Testing the existence of moments for GARCH processes

Francq, Christian and Zakoian, Jean-Michel

CREST, CREST

2 December 2019

Testing the existence of moments for GARCH processes

Christian Francq*and Jean-Michel Zakoian†

Abstract

It is generally admitted that many financial time series have heavy tailed marginal distributions. When time series models are fitted on such data, the non-existence of appropriate moments may invalidate standard statistical tools used for inference. Moreover, the existence of moments can be crucial for risk management, for instance when risk is measured through the expected shortfall. This paper considers testing the existence of moments in the framework of GARCH processes. While the second-order stationarity condition does not depend on the distribution of the innovation, higher-order moment conditions involve moments of the independent innovation process. We propose tests for the existence of high moments of the returns process which are based on the joint asymptotic distribution of the Quasi-Maximum Likelihood (QML) estimator of the volatility parameters and empirical moments of the residuals. A bootstrap procedure is proposed to improve the finite-sample performance of our test. To achieve efficiency gains we consider non Gaussian QML estimators founded on reparametrizations of the GARCH model, and we discuss optimality issues. Monte-Carlo experiments and an empirical study illustrate the asymptotic results.

JEL Classification: C12, C13 and C22

Keywords: Conditional heteroskedasticity, Efficiency comparisons, Non-Gaussian QMLE, Residual Bootstrap, Stationarity tests.

---

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

†Corresponding author: Jean-Michel Zakoïan, University of Lille and CREST, 5 Avenue Henri Le Chatelier, 91120 Palaiseau, France. E-mail: zakoian@ensae.fr.

1 Introduction

Testing for the existence of moments of financial time series is of crucial importance. A standard assumption is that prices are nonstationary while returns (or log returns) are (strictly) stationary. However, there is no commonly accepted assumption concerning the existence of moments of such returns. Many searchers in financial econometrics argue that stock returns might not admit 4th-order moments (see e.g. Politis (2007)), while some of them even question the existence of second-order moments. The existence of moments is central to many applications: in presence of heavy tails, many statistical tools developed for the analysis of financial time series become invalid. For instance, using the expected shortfall in risk analysis requires finiteness of the first absolute moment. Long-run horizons predictions of the squared returns require finite unconditional variance of the returns, and their confidence intervals require finite fourth-order moments.

The problem of testing the stationarity, or the finiteness of moments, of financial series has been tackled in different ways in the econometric literature. Loretan and Phillips (1994) investigated nonparametric methods for testing the constancy of the unconditional variance when the fourth unconditional moment is infinite. Trapani (2016) proposed a test for testing existence of the k-th moment of a random variable. A test for second-order stationarity of a time series based on the discrete Fourier transform was developed by Dwivedi and Subba Rao (2011). Other articles focused on the estimation of the tail index, as for instance Kearns and Pagan (1997), Jondeau and Rockinger (2003).

For the log returns, denoted ϵ_t throughout, the most widely used models are arguably the generalized autoregressive conditional heteroscedasticity (GARCH) models introduced by Engle (1982) and Bollerslev (1986), and extended by many authors. Such models are of the form $\epsilon_t = \sigma_t \eta_t$ where σ_t is a positive parametric function of the past returns, and (η_t) is an independent and identically distributed (i.i.d.) sequence, η_t being independent of the past returns. Importantly, the distribution of η_t is generally unspecified - the model can thus be viewed as a semi-parametric formulation. The existence of moments for GARCH-type processes were investigated in several articles. Chen and An (1998) provided sufficient conditions, and Ling and McAleer (2002a) established necessary and sufficient conditions for the existence of fourth and higher moments for the standard and asymmetric GARCH(p, q) models. He and Teräsvita (1999), and Ling and McAleer (2002b) derived such condition for a general family of non-linear GARCH(1,1) models.

A variety of econometric tools, such as the unit root tests, are available for testing the non-stationarity of prices. As far as the returns are concerned, strict stationarity testing as well as the estimation of nonstationary GARCH-type models have been studied by Jensen and Rahbek (2014a, 2014b), Francq and Zakoïan (2012, 2013a), Pedersen and Rahbek (2016), Li, Zhang, Zhu and Ling (2018). To our knowledge, no statistical procedure is available for testing the existence of unconditional moments in the GARCH framework. The main aim of this paper is to develop such procedures for the classical GARCH model. The problem is nonstandard because, except for the second-order moment condition which solely depends on the volatility parameters, the moments conditions for GARCH models involve the distribution of the underlying i.i.d. sequence.

We first use Gaussian QML to derive the joint asymptotic distribution of estimators of the volatility parameters and of moments of the rescaled residuals. A test of the existence of moments of the squared returns will be deduced. A resampling procedure will be considered in order to improve the finite sample properties of the test. The validity of this residual bootstrap procedure will be established. Next, we will show how to improve the power of our tests by using non-Gaussian QML. In particular, optimality properties will be studied.

The paper is structured as follows. Section 2 is devoted to tests of moment existence based on the Gaussian QML estimator for the GARCH(p,q) model. In Section 2.1, the joint distribution of the Gaussian QML estimator and a vector of moments of the residuals is derived. Wald Tests of the $2m$th-order stationarity are deduced in Section 2.2. Bootstrap-based test are studied in Section 2.3. Section 3 is devoted to the efficient testing of the 2nd-order stationarity. Tests based on generalized QML are considered in Section 3.1. Local alternatives and optimality issues are discussed in Section 3.2. Evidence from simulations and real financial time series are provided in Section 4. Concluding remarks are in Section 5. The proofs are presented in the Appendix.

2 Moment testing based on the Gaussian QML

Consider the standard GARCH(p,q) model

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega_0 + \sum_{i=1}^q \alpha_{0i}\epsilon_{t-i}^2 + \sum_{j=1}^p \beta_{0j}\sigma_{t-j}^2 \end{cases} \tag{2.1}$$

where (η_t) is a sequence of i.i.d. variables, and $\theta_0 = (\omega_0, \alpha_{01}, \dots, \beta_{0p})'$ satisfies $\omega_0 > 0, \alpha_{0i} \geq 0$, $\beta_{0j} \geq 0$. Under the assumption $\sum_{j=1}^p \beta_{0j} < 1$, the variable σ_t^2 can be expressed as a function of the

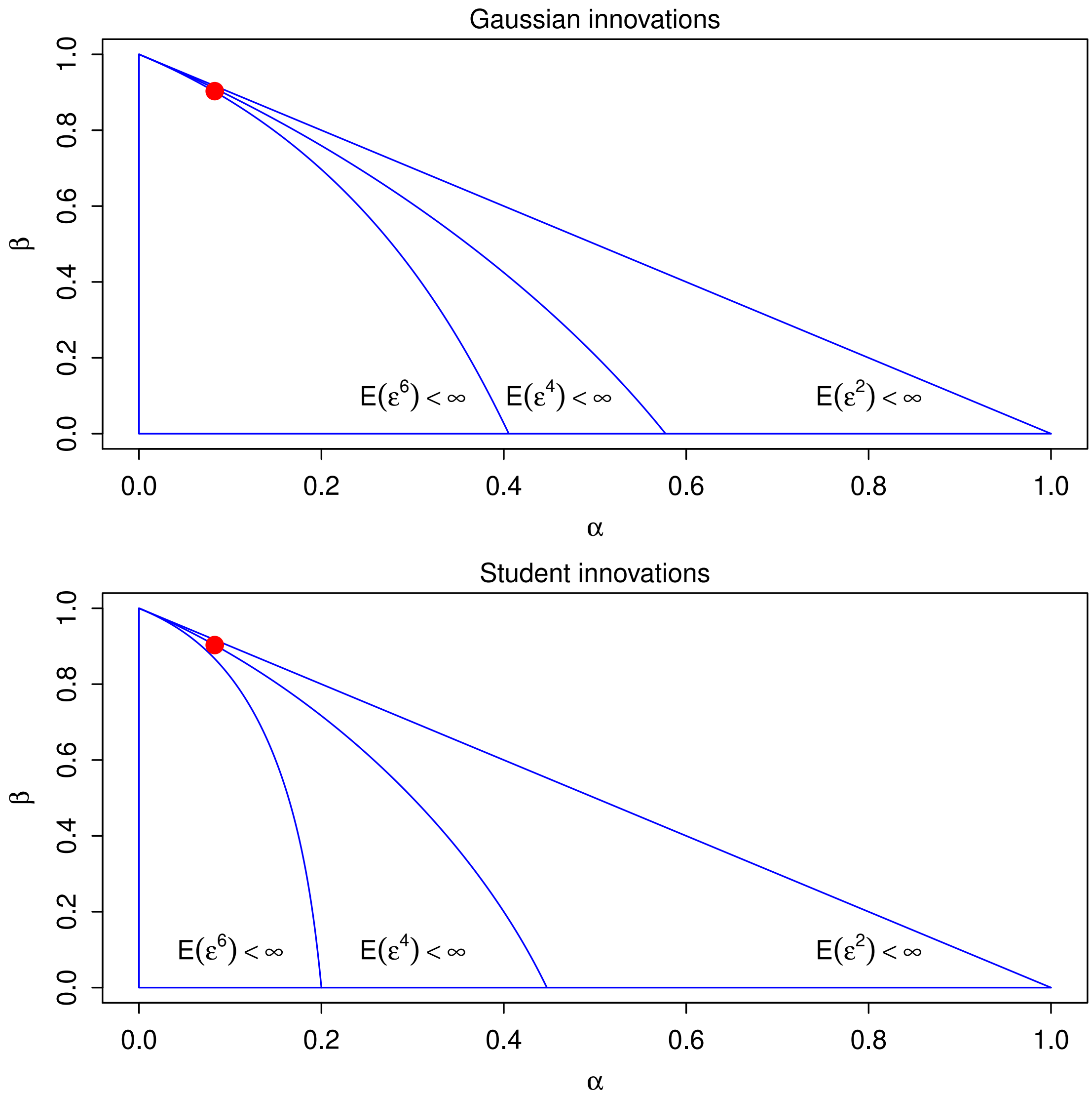

Figure 1: Existence of moments for the GARCH(1,1) model with Gaussian (top panel) and Student(7) (low panel) errors. The bullet indicates a typical value obtained for real stock returns (more precisely, the value estimated in Section 4.2)

infinite past of ϵ_t, as $\sigma_t^2 = \sigma_t^2(\boldsymbol{\theta}_0) = \sigma^2(\epsilon_{t-1}^2, \epsilon_{t-2}^2, \dots; \boldsymbol{\theta}_0)$. Figure 1 displays the regions of existence of the moments, up to the order 6, for the GARCH(1,1) model with two distributions for the error terms: standard Gaussian (top panel) and standardized Student with 7 degrees of freedom (bottom panel). While the 2nd-order moment condition ($\alpha_0 + \beta_0 < 1$) does not depend on the law of η_t, it is seen that the existence of higher-order moments is very sensitive to the moments of η_t. Note also that for small values of α_0, and for β_0 close to 1 (a situation typically reported in empirical studies and marked in the figure by a bullet), the existence of moments is very sensitive to any small variation of the parameters. This shows that testing the existence of moments in the GARCH framework may entail formidable statistical difficulties.

To develop such tests, we turn to the joint estimation of the volatility parameters $\boldsymbol{\theta}_0$ and a vector of moments of the i.i.d. noise (η_t). Given observations $\epsilon_1, \dots, \epsilon_n$, and arbitrary initial values $\tilde{\epsilon}_i$ and $\tilde{\sigma}_j$ for $i \in \{1-q, 2-q, \dots, 0\}$ and $j \in \{1-p, 2-p, \dots, 0\}$, we define, for $t = 1, \dots, n$ and any $\boldsymbol{\theta}$ belonging to a parameter set Θ, $\tilde{\sigma}_t^2(\boldsymbol{\theta}) = \omega_0 + \sum_{i=1}^q \alpha_{0i}\epsilon_{t-i}^2 + \sum_{j=1}^p \beta_{0j}\tilde{\sigma}_{t-j}^2(\boldsymbol{\theta})$, where $\tilde{\sigma}_{t-j}^2(\boldsymbol{\theta}) = \tilde{\sigma}_{t-j}$ for $t \leq j$, and $\epsilon_{t-i} = \tilde{\epsilon}_{t-i}$ for $t \leq i$.

Define the Gaussian QMLE by

$$\hat{\boldsymbol{\theta}}_n = \arg\min_{\boldsymbol{\theta}\in\Theta} \frac{1}{n}\sum_{t=1}^n \tilde{\ell}_t(\boldsymbol{\theta}), \quad \text{where} \quad \tilde{\ell}_t(\boldsymbol{\theta}) = \frac{\epsilon_t^2}{\tilde{\sigma}_t^2(\boldsymbol{\theta})} + \log\tilde{\sigma}_t^2(\boldsymbol{\theta}). \tag{2.2}$$

The following assumptions are required for the strong consistency and asymptotic normality of the Gaussian QMLE. Let $\gamma(\mathbf{A}_0)$ denote the top-Lyapunov exponent associated with Model (2.1) (see Bougerol and Picard (1992)).

A1: $\theta_0 \in \Theta$ and Θ is compact.

A2: $\gamma(\mathbf{A}_0) < 0$, and for all $\theta \in \Theta$, $\sum_{j=1}^p \beta_j < 1$.

A3: η_t^2 has a nondegenerate distribution and $E\eta_t^2 = 1$.

A4: If $p > 0$, $\mathcal{A}_{\theta_0}(z)$ and $\mathcal{B}_{\theta_0}(z)$ have no common roots, $\mathcal{A}_{\theta_0}(1) \neq 0$, and $\alpha_{0q} + \beta_{0p} \neq 0$.

A5: $\theta_0 \in \overset{\circ}{\Theta}$, where $\overset{\circ}{\Theta}$ denotes the interior of Θ.

A6: $E\eta_t^4 < \infty$.

The first part of **A2** is the necessary and sufficient condition established by Bougerol and Picard (1992) for the existence of a strictly stationary solution to the GARCH(p, q) model. Assumptions

A3 and **A4** are made for identifiability reasons in order to get the consistency of $\hat{\boldsymbol{\theta}}_n$. Assumptions **A5** and **A6** are required for the asymptotic normality of the QMLE.

2.1 Asymptotic law of the empirical moments of the rescaled GARCH returns

Let the residuals $\hat{\eta}_t = \epsilon_t/\hat{\sigma}_t$, where $\hat{\sigma}_t = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots, \tilde{\epsilon}_0, \tilde{\epsilon}_{-1}, \dots; \hat{\boldsymbol{\theta}}_n)$. We define, for any $r \geq 0$,

$$\hat{\mu}_r = \frac{1}{n}\sum_{t=1}^{n} |\hat{\eta}_t|^r, \qquad \mu_r = E|\eta_t|^r.$$

For any integer m, let $\hat{\boldsymbol{\mu}}_m = (\hat{\mu}_2, \hat{\mu}_4, \dots, \hat{\mu}_{2m})'$ and $\boldsymbol{\mu}_m = (\mu_2, \mu_4, \dots, \mu_{2m})'$. The following result provides the joint asymptotic distribution of the QMLE and the vector of sample moments of the residuals.

Theorem 2.1. *If* **A1**-**A6** *hold, and if* $\mu_{4m} < \infty$ *then*

$$\begin{pmatrix} \sqrt{n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right) \\ \sqrt{n}(\hat{\boldsymbol{\mu}}_m - \boldsymbol{\mu}_m) \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}\left\{0, \Sigma_m := \begin{pmatrix} (\mu_4 - 1)\boldsymbol{J}^{-1} & -\overline{\boldsymbol{\theta}}_0 \boldsymbol{b}_m' \\ -\boldsymbol{b}_m \overline{\boldsymbol{\theta}}_0' & \boldsymbol{A}_m \end{pmatrix}\right\}, \tag{2.3}$$

where $\overline{\boldsymbol{\theta}}_0 = (\omega_0, \alpha_{01}, \dots, \alpha_{0q}, 0, \dots, 0)'$,

$$\boldsymbol{J} = E\left(\boldsymbol{\phi}_t \boldsymbol{\phi}_t'\right), \quad \boldsymbol{\phi}_t = \boldsymbol{\phi}_t(\boldsymbol{\theta}_0), \quad \boldsymbol{\phi}_t(\boldsymbol{\theta}) = \frac{1}{\sigma_t^2(\boldsymbol{\theta})}\frac{\partial \sigma_t^2}{\partial \boldsymbol{\theta}}(\boldsymbol{\theta}),$$

and $\boldsymbol{A}_m = (a_{ij})_{1\leq i,j\leq m}$, $\boldsymbol{b}_m = (b_i)_{1\leq i\leq m}$, *with*

$$a_{ij} = \mu_{2(i+j)} + \mu_{2i}\mu_{2j}[i + j + (\mu_4 - 1)ij - 1] - i\mu_{2i}\mu_{2(j+1)} - j\mu_{2j}\mu_{2(i+1)}, \quad 1 \leq i, j \leq m,$$

$$b_i = \mu_{2i} - \mu_{2(i+1)} + (\mu_4 - 1)i\mu_{2i}, \qquad 1 \leq i \leq m.$$

Remark 2.1. It is worth noting that the asymptotic variance-covariance matrix $\boldsymbol{A}_m$ of the vector of empirical moments of the rescaled returns does not depend on the parameter $\boldsymbol{\theta}_0$. It solely depends on the moments, up to the order $2m$, of η_t.

Note that $\hat{\boldsymbol{\mu}}_2 = 1$ whence the initial values are such that, for any positive constant K, $K\tilde{\sigma}_t^2(\hat{\boldsymbol{\theta}}_n) = \tilde{\sigma}_t^2(\hat{\boldsymbol{\theta}}_n^*)$ for some $\hat{\boldsymbol{\theta}}_n^* \in \Theta$ (see Francq and Zakoïan (2013b), Remark 4). For more general initial values, the previous theorem yields the following result.

Corollary 2.1. *Under the assumptions of Theorem 2.1, we have*

$$\sqrt{n}(\hat{\boldsymbol{\mu}}_2 - 1) \to 0, \quad \text{in probability as } n \to \infty.$$

2.2 Testing the existence of 2mth-order moments in the GARCH (1,1)

In the GARCH(1,1) case, $\sigma_t^2 = \omega_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2$, the necessary and sufficient condition for the existence of $E(\epsilon_t^{2m})$, where $m \geq 1$ is an integer, is

$$\sum_{i=0}^{m} \binom{m}{i} \alpha_0^i \beta_0^{m-i} \mu_{2i} < 1$$

(see He and Teräsvirta (1999)). Let $G(\boldsymbol{\theta}, \boldsymbol{\mu}) = \sum_{i=0}^{m} \binom{m}{i} \alpha^i \beta^{m-i} \mu_{2i}$ (with $\mu_0 = 1$). Under the assumptions of Theorem 2.1[1] we have

$$\sqrt{n}\{G(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\mu}}_m) - G(\boldsymbol{\theta}_0, \boldsymbol{\mu}_m)\} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_m^2), \tag{2.4}$$

where

$$\sigma_m^2 = \frac{\partial G(\boldsymbol{\theta}_0, \boldsymbol{\mu}_m)}{\partial(\boldsymbol{\theta}', \boldsymbol{\mu}')} \boldsymbol{\Sigma}_m \frac{\partial G(\boldsymbol{\theta}_0, \boldsymbol{\mu}_m)}{\partial \binom{\boldsymbol{\theta}}{\boldsymbol{\mu}}}.$$

Consider the $2m$-th order stationarity problems

$$\boldsymbol{H}_0: \quad E(\epsilon_t^{2m}) < \infty \quad \text{against} \quad \boldsymbol{H}_1: \quad E(\epsilon_t^{2m}) = \infty, \tag{2.5}$$

and

$$\boldsymbol{H}_0^*: \quad E(\epsilon_t^{2m}) = \infty \quad \text{against} \quad \boldsymbol{H}_1^*: \quad E(\epsilon_t^{2m}) < \infty. \tag{2.6}$$

Let the Wald test statistic, with by convention $\hat{\mu}_0 = 1$,

$$T_n = \frac{\sqrt{n}\left\{\sum_{i=0}^{m} \binom{m}{i} \hat{\alpha}_n^i \hat{\beta}_n^{m-i} \hat{\mu}_{2i} - 1\right\}}{\hat{\sigma}_m}, \quad \text{where} \quad \hat{\sigma}_m^2 = \frac{\partial G(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\mu}}_m)}{\partial(\boldsymbol{\theta}', \boldsymbol{\mu}')} \hat{\boldsymbol{\Sigma}}_m \frac{\partial G(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\mu}}_m)}{\partial \binom{\boldsymbol{\theta}}{\boldsymbol{\mu}}}$$

and $\hat{\boldsymbol{\Sigma}}_m$ is a consistent estimator of $\boldsymbol{\Sigma}_m$. The following result is an immediate consequence of the convergence of T_n to the $\mathcal{N}(0,1)$ distribution when $\sum_{i=0}^{m} \binom{m}{i} \alpha_0^i \beta_0^{m-i} \mu_{2i} = 1$.

Proposition 2.1. *Under the assumptions of Theorem 2.1, a test of (2.5) [resp. (2.6)] at the asymptotic level $\alpha \in (0,1)$ is defined by the rejection region*

$$\{T_n > \Phi^{-1}(1-\alpha)\}, \qquad [\textit{resp.} \quad \{T_n < \Phi^{-1}(\alpha)\}], \tag{2.7}$$

where Φ is the $\mathcal{N}(0,1)$ cumulative distribution function.

[1]In the GARCH(1,1) case, the first part of **A2** reduces to $E \log(\alpha_0 \eta_t^2 + \beta_0) < 0$ and **A4** vanishes.

Remark 2.2. As is usual in problems where the null assumption defines an open subset of the parameter set, the test is in fact constructed for the closure of the null assumption. In other words, for $\overline{\boldsymbol{H}}_0: \quad \sum_{i=0}^{m}\binom{m}{i}\alpha_0^i\beta_0^{m-i}\mu_{2i} \leq 1$, the asymptotic region satisfies

$$\sup_{\overline{\boldsymbol{H}}_0} \lim_{n\to\infty} P\{T_n > \Phi^{-1}(1-\alpha)\} = \alpha,$$

where the sup has to be understood as the supremum over all values of $\boldsymbol{\theta}_0$ and error distributions such that $\overline{\boldsymbol{H}}_0$ be satisfied.

Remark 2.3. Proposition 2.1 can in particular be applied for testing the second-order moment condition, $\alpha_0 + \beta_0 < 1$. In this case, the test statistic is given by $T_n = \sqrt{n}(\hat{\alpha} + \hat{\beta} - 1)/\{(\hat{\mu}_4 - 1)\boldsymbol{e}'\hat{\boldsymbol{J}}^{-1}\boldsymbol{e}\}^{1/2}$ where $\boldsymbol{e} = (1,1)'$, and $\hat{\mu}_4$ and $\hat{\boldsymbol{J}}$ are consistent estimators of μ_4 and $\boldsymbol{J}$, respectively.

2.3 Bootstrap-based tests

As we will see in the numerical section, the finite sample distributions of the test statistics are not always in par with the asymptotic results. With the aim of improving the finite sample performance of our tests, we will approximate the test statistic distributions by means of a residual-based bootstrap procedure. Recent papers dealing with bootstrap inference for GARCH-type models are Leucht, Kreiss and Neumann (2015), Beutner, Heinemann and Smeekes (2018), Cavaliere, Nielsen, Pedersen and Rahbek (2018), Heinemann (2019).

We start by presenting the resampling scheme when $m = 1$ (for simplicity in the GARCH(1,1) case).

1. For a GARCH(1,1) model, let a compact parameter space Θ^c whose generic elements are constrained parameters of the form $\boldsymbol{\theta}' = (\omega, \alpha, 1-\alpha)$ with $\omega > 0$ and $0 < \alpha < 1$. Compute the constrained QMLE

$$\hat{\boldsymbol{\theta}}_c' = (\hat{\omega}_c, \hat{\alpha}_c, 1-\hat{\alpha}_c) = \arg\min_{\boldsymbol{\theta}\in\Theta^c} \sum_{t=1}^{n} \tilde{\ell}_t(\boldsymbol{\theta})$$

   and the standardized residuals $\hat{\eta}_t = \tilde{\eta}_t/s_n$, where $\tilde{\eta}_t = \epsilon_t/\tilde{\sigma}_t(\hat{\boldsymbol{\theta}}_c)$ and $s_n^2 = n^{-1}\sum_{t=1}^{n}\tilde{\eta}_t^2$. Denote by F_n^* the empirical distribution of these residuals.

2. Simulate a trajectory of length n of a GARCH model with the parameter $\hat{\boldsymbol{\theta}}_c$ and distribution F_n^* for the i.i.d. noise η_t^*, compute the unconstrained QMLE $\hat{\boldsymbol{\theta}}^* = (\hat{\omega}^*, \hat{\alpha}^*, \hat{\beta}^*)'$ of the GARCH parameter, and compute the statistic $S_n^* = \hat{\alpha}^* + \hat{\beta}^*$.

3. On the observations $\epsilon_1, \dots, \epsilon_n$, compute the unconstrained QMLE $\hat{\boldsymbol{\theta}} = (\hat{\omega}, \hat{\alpha}, \hat{\beta})$ and the statistic $S_n = \hat{\alpha} + \hat{\beta}$.

4. Repeat B times step 2, and denote by $S_n^{*1}, \dots, S_n^{*B}$ the bootstrap test statistic. Approximate the p-value of the test $H_0 : E\epsilon_t^2 < \infty$ against $H_1 : E\epsilon_t^2 = \infty$ by $\#\{S_n^{*j} \geq S_n; j = 1, \dots, B\}/B$, and approximate the p-value of the test $H_0^* : E\epsilon_t^2 = \infty$ against $H_1^* : E\epsilon_t^2 < \infty$ by $\#\{S_n^{*j} \leq S_n; j = 1, \dots, B\}/B$

The numerical optimization required for the computation of the QMLE in Step 2, repeated a large number of times B, is the most time-consuming part of the algorithm. Instead of this step, in view of (A.1), one can mimic the distribution of the QMLE by using a Newton-Raphson type iteration (see *e.g.* Kreiss et al. (2011), Shimizu (2013)). Set

$$\hat{\boldsymbol{\theta}}^* = \hat{\boldsymbol{\theta}}_c + \boldsymbol{J}_n^{-1} \frac{1}{n} \sum_{t=1}^{n} \left(\eta_t^{*\,2} - 1\right) \tilde{\phi}_t(\hat{\boldsymbol{\theta}}_c), \tag{2.8}$$

where

$$\tilde{\phi}_t(\boldsymbol{\theta}) = \frac{1}{\tilde{\sigma}_t(\boldsymbol{\theta})} \frac{\partial \tilde{\sigma}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \qquad \boldsymbol{J}_n = \frac{1}{n} \sum_{t=1}^{n} \tilde{\phi}_t \tilde{\phi}_t'(\hat{\boldsymbol{\theta}}_c)$$

and $\eta_1^*, \dots, \eta_n^*$ are independent and F_n^*-distributed. That resampling algorithm is valid in the following sense.

Theorem 2.2 (Asymptotic validity of the bootstrap procedure)**.** *Let a GARCH(p, q) process* (ϵ_t) *with parameter* $\boldsymbol{\theta}_0$ *such that* $\boldsymbol{c}'\boldsymbol{\theta}_0 = 1$ *with* $\boldsymbol{c}' = (0, 1, \dots, 1)$*, and i.i.d. sequence* (η_t) *satisfying* **A1**-**A6**. *Assume also that the distribution of* η_t *admits a bounded density with respect to the Lebesgue measure. Let* $\hat{\boldsymbol{\theta}}^*$ *be defined by* (2.8)*. For almost all realization* (ϵ_t)*, as* $n \to \infty$ *we have, given* (ϵ_t)*,*

$$\sqrt{n}\,(S_n^* - 1) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2), \qquad \sigma^2 = (\mu_4 - 1)\boldsymbol{c}'\boldsymbol{J}^{-1}\boldsymbol{c}.$$

Note that, in Theorem 2.2, σ^2 corresponds to σ_m^2 in (2.4) with $m = 1$. The previous result thus shows that the distribution of S_n^* given (ϵ_t) well mimics the (unconditional) distribution of S_n at the boundary of H_0, *i.e.* in the case $\boldsymbol{c}'\boldsymbol{\theta}_0 = 1$, at least when n is large. It is also expected that in finite samples the bootstrap distribution of S_n^* better approaches the distribution of S_n than its asymptotic distribution.

We also give informal arguments for the consistency of the bootstrap: under the alternative $\boldsymbol{c}'\boldsymbol{\theta}_0 > 1$, the constrained estimator $\hat{\boldsymbol{\theta}}_c$ should converge to a pseudo-true value $\boldsymbol{\theta}_0^*$, or a set a pseudo-true values (see *e.g.* White, 1994), solution of

$$\boldsymbol{\theta}_0^* = \arg\min_{\boldsymbol{\theta}\in\Theta^c} E\frac{\epsilon_t^2}{\sigma_t^2(\boldsymbol{\theta})} + \log\sigma_t^2(\boldsymbol{\theta})$$

and the distribution of $\sqrt{n}\,(S_n^* - 1) = \sqrt{n}\boldsymbol{c}'\left(\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}}_c\right) = \sqrt{n}\boldsymbol{c}'\left(\hat{\boldsymbol{\theta}}^* - \boldsymbol{\theta}_0^*\right)$ is also expected to be bounded in probability under the alternative, whereas $\sqrt{n}\,(S_n - 1) = \sqrt{n}\boldsymbol{c}'\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right) + \sqrt{n}(\boldsymbol{c}'\boldsymbol{\theta}_0 - 1)$ tends also surely to $+\infty$. Hence the consistency of the bootstrap.

For testing the existence of $E\epsilon_t^{2m}$ when $m > 1$, we generalize the previous resampling scheme as follows.

5. Estimate a GARCH(1,1) model and compute $\hat{\mu}_{2i} = n^{-1}\sum_{t=1}^n \hat{\eta}_t^{2i}$ on the recentred and rescaled residuals.

6. Estimate a GARCH(1,1) model of parameter $\boldsymbol{\theta}_c = (\omega_c, \alpha_c, \beta_c)$ under the constraint H_0 : $\sum_{i=0}^m \binom{m}{i}\alpha_c^i\beta_c^{m-i}\hat{\mu}_{2i} = 1$.

7. Simulate a trajectory of length n of a GARCH model with the parameter $\hat{\boldsymbol{\theta}}_c$ of the previous step, and the empirical distribution of the unconstrained residuals for the i.i.d. noise. Compute the unconstrained QMLE $\hat{\boldsymbol{\theta}}^* = (\hat{\omega}^*, \hat{\alpha}^*, \hat{\beta}^*)'$ and the statistic $S_n^* = \sum_{i=0}^m \binom{m}{i}\hat{\alpha}^{*\,i}\hat{\beta}^{*\,m-i}\hat{\mu}_{2i}^*$ where $\hat{\mu}_{2i}^*$ is computed on the residuals based on $\hat{\boldsymbol{\theta}}^*$.

8. Compute $S_n = \sum_{i=0}^m \binom{m}{i}\hat{\alpha}^i\hat{\beta}^{m-i}\hat{\mu}_{2i}$.

9. As Step 4.

The validity of this bootstrap procedure should follow from the same arguments as those used to prove Theorem 2.2. A recent paper by Heinemann (2019) establishes the validity of a *fixed-design* bootstrap for testing the existence of moments for GARCH processes.

3 Efficient testing of 2nd-order stationarity

In this section, we focus on the second-order stationarity test for the GARCH(p, q) model. Contrary to the higher-order moment conditions, the second-order moment condition does not depend on the distribution of the i.i.d. process. To achieve efficiency gains we do not only consider the *Gaussian* QML, but also alternative QML estimators founded on reparametrizations of the GARCH model.

The estimator of the original parametrization (2.1) is estimated in two steps, as in Francq, Lepage and Zakoïan (2011) (hereafter FLZ).

3.1 Generalized QML based tests

Provided that $E|\eta_t|^r < \infty$, Model (2.1) can be equivalently rewritten as

$$\epsilon_t = \sigma_t(\boldsymbol{\theta}_0^{(r)})\eta_t^{(r)}, \qquad E|\eta_t^{(r)}|^r = 1, \tag{3.1}$$

where $\eta_t^{(r)} = \eta_t/\{E|\eta_t|^r\}^{1/r}$. The link between the parameters of the two formulations, (2.1) and (3.1), is given by

$$\boldsymbol{\theta}_0 = B^{(r)}\boldsymbol{\theta}_0^{(r)}, \quad B^{(r)} = \begin{pmatrix} \mu_r^{-2/r} I_{q+1} & 0 \\ 0 & I_p \end{pmatrix} = \begin{pmatrix} \mu_2^{(r)} I_{q+1} & 0 \\ 0 & I_p \end{pmatrix}. \tag{3.2}$$

In particular, the GARCH persistence coefficients β_{0j} are unchanged in the reparametrization. Let $\mu_s^{(r)} = E|\eta_t^{(r)}|^s$ for any $s > 0$. In the sequel, we omit the upper-script (r) when $r = 2$. Let $\Theta^{(r)}$ such that $\Theta = \{B^{(r)}\boldsymbol{\theta}, \quad \boldsymbol{\theta} \in \Theta^{(r)}\}$. We consider the *generalized* QMLE of $\boldsymbol{\theta}_0^{(r)}$,

$$\widehat{\boldsymbol{\theta}}_n^{(r)} = \underset{\boldsymbol{\theta}\in\Theta^{(r)}}{\operatorname{argmin}}\, \tilde{\mathbf{I}}_n(\boldsymbol{\theta}),$$

where, for $\boldsymbol{\theta} \in \Theta^{(r)}$,

$$\tilde{\mathbf{I}}_n(\boldsymbol{\theta}) = \frac{1}{n}\sum_{t=1}^{n} \tilde{l}_t(\boldsymbol{\theta}) \quad \text{with} \quad \tilde{l}_t(\boldsymbol{\theta}) = \log \tilde{\sigma}_t^2(\boldsymbol{\theta}) + \frac{2}{r}\frac{|\epsilon_t|^r}{\tilde{\sigma}_t^r(\boldsymbol{\theta})}.$$

It was shown in Francq and Zakoïan (2013b), that under the identifiability constraint $E|\eta_t^{(r)}|^r = 1$, the only QMLE which is strongly consistent whatever the error distribution is of the above form.

Define the standardized returns $\widehat{\eta}_t^{(r)} = \frac{\epsilon_t}{\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_n^{(r)})}, \quad t = 1, \ldots, n$. For any $s > 0$ let $\widehat{\mu}_{s,n}^{(r)} = \frac{1}{n}\sum_{t=1}^{n} \left|\widehat{\eta}_t^{(r)}\right|^s$, and let

$$\widehat{B}_n^{(r)} = \begin{pmatrix} \widehat{\mu}_{2,n}^{(r)} I_{q+1} & 0 \\ 0 & I_p \end{pmatrix}.$$

Note that, under appropriate conditions, the generalized QMLE $\widehat{\boldsymbol{\theta}}_n^{(r)}$ converges to $\boldsymbol{\theta}_0^{(r)}$, not to the parameter $\boldsymbol{\theta}_0$ of the standard parametrization. Let $\widehat{\boldsymbol{\theta}}_{n,r}$ be the *two-stage* QMLE (2QMLE) of $\boldsymbol{\theta}_0$ defined as

$$\widehat{\boldsymbol{\theta}}_{n,r} = \widehat{B}_n^{(r)}\widehat{\boldsymbol{\theta}}_n^{(r)}. \tag{3.3}$$

The next result provides the asymptotic properties of this estimator.

Lemma 3.1 (FLZ, Theorem 2.1). *Let $r > 0$. Under Assumptions* **A1-A6***, and if $\mu_{2r} < \infty$, the 2QMLE of $\boldsymbol{\theta}_0$ satisfies*

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_{n,r} - \boldsymbol{\theta}_0\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \Sigma^{(r)}\right) \tag{3.4}$$

with

$$\Sigma^{(r)} = g(r)J^{-1} + \{\mu_4 - 1 - g(r)\}\,\overline{\boldsymbol{\theta}}_0\overline{\boldsymbol{\theta}}_0', \qquad g(r) = \left(\frac{2}{r}\right)^2\left(\frac{\mu_{2r}}{\mu_r^2} - 1\right),$$

and $\overline{\boldsymbol{\theta}}_0 = (\omega_0, \alpha_{01}, \ldots, \alpha_{0q}, 0, \ldots, 0)'$.

Let the null assumption of second-order stationarity

$$\mathbf{H_0}: \quad \sum_{i=1}^{q}\alpha_{0i} + \sum_{j=1}^{p}\beta_{0j} < 1, \quad \text{or, equivalently} \quad \mathbf{H_0}: \quad \boldsymbol{c}'\boldsymbol{\theta}_0 < 1,$$

where $\boldsymbol{c} = (0, 1, \ldots, 1) \in \mathbb{R}^{p+q+1}$, and let $\mathbf{H_1}:\ \boldsymbol{c}'\boldsymbol{\theta}_0 \geq 1$. Let also the null assumption of infinite variance: $\mathbf{H_0^*}:\ \boldsymbol{c}'\boldsymbol{\theta}_0 \geq 1$, and let $\mathbf{H_1^*}:\ \boldsymbol{c}'\boldsymbol{\theta}_0 < 1$. From (3.4) we have

$$\sqrt{n}\boldsymbol{c}'(\widehat{\boldsymbol{\theta}}_{n,r} - \boldsymbol{\theta}_0) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \sigma^{(r)2} := \boldsymbol{c}'\Sigma^{(r)}\boldsymbol{c}\right).$$

Let $\hat{\sigma}^{(r)}$ a consistent estimator of $\sigma^{(r)}$ and let the Wald statistic

$$T_{n,r} = \frac{\sqrt{n}(\boldsymbol{c}'\widehat{\boldsymbol{\theta}}_{n,r} - 1)}{\hat{\sigma}^{(r)}}.$$

The next result is a direct consequence of Lemma 3.1.

Proposition 3.1. *Under the assumptions of Lemma 3.1, a test of $\boldsymbol{H}_0$ [resp. $\boldsymbol{H}_0^*$] at the asymptotic level $\alpha \in (0,1)$ is defined by the rejection region*

$$\mathrm{C}_r = \{T_{n,r} > \Phi^{-1}(1-\alpha)\}, \quad [\textit{resp. } \mathrm{C}_r^* = \{T_{n,r} < \Phi^{-1}(\alpha)\}]. \tag{3.5}$$

3.2 Asymptotic properties under local alternatives

To compare the powers of the different statistic $T_{n,r}$ when r varies, we introduce a sequence of local alternatives. Around $\boldsymbol{\theta}_0$ such that $\boldsymbol{c}'\boldsymbol{\theta}_0 = 1$, let a sequence of local parameters of the form:

$$\boldsymbol{\theta}_n = \boldsymbol{\theta}_0 + \frac{\boldsymbol{\tau}}{\sqrt{n}}$$

where $\boldsymbol{\tau} \in \mathbb{R}^{p+q+1}$. Without loss of generality, assume that n is sufficiently large so that $\boldsymbol{\theta}_n \in \Theta$. We denote by $P_{n,\boldsymbol{\tau}}$ the distribution of the observations $(\epsilon_1, \ldots, \epsilon_n)$ when the parameter is $\boldsymbol{\theta}_0 + \boldsymbol{\tau}/\sqrt{n}$.

3.2.1 Asymptotic local powers

Assume that η_t has a density f which is positive everywhere, with third-order derivatives such that

$$\lim_{|y|\to\infty} yf(y) = 0 \quad \text{and} \quad \lim_{|y|\to\infty} y^2 f'(y) = 0, \tag{3.6}$$

and that, for some positive constants K and δ,

$$|y|\left|\frac{f'}{f}(y)\right| + y^2\left|\left(\frac{f'}{f}\right)'(y)\right| + y^2\left|\left(\frac{f'}{f}\right)''(y)\right| \leq K\left(1+|y|^\delta\right), \tag{3.7}$$

$$E\,|\eta_1|^{2\delta} < \infty. \tag{3.8}$$

These regularity conditions are satisfied for numerous distributions[2].

Proposition 3.2. *Under the assumptions of Proposition 3.1 and under (3.6)-(3.8), the local asymptotic powers of the second-order stationarity tests (3.5) are given by*

$$\lim_{n\to\infty} P_{n,\boldsymbol{\tau}}\left(\mathrm{C}_r\right) = \Phi\left\{\Phi^{-1}(\alpha) + \frac{\boldsymbol{c}'\boldsymbol{\tau}}{\sigma^{(r)}}\right\} \quad \textit{for } \boldsymbol{c}'\boldsymbol{\tau} \geq 0, \tag{3.9}$$

and

$$\lim_{n\to\infty} P_{n,\boldsymbol{\tau}}\left(\mathrm{C}_r^*\right) = \Phi\left\{\Phi^{-1}(\alpha) - \frac{\boldsymbol{c}'\boldsymbol{\tau}}{\sigma^{(r)}}\right\} \quad \textit{for } \boldsymbol{c}'\boldsymbol{\tau} \leq 0.$$

Comparison of the asymptotic powers of the second-order stationarity tests (3.5) when r varies thus boils down to comparing the coefficients $\sigma^{(r)}$: the smaller the latter, the more powerful the test C_r.

Corollary 3.1. *Let $[\underline{r},\overline{r}]$ such that $r_0 = \arg\min_{[\underline{r},\overline{r}]} g(r)$ is well defined. Then, within the family $\{C_r, r \in [\underline{r},\overline{r}]\}$ (resp. $\{C_r^*, r \in [\underline{r},\overline{r}]\}$), for testing $\mathbf{H_0}$ (resp. $\mathbf{H_0^*}$), the test C_{r_0} has the highest local asymptotic power, uniformly in $\boldsymbol{\tau}$.*

Remark 3.1. The optimal value r_0 of r depends on the errors distribution, and is also optimal for the estimator $\widehat{\boldsymbol{\theta}}_{n,r}$ of $\boldsymbol{\theta}_0$ (see FLZ). In the Gaussian case, unsurprisingly, $r_0 = 2$, but for other distributions, the tests based on the *Gaussian* QMLE are far from optimal. For instance, in the case of a Student $t(\nu)$ distribution, r_0 is strictly less than 1 for small values of the degree of freedom ν, and increases to 2 as ν goes to infinity.

Remark 3.2. It has to be noted that a minimum of g over the positive real line may not exist for particular distributions of η_t (see FLZ, Example 2.3). In practice, r_0 is not known but can be consistently estimated under appropriate assumptions (see FLZ, Theorem 3.1).

[2] in particular the Gaussian distribution ($\delta = 2$), the Student's distributions with $\nu > 4$ degrees of freedom ($\delta = 2$).

3.2.2 Optimality issues

Corollary 3.1 allows to determine optimal tests within the class of QML tests of critical regions C_r (or C_r^*). In this section we provide an upper bound for the local powers which, if it is reached, characterizes optimal tests. Optimality means "uniformly most powerful unbiased (UMPU)" (see van der Vaart (1998)).

Proposition 3.3. *Let a strictly stationary GARCH(p,q) model and assume that the error density f satisfies (3.6)-(3.8). Let ι_f the Fisher information for scale*

$$\iota_f = \int \{1 + yf'(y)/f(y)\}^2 f(y)dy < \infty.$$

Then, any test whose critical region satisfies

$$\lim_{n\to\infty} P_{n,\boldsymbol{\tau}}(C) = \Phi\left\{\Phi^{-1}(\alpha) + \frac{\boldsymbol{c}'\boldsymbol{\tau}\sqrt{\iota_f}}{2\sqrt{\boldsymbol{c}'\boldsymbol{J}^{-1}\boldsymbol{c}}}\right\} \quad \text{for } \boldsymbol{c}'\boldsymbol{\tau} \geq 0, \tag{3.10}$$

is UMPU for testing $\mathbf{H_0}$ against $\mathbf{H_1}$.

As a consequence, the test based on the Gaussian QML density is optimal in the following case.

Proposition 3.4. *Under the assumptions of Proposition 3.2, the second-order stationarity test (3.5) with $r = 2$ is asymptotically locally UMPU when the density of η_t has the form*

$$f(y) = \frac{a^a}{\Gamma(a)} e^{-ay^2} |y|^{2a-1}, \quad a > 0, \quad \Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt. \tag{3.11}$$

4 Numerical illustrations

To illustrate the finite sample properties of our test statistics we consider simulated and real financial data.

4.1 Monte-Carlo experiments

In this section, our aims are to (i) study the performance of the tests of Section 2.2 for the existence of $2m$th-order moments; (ii) use the bootstrap procedure of Section 2.3 to see whether the finite sample properties of the tests are improved; (iii) look for efficiency gains by implementing the generalized QML of Section 3.

We first simulated $N = 1000$ independent trajectories of size $n = 2000,\ 4000,\ 8000$ of a GARCH(1,1) process with parameter $(\omega_0, \alpha_0, \beta_0) = (0.5, 0.105, 0.87)$ and $\eta_t \sim \mathcal{N}(0,1)$. In this setting, we have

$$\sum_{i=0}^{m} \binom{m}{i} \alpha_0^i \beta_0^{m-i} \mu_{2i} - 1 = -0.025, -0.027, 0.001, 0.073, 0.216, 0.482$$

for $m = 1,2,3,4,5,6$ respectively. Therefore the moments of order $2m$ are finite for $m \leq 2$ and they are infinite for $m \geq 3$. Table 1 shows that, very often, the tests defined by (2.7) correctly detect that $E\epsilon_t^{2m}$ is finite for $m \leq 2$ and infinite for $m \geq 4$. For $m = 3$, one cannot conclude in general, which is not surprising since $S := \sum_{i=0}^{m} \binom{m}{i} \alpha^i \beta^{m-i} \mu_{2i}$ is very close to 1 when $m = 3$. Note also that, for a correct decision, the sample size n needs to be quite large. A first explanation for the need of large samples is that the parameter $(\alpha_0, \beta_0) = (0.105, 0.87)$ of the generated GARCH model is located in a region where a slight variation of the parameter may entail important modifications in the moments existence (see our comments of Figure 1). Another possible explanation is that the finite sample distribution of the test statistic S_n is far from its Gaussian asymptotic approximation, as will be seen in the following experiment. We simulated $N = 1000$ independent trajectories of a GARCH(1,1) process with parameter $(\omega, \alpha, \beta) = (0.5, 0.10, 0.90)$ and $\eta_t \sim \mathcal{N}(0,1)$. Note that the parameter of the simulated model stands at the boundary of the region of existence of the second-order moment. On each simulation, the GARCH model has been estimated and the statistic $S_n = \hat{\alpha} + \hat{\beta}$ used to test the existence of $E\epsilon_t^2$ has been computed. Figure 2 shows a kernel density estimation of the distribution of the estimator S_n of $S = 1$ for $n = 2000$ and $n = 8000$. Even for the large sample size $n = 8000$, the distribution is clearly negatively skewed, and thus is not well estimated by the Gaussian asymptotic distribution. Other numerical experiments, not presented here, reveal that the problem may be even more pronounced when testing moments moments of order $2m > 2$ and/or when η is not Gaussian.

Table 2 is the analogue of Table 1, but uses the resampling algorithm and rejects the null when the estimated p-value is smaller than the nominal level. The two tables are quite similar but, as expected, the empirical relative frequency of rejection is closer to the nominal level when $m = 3$ (*i.e* S is very close to 1).

Table 1: Relative frequency of rejection of $H_0: E\epsilon_t^{2m} < \infty$ against $H_1: E\epsilon_t^{2m} = \infty$ or of $H_0^*: E\epsilon_t^{2m} = \infty$ against $H_1^*: E\epsilon_t^{2m} < \infty$ at the nominal level $\alpha = 5\%$ or 10%. The null hypothesis H_0 is true for $m = 1, 2$ and false for $m = 3, \ldots, 6$, the null H_0^* is true for $m = 3, \ldots, 6$ and false for $m = 1, 2$.

Null	n	α	$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$	$m = 6$
H_0	2000	5%	0.0	0.0	1.2	14.4	35.8	48.9
		10%	0.0	0.0	4.5	30.6	60.5	80.6
	4000	5%	0.0	0.0	2.4	35.9	77.1	93.1
		10%	0.0	0.0	6.4	53.4	90.0	98.5
	8000	5%	0.0	0.0	3.0	66.8	99.0	99.9
		10%	0.0	0.0	6.9	79.6	99.6	100.0
H_0^*	2000	5%	97.5	48.1	7.9	0.7	0.1	0.1
		10%	99.8	65.9	15.7	1.8	0.1	0.1
	4000	5%	100.0	72.7	7.3	0.1	0.0	0.0
		10%	100.0	85.3	14.7	0.4	0.0	0.0
	8000	5%	100.0	94.1	6.7	0.0	0.0	0.0
		10%	100.0	97.3	14.5	0.0	0.0	0.0

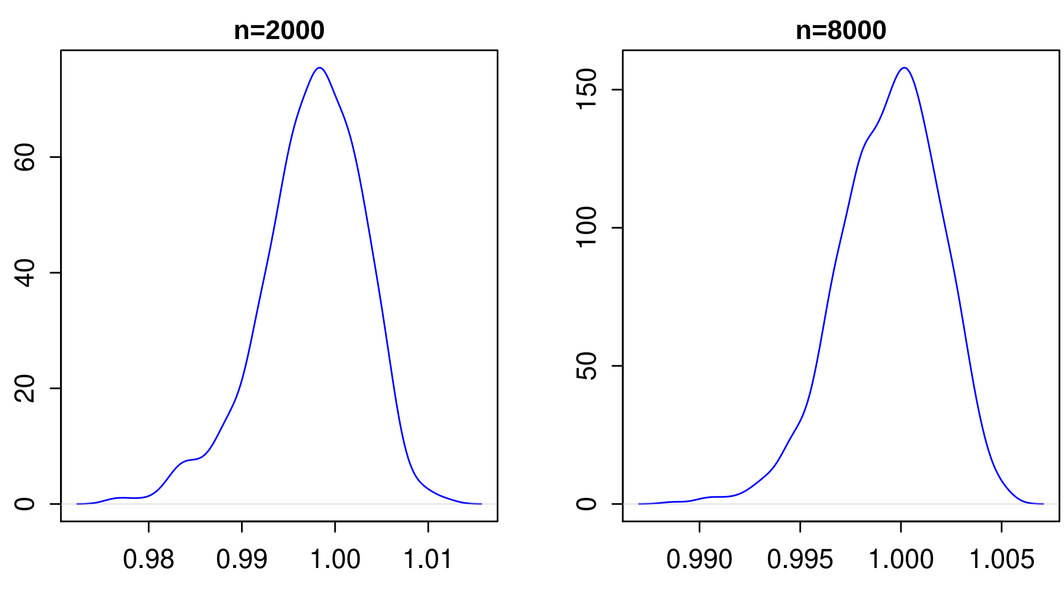

Figure 2: Empirical distribution of S_n.

Table 2: As Table 1, but the resampling algorithm is used instead of the asymptotic distribution.

Null	n	α	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$
H_0	2000	5%	0.0	0.1	3.6	24.8	50.2	72.9
		10%	0.0	0.1	8.3	38.4	67.6	86.8
	4000	5%	0.0	0.0	6.3	42.9	81.5	94.7
		10%	0.0	0.1	11.0	60.2	89.7	98.6
	8000	5%	0.0	0.0	4.3	68.3	97.9	99.8
		10%	0.0	0.0	9.1	81.5	99.4	100.0
H_0^*	2000	5%	83.3	31.2	4.3	0.6	0.0	0.0
		10%	95.1	48.9	9.7	1.3	0.1	0.0
	4000	5%	98.9	51.9	4.5	0.1	0.0	0.0
		10%	100.0	69.8	10.2	0.7	0.0	0.0
	8000	5%	100.0	81.8	5.3	0.0	0.0	0.0
		10%	100.0	93.3	10.2	0.1	0.0	0.0

Now we turn to tests based on non-Gaussian QML. Figure 3 displays the function

$$r \mapsto \hat{g}(r) = \left(\frac{2}{r}\right)^2 \left(\frac{\hat{\mu}_{2r}}{\hat{\mu}_r^2}\right)$$

for $r \in [\underline{r}, \overline{r}]$ when $\eta_t \sim \mathcal{N}(0,1)$. In this distribution, the optimal value r_0 of r, *i.e.* the point where the minimum value of $g(r)$ is reached, is $r_0 = 2$. One can see that $\arg\min_r \hat{g}(r)$ is indeed close to 2 when n is large enough and $\overline{r}$ is not chosen too large. It is actually necessary to impose an upper bound for r because, as shown in Lemma 3.1 of FLZ, when n is fixed, $\hat{g}(r)$ tends to zero as $r \to \infty$.

Table 3 presents results for tests of the existence of second-order moments on 1000 independent simulations of length n of a GARCH(1,1) process when η_t follows a GED(0.3) distribution (normalized so that $E\eta_t^2 = 1$). When $\alpha_0 = 0.1$ and $\beta_0 = 0.8$ we have $\alpha_0 + \beta_0 = 0.9$ (thus $H_0 := E\epsilon_t^2 < \infty$ is true), when $\alpha_0 = 0.105$ and $\beta_0 = 0.87$ we have $\alpha_0 + \beta_0 = 0.975$ (thus H_0 is true), when $\alpha_0 = 0.105$ and $\beta_0 = 0.895$ we have $\alpha_0 + \beta_0 = 1$ (thus we are at the boundary of H_0), when $\alpha_0 = 0.145$ and $\beta_0 = 0.88$ we have $\alpha_0 + \beta_0 = 1.025$ (thus H_0 is false) and when $\alpha_0 = 0.15$ and $\beta_0 = 0.9$ we have $\alpha_0 + \beta_0 = 1.05$ (thus H_0 is false). The columns "QML" are obtained by applying the tests defined in Proposition 2.1 in the case $m = 1$, based on the Gaussian QMLE (see Remark 2.3). For the columns "gQML", we consider the test defined in Proposition 3.1, based on the generalized QMLE

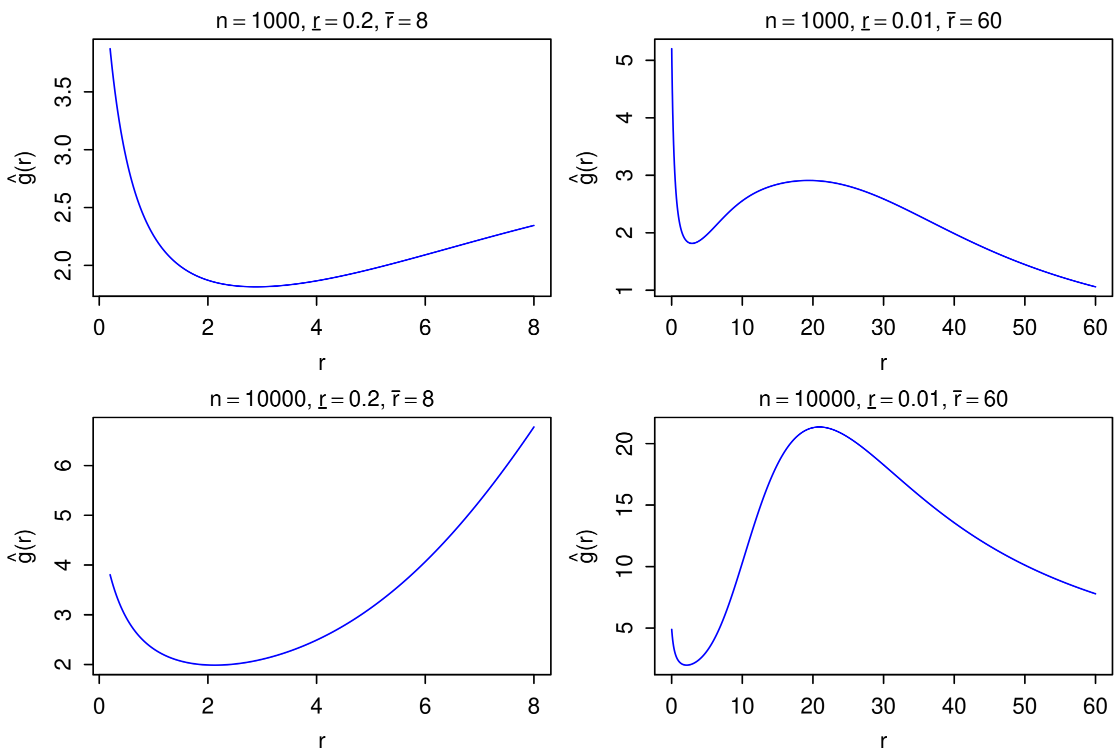

Figure 3: Empirical estimate of the function $g(r)$ when the GARCH innovation $\eta_t \sim \mathcal{N}(0, 1)$.

Table 3: Relative frequency of rejection of $H_0: \ E\epsilon_t^2 < \infty$ against $H_1: \ E\epsilon_t^2 = \infty$ or of $H_0^*: \ E\epsilon_t^2 = \infty$ against $H_1^*: \ E\epsilon_t^2 < \infty$ at the nominal level $\alpha = 5\%$ or 10%, using the Gaussian QML or the generalized QML methods.

(α_0, β_0)			$(0.1, 0.8)$		$(0.105, 0.87)$		$(0.105, 0.895)$		$(0.145, 0.88)$		$(0.15, 0.9)$	
Null	n	α	QML	gQML	QML	gQML	QML	gQML	QML	gQML	QML	gQML
H_0	2000	5%	0.0	0.0	0.2	0.0	0.4	1.8	2.6	8.9	9.9	41.3
		10%	0.2	0.0	0.8	0.4	2.8	5.3	9.7	22.4	27.1	63.2
	4000	5%	0.0	0.0	0.1	0.0	1.2	1.6	6.3	21.0	33.0	76.7
		10%	0.0	0.0	0.8	0.2	4.5	6.3	19.3	37.1	56.9	88.3
	8000	5%	0.0	0.0	0.2	0.0	2.1	3.1	14.8	43.3	67.5	96.4
		10%	0.1	0.0	0.8	0.1	6.2	7.8	31.2	61.6	83.1	98.6
H_0^*	2000	5%	6.5	84.4	2.2	33.5	0.7	10.4	0.5	2.9	0.4	0.4
		10%	25.6	91.1	16.6	47.4	6.8	15.6	4.2	5.0	1.4	0.8
	4000	5%	35.1	98.4	13.7	44.1	5.5	10.1	1.3	1.3	0.1	0.0
		10%	69.8	98.7	35.6	56.1	17.5	16.2	4.3	1.9	0.5	0.0
	8000	5%	87.2	100.0	31.3	58.7	8.2	7.6	1.4	0.2	0.1	0.0
		10%	94.6	100.0	46.5	69.0	15.4	13.3	2.5	0.9	0.2	0.0

where r is replaced by the minimizer of $\hat{g}(r)$ for $r \in [0.001, 2]$) (see Remark 3.2). For both tests, except on the boundary, the rejection frequencies are satisfactory with a clear advantage (for all except 2 cases) for the gQML. For parameters sufficiently far from the boundary, frequencies of rejection of the alternative hypotheses are high. The tests of H_0 appear conservative, the empirical probabilities of incorrect rejection being never greater than the nominal level. On the contrary, the tests of H_0^* generally over-reject the null. A bootstrap procedure was implemented, with the aim of improving the results under the null assumptions. To reduce the computational time, we only implemented the bootstrap for a subset of the parameters and sample sizes. The results reported in Table 4 show that, as expected, the errors of first kind are better controlled.

Table 4: As Table 3, but resampling algorithms are used instead of the asymptotic distributions.

(α_0, β_0)			$(0.1, 0.8)$		$(0.105, 0.895)$		$(0.15, 0.9)$	
Null	n	α	QML	gQML	QML	gQML	QML	gQML
H_0	2000	5%	0.3	0.0	2.7	4.3	21.0	41.0
		10%	1.0	0.1	6.7	8.8	40.0	59.6
H_0^*	2000	5%	14.2	31.9	3.6	3.1	0.2	0.5
		10%	30.7	51.1	8.1	7.1	0.7	0.6

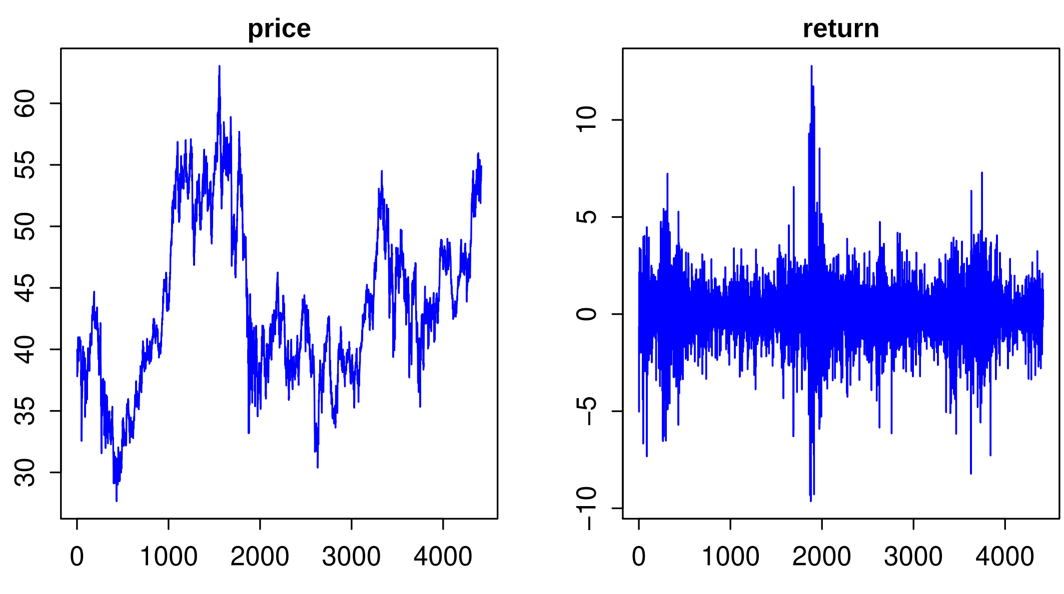

Figure 4: Total stock price and return from 2001-07-16 to 2018-09-21.

4.2 Empirical study

In this section, we consider the daily stock returns of the French energy company Total SA, which constitutes one of the main components of the CAC40 index. The sample path over the period 2001-07-16 to 2018-09-21 is displayed in Figure 4. On the return series, the estimated GARCH(1,1) model is the following (the estimated standard deviations are into brackets):

$$\hat{\omega} = 0.035(0.009), \quad \hat{\alpha} = 0.083(0.011), \quad \hat{\beta} = 0.903(0.011)$$
$$\hat{\mu}_4 = 4.1(0.3), \quad \hat{\mu}_6 = 41.0(12.5), \quad \hat{\mu}_8 = 833.2(482.5),$$
$$\hat{\mu}_{10} = 24572.4(18530.0), \quad \hat{\mu}_{12} = 844199.0(711993.3).$$

The statistics T_n are respectively equal to $-2.96, -0.69, 1.15, 1.62, 1.45, 1.19$ for $m = 1, \ldots, 6$. This provides strong evidence for the existence of moments of order 2, and some evidence of non existence of moments of order 8. Figure 5 displays, for $m = 1, \ldots, 6$, the kernel density estimator of the distribution of S_n under the null that $S = 1$. These estimators were obtained by using $B = 1000$ replications in the above-described resampling algorithm. The value of S_n computed from the observations is represented by the vertical line on the plots. A value of S_n on the left tail of the distribution is a sign that $E\epsilon_t^{2m}$ is finite. Conversely, a value of S_n in the extreme right tail of the distribution indicates that $E\epsilon_t^{2m}$ is likely to be infinite. From this figure, we conclude that $E\epsilon_t^2$ should be finite and $E\epsilon_t^8$ should be infinite, which reinforces the conclusions drawn from the asymptotic theory. In view of Figure 1, it is not surprising that we cannot conclude concerning the existence of moments of order 4 and 6. Indeed, the estimated value belongs to a zone of the parameter space where the different moment conditions are almost undistinguishable.

5 Concluding remarks

Testing for the existence of moments is particularly important for financial times series, whose distributions are thought of as being heavily tailed, even if there is no consensus in the literature about how moments really exist. GARCH models offer a framework for such tests because: i) the existence of moments is explicitly characterized in terms of the volatility parameters and the moments of the errors distribution and ii) a sound theory of estimation is available for such models. Contrary to alternative approaches (e.g. the extreme value theory) for studying the tails of returns,

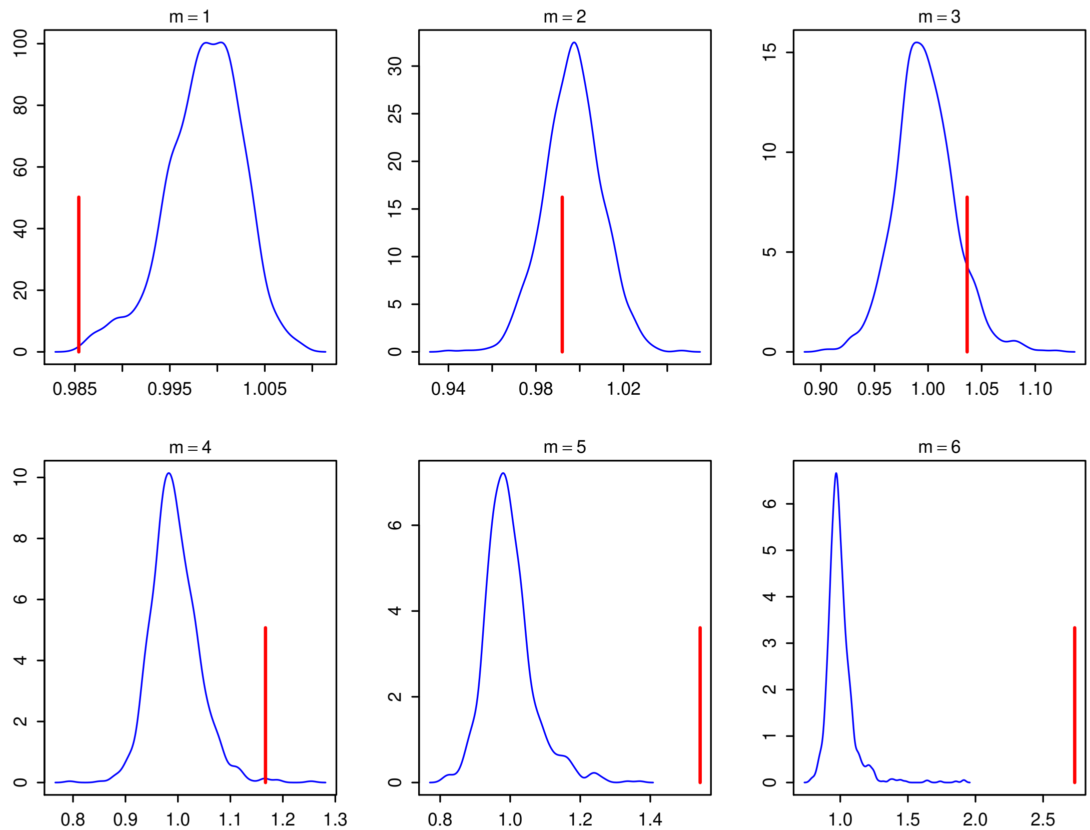

Figure 5: Bootstrap estimates of the distribution of S_n when $S = 1$ (kernel density estimator) and observed value of S_n (vertical line).

the dynamics does not constitute a "nuisance parameter": on the contrary, the dynamics of the series (i.e. the serial dependence) is used to estimate characteristics of the marginal distribution.

In this paper, we have proposed tests for detecting whether the $2m$th moment of a GARCH process is finite. We used QML approaches which do not rely on any distributional assumption on the error process. We derived the asymptotic distribution of tests based on the Gaussian QML, as well as tests relying on a reparametrization of the model enabling the use of alternative QML. We also discussed the choice of an optimal reparameterization. In this article, we focused on the classical GARCH(p, q) model but it is clear that various alternative specifications of the conditional variance (GJR-GARCH, TGARCH, ...) could be handled in a similar fashion.

A general conclusion from our study is that determining if a given moment of a GARCH series exists is a difficult statistical problem. [3] The bootstrap versions of our tests bring significant improvements in terms of size but, as expected, do not improve powers. Even locally optimal tests may be far from conclusive for moderate sample sizes. This suggests that one has to be cautious in assessing the existence, or non-existence, of moments of financial time series.

Acknowledgements

We are most thankful to the Editors of the Special Issue, Chia-Lin Chang, Shiqing Ling and Michael McAleer, and two anonymous referees for insightful comments and suggestions. We are grateful to the Agence Nationale de la Recherche (ANR), which supported this work via the Project MultiRisk (ANR CE26 2016 - CR), and the Labex ECODEC. The second author gratefully thanks the IDR ACP "Régulation et risques systémiques" for financial support, and acknowledges financial support from the Research Initiative "Quantitative Management Research Initiative (QMI)" under the aegis of the Fondation du Risque a joint initiative by University Paris-Dauphine, ENSAE ParisTech, Addstones-GFI and La Française Investment Solutions.

[3] In practice, the situation can even be complicated when the series is contaminated by the presence of outliers (e.g. due to market crashes or rallies). Several authors have proposed statistical methods for detecting the presence of outliers (see for instance Franses and Ghijsels, 1999, Franses and van Dijk, 2011). In such situations, estimation methods that are resistant to outliers are called for (see e.g. Sakata and White, 1998) but it is clear that our tests will no longer be reliable when the model becomes misspecified.

References

Beutner, E., Heinemann, A. and S. Smeekes (2018) A residual bootstrap for conditional Value-at-Risk. *Preprint arXiv:1808.09125.*

Billingsley P. (1999) *Convergence of Probability Measures* 2nd ed. John Wiley, New York.

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

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

Cavaliere, G., Nielsen, H.B., Pedersen R.S. and A. Rahbek (2018) Bootstrap inference on the boundary of the parameter space with application to conditional volatility models. Discussion paper No. 18-10, University of Copenhagen.

Chen, M. and H. Z. An (1998) A note on the stationarity and the existence of moments of the GARCH model. *Statistica Sinica* 8, 505–510.

Drost, F. C. and C. A. J. Klaassen (1997) Efficient estimation in semiparametric GARCH models. *Journal of Econometrics* 81, 193–221.

Drost, F.C., Klaassen, C.A.J. and B.J.M. Werker (1997) Adaptive estimation in time-series models. *Annals of Statistics* 25, 786–817.

Dwivedi, Y. and S. Subba Rao (2011) A test for second-order stationarity of a time series based on the discrete Fourier transform. *Journal of time Series Analysis* 32, 68–91.

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

Francq, C., Jiménez-Gamero, M.D., and S.G. Meintanis (2017) Tests for conditional ellipticity in multivariate GARCH models. *Journal of Econometrics* 196, 305–319.

Francq, C., Lepage, G. and J-M. Zakoïan (2013) 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 (2004) Maximum Likelihood Estimation of Pure GARCH and ARMA-GARCH Processes. *Bernoulli* 10, 605–637.

Francq, C. and J.M. Zakoïan (2006) On efficient inference in GARCH processes. *In: Bertail P, Doukhan P, Soulier P. (eds) Statistics for dependent data.* Springer, New-York: 305–327.

Francq, C. and J.M. Zakoïan (2012) Strict stationarity testing and estimation of explosive and stationary GARCH models. *Econometrica* 80, 821–861.

Francq, C. and J.M. Zakoïan (2013a) Inference in non stationary asymmetric GARCH models. *Annals of Statistics* 41, 1970–1998.

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

Franses, P.H. and H. Ghijsels (1999) Additive outliers, GARCH and forecasting volatility. *International Journal of Forecasting* 15, 1–9.

Franses, P.H. and D. van Dijk (2011) GARCH, outliers, and forecasting volatility. *In: Gregoriou G.N., Pascalau R. (Eds) Nonlinear Financial Econometrics: Forecasting Models, Computational and Bayesian Models.* Palgrave Macmilan, London.

He, C. and T. TerÃď svirta (1999) Properties of moments of a family of GARCH processes. *Journal of Econometrics* 92, 173–192.

Heinemann, A. (2019) A bootstrap test for the existence of moments for GARCH processes. *Preprint arXiv:1902.01808v3.*

Jondeau, E. and M. Rockinger (2003) Testing for differences in the tails of stock-market returns. *Journal of empirical Finance*, 10, 559–581.

Jensen, S.T. and A. Rahbek (2004a) Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case. *Econometrica* 72, 641–646.

Jensen, S.T. and A. Rahbek (2004b) Asymptotic inference for nonstationary GARCH. *Econometric Theory* 20, 1203–1226.

Kreiss, J.P., Paparoditis, E. and D.N. Politis On the range of validity of the autoregressive sieve bootstrap. *Annals of Statistics*, 39, 2103–2130.

Kearns, P. and A. Pagan (1997) Estimating the density tail index for financial time series. *The Review of Economics and Statistics* 79, 171–175.

Lee, S. and M. Taniguchi (2005) Asymptotic theory for ARCH-SM models: LAN and residual empirical processes. *Statistica Sinica* 15, 215–234.

Leucht, A., Kreiss, J. P., and M.H. Neumann (2015) A model specification test for GARCH (1, 1) processes. *Scandinavian Journal of Statistics* 42, 1167–1193.

Li, D., Zhang, X., Zhu, K., and S. Ling (2018) The ZD-GARCH model: A new way to study heteroscedasticity. *Journal of Econometrics* 202, 1–17.

Ling, S. and M. McAleer (2002a) Necessary and sufficient moment conditions for the GARCH(r, s) and asymmetric power GARCH(r, s) processes. *Econometric Theory* 18, 722–729.

Ling, S. and M. McAleer (2002b) Stationarity and the existence of moments of a family of GARCH processes. *Journal of Econometrics* 106, 109–117.

Loretan, M. and P.C.B. Phillips (1994) Testing the covariance stationarity of heavy-tailed time series. *Journal of Empirical Finance* 1, 211–248.

Pedersen, R. S. and A. Rahbek (2016) Nonstationary GARCH with t-distributed innovations. *Economics Letters* 138, 19–21.

Politis, D.N. (2007) Model-free versus model-based volatility prediction. *Journal of Financial Econometrics* 5, 358–359.

Sakata, S. and H. White (1998) High breakdown point conditional dispersion estimation with application to S&P500 daily returns volatility. *Econometrica* 66, 529–567.

Shimizu, K. (2013) The bootstrap does not always work for heteroscedastic models. *Statistics & Risk Modeling*, 30, 189–204.

Trapani, L. (2016) Testing for (in)finite moments. *Journal of Econometrics* 191, 57–68.

van der Vaart, A.W. (1998) *Asymptotic statistics.* Cambridge University Press, United Kingdom.

Wagner, N. and T.A. Marsh (2003) Measuring tail thickness under GARCH and an application to extreme exchange rate changes. *Journal of empirical Finance*, 12, 165–185.

White, H. (1994) *Estimation, Inference and Specification Analysis.* New York: Cambridge University Press.

Appendix: proofs

Proof of Theorem 2.1

By Francq and Zakoian (2004) Theorems 2.1 and 2.2, $\hat{\boldsymbol{\theta}}_n \to \boldsymbol{\theta}_0$ a.s. and

$$\begin{aligned}\sqrt{n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right) &= -\boldsymbol{J}^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(1-\eta_t^2\right)\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\theta}} + o_P(1)\\ &\xrightarrow{\mathcal{L}} \mathcal{N}(0,(\kappa_4-1)\boldsymbol{J}^{-1}).\end{aligned} \tag{A.1}$$

Let $\quad \eta_t(\theta) = \epsilon_t\sigma_t^{-1}(\epsilon_{t-1},\epsilon_{t-2},\dots;\theta), \quad \tilde{\eta}_t(\theta) = \epsilon_t\sigma_t^{-1}(\epsilon_{t-1},\dots,\epsilon_1,\tilde{\epsilon}_0,\tilde{\epsilon}_{-1},\dots;\theta),$

$$\mu_r(\boldsymbol{\theta}) = \frac{1}{n}\sum_{t=1}^{n}|\eta_t(\boldsymbol{\theta})|^r, \qquad \tilde{\mu}_r(\boldsymbol{\theta}) = \frac{1}{n}\sum_{t=1}^{n}|\tilde{\eta}_t(\boldsymbol{\theta})|^r.$$

Using (4.6) in Francq and Zakoian (2004), and arguments similar to those used to prove i) in their Theorem 2.1, it can be shown that

$$\hat{\mu}_r = \tilde{\mu}_r(\hat{\boldsymbol{\theta}}_n) = \mu_r(\hat{\boldsymbol{\theta}}_n) + o_P(n^{-1/2}). \tag{A.2}$$

A Taylor expansion gives, for $\boldsymbol{\theta}^*$ between $\hat{\boldsymbol{\theta}}_n$ and $\boldsymbol{\theta}_0$,

$$\mu_r(\hat{\boldsymbol{\theta}}_n) = \mu_r(\boldsymbol{\theta}_0) + \frac{\partial\mu_r(\boldsymbol{\theta}^*)}{\partial\boldsymbol{\theta}'}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) = \mu_r(\boldsymbol{\theta}_0) + \frac{\partial\mu_r(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\theta}'}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) + o_P(n^{-1/2}) \tag{A.3}$$

where the second equality follows from (A.1), with

$$\frac{\partial\mu_r(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\theta}'} = \frac{-r}{2n}\sum_{t=1}^{n}|\eta_t|^r\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\theta}'} = \frac{-r}{2}\mu_r\boldsymbol{\phi}' + o_P(1),$$

and $\boldsymbol{\phi} = E\left(\boldsymbol{\phi}_t\right)$. This expansion, together with (A.3)-(A.2), gives

$$\sqrt{n}(\hat{\mu}_r - \mu_r(\boldsymbol{\theta}_0)) = \frac{-r}{2}\mu_r\boldsymbol{\phi}'\sqrt{n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right) + o_P(n^{-1/2}),$$

and thus

$$\sqrt{n}(\hat{\mu}_r - \mu_r) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}(|\eta_t|^r - \mu_r) - \frac{r}{2}\mu_r\boldsymbol{\phi}'\sqrt{n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right) + o_P(n^{-1/2}).$$

In view of (A.1) we thus have,

$$\sqrt{n}(\hat{\boldsymbol{\mu}}_m - \boldsymbol{\mu}_m) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}(\boldsymbol{Z}_{t,m} - \boldsymbol{\mu}_m) + \boldsymbol{\nu}_m\boldsymbol{\phi}'\boldsymbol{J}^{-1}\left(1 - \eta_t^2\right)\boldsymbol{\phi}_t + o_P(n^{-1/2}),$$

where $\boldsymbol{Z}_{t,m} = (\eta_t^2, \eta_t^4, \ldots, \eta_t^{2m})'$, $\boldsymbol{\nu}_m = (\mu_2, 2\mu_4, \ldots, m\mu_{2m})'$.

The asymptotic normality in Theorem 2.1 follows by the Wold-Cràmer device and the central limit theorem for martingale differences. Using the equality $\boldsymbol{\phi}'\boldsymbol{J}^{-1}\boldsymbol{\phi} = 1$ (see Remark 3 in Francq and Zakoian (2013b)) we have,

$$\mathrm{Var}_{as}\{\sqrt{n}(\hat{\boldsymbol{\mu}}_m - \boldsymbol{\mu}_m)\} = \mathrm{Var}(\boldsymbol{Z}_{t,m}) + E[\boldsymbol{Z}_{t,m}(1-\eta_t^2)]\boldsymbol{\nu}_m' + \boldsymbol{\nu}_m E[\boldsymbol{Z}_{t,m}'(1-\eta_t^2)] + \boldsymbol{\nu}_m\boldsymbol{\nu}_m'(\boldsymbol{\mu}_4 - 1),$$

and

$$\mathrm{Cov}_{as}\{\sqrt{n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right), \sqrt{n}(\hat{\boldsymbol{\mu}}_m - \boldsymbol{\mu}_m)\} = -\boldsymbol{J}^{-1}\boldsymbol{\phi}[E\{(1-\eta_t^2)\boldsymbol{Z}_{t,m}'\} + (\mu_4 - 1)\boldsymbol{\nu}_m'].$$

The conclusion follows by noting that $\boldsymbol{J}^{-1}\boldsymbol{\phi} = \overline{\boldsymbol{\theta}}_0$. □

Proof of Corollary 2.1

It suffices to remark that the asymptotic law of $\sqrt{n}(\hat{\boldsymbol{\mu}}_2 - 1)$ is degenerate: indeed, $b_1 = a_{i1} = a_{1j} = 0$ for all i and j. □

Proof of Theorem 2.2

We start by showing a lemma.

Lemma A.1. *Suppose that the assumptions of Theorem 2.2 are satisfied. Conditionally on almost all realizations* (ϵ_t) *of the GARCH(p, q) process, the distribution* F_n^* *of the standardized residuals tends to the unconditional distribution* F *of* η_t*. Moreover, for almost all realizations* (ϵ_t) *and any* $A \in [-\infty, \infty)$*, as* $n \to \infty$

$$\frac{1}{n}\sum_{t=1}^{n} \hat{\eta}_t^4 1_{\hat{\eta}_t \geq A} = \int_A^\infty x^4 F_n^*(dx) \to \int_A^\infty x^4 F(dx). \tag{A.4}$$

More generally, for any real function h *such that* $|h(x| \leq ax^4 + b$ *where* $a, b > 0$*, and the set* D_h *of its discontinuities verifies* $P(\eta_t \in D_h) = 0$*, we have*

$$\int h(x) F_n^*(dx) \to \int h(x) F(dx). \tag{A.5}$$

Proof of Lemma A.1. The proof is inspired by that of Lemmas 8.6 and 8.7 in Francq, JimÃĺnez-Gamero and Meintanis (2017). Let $\eta_t(\boldsymbol{\theta}) = \epsilon_t/\sigma_t(\boldsymbol{\theta})$ and $\tilde{\eta}_t(\boldsymbol{\theta}) = \epsilon_t/\tilde{\sigma}_t(\boldsymbol{\theta})$, so that $\tilde{\eta}_t = \tilde{\eta}_t(\hat{\boldsymbol{\theta}}_c)$ and $\eta_t = \eta_t(\boldsymbol{\theta}_0)$. In Francq and Zakoian (2004), it is shown that

$$\sup_{\theta \in \Theta} |\sigma_t(\boldsymbol{\theta}) - \tilde{\sigma}_t(\boldsymbol{\theta})| \leq K\rho^t, \tag{A.6}$$

where, here and in the sequel, K denotes a generic positive variable depending on $\{\eta_t, t \leq 0\}$ and ρ denotes a generic constant belonging to $[0, 1)$. We thus have

$$\sup_{\theta \in \Theta} |\eta_t(\boldsymbol{\theta}) - \tilde{\eta}_t(\boldsymbol{\theta})| \leq \frac{K}{\underline{\omega}} \rho^t |\epsilon_t|$$

where $\underline{\omega}$ is a positive lower bound for ω over the compact set Θ. By the mean value theorem

$$\eta_t(\hat{\boldsymbol{\theta}}_c) = \eta_t + \frac{\partial \eta_t(\boldsymbol{\theta}_n)}{\partial \boldsymbol{\theta}'} \left(\hat{\boldsymbol{\theta}}_c - \boldsymbol{\theta}_0\right),$$

with $\boldsymbol{\theta}_n$ between $\hat{\boldsymbol{\theta}}_c$ and $\boldsymbol{\theta}_0$. In Francq and Zakoian (2004), it is also shown that for any d there exists a neighborhood $V(\boldsymbol{\theta}_0)$ of $\boldsymbol{\theta}_0$ such that

$$E \sup_{\theta \in V(\boldsymbol{\theta}_0)} \left|\frac{\sigma_t(\boldsymbol{\theta}_0)}{\sigma_t(\boldsymbol{\theta})}\right|^d < \infty, \quad E \sup_{\theta \in \Theta} \|\phi_t(\boldsymbol{\theta})\|^d < \infty. \tag{A.7}$$

This entails that

$$\sup_{\theta \in V(\boldsymbol{\theta}_0)} \left\|\frac{\partial \eta_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right\| = \sup_{\theta \in V(\boldsymbol{\theta}_0)} \left\|\frac{\sigma_t(\boldsymbol{\theta}_0)}{\sigma_t(\boldsymbol{\theta})} \phi_t(\boldsymbol{\theta}) \eta_t\right\| = u_t |\eta_t|,$$

where $u_t \in \mathcal{F}_{t-1}$ and $Eu_t^{d/2} < \infty$. We thus have

$$|\tilde{\eta}_t - \eta_t| \leq K \left(\rho^t + \left\| \hat{\boldsymbol{\theta}}_c - \boldsymbol{\theta}_0 \right\| \right) u_t |\eta_t|, \tag{A.8}$$

for n large enough. It follows that, for almost all sequence (ϵ_t), or equivalently almost all sequence (η_t),

$$s_n^2 = \frac{1}{n} \sum_{t=1}^{n} \eta_t^2 + \frac{1}{n} \sum_{t=1}^{n} (\tilde{\eta}_t - \eta_t)^2 + \frac{2}{n} \sum_{t=1}^{n} \eta_t (\tilde{\eta}_t - \eta_t) \to 1$$

as $n \to \infty$. Since

$$\hat{\eta}_t - \eta_t = \frac{1}{s_n} (\tilde{\eta}_t - \eta_t) + \left(\frac{1}{s_n} - 1 \right) \eta_t, \tag{A.9}$$

we have

$$|\hat{\eta}_t - \eta_t| \leq \left(\rho^t + a_n \right) v_t |\eta_t|,$$

for n large enough, where $v_t = 2Ku_t + 1$ and $a_n = \left\| \hat{\boldsymbol{\theta}}_c - \boldsymbol{\theta}_0 \right\| + \left(\frac{1}{s_n} - 1 \right)$ tends to 0. For all $x \in \mathbb{R}$, all $\varepsilon > 0$ and all $M > 0$, we then have

$$\begin{aligned} \left| 1_{\{\hat{\eta}_t \leq x\}} - 1_{\{\eta_t \leq x\}} \right| &\leq 1_{\{x - (\rho^t + a_n) v_t |\eta_t| \leq \eta_t \leq x + (\rho^t + a_n) v_t |\eta_t|\}} \\ &\leq 1_{A_{t,\varepsilon,M}} + 1_{a_n > \varepsilon} + 1_{|\eta_t| > M}, \end{aligned}$$

with the event

$$A_{t,\varepsilon,M} = \left\{ x - \left(\rho^t + \varepsilon \right) v_t M \leq \eta_t \leq x + \left(\rho^t + \varepsilon \right) v_t M \right\}.$$

For $t \geq \log \varepsilon / \log \rho$, we have $A_{t,\varepsilon,M} \subset A_{2\varepsilon,M}$ with

$$A_{\varepsilon,M} = \{ x - \varepsilon v_t M \leq \eta_t \leq x + \varepsilon v_t M \}.$$

Taking $d \geq 2$, we have

$$E 1_{A_{\varepsilon,M}} = EE \left(1_{A_{\varepsilon,M}} \mid \mathcal{F}_{t-1} \right) = E \int_{x - \varepsilon v_t M}^{x + \varepsilon v_t M} f(y) dy \leq 2 \max_{y \in \mathbb{R}} f(y) \varepsilon M E v_t.$$

For all $\kappa > 0$, we thus have a small $\varepsilon > 0$ and a large $M > 0$ such that

$$E \left\{ 1_{A_{2\varepsilon,M}} + 1_{a_n > \varepsilon} + 1_{|\eta_t| > M} \right\} \leq \kappa.$$

It follows that, for almost all sequences (ϵ_t), we have

$$\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}1_{\{\hat{\eta}_t\leq x\}}=\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}1_{\{\eta_t\leq x\}}=P\left(\eta_t\leq x\right),\quad \forall x\in\mathbb{R}.$$

We have shown that, for almost all (ϵ_t), F_n^* weakly converges to F.

Now note that by (A.8) we have

$$\frac{1}{n}\sum_{t=1}^{n}\left|\tilde{\eta}_t-\eta_t\right|^k\to 0$$

for $k=1,2,3,4$, assuming without loss of generality that $d\geq 8$. Since $\frac{1}{n}\sum_{t=1}^{n}\eta_t^4\to\mu_4$, this implies

$$\frac{1}{n}\sum_{t=1}^{n}\tilde{\eta}_t^4\to\mu_4.$$

We then obtain (A.4) with $A=-\infty$ from (A.9) and the convergence of s_n to 1.

By the continuous mapping theorem, given almost all sequence (ϵ_t), a random sequence (X_n,Y_n) with uniform distribution on $\{(\hat{\eta}_t^4,h(\hat{\eta}_t)),\,t=1,\dots,n\}$ converges in distribution to a random vector $(X,Y)=(\eta^4,h(\eta))$ where $\eta\sim F$. Having shown (A.4) with $A=-\infty$, we already know that $E(X_n\mid(\epsilon_t))\to EX$. Theorem 3.6 in Billingsley (1999) then shows that the sequence X_n is uniformly integrable, given (ϵ_t). By Theorem 3.5 in Billingsley (1999), to show (A.5), that is $E(Y_n\mid(\epsilon_t))\to EY$, it remains to show that Y_n is uniformly integrable, which is obvious because $|Y_n|\leq aX_n+b$. The proof of Lemma A.1 is complete.

Now we turn to the proof of Theorem 2.2. We have, in view of (2.8),

$$\sqrt{n}\left(S_n^*-1\right)=\boldsymbol{c}'\sqrt{n}\left(\hat{\boldsymbol{\theta}}^*-\hat{\boldsymbol{\theta}}_c\right)=\boldsymbol{c}'\boldsymbol{J}_n^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\tilde{\boldsymbol{x}}_{t,n},$$

with $\tilde{\boldsymbol{x}}_{t,n}=(\eta_t^{*2}-1)\,\tilde{\phi}_t(\hat{\boldsymbol{\theta}}_c)$. The index n in $\tilde{\boldsymbol{x}}_{t,n}$ emphasizes that the distribution F_n^* of η_t^*, as well as $\hat{\boldsymbol{\theta}}_c$, depend on n. Let $\boldsymbol{x}_{t,n}=(\eta_t^{*2}-1)\,\phi_t(\hat{\boldsymbol{\theta}}_c)$. In view of (A.6) and a similar inequality for the derivatives, we have

$$\sup_{\boldsymbol{\theta}\in\Theta}\left|\phi_t(\boldsymbol{\theta})-\tilde{\phi}_t(\boldsymbol{\theta})\right|\leq K\rho^t u_t,$$

where $u_t:=\sup_{\boldsymbol{\theta}\Theta}\phi_t(\boldsymbol{\theta})+1$ admits moments of any order. It follows that

$$\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\boldsymbol{x}_{t,n}-\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\tilde{\boldsymbol{x}}_{t,n}\right|\leq\frac{K}{\sqrt{n}}\sum_{t=1}^{\infty}\left|\eta_t^{*2}-1\right|\rho^t u_t\to 0\quad\text{ a.s.}$$

as $n \to \infty$, noting that the previous series is a.s. finite because its expectation is finite. Moreover, by the standard arguments of Francq and Zakoian (2004), it can be shown that $\boldsymbol{J}_n$ converges to the invertible matrix $\boldsymbol{J} = E\boldsymbol{\phi}_t\boldsymbol{\phi}_t'(\boldsymbol{\theta}_0)$ as $n \to \infty$.

It thus remains to show that, conditional on (ϵ_t),

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\boldsymbol{x}_{t,n} \xrightarrow{d} \mathcal{N}\left(0, (\mu_4 - 1)\boldsymbol{J}\right). \tag{A.10}$$

Note that, conditional on (ϵ_t), for each n the random vectors $\boldsymbol{x}_{1,n}, \boldsymbol{x}_{2,n}, \ldots$ are independent and centered, with finite second-order moments. From Lindeberg's CLT for triangular arrays of square integrable martingale increments, and the Wold-Cramer device, it suffices to show that for any $\boldsymbol{\lambda} \in \mathbb{R}^3$, $\boldsymbol{\lambda} \neq \boldsymbol{0}$,

$$\frac{1}{n}\sum_{t=1}^{n}\mathrm{Var}\left(\boldsymbol{\lambda}'\boldsymbol{x}_{t,n}\right) \to (\mu_4 - 1)\boldsymbol{\lambda}'\boldsymbol{J}\boldsymbol{\lambda} > 0 \quad \text{ as } n \to \infty, \tag{A.11}$$

and for all $\varepsilon > 0$

$$\frac{1}{n}\sum_{t=1}^{n} E\left(\left\{\boldsymbol{\lambda}'\boldsymbol{x}_{t,n}\right\}^2 1_{\{|\boldsymbol{\lambda}'\boldsymbol{x}_{t,n}| \geq \sqrt{n}\varepsilon\}}\right) \to 0 \quad \text{ as } n \to \infty. \tag{A.12}$$

Note that, given (ϵ_t), only the term η_t^* is random in $\boldsymbol{x}_{t,n}$. Moreover, if $\eta \sim F_n^*$, then $E\eta = 0$, $E\eta^2 = 1$ and, by (A.4) in Lemma A.1, $E\eta^4 \to \mu_4$ as $n \to \infty$. Given (ϵ_t), as $n \to \infty$ we thus have

$$\mathrm{Var}\boldsymbol{\lambda}'\boldsymbol{x}_{t,n} = \left\{\boldsymbol{\lambda}'\boldsymbol{\phi}_t(\hat{\boldsymbol{\theta}}_c)\right\}^2 (E\eta^4 - 1) \to \left\{\boldsymbol{\lambda}'\boldsymbol{\phi}_t\right\}^2 (\mu_4 - 1).$$

Moreover, for all $\varepsilon > 0$ there exists a neighborhood $V(\boldsymbol{\theta}_0)$ of $\boldsymbol{\theta}_0$ such that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0)}\left|\left\{\boldsymbol{\lambda}'\boldsymbol{\phi}_t(\boldsymbol{\theta})\right\}^2 - \left\{\boldsymbol{\lambda}'\boldsymbol{\phi}_t(\boldsymbol{\theta}_0)\right\}^2\right| \leq \varepsilon.$$

The previous result is obtained by using the ergodic theorem, the continuity of $\boldsymbol{\theta} \mapsto E\left|\left\{\boldsymbol{\lambda}'\boldsymbol{\phi}_t(\boldsymbol{\theta})\right\}^2 - \left\{\boldsymbol{\lambda}'\boldsymbol{\phi}_t(\boldsymbol{\theta}_0)\right\}^2\right|$, the dominated convergence theorem, and by shrinking the neighborhood. Now the consistency of $\hat{\boldsymbol{\theta}}_c$ and the ergodic theorem entail (A.11), noting that under **A2** and **A4**, $0 < \mu_4 - 1 < \infty$.

Now we turn to the proof of (A.12). Given (ϵ_t), for some neighborhood $V(\boldsymbol{\theta}_0)$ of $\boldsymbol{\theta}_0$ and

n large enough we have

$$
\begin{aligned}
& E\left\{\boldsymbol{\lambda}'\boldsymbol{x}_{t,n}\right\}^2 1_{\{|\boldsymbol{\lambda}'\boldsymbol{x}_{t,n}|\geq\sqrt{n}\varepsilon\}} \\
& \leq 1_{\left\{\sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0)}\sup_{t\geq 1}|\boldsymbol{\lambda}'\boldsymbol{\phi}_t(\boldsymbol{\theta})|>0\right\}} \sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0)} \sup_{t\geq 1} \left\{\boldsymbol{\lambda}'\boldsymbol{\phi}_t(\boldsymbol{\theta})\right\}^2 \\
& \times E\left|\eta_t^{*2}-1\right|^2 1_{\left\{\left|\eta_t^{*2}-1\right|\geq\frac{\sqrt{n}\varepsilon}{\sqrt{\sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0)}\sup_{t\geq 1}|\boldsymbol{\lambda}'\boldsymbol{\phi}_t(\boldsymbol{\theta})|}}\right\}}.
\end{aligned}
\tag{A.13}
$$

For any $A > 0$ there exists n_A such that if $n > n_A$ then the expectation in the right-hand side of (A.13) is bounded by

$$
E\left|\eta_t^{*2}-1\right|^2 1_{\left\{\left|\eta_t^{*2}-1\right|\geq A\right\}}.
$$

By Lemma A.1, this terms tends to

$$
\int_{|x^2-1|\geq A} \left|x^2-1\right|^2 F(dx)
$$

which is arbitrarily small when A is sufficiently large. We then obtain (A.12) by already given arguments. □

Proof of Proposition 3.2

Under the strict stationarity condition $\gamma_0 < 0$, Drost and Klaassen (1997) showed that, for standard GARCH, the log-likelihood ratio $\Lambda_{n,f}(\theta_n, \theta_0) = \log L_{n,f}(\theta_n)/L_{n,f}(\theta_0)$ satisfies the LAN property

$$
\Lambda_{n,f}(\theta_n, \theta_0) = \boldsymbol{\tau}' S_{n,f}(\theta_0) - \frac{1}{2}\boldsymbol{\tau}'\mathfrak{I}_f\boldsymbol{\tau} + o_{P_{\theta_0}}(1),
\tag{A.14}
$$

where $S_{n,f}(\boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}\left\{0, \mathfrak{I}_f\right\}$ under $P_{\boldsymbol{\theta}_0}$ as $n \to \infty$ and $\mathfrak{I}_f = \frac{\iota_f}{4}\boldsymbol{J}$.

Note that the so-called central sequence $S_{n,f}$ is conditional on the initial values. In the stationary case, Lee and Taniguchi (2005) showed that the initial values have no influence on the LAN property. Let the functions

$$
g_1(y) = 1 + y\frac{f'}{f}(y) \quad \text{and} \quad g_2(y) = 1 + 2y\frac{f'}{f}(y) + y^2\left(\frac{f'}{f}\right)'(y).
$$

We have

$$
S_{n,f}(\theta_0) = \frac{-1}{2\sqrt{n}}\sum_{t=1}^{n} g_1(\eta_t)\boldsymbol{\phi}_t
\tag{A.15}
$$

and thus

$$\Lambda_{n,f}(\theta_n,\theta_0)=\frac{-\boldsymbol{\tau}'}{2\sqrt{n}}\sum_{t=1}^{n}g_1(\eta_t)\boldsymbol{\phi}_t-\frac{1}{2}\boldsymbol{\tau}'\mathfrak{I}_f\boldsymbol{\tau}+o_{P_{\theta_0}}(1). \tag{A.16}$$

By FLZ, letting $\boldsymbol{\phi}_t^{(r)}=\boldsymbol{\phi}_t(\boldsymbol{\theta}_0^{(r)})$ and $\boldsymbol{\phi}^{(r)}=E\boldsymbol{\phi}_t^{(r)}$,

$$\sqrt{n}\begin{pmatrix}\hat{\mu}_{2,n}^{(r)}-\mu_2^{(r)}\\ \widehat{\boldsymbol{\theta}}_n^{(r)}-\boldsymbol{\theta}_0^{(r)}\end{pmatrix}=\begin{pmatrix}1 & -\mu_2^{(r)}\boldsymbol{\phi}^{(r)'}\{\boldsymbol{J}^{(r)}\}^{-1}\\ \boldsymbol{0} & \{\boldsymbol{J}^{(r)}\}^{-1}\end{pmatrix}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\begin{pmatrix}|\eta_t^{(r)}|^2-\mu_2^{(r)}\\ \boldsymbol{\phi}_t^{(r)}\left(|\eta_t^{(r)}|^r-1\right)\end{pmatrix}+o_P(1),$$

where $\boldsymbol{J}^{(r)}=E\left(\frac{r}{2}\boldsymbol{\phi}_t^{(r)}\boldsymbol{\phi}_t^{(r)'}\right)$. We also have

$$\begin{aligned}
&\sqrt{n}(\boldsymbol{c}'\widehat{\boldsymbol{\theta}}_{n,r}-1)\\
=\;&\sqrt{n}\boldsymbol{c}'\widehat{B}_n^{(r)}\widehat{\boldsymbol{\theta}}_n^{(r)}-\sqrt{n}\boldsymbol{c}'B^{(r)}\boldsymbol{\theta}_0^{(r)}\\
=\;&\boldsymbol{c}'B^{(r)}\sqrt{n}(\widehat{\boldsymbol{\theta}}_n^{(r)}-\boldsymbol{\theta}_0^{(r)})+\boldsymbol{c}'\sqrt{n}(\widehat{B}_n^{(r)}-B^{(r)})\widehat{\boldsymbol{\theta}}_n^{(r)}\\
=\;&\left(\sum_{i=1}^{q}\alpha_{0i}^{(r)}\quad \boldsymbol{c}'B^{(r)}\right)\sqrt{n}\begin{pmatrix}\hat{\mu}_{2,n}^{(r)}-\mu_2^{(r)}\\ \widehat{\boldsymbol{\theta}}_n^{(r)}-\boldsymbol{\theta}_0^{(r)}\end{pmatrix}+o_P(1)\\
=\;&\left(\underline{\alpha}_0^{(r)}\quad \left\{-\underline{\alpha}_0^{(r)}\mu_2^{(r)}\boldsymbol{\phi}^{(r)'}+\boldsymbol{c}'B^{(r)}\right\}\{\boldsymbol{J}^{(r)}\}^{-1}\right)\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\begin{pmatrix}|\eta_t^{(r)}|^2-\mu_2^{(r)}\\ \boldsymbol{\phi}_t^{(r)}\left(|\eta_t^{(r)}|^r-1\right)\end{pmatrix}+o_P(1)
\end{aligned}$$

where $\underline{\alpha}_0^{(r)}=\sum_{i=1}^{q}\alpha_{0i}^{(r)}$. Let $\overline{\boldsymbol{\theta}}_0=(\omega_0,\alpha_{01},\dots,\alpha_{0q},0,\dots,0)'$. Noting that

$$\mu_2^{(r)}=\mu_r^{-2/r},\quad \underline{\alpha}_0^{(r)}=\mu_r^{2/r}\underline{\alpha}_0,\quad \boldsymbol{\phi}_t^{(r)}=B^{(r)}\boldsymbol{\phi}_t,\quad \boldsymbol{J}^{(r)}=\frac{r}{2}B^{(r)}\boldsymbol{J}B^{(r)},\quad \mu^{2/r}\overline{\boldsymbol{\theta}}_0'B^{(r)}=\overline{\boldsymbol{\theta}}_0'$$

we get

$$\begin{aligned}
&\sqrt{n}(\boldsymbol{c}'\widehat{\boldsymbol{\theta}}_{n,r}-1)\\
=\;&\left(\mu_r^{2/r}\underline{\alpha}_0\quad \frac{2}{r}\left\{-\underline{\alpha}_0\mu_r^{2/r}\overline{\boldsymbol{\theta}}_0'+\boldsymbol{c}'\boldsymbol{J}^{-1}(B^{(r)})^{-1}\right\}\right)\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\begin{pmatrix}\mu_r^{-2/r}(\eta_t^2-1)\\ B^{(r)}\boldsymbol{\phi}_t\left(\frac{|\eta_t|^r}{\mu_r}-1\right)\end{pmatrix}+o_P(1)\\
=\;&\frac{\underline{\alpha}_0}{\sqrt{n}}\sum_{t=1}^{n}(\eta_t^2-1)+\frac{2}{r}\left\{-\underline{\alpha}_0\overline{\boldsymbol{\theta}}_0'+\boldsymbol{c}'\boldsymbol{J}^{-1}\right\}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\boldsymbol{\phi}_t\left(\frac{|\eta_t|^r}{\mu_r}-1\right).
\end{aligned}$$

Let $P_{n,\boldsymbol{\tau}}$ the distribution of the observations $(\epsilon_1,\dots,\epsilon_n)$ when the parameter is $\theta_0+\boldsymbol{\tau}/\sqrt{n}$. Under $P_{n,0}$

$$\begin{pmatrix}T_{n,r}\\ \Lambda_{n,f}(\theta_0+\boldsymbol{\tau}/\sqrt{n},\theta_0)\end{pmatrix}\xrightarrow{d}\mathcal{N}\left\{\begin{pmatrix}0\\ -\frac{\iota_f}{8}\boldsymbol{\tau}'\boldsymbol{J}\boldsymbol{\tau}\end{pmatrix},\begin{pmatrix}1 & c_r\\ c_r & \frac{\iota_f}{4}\boldsymbol{\tau}'\boldsymbol{J}\boldsymbol{\tau}\end{pmatrix}\right\}, \tag{A.17}$$

where, using the equality $\boldsymbol{J}\overline{\boldsymbol{\theta}}_0 = \boldsymbol{\phi}$,

$$c_r = -\frac{1}{2\sigma^{(r)}}\left\{\underline{\alpha}_0\boldsymbol{\tau}'\boldsymbol{\phi}\left(k_2 - \frac{2}{r}k_r\right) + \boldsymbol{c}'\boldsymbol{\tau}\frac{2}{r}k_r\right\},$$

and $k_r = E\left\{\left(\frac{|\eta_1|^r}{\mu_r} - 1\right)g_1(\eta_1)\right\} = \frac{1}{\mu_r}E\left\{|\eta_1|^r g_1(\eta_1)\right\} = -r$ (the latter equality is straightforwardly obtained by integration by part). Therefore, $c_r = \frac{\boldsymbol{c}'\boldsymbol{\tau}}{\sigma^{(r)}}$.

Proposition 3.1 shows that

$$\lim_{n\to\infty} P_{n,\mathbf{0}}(\mathrm{C}_r) = \alpha.$$

Le Cam's third lemma (see e.g. van der Vaart, 1998, page 90) shows that

$$T_{n,r} \xrightarrow{d} \mathcal{N}(c_r, 1), \quad \text{under } P_{n,\boldsymbol{\tau}}.$$

The conclusion follows. $\square$

Proof of Corollary 3.1

Using the fact that $J^{-1} - \overline{\theta}_0\overline{\theta}_0'$ is semi-definite positive (see FLZ), minimizing $\sigma^{(r)}$ is equivalent to minimizing $g(r)$ with respect to r. $\square$

Proof of Proposition 3.3

By (A.17), we have

$$\Lambda_{n,f}(\theta_0 + \boldsymbol{\tau}/\sqrt{n}, \theta_0) \xrightarrow{d} \mathcal{N}\left(-\frac{\iota_f}{8}\boldsymbol{\tau}'\boldsymbol{J}\boldsymbol{\tau}, \frac{\iota_f}{4}\boldsymbol{\tau}'\boldsymbol{J}\boldsymbol{\tau}\right) \quad \text{under } P_{\theta_0},$$

which is the distribution of the log-likelihood ratio in the statistical model $\mathcal{N}\left\{\boldsymbol{\tau}, 4\boldsymbol{J}^{-1}/\iota_f\right\}$ of parameter $\boldsymbol{\tau}$. The so-called local experiments $\{L_{n,f}(\theta_0 + \boldsymbol{\tau}')/\sqrt{n}), \boldsymbol{\tau} \in \mathbb{R}^{p+q+1}\}$ converge to the gaussian experiment $\left\{\mathcal{N}\left(\boldsymbol{\tau}, 4\boldsymbol{J}^{-1}/\iota_f\right), \boldsymbol{\tau} \in \mathbb{R}^{p+q+1}\right\}$ (see van der Vaart (1998) for details about the notion of statistical experiments).

The second-order stationarity test in (3.5) corresponds to the test

$$\boldsymbol{H}_{0,\tau}: \quad \boldsymbol{c}'\boldsymbol{\tau} = 0 \quad \text{against} \quad \boldsymbol{H}_{1,\tau}: \boldsymbol{c}'\boldsymbol{\tau} > 0.$$

in the limiting experiment. The UMPU test based on $\boldsymbol{X} \sim \mathcal{N}\left(\boldsymbol{\tau}, 4\boldsymbol{J}^{-1}/\iota_f\right)$ is the test of rejection region

$$C = \left\{\boldsymbol{c}'\boldsymbol{X}/\sqrt{4\boldsymbol{c}'\boldsymbol{J}^{-1}\boldsymbol{c}/\iota_f} > \Phi^{-1}(1-\underline{\alpha})\right\}.$$

This UMPU test has the power

$$P_{\boldsymbol{H}_{1,\tau}}(C) = \Phi\left\{\frac{\boldsymbol{c}'\boldsymbol{\tau}\sqrt{\iota_f}}{2\sqrt{\boldsymbol{c}'\boldsymbol{J}^{-1}\boldsymbol{c}}} - \Phi^{-1}(1-\alpha)\right\}. \tag{A.18}$$

The conclusion follows. □

Proof of Proposition 3.4

In view of (3.9) and (3.10), the test (3.5) with $r = 2$ is asymptotically locally UMPU iff

$$c_2 = \frac{\boldsymbol{c}'\boldsymbol{\tau}}{\sqrt{(\mu_4 - 1)\boldsymbol{c}'\boldsymbol{J}^{-1}\boldsymbol{c}}} = \frac{\boldsymbol{c}'\boldsymbol{\tau}\sqrt{\iota_f}}{2\sqrt{\boldsymbol{c}'\boldsymbol{J}^{-1}\boldsymbol{c}}},$$

that is, iff $(\mu_4 - 1)\iota_f = 4$. By Corollary 1 in Francq and ZakoÃŕan (2006), the solutions of this equation are given by (3.11). □