Risk-parameter estimation in volatility models

Christian Francq Jean-Michel Zakoïan

CREST and Université Lille 3

JSTAR 2012, October 25-26

Model risk/Estimation risk

Risk assessment framework defined by "Pillar II" directives: panel of risks including market risk.

In July 2009, the Basel Committee issued a directive requiring that financial institutions quantify "model risk":

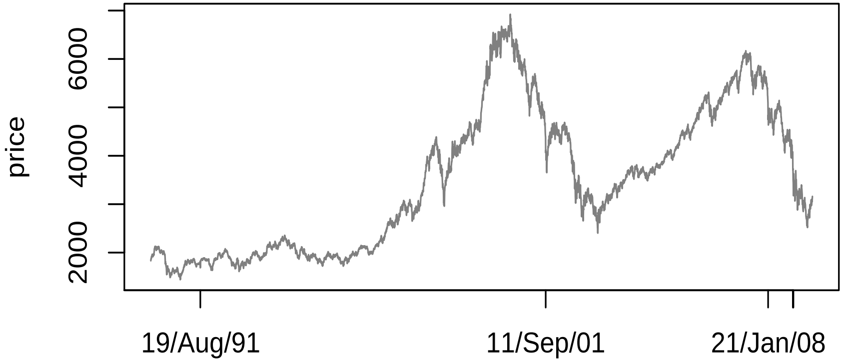
"Banks must explicitly assess the need for valuation adjustments to reflect two forms of model risk: the model risk associated with using a possibly incorrect valuation methodology; and the risk associated with using unobservable (and possibly incorrect) calibration parameters in the valuation model."

This talk is about quantifying the estimation risk in some dynamic models.

Outline

Stylized Facts (Mandelbrot (1963))

Non stationarity of the prices

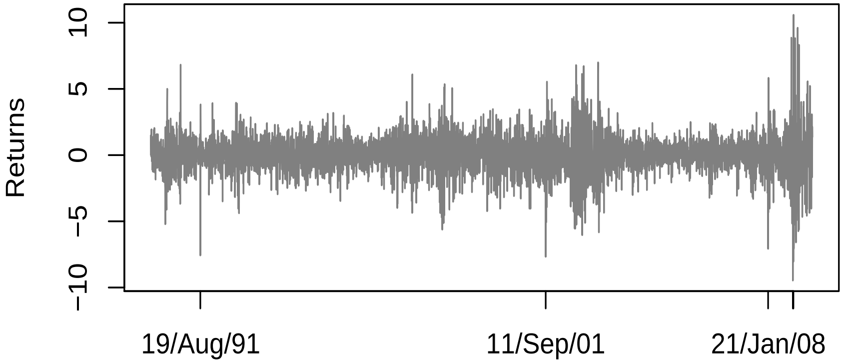


CAC 40, from March 1, 1992 to April 30, 2009

SP 500

Stylized Facts

Possible stationarity, unpredictability and volatility clustering of the returns

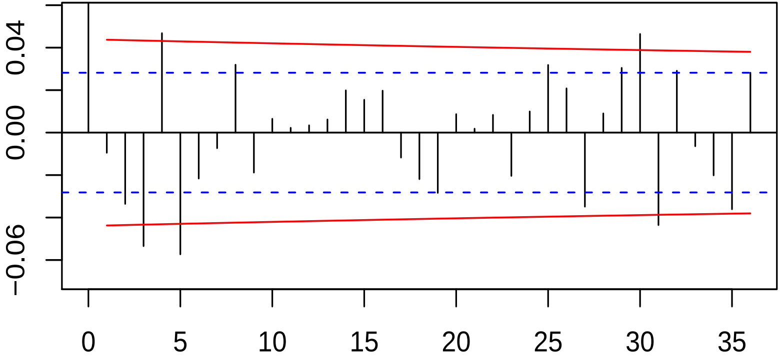


CAC 40 returns, from March 2, 1990 to February 20, 2009

SP 500 | zoom CAC40 | zoom SP500

Stylized Facts

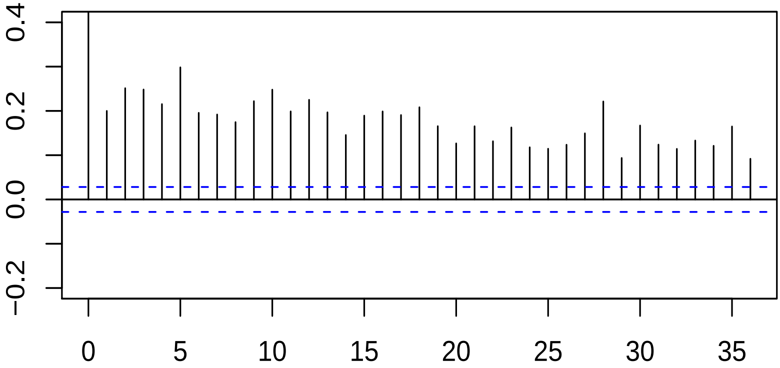
Dependence without correlation (warning: interpretation of the dotted lines)



Empirical autocorrelations of the CAC returns

SP 500
Other indices

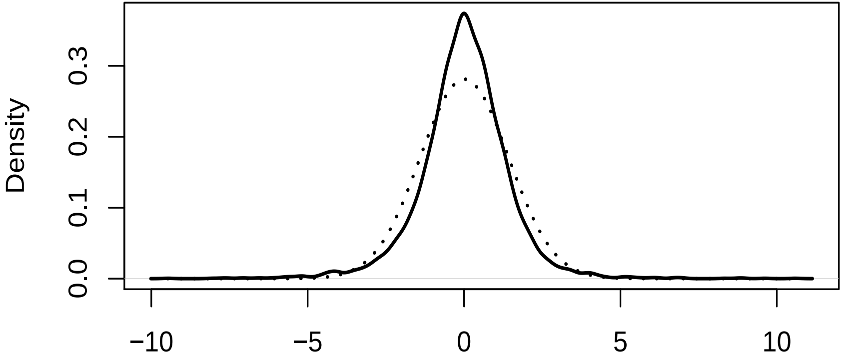
Stylized Facts
Correlation of the squares



Autocorrelations of the squares of the CAC returns

SP 500

Stylized Facts
Tail heaviness of the distributions



Density estimator for the CAC returns (normal in dotted line)

SP 500

Main properties of daily stock returns

- Unpredictability of the returns (martingale difference assumption), but non-independence.
- Strong positive autocorrelations of the squares or of the absolute values (even for large lags).
- Volatility clustering.
- Leptokurticity of the marginal distribution.
- Decreases of prices have an higher impact on the future volatility than increases of the same magnitude (leverage effects).

▸ Leverage effects

- Seasonalities.

Volatility Models

Almost all the volatility models are of the form

$$\epsilon_t = \sigma_t \eta_t$$

where (η_t) is iid (0,1), $\sigma_t > 0$, σ_t and η_t are independent.

For GARCH-type (Generalized Autoregressive Conditional Heteroskedasticity) models, $\sigma_t \in \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots)$.

See Bollerslev (Glossary to ARCH (GARCH), 2009) for an impressive list of more than one hundred GARCH-type models.

Definition: GARCH(p, q)

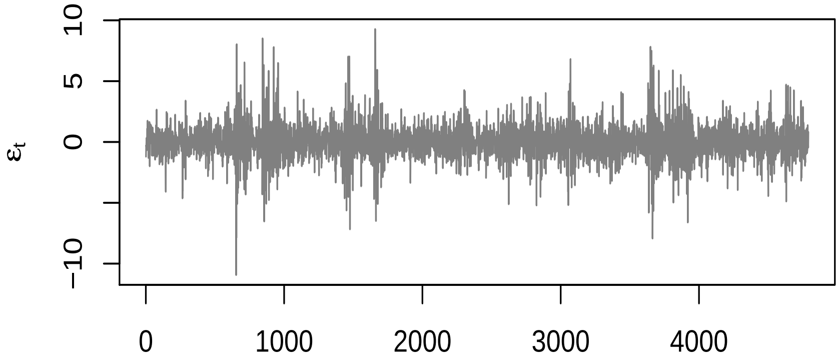
Definition (Engle (1982), Bollerslev (1986))

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega_0 + \sum_{i=1}^{q} \alpha_{0i} \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_{0j} \sigma_{t-j}^2, \quad \forall t \in \mathbb{Z} \end{cases}$$

where

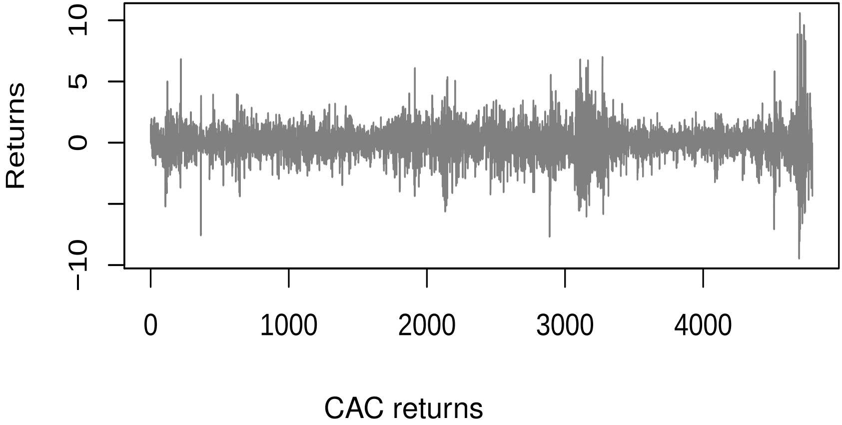
(η_t) iid, $\quad E\eta_1 = 0, \quad E\eta_1^2 = 1, \quad \omega_0 > 0, \quad \alpha_{0i} \geq 0, \quad \beta_{0j} \geq 0.$

GARCH(1,1) simulation

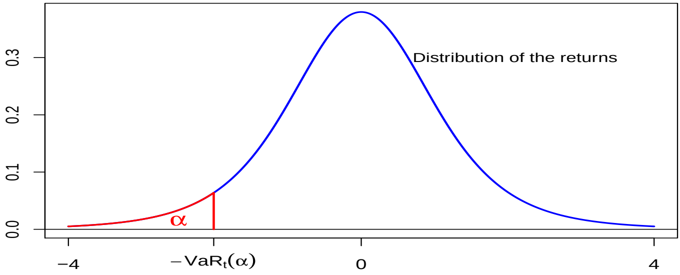


$$\epsilon_t = \sigma_t \eta_t,\ \eta_t \text{ iid } St_5,\ \sigma_t^2 = 0.033 + 0.090\epsilon_{t-1}^2 + 0.893\sigma_{t-1}^2,$$
$$t = 1, \dots, n = 4791$$

... resembles real financial series



Value at Risk and other risk measures



- Other risk measures, for instance

$$ES_t(\alpha) = \alpha^{-1} \int_0^{\alpha} \text{VaR}_t(u)du.$$

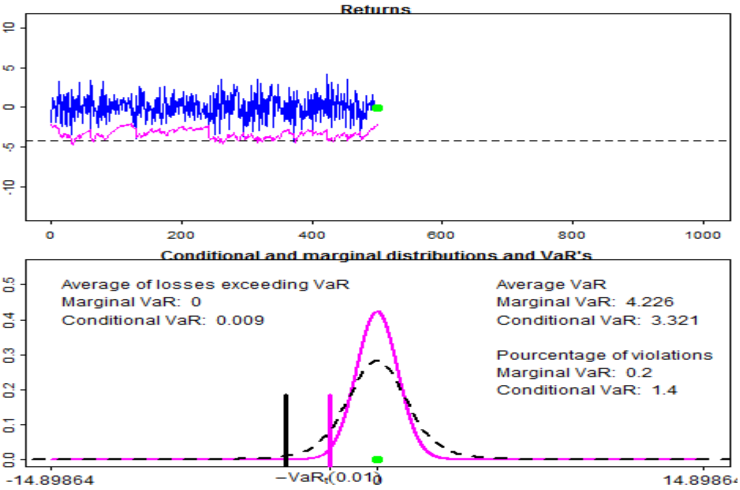
- Conditional versus marginal distribution distribution.

Conditional risk

Modern financial risk management focuses on risk measures based on distributional information.

- Traditional approaches:
  - marginal distributions of (log) returns
  - risk $=$ a parameter
- More sophisticated approaches:
  - conditional distributions of (log) returns
  - risk $=$ a stochastic process

Conditional VaR for a simulated process



A general conditional volatility model

$$\begin{cases} \epsilon_t = \sigma_t(\theta_0)\eta_t, \\ \\ \sigma_t(\theta_0) = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \theta_0) > 0, \end{cases}$$

- $\theta_0 \in \mathbb{R}^m$ is a parameter and $\sigma_t(\theta_0)$ is the volatility;
- (η_t) is a sequence of iid r.v. with $\eta_t \perp \epsilon_{t-j}, j > 0$.

Examples

No specification of the distribution of η_t
(semi-parametric model)

For the (statistical) identification of the "volatility parameter" θ_0, an assumption is needed.

On the role of an identifiability assumption on η_t

For any constant $K > 0$,

$$\epsilon_t = \underbrace{K\sigma_t(\theta_0)}_{\sigma_t(\theta_0^*)} \times \underbrace{K^{-1}\eta_t}_{\eta_t^*}$$

$\rightarrow$ a moment, a quantile, or another characteristic of the distribution of η_t must be fixed.

Standard identifiability assumption: $E\eta_1^2 = 1$.

Under this condition and $E\eta_1 = 0$, the volatility $\sigma_t^2(\theta_0)$ is the conditional variance of ϵ_t.

Estimators of the GARCH parameters

Vast literature on the estimation of the volatility parameter (under $E\eta_1^2 = 1$).

The most widely used method is the QML (Quasi-maximum likelihood):

- asymptotic theory valid under mild assumptions (strict stationarity but no moments of the observed process);
- does not require to know the distribution of η_t.

Principle of the Gaussian QMLE

Under $E\eta_t^2 = 1$, the Gaussian QML criterion (to be minimized)

$$\frac{1}{n}\sum_{t=1}^{n} \log \sigma_t^2(\theta) + \frac{\epsilon_t^2}{\sigma_t^2(\theta)}$$

gives a consistent estimator because the limit criterion

$$E\left(\log \sigma_t^2(\theta) + \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)}\eta_t^2\right) = E\left(\log \sigma_t^2(\theta) + \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)}\right)$$

is uniquely minimized at θ_0
(assuming $\sigma_t^2(\theta) = \sigma_t^2(\theta_0) \quad \Rightarrow \quad \theta = \theta_0$ + regularity conditions).

Some references on QML estimation for GARCH

- **ARCH(q) or GARCH(1,1):** Weiss (Ec. Theory, 1986), Lee and Hansen (Ec. Theory, 1994), Lumsdaine (Econometrica, 1996): CAN under moment assumptions on (ϵ_t), or a density for η_t.
- **Standard GARCH(p, q):**
  - Berkes et al. (Bernoulli, 2003), F&Z (Bernoulli, 2004): Consistency and AN under (mainly) the strict stationarity of (ϵ_t) and $E\eta_t^4 < \infty$.
  - Berkes and Horváth (AOS, 2003 and 2004) ML and non-Gaussian QML under different identifiability assumptions.
  - Hall and Yao (Econometrica, 2003): Asymptotic distribution of the QMLE when $E\eta_t^4 = \infty$ and $E\epsilon_t^2 < \infty$.
  - F&Z (SPA, 2007): Asymptotic distribution of the QMLE when θ_0 has null coefficients.

Some references on QML estimation for GARCH

- **ARMA-GARCH:** Ling and Li(JASA, 1997), F&Z (Bernoulli, 2004), Ling (J. of Econometrics, 2007):
  Consistency and AN of the QMLE under $E\eta_t = 0$ and $E\epsilon_t^4 < \infty$.
  Self-weighted QMLE to avoid the moment condition.
- **More general stationary GARCH models:** Straumann and Mikosch (AOS, 2006), Robinson and Zaffaroni (AOS, 2006), Bardet and Wintenberger (AOS, 2009), Meitz and Saikkonen (Ec. Theory, 2011):
  Non linear or long-memory GARCH models.
- **Explosive GARCH(1,1):** Jensen and Rahbek (Econometrica, 2004 and Ec. Theory, 2004), F&Z (Econometrica, 2012).
  CAN of the QMLE (except ω) when θ_0 is outside the strict stationarity region.

Conditional risk measures

Consider a risk measure, r, that is, a mapping from the set of the real random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ to $\mathbb{R}$.

Assume that r is nonnegative, and

- Positively homogenous: $r(\lambda X) = \lambda r(X)$ for any variable X and any $\lambda \geq 0$,
- Law invariant: $r(X) = r(Y)$ if X and Y have the same distribution.

The conditional risk of $\epsilon_t = \sigma_t(\theta_0)\eta_t$ is given by

$$r_{t-1}(\epsilon_t) = \sigma_t(\theta_0) r(\eta_1).$$

Example: conditional VaR

The conditional VaR of the process (ϵ_t) at risk level $\alpha \in (0,1)$, denoted by $\text{VaR}_t(\alpha)$, is defined, in the continuous case, by

$$P_{t-1}[\epsilon_t < -\text{VaR}_t(\alpha)] = \alpha,$$

where P_{t-1} denotes the historical distribution conditional on $\{\epsilon_u, u < t\}$.

For the conditional volatility model, the conditional VaR is

$$\text{VaR}_t(\alpha) = -\sigma(\epsilon_{t-1}, \epsilon_{t-2}, \ldots; \theta_0) F_\eta^{-1}(\alpha)$$

where F_η is the c.d.f. of η_t.

Assumption on the volatility function

The goal is to define a risk parameter (for a given risk r), similar to the volatility parameter.

A0: There exists a function H such that for any $\theta \in \mathbb{R}^m$, for any $K > 0$, and any sequence $(x_i)_i$

$$K\sigma(x_1, x_2, \dots; \theta) = \sigma(x_1, x_2, \dots; \theta^*), \quad \text{where } \theta^* = H(\theta, K).$$

For instance, in the GARCH(1,1) case $\theta^* = (K^2\omega, K^2\alpha, \beta)'$.

Conditional risk parameter

We have $r_{t-1}(\epsilon_t) = \sigma_t(\theta_0) r(\eta_1)$.

If $r(\eta_1) > 0$, let $\eta_t^* = \eta_t / r(\eta_1)$ and let $\theta_0^* = H(\theta_0, r(\eta_1))$.

Under **A0**, the model can be reparameterized as

$$\begin{cases} \epsilon_t = \sigma_t^* \eta_t^*, & r(\eta_1^*) = 1, \\ \sigma_t^* = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \ldots; \theta_0^*). \end{cases}$$

Because the conditional risk of ϵ_t is now simply

$$r_{t-1}(\epsilon_t) = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \ldots; \theta_0^*),$$

θ_0^* will be called the risk parameter.

Example: Conditional VaR for a GARCH(1,1)

GARCH(1,1) Model:

$$\begin{cases} \epsilon_t = \sigma_t(\theta_0)\eta_t, \\ \sigma_t^2(\theta_0) = \omega_0 + \alpha_0\epsilon_{t-1}^2 + \beta_0\sigma_{t-1}^2(\theta_0) \end{cases}$$

with $\theta_0 = (\omega_0, \alpha_0, \beta_0) \in (0, \infty) \times (\mathbb{R}^+)^2$ and $E\eta_1^2 = 1$.

$\text{VaR}_t(\alpha) = -\sigma_t(\theta_0)F_\eta^{-1}(\alpha).$

VaR parameter at level α (with $F_\eta^{-1}(\alpha) < 0$):

$$\theta_0^* = (K^2\omega_0, K^2\alpha_0, \beta_0)'$$

with $K = -F_\eta^{-1}(\alpha)$.

This coefficient takes into account the dynamics of the GARCH process, but also the lower tail of the innovations distribution.

Example: Conditional VaR for a GARCH(1,1)

Numerical illustration:

$$\begin{cases} \epsilon_t = \sigma_t \eta_t, & \eta_t \sim \mathcal{N}(0,1) \\ \sigma_t^2 = 1 + 0.05\epsilon_{t-1}^2 + 0.9\sigma_{t-1}^2 \end{cases} \quad \text{and} \quad \begin{cases} \epsilon_t = \sigma_t \eta_t, & \eta_t \sim \frac{1}{\sqrt{2}} St_4 \\ \sigma_t^2 = 1 + 0.04\epsilon_{t-1}^2 + 0.9\sigma_{t-1}^2. \end{cases}$$

The volatility parameter of the Gaussian model is larger than that of the Student-innovation model.

Now consider the VaR's at level 1%.

The risk parameter of the first model is $\theta_0^* = (5.41, 0.27, 0.9)$, whereas that of the second model is $\theta_0^* = (7.01, 0.28, 0.9)$.

The first model is more volatile but less risky than the second one for the VaR at 1%.

Two strategies for estimating the conditional risk parameter

Based on two formulations of the conditional risk:

$$r_{t-1}(\epsilon_t) = \begin{cases} \sigma_t(\theta_0) r(\eta_1), & \text{with } E\eta_1^2 = 1, \\ \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \theta_0^*), & \text{with } r(\eta_1^*) = 1. \end{cases}$$

1. Standard Gaussian QML estimation + nonparametric estimation of $r(\eta_1)$.
2. Non Gaussian QML estimation under the identifiability assumption $r(\eta_1^*) = 1$.

Non Gaussian QML estimator under $r(\eta_1^*) = 1$.

Given observations $\epsilon_1, \dots, \epsilon_n$, and arbitrary initial values $\tilde{\epsilon}_i$ for $i \leq 0$, let

$$\tilde{\sigma}_t(\theta) = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_1, \tilde{\epsilon}_0, \tilde{\epsilon}_{-1}, \dots; \theta).$$

This random variable will be used to approximate

$$\sigma_t(\theta) = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_1, \epsilon_0, \epsilon_{-1}, \dots; \theta).$$

We choose an arbitrary, *instrumental*, positive density h, and we define the QML criterion

$$\tilde{Q}_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} g(\epsilon_t, \tilde{\sigma}_t(\theta)), \qquad g(x, \sigma) = \log \frac{1}{\sigma} h\left(\frac{x}{\sigma}\right).$$

Let the QMLE, for some compact space $\Theta \subset \mathbb{R}^m$,

$$\hat{\theta}_n^* = \arg\max_{\theta \in \Theta} \tilde{Q}_n(\theta).$$

Technical assumptions for the consistency:

A1: (ϵ_t) is a strictly stationary and ergodic solution of the model.

A2: Almost surely, $\sigma_t(\theta) \in (\underline{\omega}, \infty]$ for any $\theta \in \Theta$ and for some $\underline{\omega} > 0$. Moreover, $\sigma_t(\theta_0^*)/\sigma_t(\theta) = 1 \; a.s.$ iff $\theta = \theta_0^*$.

A3: $Eg(\eta_1^*, \sigma) < Eg(\eta_1^*, 1), \qquad \forall \sigma > 0, \quad \sigma \neq 1.$ Interpretation of A3

A4: h is continuous on $\mathbb{R}$, differentiable except on a finite set A, and there exist constants $\delta \geq 0$ and $C_0 > 0$ such that for all $u \in A^c$, $|uh'(u)/h(u)| \leq C_0(1 + |u|^\delta)$ with $E|\epsilon_0|^\delta < \infty$.

A5: $\sup_{\theta \in \Theta} |\sigma_t(\theta) - \tilde{\sigma}_t(\theta)| \leq C_1 \rho^t, \quad$ where $\rho \in (0, 1)$.

Consistency of the risk parameter estimator

Consistency

If **A0-A5** hold, the non-Gaussian QML estimator satisfies

$$\hat{\theta}_n^* \to \theta_0^*, \quad a.s.$$

Remark: the innovation distribution is subject to two conditions

$$r(\eta_1^*) = 1 \quad \text{and} \quad Eg(\eta_1^*, \sigma) < Eg(\eta_1^*, 1).$$

Can we find a density h making them compatible?

Choice of the QML density h

Assume that, for some measurable function $\psi : \mathbb{R} \to \mathbb{R}$,

$$r(\eta) = 1 \quad \text{iff} \quad E\{\psi(\eta)\} = 0.$$

More explicit condition on h

Assume **A4** holds with $A = \emptyset$. Then **A3** holds for any distribution of η_1^* satisfying $r(\eta_1^*) = 1$ iff the density h is such that

$$x\{\log h(x)\}' = \lambda\psi(x) - 1, \quad \text{for all } x,$$

for some constant $\lambda \neq 0$.

Examples

- $r(\eta) = \|\eta\|_s = (E|\eta|^s)^{1/s}, \quad s > 0.$

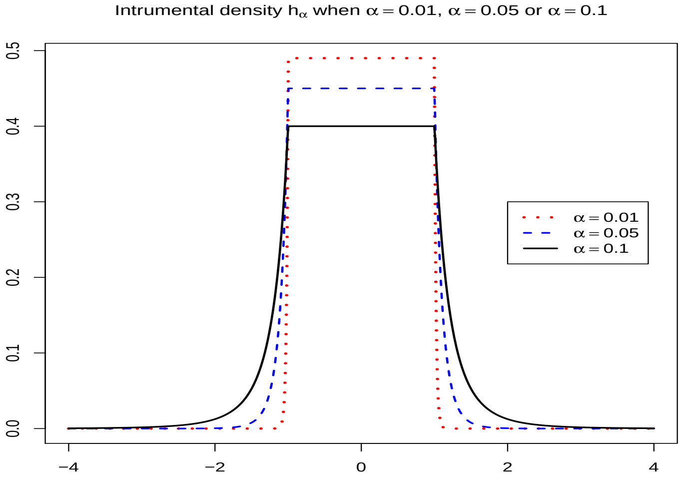
  We have $\psi(\eta) = |\eta|^s - 1$ and we find

$$h(x) \; \propto \; x^{-(1-\lambda)} \exp(-\lambda |x|^s / s), \quad \forall \lambda > 0.$$

- VaR at level α: $\quad r(\eta) = -F_\eta^{-1}(\alpha)$.

  If $\mathbb{P}^\eta = \mathbb{P}^{-\eta}$ and $\alpha \in (0, 0.5)$, $\psi(\eta) = \mathbf{1}_{\{|\eta|>1\}} - 2\alpha$, we find

$$h_\alpha(x) \; \propto \; |x|^{2\lambda\alpha - 1} \{ |x|^{-\lambda} \mathbf{1}_{\{|x|>1\}} + \mathbf{1}_{\{|x| \le 1\}} \}, \quad \forall \lambda > 0.$$


Intrumental density hα when α = 0.01, α = 0.05 or α = 0.1
0.0
0.1
0.2
0.3
0.4
0.5
−4
−2
0
2
4
α = 0.01
α = 0.05
α = 0.1

Additional assumptions for the asymptotic normality

A6: $\theta_0^* \in \overset{\circ}{\Theta}$.

A7: $x' \frac{\partial \sigma_t(\theta_0^*)}{\partial \theta} = 0, \quad a.s. \quad \Rightarrow \quad x = 0.$

A8: The function $\theta \mapsto \sigma(x_1, x_2, \dots; \theta)$ has continuous second-order derivatives, and for C_1, ρ as in **A5**,

$$\sup_{\theta \in \Theta} \left\| \frac{\partial \sigma_t(\theta)}{\partial \theta} - \frac{\partial \tilde{\sigma}_t(\theta)}{\partial \theta} \right\| + \left\| \frac{\partial^2 \sigma_t(\theta)}{\partial \theta \partial \theta'} - \frac{\partial^2 \tilde{\sigma}_t(\theta)}{\partial \theta \partial \theta'} \right\| \leq C_1 \rho^t.$$

A9: h is twice differentiable with $\left|u^2 \left(h'(u)/h(u)\right)'\right| \leq C_0(1 + |u|^\delta)$ for all $u \in \mathbb{R}$ and $E|\epsilon_1|^{2\delta} < \infty$.

A10: There exists a neighborhood $V(\theta_0^*)$ of θ_0^* such that

$$E \sup_{\theta \in V(\theta_0^*)} \left\| \frac{1}{\sigma_t(\theta)} \frac{\partial^2 \sigma_t(\theta)}{\partial \theta \partial \theta'} \right\|^2 < \infty.$$

Asymptotic normality of the risk parameter estimator

Let $g_1(x,\sigma) = \partial g(x,\sigma)/\partial\sigma$ and $g_2(x,\sigma) = \partial g_1(x,\sigma)/\partial\sigma$.

Asymptotic normality

Under **A0-A10** and if $Eg_2(\eta_1^*, 1) \neq 0$,

$$\sqrt{n}\left(\hat{\theta}_n^* - \theta_0^*\right) \xrightarrow{d} \mathcal{N}(0, 4\tau_{h,f}^2 I^{-1})$$

where

$$I = I(\theta_0^*) = E\left(\frac{1}{\sigma_t^4}\frac{\partial\sigma_t^2}{\partial\theta}\frac{\partial\sigma_t^2}{\partial\theta'}(\theta_0^*)\right) \quad \text{and} \quad \tau_{h,f}^2 = \frac{Eg_1^2(\eta_1^*, 1)}{\left\{Eg_2(\eta_1^*, 1)\right\}^2}.$$

But this does not apply to the VaR (**A9** not satisfied).

Definition of the VaR parameter

Model reparameterization:

$$\begin{cases} \epsilon_t = \sigma_t^* \eta_t^*, \quad P[\eta_t^* < -1] = \alpha, \\ \sigma_t^* = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \ldots ; \theta_{0,\alpha}). \end{cases}$$

where

- $\eta_t^* = -\eta_t / F^{-1}(\alpha)$ (provided $F^{-1}(\alpha) < 0$)
- $\theta_{0,\alpha} = \theta_0^* = H(\theta_0, -F^{-1}(\alpha))$: the *VaR parameter* at level α.

The theoretical VaR is now given by

$$\text{VaR}_t(\alpha) = \sigma_t^*.$$

Definition of the VaR parameter estimator

QML estimator of $\theta_{0,\alpha}$:

$$\hat{\theta}_{n,\alpha} = \arg\max_{\theta\in\Theta} \sum_{t=1}^{n} \log \frac{1}{\tilde{\sigma}_t(\theta)} h_\alpha\left(\frac{\epsilon_t}{\tilde{\sigma}_t(\theta)}\right)$$

where

$$h_\alpha(x) = \frac{1}{2}(1-2\alpha)\{|x|^{-\frac{1}{2\alpha}}\mathbf{1}_{\{|x|>1\}} + \mathbf{1}_{\{|x|\leq 1\}}\}$$

Interpretation as a quantile regression estimator

If the distribution of η_1^* is symmetric, we have

$$\log|\epsilon_t| = \log\sigma_t^* + \log|\eta_t^*|, \qquad P[\log|\eta_1^*| < 0] = 1 - 2\alpha,$$

Let $\rho_\alpha(u) = u(\alpha - \mathbf{1}_{\{u\leq 0\}})$.

Then

$$\begin{aligned}\hat{\theta}_{n,\alpha} &= \arg\min_{\theta\in\Theta} \frac{1}{n}\sum_{t=1}^{n}\rho_{1-2\alpha}\left\{\log\left(\frac{|\epsilon_t|}{\tilde{\sigma}_t(\theta)}\right)\right\} \\ &= \arg\min_{\theta\in\Theta} \frac{1}{n}\sum_{t=1}^{n}\left|\log\left(\frac{|\epsilon_t|}{\tilde{\sigma}_t(\theta)}\right)\right| \\ &\qquad \times\left((1-2\alpha)\mathbf{1}_{\{|\epsilon_t|>\tilde{\sigma}_t(\theta)\}} + 2\alpha\mathbf{1}_{\{|\epsilon_t|<\tilde{\sigma}_t(\theta)\}}\right).\end{aligned}$$

Consistency and A.N. of the VaR parameter estimator

B1: The law of η_1^* is symmetric, admits a density in a neighborhood of 1 and satisfies $E|\log|\eta_1^*|| < \infty$.

If **A0-A2**, **A5** and **B1** hold, for all $\alpha \in (0, 1/2)$,

$$\hat{\theta}_{n,\alpha} \to \theta_{0,\alpha}, \quad a.s.$$

Under additional technical assumptions, there exists a sequence of local minimizers $\hat{\theta}_{n,\alpha}$ of the criterion satisfying

$$\sqrt{n}(\hat{\theta}_{n,\alpha} - \theta_{0,\alpha}) \xrightarrow{d} \mathcal{N}\left(0, \Xi_\alpha := \frac{2\alpha(1-2\alpha)}{4f^{*2}(1)} J_\alpha^{-1}\right)$$

where $J_\alpha = ED_t(\theta_{0,\alpha})D_t'(\theta_{0,\alpha})$ and $D_t(\theta) = \sigma_t^{-1}(\theta)\partial\sigma_t(\theta)/\partial\theta$.

Confidence intervals for the VaR

One-step consistent estimator of the VaR parameter:

$$\widehat{\text{VaR}}_t(\alpha) = \tilde{\sigma}_t(\hat{\theta}_{n,\alpha}).$$

Asymptotic confidence interval for $\text{VaR}_t(\alpha)$ (at the level $(1-\alpha_0)\%$):

$$\tilde{\sigma}_t(\hat{\theta}_{n,\alpha}) \pm \frac{\Phi^{-1}_{1-\alpha_0}}{\sqrt{n}} \left\{ \frac{\partial \tilde{\sigma}_t(\hat{\theta}_{n,\alpha})}{\partial \theta'} \widehat{\Xi}_\alpha \frac{\partial \tilde{\sigma}_t(\hat{\theta}_{n,\alpha})}{\partial \theta} \right\}^{1/2},$$

Interpretation: The VaR evaluation is subject to estimation risk. Even when the model is correctly specified, the market risk, as measured by the theoretical VaR, is not perfectly known, but is likely to belong to the confidence interval.

Confidence intervals for the VaR

One-step consistent estimator of the VaR parameter:

$$\widehat{\text{VaR}}_t(\alpha) = \tilde{\sigma}_t(\hat{\theta}_{n,\alpha}).$$

Asymptotic confidence interval for $\text{VaR}_t(\alpha)$
(at the level $(1-\alpha_0)\%$):

$$\tilde{\sigma}_t(\hat{\theta}_{n,\alpha}) \pm \frac{\Phi^{-1}_{1-\alpha_0}}{\sqrt{n}} \left\{ \frac{\partial \tilde{\sigma}_t(\hat{\theta}_{n,\alpha})}{\partial \theta'} \widehat{\Xi}_\alpha \frac{\partial \tilde{\sigma}_t(\hat{\theta}_{n,\alpha})}{\partial \theta} \right\}^{1/2},$$

Interpretation: The VaR evaluation is subject to estimation risk. Even when the model is correctly specified, the market risk, as measured by the theoretical VaR, is not perfectly known, but is likely to belong to the confidence interval.

Two-step estimator

$$\text{VaR}_t(\alpha) = -\sigma_t(\theta_0)F_\eta^{-1}(\alpha).$$

Under the usual condition $E\eta_t^2 = 1$, and $E\eta_t^4 < \infty$,

- θ_0 is estimated by standard QML (estimator $\hat{\theta}_n$);
- the theoretical quantile $\xi_\alpha := F_\eta^{-1}(\alpha)$ is estimated using the estimated rescaled innovations:

$$\hat{\eta}_t = \frac{\epsilon_t}{\tilde{\sigma}_t(\hat{\theta}_n)}.$$

Let $\xi_{n,\alpha}$ denote the empirical α-quantile of $\hat{\eta}_1, \dots, \hat{\eta}_n$.

An estimator of the VaR at level α is then given by

$$\widetilde{\text{VaR}}_t(\alpha) = -\tilde{\sigma}_t(\hat{\theta}_n)\xi_{n,\alpha}.$$

Comparing the one-step and two-step estimators

Estimators of the VaR: the one-step estimator

$$\widehat{\mathrm{VaR}}_t(\alpha) = \tilde{\sigma}_t(\hat{\theta}_{n,\alpha})$$

and the two-step estimators

$$\widetilde{\mathrm{VaR}}_t(\alpha) = -\tilde{\sigma}_t(\hat{\theta}_n)\xi_{n,\alpha} = \tilde{\sigma}_t\{H(\hat{\theta}_n, -\xi_{n,\alpha})\}.$$

and, under the assumption of symmetric errors distribution,

$$\widetilde{\widetilde{\mathrm{VaR}}}_t(\alpha) = \tilde{\sigma}_t(\hat{\theta}_n)\tilde{\xi}_{n,1-2\alpha} = \tilde{\sigma}_t\{H(\hat{\theta}_n, \tilde{\xi}_{n,1-2\alpha})\},$$

where $\tilde{\xi}_{n,1-2\alpha}$ is the empirical $(1-2\alpha)$-quantile of $|\hat{\eta}_1|, \dots, |\hat{\eta}_n|$.

A comparison of the VaR estimators can then be based on the asymptotic accuracies of the estimators of $\theta_{0,\alpha}$:

$$\hat{\theta}_{n,\alpha}, \quad \hat{\theta}_{n,\alpha}^{2step} := H(\hat{\theta}_n, -\xi_{n,\alpha}), \quad \hat{\theta}_{n,\alpha}^{S2step} := H(\hat{\theta}_n, \tilde{\xi}_{n,1-2\alpha}).$$

Asymptotic distribution of the two-step estimators

Requires deriving the joint asymptotic distributions of $(\hat{\theta}_n', -\xi_{n,\alpha})$ and $(\hat{\theta}_n', \tilde{\xi}_{n,1-2\alpha})$ under $E\eta_t^2 = 1$.

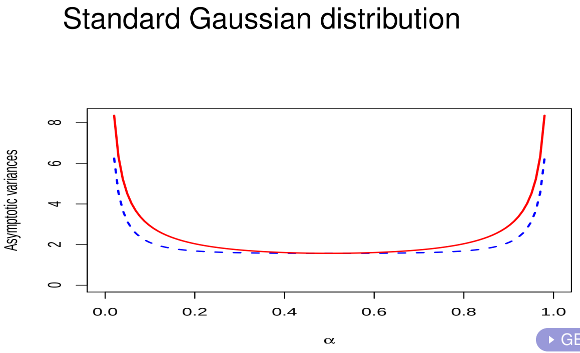
An additional assumption is needed: $\kappa_4 = E\eta_t^4 < \infty$.

We find that the asymptotic variance of $-\xi_{n,\alpha}$, the empirical quantile of $\hat{\eta}_t$'s, is

$$\zeta_\alpha = \underbrace{\frac{\alpha(1-\alpha)}{f^2(\xi_\alpha)}}_{\text{If no estimation}} + \underbrace{\frac{\xi_\alpha p_\alpha}{f(\xi_\alpha)} + \xi_\alpha^2 \frac{\kappa_4 - 1}{4}}_{\text{Effect of estimation}}.$$

where $\xi_\alpha = F_\eta^{-1}(\alpha)$ and $p_\alpha = E\left(\eta_1^2 \mathbf{1}_{\{\eta_1 < \xi_\alpha\}}\right) - \alpha$.

Asymptotic variances of empirical quantiles, with or without estimation (dotted and full lines)



GED example

Comparison of VaR estimators for standard GARCH with symmetric innovations

Under the previous assumptions (in particular $E\eta_t^4 < \infty$ for the two-step estimator),

$$\begin{aligned}\text{Var}_{as}\{\sqrt{n}(\hat{\theta}_{n,\alpha} - \theta_{0,\alpha})\} \quad &\preceq \quad \text{Var}_{as}\left\{\sqrt{n}\left(\hat{\theta}_{n,\alpha}^{S2tep} - \theta_{0,\alpha}\right)\right\} \\ \text{iff} \quad & \quad \Delta_\alpha \leq 0,\end{aligned}$$

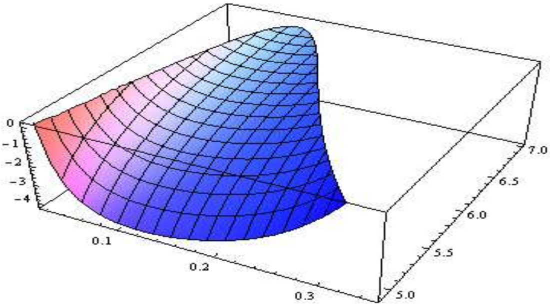
where $\Delta_\alpha = \frac{2\alpha(1-2\alpha)}{\xi_\alpha^2 f^2(\xi_\alpha)} - (\kappa_4 - 1)$.

Remark: the coefficient Δ_α only depends on the distribution of η_t, not on the true parameter value.

For fat-tailed distributions the one-step estimator will be better.

Surface $\Delta_\alpha \leq 0$ (1-step estimator better than 2-step)

Student(ν) distribution (standardized)
$\nu \in [4.9, 7]$ and $\alpha \in [0.01, 0.35]$



GED example

Simulation experiments

Table: Empirical relative efficiency of the 1-step method with respect to the 2-step method for estimating the VaR parameter. ARCH(1) model with Student innovations. Number of replication: $N = 1,000$.

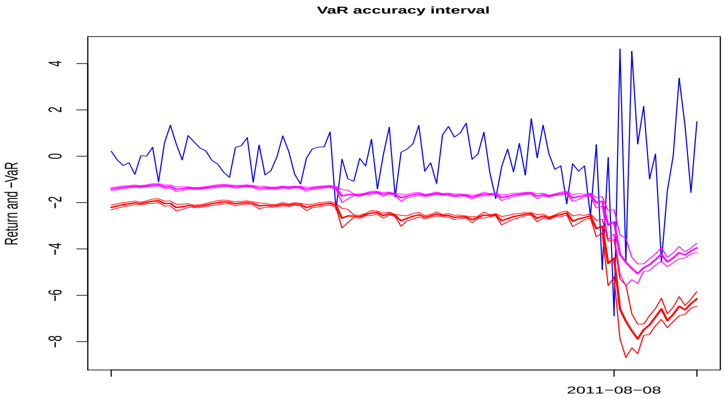
	$n = 500$							$n = 5,000$						
	ν							ν						
	1	2	3	4	5	6	∞	1	2	3	4	5	6	∞
$\alpha = 5\%$														
$\omega_{0,\alpha}$	7.5	2.8	1.7	1.3	1	0.9	0.9	13.9	6.6	2.7	1.3	1.1	1	0.8
$\alpha_{0,\alpha}$	7.3	3.6	1.7	1.3	1	1.0	0.8	22.2	8.7	3.2	1.3	1.1	1	0.9
$\alpha = 1\%$														
$\omega_{0,\alpha}$	6.1	1.6	1.0	0.8	0.7	0.7	0.7	41.1	3.6	1.6	0.9	0.8	0.8	0.7
$\alpha_{0,\alpha}$	3.8	1.8	2.6	0.8	0.7	0.7	0.7	13.7	6.0	2.1	0.9	0.8	0.7	0.7

Real data: January, 2, 1991 to August, 26, 2011

Table: Estimators of the VaR parameter $\theta_{0,\alpha}$ at level $\alpha = 5\%$ of GARCH(1,1) models. Estimations of Δ_α based on residuals of the 2-step and 1-step methods: $\Delta_\alpha < 0$ indicates superiority of the 1-step method.

		$\omega_{0,\alpha}$	$\alpha_{0,\alpha}$	$\beta_{0,\alpha}$	$\hat{\Delta}_{5\%}^{S2step}$	$\hat{\Delta}_{5\%}$
Nikkei	$\hat{\theta}_{n,5\%}^{S2step}$	0.08 (0.02)	0.33 (0.05)	0.87 (0.02)	-3.86	-4.54
	$\hat{\theta}_{n,5\%}$	0.04 (0.01)	0.30 (0.03)	0.88 (0.01)		
NSE	$\hat{\theta}_{n,5\%}^{S2step}$	0.16 (0.06)	0.26 (0.06)	0.87 (0.03)	-3.11	-3.30
	$\hat{\theta}_{n,5\%}$	0.18 (0.05)	0.31 (0.05)	0.85 (0.02)		
SP500	$\hat{\theta}_{n,5\%}^{S2step}$	0.02 (0.00)	0.20 (0.02)	0.92 (0.01)	-2.10	-2.31
	$\hat{\theta}_{n,5\%}$	0.02 (0.00)	0.19 (0.01)	0.92 (0.01)		
SPTSX	$\hat{\theta}_{n,5\%}^{S2step}$	0.02 (0.01)	0.17 (0.03)	0.93 (0.01)	-0.06	-0.52
	$\hat{\theta}_{n,5\%}$	0.04 (0.01)	0.23 (0.03)	0.90 (0.01)		

Estimated VaR's and VaR accuracy intervals



Log returns of the SP500 and estimated VaR's at the 5% and 1% levels, from April 6, 2011 to August 26, 2011. Estimation of the VaR parameter is based on the 500 previous values.

Conclusions

- Notion of conditional risk/VaR parameter.
- Facilitates the asymptotic comparison of risk evaluation procedures.
- Reparameterization allows for one-step estimation and easier evaluation of confidence intervals for the VaR.
- For standard GARCH models the ranking of the two methods only depends on the sign of the scalar Δ_α.
- This coefficient involves α and simple characteristics of the innovations distribution, and thus can be easily estimed.
- The one-step method is typically more efficient in presence of fat tailed innovations.

Examples

- Standard GARCH(p,q) (Engle (82), Bollerslev (86)):

$$\sigma_t^2 = \omega_0 + \sum_{i=1}^{q} \alpha_{0i}\epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_{0j}\sigma_{t-j}^2.$$

- Asymmetric Power GARCH model: for $\delta > 0$,

$$\sigma_t^\delta = \omega_0 + \sum_{i=1}^{q} \alpha_{0i+}(\epsilon_{t-i}^+)^\delta + \alpha_{0i-}(-\epsilon_{t-i}^-)^\delta + \sum_{j=1}^{p} \beta_{0j}\sigma_{t-j}^\delta$$

- ARCH(∞) (Robinson (91)), introduced to capture long memory:

$$\sigma_t^2 = \psi_{00} + \sum_{i=1}^{\infty} \psi_{0i}\epsilon_{t-i}^2$$

Return

Examples

- Standard GARCH(p,q) (Engle (82), Bollerslev (86)):

$$\sigma_t^2 = \omega_0 + \sum_{i=1}^{q} \alpha_{0i}\epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_{0j}\sigma_{t-j}^2.$$

- Asymmetric Power GARCH model: for $\delta > 0$,

$$\sigma_t^\delta = \omega_0 + \sum_{i=1}^{q} \alpha_{0i+}(\epsilon_{t-i}^+)^\delta + \alpha_{0i-}(-\epsilon_{t-i}^-)^\delta + \sum_{j=1}^{p} \beta_{0j}\sigma_{t-j}^\delta$$

- ARCH(∞) (Robinson (91)), introduced to capture long memory:

$$\sigma_t^2 = \psi_{00} + \sum_{i=1}^{\infty} \psi_{0i}\epsilon_{t-i}^2$$

Return

Examples

- Standard GARCH(p, q) (Engle (82), Bollerslev (86)):

$$\sigma_t^2 = \omega_0 + \sum_{i=1}^{q} \alpha_{0i}\epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_{0j}\sigma_{t-j}^2.$$

- Asymmetric Power GARCH model: for $\delta > 0$,

$$\sigma_t^\delta = \omega_0 + \sum_{i=1}^{q} \alpha_{0i+}(\epsilon_{t-i}^+)^\delta + \alpha_{0i-}(-\epsilon_{t-i}^-)^\delta + \sum_{j=1}^{p} \beta_{0j}\sigma_{t-j}^\delta$$

- ARCH(∞) (Robinson (91)), introduced to capture long memory:

$$\sigma_t^2 = \psi_{00} + \sum_{i=1}^{\infty} \psi_{0i}\epsilon_{t-i}^2$$

Return

Interpretation of the identifiability assumption

A3: $Eg(\eta_1^*, \sigma) < Eg(\eta_1^*, 1) \qquad \forall \sigma > 0, \quad \sigma \neq 1.$

If η_1^* has a density f, let $h_\sigma(x) = \sigma^{-1}h(\sigma^{-1}x)$ and the Kullback-Leibler "distance" $K(f, f^*) = E\log(f/f^*)(\eta_0)$. Then

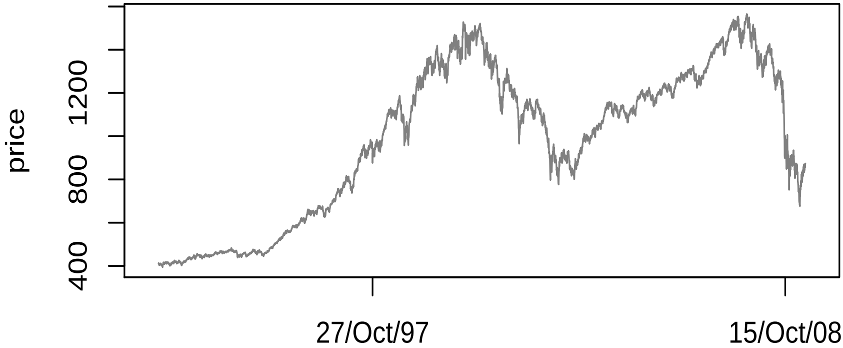
A3: $K(f,h) < K(f,h_\sigma) \qquad \forall \sigma > 0, \quad \sigma \neq 1$

Remark: When $h = f$ (MLE), **A3** vanishes.

Return

Stylized Facts (Mandelbrot (1963))

Non stationarity of the prices

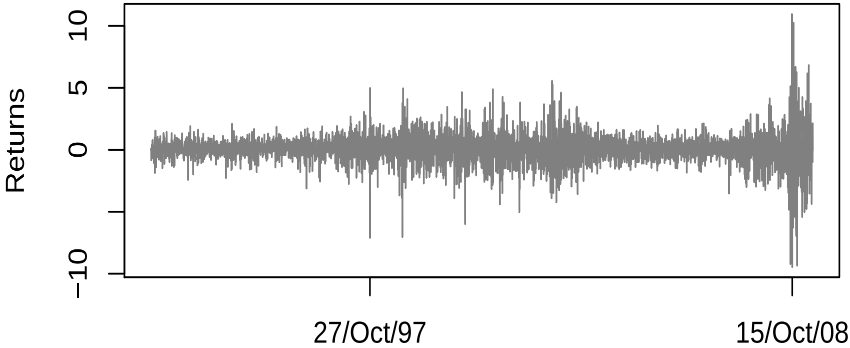


S&P 500, from March 2, 1992 to April 30, 2009

Return

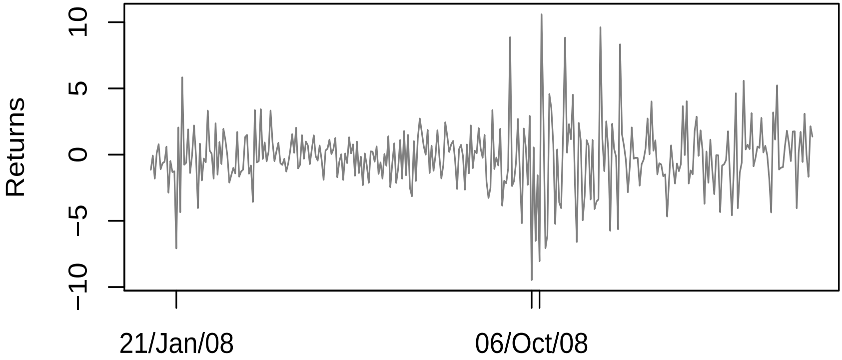
Stylized Facts

Possible stationarity of the returns



S&P 500 returns, from March 2, 1992 to April 30, 2009 Return

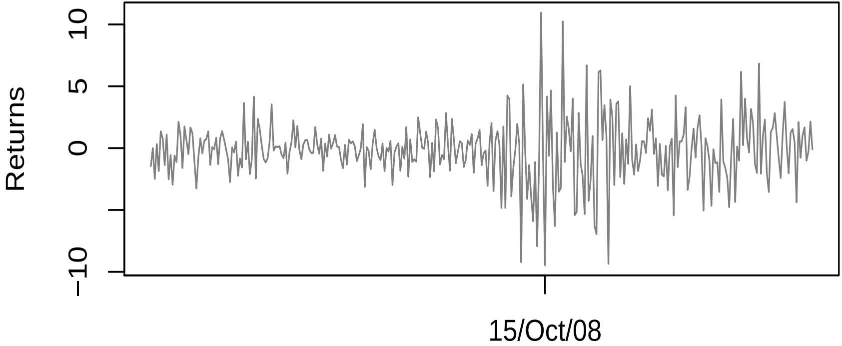
Stylized Facts
Volatility clustering



CAC 40 returns, from January 2, 2008 to April 30, 2009 Return

Stylized Facts

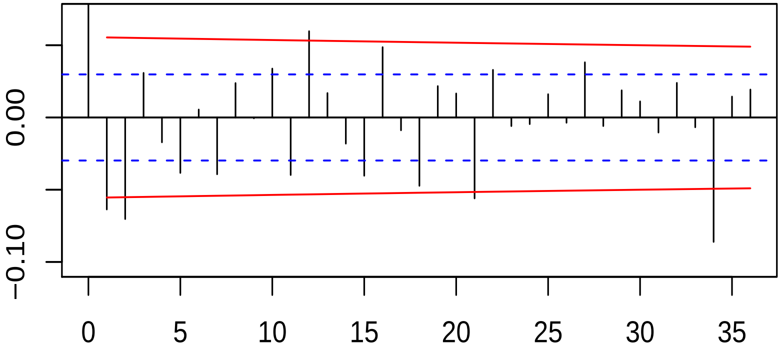
Conditional heteroskedasticity (compatible with marginal homoscedasticity and even stationarity)



S&P 500 returns, from January 2, 2008 to April 30, 2009 Return

Stylized Facts

Dependence without correlation (see FZ 2009 for the interpretation of the red lines)

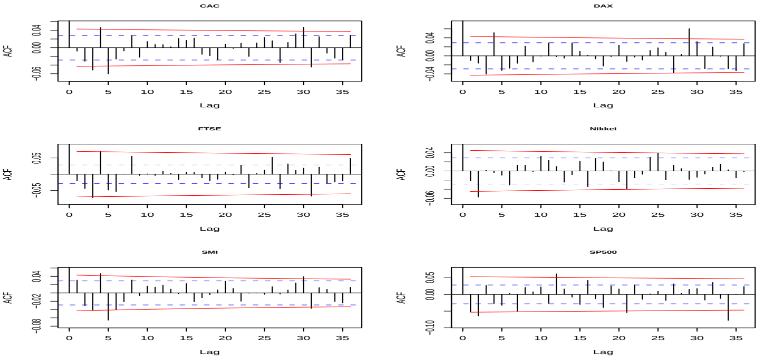


Empirical autocorrelations of the S&P 500 returns

Return

Stylized Facts

Dependence without correlation (significance bands under the GARCH(1,1) assumption)

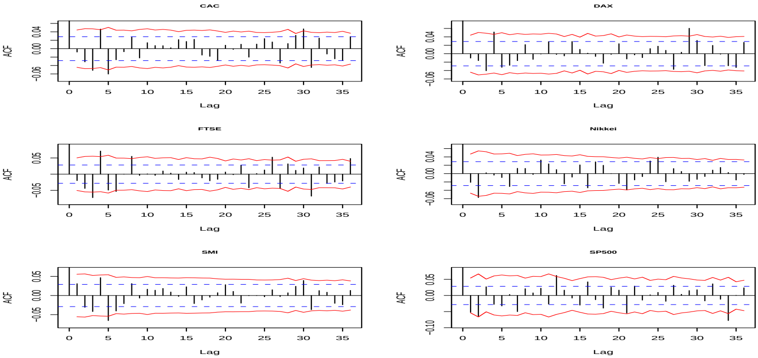


Empirical autocorrelations of daily stock returns

Return

Stylized Facts

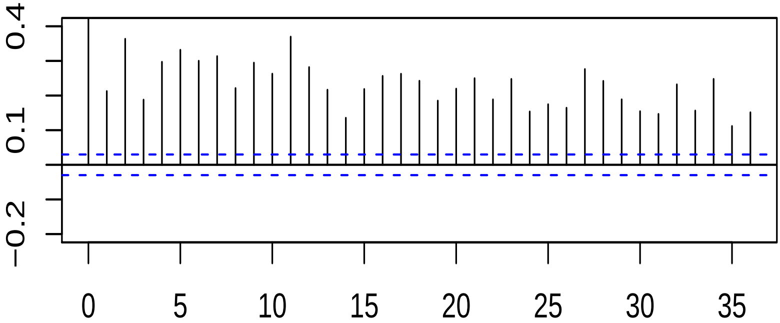
Dependence without correlation (the significance bands in red are estimated nonparametrically)



Empirical autocorrelations of daily stock returns

Return

Stylized Facts
Correlation of the squares

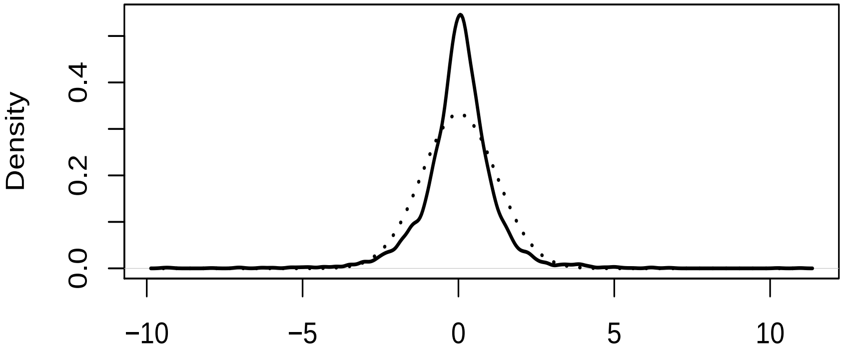


Autocorrelations of the squares of the S&P 500 returns

Return

Stylized Facts

Tail heaviness of the distributions



Density estimator for the S&P 500 returns (normal in dotted line)

Return

Stylized Facts

Decreases of prices have an higher impact on the future volatility than increases of the same magnitude

Table: Autocorrelations of tranformations of the CAC returns ϵ

h	1	2	3	4	5	6
$\hat{\rho}(\epsilon^+_{t-h}, \vert\epsilon_t\vert)$	0.03	0.07	0.07	0.08	0.08	0.12
$\hat{\rho}(-\epsilon^-_{t-h}, \vert\epsilon_t\vert)$	0.18	0.20	0.22	0.18	0.21	0.15

SP 500
Return

Stylized Facts

Decreases of prices have an higher impact on the future volatility than increases of the same magnitude

Table: Autocorrelations of tranformations of the S&P 500 returns ϵ

h	1	2	3	4	5	6
$\hat{\rho}_{\epsilon}(h)$	-0.06	-0.07	0.03	-0.02	-0.04	0.01
$\hat{\rho}_{\vert\epsilon\vert}(h)$	0.26	0.34	0.29	0.32	0.36	0.32
$\hat{\rho}(\epsilon_{t-h}^{+}, \vert\epsilon_t\vert)$	0.06	0.12	0.11	0.14	0.15	0.16
$\hat{\rho}(-\epsilon_{t-h}^{-}, \vert\epsilon_t\vert)$	0.25	0.28	0.23	0.24	0.28	0.23

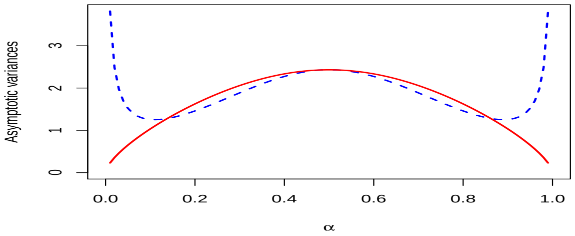
Return

Asymptotic variances of empirical quantiles, with or without estimation (dotted and full lines)

GED(ν) distribution

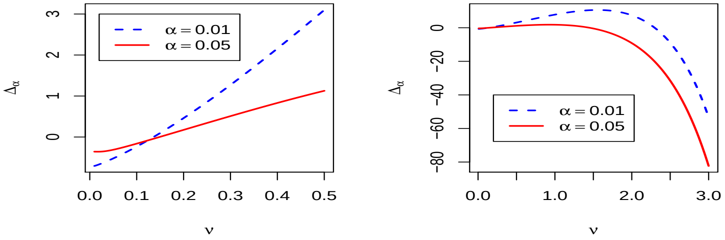
density $f(x) \propto \exp\{-0.5|x|^{1/\nu}\}$

$\nu = 0.25$



Return

Δ_α for GED(ν)



Student example