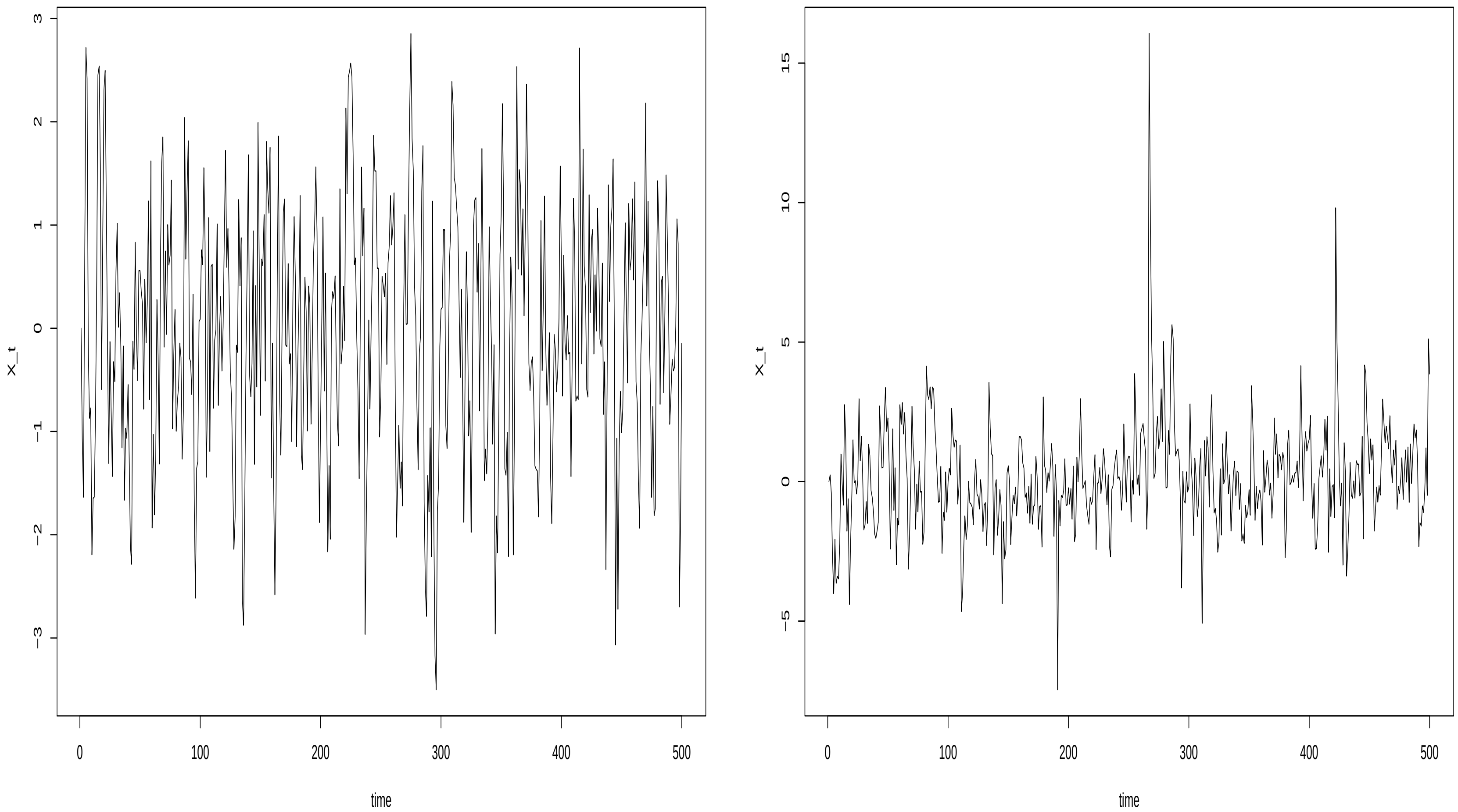

Figure 4. Simulation of AR(1) process $\boldsymbol{X_t = 0.5X_{t-1} + Z_t}$ with iid standard normal noise (left) and student noise with 3 degrees of freedom (right).

Which stylized facts can be explained by linear processes?

- **Heavy tails of X_t are only possible if the noise Z_t has heavy tails.**
- **ACF behavior. (X_t) must be iid or MA(1) (or an "all pass" process).**

Example. If (X_t) were iid or an MA(1) process, the ACFs of (X_t), $(|X_t|)$ and (X_t^2) would vanish at lags $h \geq 2$.

- **A linear process cannot explain the complicated dependence structure of the sequences $(|X_t|)$ and (X_t^2).**
- **Conclusion: We need a "non-linear" model !**

2.1. Multiplicative models.

$$X_t = \sigma_t \, Z_t \,, \quad t \in \mathbb{Z} \,.$$

We assume

- (Z_t) iid mean zero (or symmetric) noise, $EZ_t^2 = 1$
- $\sigma_t > 0$ and Z_t independent for every t
- (σ_t) volatility sequence (unobservable) strictly stationary
- (X_t) strictly stationary

Why this model?

- Conditional forecast of X_t given $\sigma_t = f(\text{past})$. Then $\mathcal{L}(X_t \mid \text{past})$ is known, e.g. $Z_t \sim N(0,1)$ and $X_t \mid \text{past} \sim N(0, \sigma_t^2)$ (Conditional VaR)
- $\rho_X(h) = \text{corr}(X_0, X_h) = 0$, $h \neq 0$, in agreement with stylized facts.
- To some extent, it can explain that $\rho_{|X|}(h)$ and $\rho_{X^2}(h)$ are different from zero and decay "slowly".
- Although the whole time series (X_t) is stationary one can model changing *conditional variance* over time quite flexibly. ("volatility clusters", "conditional heteroscedasticity")

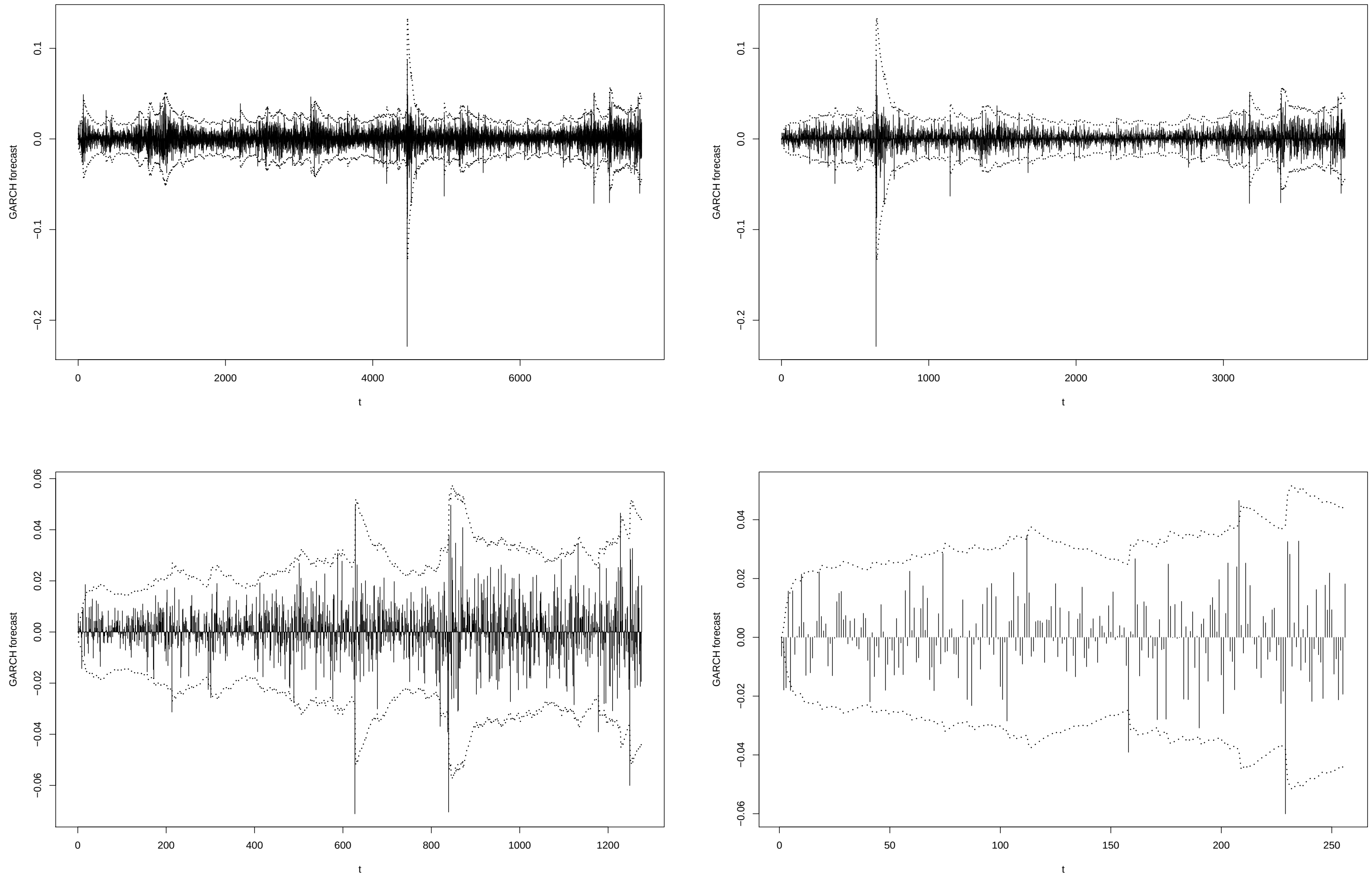

Figure 5. One day $\mathbf{95\%}$ distributional forecasts of log-returns of the S&P500 composite stock index $\mathbf{(}$from top left, top right, bottom left to bottom right: $\mathbf{30}$, $\mathbf{15}$, $\mathbf{5}$, $\mathbf{1}$ years of data$\mathbf{)}$ based on a GARCH$\mathbf{(1,1)}$ model with iid standard normal noise and parameters $\boldsymbol{\alpha_0 = 10^{-6}}$, $\boldsymbol{\alpha_1 = 0.07}$, $\boldsymbol{\beta_1 = 0.96}$. The extreme values of the log-returns are not correctly captured by the model.

2.2. The ARCH family. Engle (1982)

$$X_t = \sigma_t Z_t, \quad (Z_t) \text{ iid}, \ EZ = 0, \ \text{var}(Z) = 1$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2, \quad t \in \mathbb{Z}.$$

for $\alpha_0 > 0$, certain $\alpha_i \geq 0$, $\alpha_p > 0$.

ARCH(p) autoregressive conditionally heteroscedastic process of order p

Why "autoregressive" process?

- $\nu_t = X_t^2 - \sigma_t^2 = \sigma_t^2\,(Z_t^2 - 1)$ is white noise (zero correlations, constant variance) if (σ_t^2) strictly stationary and $EX_t^4 < \infty$.

- $\varphi(B)X_t^2 = \alpha_0 + \nu_t\,, \quad t \in \mathbb{Z}\,,$ where

$$\varphi(z) = 1 - \sum_{i=1}^{p} \alpha_i\, z^i\,.$$

- **Problem:** ARCH(p) does not fit returns well unless p is large.

2.3. The GARCH model. Bollerslev (1986), Taylor (1986)

- $$\varphi(B)\, X_t^2 = \alpha_0 + \beta(B)\, \nu_t\,, \quad t \in \mathbb{Z}\,,$$

  where $\nu_t = X_t^2 - \sigma_t^2$ and

  $$\varphi(z) = 1 - \sum_{i=1}^{p} \alpha_i\, z^i - \sum_{j=1}^{q} \beta_j\, z^j\,,\ \beta(z) = 1 + \sum_{j=1}^{q} \beta_j\, z^j\,,$$

  for certain $\alpha_i, \beta_j \geq 0$, $\alpha_p \beta_q > 0$.

- Generalized ARCH(p, q) (GARCH(p, q))

  $$X_t = \sigma_t\, Z_t\,,$$

  $$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i\, X_{t-i}^2 + \sum_{j=1}^{q} \beta_j\, \sigma_{t-j}^2\,, \quad t \in \mathbb{Z}\,.$$

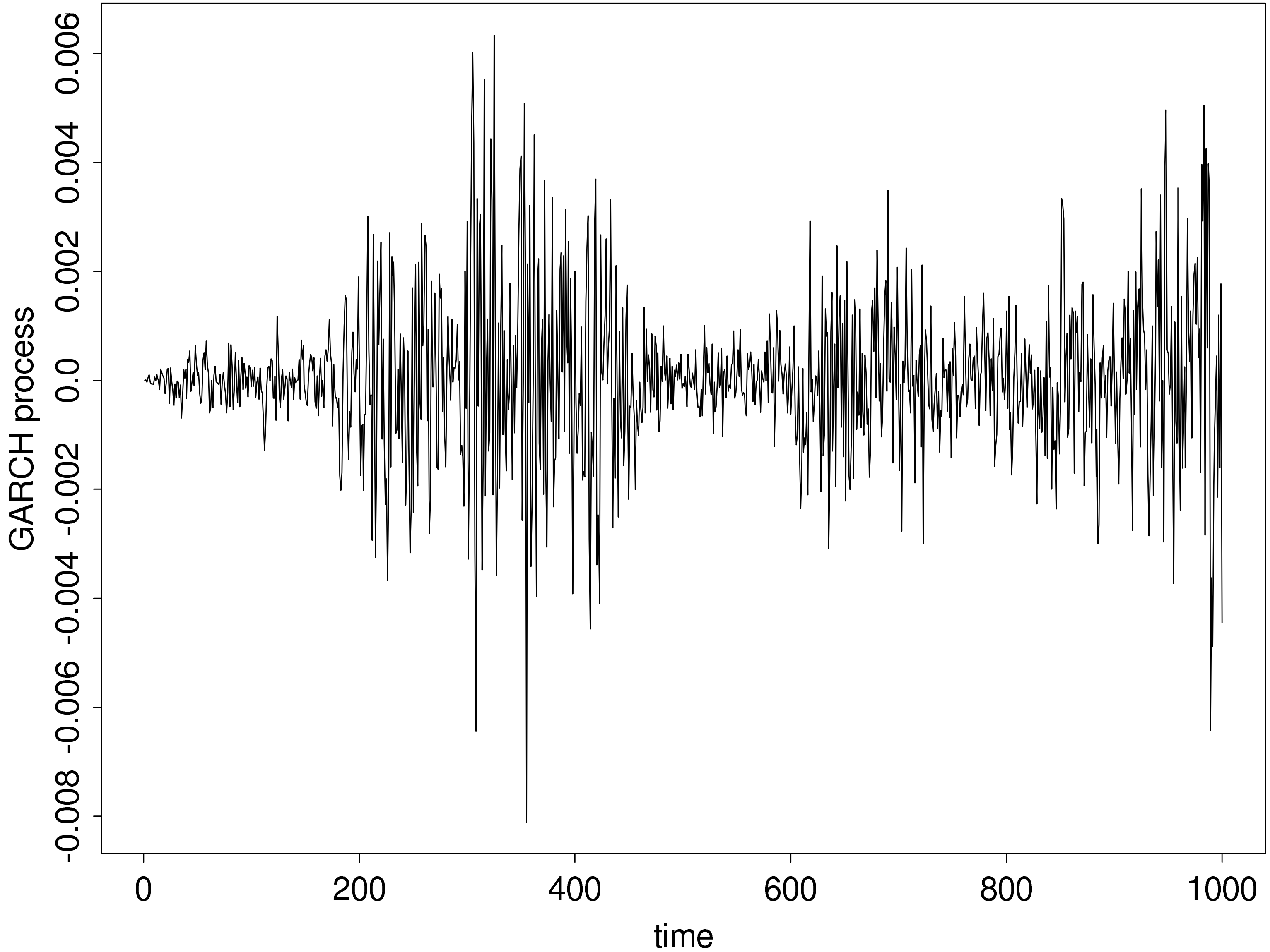

Figure 6. Simulation of a GARCH(1,1) process with Gaussian noise.

2.4. An easier model: The stochastic volatility model.

$$X_t = \sigma_t \, Z_t \,, \qquad (Z_t) \text{ iid centered or symmetric}$$

- (σ_t) strictly stationary
- (Z_t) and (σ_t) independent
- We always assume that $(\log \sigma_t)$ is a Gaussian linear process:
$$\log \sigma_t = \sum_{j=0}^{\infty} c_j \, \eta_{t-j} \,, \qquad \eta_j \sim N(0,1) \text{ iid}$$
- No feedback between (σ_t) and (Z_t)
- Dependence modeled via (σ_t), tails via (Z_t)

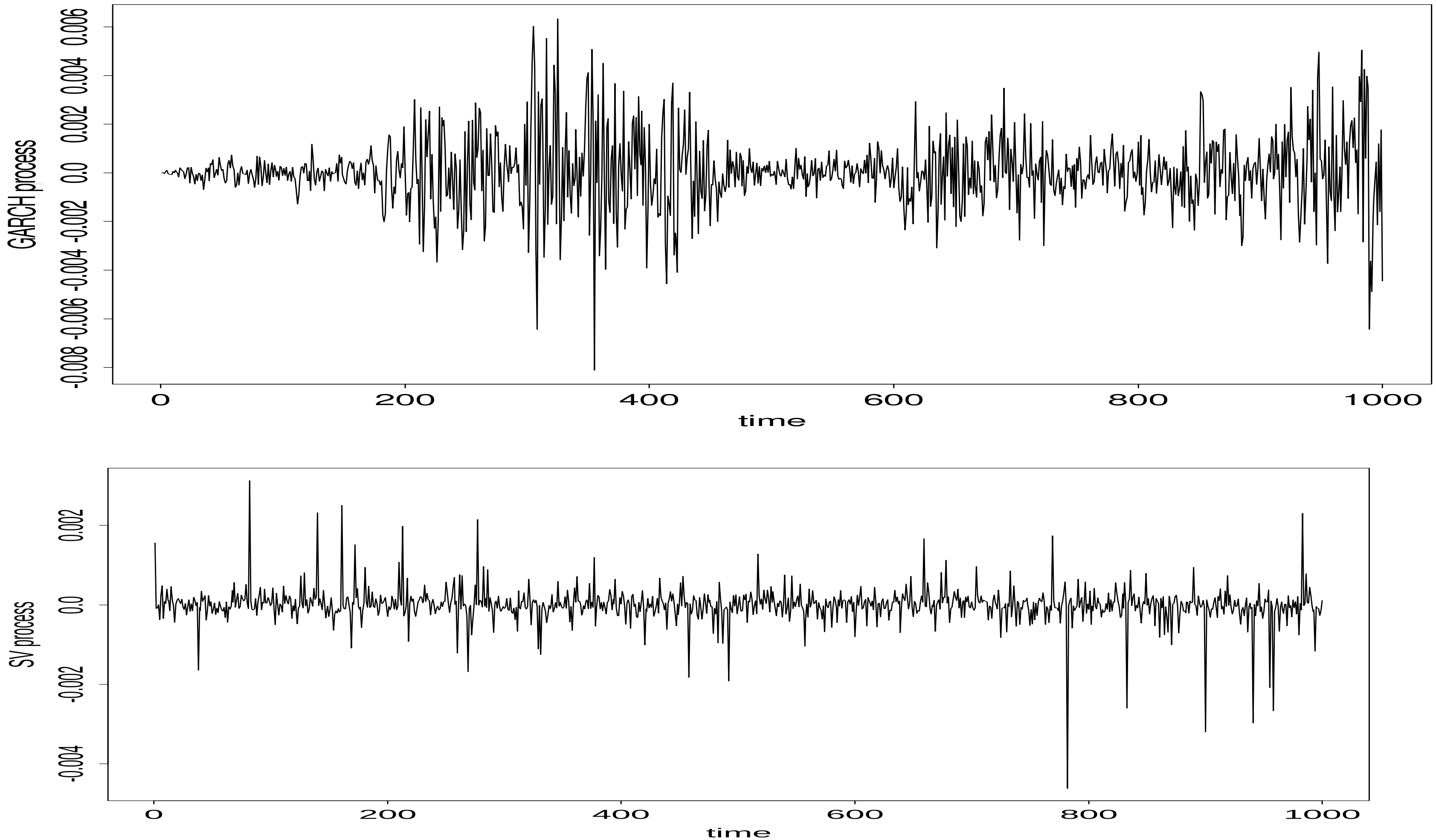

FIGURE 7. *Top:* 1000 simulated values from the GARCH$(1,1)$ model $X_t = (0.0001 + 0.1X_{t-1}^2 + 0.9\sigma_{t-1}^2)^{0.5}Z_t$ for iid standard normal (Z_t). *Bottom:* 1000 simulated values from the stochastic volatility model $X_t = \mathrm{e}^{Y_t}\, Z_t$ for iid student noise (Z_t) with 4 degrees of freedom, $Y_t = 0.5Y_{t-1} + 0.3\eta_{t-1} + \eta_t$ is an ARMA(1,1) process with iid standard normal noise (η_t).

2.5. The stationarity problem.

The SV model.

$$X_t = \sigma_t\, Z_t\,, \quad (Z_t) \text{ iid, } (\sigma_t) \text{ and } (Z_t) \text{ independent.}$$

The log-volatility sequence

$$\log \sigma_t = \sum_{i=1}^{\infty} c_i\, \eta_{t-i}\,, \quad \text{iid } \eta_i \sim N(0,1)$$

is *strictly stationary* if and only if $\sum_i c_i^2 < \infty$.

(X_t) is stationary if and only if (σ_t) is stationary.

The GARCH model.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 .$$

- The sequence $X_t = \sigma_t Z_t$ is stationary if (σ_t) is stationary.
- Example GARCH(1, 1). Write

$$A_t = \alpha_1 Z_{t-1}^2 + \beta_1 , \quad B_t = \alpha_0 \quad \text{and} \quad Y_t = \sigma_t^2 .$$

Then

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta_1) \sigma_{t-1}^2$$

or

$$Y_t = A_t Y_{t-1} + B_t . \tag{1}$$

- A_t and Y_{t-1} are independent and $((A_t, B_t))$ is iid.
- (1) is a stochastic recurrence equation (SRE).

- Iterate back and notice that $B_t = B_1$,

$$Y_t = A_t \cdots A_{t-r} Y_{t-r-1} + \sum_{i=t-r}^{t} A_t \cdots A_{i+1} B_1$$

$$= \sum_{i=-\infty}^{t} A_t \cdots A_{i+1} B_1 \,, \quad t \in \mathbb{Z} \,, \tag{2}$$

Hence, if $-\infty \leq E \log A_1 < 0$, (2) converges a.s. for every fixed t and (Y_t) is the a.s. unique strictly stationary solution to SRE.

- **Theorem.** (Nelson (1990), Bougerol and Picard (1992a,b)) There exists an a.s. unique non-vanishing strictly stationary causal (i.e., depending only on past and present values of the Z's) solution of SRE with $B_t \equiv \alpha_0$ if and only if $\alpha_0 > 0$ and $E \log(\alpha_1 Z_1^2 + \beta_1) < 0$.

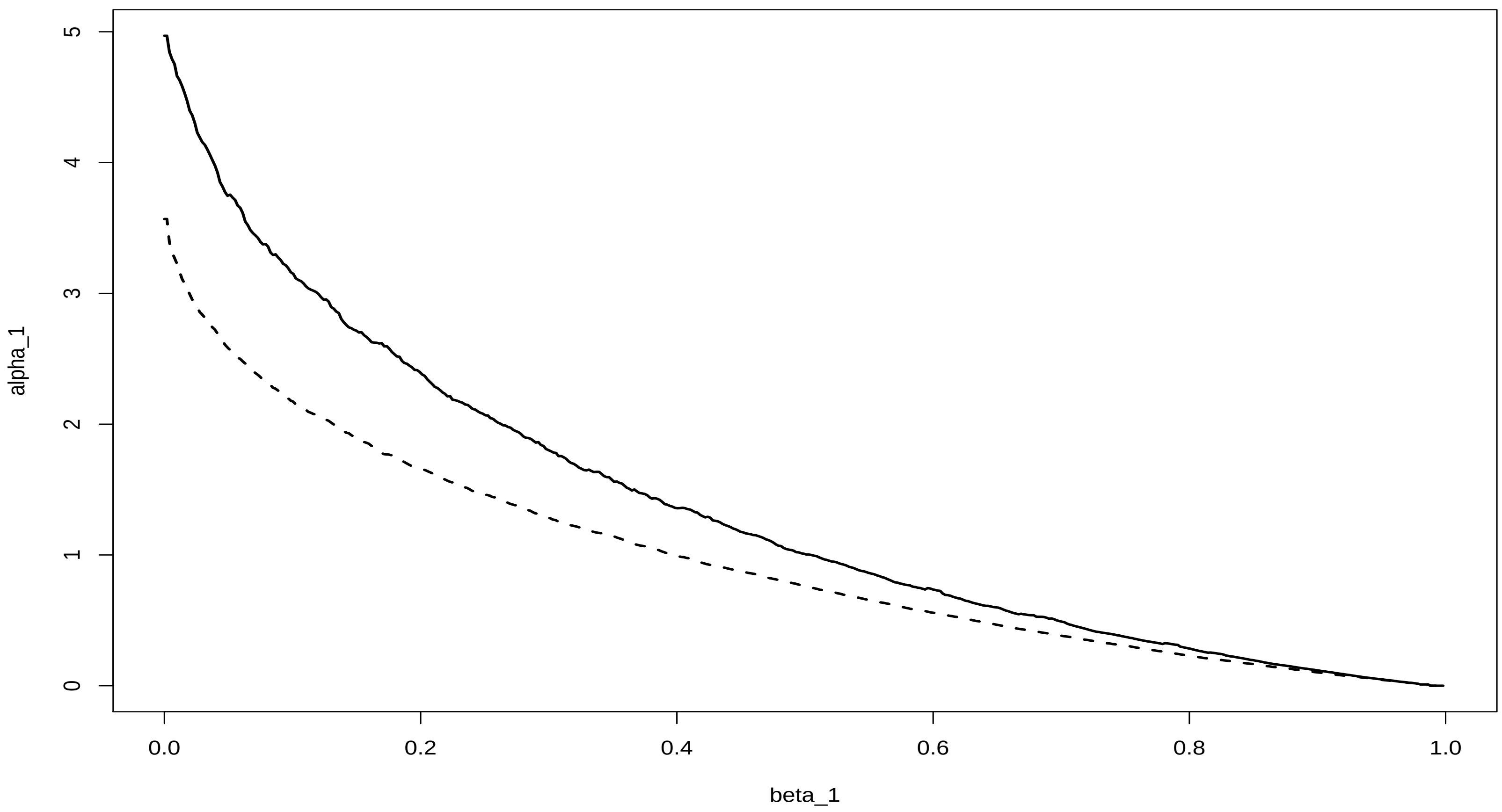

Figure 8. The $(\boldsymbol{\alpha_1}, \boldsymbol{\beta_1})$-areas below the two curves guarantee the existence of a stationary GARCH$(\mathbf{1}, \mathbf{1})$ process. *Solid line:* IID student noise with $\mathbf{4}$ degrees of freedom with variance $\mathbf{1}$. *Dotted line:* IID standard normal noise.

The general GARCH case

$$\mathbf{Y}_t = \left(\sigma^2_{t+1}, \ldots, \sigma^2_{t-q+2}, X^2_t, \ldots, X^2_{t-p+2}\right)',$$

$$\mathbf{A}_t = \begin{pmatrix} \alpha_1 Z_t^2 + \beta_1 & \beta_2 & \cdots & \beta_{q-1} & \beta_q & \alpha_2 & \alpha_3 & \cdots & \alpha_p \\ 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 \\ Z_t^2 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

$$\mathbf{B}_t = (\alpha_0, 0, \ldots, 0)',$$

Then

$$\mathbf{Y}_t = \mathbf{A}_t \mathbf{Y}_{t-1} + \mathbf{B}_t, \quad t \in \mathbb{Z}.$$

- This is a multivariate stochastic recurrence equation.
- Existence of stationary solution depends on top Lyapunov exponent

$$\gamma = \inf \left\{ n^{-1} \, E \log \| \mathrm{A}_n \cdots \mathrm{A}_1 \| \right\} < 0 \, ,$$

$\| \cdot \|$ operator norm corresponding to norm $| \cdot |$.

- In general, γ cannot be calculated explicitly.
- For GARCH(1,1), $A_n \cdots A_1 = \prod_{t=1}^{n} (\alpha_1 Z_t^2 + \beta_1)$, and so $\gamma = E \log(\alpha_1 Z_1^2 + \beta_1)$.

- **Theorem.** (Bougerol and Picard (1992a,b)) The GARCH(p, q) SRE has the a.s. unique strictly stationary non-vanishing causal solution

$$\mathbf{Y}_t = \sum_{i=-\infty}^{t} \mathbf{A}_t \cdots \mathbf{A}_{i+1} \mathbf{B}_i = \mathbf{B}_t + \sum_{i=-\infty}^{t-1} \mathbf{A}_t \cdots \mathbf{A}_{i+1} \mathbf{B}_i \,,$$

if and only if $\alpha_0 > 0$ and $\gamma < 0$.

- $\sum_{j=1}^{q} \beta_j < 1$ is a necessary condition for $\gamma < 0$.

- $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ is sufficient for $\gamma < 0$ if $EZ_1^2 = 1$ and $EZ_1 = 0$, and ensures that $EX_t^2 < \infty$.

- **Example: The integrated GARCH process (IGARCH)** Engle and Bollerslev (1986)

$$\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j = 1,$$

IGARCH process has infinite variance; **see p. 46.** **This is not desirable (e.g. ACF would not make sense) and is in contrast to statistical evidence.**

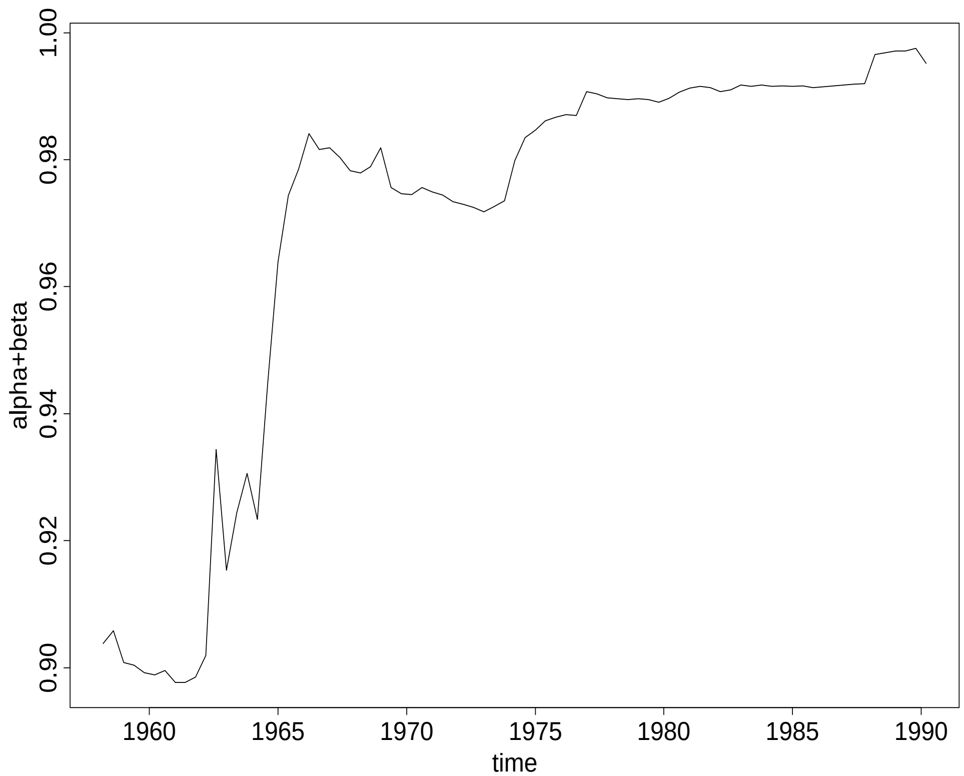

FIGURE 9. The estimated values of $\boldsymbol{\alpha_1} + \boldsymbol{\beta_1}$ for an increasing sample of the *S&P500* log-returns.

2.6. The tails and extremes of return models.

SV models. Davis and M. (2001a,b, 2009a,c)

$$X_t = \sigma_t Z_t\,, \quad \log \sigma_t = \sum_{i=0}^{\infty} c_i \,\eta_{t-i}\,.$$

$\eta_i \sim N(0,1)$ iid, independent of iid (Z_t) regularly varying with index $\alpha > 0$ such that $P(Z_1 > x)/P(|Z_1| > x) \to p > 0$.

- Then σ_t is log-normal, hence $E\sigma_t^p < \infty$ for all $p > 0$.
- One-dimensional Breiman lemma implies (see Breiman (1965))

$$P(X_1 > x) \;=\; P(\sigma_1 Z_1 > x) \sim E\sigma_1^\alpha \, P(Z_1 > x)$$

$$P(X_1 \le -x) \;\sim\; E\sigma_1^\alpha \, P(Z_1 \le -x)\,.$$

The one-dimensional marginals of an SV model are regularly varying with index α.[3]

[3] Breidt and Davis (1998) consider the case of a *light-tailed SV model* with normal Z.

- Multivariate Breiman implies (see Basrak, Davis, M. (2002a))

$$(X_1,\ldots,X_n)' = \mathrm{diag}(\sigma_1,\ldots,\sigma_n)\,(Z_1\ \ldots,Z_n)'$$

is regularly varying since $\mathrm{Z} = (Z_1\ \ldots,Z_n)'$ is regularly varying, independent of $\mathrm{diag}(\sigma_1,\ldots,\sigma_n)$.

- Regular variation of a vector $\mathrm{X} \in \mathbb{R}^n$ means that $|\mathrm{X}|$ is regularly varying with some index $\alpha > 0$ and

$$P(\mathrm{X}/|\mathrm{X}| \in \cdot \mid |\mathrm{X}| > x) \stackrel{w}{\to} P(\Theta \in \cdot) = P_\Theta(\cdot)\,, \quad x \to \infty\,,$$

where the distribution P_Θ is the spectral measure of X on the unit sphere $\mathbb{S}^{n-1}$ of $\mathbb{R}^n$.

- For a SV model, the spectral measure of the lagged vector $(X_1,\ldots,X_n)$ is concentrated at the axes as in the iid case. (It is very unlikely that two values X_i and X_i, $i \neq j$, are large at the same time.)

- We have for $i \neq j$, as $x \to \infty$,

$$\frac{P(X_i > x\,, X_j > x)}{P(X_j > x)} = \frac{E[P(\sigma_i Z_i > x\,, \sigma_j Z_j > x \mid (\sigma_t))]}{P(X_j > x)}$$
$$= \frac{E[P(\sigma_i Z_1 > x \mid \sigma_i)\, P(\sigma_j Z_1 > x \mid \sigma_j)]}{P(\sigma_1 Z_1 > x)}$$
$$\sim \frac{E(\sigma_i \sigma_j)^\alpha [P(Z_1 > x)]^2}{E\sigma_1^\alpha P(Z_1 > x)} \to 0\,.$$

- Similar calculations for left/right tails.

- Tail dependence coefficient for $i \neq j$ is zero:

$$\lambda(X_i, X_j) = \lim_{x \to \infty} P(X_i > x \mid X_j > x) = 0\,.$$

as in the iid case.

- For an iid sequence $(\widetilde{X}_t)$ with the same marginal distribution as X_t, with (a_n) such that $P(X_1 > a_n) \sim n^{-1}$,

$$\lim_{n\to\infty} P(a_n^{-1}\widetilde{M}_n \le x) = \lim_{n\to\infty} P(a_n^{-1}M_n \le x)$$
$$= \Phi_\alpha(x) = \mathrm{e}^{-x^{-\alpha}}, \quad x > 0\,.$$

- The extremal behavior of a regularly varying SV process is very much like the extremal behavior of an iid sequence with the same marginal distribution.
- On the other hand, the ACFs of $(|X_t|)$ and (X_t^2) can decay to zero arbitrarily slowly, i.e., extremal dependence is not related to the ACF.

The tails of a solution to a SRE. Kesten (1973)

- Recall from p. 30 that squares of a GARCH process can be embedded in the SRE:

(3) $$\mathbf{Y}_n = \mathbf{A}_n \mathbf{Y}_{n-1} + \mathbf{B}_n\,, \quad n \in \mathbb{Z}\,,$$

- **Assumptions.** $((\mathbf{A}_n, \mathbf{B}_n))$ is an iid sequence of $d \times d$ matrices $\mathbf{A}_n$ with non-negative entries and d-dimensional non-negative-valued random vectors $\mathbf{B}_n$.
  - For some $\epsilon > 0$, $E\|\mathbf{A}_1\|^\epsilon < 1$.
  - The set

    $$\{\log \|\mathbf{a}_n \cdots \mathbf{a}_1\| : \ n \geq 1,\ \mathbf{a}_n \cdots \mathbf{a}_1 > 0 \text{ and } \mathbf{a}_n, \ldots, \mathbf{a}_1 \in \text{the support of } P_{\mathbf{A}_1}\}$$

    generates a dense group in $\mathbb{R}$.
  - There exists a $\kappa_0 > 0$ such that $E\Big(\min_{i=1,\ldots,d} \sum_{j=1}^d A_{ij}\Big)^{\kappa_0} \geq d^{\kappa_0/2}$ and

    $E\Big(\|\mathbf{A}_1\|^{\kappa_0} \log^+ \|\mathbf{A}_1\|\Big) < \infty\,.$

- **Then the following statements hold:**

a) **There exists a unique solution $\kappa_1 \in (0, \kappa_0]$ to the equation**

$$0 = \lim_{n\to\infty} \frac{1}{n} \log E\|\mathrm{A}_n \cdots \mathrm{A}_1\|^{\kappa_1} .$$

b) **If $E|\mathrm{B}_1|^{\kappa_1} < \infty$, there exists a unique strictly stationary causal solution (Y_n) to the stochastic recurrence equation (3).**

c) **If $E|\mathrm{B}_1|^{\kappa_1} < \infty$, then Y_1 satisfies the following regular variation condition: for all $\mathrm{x} \in \mathbb{R}^d \setminus \{0\}$**

$$\lim_{u\to\infty} u^{\kappa_1} P((\mathrm{x}, \mathrm{Y}_1) > u) = w(\mathrm{x})$$

exists and is positive for all non-negative-valued vectors $\mathrm{x} \neq 0$.

In particular, all components of Y_1 are regularly varying with index κ_1.

- **The case $d = 1$.** κ_1 is the solution to

$$0 = \lim_{n\to\infty} \frac{1}{n} \log E[(A_n \cdots A_1)^{\kappa_1}] \quad \text{i.e., } EA_1^{\kappa_1} = 1.$$

- **For GARCH(1, 1).**

$$\sigma_t^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta_1)\sigma_{t-1}^2$$

implies that $EA_1^{\kappa_1} = E[(\alpha_1 Z^2 + \beta_1)^{\kappa_1}] = 1$.

- Hence with $\alpha = 2\kappa_1$

$$P(\sigma_t > x) \sim c\, x^{-\alpha}\,, \quad x \to \infty\,,$$

and by Breiman,

$$P(X_t > x) = P(\sigma_t Z_t > x) \sim E[Z_+^\alpha]\, P(\sigma > x) \sim E[Z_+^\alpha]\, c\, x^{-\alpha}\,.$$

- **Moreover, the finite-dimensional distributions of a GARCH$(1,1)$ (and of any GARCH(p,q)) process are regularly varying.**[4]

- **Due to regular variation of (X_t), for $i \neq j$, the tail dependence coefficient**

$$\lambda(X_i, X_j) = \lim_{x\to\infty} P(X_j > x \mid X_i > x) = \lim_{x\to\infty} \frac{P(X_i > x, X_j > x)}{P(X_i > x)}$$

exists and is positive.

[4] It takes some efforts to prove that regular variation of all linear combinations of a random vector implies multivariate regular variation; see Basrak, Davis, M. (2002a,b).

- For an iid sequence $(\widetilde{X}_t)$ with the same marginal distribution as X_t, with (a_n) such that $P(X_1 > a_n) \sim n^{-1}$, i.e., $a_n \sim (c\,n)^{1/\alpha}$,

$$\lim_{n\to\infty} P(a_n^{-1}\widetilde{M}_n \le x) = \Phi_\alpha(x) = \mathrm{e}^{-x^{-\alpha}}\,,$$
$$\lim_{n\to\infty} P(a_n^{-1}M_n \le x) = \Phi_\alpha^{\theta_X}(x) = \mathrm{e}^{-\theta_X\,x^{-\alpha}}\,, \qquad x > 0\,,$$

for some $\theta_X \in (0,1)$.

- θ_X is the extremal index of the sequence (X_t). It is the reciprocal of the expected cluster size above high thresholds.
- Recall. The extremal behavior of a regularly varying SV model is similar to the extremal behavior of an iid sequence with the same marginal distribution, e.g. $\theta_X = 1$.

- **The extremal behavior of a GARCH process is characterized by clusters of extremes above high thresholds: there is dependence in the tails.**
- **Note: There is a crucial difference between the GARCH and SV models:**

  Regular variation of the SV model is a consequence of the regular variation of the noise (Z_t), while regular variation of a GARCH model is due to regular variation of the volatility sequence (σ_t).

- **For the GARCH(1, 1) case the tail index $\alpha = 2\kappa_1$ of X_t^2 and σ_t^2 is determined through the equation**

$$EA_1^{\kappa_1} = E(\alpha_1 Z_1^2 + \beta_1)^{\kappa_1} = 1\,.$$

- **This equation can be solved numerically or by Monte Carlo methods.**

Examles. $\alpha_1 = 0.1,\ \alpha = 2\kappa_1$

Table 2.1. Results for α, standard normal noise.

β_1	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0.0
α	2.0	12.5	16.2	18.5	20.2	21.7	23.0	24.2	25.4	26.5
β_1	0.81	0.82	0.83	0.84	0.85	0.86	0.87	0.88	0.89	
α	11.9	11.3	10.7	9.9	9.1	8.1	7.0	5.6	4.0	

Table 2.2. Results for α, student noise with 4 degrees of freedom and variance 1.

β_1	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0.0
α	2.0	3.68	3.83	3.88	3.91	3.92	3.93	3.93	3.94	3.94
β_1	0.81	0.82	0.83	0.84	0.85	0.86	0.87	0.88	0.89	
α	3.65	3.61	3.56	3.49	3.41	3.29	3.13	2.90	2.54	

THE IGARCH(1,1) CASE $\alpha_1 + \beta_1 = 1$

- Since $EZ_1^2 = 1$,

$$EA_1^{\kappa_1} = E(\alpha_1 Z_1^2 + \beta_1)^{\kappa_1} = 1$$

has the unique solution $\kappa_1 = 1$.

- Kesten's result yields regular variation with index 2 for $\mathrm{X}_d = (X_1, \ldots, X_d)$:

$$P(|\mathrm{X}_d| > x) \sim c\, x^{-2},$$

$$P\big(\mathrm{X}_d/|\mathrm{X}_d| \in \cdot \mid |\mathrm{X}_d| > x\big) \xrightarrow{w} P(\Theta \in \cdot) \quad \text{as } x \to \infty.$$

- In particular, X_t has infinite variance.

3. Asymptotic theory for the sample ACVF of financial time series models

- The SV and GARCH model satisfy the strong mixing condition under mild extra conditions.
- If $E|X_1 X_{1+h}|^{2+\delta} < \infty$ for $h = 0, 1, \ldots, m$, some $\delta > 0$ and $EX^4 < \infty$ for the GARCH(p, q) process, then the central limit theorem holds:

$$\left(n^{1/2}(\gamma_{n,X}(h) - \gamma_X(h))\right)_{h=0,\ldots,m} \xrightarrow{d} (V_h)_{h=0,\ldots,m} \,,$$

$$\left(n^{1/2}(\rho_{n,X}(h) - \rho_X(h))\right)_{h=1,\ldots,m} \xrightarrow{d} \gamma_X^{-1}(0)\,(V_h - \rho_X(h)\,V_0)_{h=1,\ldots,m} \,,$$

where $(V_1, \ldots, V_m)$ is multivariate normal with mean zero and covariance matrix

$$\left[\sum_{k=-\infty}^{\infty} \mathrm{cov}(X_0 X_i, X_k X_{k+j})\right]_{i,j=1,\ldots,m} \quad \text{and } V_0 = E(X_0^2).$$

- **For the $|X|$, X^2- sequences the same theory applies with the corresponding moments, covariances and correlations adjusted.**

- **For example, if $EX_t^8 < \infty$ and (X_t) is GARCH, the CLT applies to the sample ACF ρ_{n,X^2}.**

- **Since $EX_t^8 = \infty$ or even $EX_t^4 = \infty$ are not unusual for return data, the question arises as to the weak limits of the sample ACF in this case.**

- **If (X_t) GARCH and $EX_t^4 = \infty$ or (X_t) SV and $EX_t^2 = \infty$: standard central limit theory for the sample ACFs $\rho_{n,X}$ and $\rho_{n,|X|}$ breaks down.**
- **(X_t) GARCH and $EX_t^8 = \infty$ or (X_t) SV and $EX_t^4 = \infty$: standard central limit theory for the sample ACF ρ_{n,X^2} breaks down.**
- **Then regular variation of the sequence (X_t) leads to unfamiliar limits of the sample ACFs involving infinite variance stable distributions.**
- **For the GARCH process, these asymptotic results yield confidence bands for the sample ACFs much wider than**

prescribed by the normal central limit theorem. Davis, M. (1998), M., Stărică (2000), Basrak, Davis, M. (2002b).

- **For the SV model, the sample ACF approaches zero at a rate much faster than $1/\sqrt{n}$.** Davis, M. (2001a,b) **This is same as for an iid regularly varying sequence with infinite variance.**

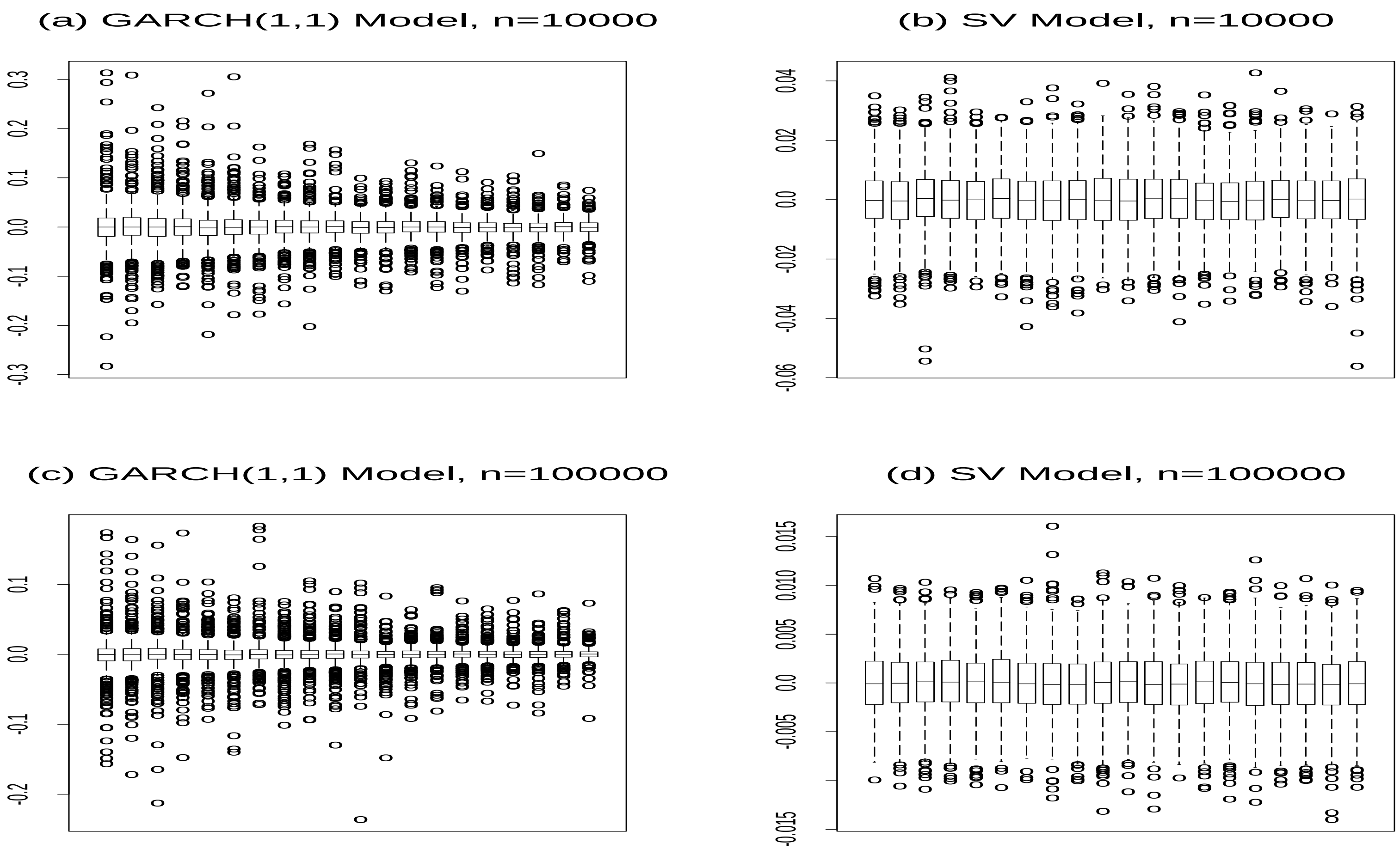

Figure 10. Boxplot comparison of the asymptotic behavior of the sample ACF for a GARCH *(left)* and a stochastic volatility model *(right)*. The parameters and noise distributions are chosen in such a way that both time series have tail index **3** which is not untypical for return series.

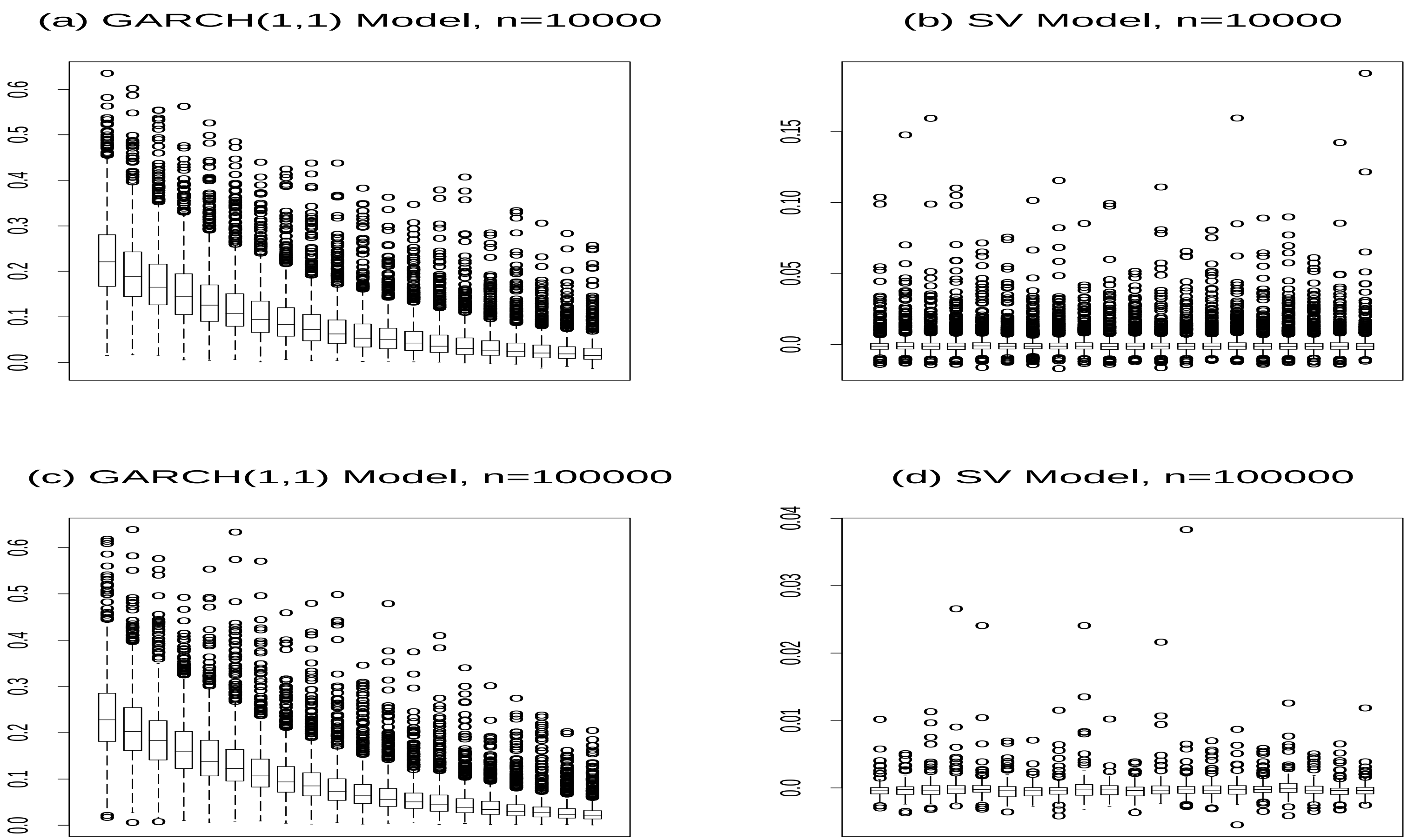

Figure 11. Boxplot comparison of the asymptotic behavior of sample ACF for the squares of the GARCH and the stochastic volatility models. Since the tail index of the model is **3** in both cases, it is **1.5** for the squared time series.

References

[1] Basrak, B., Davis, R.A. and Mikosch, T. (2002a) A characterization of multivariate regular variation. *Ann. Appl. Probab.* **12**, 908–920.

[2] Basrak, B., Davis, R.A. and Mikosch. T. (2002b) Regular variation of GARCH processes. *Stoch. Proc. Appl.* **99**, 95–116 .

[3] Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31**, 307–327.

[4] Bougerol, P. and Picard, N. (1992a) Strict stationarity of generalized autoregressive processes. *Ann. Probab.* **20**, 1714–1730.

[5] Bougerol, P. and Picard, N. (1992b) Stationarity of GARCH processes and of some nonnegative time series. *J. Econometrics* **52**, 115–127.

[6] Breidt, F.J. and Davis, R.A. (1998) Extremes of stochastic volatility models. *Ann. Appl. Probab* **8**, 664–675.

[7] Breiman, L. (1965) On some limit theorems similar to the arc-sin law. *Theory Probab. Appl.* **10**, 323–331.

[8] Brockwell, P.J. and Davis, R.A. (1991) *Time Series: Theory and Methods,* 2nd edition. Springer, New York.

[9] Brockwell, P.J. and Davis, R.A. (1996) *Introduction to Time Series and Forecasting.* Springer, New York.

[10] Davis, R.A. and Mikosch, T. (1998) The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Ann. Statist.* **26**, 2049–2080.

[11] Davis, R.A. and Mikosch, T. (2001a) Point process convergence of stochastic volatility processes with application to sample autocorrelations. *J. Appl. Probab.* Special Volume **38A**, A Festschrift for David Vere-Jones, 93–104.

[12] Davis, R.A. and Mikosch, T. (2001b) The sample autocorrelations of financial time series models. In: Fitzgerald, W.J., Smith, R.L., Walden, A.T. and Young, P.C. (Eds.) *Nonlinear and Nonstationary Signal Processing,* pp. 247–274. Cambridge University Press, Cambridge (U.K.).

[13] Davis, R.A. and Mikosch, T. (2009a) Probabilistic properties of stochastic volatility models. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.) *The Handbook of Financial Time Series*. Springer, Heidelberg, pp. 255–268.

[14] Davis, R.A. and Mikosch, T. (2009b) Extreme value theory for GARCH processes. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.) *The Handbook of Financial Time Series*. Springer, Heidelberg, pp. 187–200.

[15] Davis, R.A. and Mikosch, T. (2009c) Extremes of stochastic volatility models. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.) *The Handbook of Financial Time Series*. Springer, Heidelberg, pp. 355–364.

[16] Doukhan, P., Oppenheim, G. and Taqqu, M.S. (Eds.) (2003) *Long Range Dependence*. Birkhäuser, Boston.

[17] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.

[18] Engle, R.F. (1982) Autoregressive conditional heteroscedastic models with estimates of the variance of United Kingdom inflation. *Econometrica* **50**, 987–1007.

[19] Engle, R.F. and Bollerslev, T. (1986) Modelling the persistence of conditional variances. With comments and a reply by the authors. *Econometric Rev.* **5**, 1–87

[20] Goldie, C.M. (1991) Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1**, 126–166.

[21] Kesten, H. (1973) Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131**, 207–248.

[22] Mikosch, T. (2003) Modelling dependence and tails of financial time series. In: B. Finkenstädt and H. Rootzén (Eds.) *Extreme Values in Finance, Telecommunications and the Environment*. Chapman and Hall, pp. 185–286.

[23] Mikosch, T. and Stărică, C. (2000) Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. *Ann. Statist.* **28**, 1427–1451. An extended version is available under www.math.ku.dk/∼mikosch.

[24] NELSON, D.B. (1990) Stationarity and persistence in the GARCH$(\mathbf{1,1})$ model. *Econometric Theory* **6**, 318–334.

[25] SAMORODNITSKY, G. AND TAQQU, M.S. (1994) *Stable Non–Gaussian Random Processes. Stochastic Models with Infinite Variance.* Chapman and Hall, London.

[26] TAYLOR, S.J. (1986) *Modelling Financial Time Series.* Wiley, Chichester.