



Munich Personal RePEc Archive

# Looking for efficient qml estimation of conditional value-at-risk at multiple risk levels

Francq, Christian and Zakoian, Jean-Michel

CREST, Université Lille 3

October 2015

Online at <https://mpra.ub.uni-muenchen.de/67195/>  
MPRA Paper No. 67195, posted 17 Oct 2015 11:18 UTC

# Looking for efficient QML estimation of conditional VaRs at multiple risk levels

CHRISTIAN FRANCO\* AND JEAN-MICHEL ZAKOÏAN†

## Abstract

We consider joint estimation of conditional Value-at-Risk (VaR) at several levels, in the framework of general GARCH-type models. The conditional VaR at level  $\alpha$  is expressed as the product of the volatility and the opposite of the  $\alpha$ -quantile of the innovation. A standard method is to estimate the volatility parameter by Gaussian Quasi-Maximum Likelihood (QML) in a first step, and to use the residuals for estimating the innovations quantiles in a second step. We argue that the Gaussian QML may be inefficient with respect to more general QML and can even be in failure for heavy tailed conditional distributions. We therefore study, for a vector of risk levels, a two-step procedure based on a generalized QML. For a portfolio of VaR's at different levels, confidence intervals accounting for both market and estimation risks are deduced. An empirical study based on stock indices illustrates the theoretical results.

*JEL Classification:* C13, C22 and C58.

*Keywords:* Asymmetric Power GARCH, Distortion Risk Measures, Estimation risk, Non-Gaussian Quasi-Maximum Likelihood, Value-at-Risk.

## 1 Introduction

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 associated 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 risk estimates remain reliable in very turbulent periods.

To this aim, the recent econometric literature on risk has focused on the concept of *estimation risk*. Whatever the risk measure, it depends on unknown characteristics of the loss distribution

---

\*CREST and University Lille 3, BP 60149, 59653 Villeneuve d'Ascq cedex, France. E-Mail: christian.franco@univ-lille3.fr

†Corresponding author: Jean-Michel Zakoïan, University Lille 3 and CREST, 15 boulevard Gabriel Péri, 92245 Malakoff Cedex, France. E-mail: zakoian@ensae.fr, Phone number: 33.1.41.17.77.25. The second author gratefully thanks the IDR ACP "Régulation et risques systémiques" for financial support.

which, for practical use, have to be estimated. For instance, the Value-at-Risk (VaR) at a given level  $\alpha$ , can be defined as the opposite of the  $\alpha$ -quantile of the loss distribution. In advanced approaches of risk measurement, the statistical framework is complicated by the dynamic nature of the loss variables. The corresponding risk measures have to be considered conditional on the past losses, and are therefore called conditional risk measures.

In recent research, different approaches were proposed to account for the presence of estimation risk in conditional risk measurement. Chan, Deng, Peng, Xia (2007) constructed confidence intervals for conditional VaRs under the assumption that the errors have heavy tails, using the Extreme-Value Theory, while Spierdijk (2013) proposed a residual subsample bootstrap approach. A bootstrap testing procedure, for the equality of conditional VaRs in a multivariate setting, was recently studied by Hurlin, Laurent, Quaadvlieg and Smeekes (2013). Francq and Zakoïan (2015) 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. They derive an asymptotic theory for a Gaussian Quasi-Maximum Likelihood (QML) of this parameter. For the same problem, a more general approach based on non-Gaussian QMLs was studied by El Ghourabi, Francq and Telmoudi (2015). Gouriéroux and Zakoïan (2013) investigated the bias induced by estimation in the coverage probabilities associated with VaR. Escanciano and Olmo (2010) studied the effect of estimation on backtesting VaR.

In the inference of GARCH-type models, recent articles underlined the possible efficiency loss of the QML estimator (QMLE) due the use of an inappropriate Gaussian error distribution (see Berkes and Horváth (2004), Francq, Lepage and Zakoïan (2011), Francq and Zakoïan (2013), Fan, Qi and Xiu (2014)). In the present paper, we study the estimation of a vector of conditional VaRs based on generalized QMLEs of the volatility. We extend the article by Francq and Zakoïan (2014) devoted to the Gaussian QML by considering QML criteria based on "instrumental densities" which, in general, will not coincide with the errors distribution. We consider a general GARCH-type framework which does not impose a specific form for the volatility. Our approach is aimed at, not only providing VaR estimates, but also confidence intervals based on asymptotic results. The introduction of several risk levels provides a better account of the tail properties of the conditional loss distribution. The VaRs at different levels can also be combined to construct a *portfolio* of VaR's, which can be interpreted as a Distortion Risk Measure (DRM). Deriving confidence intervals for such a portfolio of VaRs

is the main aim of the paper. We also show how efficiency gains can be reached by selecting an appropriate instrumental density.

This paper is structured as follows. In Section 2, we derive the asymptotic joint distribution of the generalized QMLE of the volatility parameter, and a vector of empirical quantiles of the residuals. In Section 3, we deduce asymptotic confidence intervals for the VaR portfolios. The choice of an optimal criterion is also discussed. An empirical illustration based on major stock indices is proposed in Section 4. Section 5 concludes.

## 2 Non-Gaussian QMLE of vectors of VaRs

### 2.1 Conditional VaR in a general model

GARCH-type models are arguably the most widely used discrete-time volatility models. Most of them can be written under the form

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

where  $(\eta_t)$  is a sequence of iid random variables,  $\eta_t$  is independent of  $\{\epsilon_u, u < t\}$ ,  $\boldsymbol{\theta}_0 \in \mathbb{R}^d$  is a parameter belonging to a parameter space  $\Theta$ , and  $\sigma : \mathbb{R}^\infty \times \Theta \rightarrow (0, \infty)$ . A standard assumption is that  $E\eta_t^2 = 1$  but, unless otherwise stated, we do not make this assumption in the present article.

An example of widely used specification is the Asymmetric Power GARCH (APARCH) model introduced by Ding, Granger and Engle (1993). Letting  $x^+ = \max(x, 0)$  and  $x^- = \max(-x, 0)$ , the APARCH( $p, q$ ) model is defined by

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t^\delta = \omega_0 + \sum_{i=1}^q \{\alpha_{0i+} (\epsilon_{t-i}^+)^{\delta} + \alpha_{0i-} (\epsilon_{t-i}^-)^{\delta}\} + \sum_{j=1}^p \beta_{0j} \sigma_{t-j}^\delta \end{cases} \quad (2.2)$$

where the coefficients satisfy  $\alpha_{0i+} \geq 0$ ,  $\alpha_{0i-} \geq 0$ ,  $\beta_{0j} \geq 0$ ,  $\omega_0 > 0$  and  $\delta > 0$ . The standard GARCH model is obtained for  $\delta = 2$  and  $\alpha_{0i-} = \alpha_{0i+}$ . When  $\alpha_{0i-} > \alpha_{0i+}$ , negative returns have more impact on future volatilities than positive returns of the same magnitude, which is the well-documented "leverage effect".

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

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

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

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

where  $\xi_\alpha$  is the  $\alpha$ -quantile of  $\eta_t$ .

**Remark 2.1** It can be noted that in the PARCH( $p, q$ ) model, the conditional VaR at level  $\alpha$  satisfies the stochastic recurrence equation

$$\begin{aligned} \text{VaR}_t^\delta(\alpha) &= \omega_0(-\xi_\alpha)^\delta + \sum_{i=1}^q \left\{ \alpha_{0i+}(\epsilon_{t-i}^+)^\delta + \alpha_{0i-}(\epsilon_{t-i}^-)^\delta \right\} (-\xi_\alpha)^\delta \\ &\quad + \sum_{j=1}^p \beta_{0j} \text{VaR}_{t-j}^\delta(\alpha). \end{aligned} \quad (2.4)$$

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 with conditional VaR dynamic models 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 (2.4).

It will be convenient to assume that the parametric form of the volatility is stable by scaling.

**A0:** There exists a continuous function  $H$  such that for any  $\boldsymbol{\theta} \in \Theta$ , for any  $K > 0$ , and any sequence  $(x_i)_i$

$$K\sigma(x_1, x_2, \dots; \boldsymbol{\theta}) = \sigma(x_1, x_2, \dots; H(\boldsymbol{\theta}, K)).$$

This assumption is clearly satisfied for all commonly used GARCH models (see Section 2.3 for the APARCH model).

When  $\xi_\alpha < 0$ , **A0** and (2.3) entail

$$\text{VaR}_t(\alpha) = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_0^{(\alpha)}), \quad \boldsymbol{\theta}_0^{(\alpha)} = H(\boldsymbol{\theta}_0, -\xi_\alpha). \quad (2.5)$$

This parameter  $\boldsymbol{\theta}_0^{(\alpha)}$ , introduced by Francq and Zakoïan (2015) for more general risk measures, can be called the *VaR parameter* at level  $\alpha$ .

## 2.2 Asymptotic properties of a vector of conditional VaR estimators

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  $\xi_{\alpha_i}$  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). El Ghourabi, Francq and Telmoudi (2015) showed that, in this two-step procedure, the Gaussian QML can be replaced by any non-Gaussian QML. We now extend this approach to a vector of VaRs at different risk levels.

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

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

which is used to approximate  $\sigma_t(\boldsymbol{\theta}) = \sigma(\epsilon_{t-1}, \dots, \epsilon_1, \epsilon_0, \epsilon_{-1}, \dots; \boldsymbol{\theta})$ .

Given an *instrumental* density  $h > 0$ , consider the QML criterion

$$\tilde{Q}_n(\boldsymbol{\theta}) = \frac{1}{n} \sum_{t=1}^n g(\epsilon_t, \tilde{\sigma}_t(\boldsymbol{\theta})), \quad g(x, \sigma) = \log \left\{ \frac{1}{\sigma} h \left( \frac{x}{\sigma} \right) \right\}, \quad (2.6)$$

and the (generalized) QMLE

$$\hat{\boldsymbol{\theta}}_{n,h} = \arg \max_{\boldsymbol{\theta} \in \Theta} \tilde{Q}_n(\boldsymbol{\theta}). \quad (2.7)$$

This estimator is the standard Gaussian QMLE if  $h$  is the standard Gaussian density  $\phi$ . We emphasize that the parametric form of  $\sigma_t(\cdot)$  is assumed to be correctly specified, but we do not make precise assumptions on the distribution of  $\eta_t$ . In particular, we do not assume that  $\text{Var}(\eta_t) = 1$ .<sup>1</sup> Consequently,  $\hat{\boldsymbol{\theta}}_{n,h}$  will be a consistent estimator of some pseudo-value (to be defined below)  $\boldsymbol{\theta}_{0,h}$ .

The following assumptions will be used to derive the asymptotic properties of the QMLE  $\hat{\boldsymbol{\theta}}_{n,h}$ .

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

**A2:** Almost surely,  $\sigma_t(\boldsymbol{\theta}) \in (\underline{\omega}, \infty]$  for any  $\boldsymbol{\theta} \in \Theta$  and for some  $\underline{\omega} > 0$ . For  $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta$ , we have  $\sigma_t(\boldsymbol{\theta}_1) = \sigma_t(\boldsymbol{\theta}_2)$  *a.s.* if and only if  $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$ .

Note that

$$g(\epsilon_t, \sigma_t(\boldsymbol{\theta})) = g \left( \eta_t, \frac{\sigma_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta}_0)} \right) - \log \sigma_t(\boldsymbol{\theta}_0). \quad (2.8)$$

---

<sup>1</sup>Thus,  $\sigma_t^2$  is not, in general, the conditional variance  $\text{Var}(\epsilon_t | \epsilon_u, u < t)$ .

**A3:** The function  $\sigma \rightarrow Eg(\eta_0, \sigma)$  takes its values in  $[-\infty, +\infty)$  and has a unique maximum at some point  $\sigma_h \in (0, \infty)$ .

**A4:** The instrumental density  $h$  is twice continuously differentiable on  $\mathbb{R}$ , except possibly in 0, and there exist constants  $r \geq 0$  and  $C_0 > 0$  such that, for all  $u \in \mathbb{R} \setminus \{0\}$ ,

$$\max \left\{ \left| u \frac{h'(u)}{h(u)} \right|, u^2 \left| \left( \frac{h'(u)}{h(u)} \right)' \right| \right\} \leq C_0(1 + |u|^r), \quad \text{with } E|\eta_0|^{2r} < \infty.$$

**A5:** The function  $\boldsymbol{\theta} \mapsto \sigma(x_1, x_2, \dots; \boldsymbol{\theta})$  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.

**A6:**  $\boldsymbol{\theta}_{0,h} = H(\boldsymbol{\theta}_0, \sigma_h)$  belongs to the interior of  $\Theta$ .

**A7:** There exist no non-zero  $\mathbf{x} \in \mathbb{R}^d$  such that  $\mathbf{x}' \frac{\partial \sigma_t(\boldsymbol{\theta}_{0,h})}{\partial \boldsymbol{\theta}} = 0$ , *a.s.*

**A8:** There exists a neighborhood  $V(\boldsymbol{\theta}_{0,h})$  of  $\boldsymbol{\theta}_{0,h}$  such that the following variables have finite expectation:

$$\sup_{\boldsymbol{\theta} \in V(\boldsymbol{\theta}_{0,h})} \left\| \frac{1}{\sigma_t(\boldsymbol{\theta})} \frac{\partial \sigma_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right\|^4, \quad \sup_{\boldsymbol{\theta} \in V(\boldsymbol{\theta}_{0,h})} \left\| \frac{1}{\sigma_t(\boldsymbol{\theta})} \frac{\partial^2 \sigma_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right\|^2, \quad \sup_{\boldsymbol{\theta} \in V(\boldsymbol{\theta}_{0,h})} \left| \frac{\sigma_t(\boldsymbol{\theta}_{0,h})}{\sigma_t(\boldsymbol{\theta})} \right|^{2r}.$$

**Remark 2.2** Assumption **A3** reduces to  $E|\eta_0|^r < \infty$ , and Assumption **A4** reduces to  $E|\eta_0|^{2r} < \infty$  for instrumental densities of the form  $h(u) = K_1 |u|^\lambda \exp\{K_2 |u|^r\}$ , for some constants  $\lambda, K_1, K_2$ .

**Remark 2.3** The number  $\sigma_h$  involved in Assumption **A3** depends on both the density of  $\eta_t$  and the instrumental density  $h$ . It can be made explicit for classes of density  $h$  (see El Ghourabi, Francq and Telmoudi (2015)). For instance, when  $h$  belongs to the class of the Generalized Error Distributions with shape parameter  $\kappa > 0$ , defined by

$$h_\kappa(x) = \frac{\kappa}{\Gamma(1/\kappa) 2^{1+1/\kappa}} e^{-\frac{|x|^\kappa}{2}},$$

we have

$$\sigma_{h_\kappa} = \left( \frac{\kappa}{2} E|\eta_1|^\kappa \right)^{1/\kappa}.$$

Now let  $\eta_{t,h} = \frac{1}{\sigma_h} \eta_t$  and let  $\xi_{\alpha,h} = \frac{1}{\sigma_h} \xi_\alpha$  the  $\alpha$ -quantile of  $\eta_{t,h}$ . Noting that

$$\epsilon_t = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_0) \eta_t = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_{0,h}) \eta_{t,h},$$

we have, by (2.5) and **A3**,

$$\text{VaR}_t(\alpha) = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_0^{(\alpha)}), \quad \boldsymbol{\theta}_0^{(\alpha)} = H(\boldsymbol{\theta}_0, -\xi_\alpha) = H(\boldsymbol{\theta}_{0,h}, -\xi_{\alpha,h}). \quad (2.9)$$

In view of the last equality, a strategy for consistently estimating  $\boldsymbol{\theta}_0^{(\alpha)}$  is thus to estimate  $\boldsymbol{\theta}_{0,h}$  by generalized QML in the first step, and to estimate the quantile  $\xi_{\alpha,h}$  in the second step. Let the residuals of the QML estimation

$$\hat{\eta}_{t,h} = \frac{\epsilon_t}{\tilde{\sigma}_t(\hat{\boldsymbol{\theta}}_{n,h})}, \quad t = 1, \dots, n,$$

and let  $\xi_{n,\alpha_i,h}$  denote the empirical  $\alpha_i$ -quantile of  $\hat{\eta}_{1,h}, \dots, \hat{\eta}_{n,h}$ . Let  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m)'$ ,  $\boldsymbol{\xi}_{n,\alpha,h} = (\xi_{n,\alpha_1,h}, \dots, \xi_{n,\alpha_m,h})'$  and let  $\boldsymbol{\xi}_{\alpha,h} = (\xi_{\alpha_1,h}, \dots, \xi_{\alpha_m,h})'$  denote the vector of population quantiles.

The next result gives the joint asymptotic distributions of  $(\hat{\boldsymbol{\theta}}_{n,h}, \boldsymbol{\xi}'_{n,\alpha})$ . Let  $\mathbf{D}_t(\boldsymbol{\theta}) = \sigma_t^{-1}(\boldsymbol{\theta}) \frac{\partial \sigma_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ ,  $g_1(x, \sigma) = \frac{\partial g(x, \sigma)}{\partial \sigma}$  and  $g_2(x, \sigma) = \frac{\partial g_1(x, \sigma)}{\partial \sigma}$ .

**Theorem 2.1** *Assume  $\xi_{\alpha_i,h} < 0$ , for  $i = 1, \dots, m$ . Suppose  $\eta_{0,h}$  admits a density  $f_h$  which is continuous and strictly positive in a neighborhood of  $\xi_{\alpha_i,h}$ , for  $i = 1, \dots, m$ . Assume  $Eg_2(\eta_{0,h}, 1) \neq 0$ . Let **A0-A8** hold, with  $r > 1$  in **A4** and **A8**. Then*

$$\begin{pmatrix} \sqrt{n}(\hat{\boldsymbol{\theta}}_{n,h} - \boldsymbol{\theta}_{0,h}) \\ \sqrt{n}(\boldsymbol{\xi}_{n,\alpha,h} - \boldsymbol{\xi}_{\alpha,h}) \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \boldsymbol{\Sigma}_{\alpha,h}), \quad \boldsymbol{\Sigma}_{\alpha,h} = \begin{pmatrix} \tau_h \mathbf{J}_h^{-1} & \boldsymbol{\lambda}'_{\alpha,h} \otimes \mathbf{J}_h^{-1} \boldsymbol{\Omega}_h \\ \boldsymbol{\lambda}_{\alpha,h} \otimes \boldsymbol{\Omega}'_h \mathbf{J}_h^{-1} & \boldsymbol{\zeta}_{\alpha,h} \end{pmatrix},$$

where

$$\tau_h = \frac{4Eg_1^2(\eta_{0,h}, 1)}{\{Eg_2(\eta_{0,h}, 1)\}^2}, \quad \boldsymbol{\Omega}_h = E\mathbf{D}_t(\boldsymbol{\theta}_{0,h}), \quad \mathbf{J}_h = 4E\mathbf{D}_t(\boldsymbol{\theta}_{0,h})\mathbf{D}'_t(\boldsymbol{\theta}_{0,h}),$$

$\boldsymbol{\lambda}_{\alpha,h} = (\lambda_{\alpha_1,h}, \dots, \lambda_{\alpha_m,h})'$ ,  $\boldsymbol{\zeta}_{\alpha,h} = (\zeta_{ij,h})_{1 \leq i,j \leq m}$  and

$$\begin{aligned} \lambda_{\alpha_i,h} &= -\xi_{\alpha_i,h} \tau_h + \frac{4p_{\alpha_i,h}}{f_h(\xi_{\alpha_i,h}) Eg_2(\eta_{0,h}, 1)}, \\ \zeta_{ij,h} &= \xi_{\alpha_i,h} \xi_{\alpha_j,h} \frac{\tau_h}{4} - \frac{1}{Eg_2(\eta_{0,h}, 1)} \left( \frac{\xi_{\alpha_i,h} p_{\alpha_j,h}}{f_h(\xi_{\alpha_j,h})} + \frac{\xi_{\alpha_j,h} p_{\alpha_i,h}}{f_h(\xi_{\alpha_i,h})} \right) + \frac{\alpha_i \wedge \alpha_j - \alpha_i \alpha_j}{f_h(\xi_{\alpha_i,h}) f_h(\xi_{\alpha_j,h})}, \end{aligned}$$

with  $p_{\alpha,h} = \text{Cov} \left\{ \mathbf{1}_{\{\eta_{0,h} < \xi_{\alpha,h}\}}, g_1(\eta_{0,h}, 1) \right\}$ .



**Proof.** In view of El Ghourabi, Francq and Telmoudi (Proof of Theorem 1, 2015), we have, for  $i = 1, \dots, m$ ,

$$\begin{aligned}\sqrt{n}(\xi_{\alpha_i, h} - \xi_{n, \alpha_i, h}) &= \xi_{\alpha_i, h} \boldsymbol{\Omega}'_h \sqrt{n}(\widehat{\boldsymbol{\theta}}_{n, h} - \boldsymbol{\theta}_{0, h}) \\ &\quad + \frac{1}{f_h(\xi_{\alpha_i, h})} \frac{1}{\sqrt{n}} \sum_{t=1}^n (\mathbf{1}_{\{\eta_{t, h} < \xi_{\alpha_i, h}\}} - \alpha_i) + o_P(1),\end{aligned}$$

and

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_{n, h} - \boldsymbol{\theta}_{0, h}) = \frac{-4}{Eg_2(\eta_{0, h}, 1)} \mathbf{J}_h^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^n g_1(\eta_{t, h}, 1) \mathbf{D}_t(\boldsymbol{\theta}_{0, h}) + o_P(1).$$

Hence

$$\text{Cov}_{as} \left( \sqrt{n}(\widehat{\boldsymbol{\theta}}_{n, h} - \boldsymbol{\theta}_{0, h}), \frac{1}{\sqrt{n}} \sum_{t=1}^n (\mathbf{1}_{\{\eta_{t, h} < \xi_{\alpha_i, h}\}} - \alpha_i) \right) = \frac{-4p_{\alpha_i, h}}{Eg_2(\eta_{0, h}, 1)} \mathbf{J}_h^{-1} \boldsymbol{\Omega}_h.$$

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

$$\begin{aligned}&\text{Cov}_{as} \{ \sqrt{n}(\xi_{\alpha_i, h} - \xi_{n, \alpha_i, h}), \sqrt{n}(\xi_{\alpha_j, h} - \xi_{n, \alpha_j, h}) \} \\ &= \left\{ \xi_{\alpha_i, h} \xi_{\alpha_j, h} \tau_h - \frac{4}{Eg_2(\eta_{0, h}, 1)} \left( \frac{\xi_{\alpha_i, h} p_{\alpha_j, h}}{f_h(\xi_{\alpha_j, h})} + \frac{\xi_{\alpha_j, h} p_{\alpha_i, h}}{f_h(\xi_{\alpha_i, h})} \right) \right\} \boldsymbol{\Omega}'_h \mathbf{J}_h^{-1} \boldsymbol{\Omega}_h \\ &\quad + \frac{\alpha_i(1 - \alpha_j)}{f_h(\xi_{\alpha_i, h}) f_h(\xi_{\alpha_j, h})}, \\ &\text{Cov}_{as} \left( \sqrt{n}(\widehat{\boldsymbol{\theta}}_{n, h} - \boldsymbol{\theta}_{0, h}), \sqrt{n}(\xi_{\alpha_i, h} - \xi_{n, \alpha_i, h}) \right) \\ &= \lambda_{\alpha_i, h} \mathbf{J}_h^{-1} \boldsymbol{\Omega}_h.\end{aligned}$$

We have  $\boldsymbol{\Omega}'_h \mathbf{J}_h^{-1} \boldsymbol{\Omega}_h = 1/4$  (see Remark 3.1 in Francq and Zakoïan, 2013) and thus we obtain

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

By the CLT for martingale differences, we get the announced result.  $\square$

Let  $\mathbf{VaR}_t(\boldsymbol{\alpha}) = (\text{VaR}_t(\alpha_1), \dots, \text{VaR}_t(\alpha_m))'$ , the vector of conditional VaRs at levels  $\alpha_i$ . In view of

$$\mathbf{VaR}_t(\boldsymbol{\alpha}) = -\sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_{0, h}) \boldsymbol{\xi}_{\boldsymbol{\alpha}, h}, \quad (2.10)$$

the vector of conditional VaRs can be estimated by

$$\widehat{\mathbf{VaR}}_t(\boldsymbol{\alpha}) = -\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n, h}) \boldsymbol{\xi}_{n, \boldsymbol{\alpha}, h}. \quad (2.11)$$

**Remark 2.4** In classical quantile regression, a serious problem is that the estimated quantile curves can cross, leading to an invalid inference at multiple percentiles (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  $\alpha$ 's.

**Remark 2.5** The coefficient  $\tau_h$  can be made explicit for different classes of density functions  $h$ . For instance, for the GED distribution of Remark 2.3, simple computation shows that for the density  $h_\kappa$ ,

$$\tau_h = \frac{4}{\kappa^2} \left( \frac{E|\eta_1|^{2\kappa}}{(E|\eta_1|^\kappa)^2} - 1 \right).$$

In other cases, such as the class of the Student densities, coefficients  $\tau_h$  and  $\sigma_h$  do not have an explicit expression but can be obtained numerically.

**Remark 2.6** It should be noted that the coefficient  $\tau_h$  appearing in the asymptotic distribution only depends on i) the distributional properties of the variable  $\eta_0$  and ii) the choice of the QML density  $h$ . For a given  $h$ , the coefficient  $\tau_h$  can be estimated using the residuals  $\hat{\eta}_{t,h}$  by

$$\hat{\tau}_h = 4 \frac{n^{-1} \sum_{t=1}^n g_1^2(\hat{\eta}_{t,h}, 1)}{\{n^{-1} \sum_{t=1}^n g_2(\hat{\eta}_{t,h}, 1)\}^2}. \quad (2.12)$$

The residuals  $\hat{\eta}_{t,h}$  can be used to estimate the density  $f_h$ , as well as all other quantities involved in the asymptotic distribution.

### 2.3 Application to the APARCH model

For the APARCH( $p, q$ ) model with  $\delta$  fixed<sup>2</sup>, we have  $\boldsymbol{\theta} = (\omega, \alpha_{1+}, \dots, \alpha_{q-}, \beta_1, \dots, \beta_p)'$  and  $\mathbf{A}\mathbf{0}$  is satisfied with  $H(\boldsymbol{\theta}, K) = (K^\delta \omega, K^\delta \alpha_{1+}, \dots, K^\delta \alpha_{q-}, \beta_1, \dots, \beta_p)'$ . The parameter is assumed to belong to a compact set  $\Theta \subset ]0, +\infty[ \times ]0, +\infty[^{p+2q}$ . Let  $\mathcal{A}_{\boldsymbol{\theta}+}(z) = \sum_{i=1}^q \alpha_{i+} z^i$  and  $\mathcal{A}_{\boldsymbol{\theta}-}(z) = \sum_{i=1}^p \alpha_{i-} z^i$   $\mathcal{B}_{\boldsymbol{\theta}}(z) = 1 - \sum_{j=1}^p \beta_j z^j$  with, by convention,  $\mathcal{A}_{\boldsymbol{\theta}+}(z) = \mathcal{A}_{\boldsymbol{\theta}-}(z) = 0$  if  $q = 0$  and  $\mathcal{B}_{\boldsymbol{\theta}}(z) = 1$  if  $p = 0$ . Let  $\gamma$  denote the Lyapunov coefficient of the sequence  $(A_t)$  associated with the vector representation of the model. Hamadeh and Zakoïan (2011) showed the CAN of the Gaussian QMLE of  $\boldsymbol{\theta}_0$  under the assumption:

---

<sup>2</sup>Hamadeh and Zakoïan (2011) showed that estimating the power  $\delta$  is feasible though complicated. We therefore consider  $\delta$  as fixed. In most applications,  $\delta$  is either equal to 1 (as in the TARCh of Zakoïan (1994)) or to 2 (as in the GJR model of Glosten, Jagannathan and Runkle (1993)).

**D**( $\theta_0$ ):  $\gamma < 0$ ; the true parameter value  $\theta_0$  belongs to the interior of  $\Theta$ ; there exists  $\underline{\omega} > 0$  such that,  $\forall \theta \in \Theta$ ,  $\omega > \underline{\omega}$  and  $\sum_{j=1}^p \beta_j < 1$ ; the support of the distribution of  $\eta_0$  contains at least 3 points;  $P[\eta_t > 0] \in (0, 1)$ ; if  $p > 0$ ,  $\mathcal{B}_{\theta_0}(z)$  has no common root with  $\mathcal{A}_{\theta_0+}(z)$  and  $\mathcal{A}_{\theta_0-}(z)$ ;  $\mathcal{A}_{\theta_0+}(1) + \mathcal{A}_{\theta_0-}(1) \neq 0$  and  $\alpha_{0q,+} + \alpha_{0q,-} + \beta_{0p} \neq 0$

and under the identifiability condition  $E\eta_1^2 = 1$  (which we do not assume in our framework).

For any  $\theta \in \Theta$ , let  $\bar{\theta}' = (\omega, \alpha_{1+}, \dots, \alpha_{q-}, 0, \dots, 0)$ . We have,

$$\begin{aligned} \bar{\theta}' \frac{\partial \sigma_t^\delta(\theta)}{\partial \theta} &= \omega + \sum_{i=1}^q \left\{ \alpha_{i+} (\epsilon_{t-i}^+)^{\delta} + \alpha_{i-} (\epsilon_{t-i}^-)^{\delta} \right\} + \sum_{j=1}^p \beta_j \bar{\theta}' \frac{\partial \sigma_{t-j}^\delta(\theta)}{\partial \theta} \\ &= \mathcal{B}_{\theta}^{-1}(L) \left( \omega + \sum_{i=1}^q \left\{ \alpha_{i+} (\epsilon_{t-i}^+)^{\delta} + \alpha_{i-} (\epsilon_{t-i}^-)^{\delta} \right\} \right) = \sigma_t^\delta(\theta), \end{aligned}$$

where  $L$  denotes the usual lag operator. Therefore,

$$\bar{\theta}' \frac{1}{\sigma_t} \frac{\partial \sigma_t^\delta(\theta)}{\partial \theta} = \frac{1}{\delta},$$

and thus

$$\bar{\theta}'_{0,h} \Omega_h = \frac{1}{\delta}, \quad \mathbf{J}_h^{-1} \Omega_h = \frac{\delta}{4} \bar{\theta}_{0,h}.$$

It follows that Theorem 2.1 can be simplified as follows in the case of the APARCH model.

**Corollary 2.1** *Consider the APARCH( $p, q$ ) model (2.2) under Assumption **D**( $\theta_{0,h}$ ). Assume  $\xi_{\alpha_i, h} < 0$ , for  $i = 1, \dots, m$ . Suppose  $\eta_{0,h}$  admits a density  $f_h$  which is continuous and strictly positive in a neighborhood of  $\xi_{\alpha_i, h}$ , for  $i = 1, \dots, m$ . If the instrumental density  $h$  satisfies **A3**, **A4**, and if  $Eg_2(\eta_{0,h}, 1) \neq 0$ , then*

$$\begin{pmatrix} \sqrt{n} (\hat{\theta}_{n,h} - \theta_{0,h}) \\ \sqrt{n} (\hat{\xi}_{n,\alpha,h} - \xi_{\alpha,h}) \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_{\alpha,h}), \quad \Sigma_{\alpha,h} = \begin{pmatrix} \tau_h \mathbf{J}_h^{-1} & \lambda'_{\alpha,h} \otimes \frac{\delta}{4} \bar{\theta}_{0,h} \\ \lambda_{\alpha,h} \otimes \frac{\delta}{4} \bar{\theta}'_{0,h} & \zeta_{\alpha,h} \end{pmatrix}.$$

## 2.4 Optimal choice of the density $h$

El Ghourabi, Francq and Telmoudi (2015) showed that, for the VaR estimation at a single level, an optimal choice of the density  $h$  is obtained by minimizing (within a class of densities) the coefficient  $\tau_h$ . For such an optimal density  $h^*$ , the accuracy of the VaR estimation is maximal. Interestingly, this density  $h^*$  does not depend on  $\alpha$ , nor on the volatility specification.

For instance, when the density  $h$  is chosen among the GED densities, in view of Remark 2.5 an optimal value for  $\kappa$  can be estimated by taking

$$\hat{\kappa} = \arg \min_{\kappa \in \mathcal{K}} \frac{1}{\kappa^2} \left( \frac{\hat{\mu}_{2\kappa}}{\hat{\mu}_{\kappa}^2} - 1 \right), \quad \hat{\mu}_r = \frac{1}{n} \sum_{t=1}^n |\hat{\eta}_{t,h}|^r, \quad (2.13)$$

for some compact set  $\mathcal{K} \subset \mathbb{R}^+$ . Practical implementation, for any given class  $\mathcal{H}$  of densities  $h$ , thus involves the following steps:

1. For any  $h_0 \in \mathcal{H}$ , compute  $\hat{\boldsymbol{\theta}}_{n,h_0}$  by solving (2.7) (for  $h = h_0$ ).
2. Using the residuals  $\hat{\eta}_{t,h_0}$  of the first step, compute the coefficients

$$\hat{\tau}_h = 4 \frac{n^{-1} \sum_{t=1}^n g_1^2(\hat{\eta}_{t,h_0}, 1)}{\{n^{-1} \sum_{t=1}^n g_2(\hat{\eta}_{t,h_0}, 1)\}^2}, \quad (2.14)$$

where  $g_1, g_2$  are defined before Theorem 2.1. Solve  $h^* = \arg \max_{h \in \mathcal{H}} \hat{\tau}_h$ .

3. Compute  $\hat{\boldsymbol{\theta}}_{n,h^*}$  and deduce the conditional VaR from (2.11) (with  $h^*$  instead of  $h$ ).

### 3 Portfolios of VaR's

Risk measurement based on a single VaR at a given level can be misleading since it gives a limited view of the loss distribution. To circumvent this problem, DRMs have been introduced in the insurance literature, in a series of papers by Wang and coauthors [see Wang (2000) and the references therein]. General conditional DRMs take the form

$$\text{DRM}_t = \int_0^1 \text{VaR}_t(u) dG(u), \quad (3.1)$$

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

$$\text{DRM}_t = -\sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_0) \int_0^1 \xi_u dG(u). \quad (3.2)$$

#### 3.1 Estimating the discrete DRM parameter

If  $-\int_0^1 \xi_u dG(u) > 0$ , the DRM in (3.2) can be written as

$$\text{DRM}_t = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_0^G) \quad (3.3)$$

where  $\boldsymbol{\theta}_0^G$  defined by

$$\boldsymbol{\theta}_0^G = H \left( \boldsymbol{\theta}_0, - \int_0^1 \xi_u dG(u) \right) \quad (3.4)$$

is called DRM-parameter (similarly to the VaR parameter in (2.5)).

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

$$\mathbf{p}' \mathbf{VaR}_t(\boldsymbol{\alpha}) = \sum_{i=1}^m p_i \mathbf{VaR}_t(\alpha_i), \quad (3.5)$$

where  $\mathbf{p} = (p_1, \dots, p_m)$  with  $p_i \geq 0$  for  $i = 1, \dots, m$  and  $\sum_{i=1}^m p_i = 1$ . This risk measure can be interpreted as a *discrete DRM* with associated distortion function corresponding to Dirac masses at the points  $\alpha_i$ . By (2.3) we have

$$\begin{aligned} \mathbf{p}' \mathbf{VaR}_t(\boldsymbol{\alpha}) &= -\sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_0) \sum_{i=1}^m p_i \xi_{\alpha} \\ &= \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}_0^{DRM}), \quad \text{where} \quad \boldsymbol{\theta}_0^{DRM} = H(\boldsymbol{\theta}_0, -\mathbf{p}' \boldsymbol{\xi}_{\alpha}) \end{aligned} \quad (3.6)$$

can be called *discrete DRM-parameter*. In view of (2.9), we also have, for any instrumental density  $h$ ,

$$\boldsymbol{\theta}_0^{DRM} = H(\boldsymbol{\theta}_{0,h}, -\mathbf{p}' \boldsymbol{\xi}_{\alpha,h}),$$

from which we deduce an estimator of the *discrete DRM-parameter* given by

$$\widehat{\boldsymbol{\theta}}_{n,h}^{DRM} = H(\widehat{\boldsymbol{\theta}}_{n,h}, -\mathbf{p}' \boldsymbol{\xi}_{n,\alpha,h}),$$

whose asymptotic distribution is a straightforward consequence of Theorem 2.1. Denoting by  $(\boldsymbol{\theta}, x)$  the generic arguments of the function  $H$ , let the  $d \times d$  matrix  $\mathbf{A}_{0,h} = \frac{\partial H}{\partial \boldsymbol{\theta}'}(\boldsymbol{\theta}_{0,h}, -\mathbf{p}' \boldsymbol{\xi}_{\alpha,h})$  and the  $d \times 1$  vector  $\mathbf{b}_{0,h} = \frac{\partial H}{\partial x}(\boldsymbol{\theta}_{0,h}, -\mathbf{p}' \boldsymbol{\xi}_{\alpha,h})$ .

**Corollary 3.1** *Under the assumptions of Theorem 2.1,*

$$\begin{aligned} \sqrt{n} \left( \widehat{\boldsymbol{\theta}}_{n,h}^{DRM} - \boldsymbol{\theta}_{0,h}^{DRM} \right) &\xrightarrow{\mathcal{L}} \mathcal{N}(0, \boldsymbol{\Sigma}_{\alpha,h}^{DRM}), \\ \boldsymbol{\Sigma}_{\alpha,h}^{DRM} &= [\mathbf{A}_{0,h} - \mathbf{b}_{0,h} \mathbf{p}'] \boldsymbol{\Sigma}_{\alpha,h} \begin{pmatrix} \mathbf{A}'_{0,h} \\ -\mathbf{p} \mathbf{b}'_{0,h} \end{pmatrix}. \end{aligned}$$

**Proof.** We have, by a Taylor expansion of the function  $H(\boldsymbol{\theta}, x)$  around  $(\boldsymbol{\theta}_{0,h}, -\mathbf{p}' \boldsymbol{\xi}_{\alpha,h})$ ,

$$\sqrt{n} \left( \widehat{\boldsymbol{\theta}}_{n,h}^{DRM} - \boldsymbol{\theta}_{0,h}^{DRM} \right) = \mathbf{A}_{0,h} \sqrt{n} \left( \widehat{\boldsymbol{\theta}}_{n,h} - \boldsymbol{\theta}_{0,h} \right) - \mathbf{b}_{0,h} \sqrt{n} \{ \mathbf{p}' (\boldsymbol{\xi}_{\alpha,h} - \boldsymbol{\xi}_{n,\alpha,h}) \}$$

and the conclusion follows.  $\square$

It follows from (3.6) that the discrete DRM can be estimated by

$$\mathbf{p}'\widehat{\text{VaR}}_t(\boldsymbol{\alpha}) = \tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n,h}^{DRM}).$$

### 3.2 Constructing confidence intervals for the portfolio of VaR's

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

Let also  $\widehat{\mathbf{A}}_{n,h} = \frac{\partial H}{\partial \boldsymbol{\theta}'}(\widehat{\boldsymbol{\theta}}_{n,h}, -\mathbf{p}'\boldsymbol{\xi}_{n,\alpha,h})$ ,  $\widehat{\mathbf{b}}_{n,h} = \frac{\partial H}{\partial x}(\widehat{\boldsymbol{\theta}}_{n,h}, -\mathbf{p}'\boldsymbol{\xi}_{n,\alpha,h})$  and let

$$\widehat{\boldsymbol{\Sigma}}_{n,\alpha,h}^{DRM} = [\widehat{\mathbf{A}}_{n,h} - \widehat{\mathbf{b}}_{n,h}\mathbf{p}']\widehat{\boldsymbol{\Sigma}}_{n,\alpha,h} \begin{pmatrix} \widehat{\mathbf{A}}_{n,h}' \\ -\mathbf{p}\widehat{\mathbf{b}}_{n,h}' \end{pmatrix}.$$

Corollary 3.1 and the delta method thus suggests a  $(1 - \alpha_0)\%$  confidence interval (CI) for  $\mathbf{p}'\widehat{\text{VaR}}_t(\boldsymbol{\alpha})$  whose bounds are

$$\mathbf{p}'\widehat{\text{VaR}}_t(\boldsymbol{\alpha}) \pm \frac{\Phi_{1-\alpha_0/2}^{-1}}{\sqrt{n}} \left\{ \frac{\partial \tilde{\sigma}_t}{\partial \boldsymbol{\theta}'}(\widehat{\boldsymbol{\theta}}_{n,h}^{DRM}) \widehat{\boldsymbol{\Sigma}}_{\alpha,h}^{DRM} \frac{\partial \tilde{\sigma}_t}{\partial \boldsymbol{\theta}}(\widehat{\boldsymbol{\theta}}_{n,h}^{DRM}) \right\}^{1/2}, \quad (3.7)$$

where  $\Phi_{\alpha_0}^{-1}$  denotes the  $\alpha_0$ -quantile of the standard Gaussian distribution. It should be noted  $\alpha_0$  (the risk estimation level) can be chosen independently from the  $\alpha_i$ 's (the financial risk levels). Drawing such CIs allows to underline the importance of the estimation risk for VaR evaluation.

In particular, a  $(1 - \alpha_0)\%$  confidence interval (CI) for the  $\text{VaR}_t(\alpha_i)$  is given by

$$-\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n,h})\xi_{n,\alpha_i} \pm \frac{\Phi_{1-\alpha_0/2}^{-1}}{\sqrt{n}} \left\{ \left( \widehat{\boldsymbol{\Delta}}_{t,\alpha,h} \widehat{\boldsymbol{\Sigma}}_{\alpha,h} \widehat{\boldsymbol{\Delta}}_{t,\alpha,h}' \right)_{ii} \right\}^{1/2},$$

where

$$\widehat{\boldsymbol{\Delta}}_{t,\alpha,h} = \left( \boldsymbol{\xi}_{n,\alpha,h} \frac{\partial \tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n,h})}{\partial \boldsymbol{\theta}'}, \tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n,h}) \mathbf{I}_m \right), \quad (3.8)$$

where  $\mathbf{I}_m$  denotes the  $m \times m$  identity matrix

### 3.3 On the moment assumptions

We have seen in Section 2.4 that the choice of the instrumental density  $h$  has an impact on the asymptotic accuracy of the VaR estimator. This will be illustrated in the empirical section. It is also

important to note that the moments assumptions on the iid process required for the validity of the asymptotic results depend on the choice of  $h$ . Such moments assumptions appear in Assumptions **A3** and **A4** through the conditions

$$Eg(\eta_0, \sigma) < \infty \quad \text{and} \quad E|\eta_0|^{2r} < \infty,$$

where the  $r > 0$  is determined, in the first part of **A4**, by the choice of  $h$ .

To be more specific, consider instrumental densities of the form  $h(u) = K_1|u|^\lambda \exp\{K_2|u|^r\}$ , for some constants  $\lambda, K_1, K_2$ . By Remark 2.2, the moment assumptions reduce to  $E|\eta_0|^{2r} < \infty$ .

For instance, consider the usual Gaussian QMLE ( $r = 2$  and  $\lambda = 0$ ). If  $E|\eta_0|^4 < \infty$ , the two-step VaR and discrete DRM parameter estimators are consistent and asymptotically normal. In view of Corollary 3.1, valid confidence intervals for the portfolio of VaR's can be constructed. Now suppose that  $E|\eta_0|^2 = 1$  but  $E|\eta_0|^4 = \infty$ . Then, under appropriate assumptions, the Gaussian QMLE of the volatility parameter is well known to be consistent, and it could probably be established that the discrete DRM parameter is also consistent. Hall and Yao (2003) derived a non standard asymptotic distribution for the estimator of the volatility parameter. However, establishing the analogous of Theorem 2.1 in this situation would be a formidable task. Finally, if  $E|\eta_0|^2 = \infty$ , the Gaussian QML estimator of the volatility parameter is probably not even consistent.

If, instead, an instrumental density of the form  $h(u) = K_1|u|^\lambda \exp\{K_2|u|^r\}$  with  $r < 2$  such that  $E|\eta_0|^{2r} < \infty$  is chosen, then the estimator of the discrete DRM parameter has a standard asymptotic distribution given by Corollary 3.1. In particular, our theory allows to handle GARCH models with Lévy alpha-stable conditional distributions.

To conclude this section, it is worth noting that estimating a portfolio of VaRs with a non-Gaussian instrumental density is not more demanding, in terms of computational burden, than with the usual QMLE.

## 4 Empirical illustration

We now illustrate our theoretical results by considering a set of daily returns of 9 world stock market indices: CAC (Paris), DAX (Frankfurt), FTSE (London), Nikkei (Tokyo), NSE (Bombay), SMI (Switzerland), SP500 (New York), SPTSX (Toronto), and SSE (Shanghai), collected from early January 1990 to the end of June 2013 (the data have been downloaded from Yahoo Finance website).

We compared two estimators of the discrete DRM (3.5), in which the  $m$  levels are equally spaced from  $\alpha_1 = 0.01$  to  $\alpha_m = 0.10$ , and the weights are defined, for some  $r > 0$ , by

$$p_1 = \frac{\alpha_1^r}{\alpha_m^r}, \quad p_i = \frac{\alpha_i^r - \alpha_{i-1}^r}{\alpha_m^r}, \quad i = 2, \dots, m.$$

Note that the weights are derived from the so-called "proportional hazard" DRM. In particular the weights decrease with  $i$  when  $r < 1$  (which reflects risk aversion). The results presented in Table 1 correspond to  $r = 1/2$  and  $m = 5$ , but the outputs are qualitatively similar for other choices of these coefficients. For the volatility specification we used a standard GARCH(1,1). The estimator displayed in the columns "Gaussian QMLE" is simply obtained with the usual Gaussian instrumental density  $h = \phi$ . For the estimator displayed in the columns "GED QMLE", we used an instrumental density within the GED( $\kappa$ ) class. We applied the three steps of Section 2.4 using the Gaussian QML in the first step.

Table 1 shows that the estimates of the DRM parameters produced by the two methods are similar, with non empty intersections for confidence bounds whose widths are two estimated standard deviations. The estimated standard deviations are however systematically larger with the method based on the Gaussian QMLE than with that based on the optimal GED instrumental density, which leads us to think that the second method is preferable.

Although the estimated VaR parameters appear similar for all indices, some differences can be underlined. For instance, let us compare the VaR parameters estimated by the GED QML for the DAX and the SMI. By (3.6), the conditional risk of the portfolio,  $\mathbf{p}'\mathbf{VaR}_t(\boldsymbol{\alpha})$  is estimated by

$$\tilde{\sigma}_t(\hat{\boldsymbol{\theta}}_{n,h}^{DRM}) = \left( \frac{\hat{\omega}_n}{1 - \hat{\beta}_n} + \sum_{i=1}^{t-1} \hat{\alpha}_n \hat{\beta}_n^i \epsilon_{t-i}^2 \right)^{1/2}, \quad (4.1)$$

where  $\hat{\boldsymbol{\theta}}_{n,h}^{DRM} = (\hat{\omega}_n, \hat{\alpha}_n, \hat{\beta}_n)'$ . It is seen from Table 2 that the effect of a shock at time  $t - 1$  on the estimated VaR of the portfolio will be much larger for the SMI than for the DAX. At longer horizons, however, the effects are reversed. It is worthnoting that this kind of interpretation is made possible by the notion of VaR parameter, which jointly incorporates the effects of volatility and the tails of the conditional distribution.

Figure 1 displays, for the CAC and DAX indices, the conditional DRM and its estimated 95% CI, as defined by (3.8), obtained from each of the two methods. The DRM's are estimated over a period of 100 days, from April, 7, 2011 to August, 26, 2011 for the CAC index and from April,



Table 1: Estimation of the conditional discrete DRM parameter for 9 stock market indices. The estimated standard deviations are displayed into brackets.

Index	$n$	Gaussian QMLE			GED QMLE		
		$\omega^{DRM}$	$\alpha^{DRM}$	$\beta^{DRM}$	$\omega^{DRM}$	$\alpha^{DRM}$	$\beta^{DRM}$
CAC	5229	0.11 (0.03)	0.31 (0.04)	0.9 (0.01)	0.09 (0.02)	0.29 (0.03)	0.91 (0.01)
DAX	5226	0.12 (0.04)	0.31 (0.06)	0.9 (0.02)	0.07 (0.02)	0.31 (0.04)	0.91 (0.01)
FTSE	5217	0.04 (0.01)	0.32 (0.04)	0.91 (0.01)	0.04 (0.01)	0.30 (0.03)	0.91 (0.01)
Nikkei	5078	0.20 (0.04)	0.36 (0.05)	0.88 (0.01)	0.14 (0.03)	0.32 (0.04)	0.90 (0.01)
NSE	2265	0.25 (0.09)	0.42 (0.10)	0.87 (0.03)	0.22 (0.08)	0.42 (0.12)	0.87 (0.02)
SMI	5209	0.17 (0.04)	0.45 (0.08)	0.84 (0.03)	0.10 (0.02)	0.42 (0.06)	0.87 (0.01)
SP500	5206	0.03 (0.01)	0.27 (0.03)	0.92 (0.01)	0.02 (0.01)	0.24 (0.03)	0.93 (0.01)
SPTSX	2934	0.03 (0.01)	0.27 (0.04)	0.93 (0.01)	0.04 (0.01)	0.30 (0.05)	0.92 (0.01)
SSE	2982	0.11 (0.04)	0.26 (0.05)	0.93 (0.01)	0.14 (0.04)	0.33 (0.06)	0.91 (0.01)

8, 2011 to August, 26, 2011 for the DAX index. The estimation of the DRM parameters is based on the 1000 previous values. It can be seen that the estimated DRM's are very close, but the CI's can be quite different. This is not surprising because we know from the asymptotic theory that the two methods are consistent, but that the method based on the optimal GED can be more efficient than that based on the Gaussian instrumental density (with corresponds to the particular GED of parameter  $\tau = 2$ ). The difference is particularly important during turbulent periods (near the August 2011 stock markets fall).

It can be noted that in turbulent periods, both the market risk and the estimation risk increase. This is due to the fact that, as can be seen from (3.8), the derivatives of the VaR, with respect to  $\theta$  and to the quantiles of the innovations, increase with volatility. Participants of financial markets are well aware that the reserves should be increased in turbulent periods, but our conclusion is that even the surplus of reserves should be increased due to the estimation risk.

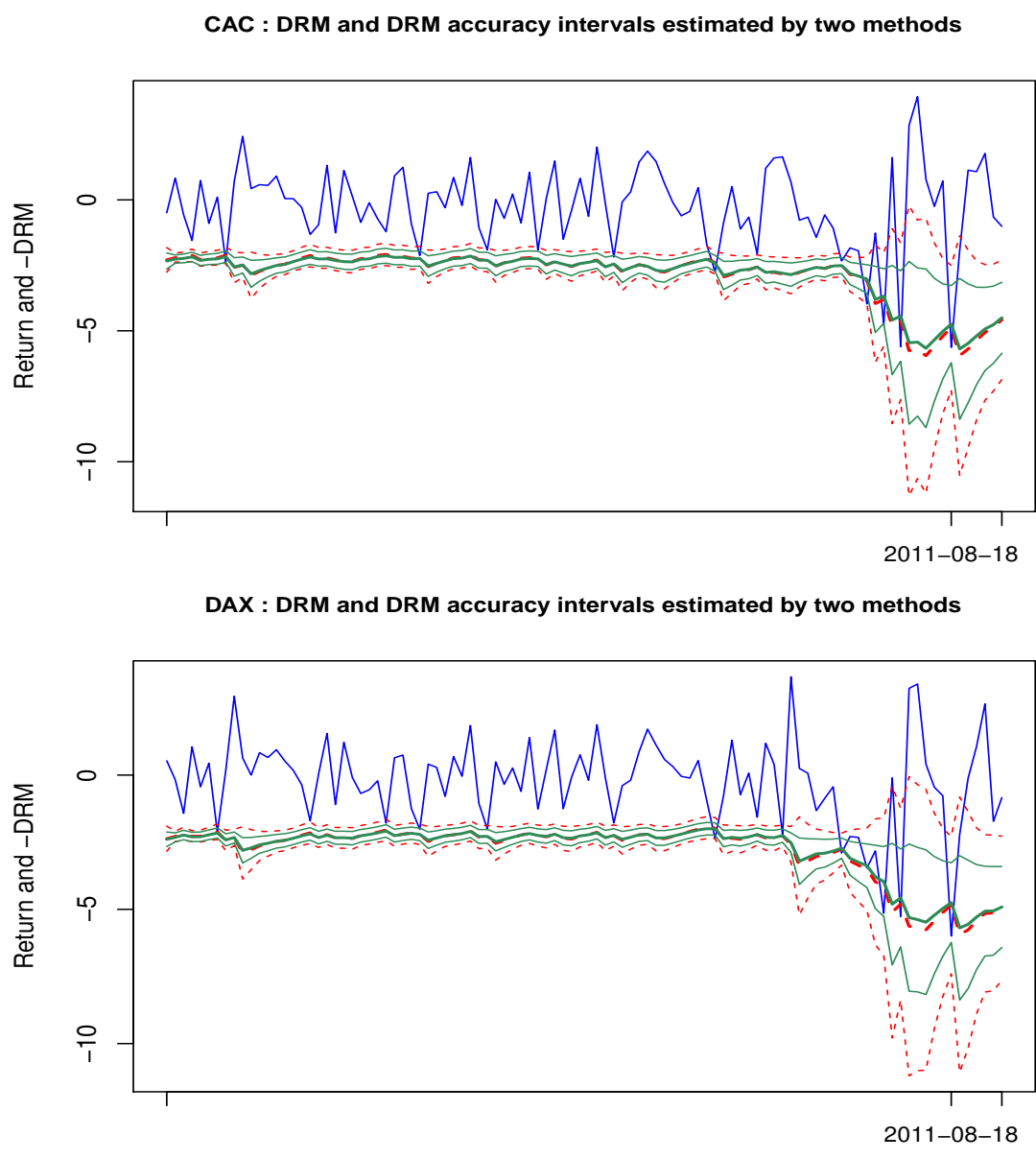


Figure 1: Returns (in blue) and (opposite of the) estimated discrete DRM and its 95% CI, with the Gaussian QMLE (in dotted red lines) and the GED-based method (in full green lines).

Table 2: Coefficients of the  $\epsilon_{t-i}^2$ , in the estimated discrete DRM displayed in (4.1), for two indices.

Index	$\epsilon_{t-1}^2$	$\epsilon_{t-2}^2$	$\epsilon_{t-3}^2$	$\epsilon_{t-4}^2$	$\epsilon_{t-5}^2$	$\epsilon_{t-6}^2$	$\epsilon_{t-7}^2$	$\epsilon_{t-8}^2$	$\epsilon_{t-9}^2$
SMI	0.42	0.37	0.32	0.28	0.24	0.21	0.18	0.16	0.14
DAX	0.31	0.28	0.26	0.23	0.21	0.19	0.18	0.16	0.15

## 5 Conclusion

In this paper, we considered the joint estimation of conditional VaRs at different levels, in the framework of conditionally heteroskedastic models. By considering a QML approach, we avoided strong distributional assumptions on the noise sequence. The asymptotic results were established for a general class of QMLEs, including the usual Gaussian QMLE. The generalized QMLE converges to a volatility parameter which is specific to the chosen instrumental density  $h$ . The true conditional VaR is obviously independent of the chosen parameterization and, interestingly, it can be estimated by any QML contrary to the volatility parameter. The VaR estimator and its asymptotic accuracy depend on the specific QML, however. We showed how the choice of  $h$  can be optimized, based on a preliminary QML estimation of the model, to gain in asymptotic accuracy.

We also introduced discrete DRM based on a finite number of VaRs. Our empirical analysis showed that confidence intervals for portfolios of VaRs crucially depend on the chosen density  $h$ .

In particular, we have seen that, for heavy tailed error distributions, the Gaussian QML may not be reliable for estimating the conditional VaR, or at least for determining its confidence intervals. An estimator based on an alternative instrumental density may be reliable in such situations. Even for error distributions with finite fourth moments, non Gaussian QML estimators of portfolios of VaRs can provide important efficiency gains without cost in terms of computational time.

## References

- Berkes, I. and L. Horváth** (2004) The efficiency of the estimators of the parameters in GARCH processes. *The Annals of Statistics* 32, 633–655.
- 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.

- 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.
- El Ghourabi, M., Francq, C. and F. Telmoudi** (2015) Consistent estimation of the Value-at-Risk when the error distribution of the volatility model is misspecified. *Journal of Time Series Analysis*, forthcoming.
- Engle, R.F. and S. Manganelli** (2004) CAViaR: Conditional Value-at-Risk by quantile regression. *Journal of Business and Economic Statistics* 22, 367–381.
- Escanciano, J.C. and J. Olmo** (2010) Backtesting parametric Value-at-Risk with estimation risk. *Journal of Business and Economic Statistics* 28, 36–51.
- Fan, J., Qi, L. and D. Xiu** (2014) Quasi-maximum likelihood estimation of GARCH Models with heavy-tailed likelihoods. *Journal of Business and Economic Statistics* 32, 178–191.
- 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** (2013) Optimal predictions of powers of conditionally heteroskedastic processes. *Journal of the Royal Statistical Society - Series B* 75, 345–367.
- Francq, C. and J-M. Zakoïan** (2014) Multi-level conditional VaR estimation in dynamic models. In *Modeling Dependence in Econometrics. Advances in Intelligent Systems and Computing Volume 251*. Edts: V-N. Huynh et al., Springer.
- Francq, C. and J-M. Zakoïan** (2015) Risk-parameter estimation in volatility models. *Journal of Econometrics* 184, 158–173.
- Glosten, L.R., Jaganathan, R. and 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.

- Gouriéroux, C. and J.M. Zakoïan** (2013) Estimation adjusted VaR. *Econometric Theory* 29, 735–770.
- Hall, P. and Q. Yao** (2003) Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica* 71, 285–317.
- Hamadeh, T. and J-M. Zakoïan** (2011) Asymptotic properties of LS and QML estimators for a class of nonlinear GARCH Processes. *Journal of Statistical Planning and Inference* 141, 488–507.
- Hurlin, C., Laurent, S., Quaadvlieg, R., and S. Smeeke** (2013) Risk measure inference. Available at SSRN 2345299.
- 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.
- 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.
- Zakoïan J.M.** (1994) Threshold Heteroskedastic Models. *Journal of Economic Dynamics and Control*, 18, 931–955.