Centre de Recherche en Économie et Statistique
CREST

Série des Documents de Travail

n° 2014-01

Multi-level Conditional VaR Estimation in Dynamic Models

C. FRANCQ[1]
J.-M. ZAKOIAN[2]

[1] CREST, 15 Boulevard Gabriel Péri, 92245 Malakoff Cedex, France. christian.francq@ensae.fr
[2] CREST, 15 Boulevard Gabriel Péri, 92245 Malakoff Cedex, France. zakoian@ensae.fr

Multi-level conditional VaR estimation in dynamic models

Christian Francq* and Jean-Michel Zakoïan†‡

Abstract

We consider joint estimation of conditional Value-at-Risk (VaR) at several levels, in the framework of general conditional heteroskedastic models. The volatility is estimated by Quasi-Maximum Likelihood (QML) in a first step, and the residuals are used to estimate the innovations quantiles in a second step. The joint limiting distribution of the volatility parameter and a vector of residual quantiles is derived. We deduce confidence intervals for general Distortion Risk Measures (DRM) which can be approximated by a finite number of VaR's. We also propose an alternative approach based on non Gaussian QML which, although numerically more cumbersome, has interest when the innovations distribution is fat tailed. An empirical study based on stock indices illustrates the theoretical findings.

JEL Classification: C13, C22 and C58.
Keywords: GARCH, Distortion Risk Measures, Quasi-Maximum Likelihood, Value-at-Risk.

*CREST, 15 Boulevard G. Péri, 92245 Malakoff Cedex, France. E-mail: Christian.Francq@ensae.fr

†CREST, 15 Boulevard G. Péri, 92245 Malakoff Cedex, France. E-mail: zakoian@ensae.fr

‡The authors gratefully acknowledge financial support of the Labex ECODEC. The second author gratefully thanks the IDR ACP "Régulation et risques systémiques" for financial support.

1 Introduction

Under the regulations introduced in Finance since Basel 2, bank capital is risk-sensitive. Financial institutions are required to measure the riskiness of their assets and, for instance, to hold more capital to compensate more risk. While the Value-at-Risk (VaR), defined as a quantile of some loss distribution, continues to play a prominent role in the mainstream financial risk management, a variety of alternative risk measures have been introduced and studied in recent years. The Expected Shortfall (ES), and more generally the Distortion Risk Measures (DRM), are quantile-based measures which, by comparison with the VaR at a given level, give further insight on the shape of the loss distribution[1].

Whatever the choice of a risk measure, it depends on unknown characteristics of the loss distribution which, for practical use, have to be estimated. In the so-called standard approach, the quantity of interest is a parameter, defined as a characteristic of the *marginal* loss distribution. In the so-called advanced approaches, the focus is on *conditional* characteristics of the loss distributions, that is, characteristics which, at the current date, take into account the available past information. The conditional VaR, and more generally conditional risk measures, are stochastic processes, which are not directly observable just like volatility. This complicates the statistical inference of risk measures. The problem is not only to get consistent estimators of conditional risks but also to evaluate the accuracy of such estimators[2].

[1] These measures are also advocated because, contrary to the VaR, they satisfy a set of "coherence requirements" for a large family of distributions.

[2] In July 2009, the Basel Committee issued a directive requiring that financial institutions quantify "model risk". The Committee states that *"Banks must explicitly assess the need for valuation adjustments to reflect two forms of model risk: the model risk asso-*

Confidence intervals for conditional VaR's were derived, in the recent econometric literature, using different approaches. Chan, Deng, Peng, Xia (2007) constructed confidence intervals under the assumption that the errors have heavy tails, using the Extreme-Value Theory, while Spierdijk (2013) proposed a residual subsample bootstrap approach. Francq and Zakoïan (2012) used a QML approach. They showed that the problem of estimating a conditional risk measure, for instance a VaR at a given level, in GARCH-type models reduced to the estimation of a parameter, called risk parameter.

In the present article we extend those results to the joint estimation of several conditional risks. In practice, it is often important to handle several risk levels, in order to have a better view on the tail properties of the conditional distribution. We will provide statistical tools for jointly estimating conditional VaR's corresponding to different levels, in a general GARCH-type framework which does not impose a specific form for the volatility, and for estimating the accuracy of such VaR estimates. Our approach is aimed at, not only providing VaR estimates, but also confidence intervals based on asymptotic results. A tractable risk measure based on a vector of risk levels can be defined by weighting the corresponding VaR's, that is, by defining a *portfolio* of VaR's. This approach can be connected with DRMs through an appropriate choice of the weights. For a given DRM, our asymptotic results allow us to construct upper and lower bounds based on a finite number of VaR's.

This paper is organized as follows. In Section 2, we start by introduc-

ciated with using a possibly incorrect valuation methodology; and the risk associated with using unobservable (and possibly incorrect) calibration parameters in the valuation model." For instance, an important issue in determining the reserves of a financial institution is whether VaR estimates remain reliable in very hectic periods.

ing a general class of GARCH-type models. Then we derive the asymptotic joint distribution of the Quasi-Maximum Likelihood Estimator (QMLE) and a vector of empirical quantiles of the residuals. We deduce asymptotic confidence intervals for the VaR's and for VaR portfolios. Section 3 proposes another approach for conditional VaR estimation based on non Gaussian QMLEs. An empirical illustration based on major stock indices is proposed in Section 4. Section 5 concludes.

2 Two-step VaR estimation in volatility models

2.1 Conditional VaR in a general model

Consider a GARCH-type model of the form

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_0) \end{cases} \tag{2.1}$$

where (η_t) is a sequence of iid random variables, η_t is independent of $\{\epsilon_u, u < t\}$, $\boldsymbol{\theta}_0 \in \mathbb{R}^d$ is a parameter belonging to a parameter space Θ, and $\sigma : \mathbb{R}^\infty \times \Theta \to (0, \infty)$. The most widely used specifications of volatility belong to this class, in particular the GARCH(p, q) model of Engle (1982) and Bollerslev (1986),

$$\begin{cases} \epsilon_t = \sigma_t \eta_t, \\ \sigma_t^2 = \omega_0 + \sum_{i=1}^q a_{0i} \epsilon_{t-i}^2 + \sum_{j=1}^p b_{0j} \sigma_{t-j}^2, \end{cases} \tag{2.2}$$

where $\boldsymbol{\theta}_0 = (\omega_0, a_{01}, \dots, b_{0p})'$ satisfies $\omega_0 > 0, a_{0i} \geq 0,\ b_{0j} \geq 0$. For this model, if the lag polynomial $\beta(L) = 1 - \sum_{j=1}^p L^j$ has its roots outside the unit disk, we have a representation of the form (2.1) given by

$$\sigma_t^2 = \beta(1)^{-1} \omega_0 + \sum_{i=1}^\infty \gamma_i \epsilon_{t-i}^2,$$

where $\beta(L)^{-1}\sum_{i=1}^{q} a_i L^i = \sum_{i=1}^{\infty} \gamma_i L^i$. Other classical examples of models belonging to the class (2.1) are the EGARCH, GJR-GARCH, TGARCH, QGARCH, APARCH, Log-GARCH, models introduced, respectively, by Nelson (1991), Glosten, Jagannathan and Runkle (1993), Zakoïan (1994), Sentana (1995), Ding, Granger and Engle (1993), and for the log-GARCH, under slightly different forms, by Geweke (1986), Pantula (1986) and Milhøj (1987). See Francq and Zakoïan (2010) for an overview on GARCH models.

The *conditional* VaR of a process (ϵ_t) at risk level $\alpha \in (0,1)$, denoted by $\mathrm{VaR}_t(\alpha)$, is defined by

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

where P_{t-1} denotes the historical distribution conditional on $\{\epsilon_u, u < t\}$. When (ϵ_t) satisfies (2.1), the theoretical VaR is then given by

$$\mathrm{VaR}_t(\alpha) = -\sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_0)\xi_\alpha \tag{2.3}$$

where ξ_α is the α-quantile of η_t.

Remark 2.1 It can be noted that in the standard GARCH(p,q) model, the conditional VaR at level α automatically satisfies the stochastic recurrence equation

$$\mathrm{VaR}_t^2(\alpha) = \omega_0\xi_\alpha^2 + \sum_{i=1}^{q} a_{0i}\xi_\alpha^2\epsilon_{t-i}^2 + \sum_{j=1}^{p} b_{0j}\mathrm{VaR}_{t-j}^2(\alpha).$$

Direct modelling of the conditional VaR has been proposed in several papers, for instance Engle and Manganelli (2004), Koenker and Xiao (2006), Gouriéroux and Jasiak (2008). A difficulty in this approach is to constrain the model so as to guarantee the monotonicity of the conditional VaR as a function of the risk level. Monotonicity is automatically satisfied in our approach.

2.2 Asymptotic properties of the multi-level two-step VaR estimator

A two-step standard method for evaluating the VaR at different levels $\alpha_i \in (0,1)$, for $i = 1, \dots, m$ consists in estimating the volatility parameter $\boldsymbol{\theta}_0$ by Gaussian QMLE, and then estimating the ξ_{α_i} by the corresponding empirical quantiles of the residuals; see, for instance, Chapter 2 in McNeil, Frey and Embrechts (2005). For a comparison of alternative strategies based on residuals following a preliminary volatility estimation, see Kuester, Mittnik and Paolella (2006).

Given observations $\epsilon_1, \dots, \epsilon_n$, and arbitrary initial values $\tilde{\epsilon}_i$ for $i \leq 0$, we define, under assumptions given below,

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

which is used to approximate $\sigma_t(\boldsymbol{\theta}) = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_1, \epsilon_0, \epsilon_{-1}, \dots; \boldsymbol{\theta})$. A QMLE of $\boldsymbol{\theta}_0$ in Model (2.1) is defined as any measurable solution $\widehat{\boldsymbol{\theta}}_n$ of

$$\widehat{\boldsymbol{\theta}}_n = \arg\min_{\boldsymbol{\theta} \in \Theta} \widetilde{Q}_n(\boldsymbol{\theta}), \tag{2.4}$$

with

$$\widetilde{Q}_n(\boldsymbol{\theta}) = n^{-1} \sum_{t=1}^{n} \widetilde{\ell}_t(\boldsymbol{\theta}), \qquad \widetilde{\ell}_t(\boldsymbol{\theta}) = \frac{\epsilon_t^2}{\widetilde{\sigma}_t^2(\boldsymbol{\theta})} + \log \widetilde{\sigma}_t^2(\boldsymbol{\theta}).$$

The following assumptions are required to derive the asymptotic properties of the QMLE $\widehat{\boldsymbol{\theta}}_n$.

A1: (ϵ_t) is a strictly stationary and ergodic solution of Model (2.1). Moreover, $E|\epsilon_0|^s < \infty$ for some $s > 0$.

A2: For any real sequence (x_i), the function $\boldsymbol{\theta} \mapsto \sigma(x_1, x_2, \dots; \boldsymbol{\theta})$ is continuous. Almost surely, $\sigma_t(\boldsymbol{\theta}) \in (\underline{\omega}, \infty]$ for any $\boldsymbol{\theta} \in \Theta$ and for some $\underline{\omega} > 0$.

A3: The function $\boldsymbol{\theta} \mapsto \sigma(x_1, x_2, \ldots; \boldsymbol{\theta})$ has continuous second-order derivatives, and

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

where C_1 is a random variable which is measurable with respect to $\{\epsilon_u, u < 0\}$ and $\rho \in (0,1)$ is a constant.

A4($\boldsymbol{\theta}_0^*$): $\boldsymbol{\theta}_0^*$ belongs to the interior of Θ and $\sigma_t(\boldsymbol{\theta}_0^*)/\sigma_t(\boldsymbol{\theta}) = 1$ *a.s.* iff $\boldsymbol{\theta} = \boldsymbol{\theta}_0^*$.

A5($\boldsymbol{\theta}_0^*$): There exist no non-zero $x \in \mathbb{R}^d$ such that $x' \frac{\partial \sigma_t(\boldsymbol{\theta}_0^*)}{\partial \boldsymbol{\theta}} = 0$, *a.s.*

A6($\boldsymbol{\theta}_0^*$): There exists a neighborhood $V(\boldsymbol{\theta}_0^*)$ of $\boldsymbol{\theta}_0^*$ such that the following variables have finite expectation:

$$\sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0^*)} \left\| \frac{1}{\sigma_t(\boldsymbol{\theta})} \frac{\partial \sigma_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right\|^4, \quad \sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0^*)} \left\| \frac{1}{\sigma_t(\boldsymbol{\theta})} \frac{\partial^2 \sigma_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right\|^2, \quad \sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0^*)} \left| \frac{\sigma_t(\boldsymbol{\theta}_0^*)}{\sigma_t(\boldsymbol{\theta})} \right|^{2\delta}.$$

Note that Assumptions **A2**, **A3**, **A5** and **A6** can be simplified for specific forms of σ_t: for instance if the model is the GARCH (p, q) Model (2.2), **A2** reduces to standard assumptions on the lag polynomials of the volatility and **A3**, **A5**, **A6** can be directly verified. Note also that the only moment assumption on the observed process is the existence of a small moment in **A1**, which is automatically satisfied for standard models such as the classical GARCH(p, q).

Now let the residuals of the QML estimation

$$\hat{\eta}_t = \frac{\epsilon_t}{\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_n)}, \qquad t = 1, \ldots, n,$$

and let ξ_{n,α_i} denote the empirical α_i-quantile of $\hat{\eta}_1, \ldots, \hat{\eta}_n$. Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)'$, $\boldsymbol{\xi}_{n,\boldsymbol{\alpha}} = (\xi_{n,\alpha_1}, \ldots, \xi_{n,\alpha_m})'$ and let $\boldsymbol{\xi}_{\boldsymbol{\alpha}} = (\xi_{\alpha_1}, \ldots, \xi_{\alpha_m})'$ denote the vector of population quantiles.

Remark 2.2 The derivation of the joint asymptotic properties of sample quantiles goes back to Cramér (1946) in the iid case. Different articles have extended these results for the marginal quantiles of stationary processes, under different dependence assumptions. See Dominicy, Hörmann, Ogata and Veredas (2013) and the references therein. We cannot apply their results because $(\hat{\eta}_t)$ is not a stationary process.

The next result gives the joint asymptotic distributions of $(\widehat{\boldsymbol{\theta}}'_n, \boldsymbol{\xi}'_{n,\boldsymbol{\alpha}})$. Let $\boldsymbol{D}_t(\boldsymbol{\theta}) = \sigma_t^{-1}(\boldsymbol{\theta})\partial\sigma_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$.

Theorem 2.1 *Assume $\xi_{\alpha_i} < 0$, for $i = 1, \dots, m$, $E\eta_t^2 = 1$ and $\kappa_4 := E\eta_t^4 < \infty$. Suppose that η_1 admits a density f which is continuous and strictly positive in a neighborhood of ξ_{α_i}, for $i = 1, \dots, m$. Let* **A1**-**A3** *and* **A4**$(\boldsymbol{\theta}_0)$-**A6**$(\boldsymbol{\theta}_0)$ *hold. Then*

$$\begin{pmatrix} \sqrt{n}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right) \\ \sqrt{n}(\boldsymbol{\xi}_{n,\boldsymbol{\alpha}} - \boldsymbol{\xi}_{\boldsymbol{\alpha}}) \end{pmatrix} \overset{\mathcal{L}}{\to} \mathcal{N}(0, \boldsymbol{\Sigma}_\alpha), \qquad \boldsymbol{\Sigma}_\alpha = \begin{pmatrix} \frac{\kappa_4-1}{4}\boldsymbol{J}^{-1} & \boldsymbol{\lambda}'_{\boldsymbol{\alpha}} \otimes \boldsymbol{J}^{-1}\boldsymbol{\Omega} \\ \boldsymbol{\lambda}_{\boldsymbol{\alpha}} \otimes \boldsymbol{\Omega}'\boldsymbol{J}^{-1} & \boldsymbol{\zeta}_{\boldsymbol{\alpha}} \end{pmatrix},$$

where $\boldsymbol{\Omega} = E(\boldsymbol{D}_t)$, $\boldsymbol{J} = E(\boldsymbol{D}_t\boldsymbol{D}'_t)$ with $\boldsymbol{D}_t = \boldsymbol{D}_t(\boldsymbol{\theta}_0)$, $\boldsymbol{\lambda}_{\boldsymbol{\alpha}} = (\lambda_{\alpha_1}, \dots, \lambda_{\alpha_m})'$, $\boldsymbol{\zeta}_{\boldsymbol{\alpha}} = (\zeta_{ij})_{1\le i,j\le m}$ and

$$\begin{aligned} \lambda_{\alpha_i} &= \xi_{\alpha_i}\frac{\kappa_4 - 1}{4} + \frac{p_{\alpha_i}}{2f(\xi_{\alpha_i})}, \\ \zeta_{ij} &= \xi_{\alpha_i}\xi_{\alpha_j}\frac{\kappa_4 - 1}{4} + \frac{\xi_{\alpha_i}p_{\alpha_j}}{2f(\xi_{\alpha_j})} + \frac{\xi_{\alpha_j}p_{\alpha_i}}{2f(\xi_{\alpha_i})} + \frac{(\alpha_i \wedge \alpha_j)(1 - \alpha_i \vee \alpha_j)}{f(\xi_{\alpha_i})f(\xi_{\alpha_j})}, \end{aligned}$$

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

Proof. In view of Francq and Zakoïan (Proof of Theorem 4, 2012), we have, for $i = 1, \dots, m$,

$$\sqrt{n}(\xi_{\alpha_i} - \xi_{n,\alpha_i}) = \xi_{\alpha_i}\boldsymbol{\Omega}'\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) + \frac{1}{f(\xi_{\alpha_i})}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(\mathbf{1}_{\{\eta_t<\xi_{\alpha_i}\}} - \alpha_i) + o_P(1),$$

and

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) = \frac{-\boldsymbol{J}^{-1}}{2\sqrt{n}} \sum_{t=1}^{n} (1-\eta_t^2)\boldsymbol{D}_t + o_P(1).$$

Hence

$$\mathrm{Cov}_{as}\left(\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0), \frac{1}{\sqrt{n}}\sum_{t=1}^{n}(\mathbf{1}_{\{\eta_t<\xi_{\alpha_i}\}} - \alpha_i)\right) = \frac{1}{2}p_{\alpha_i}\boldsymbol{J}^{-1}\boldsymbol{\Omega}.$$

It follows that, for $i \leq j$,

$$\begin{aligned}
& \mathrm{Cov}_{as}\{\sqrt{n}(\xi_{\alpha_i} - \xi_{n,\alpha_i}), \sqrt{n}(\xi_{\alpha_j} - \xi_{n,\alpha_j})\} \\
= & \left\{\xi_{\alpha_i}\xi_{\alpha_j}\frac{\kappa_4 - 1}{4} + \frac{\xi_{\alpha_i}p_{\alpha_j}}{2f(\xi_{\alpha_j})} + \frac{\xi_{\alpha_j}p_{\alpha_i}}{2f(\xi_{\alpha_i})}\right\}\boldsymbol{\Omega}'J^{-1}\boldsymbol{\Omega} + \frac{\alpha_i(1-\alpha_j)}{f(\xi_{\alpha_i})f(\xi_{\alpha_j})}, \\
& \mathrm{Cov}_{as}\left(\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0), \sqrt{n}(\xi_{\alpha_i} - \xi_{n,\alpha_i})\right) \\
= & \lambda_{\alpha_i}\boldsymbol{J}^{-1}\boldsymbol{\Omega}.
\end{aligned}$$

We have $\boldsymbol{\Omega}'\boldsymbol{J}^{-1}\boldsymbol{\Omega} = 1$ (see Remark 3.1 in Francq and Zakoian, 2013) and thus we obtain

$$\mathrm{Cov}_{as}\{\sqrt{n}(\xi_{\alpha_i} - \xi_{n,\alpha_i}), \sqrt{n}(\xi_{\alpha_j} - \xi_{n,\alpha_j})\} = \zeta_{ij}.$$

By the CLT for martingale differences, we get the announced result. □

Let $\mathbf{VaR}_t(\boldsymbol{\alpha}) = (\mathrm{VaR}_t(\alpha_1), \ldots, \mathrm{VaR}_t(\alpha_m))'$, the vector of VaR's at levels α_i. We have

$$\mathbf{VaR}_t(\boldsymbol{\alpha}) = -\sigma(\epsilon_{t-1}, \epsilon_{t-2}, \ldots; \boldsymbol{\theta}_0)\boldsymbol{\xi}_{\boldsymbol{\alpha}}. \tag{2.5}$$

A natural estimator of $\mathbf{VaR}_t(\boldsymbol{\alpha})$ is thus

$$\widehat{\mathbf{VaR}}_t(\boldsymbol{\alpha}) = -\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_n)\boldsymbol{\xi}_{n,\boldsymbol{\alpha}}.$$

Remark 2.3 A classical problem, called quantile crossing, in quantile regression is that two or more estimated conditional quantile functions can cross or overlap. This drawback occurs because each conditional quantile function is independently estimated (see Koenker (2005)). It is thus worth noting that our estimation procedure does not face this problem. By construction, the estimated conditional VaR are monotonous functions of the α's.

Remark 2.4 For the standard GARCH(p,q) model, we have $\boldsymbol{J}^{-1}\boldsymbol{\Omega} = 2\overline{\boldsymbol{\theta}}_0$, where

$$\overline{\boldsymbol{\theta}}_0 = \begin{pmatrix} \boldsymbol{\theta}_0^{[1:q+1]} \\ 0_p \end{pmatrix}, \quad \boldsymbol{\theta}_0^{[1:q+1]} = (\omega_0, a_{01}, \dots, a_{0q})',$$

(see Francq and Zakoian (2013)), and the asymptotic variance in Theorem 2.1 takes the more explicit form

$$\boldsymbol{\Sigma}_{\boldsymbol{\alpha}} = \begin{pmatrix} \frac{\kappa_4-1}{4}\boldsymbol{J}^{-1} & 2\boldsymbol{\lambda}'_{\boldsymbol{\alpha}} \otimes \overline{\boldsymbol{\theta}}_0 \\ 2\boldsymbol{\lambda}_{\boldsymbol{\alpha}} \otimes \overline{\boldsymbol{\theta}}'_0 & \boldsymbol{\zeta}_{\boldsymbol{\alpha}} \end{pmatrix}.$$

2.3 Constructing confidence intervals for the VaR's

Let $\widehat{\boldsymbol{\Sigma}}_\alpha$ denote a consistent estimator of the asymptotic variance $\boldsymbol{\Sigma}_\alpha$. Such an estimator can be constructed by i) replacing $\boldsymbol{J}$ by $\hat{\boldsymbol{J}} = n^{-1}\sum_{t=1}^n \boldsymbol{D}_t(\widehat{\boldsymbol{\theta}}_n)\boldsymbol{D}_t(\widehat{\boldsymbol{\theta}}_n)'$; ii) using the residuals $\hat{\eta}_t$ to construct an estimator $\hat{f}$ of the density function f of the innovation, and to replace the theoretical moments of the process (η_t) by their empirical counterpart.

The delta method thus suggests a $(1-\alpha_0)\%$ confidence interval (CI) for the $\mathrm{VaR}_t(\alpha_i)$ whose bounds are

$$-\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n,\alpha_i})\xi_{n,\alpha_i} \pm \frac{\Phi^{-1}_{1-\alpha_0/2}}{\sqrt{n}} \left\{ \left(\widehat{\boldsymbol{\Delta}}_\alpha \widehat{\boldsymbol{\Sigma}}_\alpha \widehat{\boldsymbol{\Delta}}'_\alpha \right)_{ii} \right\}^{1/2}, \tag{2.6}$$

where

$$\widehat{\boldsymbol{\Delta}}_\alpha = \left(\boldsymbol{\xi}_{n,\boldsymbol{\alpha}} \frac{\partial \tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n,\alpha})}{\partial \boldsymbol{\theta}'}, \tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_n)\boldsymbol{I}_m \right),$$

$\Phi_{\alpha_0}^{-1}$ denotes the α_0-quantile of the standard Gaussian distribution, and $\boldsymbol{I}_m$ denotes the $m \times m$ identity matrix. Note that the choice of α_0 (the risk estimation level) is independent from that of the α_i's (the financial risk levels). Drawing such CI allows to underline the importance of the estimation risk for VaR evaluation.

2.4 A portfolio of VaR's

Focusing only on VaR at a given level for measuring risk can be misleading since it gives a limited view of the distribution, which may result in lack of robustness for risk management and risk control. To circumvent this problem, several risk measures have to be jointly considered in practice. To this aim, Distortion Risk Measures (DRM) have been introduced in the insurance literature, in a series of papers by Wang and coauthors [see Wang (2000) and the references therein]. A particular case is the conditional expected shortfall (ES) which, at level $\alpha \in (0,1)$, can be written as

$$\mathrm{ES}_t(\alpha) = -E_{t-1}[\epsilon_t \mid \epsilon_t < -\mathrm{VaR}_t(\alpha)] = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_t(u)du.$$

More general DRM take the form

$$\mathrm{DRM}_t = \int_0^1 \mathrm{VaR}_t(u)dG(u), \tag{2.7}$$

where the distortion function, G, is a given cumulative distribution function (cdf) on $[0,1]$[3]. It follows from (2.3) that, for Model (2.1),

$$\mathrm{DRM}_t = -\sigma(\epsilon_{t-1}, \epsilon_{t-2}, \ldots; \boldsymbol{\theta}_0)\int_0^1 \xi_u dG(u). \tag{2.8}$$

[3] Examples of DRM are the Proportional Hazard DRM, defined with $G(u) = u^r$, and the Exponential DRM defined with $G(u) = (1 - e^{ru})/(1 - e^r)$, both of them defined for $r > 0$. These distortion functions are concave for $0 < r < 1$ and $r > 0$, respectively, which corresponds to coherent risk measure in the sense of Artzner, Delbaen, Eber and Heath (1999) [see e.g. Wirch and Hardy (1999)].

In the spirit of DRM, a risque measure which can be interpreted as a portfolio of VaR's at different levels is defined by

$$\boldsymbol{p}'\mathbf{VaR}_t(\boldsymbol{\alpha}) = \sum_{i=1}^{m} p_i \mathrm{VaR}_t(\alpha_i)$$

where $\boldsymbol{p} = (p_1, \ldots, p_m)$ with $p_i \geq 0$ for $i = 1, \ldots, m$ and $\sum_{i=1}^{m} p_i = 1$. This risk measure can be interpreted as a special DRM with associated distortion function corresponding to Dirac masses at the points α_i. In view of (2.6), an asymptotic CI at level α_0 for this risk measure is

$$-\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n,\alpha_i})\boldsymbol{p}'\boldsymbol{\xi}_{n,\boldsymbol{\alpha}} \pm \frac{\Phi^{-1}_{1-\alpha_0/2}}{\sqrt{n}} \left\{\boldsymbol{p}'\widehat{\boldsymbol{\Delta}}_\alpha \widehat{\boldsymbol{\Sigma}}_\alpha \widehat{\boldsymbol{\Delta}}'_\alpha \boldsymbol{p}\right\}^{1/2}. \tag{2.9}$$

2.5 Choosing the weights to approximate DRMs

An estimator of the DRM in (2.8) can be constructed as follows:

$$\widehat{\mathrm{DRM}}_t = -\sigma(\epsilon_{t-1}, \epsilon_{t-2}, \ldots; \widehat{\boldsymbol{\theta}}_n) \sum_{i=1}^{n} \left\{G\left(\frac{i}{n}\right) - G\left(\frac{i-1}{n}\right)\right\} \hat{\eta}_{n,i}, \tag{2.10}$$

where $(\hat{\eta}_{n,n-i})$ denotes the order statistics, obtained by ranking the $\hat{\eta}_t$ in ascending order: $\hat{\eta}_{n,1} < \cdots < \hat{\eta}_{n,n}$.

However, deriving the asymptotic distribution of this estimator, might be a formidable task. To our knowledge such results do not exist in the literature. In this section, we use VaR portfolios to obtain lower and upper bounds for a class of DRM, leading to (approximate) asymptotic CI's for such DRM.

It is not restrictive to assume $\alpha_1 < \alpha_2 < \ldots < \alpha_m$. Suppose that the support of the distortion cdf G is $[\alpha_1, \alpha_m]$, that is

$$\mathrm{DRM}_t = \int_{\alpha_1}^{\alpha_m} \mathrm{VaR}_t(u) dG(u). \tag{2.11}$$

In other words, we focus on "moderate risks": we do not consider extreme risks, corresponding to values of α approaching 0. An example of class of such DRM, parameterized by the coefficient $r > 0$ and adapted from the so-called "proportional hazard" DRM, is defined by

$$G(u) = \left(\frac{u - \alpha_1}{\alpha_m - \alpha_1}\right)^r \mathbf{1}_{u \in (\alpha_1, \alpha_m)} + \mathbf{1}_{u \in (\alpha_m, 1)}, \tag{2.12}$$

where $\mathbf{1}_A$ denotes the indicator function of any set A.

Lower and upper bounds for the DRM in (2.11), can be constructed as follows. Because $u \mapsto \mathrm{VaR}_t(u)$ is decreasing we have, noting that $G(\alpha_1) = 0$ and $G(\alpha_m) = 1$,

$$\boldsymbol{p}_L' \mathbf{VaR}_t(\boldsymbol{\alpha}) \leq \mathrm{DRM}_t(\alpha) \leq \boldsymbol{p}_U' \mathbf{VaR}_t(\boldsymbol{\alpha})$$

where

$$\begin{aligned} \boldsymbol{p}_L &= (0, G(\alpha_2), G(\alpha_3) - G(\alpha_2), \dots, 1 - G(\alpha_{m-1}))\,, \\ \boldsymbol{p}_U &= (G(\alpha_2), G(\alpha_3) - G(\alpha_2), \dots, 1 - G(\alpha_{m-1}), 0)\,. \end{aligned}$$

It follows that a CI at significance level $\alpha_0^* \leq \alpha_0$ for this risk measure is

$$\Bigg[-\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n,\alpha_i}) \boldsymbol{p}_L' \boldsymbol{\xi}_{n,\boldsymbol{\alpha}} - \frac{\Phi^{-1}_{1-\alpha_0/2}}{\sqrt{n}} \left\{ \boldsymbol{p}_L' \widehat{\boldsymbol{\Delta}}_\alpha \widehat{\boldsymbol{\Sigma}}_\alpha \widehat{\boldsymbol{\Delta}}_\alpha' \boldsymbol{p}_L \right\}^{1/2}, \\ -\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n,\alpha_i}) \boldsymbol{p}_U' \boldsymbol{\xi}_{n,\boldsymbol{\alpha}} + \frac{\Phi^{-1}_{1-\alpha_0/2}}{\sqrt{n}} \left\{ \boldsymbol{p}_U' \widehat{\boldsymbol{\Delta}}_\alpha \widehat{\boldsymbol{\Sigma}}_\alpha \widehat{\boldsymbol{\Delta}}_\alpha' \boldsymbol{p}_U \right\}^{1/2} \Bigg]. \tag{2.13}$$

3 NonGaussian QML estimation of VaR's

In this section we develop an alternative method for estimating the conditional VaR's. This method is based on a reparameterization of model (2.1). QML inferences based on similar reparameterizations were proposed by Francq, Lepage and Zakoian (2011), Fan, Qi and Xiu (2012), Francq and Zakoian (2013).

3.1 Reparameterization and VaR parameter

The approach of this section requires the following assumption

A7: There exists a function H such that for any $\boldsymbol{\theta} \in \Theta$, for any $K > 0$, and any sequence $(x_i)_i$

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

This assumption is not very restrictive as it is satisfied by all commonly used GARCH-type formulations, in particular those mentioned in Section 2. It means that scaling the volatility is equivalent to a change of parameter. In general, the new parameter satisfies $\boldsymbol{\theta}^* \geq \boldsymbol{\theta}$, componentwise, when $K \geq 1$. For instance, in the GARCH(p, q) model (2.2) we have $\boldsymbol{\theta}^* = (K^2\omega, K^2 a_1, \dots, K^2 a_q, b_1, \dots, b_p)'$.

In view of (2.3), we have under **A7**, provided α_i is small enough so that $-\xi_{\alpha_i} > 0$,

$$\mathrm{VaR}_t(\alpha_i) = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_{0\alpha_i}) \tag{3.1}$$

where $\boldsymbol{\theta}_{0,\alpha_i} = H(\boldsymbol{\theta}_0, -\xi_{\alpha_i})$. This parameter depends on both the dynamics of the GARCH process, through the volatility parameters, and the innovations distribution through the α-quantile. It is called VaR-parameter in Francq and Zakoïan (2012) (hereafter FZ). Similarly, if $-\int_0^1 \xi_u dG(u) > 0$, the DRM in (2.8) can be written as

$$\mathrm{DRM}_t = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_0^G) \tag{3.2}$$

where $\boldsymbol{\theta}_0^G$ is a DRM-parameter defined by

$$\boldsymbol{\theta}_0^G = H\left(\boldsymbol{\theta}_0, -\int_0^1 \xi_u dG(u)\right). \tag{3.3}$$

It follows from (3.1) that, with the notation used in (2.5),

$$\mathbf{VaR}_t(\boldsymbol{\alpha}) = \begin{pmatrix} \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_{0\alpha_1}) \\ \vdots \\ \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_{0\alpha_m}) \end{pmatrix}.$$

The approach, in this section, consists in estimating by QML the $\boldsymbol{\theta}_{0\alpha_i}$'s instead of $\boldsymbol{\theta}_0$. The idea is to interpret, for $i = 1, \dots, m$, the VaR-parameter $\boldsymbol{\theta}_{0\alpha_i}$ as a volatility parameter in a reparameterized model. We note that

$$\epsilon_t = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_{0\alpha_i})\eta_{i,t}, \qquad \text{where } \eta_{i,t} = \frac{\eta_t}{-\xi_{\alpha_i}}.$$

The problem is thus to estimate by QML the model

$$\begin{cases} \epsilon_t = \sigma_{i,t}\eta_{i,t}, \quad P[\eta_{i,t} < -1] = \alpha_i, \\ \sigma_{i,t} = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_{0,\alpha_i}). \end{cases} \tag{3.4}$$

Note that the Gaussian QML cannot be employed because it requires the assumption that $E\eta_t^2 = 1$. FZ derived the asymptotic distribution of the non-Gaussian QMLE of $\boldsymbol{\theta}_{0,\alpha_i}$ defined by

$$\widehat{\boldsymbol{\theta}}_{n,\alpha_i} = \arg\max_{\boldsymbol{\theta}\in\Theta} \sum_{t=1}^{n} \log \frac{1}{\tilde{\sigma}_t(\boldsymbol{\theta})} h_{\alpha_i}\left(\frac{\epsilon_t}{\tilde{\sigma}_t(\boldsymbol{\theta})}\right) \tag{3.5}$$

where h_{α_i} is given by

$$h_{\alpha_i}(x) = \lambda\alpha_i(1-2\alpha_i)|x|^{2\lambda\alpha_i - 1}\{|x|^{-\lambda}\mathbf{1}_{\{|x|>1\}} + \mathbf{1}_{\{|x|\le 1\}}\} \tag{3.6}$$

for some (unimportant) positive constant λ.

As noted by FZ, the non-Gaussian QML estimator in (3.5) can be interpreted as a nonlinear quantile regression estimator. Letting $\rho_\alpha(u) = u(\alpha - \mathbf{1}_{\{u\le 0\}})$, for $\alpha \in (0,1)$, we have

$$\widehat{\boldsymbol{\theta}}_{n,\alpha_i} \quad = \quad \arg\min_{\boldsymbol{\theta}\in\Theta} \frac{1}{n} \sum_{t=1}^{n} \rho_{1-2\alpha_i}\left\{\log\left(\frac{|\epsilon_t|}{\tilde{\sigma}_t(\boldsymbol{\theta})}\right)\right\}.$$

In the next section, we derive the joint distribution of the $\widehat{\boldsymbol{\theta}}_{n,\alpha_i}$'s.

3.2 Asymptotic joint distribution of the VaR parameter estimators

We introduce the following additional assumption.

A8: The density f of η_0 is symmetric, continuous and strictly positive at the points ξ_{α_i}, for $i = 1, \dots, m$, and satisfies $M = \sup_{x \in \mathbb{R}} |x| f(x) < \infty$. Moreover $E|\log|\eta_0|| < \infty$.

Let $\boldsymbol{\theta}_{0\boldsymbol{\alpha}} = (\boldsymbol{\theta}'_{0\alpha_1}, \dots, \boldsymbol{\theta}'_{0,\alpha_m})'$ and let $\widehat{\boldsymbol{\theta}}_{n,\boldsymbol{\alpha}} = (\widehat{\boldsymbol{\theta}}'_{n,\alpha_1}, \dots, \widehat{\boldsymbol{\theta}}'_{n,\alpha_m})'$.

Theorem 3.1 *Under the assumptions* **A1-A3***,* **A7***,* **A8** *and if, for $i = 1, \dots, m$, $\alpha_i \in (0, 1/2)$ and* **A4**$(\boldsymbol{\theta}_{0\alpha_i})$-**A6**$(\boldsymbol{\theta}_{0\alpha_i})$ *hold, there exists a sequence of local minimizers $\widehat{\boldsymbol{\theta}}_{n,\boldsymbol{\alpha}}$ of the QML criterion satisfying*

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_{n,\boldsymbol{\alpha}} - \boldsymbol{\theta}_{0,\boldsymbol{\alpha}}) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Xi}_{\boldsymbol{\alpha}}),$$

where $\boldsymbol{\Xi}_{\boldsymbol{\alpha}}$ is a $md \times md$ matrix whose (i,j)-block of size $d \times d$ is

$$\boldsymbol{\Xi}_{\boldsymbol{\alpha}}[i,j] = \frac{2\alpha_i \wedge \alpha_j (1 - 2\alpha_i \vee \alpha_j)}{4 f(\xi_{\alpha_i}) f(\xi_{\alpha_j})} \boldsymbol{J}^{-1}_{\alpha_i\alpha_i} \boldsymbol{J}_{\alpha_i\alpha_j} \boldsymbol{J}^{-1}_{\alpha_j\alpha_j}$$

with $\boldsymbol{J}_{\alpha_i\alpha_j} = E\boldsymbol{D}_t(\boldsymbol{\theta}_{0,\alpha_i})\boldsymbol{D}'_t(\boldsymbol{\theta}_{0,\alpha_j})$.

Proof. Note that $\hat{\boldsymbol{\nu}}_{n,\alpha_i} := \sqrt{n}(\widehat{\boldsymbol{\theta}}_{n,\alpha_i} - \boldsymbol{\theta}_{0,\alpha_i})$ is such that

$$\hat{\boldsymbol{\nu}}_{n,\alpha_i} = \arg\min_{\boldsymbol{\nu} \in \Lambda_{n,\alpha_i}} \tilde{S}_{n,\alpha_i}(\boldsymbol{\nu}),$$

where $\Lambda_{n,\alpha_i} := \sqrt{n}(\Theta - \boldsymbol{\theta}_{0,\alpha_i})$ and

$$\tilde{S}_{n,\alpha_i}(\boldsymbol{\nu}) = \sum_{t=1}^{n} \rho_{1-2\alpha_i} \left\{ \log\left(\frac{|\epsilon_t|}{\tilde{\sigma}_t(\boldsymbol{\theta}_{0,\alpha_i} + n^{-1/2}\boldsymbol{\nu})} \right) \right\} - \rho_{1-2\alpha_i} \left\{ \log\left(\frac{|\epsilon_t|}{\tilde{\sigma}_t(\boldsymbol{\theta}_{0,\alpha_i})} \right) \right\}.$$

For notational convenience, write $a \overset{c}{=} b$ when $a = b + c$. Showing that the initial values are asymptotically negligible, and noting that $\epsilon_t/\sigma_t(\boldsymbol{\theta}_{0,\alpha_i}) =$

$-\eta_t/\xi_{\alpha_i}$, it can be proven that, uniformly in ν belonging to an arbitrary compact set (see Lemma 2 in FZ),

$$\tilde{S}_{n,\alpha_i}(\boldsymbol{\nu}) \overset{o_P(1)}{=} S_{n,\alpha_i}(\boldsymbol{\nu}) \quad := \quad \sum_{t=1}^{n} \rho_{1-2\alpha_i} \left\{ \log \left(\frac{|\epsilon_t|}{\sigma_t(\boldsymbol{\theta}_{0,\alpha_i} + n^{-1/2}\boldsymbol{\nu})} \right) \right\} \\ -\rho_{1-2\alpha_i} \left\{ \log \left| \frac{\eta_t}{\xi_i} \right| \right\}.$$

Doing a Taylor expansion of $\log \sigma_t(\boldsymbol{\theta}_{0,\alpha_i} + n^{-1/2}\boldsymbol{\nu})$ around $\boldsymbol{\nu} = \mathbf{0}$, and using Lemma 2 in FZ, we obtain

$$S_{n,\alpha_i}(\boldsymbol{\nu}) \overset{o_P(1)}{=} S^*_{n,\alpha_i}(\boldsymbol{\nu}) \quad := \quad \sum_{t=1}^{n} \rho_{1-2\alpha_i} \left\{ \log \left| \frac{\eta_t}{\xi_{\alpha_i}} \right| - \frac{1}{\sqrt{n}} \boldsymbol{\nu}' \boldsymbol{D}_t(\boldsymbol{\theta}_{0,\alpha_{\alpha_i}}) \right\} \\ -\rho_{1-2\alpha_i} \left\{ \log \left| \frac{\eta_t}{\xi_{\alpha_i}} \right| \right\}.$$

Note that $S^*_{n,\alpha_i}(\cdot)$ is equal to the function $Z_n(\cdot)$ defined by Equation (17) in Koenker and Xiao (2006), when applied to the quantile regression of $\log |\eta_t/\xi_{\alpha_i}|$ on $\boldsymbol{D}_t(\boldsymbol{\theta}_{0,\alpha_{\alpha_i}})$ at the level $1 - 2\alpha_i$. Even if our framework is not that of the above-mentioned paper, similar results hold true. More precisely, FZ show that the finite-dimensional distributions of $S^*_{n,\alpha_i}(\boldsymbol{\nu})$ and

$$S^{**}_{n,\alpha_i}(\boldsymbol{\nu}) := -\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \boldsymbol{\nu}' \boldsymbol{D}_t(\boldsymbol{\theta}_{0,\alpha_i}) \left(1 - 2\alpha_i - \mathbf{1}_{\{|\eta_t| < -\xi_{\alpha_i}\}} \right) + f(\xi_{\alpha_i}) \boldsymbol{\nu}' \boldsymbol{J}_{\alpha_i\alpha_i} \boldsymbol{\nu}$$

converge to those of the same Gaussian process. Noting that the trajectories of $S^*_{n,\alpha_i}(\cdot)$ and $S^{**}_{n,\alpha_i}(\cdot)$ are convex, we also have uniform convergence over every compact set in the space of the continuous function on $\mathbb{R}^d$. By Lemma 2.2 in Davis, Knight and Liu (1992) the minima of $S^*_{n,\alpha_i}(\cdot)$ and $S^{**}_{n,\alpha_i}(\cdot)$ are asymptotically the same. By Remark 1 of the above-mentioned paper, we finally obtain

$$\hat{\boldsymbol{\nu}}_{n,\alpha_i} \overset{o_P(1)}{=} \frac{1}{2f(\xi_{\alpha_i})} \boldsymbol{J}^{-1}_{\alpha_i\alpha_i} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \boldsymbol{D}_t(\boldsymbol{\theta}_{0,\alpha_i}) \left(1 - 2\alpha_i - \mathbf{1}_{\{|\eta_t| < -\xi_{\alpha_i}\}} \right).$$

The conclusion follows easily. $\square$

Remark 3.1 Theorems 2.1 and 3.1 provide the asymptotic distributions of two estimators for VaR portfolios. At first sight, the method of this section is not attractive because it is more cumbersome, from a numerical point of view, than that of the previous section. Indeed, it requires the optimization of m QML criteria, whereas the first method requires one. However, it is important to note that the assumptions required for the asymptotic results are different. In particular, the fourth moment assumption $E\eta_t^4 < \infty$ of the first method, is not required in Theorem 3.1. On the other hand, the latter theorem is valid under a symmetry assumption on the noise distribution. To conclude, the method of this section can only be recommended in presence of very heavy-tailed errors distribution.

4 Empirical illustration

In this section we present empirical results using returns of nine major stock indices: CAC (Paris), DAX (Frankfurt), FTSE (London), Nikkei (Tokyo), NSE (Bombay), SMI (Switzerland), SP500 (New York), SPTSX (Toronto), and SSE (Shanghai). Our sample spans the period from January, 2 1991 to August, 26 2011 (but all series are not available for the whole period, see Table 1 for the sample sizes). For each series of log-returns, $\epsilon_t = \log(p_t/p_{t-1})$ where p_t denotes the value of the index, we used a GARCH(1,1) model for the volatility dynamics. We estimated the DRM parameter $\boldsymbol{\theta}_0^G = (\omega^G, a^G, b^G)$, defined in (3.3), with $r = 1/2$, $\alpha_1 = 0.01$ and $\alpha_m = 0.1$ for the DRM function G defined in (2.12). In view of (2.10) and (3.3), the DRM-parameter estimator is given by

$$\widehat{\boldsymbol{\theta}}_n^G = H\left(\widehat{\boldsymbol{\theta}}_n, -\sum_{i=1}^n \left\{G\left(\frac{i}{n}\right) - G\left(\frac{i-1}{n}\right)\right\}\hat{\eta}_{n,i}\right),$$

where $H(\omega, \alpha, \beta; K) = (K^2\omega, K^2\alpha, \beta)$. The CI's are obtained using (2.13) with $m = 20$ and $\alpha_0 = 5\%$.

We report in Table 1 our estimates of the conditional DRM parameter and the corresponding CI's. Caution is needed in the interpretation of this table because the DRM parameter is not the usual volatility parameter. In particular, the fact that $a^G + b^G > 1$ is not in contradiction with the usual empirical finding, $a + b \approx 1$, for GARCH(1,1) models. Noticeable differences appear between these series, particularly for the coefficients ω^G and a^G and their CI's. Replacing the number $m = 20$ of α_i's used for the discretization of the DRM by $m = 10$ or $m = 30$ left almost unchanged the CI's, so we did not report the results. We can depict three categories of assets: i) the FTSE, SP500 and SPTSX display similar coefficients, relatively small ω^G's, large persistence parameter b's, small CI's; ii) Nikkei, NSE and SMI provide, by comparison, larger ω^G's and a^G's, smaller persistence and much larger CI's; iii) the CAC, DAX and SSE display intermediate results. Examination of the CI's shows that the differences between parameters of series in groups i) and ii) are statistically significant. Note also that larger CI's are not always due to smaller sample size.

Figure 1 displays the returns, estimated -VaR (at the 10% and 1% levels), -DRM, and their accuracy intervals for the DAX index from April, 8, 2011 to August, 26, 2011. The $(1 - \alpha_0)\%$ confidence intervals (for $\alpha_0 = 5\%$) are obtained from formula (2.13). We reported the opposite of the conditional risks (VaR and DRM), because in terms of capital reserves, only large negative returns matter. As expected, the accuracy on VaR estimation decreases when the risk α approaches 0. Interestingly, the accuracy of the DRM is comparable to that of the VaR's, despite the more sophisticated construction of this measure of risk. Note also that, in turbulent periods, both the

Table 1: Estimation of the conditional DRM parameter for 9 stock market indices. The approximate 95% confidence intervals are displayed into brackets.

Index	n	ω^G	a^G	b
CAC	5229	0.11 [0.05,0.17]	0.31[0.22,0.41]	0.90 [0.88,0.92]
DAX	5226	0.12 [0.04,0.20]	0.31[0.18,0.45]	0.90 [0.86,0.93]
FTSE	5217	0.04 [0.02,0.07]	0.32[0.24,0.41]	0.91 [0.89,0.92]
Nikkei	5078	0.20 [0.11,0.30]	0.37[0.26,0.48]	0.88 [0.85,0.91]
NSE	2265	0.25 [0.06,0.46]	0.40[0.20,0.65]	0.87 [0.82,0.92]
SMI	5209	0.17 [0.08,0.27]	0.46[0.27,0.65]	0.84 [0.79,0.89]
SP500	5206	0.03 [0.01,0.05]	0.27[0.19,0.36]	0.92 [0.90,0.94]
SPTSX	2934	0.03 [0.01,0.06]	0.27[0.17,0.38]	0.93 [0.91,0.95]
SSE	2982	0.11 [0.03,0.20]	0.25[0.15,0.37]	0.93 [0.90,0.95]

market risks, as measured by the VaR's or the DRM, and the estimation risks, as measured by the CI's, increase.

5 Conclusion

In this paper, we proposed procedures for joint statistical inference on the VaR's at different levels, in the framework of conditionally heteroskedastic models. We also introduced an approximation of general DRM based on a finite number of VaR's. Our empirical analysis showed that confidence intervals based on this measure of risk have similar magnitude as those obtained for VaR's.

One alternative for deriving the asymptotic distribution of the DRM

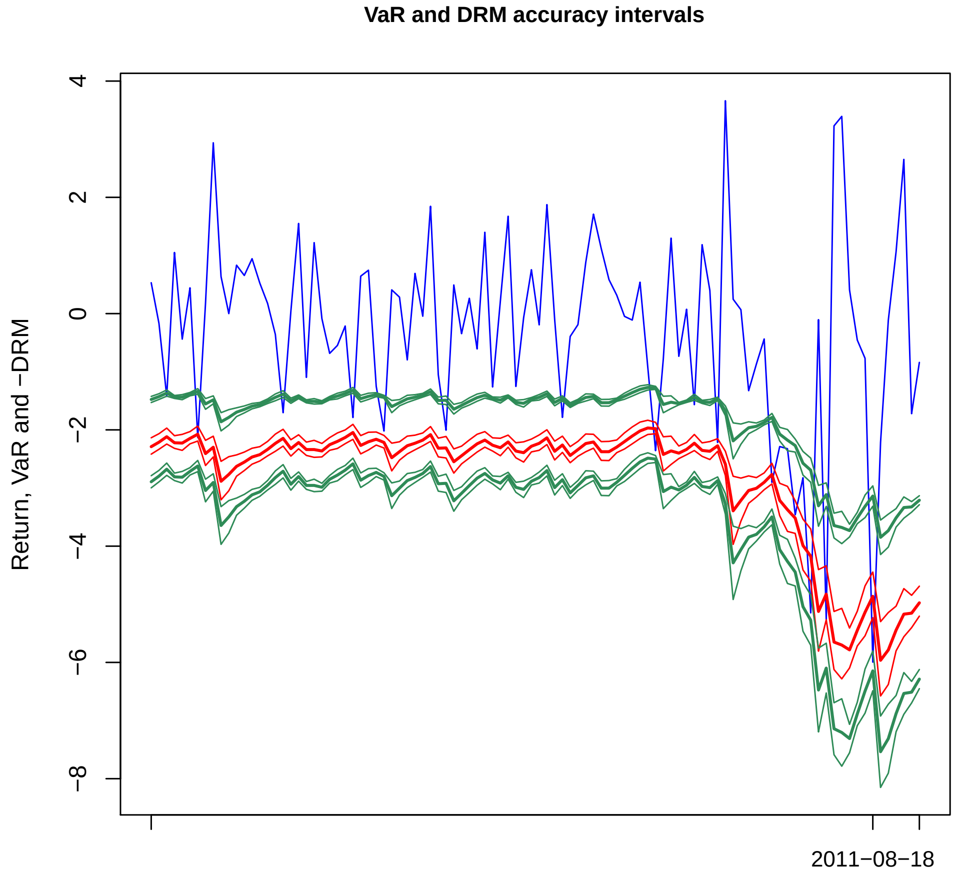

Figure 1: Returns (in blue), estimated -VaR (at the 10% and 1% levels, in green), -DRM (in red), and CI's of the VaR's and DRMs, for the DAX index from April, 8, 2011 to August, 26, 2011. Estimation of the volatility and risk parameters is based on the 1000 previous values.

estimator would be to establish a functional CLT, in function of α, for the vector of the volatility parameter estimator and the empirical quantile of the residuals. Deriving this asymptotic distribution could be a formidable challenge. Moreover, the asymptotic distribution would certainly be non explicit. The approximation proposed in this article, which provides an explicit and easily estimable asymptotic distribution, thus has the advantage of simplicity.

One object of this study was also to draw attention on the estimation risk, in other words the effects of parameter estimation on the accuracy of VaR's evaluations. We showed that estimation risk can be explicitly taken into account, leading to confidence bounds for portfolios, or more generally any smooth function, of VaR's. For risk management purposes, or from a regulation point of view, such confidence intervals could be used to increase the capital reserve in order to account for the underlying estimation uncertainty.

References

Artzner, P., Delbaen, F., Eber, J-M. and D. Heath (1999) Coherent measures of risk. *Mathematical Finance* 9, 203–228.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307–327.

Chan, N.H., Deng, S.J., Peng, L. and Z. Xia (2007) Interval estimation of value-at-risk based on GARCH models with heavy-tailed innovations. *Journal of Econometrics* 137, 556–576.

Cramér, H. (1946) *Mathematical Methods of Statistics*. Princeton Mathematical Series, vol.9. Princeton University Press, Princeton.

Davis, R.A., Knight, K. and J. Liu (1992) M-estimation for autoregressions with infinite variance. *Stochastic Processes and their Applications* 40, 145–180.

Ding, Z., Granger C. et R.F. Engle (1993) A long memory property of stock market returns and a new model. *Journal of Empirical Finance* 1, 83–106.

Dominicy, Y., Hörmann, S., Ogata, H., and D. Veredas (2013). Marginal Quantiles for StationaryProcesses. *Statistics and Probability Letters* 83, 28–36.

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

Engle, R.F. and S. Manganelli (2004) CAViaR: Conditional Value at risk by Quantile Regression. *Journal of Business and Economic Statistics* 22, 367–381.

Fan, J., Qi, L. and D. Xiu (2012) Quasi Maximum Likelihood Estimation of GARCH Models with Heavy-Tailed Likelihoods. Discussion Paper, available at `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1540363`.

Francq, C., Lepage, G. and J-M. Zakoïan (2011) Two-stage non Gaussian QML estimation of GARCH Models and testing the efficiency of the Gaussian QMLE. *Journal of Econometrics* 165, 246–257.

Francq, C. and J-M. Zakoïan (2010) *GARCH Models: Structure, Statistical Inference and Financial Applications.* John Wiley.

Francq, C. and J-M. Zakoïan (2012) Risk-parameter estimation in volatility models. MPRA Preprint No. 41713.

Francq, C. and J.M. Zakoïan (2013) Optimal predictions of powers of conditionally heteroskedastic processes. *Journal of the Royal Statistical Society - Series B* 75, 345–367.

Geweke, J. (1986) Modeling the persistence of conditional variances: a comment. *Econometric Review* 5, 57–61.

Glosten, L.R., Jaganathan, R. et D. Runkle (1993) On the relation between the expected values and the volatility of the nominal excess return on stocks. *Journal of Finance* 48, 1779–1801.

Gouriéroux, C. and J. Jasiak (2008) Dynamic Quantile Models. *Journal of Econometrics* 147, 198–205.

Knight, K. (1998) Limiting distributions for L_1 regression estimators under general conditions. *The Annals of Statistics* 26, 755–770.

Koenker, R. (2005) *Quantile Regression.* Cambridge: Cambridge University Press.

Koenker, R. and Z. Xiao (2006) Quantile autoregression. *Journal of the American Statistical Association* 101, 980–990.

Kuester, K., Mittnik, S. and M.S. Paolella (2006) Value-at-Risk predictions: A comparison of alternative strategies. *Journal of Financial Econometrics* 4, 53–89.

McNeil, A. J., Frey, R. and P. Embrechts (2005) *Quantitative Risk Management.* Princeton University Press.

Milhøj, A. (1987) A multiplicative parameterization of ARCH Models. *Working paper*, Department of Statistics, University of Copenhagen.

Nelson D.B. (1991) Conditional Heteroskedasticity in Asset Returns : a New Approach. *Econometrica* 59, 347–370.

Pantula, S.G. (1986) Modeling the persistence of conditional variances: a comment. *Econometric Review* 5, 71–74.

Pollard, D. (1991) Asymptotics for Least Absolute Deviation Regression Estimators *Econometric Theory* 7, 186–199.

Sentana, E. (1995) Quadratic ARCH Models. *Review of Economic Studies* 62, 639–661.

Spierdijk, L. (2013) Confidence intervals for ARMA-GARCH Value-at-Risk. *Working paper*, University of Groningen.

Wang, S. (2000) A class of distortion operators for pricing financial and insurance risks. *Journal of Risk and Insurance* 67, 15–36.

Wirch, J. L. and M. R. Hardy (1999) A Synthesis of Risk Measures for Capital Adequacy. *Insurance: Mathematics and Economics* 25, 337–347.

Zakoïan J.M. (1994) Threshold Heteroskedastic Models. *Journal of Economic Dynamics and Control*, 18, 931–955.