Merits and drawbacks of variance targeting in GARCH models

Christian Francq[1] Lajos Horvath[2] Jean-Michel Zakoïan[3]

[1]Université Lille 3
http://perso.univ-lille3.fr/~cfrancq

[2]University of Utah

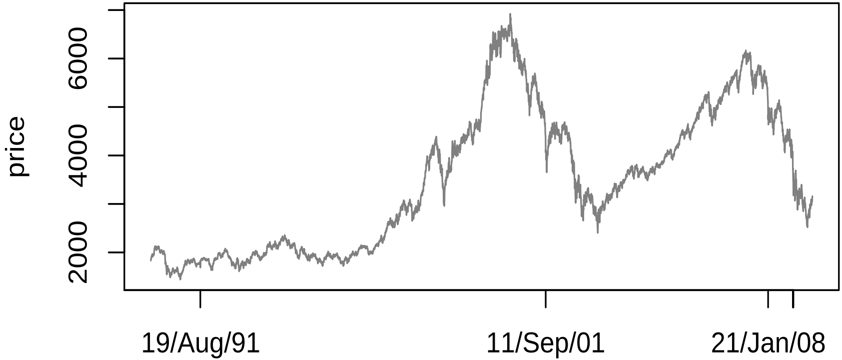
[3]CREST

USTHB, Alger

Outline

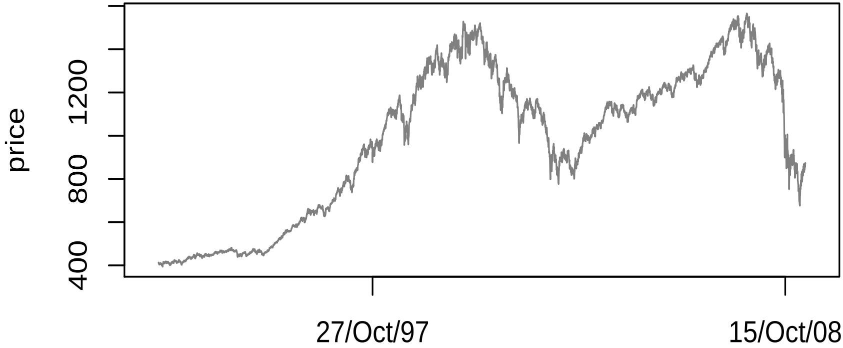
Outline

Stylized Facts (Mandelbrot (1963))
Non stationarity of the prices



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

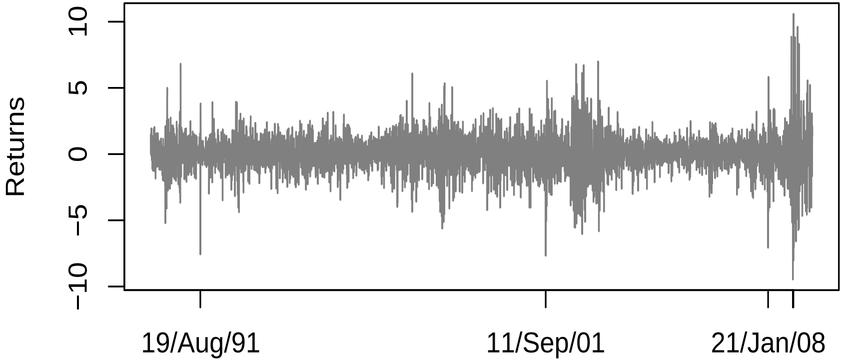
Stylized Facts (Mandelbrot (1963))
Non stationarity of the prices



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

Stylized Facts

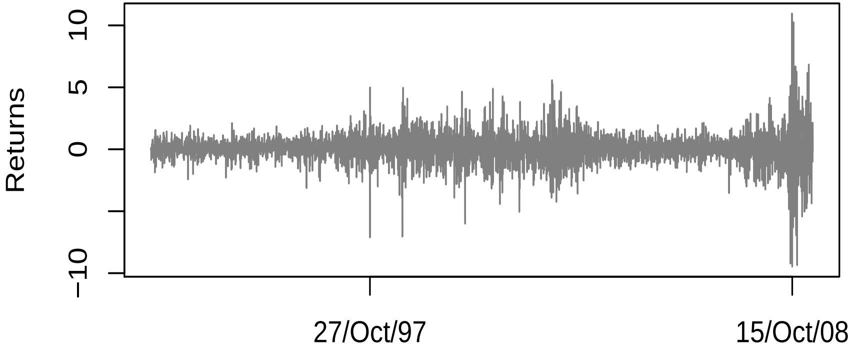
Possible stationarity of the returns



CAC 40 returns, from March 2, 1992 to February 20, 2009

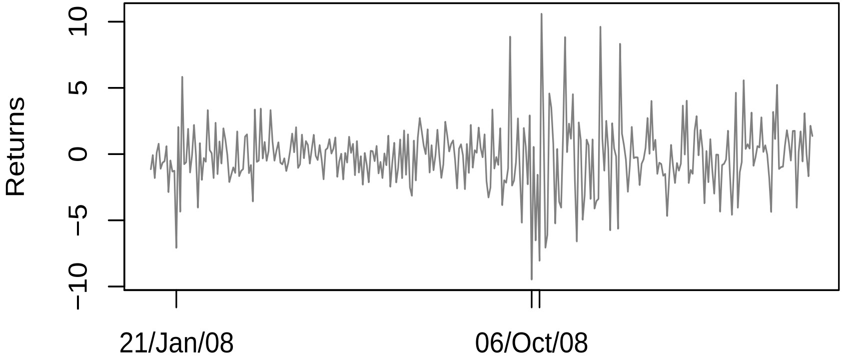
Stylized Facts

Possible stationarity of the returns



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

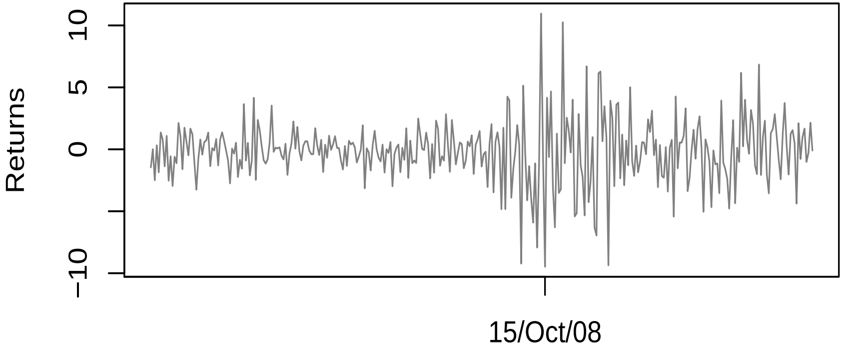
Stylized Facts
Volatility clustering



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

Stylized Facts

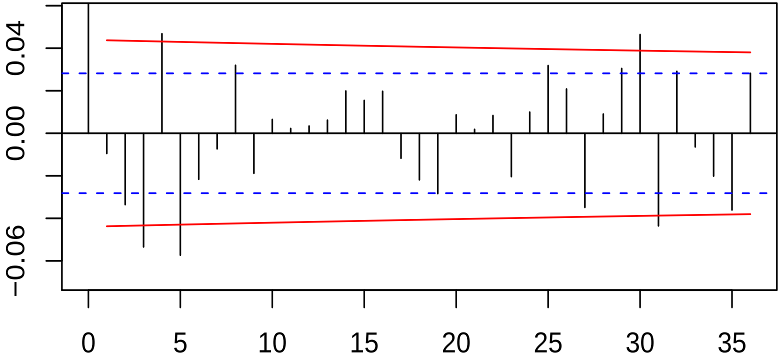
Conditional heteroskedasticity (compatible with marginal homoscedasticity and even stationarity)



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

Stylized Facts

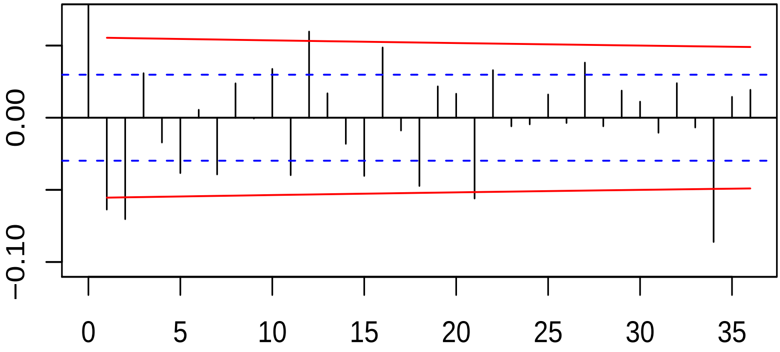
Dependence without correlation (warning: interpretation of the dotted lines)



Empirical autocorrelations of the CAC returns

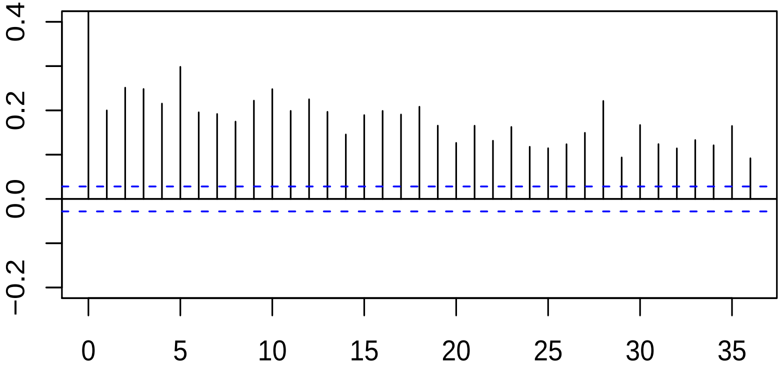
Stylized Facts

Dependence without correlation (see FZ 2009 http://perso.univ-lille3.fr/~cfrancq/Christian-Francq/Generalized-Bartlett-Formula.html for the interpretation of the red lines)



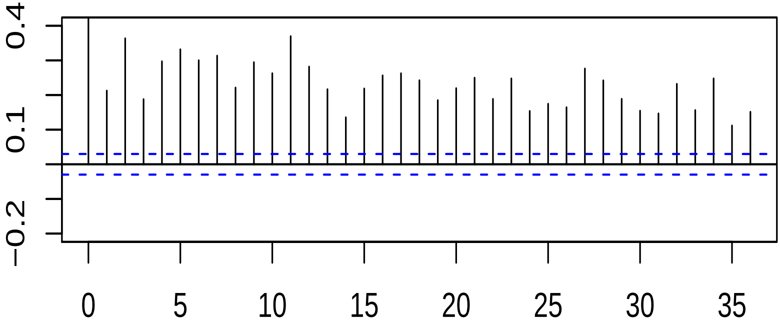
Empirical autocorrelations of the S&P 500 returns

Stylized Facts
Correlation of the squares



Autocorrelations of the squares of the CAC returns

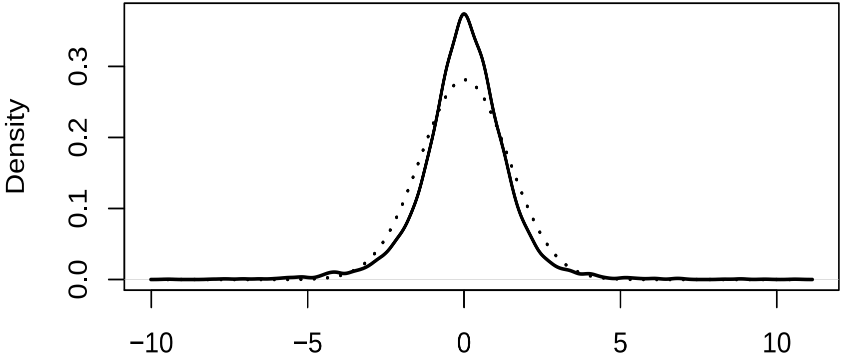
Stylized Facts
Correlation of the squares



Autocorrelations of the squares of the S&P 500 returns

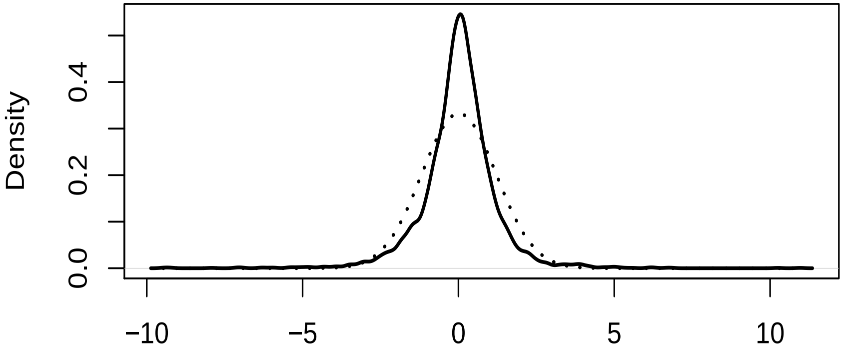
Stylized Facts

Tail heaviness of the distributions



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

Stylized Facts
Tail heaviness of the distributions



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

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
$\hat{\rho}_{\epsilon}(h)$	-0.01	-0.03	-0.05	0.05	-0.06	-0.02
$\hat{\rho}_{\|\epsilon\|}(h)$	0.18	0.24	0.25	0.23	0.25	0.23
$\hat{\rho}(\epsilon^{+}_{t-h}, \|\epsilon_t\|)$	0.03	0.07	0.07	0.08	0.08	0.12
$\hat{\rho}(-\epsilon^{-}_{t-h}, \|\epsilon_t\|)$	0.18	0.20	0.22	0.18	0.21	0.15

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
$\hat{\rho}_{\lvert\epsilon\rvert}(h)$	0.26	0.34	0.29	0.32	0.36	0.32
$\hat{\rho}(\epsilon_{t-h}^{+}, \lvert\epsilon_t\rvert)$	0.06	0.12	0.11	0.14	0.15	0.16
$\hat{\rho}(-\epsilon_{t-h}^{-}, \lvert\epsilon_t\rvert)$	0.25	0.28	0.23	0.24	0.28	0.23

Outline

Classes of Volatility Models

Amost all the models are of the form

$$\epsilon_t = \sigma_t \eta_t$$

where

- (η_t) is an iid (0,1) process
- (σ_t) is a process (volatility), $\sigma_t > 0$
- the variables σ_t and η_t are independent

Two main classes of models:

- GARCH-type (Generalized Autoregressive Conditional Heteroskedasticity): $\sigma_t \in \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots)$
- Stochastic volatility

Classes of Volatility Models

Amost all the models are of the form

$$\epsilon_t = \sigma_t \eta_t$$

where

- (η_t) is an iid (0,1) process
- (σ_t) is a process (volatility), $\sigma_t > 0$
- the variables σ_t and η_t are independent

Two main classes of models:

- GARCH-type (Generalized Autoregressive Conditional Heteroskedasticity): $\sigma_t \in \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots)$
- Stochastic volatility

Definition: GARCH(p, q)

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

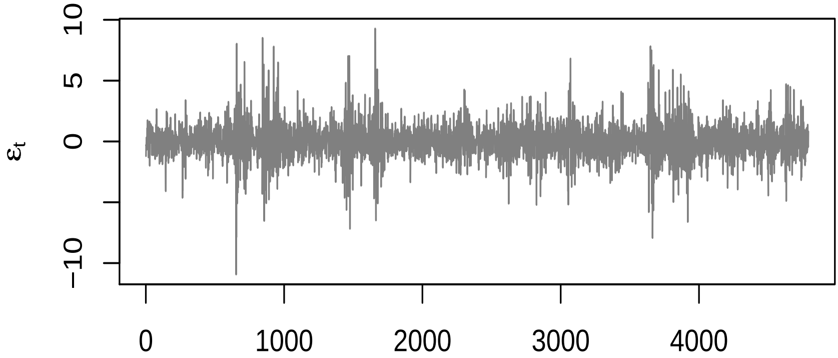
$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \\ \sigma_t^2 = \omega_0 + \sum_{i=1}^{q} \alpha_{0i} \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_{0j} \sigma_{t-j}^2, \quad \forall t \in \mathbb{Z} \end{cases}$$

where

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

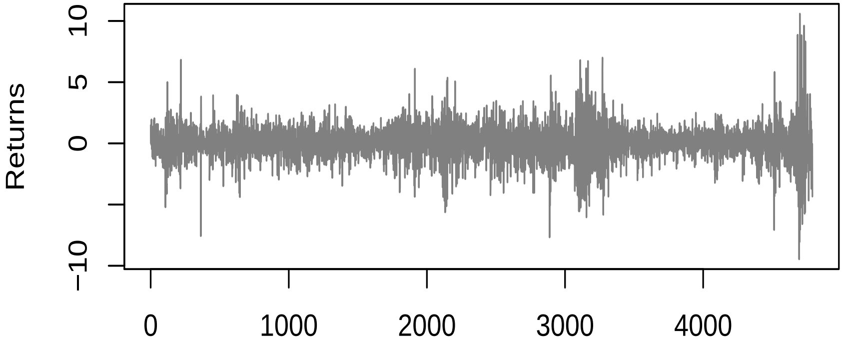
$$\theta_0 = (\omega_0, \alpha_{01}, \dots, \alpha_{0q}, \beta_{01}, \dots, \beta_{0p}).$$

GARCH(1,1) simulation



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

The previous GARCH(1,1) simulation resembles real financial series



CAC returns

Stricty Stationarity

$$A_{0t} = \begin{pmatrix} \alpha_{01}\eta_t^2 & \cdots & \alpha_{0q}\eta_t^2 & \beta_{01}\eta_t^2 & \cdots & \beta_{0p}\eta_t^2 \\ & I_{q-1} & 0 & & 0 & \\ \alpha_{01} & \cdots & \alpha_{0q} & \beta_{01} & \cdots & \beta_{0p} \\ & 0 & & & I_{p-1} & 0 \end{pmatrix}.$$

$$\gamma(\mathbf{A_0}) \quad = \quad \lim_{t\to\infty} a.s.\ \frac{1}{t} \log \|A_{0t}A_{0t-1}\ldots A_{01}\|.$$

Theorem (Bougerol & Picard, 1992)

The model has a (unique) strictly stationary non anticipative solution iff

$$\gamma(\mathbf{A_0}) < 0.$$

Quasi-Maximum Likelihood Estimation

A QMLE of θ is defined as any measurable solution $\hat{\theta}_n$ of

$$\hat{\theta}_n = \arg\min_{\theta\in\Theta} \tilde{\mathbf{l}}_n(\theta),$$

where $\tilde{\mathbf{l}}_n(\theta) = n^{-1}\sum_{t=1}^{n} \tilde{\ell}_t,$ and $\tilde{\ell}_t = \frac{\epsilon_t^2}{\tilde{\sigma}_t^2} + \log \tilde{\sigma}_t^2.$

Remark

- *The constraint $\tilde{\sigma}_t^2 > 0$ for all $\theta \in \Theta$ is necessary to compute $\tilde{\mathbf{l}}_n(\theta)$.*
- *The QMLE is always constrained: the "unrestricted" QMLE does not exist.*

Quasi-Maximum Likelihood Estimation

Theorem (Berkes, Horváth and Kokoszka (2003), FZ (2004))

Under appropriate conditions (in particular strict stationarity and $\theta_0 > 0$)

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (E\eta_1^4 - 1)J^{-1}),$$

$$J = E_{\theta_0}\left(\frac{1}{\sigma_t^4(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta'}\right).$$

Drawbacks of the QMLE

- Require a numerical optimization which is difficult when the number of parameters is large;
- The numerical optimization is sensitive to the choice of the initial value;
- The variance of the estimated model can be far from the empirical variance.

Outline

1. Volatility Models and QMLE
   - Properties of Financial Time Series
   - Models for the Volatility of Financial Returns
2. Variance Targeting Estimator
   - Description of the method
   - Asymptotic Properties of the VTE
   - Numerical comparison of the VTE and QMLE
     - On well specified GARCH models
     - On misspecified models (Long-term predictions and Value-at-Risks)
3. Conclusion

Variance Targeting Principle

Engle and Mezrich (1996)

- Principle of the two-step estimator:
  1. The unconditional variance is estimated by the sample variance;
  2. The remaining parameters are estimated by QML.
- Advantages:
  - Facilitates the numerical optimization by reducing the dimensionality of the parameter space;
  - Speeds up the convergence of the optimization routines;
  - Ensures a consistent estimate of the long-run variance even when the model is misspecified;
  - Provides reasonable initial values for the QMLE.
- Potential drawbacks:
  - Requires stronger assumptions (existence of the variance);
  - Is likely to suffer from efficiency loss.

Variance Targeting Principle
Engle and Mezrich (1996)

- Principle of the two-step estimator:
  1. The unconditional variance is estimated by the sample variance;
  2. The remaining parameters are estimated by QML.
- Advantages:
  - Facilitates the numerical optimization by reducing the dimensionality of the parameter space;
  - Speeds up the convergence of the optimization routines;
  - Ensures a consistent estimate of the long-run variance even when the model is misspecified;
  - Provides reasonable initial values for the QMLE.
- Potential drawbacks:
  - Requires stronger assumptions (existence of the variance);
  - Is likely to suffer from efficiency loss.

Variance Targeting Principle

Engle and Mezrich (1996)

- Principle of the two-step estimator:
  1. The unconditional variance is estimated by the sample variance;
  2. The remaining parameters are estimated by QML.
- Advantages:
  - Facilitates the numerical optimization by reducing the dimensionality of the parameter space;
  - Speeds up the convergence of the optimization routines;
  - Ensures a consistent estimate of the long-run variance even when the model is misspecified;
  - Provides reasonable initial values for the QMLE.
- Potential drawbacks:
  - Requires stronger assumptions (existence of the variance);
  - Is likely to suffer from efficiency loss.

Objectives

- Establish the asymptotic distribution of the VTE in univariate GARCH models
- Provide effective comparisons with the standard QML;
- Discuss the relative merits and drawbacks of the variance targeting method.

Reparameterization of the Standard GARCH(1,1)

- Standard form:

$$\epsilon_t = \sigma_t \eta_t \qquad \sigma_t^2 = \omega_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2,$$

  where $\boldsymbol{\theta}_0 = (\omega_0, \alpha_0, \beta_0)'$ is the unknown parameter.
- Alternative form: (with $\gamma_0 = \omega_0/(1 - \alpha_0 - \beta_0)$ when $\alpha_0 + \beta_0 < 1$)

$$\epsilon_t = \sigma_t \eta_t, \qquad \sigma_t^2 = \kappa_0 \gamma_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2, \quad \kappa_0 + \alpha_0 + \beta_0 = 1$$

  where $\boldsymbol{\vartheta}_0 = (\gamma_0, \alpha_0, \kappa_0)'$ is the new parameter.

Definition of the VTE of $\boldsymbol{\vartheta}_0 = (\gamma_0, \boldsymbol{\lambda}_0')'$ with $\boldsymbol{\lambda}_0 := (\alpha_0, \kappa_0)'$

1. First step: $\hat{\gamma}_n = \hat{\sigma}_n^2 := n^{-1} \sum_{t=1}^n \epsilon_t^2$,
2. Second step: $\hat{\boldsymbol{\lambda}}_n = \arg\min_{\boldsymbol{\lambda} \in \Lambda} \tilde{\mathbf{l}}_n(\boldsymbol{\lambda})$, where

$$\tilde{\mathbf{l}}_n(\boldsymbol{\lambda}) = n^{-1} \sum_{t=1}^n \ell_{t,n}, \qquad \ell_{t,n} := \ell_{t,n}(\boldsymbol{\lambda}) = \frac{\epsilon_t^2}{\sigma_{t,n}^2} + \log \sigma_{t,n}^2,$$

with

$$\sigma_{t,n}^2 = \sigma_{t,n}^2(\boldsymbol{\lambda}) = \kappa \hat{\sigma}_n^2 + \alpha \epsilon_{t-1}^2 + (1 - \kappa - \alpha)\sigma_{t-1,n}^2$$

and the initial value $\sigma_{0,n}^2 = \sigma_0^2$.

Standard QMLE of $\boldsymbol{\vartheta}_0 = (\gamma_0, \boldsymbol{\lambda}_0')'$

$$\hat{\boldsymbol{\vartheta}}_n^* = \arg\min_{\boldsymbol{\vartheta} \in \Theta} n^{-1} \sum_{t=1}^{n} \tilde{\ell}_t(\boldsymbol{\vartheta}),$$

where

$$\tilde{\ell}_t(\boldsymbol{\vartheta}) = \frac{\epsilon_t^2}{\tilde{\sigma}_t^2(\boldsymbol{\vartheta})} + \log \tilde{\sigma}_t^2(\boldsymbol{\vartheta}),$$

with

$$\tilde{\sigma}_t^2(\boldsymbol{\vartheta}) = \kappa\gamma + \alpha\epsilon_{t-1}^2 + (1 - \kappa - \alpha)\tilde{\sigma}_{t-1}^2(\boldsymbol{\vartheta})$$

and the initial value $\tilde{\sigma}_0^2(\boldsymbol{\vartheta}) = \sigma_0^2$. Note that $\tilde{\sigma}_t^2(\hat{\sigma}_n^2, \boldsymbol{\lambda}) = \sigma_{t,n}^2$.

Outline

Assumptions

in addition to η_t iid, $E\eta_t^2 = 1$, $\omega_0 > 0$, $\alpha_0 \geq 0$, $\beta_0 \geq 0$, $\alpha_0 + \beta_0 < 1$

The results are stated for the GARCH(1,1), but remain valid in the general GARCH(p, q) case.

Let the parameter space $\Lambda \subset \{(\alpha, \kappa) \mid \alpha \geq 0, \kappa > 0, \alpha + \kappa \leq 1\}$.

A1: λ_0 belongs to Λ and Λ is compact.

A2: $\alpha_0 \neq 0$ and η_t^2 has a non-degenerate distribution.

A3: $\alpha_0^2 \left(E\eta_t^4 - 1\right) + (1 - \kappa_0)^2 < 1$.

A4: λ_0 belongs to the interior of Λ.

Assumptions

in addition to η_t iid, $E\eta_t^2 = 1$, $\omega_0 > 0$, $\alpha_0 \geq 0$, $\beta_0 \geq 0$, $\alpha_0 + \beta_0 < 1$

The results are stated for the GARCH(1,1), but remain valid in the general GARCH(p, q) case.

Let the parameter space $\Lambda \subset \{(\alpha, \kappa) \mid \alpha \geq 0, \kappa > 0, \alpha + \kappa \leq 1\}$.

A1: $\boldsymbol{\lambda}_0$ belongs to Λ and Λ is compact.

A2: $\alpha_0 \neq 0$ and η_t^2 has a non-degenerate distribution.

A3: $\alpha_0^2 \left(E\eta_t^4 - 1\right) + (1 - \kappa_0)^2 < 1$.

A4: $\boldsymbol{\lambda}_0$ belongs to the interior of Λ.

Asymptotic properties of the GARCH(1,1) VTE

Theorem

Under Assumptions **A1**-**A2** *the VTE satisfies* $\hat{\vartheta}_n \to \vartheta_0$ *almost surely as* $n \to \infty$ *and, under the additional assumptions* **A3**-**A4**, *we have*

$$\sqrt{n}\left(\hat{\vartheta}_n - \vartheta_0\right) \xrightarrow{d} \mathcal{N}(0, (E\eta_1^4 - 1)\boldsymbol{\Sigma}).$$

Form of the VTE asymptotic variance

$$\boldsymbol{\Sigma} = \begin{pmatrix} b & -b\boldsymbol{K}'\boldsymbol{J}^{-1} \\ -b\boldsymbol{J}^{-1}\boldsymbol{K} & \boldsymbol{J}^{-1} + b\boldsymbol{J}^{-1}\boldsymbol{K}\boldsymbol{K}'\boldsymbol{J}^{-1} \end{pmatrix}$$

is non-singular with

$$\begin{aligned} b &= \frac{(\alpha_0 + \kappa_0)^2 \gamma^2 (2 - \kappa_0)}{\kappa_0 \left\{ 1 - \alpha_0^2 \left(E\eta_t^4 - 1 \right) - (1 - \kappa_0)^2 \right\}}, \\ \boldsymbol{J} &= E\left(\frac{1}{\sigma_t^4(\boldsymbol{\vartheta}_0)} \frac{\partial \sigma_t^2(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\lambda}} \frac{\partial \sigma_t^2(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\lambda}'} \right)_{2\times 2}, \\ \boldsymbol{K} &= E\left(\frac{1}{\sigma_t^4(\boldsymbol{\vartheta}_0)} \frac{\partial \sigma_t^2(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\lambda}} \frac{\partial \sigma_t^2(\boldsymbol{\vartheta}_0)}{\partial \gamma} \right)_{2\times 1}. \end{aligned}$$

VTE of the usual parameter $\boldsymbol{\theta}_0 = (\omega_0, \alpha_0, \beta_0)'$

Corollary

The VTE of $\boldsymbol{\theta}_0$ satisfies

$$\sqrt{n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right) \xrightarrow{d} \mathcal{N}(0, (E\eta_0^4 - 1)\boldsymbol{L}'\boldsymbol{\Sigma}\boldsymbol{L}),$$

with

$$\boldsymbol{L} = \begin{pmatrix} 1-\alpha_0-\beta_0 & 0 & 0 \\ 0 & 1 & -1 \\ \omega_0(1-\alpha_0-\beta_0)^{-1} & 0 & -1 \end{pmatrix}.$$

VTE and QMLE comparison

Theorem

The QMLE $\hat{\vartheta}_n^*$ satisfies

$$\sqrt{n}\left(\hat{\vartheta}_n^* - \vartheta_0\right) \xrightarrow{d} \mathcal{N}\left\{0, (E\eta_0^4 - 1)\boldsymbol{\Sigma}^*\right\},$$

where

$$\boldsymbol{\Sigma}^* = \boldsymbol{\Sigma} - (b-a)\boldsymbol{C}\boldsymbol{C}',$$

with

$$\boldsymbol{C} = \begin{pmatrix} 1 \\ -\boldsymbol{J}^{-1}\boldsymbol{K} \end{pmatrix}, \qquad a = \left\{\frac{\kappa_0^2}{(\alpha_0+\kappa_0)^2}E\left(\frac{1}{h_t^2}\right) - \boldsymbol{K}'\boldsymbol{J}^{-1}\boldsymbol{K}\right\}^{-1}.$$

The VTE is never asymptotically more accurate than the QMLE

Theorem

The asymptotic variance $(E\eta_0^4 - 1)\boldsymbol{\Sigma}$ of the VTE and the asymptotic variance $(E\eta_0^4 - 1)\boldsymbol{\Sigma}^*$ of the QMLE are such that

$$\boldsymbol{\Sigma} - \boldsymbol{\Sigma}^* \quad \textit{is positive semidefinite, but not positive definite.}$$

Proof that the QMLE is asymptotically more efficient than the VTE

The asymptotic variances of the two estimators are the variances of linear combinations of a same vector:

$$\boldsymbol{\Sigma}^* = \left\{E\left(\boldsymbol{G}\boldsymbol{S}_t\boldsymbol{S}_t'\boldsymbol{G}'\right)\right\}^{-1}, \qquad \boldsymbol{\Sigma} = E\left(\boldsymbol{H}\boldsymbol{S}_t\boldsymbol{S}_t'\boldsymbol{H}'\right)$$

where

$$\boldsymbol{G} = \begin{pmatrix} \boldsymbol{I}_2 & 0 \end{pmatrix}, \qquad \boldsymbol{H} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & \boldsymbol{J}^{-1} & -\boldsymbol{J}^{-1}\boldsymbol{K} \end{pmatrix},$$

and

$$\boldsymbol{S}_t = \begin{pmatrix} eh_t^{-1} \\ h_t^{-1}\frac{\partial\sigma_t^2}{\partial\boldsymbol{\lambda}}(\vartheta_0) \\ e^{-1}h_t \end{pmatrix} \qquad e = \frac{\kappa_0}{1-\beta_0}.$$

End of the Proof

We then easily show that

$$\boldsymbol{\Sigma} - \boldsymbol{\Sigma}^* = E\boldsymbol{D}_t\boldsymbol{D}_t'$$

where $\boldsymbol{D}_t = \boldsymbol{\Sigma}^*\boldsymbol{G}\boldsymbol{S}_t - \boldsymbol{H}\boldsymbol{S}_t$.

Cases where VTE and QMLE have the same asymptotic variance

Theorem

Let ϕ be a mapping from $\mathbb{R}^2$ to $\mathbb{R}$, which is continuously differentiable in a neighborhood of ϑ_0. If

$$\begin{pmatrix} 1 & -\boldsymbol{K}'\boldsymbol{J}^{-1} \end{pmatrix} \frac{\partial \phi}{\partial \vartheta}(\vartheta_0) = 0,$$

then the asymptotic distribution of the VTE of the parameter $\phi(\vartheta_0)$ is the same as that of the QMLE.

Cases where VTE and QMLE have the same asymptotic variance

Under the previous condition,

$$\sqrt{n}\left\{\phi(\hat{\vartheta}_n) - \phi(\vartheta_0)\right\} \xrightarrow{d} \mathcal{N}\left(0, s^2\right)$$

and

$$\sqrt{n}\left\{\phi(\hat{\vartheta}_n^*) - \phi(\vartheta_0)\right\} \xrightarrow{d} \mathcal{N}\left(0, s^2\right),$$

where

$$s^2 = (E\eta_0^4 - 1)\frac{\partial \phi}{\partial \boldsymbol{\vartheta}'}\boldsymbol{\Sigma}\frac{\partial \phi}{\partial \boldsymbol{\vartheta}}.$$

Outline

1. Volatility Models and QMLE
   - Properties of Financial Time Series
   - Models for the Volatility of Financial Returns
2. Variance Targeting Estimator
   - Description of the method
   - Asymptotic Properties of the VTE
   - Numerical comparison of the VTE and QMLE
     - On well specified GARCH models
     - On misspecified models (Long-term predictions and Value-at-Risks)
3. Conclusion

Numerical evaluation of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}^*$

The asymptotic variance of the VTE

$$\boldsymbol{\Sigma} = \begin{pmatrix} b & -b\boldsymbol{K}'\boldsymbol{J}^{-1} \\ -b\boldsymbol{J}^{-1}\boldsymbol{K} & \boldsymbol{J}^{-1} + b\boldsymbol{J}^{-1}\boldsymbol{K}\boldsymbol{K}'\boldsymbol{J}^{-1} \end{pmatrix}$$

and that of the QMLE, $\boldsymbol{\Sigma}^*$, are not numerically computable, even for the simplest model (the ARCH(1)).

$\rightsquigarrow$ Approximations of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}^*$ from $N = 1,000$ replications of simulations of size $n = 10,000$.

Asymptotic variances of the QMLE and VTE

for $\vartheta_0 = (\gamma_0, \alpha_0)$ in ARCH(1) models, $\gamma_0 = 1$ and $\eta_t \sim \mathcal{N}(0,1)$

	$\alpha_0 = 0.1$	$\alpha_0 = 0.55$	$\alpha_0 = 0.7$
QMLE	$\begin{pmatrix} 2.52 & 0.51 \\ 0.51 & 1.69 \end{pmatrix}$	$\begin{pmatrix} 15.94 & 7.11 \\ 7.11 & 4.20 \end{pmatrix}$	$\begin{pmatrix} 45.27 & 14.32 \\ 14.32 & 5.02 \end{pmatrix}$
VTE	$\begin{pmatrix} 2.52 & 0.51 \\ 0.51 & 1.69 \end{pmatrix}$	$\begin{pmatrix} 28.78 & 12.82 \\ 12.82 & 6.74 \end{pmatrix}$	∞

Asymptotic variances of the QMLE and VTE

for $\boldsymbol{\theta}_0 = (\omega_0, \alpha_0)$ in ARCH(1) models, $\omega_0 = 1$ and $\eta_t \sim \mathcal{N}(0, 1)$

	$\alpha_0 = 0.1$	$\alpha_0 = 0.55$	$\alpha_0 = 0.7$
QMLE	$\begin{pmatrix} 3.5 & -1.4 \\ -1.4 & 1.7 \end{pmatrix}$	$\begin{pmatrix} 5.1 & -2.2 \\ -2.2 & 4.2 \end{pmatrix}$	$\begin{pmatrix} 5.6 & -2.4 \\ -2.4 & 5.1 \end{pmatrix}$
VTE	$\begin{pmatrix} 3.5 & -1.4 \\ -1.4 & 1.7 \end{pmatrix}$	$\begin{pmatrix} 5.1 & -2.1 \\ -2.1 & 9.3 \end{pmatrix}$	∞

Sampling distribution of the two estimators

on $N = 1,000$ independent ARCH(1) simulations

$n = 500$

parameter	true value	estimator	bias	RMSE
ω	1.0	QMLE	0.013	0.102
		VTE	0.012	0.102
α	0.55	QMLE	-0.012	0.092
		VTE	-0.026	0.088
ω	1.0	QMLE	0.012	0.114
		VTE	0.036	0.111
α	0.9	QMLE	-0.012	0.110
		VTE	-0.103	0.089

Sampling distribution of the two estimators
on $N = 1,000$ independent ARCH(1) simulations

$n = 10,000$

parameter	true value	estimator	bias	RMSE
ω	1.0	QMLE	0.000	0.010
		VTE	0.000	0.010
α	0.55	QMLE	0.000	0.009
		VTE	0.000	0.013
ω	1.0	QMLE	0.000	0.012
		VTE	0.010	0.015
α	0.9	QMLE	0.000	0.011
		VTE	-0.032	0.032

Comparison on daily stock returns

Index	estimator	ω	α	β
CAC	QMLE	0.033 (0.009)	0.090 (0.014)	0.893 (0.015)
	VTE	0.033 (0.009)	0.090 (0.014)	0.893 (0.015)
DAX	QMLE	0.037 (0.014)	0.093 (0.023)	0.888 (0.024)
	VTE	0.036 (0.013)	0.095 (0.022)	0.888 (0.024)
FTSE	QMLE	0.013 (0.004)	0.091 (0.014)	0.899 (0.014)
	VTE	0.013 (0.004)	0.090 (0.013)	0.899 (0.014)
Nasdaq	QMLE	0.025 (0.006)	0.072 (0.009)	0.922 (0.009)
	VTE	0.025 (0.006)	0.072 (0.009)	0.922 (0.009)
Nikkei	QMLE	0.053 (0.012)	0.100 (0.013)	0.880 (0.014)
	VTE	0.054 (0.012)	0.098 (0.013)	0.880 (0.015)
SP500	QMLE	0.014 (0.004)	0.084 (0.012)	0.905 (0.012)
	VTE	0.014 (0.003)	0.084 (0.011)	0.905 (0.012)

Computation time comparison for estimating GARCH(1,1) models on a set of 11 stock indices

Table: Design 1 and Design 2 correspond to different initial values.

	Design 1	Design 2
VTE	39.0	55.5
QMLE	61.6	88.1
VTE+QMLE	85.1	98.9

In Design 2, for two series, the QMLE leads to a nonoptimal local maximum

Long-term predictions with misspecified models

Two GARCH(1,1) models are estimated by VTE and by QMLE and are used to compute prediction intervals for ϵ_{n+h}:

$$\left[\sqrt{\hat{\sigma}^2_{n+h|n}}\hat{F}_\eta^{-1}(\underline{\alpha}/2),\quad \sqrt{\hat{\sigma}^2_{n+h|n}}\hat{F}_\eta^{-1}(1-\underline{\alpha}/2)\right],$$

when

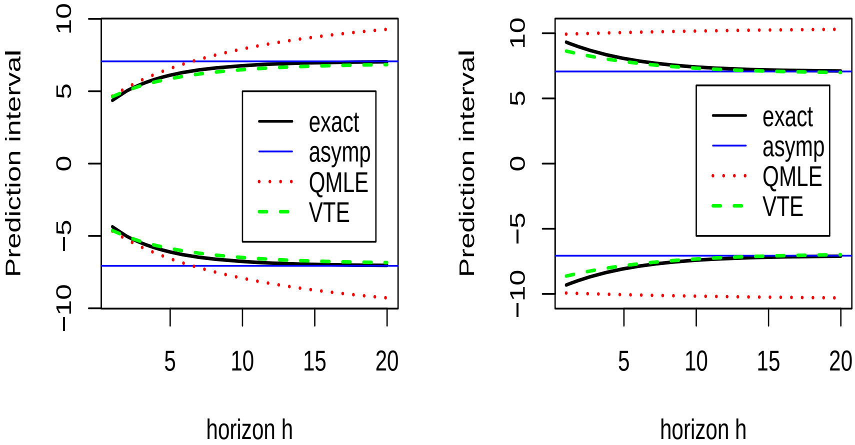
$$\epsilon_t = \omega(\Delta_t)\eta_t, \quad \eta_t \text{ iid } \mathcal{N}(0,1),$$

(Δ_t) is a Markov chain, independent of (η_t), with state-space $\{1,2\}$ and transition probabilities $P(\Delta_t = 1|\Delta_{t-1} = 1) = P(\Delta_t = 2|\Delta_{t-1} = 2) = 0.9.$

- $\rightsquigarrow$ Contrary to the QMLE, TVE should guarantee correct predictions over long horizons.

TVE guarantees correct long-term predictions

Prediction intervals of the Markov-switching model with different methods



Value-at-Risk(VaR)

Let V_t be the value of a portfolio at time t. The (conditional) VaR is the $(1-\underline{\alpha})$-quantile of the conditional distribution of $L_{t,t+h} = -(V_{t+h} - V_t)$:

$$\mathrm{VaR}_{t,h}(\underline{\alpha}) = \inf \left\{ x \in \mathbb{R} \mid P\left(L_{t,t+h} \leq x \mid V_u, u \leq t\right) \geq 1 - \underline{\alpha} \right\}.$$

Introducing the log-returns $\epsilon_t = \log(V_t / V_{t-1})$,

$$\mathrm{VaR}_{t,h}(\underline{\alpha}) = \left[1 - \exp\left\{q_{t,h}(\underline{\alpha})\right\}\right] V_t,$$

where $q_{t,h}(\underline{\alpha})$ is the $\underline{\alpha}$-quantile of the conditional distribution of the future returns $\epsilon_{t+1} + \cdots + \epsilon_{t+h}$.

Estimating long horizon VaR

Lemma

Assume that (ϵ_t) *is a strictly stationary process such that* $E\epsilon_t = 0$, $\sum_{h=1}^{\infty} \{\alpha_\epsilon(h)\}^{\nu/(2+\nu)} < \infty$ *and* $E|\epsilon_t|^{2+\nu} < \infty$ *for some* $\nu > 0$. *Let* $Var(\epsilon_t) = \overline{\omega}^2$. *We have*

$$\lim_{h\to\infty} \sqrt{h}\,\overline{\omega}\,\Phi^{-1}(\underline{\alpha})/q_{t,h}(\underline{\alpha}) = 1 \quad a.s.$$

- horizon 1:

$$\mathrm{VaR}_{t,1}(\underline{\alpha}) = \left[1 - \exp\left\{\sigma_t(\theta_0)F_\eta^{-1}(1-\underline{\alpha})\right\}\right] V_t,$$

- long horizon:

$$\widehat{\mathrm{VaR}}_{t,h}(\underline{\alpha}) = \left[1 - \exp\left\{\sqrt{h}\,\Phi^{-1}(\underline{\alpha})\,\widehat{\overline{\omega}}\right\}\right] V_t.$$

Estimating long horizon VaR

Lemma

Assume that (ϵ_t) is a strictly stationary process such that $E\epsilon_t = 0$, $\sum_{h=1}^{\infty} \{\alpha_\epsilon(h)\}^{\nu/(2+\nu)} < \infty$ and $E|\epsilon_t|^{2+\nu} < \infty$ for some $\nu > 0$. Let $Var(\epsilon_t) = \overline{\omega}^2$. We have

$$\lim_{h\to\infty} \sqrt{h}\,\overline{\omega}\,\Phi^{-1}(\underline{\alpha})/q_{t,h}(\underline{\alpha}) = 1 \quad a.s.$$

- horizon 1:

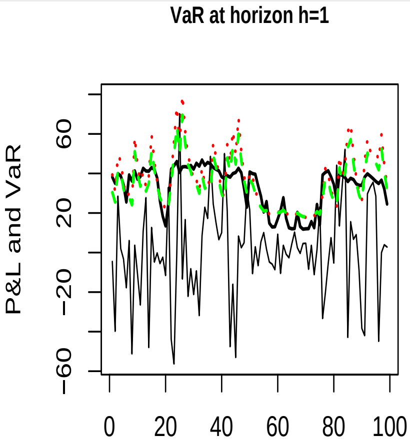
$$\mathrm{VaR}_{t,1}(\underline{\alpha}) = \left[1 - \exp\left\{\sigma_t(\boldsymbol{\theta}_0) F_\eta^{-1}(1-\underline{\alpha})\right\}\right] V_t,$$

- long horizon:

$$\widehat{\mathrm{VaR}}_{t,h}(\underline{\alpha}) = \left[1 - \exp\left\{\sqrt{h}\,\Phi^{-1}(\underline{\alpha})\,\widehat{\overline{\omega}}\right\}\right] V_t.$$

Estimating long horizon VaR with misspecified models

Comparison between the true VaR (black line) computed from the DGP (an HMM model) and the VaR's computed from a GARCH(1,1) estimated by QMLE (red) and VTE (green)



VTE can be recommended because it

- reduces the computational complexity of GARCH estimation;
- is asymptotically less efficient that the QMLE, but can work better in finite sample;
- requires fourth-order moments for asymptotic normality, but continues to work well with lower moments;
- provides good (first step) estimations of real financial series;
- guarantees a consistent estimation of the long-run variance;
- thus guarantees correct long-horizon predictions and VaR's;
- can be an indicator of misspecification if an important discrepancy with the QMLE is observed.

Directions for Future Work

- Extension to other GARCH formulations;
- Extension to multivariate models.
- Other applications where the long-term variance is essential (prediction of the realized volatility).