



Munich Personal RePEc Archive

# **An Exponential Chi-Squared QMLE for Log-GARCH Models Via the ARMA Representation**

Christian Francq and Genaro Sucarrat

Universite de Lille 3, BI Norwegian Business School

24. October 2013

Online at <http://mpa.ub.uni-muenchen.de/51783/>

MPRA Paper No. 51783, posted 29. November 2013 05:15 UTC

# An Exponential Chi-Squared QMLE for Log-GARCH Models Via the ARMA Representation\*

Christian Francq<sup>†</sup> and Genaro Sucarrat<sup>‡</sup>

24 October 2013

## Abstract

Estimation of log-GARCH models via the ARMA representation is attractive because it enables a vast amount of already established results in the ARMA literature. We propose an exponential Chi-squared QMLE for log-GARCH models via the ARMA representation. The advantage of the estimator is that it corresponds to the theoretically and empirically important case where the conditional error of the log-GARCH model is normal. We prove the consistency and asymptotic normality of the estimator, and show that, asymptotically, it is as efficient as the standard QMLE in the log-GARCH(1,1) case. We also verify and study our results in finite samples by Monte Carlo simulations. An empirical application illustrates the versatility and usefulness of the estimator.

*JEL Classification:* C13, C22, C58

*Keywords:* Log-GARCH, EGARCH, Quasi Maximum Likelihood, Exponential Chi-Squared, ARMA

---

\*The authors gratefully acknowledge financial support of the Agence Nationale de la Recherche (ANR) through the ORA Program (Project PRAM ANR-10-ORAR-008-01) and of the Ecodec Labex.

<sup>†</sup>Department of Economics, Université de Lille 3. Email: christian.francq@univ-lille3.fr

<sup>‡</sup>Corresponding author. Department of Economics, BI Norwegian Business School, Nydalsveien 37, 0484 Oslo, Norway. Phone/Fax: +47+46 41 07 79 / +47+23 26 47 88. Email: genaro.sucarrat@bi.no. Webpage: <http://www.sucarrat.net/>

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Model, estimator and main assumptions</b>	<b>4</b>
2.1	The ARMA representation of the log-GARCH model . . . . .	5
2.2	Estimator based on the ARMA representation . . . . .	5
2.3	CAN of the estimator . . . . .	6
<b>3</b>	<b>Proofs of Theorems</b>	<b>8</b>
3.1	Proof of Theorem 2.1: Strong consistency . . . . .	8
3.2	Proof of Theorem 2.2: Asymptotic normality . . . . .	12
<b>4</b>	<b>Efficiency comparison</b>	<b>18</b>
4.1	Asymptotic comparison . . . . .	18
4.2	Finite sample comparison . . . . .	19
<b>5</b>	<b>Empirical application</b>	<b>20</b>
<b>6</b>	<b>Conclusions</b>	<b>21</b>
	<b>References</b>	<b>23</b>
<b>A</b>	<b>Asymptotic variance of the standard QMLE</b>	<b>23</b>
<b>B</b>	<b>Asymptotic variance of the Gaussian ARMA-QMLE</b>	<b>27</b>

# 1 Introduction

Autoregressive conditional heteroscedasticity (ARCH) models have successfully been applied in the modelling of a wide range of phenomena, including the uncertainty of inflation (e.g. [Engle \(1982\)](#)), the uncertainty of electricity prices (e.g. [Koopman et al. \(2007\)](#)), temperature variability (e.g. [Franses et al. \(2001\)](#)) and – most commonly – the variability of financial returns, see [Francq and Zakoïan \(2010\)](#) for a survey of ARCH models. Within the ARCH class of models exponential specifications are of special interest, since the positivity of volatility is guaranteed (this is not the case for ordinary or non-exponential ARCH models), and since they enable richer dynamics compared with ordinary specifications. In particular, in ordinary ARCH models the autocorrelations of squared errors are restricted to be positive. In exponential ARCH models, by contrast, negative autocorrelations are also allowed. Finally, exponential ARCH models can also be viewed as logarithmic versions of Multiplicative Error Models (MEM), that is, models of non-negative variables, see [Brownlees et al. \(2012\)](#) for a survey of MEM models.

The logarithmic ARCH (log-ARCH) model was independently proposed by [Pantula \(1986\)](#), [Geweke \(1986\)](#) and [Milhøj \(1987\)](#). [Engle and Bollerslev \(1986\)](#) argued against log-ARCH models because of the possibility of applying the log-operator (in the log-ARCH terms) on zero-values, which occurs whenever the error is equal to zero. A solution to this problem, however, is provided in [Sucarrat and Escribano \(2013\)](#) for the case where the zero-probability is zero (e.g. because zeros are due to discreteness or missing values). The solution relies on the Expectation-Maximisation (EM) algorithm combined with QML estimation via the ARMA representation. [Nelson \(1991\)](#) proposed an alternative exponential specification, known as the EGARCH model, and more recently [Creal et al. \(2013\)](#) and [Harvey \(2013\)](#) have proposed exponential specifications driven by the score of the log-likelihood.

Although the most common ARCH specifications were put forward already in the 1980s by [Engle \(1982\)](#) and [Bollerslev \(1986\)](#), the Consistency and Asymptotic Normality (CAN) of a Quasi Maximum Likelihood Estimator (QMLE) was not proved under mild conditions before the early 2000s by [Berkes et al. \(2003\)](#), and by [Francq and Zakoïan \(2004\)](#). For the EGARCH of [Nelson \(1991\)](#), a proof of CAN for a QMLE has turned out to be very difficult, see e.g. [Straumann and Mikosch \(2006\)](#). Indeed, currently there is only a single proof (that we know of) by [Wintenberger \(2012\)](#) under the somewhat complicated condition of continuous invertibility. And this is for the first-order version only. The score-driven exponential specifications of [Creal et al. \(2013\)](#) and [Harvey \(2013\)](#) do not admit QML estimation by their very nature. By contrast, the theoretical structure of log-GARCH models is much more tractable. The first proofs were independently provided by [Sucarrat et al. \(2013\)](#) for a Gaussian QMLE via the ARMA representation (henceforth Gaussian ARMA-QMLE),<sup>1</sup> and by [Francq et al. \(2013\)](#) for a Gaussian QMLE (henceforth standard QMLE) that does not make use of the ARMA representation. An advantage of the second estimator is that maximum efficiency is achieved in the theoretically and empirically important case where the standardised error is Gaussian. This is not the case

---

<sup>1</sup>In fact, the result of [Sucarrat et al. \(2013\)](#) applies to all ARMA estimators that satisfy a set of mild assumptions, and not only the Gaussian QMLE. For convenience, however, we choose the Gaussian QMLE as representative of these ARMA estimators.

for the first estimator. But an advantage of the first estimator is that a vast number of already established results from the ARMA literature is enabled. In particular, zeros on the errors and missing error-values can readily be handled satisfactorily by using the EM-algorithm on the ARMA-representation, see [Sucarrat and Escribano \(2013\)](#). This is not possible with the second estimator. The exponential Chi-squared QMLE that we propose combines the strengths of both of these estimators: (1) Zero error-values and missing values can be handled satisfactorily since estimation is via the ARMA representation, and (2) maximum efficiency is achieved when the standardised error is Gaussian, since the log of a squared Gaussian variate is distributed as a an exponential Chi-squared.

This paper makes two contributions. First, we prove the consistency and asymptotic normality of a centred exponential Chi-squared QMLE (henceforth Cex- $\chi^2$  ARMA-QMLE) for the log-GARCH model under mild conditions, and derive expressions for the asymptotic covariance matrix. Second, we study the asymptotic and finite sample efficiency of the first order log-GARCH specification. In particular, we show that our estimator has exactly the same asymptotic variances as that of the standard QMLE in the first order case, and that our estimator is generally more efficient than the Gaussian ARMA-QMLE, both asymptotically and in finite samples.

The paper is organised as follows. The next section, Section 2, contains our main results, i.e. the consistency and asymptotic normality theorems, whereas Section 3 contains their proofs. Section 4 compares the asymptotic and finite sample efficiency of our estimator with the Gaussian ARMA-QMLE and the standard QMLE. Section 5 illustrates our estimator in an empirical application. Finally Section 6 concludes. Tables and figures are placed at the end after the appendices.

## 2 Model, estimator and main assumptions

The log-GARCH( $p, q$ ) model is defined by

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

where  $(\eta_t)$  is a sequence of independent and identically distributed (iid) variables such that  $E\eta_1 = 0$ ,  $E\eta_1^2 = 1$  and  $P(\eta_t = 0) = 0$ . We will estimate separately the intercept  $\omega_0 \in \mathbb{R}$  and the parameter  $\theta_0 = (\alpha_{01}, \dots, \alpha_{0q}, \beta_{01}, \dots, \beta_{0p})$ , which belongs to a parameter space  $\Theta \subset \mathbb{R}^{p+q}$ . Interesting features of this log-GARCH specification are: 1) absence of positivity constraints on the parameters; 2) possibility of persistence of both high and low levels of volatility; 3) absence of a lower bound for the volatility; 4) stability by scaling of the observations and power transformation of the volatility. Point 4) means that if  $(\epsilon_t)$  satisfies the log-GARCH( $p, q$ ) model (2.1), then for any power  $\delta \neq 0$  and any  $a \neq 0$ , the process  $\epsilon_t^* = a\epsilon_t$  satisfies also an equation of the form (2.1) where, in the volatility equation, the powers  $x^2$  can be replaced by  $|x|^\delta$ .

The log-GARCH model (2.1) can be estimated in the usual way, by using a QMLE for scale-parameter (see [Francq et al. \(2013\)](#)). [Sucarrat et al. \(2013\)](#) propose a class of alternative estimators that exploit the existence of an ARMA representation with iid errors.

## 2.1 The ARMA representation of the log-GARCH model

The process  $(\ln \epsilon_t^2)$  satisfies an ARMA-type equation of the form

$$\mathcal{A}_{\theta_0}(L) \ln \epsilon_t^2 = \omega_0 + \mathcal{B}_{\theta_0}(L)v_t \quad (2.2)$$

where  $L$  denotes the lag operator,  $v_t = \ln \eta_t^2$ , and for all  $\theta = (\alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p) \in \Theta$ , the AR and MA polynomials are respectively defined by  $\mathcal{A}_{\theta}(z) = 1 - \sum_{i=1}^r (\alpha_i + \beta_i) z^i$ ,  $\mathcal{B}_{\theta}(z) = 1 - \sum_{i=1}^p \beta_i z^i$ ,  $r = \max\{p, q\}$ ,  $\alpha_i = 0$  for  $i > q$  and  $\beta_i = 0$  for  $i > p$ . Assuming  $E \ln^+ |\ln \eta_1^2| < \infty$ , it is well-known that both (2.2) and (2.1) admit a strictly stationary solution if the AR polynomial satisfies

$$\mathcal{A}_{\theta_0}(z) \neq 0 \text{ when } |z| \leq 1. \quad (2.3)$$

This condition is also necessary for the existence of a stationary and nonanticipative solution to (2.2) (and/or (2.1)), under the additional condition that  $P(\eta_1^2 = 1) \neq 1$  or  $\omega_0 \neq 0$  (otherwise there is the trivial solution  $\ln \epsilon_t^2 = 0$ , regardless of the value of  $\theta_0$ ). Under the moment condition  $E(\ln \eta_1^2)^2 < \infty$ , Equation (2.2) is a standard ARMA( $r, p$ ) equation of the form

$$\mathcal{A}_{\theta_0}(L) (\ln \epsilon_t^2 - \nu_0) = \mathcal{B}_{\theta_0}(L)u_t, \quad (2.4)$$

with  $\nu_0 = E \ln \epsilon_1^2$  and the white noise  $u_t = v_t - \mu_0$ , where  $\mu_0 = E \ln \eta_1^2$ . It is well known that the squares of a standard GARCH model also satisfy ARMA representations. But these ARMA representations are rarely used, in particular for the inference, because the innovations  $u_t = (\eta_t^2 - 1)\sigma_t^2$  are not independent in the standard GARCH case.

## 2.2 Estimator based on the ARMA representation

It is also well known that GARCH-type models can be consistently estimated by quasi-maximum likelihood estimation (QMLE) based on the instrumental  $\mathcal{N}(0, 1)$  density for  $\eta_t$  (see *e.g.* [Gourieroux et al. \(1984\)](#) for a general reference on QMLE, and [Berkes et al. \(2003\)](#) and [Francq and Zakoïan \(2004\)](#) for applications to GARCH models).<sup>2</sup> If  $\eta_t \sim \mathcal{N}(0, 1)$  then  $x_t = \ln \eta_t^2$  follows the exponential Chi-squared distribution with density  $\chi_0(x)$ , whereas the centred exponential Chi-squared distribution (Cex- $\chi^2$ ) is given by

$$\chi_{\mu}(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{x+\mu}{2} - \frac{e^{x+\mu}}{2}}.$$

One can wonder if a QMLE estimator based on the ARMA equation (2.4) would be consistent by taking the instrumental density  $\chi_{\mu}(x)$  for  $u_t$ . The answer is generally negative if one tries to estimate simultaneously  $\theta_0$ ,  $\nu_0$  and  $\mu_0$ . As we will show, however, the answer is positive if  $\nu_0$  is empirically estimated in a first step.

Under the invertibility condition

$$\forall \theta \in \Theta, \quad \mathcal{B}_{\theta}(z) \neq 0 \text{ when } |z| \leq 1, \quad (2.5)$$

<sup>2</sup>The term "quasi" is employed because the estimator is consistent even if the unknown density of  $\eta_t$  is not Gaussian.

the innovations of the de-meaned ARMA representation (2.4) are defined by

$$u_t(\theta) = \mathcal{B}_\theta^{-1}(L)\mathcal{A}_\theta(L) (\ln \epsilon_t^2 - \nu_0) := \sum_{i=0}^{\infty} \psi_i(\theta) (\ln \epsilon_{t-i}^2 - \nu_0).$$

These innovations can be approximated by

$$\tilde{u}_{t,n}(\theta) = \sum_{i=0}^{t-1} \psi_i(\theta) (\ln \epsilon_{t-i}^2 - \nu_n), \quad \nu_n = \frac{1}{n} \sum_{t=1}^n \ln \epsilon_t^2.$$

In practice, the  $\tilde{u}_{t,n}(\theta)$ 's can be obtained by taking the initial values  $\tilde{u}_{0,n}(\theta) = \dots = \tilde{u}_{1-q,n}(\theta) = 0$  and  $\ln \epsilon_0^2 = \dots = \ln \epsilon_{1-r}^2 = \nu_n$ , and by computing recursively

$$\tilde{u}_{t,n}(\theta) = \mathcal{A}_\theta(L) (\ln \epsilon_t^2 - \nu_n) + \sum_{i=1}^q \beta_i \tilde{u}_{t-i,n}(\theta), \quad t = 1, \dots, n. \quad (2.6)$$

Now consider the estimator defined by

$$\hat{\vartheta}_n = (\hat{\theta}_n, \hat{\mu}_n) = \arg \max_{\vartheta \in \Xi} \tilde{Q}_n(\vartheta), \quad \tilde{Q}_n(\vartheta) = \frac{1}{n} \sum_{t=1}^n \tilde{\ell}_{t,n}(\vartheta) \quad (2.7)$$

where  $\vartheta = (\theta, \mu)$ ,  $\Xi$  is a compact set of the form  $\Theta \times [a, b]$ , and

$$\tilde{\ell}_{t,n}(\vartheta) = \ln \chi_\mu \{ \tilde{u}_{t,n}(\theta) \} + \frac{1}{2} \ln(2\pi) = \frac{1}{2} (\tilde{u}_{t,n}(\theta) + \mu - e^{\tilde{u}_{t,n}(\theta) + \mu}).$$

The intercept can then be estimated by

$$\hat{\omega}_n = \mathcal{A}_{\hat{\theta}_n}(1)\nu_n - \mathcal{B}_{\hat{\theta}_n}(1)\hat{\mu}_n. \quad (2.8)$$

Note that, for the log-GARCH model (2.1), the estimator of the parameters of interest is  $(\hat{\omega}_n, \hat{\theta}_n)$ . This is a multi-step estimator, which is in the spirit of the variance targeting estimator (see *e.g.* Francq et al. (2011)). Indeed, it involves the estimation of a parameter by an empirical mean in a first step and a QML estimation of the remaining parameters in a second step,

### 2.3 CAN of the estimator

**Theorem 2.1 (Strong consistency)** *Let  $\hat{\vartheta}_n$  and  $\hat{\omega}_n$  be sequences of estimators satisfying (2.7) and (2.8), where the  $\epsilon_t$ 's follow the log-GARCH model (2.1). Assume that  $\theta_0 \in \Theta$  and  $\Theta$  is compact, that the stationary condition (2.3) and the invertibility condition (2.5) hold, that the distribution of  $\ln \eta_1^2$  is not degenerate with  $E|\ln \eta_1^2| < \infty$ , that  $\mathcal{A}_{\theta_0}(z)$  and  $\mathcal{B}_{\theta_0}(z)$  have no common roots, that  $p + q > 0$  with  $\alpha_{0r} + \beta_{0r} \neq 0$  or  $\beta_{0p} \neq 0$  (with the convention  $\alpha_{00} = \beta_{00} = 1$ ).*

*Then, almost surely  $\hat{\vartheta}_n \rightarrow \vartheta_0 = (\theta_0, \mu_0)$  and  $\hat{\omega}_n \rightarrow \omega_0$  as  $n \rightarrow \infty$ .*

**Proof:** See Section 3.1.

Francq et al. (2013) showed the consistency of the standard QMLE under the same assumptions Theorem 2.1. The consistency of the Gaussian ARMA-QMLE is obtained by Sucarrat et al. (2013) under the same assumptions as, plus the very mild additional assumption that  $E|\eta_t|^{2+\varepsilon} < \infty$  for some  $\varepsilon > 0$  (see their Assumption A4 a)).

Let us turn to the asymptotic distribution. We need a condition which ensures the existence of a moment of order  $r > 0$  for  $e^{u_t(\theta)}$  in a neighborhood of  $\theta_0$ .<sup>3</sup> First define the sequence  $\{\pi_i(\theta)\}_i$  by

$$u_t(\theta) = \mathcal{B}_\theta^{-1}(L)\mathcal{A}_\theta(L)\mathcal{A}_{\theta_0}^{-1}(L)\mathcal{B}_{\theta_0}(L)u_t := \sum_{i=0}^{\infty} \pi_i(\theta)u_{t-i}.$$

Assume that there exists a compact set  $\bar{V}(\theta_0)$  which contains a neighborhood  $V(\theta_0)$  of  $\theta_0$ , and which is such that

$$\prod_{i=0}^{\infty} E|\eta_t|^{2r\pi_i(\theta)} < \infty, \quad \forall \theta \in \bar{V}(\theta_0). \quad (2.9)$$

With some abuse of notation, we set  $\bar{V}(\vartheta_0) = \bar{V}(\theta_0) \times [a, b]$ .

**Remark 2.1** Note that  $\prod_{i=0}^{\infty} E|\eta_t|^{2r\pi_i(\theta)} = E|\eta_t|^{2r}$  at  $\theta = \theta_0$ . Thus, by a continuity argument, one can expect (2.9) be satisfied when  $E|\eta_1|^{2r+\nu} < \infty$  for some  $\nu > 0$ . This is however a result that we did not succeed to show.

It is possible to check (2.9) for specific distributions of  $\eta_1$ .

**Lemma 2.1** If  $\eta_1 \sim \mathcal{N}(0, 1)$  then (2.9) holds true.

**Proof of Lemma 2.1.** We show that  $\sum_{i=0}^{\infty} \ln E|\eta_t|^{2r\pi_i(\theta)} < \infty$ , by noting that, as  $s \rightarrow 0$ ,

$$\ln E|\eta_1|^s = \ln \left\{ \frac{2^{s/2}}{\sqrt{\pi}} \Gamma\left(\frac{1+s}{2}\right) \right\} = \frac{s}{2} \ln 2 + \frac{s}{2} \{-\gamma_0 + o(s)\}$$

where  $\gamma$  is the Euler constant. □

Note that (2.9) entails  $Ee^{ru_t(\theta)} < \infty$ . Because we need the existence of this moment for a uniform bound of  $e^{ru_t(\theta)}$ , we slightly reinforce (2.9) by assuming that

$$\prod_{i=0}^{\infty} E \sup_{\theta \in \bar{V}(\theta_0)} |\eta_t|^{2r\pi_i(\theta)} < \infty. \quad (2.10)$$

**Remark 2.2** Under the previous assumption, we have

$$E \sup_{\theta \in \bar{V}(\theta_0)} |e^{u_t(\theta)}|^r = E \sup_{\theta \in \bar{V}(\theta_0)} \prod_{i=0}^{\infty} |\eta_{t-i}|^{2r\pi_i(\theta)} < \infty.$$

<sup>3</sup>At  $\theta_0$  the existence of the moment of order  $r$  is ensured when  $E|\eta_1|^{2r} < \infty$ .



**Theorem 2.2 (Asymptotic normality)** *Let the assumptions of Theorem 2.1, and the additional assumptions that (2.10) holds true for some  $r > 2$ ,  $\vartheta_0$  belongs to the interior of  $\Xi$ ,  $E\eta_1^4 < \infty$ , and  $E(\ln \eta_1^2)^2 < \infty$ .*

*Then, as  $n \rightarrow \infty$ ,*

$$\sqrt{n} \begin{pmatrix} \hat{\omega}_n - \omega_0 \\ \hat{\theta}_n - \theta_0 \end{pmatrix} \xrightarrow{d} \mathcal{N} \left\{ 0, (E\eta_1^4 - 1) \begin{pmatrix} \mathcal{B}_{\theta_0}^2(1) + \gamma' \Sigma_u^{-1} \gamma & \gamma' \Sigma_u^{-1} \\ \Sigma_u^{-1} \gamma & \Sigma_u^{-1} \end{pmatrix} \right\}.$$

where  $\gamma = (-\nu_0 1'_q, (\mu_0 - \nu_0) 1'_p)'$ ,  $1_s$  denoting the all-ones vector of dimension  $s$ , and

$$\Sigma_u = E \frac{\partial u_t}{\partial \theta} \frac{\partial u_t}{\partial \theta'} (\theta_0).$$

**Proof:** See Section 3.2.

To show the asymptotic normality of the standard QMLE Francq et al. (2013) do not need (2.10), but they assume the existence of an exponential moment for  $|\log \eta_1^2|$ . For the asymptotic normality of Gaussian QMLE, Sucarrat et al. (2013) also avoid (2.10), but they have to assume  $E|\eta_1|^{4+\varepsilon} < \infty$  for some  $\varepsilon > 0$ .

Even if the CAN of the standard QMLE, Gaussian ARMA-QMLE and Cex- $\chi^2$  ARMA-QMLE are obtained under similar assumptions, for the three estimators the techniques of proof are quite different and, as we will see in Section 4.1 below, the asymptotic variances may be different. More importantly, Section 4 show that the three estimators are not equivalent in practice. In particular, they generally have different accuracies in finite samples, and the estimators based on the ARMA representations are simpler because they inherit the usual techniques developed for linear time series analysis.

## 3 Proofs of Theorems

### 3.1 Proof of Theorem 2.1: Strong consistency

Let the random variables

$$O_n(\vartheta) = \frac{1}{n} \sum_{t=1}^n \ell_t(\vartheta), \quad Q_n(\vartheta) = \frac{1}{n} \sum_{t=1}^n \ell_{t,n}(\vartheta),$$

where, up to the unimportant constant  $\ln(2\pi)/2$ ,  $\ell_t(\vartheta) = \ln \chi_\mu \{u_t(\theta)\}$ ,  $\ell_{t,n}(\vartheta) = \ln \chi_\mu \{u_{t,n}(\theta)\}$  and  $u_{t,n}(\theta) = \sum_{i=0}^{\infty} \psi_i(\theta) (\ln \epsilon_{t-i}^2 - \nu_n)$ . We also need to introduce the set

$$\Lambda = \{ \vartheta \in \Xi : Ee^{u_t(\theta)} < \infty \}.$$

Note that  $\vartheta_0 \in \Lambda$  because  $Ee^{u_t(\theta_0)} = Ee^{u_t} = Ee^{\ln \eta_t^2 - \mu_0} = e^{-\mu_0}$ . The relative complement of  $\Lambda$  in  $\Xi$  is denoted by  $\Lambda^c$ . The proof of the consistency of  $\hat{\vartheta}_n$  is split into the following

steps.

- i)* for any compact subset  $\Lambda_0$  of  $\Lambda$ 

$$\lim_{n \rightarrow \infty} \sup_{\vartheta \in \Lambda_0} |O_n(\vartheta) - Q_n(\vartheta)| = 0 \text{ a.s.};$$
- ii)*  $\lim_{n \rightarrow \infty} \sup_{\vartheta \in \Xi} |Q_n(\vartheta) - \tilde{Q}_n(\vartheta)| = 0 \text{ a.s.};$
- iii)* if  $\vartheta \in \Lambda^c$  then  $\tilde{Q}_n(\vartheta) \rightarrow -\infty \text{ a.s.};$
- iv)* if  $u_1(\theta) + \mu = u_1(\theta_0) + \mu_0 \text{ a.s.}$  then  $\theta = \theta_0$  and  $\mu = \mu_0$ ;
- v)* if  $\vartheta \neq \vartheta_0$ ,  $E\ell_1(\vartheta) < E\ell_1(\vartheta_0)$ ;
- vi)* any  $\vartheta \neq \vartheta_0$  has a neighborhood  $V(\vartheta)$  such that
$$\limsup_{n \rightarrow \infty} \sup_{\vartheta^* \in V(\vartheta)} \tilde{Q}_n(\vartheta^*) < E\ell_1(\vartheta_0) = \lim_{n \rightarrow \infty} \tilde{Q}_n(\vartheta_0) \text{ a.s.}$$

In the sequel of the paper, the letters  $K \geq 0$  and  $\rho \in [0, 1)$  denote generic constants, or measurable functions of  $\{\epsilon_u, u \leq 0\}$ , that do not vary with  $n$ .

We first show *i)*. Note that (2.3) and  $E|\ln \eta_1^2| < \infty$  entail  $E|\ln \epsilon_1^2| < \infty$ , so that  $\nu_0$  is well defined. By the invertibility condition (2.5), we immediately have

$$u_t(\theta) - u_{t,n}(\theta) = K_\theta(\nu_n - \nu_0),$$

where  $K_\theta = \mathcal{B}_\theta^{-1}(1)\mathcal{A}_\theta(1) = \sum_{i=0}^{\infty} \psi_i(\theta)$ . The compactness of  $\Theta$  then entails that

$$\sup_{\theta \in \Theta} |u_t(\theta) - u_{t,n}(\theta)| \leq K|\nu_n - \nu_0|, \quad \text{a.s.} \quad (3.1)$$

We now need to show the same bound for  $|e^{u_t(\theta)} - e^{u_{t,n}(\theta)}|$  when  $\vartheta \in \Lambda$ . We have

$$e^{u_t(\theta)} - e^{u_{t,n}(\theta)} = X_t(\theta) \{e^{-\nu_0 K_\theta} - e^{-\nu_n K_\theta}\},$$

where  $X_t(\theta) = e^{\sum_{i=0}^{\infty} \psi_i(\theta) \ln \epsilon_{t-i}^2}$ . By a Taylor expansion, we then obtain

$$|e^{u_t(\theta)} - e^{u_{t,n}(\theta)}| = K_\theta |\nu_n - \nu_0| e^{-\nu^* K_\theta} X_t(\theta),$$

where  $\nu^*$  stands between  $\nu_0$  and  $\nu_n$ . By the ergodic theorem, we have strong convergence of  $\nu_n$  to  $\nu_0$  as  $n \rightarrow \infty$ . We thus obtain

$$|e^{u_t(\theta)} - e^{u_{t,n}(\theta)}| \leq K |\nu_n - \nu_0| X_t(\theta) \quad (3.2)$$

where  $E \sup_{\vartheta \in \Lambda_0} X_t(\theta) < \infty$ . Since

$$|\ell_t(\vartheta) - \ell_{t,n}(\vartheta)| \leq K (|u_t(\theta) - u_{t,n}(\theta)| + |e^{u_t(\theta)} - e^{u_{t,n}(\theta)}|),$$

we obtain *i)* from (3.1) and (3.2), together with the ergodic theorem.

We now show *ii)*. The compactness of  $\Theta$  and the invertibility condition (2.5) entail

that

$$\mathcal{B}_\theta^{-1}(L) = \sum_{i=0}^{\infty} \varphi_i(\theta) L^i \quad \text{where} \quad \sup_{\theta \in \Theta} |\varphi_i(\theta)| \leq K \rho^i. \quad (3.3)$$

Note that (2.6) still holds true for any  $t$  when  $\tilde{u}_{t,n}(\theta)$  is replaced by  $u_{t,n}(\theta)$ . For all  $t > r$  we thus have

$$\mathcal{B}_\theta(L) \{u_{t,n}(\theta) - \tilde{u}_{t,n}(\theta)\} = 0.$$

Iterating this relation, we obtain in the log-GARCH(1,  $q$ ) case

$$u_{t,n}(\theta) - \tilde{u}_{t,n}(\theta) = \beta^{t-r} \{u_{r,n}(\theta) - \tilde{u}_{r,n}(\theta)\},$$

with the simplified notation  $\beta = \beta_1$ . For the general log-GARCH( $p$ ,  $q$ ) model, by (3.3) we also obtain

$$\sup_{\theta \in \Theta} |u_{t,n}(\theta) - \tilde{u}_{t,n}(\theta)| \leq K \rho^t, \quad \text{a.s.} \quad (3.4)$$

We now study the difference  $e^{u_{t,n}(\theta)} - e^{\tilde{u}_{t,n}(\theta)}$ . For simplicity, we focus on the log-GARCH(1,1) case, but the same arguments apply to the general model (however, with more complex notations). We then have, for  $t \geq 2$ ,

$$u_{t,n}(\theta) = d_{t,n}(\theta) + \beta^{t-1} u_{1,n}(\theta), \quad d_{t,n}(\theta) = \sum_{i=0}^{t-2} \beta^i \mathcal{A}_\theta(L) \{\ln \epsilon_{t-i}^2 - \nu_n\}.$$

The same expression holds true for  $\tilde{u}_{t,n}(\theta)$  when  $u_{1,n}(\theta)$  is replaced by  $\tilde{u}_{1,n}(\theta)$ . Doing a Taylor expansion, it follows that

$$\begin{aligned} e^{u_{t,n}(\theta)} - e^{\tilde{u}_{t,n}(\theta)} &= \left\{ e^{\beta^{t-1} u_{1,n}(\theta)} - e^{\beta^{t-1} \tilde{u}_{1,n}(\theta)} \right\} e^{d_{t,n}(\theta)} \\ &= \beta^{t-1} \{u_{1,n}(\theta) - \tilde{u}_{1,n}(\theta)\} e^{\beta^{t-1} u^*} e^{d_{t,n}(\theta)} \end{aligned}$$

where  $u^*$  is between  $u_{1,n}(\theta)$  and  $\tilde{u}_{1,n}(\theta)$ . It follows that

$$\frac{1}{t} \ln |e^{u_{t,n}(\theta)} - e^{\tilde{u}_{t,n}(\theta)}| \leq \frac{K}{t} + \frac{d_{t,n}(\theta)}{t} + \ln |\beta|.$$

Because  $E |d_{t,n}(\theta)| < \infty$  uniformly in  $t$ , the second term of the right-hand side of the inequality tends almost surely to 0 (see Lemma 7.1 in Francq et al. (2013)). Since  $|\beta| < 1$  on the compact  $\Theta$ ,

$$\limsup_{n \rightarrow \infty} \sup_{\theta \in \Theta} \frac{1}{t} \ln |e^{u_{t,n}(\theta)} - e^{\tilde{u}_{t,n}(\theta)}| \in [-\infty, 0).$$

We thus obtain

$$\sup_{\theta \in \Theta} |e^{u_{t,n}(\theta)} - e^{\tilde{u}_{t,n}(\theta)}| \leq K \rho^t \quad \text{a.s.} \quad (3.5)$$

and the conclusion follows from (3.4) and (3.5), with the arguments used to show *i*).

To show *iii*), first note that

$$E \max \{\ell_t(\vartheta), 0\} \leq \frac{1}{2} E(|u_t(\theta)| + |\mu|) \leq K \left( 1 + \sum_{i=0}^{\infty} |\psi_i(\theta)| E |\ln \epsilon_1^2| \right) < \infty.$$

Therefore  $E\ell_t(\vartheta)$  is well defined in  $\mathbb{R} \cup \{-\infty\}$ , for any  $\vartheta \in \Xi$ . Now assume that  $\vartheta \in \Xi$  is such that  $Ee^{u_t(\theta)} = +\infty$ . We then have

$$E\ell_t(\vartheta) = \frac{1}{2} E \{u_t(\theta) + \mu - e^{u_t(\theta)+\mu}\} = -\infty.$$

Applying the ergodic theorem for stationary and ergodic processes having an expectation in  $[-\infty, 0)$  (see Billingsley (1995, pp. 284 and 495)), we have  $O_n(\vartheta) \rightarrow -\infty$  a.s. when  $n \rightarrow \infty$ . By the same arguments, it can be shown that  $Q_n(\vartheta) \rightarrow -\infty$ . In view of *ii*), we also have  $\tilde{Q}_n(\vartheta) \rightarrow -\infty$ .

Let us turn to *iv*). If  $u_t(\theta) + \mu = u_t(\theta_0) + \mu_0$  then

$$\sum_{i=0}^{\infty} \{\psi_i(\theta) - \psi_i(\theta_0)\} (\ln \epsilon_{t-i}^2 - \nu_0) = \mu_0 - \mu.$$

Since the left-hand side of the equality is a centered random variable, we must have  $\mu = \mu_0$ . If  $\psi_1(\theta) - \psi_1(\theta_0) \neq 0$  then  $\ln \epsilon_{t-1}^2$  is a linear combination of its past values. This is impossible because the linear innovations  $u_t$  are not a.s. equal to zero. By recursion,  $\psi_i(\theta) - \psi_i(\theta_0) = 0$  for all  $i$ , which entails  $\theta = \theta_0$  under the conditions on the AR and MA polynomials.

To show *v*), note that  $E\ell_1(\vartheta) = -\infty$  when  $\vartheta \in \Lambda^c$ . We can therefore assume that  $\vartheta \in \Lambda$ . Then we have

$$\begin{aligned} 2 \{E\ell_1(\vartheta) - E\ell_1(\vartheta_0)\} &= E \{u_1(\theta) + \mu - u_1(\theta_0) - \mu_0 + 1 - e^{u_1(\theta)+\mu}\} \\ &= 1 + \mu - \mu_0 - Ee^{u_1(\theta)-u_1(\theta_0)+\mu-\mu_0+u_1(\theta_0)+\mu_0} \\ &= 1 + \mu - \mu_0 - Ee^{u_1(\theta)-u_1(\theta_0)+\mu-\mu_0} \\ &\leq 0, \end{aligned}$$

with equality iff  $u_1(\theta) + \mu = u_1(\theta_0) + \mu_0$  with probability one. For the first equality, we used the fact that

$$e^{u_1(\theta_0)+\mu_0} = e^{u_1+\mu_0} = \eta_1^2, \quad (3.6)$$

for the second equality we note that  $Eu_1(\theta) = 0$  for all  $\theta$ , for the third equality we use the independence between  $u_1(\theta) - u_1(\theta_0)$  and  $u_1(\theta_0) = \ln \eta_1^2 - \mu_0$ , and for the inequality we argue that  $e^x \geq x + 1$  with equality iff  $x = 0$ . The conclusion comes from *iii*).

It remains to show *vi*). Let  $V_k(\vartheta)$  be the open ball with center  $\vartheta$  and radius  $1/k$ . In view of *iii*), and by continuity of  $\vartheta \rightarrow E\ell_1(\vartheta)$ , there exists a compact set  $\Lambda_0 \subset \Lambda$  such that if  $V_k(\vartheta) \subset \Lambda_0^c$  then

$$\limsup_{n \rightarrow \infty} \sup_{\vartheta^* \in V_k(\vartheta) \cap \Xi} \tilde{Q}_n(\vartheta^*) < E\ell_1(\vartheta_0).$$

We can therefore assume that, for all  $k$ ,  $V_k(\vartheta) \cap \Lambda_0 \neq \emptyset$  (i.e.  $\vartheta \in \Lambda_0$ ). Take another compact set  $\Lambda_1$  such that  $\Lambda_0 \subset \Lambda_1 \subset \Lambda$ . Assume  $k$  large enough so that  $V_k(\vartheta) \subset \Lambda_1$ . Using successively *i*)-*ii*), the ergodic process, the monotone convergence theorem and *v*), we then obtain almost surely

$$\begin{aligned}
& \limsup_{n \rightarrow \infty} \sup_{\vartheta^* \in V_k(\vartheta) \cap \Xi} \tilde{Q}_n(\vartheta^*) \\
& \leq \limsup_{n \rightarrow \infty} \sup_{\vartheta^* \in V_k(\vartheta) \cap \Xi} O_n(\vartheta^*) + \limsup_{n \rightarrow \infty} \sup_{\vartheta \in \Lambda_1} |\tilde{Q}_n(\vartheta) - O_n(\vartheta)| \\
& \leq \limsup_{n \rightarrow \infty} n^{-1} \sum_{t=1}^n \sup_{\vartheta^* \in V_k(\vartheta) \cap \Xi} \ell_t(\vartheta^*) \\
& = E \sup_{\vartheta^* \in V_k(\vartheta) \cap \Xi} \ell_1(\vartheta^*) \\
& < E \ell_1(\vartheta_0)
\end{aligned}$$

for  $k$  large enough, when  $\vartheta \neq \vartheta_0$ , which completes the proof of *vi*). The proof of the consistency  $\hat{\vartheta}_n$  then follows from a standard compactness argument, as in Wald (1949).

Taking expectation in both sides of (2.2), we obtain

$$\mathcal{A}_{\theta_0}(1)\nu_0 = \omega_0 + \mathcal{B}_{\theta_0}(1)\mu_0.$$

The consistency of  $\hat{\omega}_n$  follows from that of  $\hat{\vartheta}_n$  and  $\nu_n$ . □

### 3.2 Proof of Theorem 2.2: Asymptotic normality

We decompose the proof into several steps.

**a) Asymptotic distribution of  $\nu_n$ .** Note that our estimation procedure shares some similarities with the variance targeting estimator (see Francq et al. (2011)) because they are both two-step estimators, involving an empirical moment estimation in the first step. Taking the average of both sides of (2.4), for  $t$  varying from 1 to  $n$ , we obtain

$$\mathcal{A}_{\theta_0}(1)(\nu_n - \nu_0) = \mathcal{B}_{\theta_0}(1) \frac{1}{n} \sum_{t=1}^n u_t + O_P\left(\frac{1}{n}\right),$$

which is the analog of equation (A.15) in Francq et al. (2011). The central limit theorem then entails

$$\begin{aligned}
\sqrt{n}(\nu_n - \nu_0) &= \frac{\mathcal{B}_{\theta_0}(1)}{\mathcal{A}_{\theta_0}(1)} \frac{1}{\sqrt{n}} \sum_{t=1}^n u_t + o_P(1) \\
&\xrightarrow{d} \mathcal{N}(0, \sigma_\nu^2), \quad \sigma_\nu^2 = \frac{\mathcal{B}_{\theta_0}^2(1)}{\mathcal{A}_{\theta_0}^2(1)} \text{Var} \ln \eta_1^2.
\end{aligned}$$

**b) Negligible impact of the initial values.** Similarly to *ii)* in the proof of Theorem 2.1, we now show that

$$\lim_{n \rightarrow \infty} \sqrt{n} \sup_{\vartheta \in \Xi \cap \bar{V}(\vartheta_0)} \left| \frac{\partial Q_n(\vartheta)}{\partial \vartheta} - \frac{\partial \tilde{Q}_n(\vartheta)}{\partial \vartheta} \right| = 0 \quad \text{a.s.} \quad (3.7)$$

The derivatives of  $u_{t,n}(\theta)$  with respect to  $\theta$  have ARMA representations similar to that of  $u_{t,n}(\theta)$ . By the arguments used to show (3.4), we thus obtain

$$\sup_{\theta \in \Theta} \left\| \frac{\partial u_{t,n}(\theta)}{\partial \theta} - \frac{\partial \tilde{u}_{t,n}(\theta)}{\partial \theta} \right\| \leq K \rho^t, \quad \text{a.s.} \quad (3.8)$$

Using this inequality, (3.5) and

$$\frac{\partial e^{u_{t,n}(\theta)+\mu}}{\partial \vartheta'} = e^{u_{t,n}(\theta)+\mu} \left( \frac{\partial u_{t,n}(\theta)}{\partial \vartheta'}, 1 \right),$$

we obtain

$$\sup_{\vartheta \in \Xi \cap \bar{V}(\vartheta_0)} \left\| \frac{\partial e^{u_{t,n}(\theta)}}{\partial \vartheta} - \frac{\partial e^{\tilde{u}_{t,n}(\theta)}}{\partial \vartheta} \right\| \leq K \rho^t \sup_{\vartheta \in \bar{V}(\vartheta_0)} \left( e^{u_{t,n}(\theta)} + \left\| \frac{\partial u_{t,n}(\theta)}{\partial \theta} \right\| \right).$$

In view of Remark 2.2, it follows that

$$E \sup_{\vartheta \in \Xi \cap \bar{V}(\vartheta_0)} \left\| \frac{\partial e^{u_{t,n}(\theta)}}{\partial \vartheta} - \frac{\partial e^{\tilde{u}_{t,n}(\theta)}}{\partial \vartheta} \right\| \leq K \rho^t. \quad (3.9)$$

We easily obtain (3.7) from (3.8) and (3.9).

**c) A Taylor expansion for the derivative of the criterion.** Note that  $Q_n(\vartheta)$  and  $O_n(\vartheta)$  are values of the same function at the points  $(\vartheta, \nu_n)$  and  $(\vartheta, \nu_0)$ , respectively. More precisely, with some abuse of notation, we have

$$u_{t,n}(\theta) = u_t(\theta, \nu_n) \quad \text{and} \quad u_t(\theta) = u_t(\theta, \nu_0), \quad (3.10)$$

$$\ell_{t,n}(\hat{\vartheta}_n) = \ell_t(\hat{\vartheta}_n, \nu_n) \quad \text{and} \quad \ell_t(\vartheta_0) = \ell_t(\vartheta_0, \nu_0). \quad (3.11)$$

Using (3.7), the notation  $a \stackrel{c}{=} b$  when  $a = b + c$ , the consistency of  $\hat{\vartheta}_n$ , and Taylor expansions, we then obtain, for  $n$  large enough,

$$\begin{aligned} 0_d &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial \tilde{\ell}_{t,n}(\hat{\vartheta}_n)}{\partial \vartheta} \stackrel{o_P(1)}{=} \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial \ell_{t,n}(\hat{\vartheta}_n)}{\partial \vartheta} \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial \ell_t(\vartheta_0)}{\partial \vartheta} + J_n \sqrt{n} \left( \hat{\vartheta}_n - \vartheta_0 \right) + K_n \sqrt{n} (\nu_n - \nu_0), \end{aligned} \quad (3.12)$$

where  $d = p + q + 1$ , and the elements of the  $d \times d$  matrix  $J_n$  and the  $d \times 1$  vector  $K_n$  are defined by

$$J_n(i, j) = \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 \ell_t(\vartheta_i^*, \nu_i^*)}{\partial \vartheta_i \partial \vartheta_j} \quad \text{and} \quad K_n(i) = \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 \ell_t(\vartheta_i^*, \nu_i^*)}{\partial \vartheta_i \partial \nu},$$

for some  $\vartheta_i^*$  between  $\hat{\vartheta}_n$  and  $\vartheta_0$  and some  $\nu_i^*$  between  $\nu_n$  and  $\nu_0$ .

**d) A CLT for stationary martingale increments.** Noting that

$$\frac{\partial \ln \chi_\mu(x)}{\partial x} = \frac{\partial \ln \chi_\mu(x)}{\partial \mu} = \frac{1}{2} (1 - e^{x+\mu})$$

and using (3.6), we obtain

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial \ell_t(\vartheta_0)}{\partial \vartheta} = \frac{1}{2\sqrt{n}} \sum_{t=1}^n (1 - \eta_t^2) \begin{pmatrix} \frac{\partial u_t(\theta_0)}{\partial \theta} \\ 1 \end{pmatrix}.$$

Because  $E(\ln \eta_t^2)^2 < \infty$  and  $\partial u_t(\theta_0)/\partial \theta_i$  is an absolutely summable linear combination of the  $\ln \eta_u^2$ 's, for  $u \leq t - 1$ , we have  $E\|\partial u_t(\theta_0)/\partial \theta\|^2 < \infty$ . The central limit theorem for martingale difference thus implies

$$\begin{aligned} \begin{pmatrix} \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial \ell_t(\vartheta_0)}{\partial \vartheta} \\ \sqrt{n}(\nu_n - \nu_0) \end{pmatrix} &\stackrel{o_P(1)}{=} \begin{pmatrix} \frac{1}{2\sqrt{n}} \sum_{t=1}^n (1 - \eta_t^2) \begin{pmatrix} \frac{\partial u_t(\theta_0)}{\partial \theta} \\ 1 \end{pmatrix} \\ \frac{\mathcal{B}_{\theta_0(1)}}{\mathcal{A}_{\theta_0(1)}} \frac{1}{\sqrt{n}} \sum_{t=1}^n u_t \end{pmatrix} \\ &\xrightarrow{d} \mathcal{N} \left\{ 0, \begin{pmatrix} \tau_\eta E \frac{\partial u_t}{\partial \theta} \frac{\partial u_t}{\partial \theta'}(\theta_0) & 0_{d-1} & 0_{d-1} \\ 0'_{d-1} & \tau_\eta & \xi \\ 0'_{d-1} & \xi & \sigma_\nu^2 \end{pmatrix} \right\} \end{aligned}$$

where  $\tau_\eta = \frac{E\eta_1^4 - 1}{4}$  and  $\xi = \frac{\mathcal{B}_{\theta_0(1)}}{2\mathcal{A}_{\theta_0(1)}} E(1 - \eta_1^2) \ln \eta_1^2$ .

**e) Existence and invertibility of some information matrices.** We have

$$\frac{\partial^2 \ell_t(\vartheta_0)}{\partial \vartheta \partial \vartheta'} = \frac{1 - \eta_t^2}{2} \begin{pmatrix} \frac{\partial^2 u_t(\theta_0)}{\partial \theta \partial \theta'} & 0_{d-1} \\ 0'_{d-1} & 0 \end{pmatrix} - \frac{\eta_t^2}{2} \begin{pmatrix} \frac{\partial u_t(\theta_0)}{\partial \theta} \frac{\partial u_t(\theta_0)}{\partial \theta'} & \frac{\partial u_t(\theta_0)}{\partial \theta} \\ \frac{\partial u_t(\theta_0)}{\partial \theta'} & 1 \end{pmatrix}.$$

Noting that  $\partial u_t(\theta_0)/\partial \theta$  and  $\eta_t$  are independent, that  $E\|\partial^2 u_t(\theta_0)/\partial \theta \partial \theta'\| < \infty$  and  $E\|\partial u_t(\theta_0)/\partial \theta\|^2 < \infty$ , one can set

$$J := E \frac{\partial^2 \ell_t(\vartheta_0)}{\partial \vartheta \partial \vartheta'} = -\frac{1}{2} \begin{pmatrix} E \frac{\partial u_t(\theta_0)}{\partial \theta} \frac{\partial u_t(\theta_0)}{\partial \theta'} & 0_{d-1} \\ 0'_{d-1} & 1 \end{pmatrix}.$$

Arguing by contradiction, we assume that  $J$  is singular. Then there exists  $\lambda \in \mathbb{R}^{p+q}$

such that  $\lambda' \frac{\partial u_t(\theta_0)}{\partial \theta} = 0$  a.s. Taking the derivative of both sides of (2.4), we obtain

$$\mathcal{B}_{\theta_0}(L) \frac{\partial u_t(\theta_0)}{\partial \theta} - \begin{pmatrix} 0 \\ \vdots \\ 0 \\ u_{t-1}(\theta_0) \\ \vdots \\ u_{t-p}(\theta_0) \end{pmatrix} = - \begin{pmatrix} \ln \epsilon_{t-1}^2 - \nu_0 \\ \vdots \\ \ln \epsilon_{t-q}^2 - \nu_0 \\ \ln \epsilon_{t-1}^2 - \nu_0 \\ \vdots \\ \ln \epsilon_{t-p}^2 - \nu_0 \end{pmatrix}.$$

Multiplying the two sides of this equation by  $\lambda' = (\lambda_1, \dots, \lambda_{p+q})$ , it can be seen that  $\lambda_1 = 0$  (otherwise the linear innovation of  $(\ln \epsilon_t^2)$  would be degenerated) and

$$\sum_{i=2}^q \lambda_i (\ln \epsilon_{t-i}^2 - \nu_0) + \sum_{j=1}^p \lambda_{j+q} \{ \ln \epsilon_{t-j}^2 - \nu_0 - u_{t-j}(\theta_0) \}.$$

The process  $(\ln \epsilon_t^2 - \nu_0)$  thus satisfies an ARMA( $r-1, p-1$ ). This is impossible under the identifiability conditions of Theorem 2.1. Therefore  $J$  is invertible.

Introducing the notation

$$\dot{K}_\theta := \frac{\partial^2 u_t(\theta)}{\partial \theta \partial \nu} = - \frac{\partial \mathcal{A}_\theta(1)}{\partial \theta \mathcal{B}_\theta(1)},$$

we also have

$$\frac{\partial^2 \ell_t(\vartheta_0)}{\partial \vartheta \partial \nu} = \frac{1 - \eta_t^2}{2} \begin{pmatrix} \dot{K}_\theta \\ 0 \end{pmatrix} + \frac{\eta_t^2 \mathcal{A}_\theta(1)}{2 \mathcal{B}_\theta(1)} \begin{pmatrix} \frac{\partial u_t(\theta_0)}{\partial \theta} \\ 1 \end{pmatrix}.$$

Thus

$$K := E \frac{\partial^2 \ell_t(\vartheta_0)}{\partial \vartheta \partial \nu} = \begin{pmatrix} 0_{d-1} \\ \frac{\mathcal{A}_\theta(1)}{2\mathcal{B}_\theta(1)} \end{pmatrix}.$$

**f) Convergence of  $(J_n, K_n)$  to  $(J, K)$ .** With the notation (3.10), for  $i, j \in \{1, p+q\}^2$ , we have

$$\frac{\partial^2 \ell_t(\vartheta, \nu)}{\partial \theta_i \partial \theta_j} = \frac{1 - e^{u_t(\theta, \nu) + \mu}}{2} \frac{\partial^2 u_t(\theta, \nu)}{\partial \theta_i \partial \theta_j} - \frac{e^{u_t(\theta, \nu) + \mu}}{2} \frac{\partial u_t(\theta, \nu)}{\partial \theta_i} \frac{\partial u_t(\theta, \nu)}{\partial \theta_j}.$$

Write

$$2 \frac{\partial^2 \ell_t(\vartheta, \nu)}{\partial \theta_i \partial \theta_j} - 2 \frac{\partial^2 \ell_t(\vartheta_0, \nu_0)}{\partial \theta_i \partial \theta_j} = c_1 + c_2 + c_3,$$



where, for  $m = 1, 2, 3$ , the  $c_m = c_{mt}(\theta, \nu)$  are defined by

$$\begin{aligned} c_1 &= \frac{\partial^2 u_t(\theta, \nu)}{\partial \theta_i \partial \theta_j} - \frac{\partial^2 u_t(\theta_0, \nu_0)}{\partial \theta_i \partial \theta_j}, \\ c_2 &= \left\{ e^{u_t(\theta, \nu)} - e^{u_t(\theta_0, \nu_0)} \right\} \left\{ \frac{\partial^2 u_t(\theta, \nu)}{\partial \theta_i \partial \theta_j} + \frac{\partial u_t(\theta, \nu)}{\partial \theta_i} \frac{\partial u_t(\theta, \nu)}{\partial \theta_j} \right\}, \\ c_3 &= e^{u_t(\theta_0, \nu_0)} \left\{ c_1 + \frac{\partial u_t(\theta, \nu)}{\partial \theta_i} \frac{\partial u_t(\theta, \nu)}{\partial \theta_j} - \frac{\partial u_t(\theta_0, \nu_0)}{\partial \theta_i} \frac{\partial u_t(\theta_0, \nu_0)}{\partial \theta_j} \right\}. \end{aligned}$$

Introducing the notation  $\psi_k^{i,j}(\theta) = \partial^2 \psi_k(\theta) / \partial \theta_i \partial \theta_j$ , we have

$$c_1 = \sum_{\ell=0}^{\infty} \left\{ \psi_\ell^{i,j}(\theta) - \psi_\ell^{i,j}(\theta_0) \right\} (\ln \epsilon_{t-\ell}^2 - \nu) - \sum_{\ell=0}^{\infty} \psi_\ell^{i,j}(\theta_0) (\nu - \nu_0).$$

Recall that  $V_k(\vartheta_0)$  denotes the ball of center  $\vartheta_0$  and radius  $1/k$ . Since there is no risk of confusion, we also denote by  $V_k(\theta_0)$  be the ball of center  $\theta_0$  and radius  $1/k$ . Noting that  $\psi_\ell^{i,j}(\theta) \leq K \rho^\ell$ ,  $E |\ln \epsilon_t^2| < \infty$  and  $\theta \mapsto \psi_\ell^{i,j}(\theta)$  is continuous, the dominated convergence theorem entails

$$\lim_{k \rightarrow \infty} E \sup_{(\theta, \nu) \in V_k(\theta_0) \times [\nu_0 - \frac{1}{k}, \nu_0 + \frac{1}{k}]} |c_{mt}(\theta, \nu)| = 0 \quad (3.13)$$

for  $m = 1$ . We now consider the term  $c_2$ . By already given arguments, it is easy to show that

$$E \sup_{\theta \in \Theta, \nu \in [a, b]} \left\{ \frac{\partial^2 u_t(\theta, \nu)}{\partial \theta_i \partial \theta_j} + \frac{\partial u_t(\theta, \nu)}{\partial \theta_i} \frac{\partial u_t(\theta, \nu)}{\partial \theta_j} \right\}^2 < \infty.$$

In view of Remark 2.2, by the dominated convergence theorem

$$\lim_{k \rightarrow \infty} E \sup_{(\theta, \nu) \in V_k(\theta_0) \times [\nu_0 - \frac{1}{k}, \nu_0 + \frac{1}{k}]} (e^{u_t(\theta, \nu)} - e^{u_t(\theta_0, \nu_0)})^2 = 0.$$

We then obtain (3.13) for  $m = 2$  by the Cauchy-Schwarz inequality. Similar arguments show (3.13) for  $m = 3$ . Doing the same derivations when  $\partial \theta_i$  and/or  $\partial \theta_j$  are replaced by  $\partial \mu$  in the second order derivatives, we finally obtain that  $\forall \varepsilon > 0$ , there exists an integer  $k_\varepsilon$  such that

$$E \sup_{(\vartheta, \nu) \in V_k(\vartheta_0) \times [\nu_0 - \frac{1}{k}, \nu_0 + \frac{1}{k}]} \left\| \frac{\partial^2 \ell_t(\vartheta, \nu)}{\partial \vartheta \partial \vartheta'} - \frac{\partial^2 \ell_t(\vartheta_0, \nu_0)}{\partial \vartheta \partial \vartheta'} \right\| < \varepsilon$$

for all  $k \geq k_\varepsilon$ . By the ergodic theorem, with probability one we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 \ell_t(\vartheta_0, \nu_0)}{\partial \vartheta \partial \vartheta'} = J.$$

Because  $(\hat{\vartheta}_n, \nu_n) \in V_{k_\varepsilon}(\vartheta_0) \times [\nu_0 - \frac{1}{k_\varepsilon}, \nu_0 + \frac{1}{k_\varepsilon}]$  for  $n$  large enough, we have

$$\lim_{n \rightarrow \infty} \|J_n - J\| < \varepsilon \quad \text{a.s.}$$

Since  $\varepsilon$  is an arbitrary positive number, the limit is actually zero. By exactly the same arguments, it can be show that  $K_n \rightarrow K$  a.s.

**g) Joint asymptotic distribution of  $\hat{\vartheta}_n$  and  $\nu_n$ .** By (3.12), d), e) and f) we have

$$\begin{aligned} \sqrt{n} \begin{pmatrix} \hat{\vartheta}_n - \vartheta_0 \\ \nu_n - \nu_0 \end{pmatrix} &\stackrel{o_P(1)}{=} \begin{pmatrix} -J^{-1} & -J^{-1}K \\ 0'_d & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial \ell_t(\vartheta_0)}{\partial \vartheta} \\ \sqrt{n}(\nu_n - \nu_0) \end{pmatrix} \\ &\xrightarrow{d} \mathcal{N} \left\{ 0, \begin{pmatrix} \Sigma_\theta & 0_{d-1} & 0_{d-1} \\ 0'_{d-1} & \sigma_\mu^2 & \sigma_{\mu\nu} \\ 0'_{d-1} & \sigma_{\mu\nu} & \sigma_\nu^2 \end{pmatrix} \right\}, \end{aligned}$$

where

$$\Sigma_\theta = 4\tau_\eta \Sigma_u^{-1}, \quad \Sigma_u = E \frac{\partial u_t}{\partial \theta} \frac{\partial u_t}{\partial \theta'}(\theta_0), \quad \sigma_\mu^2 = \text{Var}(\eta_1^2 - \ln \eta_1^2) \quad (3.14)$$

and  $\sigma_{\mu\nu} = 2\xi + \sigma_\nu^2 \mathcal{A}_{\theta_0}(1) / \mathcal{B}_{\theta_0}(1)$ .

**h) Joint asymptotic distribution of  $(\hat{\omega}_n, \hat{\vartheta}'_n, \nu_n)$ .** By (2.8), we have

$$\begin{aligned} \hat{\omega}_n - \omega_0 &= -\mathcal{B}_{\theta_0}(1)(\hat{\mu}_n - \mu_0) - \{\mathcal{B}_{\hat{\theta}_n}(1) - \mathcal{B}_{\theta_0}(1)\} \hat{\mu}_n \\ &\quad + \mathcal{A}_{\theta_0}(1)(\nu_n - \nu_0) + \{\mathcal{A}_{\hat{\theta}_n}(1) - \mathcal{A}_{\theta_0}(1)\} \nu_n \\ &= -\mathcal{B}_{\theta_0}(1)(\hat{\mu}_n - \mu_0) + \mathcal{A}_{\theta_0}(1)(\nu_n - \nu_0) \\ &\quad + \sum_{j=1}^p (\hat{\beta}_j - \beta_{0j}) \hat{\mu}_n - \sum_{i=1}^r (\hat{\alpha}_i + \hat{\beta}_i - \alpha_{0i} - \beta_{0i}) \nu_n. \end{aligned}$$

Recalling the notation  $\gamma = (-\nu_0 1'_q, (\mu_0 - \nu_0) 1'_p)'$ , we have

$$\begin{aligned} \sqrt{n} \begin{pmatrix} \hat{\omega}_n - \omega_0 \\ \hat{\vartheta}'_n - \vartheta_0 \\ \nu_n - \nu_0 \end{pmatrix} &\stackrel{o_P(1)}{=} \begin{pmatrix} \gamma' & -\mathcal{B}_{\theta_0}(1) & \mathcal{A}_{\theta_0}(1) \\ & I_{d+1} & \end{pmatrix} \sqrt{n} \begin{pmatrix} \hat{\vartheta}_n - \vartheta_0 \\ \nu_n - \nu_0 \end{pmatrix} \\ &\xrightarrow{d} \mathcal{N} \left\{ 0, \begin{pmatrix} \sigma_\omega^2 & \gamma' \Sigma_\theta & \sigma_{\omega\mu} & \sigma_{\omega\nu} \\ \gamma' \Sigma_\theta & \Sigma_\theta & 0_{d-1} & 0_{d-1} \\ \sigma_{\omega\mu} & 0'_{d-1} & \sigma_\mu^2 & \sigma_{\mu\nu} \\ \sigma_{\omega\nu} & 0'_{d-1} & \sigma_{\mu\nu} & \sigma_\nu^2 \end{pmatrix} \right\}, \end{aligned}$$

where

$$\begin{aligned} \sigma_{\omega\mu} &= -\mathcal{B}_{\theta_0}(1) \sigma_\mu^2 + \mathcal{A}_{\theta_0}(1) \sigma_{\mu\nu} = -\mathcal{B}_{\theta_0}(1) \{ \text{Var}(\eta_1^2) - \text{Cov}(\eta_1^2, \ln \eta_1^2) \}, \\ \sigma_{\omega\nu} &= -\mathcal{B}_{\theta_0}(1) \sigma_{\mu\nu} + \mathcal{A}_{\theta_0}(1) \sigma_\nu^2 = -\frac{\mathcal{B}_{\theta_0}^2(1)}{\mathcal{A}_{\theta_0}(1)} \text{Cov}(\eta_1^2, \ln \eta_1^2), \\ \sigma_\omega^2 &= \gamma' \Sigma_\theta \gamma + \mathcal{B}_{\theta_0}^2(1) \sigma_\mu^2 - 2\mathcal{B}_{\theta_0}(1) \mathcal{A}_{\theta_0}(1) \sigma_{\mu\nu} + \mathcal{A}_{\theta_0}^2(1) \sigma_\nu^2 \\ &= \gamma' \Sigma_\theta \gamma + \mathcal{B}_{\theta_0}^2(1) \text{Var}(\eta_1^2). \end{aligned}$$

The conclusion follows by direct computation.  $\square$

## 4 Efficiency comparison

### 4.1 Asymptotic comparison

A nice feature of the log-GARCH model is that, contrary to the standard GARCH model, a closed form is available for the asymptotic covariance matrix of the QMLE's. This enables direct comparison between the asymptotic efficiency of the Cex- $\chi^2$  ARMA-QMLE, the Gaussian ARMA-QMLE and the standard QMLE. For simplicity we do this only for the log-GARCH(1,1) specification.

Proceeding as in [McLeod \(1978\)](#), and lightening the notation by omitting subscripts, we have

$$(1 - \beta L) \frac{\partial u_t(\theta)}{\partial \alpha} = -\ln \epsilon_{t-1}^2 + \nu_0, \quad (1 - \beta L) \frac{\partial u_t(\theta)}{\partial \beta} - u_{t-1}(\theta) = -\ln \epsilon_{t-1}^2 + \nu_0,$$

which gives

$$\begin{aligned} \frac{\partial u_t(\theta)}{\partial \alpha} &= -\{1 - (\alpha_0 + \beta_0)L\}^{-1} u_{t-1}, \\ \frac{\partial u_t(\theta)}{\partial \beta} &= \{1 - \beta_0 L\}^{-1} u_{t-1} - \{1 - (\alpha_0 + \beta_0)L\}^{-1} u_{t-1}. \end{aligned}$$

As in [Brockwell and Davis \(2006, p. 260\)](#), we then obtain

$$\Sigma_u = \text{Var}(\ln \eta_1^2) \cdot \begin{pmatrix} \frac{1}{1 - (\alpha_0 + \beta_0)^2} & \frac{1}{1 - (\alpha_0 + \beta_0)^2} - \frac{1}{1 - \beta_0(\alpha_0 + \beta_0)} \\ \frac{1}{1 - (\alpha_0 + \beta_0)^2} - \frac{1}{1 - \beta_0(\alpha_0 + \beta_0)} & \frac{1}{1 - \beta_0^2} + \frac{1}{1 - (\alpha_0 + \beta_0)^2} - \frac{1}{1 - \beta_0(\alpha_0 + \beta_0)} \end{pmatrix}$$

with inverse

$$\Sigma_u^{-1} = \frac{1}{\text{Var}(\ln \eta_1^2)} \cdot \begin{pmatrix} 1 - \beta_0^2(\alpha_0 + \beta_0)^2 & -\frac{(\alpha_0 + \beta_0)(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0))}{\alpha_0} \\ -\frac{(\alpha_0 + \beta_0)(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0))}{\alpha_0} & \frac{(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0))^2}{\alpha_0^2} \end{pmatrix}.$$

This gives the following explicit formula for the asymptotic variance of the Cex- $\chi^2$  estimator of  $(\alpha_0, \beta_0)$ :

$$E(\eta_t^4 - 1) \cdot \Sigma_u^{-1}. \quad (4.1)$$

It can be seen that the variance of  $\beta_0$  explodes when  $\alpha_0 \rightarrow 0$ . This is not surprising because the existence of common roots in the AR and MA polynomials is excluded for the consistency. The asymptotic variance of  $\hat{\omega}$  for the log-GARCH(1,1) is given by  $E(\eta_t^2 - 1)(\mathcal{B}_{\theta_0}^2(1) + \gamma' \Sigma_u^{-1} \gamma)$  with  $\mathcal{B}_{\theta_0}^2(1) = (1 - \beta_0)^2$  and  $\gamma = (\nu_0, \mu_0 - \nu_0)'$ , where  $\mu_0 = E(\ln \eta_t^2)$  and  $\nu_0 = (\omega_0 + (1 - \beta_0)\mu_0)/(1 - \alpha_0 - \beta_0)$ . The full asymptotic covariance matrix the Cex- $\chi^2$

estimator of  $(\omega_0, \alpha_0, \beta_0)$  is given by

$$(E\eta_1^4 - 1) \begin{pmatrix} \mathcal{B}_{\theta_0}^2(1) + \gamma' \Sigma_u^{-1} \gamma & \gamma' \Sigma_u^{-1} \\ \Sigma_u^{-1} \gamma & \Sigma_u^{-1} \end{pmatrix}, \quad (4.2)$$

where  $\gamma = (-\nu_0 1'_q, (\mu_0 - \nu_0) 1'_p)'$  and where  $1_s$  denotes the all-ones vector of dimension  $s$ , see Theorem 2.2. The asymptotic variance of the  $\mu_0$  estimate is  $Var(\eta_t - \ln \eta_t^2)$ , see (3.14).

In comparing (4.2) with the asymptotic covariance of the standard QMLE, it can be shown that their diagonal is the same (see Appendix A). This is a somewhat surprising result, since one might have expected that the two-step nature of the Cex- $\chi^2$  ARMA-QMLE would make it less efficient asymptotically. In comparing (4.2) with the asymptotic covariance of the Gaussian ARMA-QMLE, only the variances of the estimates of  $\alpha_0$ ,  $\beta_0$  and  $\mu_0$  can be compared, since the asymptotic variance of the  $\omega_0$ -estimate is not available. The variance of  $\mu_0$  is exactly the same, see Theorem 1 in Sucarrat et al. (2013). As for the variances of  $\alpha_0$  and  $\beta_0$ , it can be shown (see Appendix B) that the covariance of these two parameters is  $Var(\ln \eta_1^2) \cdot \Sigma_u^{-1}$ . In other words, the asymptotic variances are higher than those of the Cex- $\chi^2$  when  $E(\eta_t^4 - 1)/Var(\ln \eta_t^2) = Var(\eta^2)/Var(\ln \eta_t^2) < 1$ . In most cases this will indeed be the case, e.g. when  $\eta_t$  is  $N(0, 1)$  which yields a fraction equal to about 0.41, see Table 1. For some very fat-tailed and/or very skewed densities, however, e.g.  $\eta_t \sim t(5)$  and  $\eta_t \sim t(5, 0.7)$ , then the fraction is approximately 1.3 and 2.1, respectively. In fact, when  $\eta_t$  is symmetric  $t$  then the Cex- $\chi^2$  ARMA-QMLE is more efficient for degrees of freedom equal to or greater than 7. Somewhere between 6 and 7 degrees of freedom the Gaussian ARMA-QMLE becomes more efficient.

## 4.2 Finite sample comparison

An extensive set of Monte Carlo simulations were undertaken in order to verify our theoretical results, and in order to compare the finite sample properties of the Cex- $\chi^2$  ARMA-QMLE with those of the Gaussian ARMA-QMLE and the standard QMLE. The full set of results from the simulations is available on request.

Table 2 summarises the results when  $n = 1000$ . Unsurprisingly, the Cex- $\chi^2$  is substantially more efficient than the Gaussian ARMA-QMLE when  $\eta_t \sim N(0, 1)$ . This is in correspondence with the asymptotic efficiency ratio, although the empirical ratios are lower than their asymptotic counterparts for all but two estimates. When  $\eta_t$  is  $t$  or skewed  $t$  with 5 degrees of freedom, then the Gaussian ARMA-QMLE is more efficient according to the asymptotic efficiency ratios. Empirically, however, the Cex- $\chi^2$  is in fact substantially more efficient at times. This is particularly the case for  $\omega_0$ ,  $\beta_0$  and  $\mu_0$  when  $\alpha_0$  and  $\beta_0$  have empirically relevant values on the DGP (i.e. when  $\alpha_0$  and  $\beta_0$  is either 0.1 and 0.8 or 0.05 and 0.94). A possible reason for this is that the Gaussian ARMA-QMLE is in fact a three-step estimator, whereas the Cex- $\chi^2$  is a two-step estimator.

Asymptotically the Cex- $\chi^2$  and standard QMLEs are equally efficient. In finite samples, however, they are sometimes very different in relative terms. The difference is seemingly unsystematic, but an interesting exception is when  $\eta_t \sim N(0, 1)$ . In this case the Cex- $\chi^2$  is slightly more efficient for all but one estimate. Another finding of interest is

that skewness can matter a lot. For example, the standard QMLE is substantially more efficient for all three parameters in the case where volatility is at its most persistent (i.e.  $\alpha_0 = 0.05$  and  $\beta_0 = 0.94$ ) and  $\eta_t \sim t(5)$  (i.e. symmetry). By contrast, in the skewed case (i.e.  $\eta_t \sim t(5, 0.7)$ ,  $\alpha_0 = 0.05$  and  $\beta_0 = 0.94$ ) then the Cex- $\chi^2$  is substantially more efficient. Finally, as expected, the simulations confirm that when  $n$  goes towards infinity (in the simulations we studied  $n = 10000$  and  $n = 100000$ ), then the finite sample properties go towards their asymptotic counterparts.

## 5 Empirical application

Forecasts of inflation play an important role in monetary policy making. This is particularly true for inflation-targeting central banks, where the conditional forecasts in part determine policy interest rates. Accordingly, precise forecasts of the uncertainty of the conditional forecasts are of great importance.

When Engle (1982) proposed his ARCH model, he used forecasts of the uncertainty of quarterly UK inflation to illustrate the usefulness of the model. However, the ARCH(4) specification he used in his illustration was severely restricted in order to ensure the positivity of the variance estimates (see Engle (1982, p. 1002)). In fact, instead of estimating the ARCH parameters freely he imposed a linearly declining relationship. Here, we illustrate the versatility and usefulness of the log-GARCH model by fitting specifications of up to twelve orders – without any parameter restrictions – to the residuals of a dynamic model of monthly Euro-area inflation. The underlying series is the Harmonised Index of Consumer Prices (HICP) from January 2001 to June 2013 ( $n = 150$  observations), and the source of the data is the European Central Bank (<http://www.ecb.int/>). We denote the HICP index-value at  $t$  by  $p_t$ , and define the 12-month inflation or %-change as  $y_t = 100 \cdot (p_t - p_{t-12})/p_{t-12}$ . Figure 1 contains graphs of the two series.

The estimation results of the dynamic model of inflation – a simplified AR(12) specification – is contained in Table 3. General-to-Specific (Gets) model selection with the *R* package AutoSEARCH (see Sucarrat (2011) and Sucarrat and Escribano (2012)) was used to obtain the model, with the starting model being an AR(12) specification. The diagnostic tests suggests that there is little or no autocorrelation in the residuals  $\hat{\epsilon}_t$ , since the  $p$ -values of the tests of no autocorrelation up to the 12th. and 13th. lags, respectively, are 0.11 and 0.08. However, the same diagnostic tests suggest that there is significant ARCH in the residuals, since the  $p$ -values are equal to 0.01 and 0.02 in the two tests of no ARCH up to the 12th. and 13th. lags, respectively.

Table 4 contains the estimation results of a log-GARCH(1,12) model of the log-variance of  $\hat{\epsilon}_t$ . The diagnostic tests show that it successfully removes the ARCH in the standardised residuals  $\hat{\eta}_t$ , since the two  $p$ -values associated with the tests of no ARCH increase to 0.90 and 0.92, respectively. The diagnostic tests for autocorrelation also improve, since they are 0.34 and 0.35, respectively. Several of the ARCH-lags are estimated to be negative, but only the first and fifth – which are both positive – are significant at 5%. Sequential backwards elimination of regressors with  $t$ -ratios smaller than 2 in absolute value, however, leads to the model in Table 5. There, the first-order GARCH term together with ARCH-terms 1,3,5 and 7 are significant, and two of the ARCH-terms (3 and 7) are even negative. This

suggests there might be some cyclical seasonal dynamics in the uncertainty of inflation. The log-moment  $\mu_0$  is estimated to  $-1.394$ , which is close to normality (i.e.  $-1.27$ ). The test-statistic  $(-1.394 + 1.27)/0.126$  is equal to  $-0.98$ , which means the null of normality is not rejected at usual significance level.

## 6 Conclusions

We have proposed a QMLE for log-GARCH models via the ARMA representation with the centred exponential chi-squared (Cex- $\chi^2$ ) distribution as instrumental density. Estimation via the ARMA representation is attractive because it enables a vast amount of already established result. In particular, the problem of zero errors is readily resolved, see [Sucarrat and Escribano \(2013\)](#). We have proved the consistency and asymptotic normality of the Cex- $\chi^2$  ARMA-QMLE under mild conditions, and we have compared its efficiency both asymptotically and in finite samples for the log-GARCH(1,1). In finite samples the Cex- $\chi^2$  ARMA-QMLE is in general less biased and more efficient than the Gaussian ARMA-QMLE, and equally efficient as the standard QMLE. Finally, we have illustrated the usefulness and versatility of the log-GARCH in an empirical application to monthly inflation modelling.

The results in this paper are likely to be extendable in several ways. The most straightforward concerns the addition of leverage or asymmetry terms, and of additional exogenous or predetermined terms ("X"). Since the relationships between the log-GARCH and ARMA parameters are not affected by the linear addition of terms, it is straightforward to add leverage and exogenous terms to the log-GARCH specification. We conjecture that consistency and asymptotic normality results should not be too difficult to establish for a Cex- $\chi^2$  QMLE in these instances.

## References

- Berkes, I., L. Horvath, and P. Kokoszka (2003). GARCH processes: structure and estimation. *Bernoulli* 9, 201–227.
- Billingsley, P. (1995). *Probability and Measure*.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* 31, 307–327.
- Brockwell, P. J. and R. A. Davis (2006). *Time Series: Theory and Methods*. New York: Springer. 2nd. Edition.
- Brownlees, C., F. Cipollini, and G. Gallo (2012). Multiplicative Error Models. In L. Bauwens, C. Hafner, and S. Laurent (Eds.), *Handbook of Volatility Models and Their Applications*. New Jersey: Wiley.
- Creal, D., S. J. Koopmans, and A. Lucas (2013). Generalized Autoregressive Score Models with Applications. *Journal of Applied Econometrics* 28, 777–795.

- Engle, R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* 50, 987–1008.
- Engle, R. F. and T. Bollerslev (1986). Reply. *Econometric Reviews* 5, 81–87.
- Fernández, C. and M. Steel (1998). On Bayesian Modelling of Fat Tails and Skewness. *Journal of the American Statistical Association* 93, 359–371.
- Francq, C., L. Horvath, and J.-M. Zakoian (2011). Merits and drawbacks of variance targeting in GARCH models. *Journal of Financial Econometrics* 9, 619–656.
- Francq, C., O. Wintenberger, and J.-M. Zakoian (2013). GARCH Models Without Positivity Constraints: Exponential or Log-GARCH? Forthcoming in *Journal of Econometrics*, <http://dx.doi.org/10.1016/j.jeconom.2013.05.004>.
- Francq, C. and J.-M. Zakoian (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* 10, 605–637.
- Francq, C. and J.-M. Zakoian (2010). *GARCH Models*. New York: Marcel Dekker.
- Franses, P. H., J. Neele, and D. Van Dijk (2001). Modelling asymmetric volatility in weekly Dutch temperature data. *Environmental Modeling and Software* 16, 131–137.
- Geweke, J. (1986). Modelling the Persistence of Conditional Variance: A Comment. *Econometric Reviews* 5, 57–61.
- Gourieroux, C., A. Monfort, and A. Trognon (1984). Pseudo Maximum Likelihood Methods: Theory. *Econometrics* 52, 681–700.
- Harvey, A. C. (2013). *Dynamic Models for Volatility and Heavy Tails*. New York: Cambridge University Press.
- Koopman, S. J., M. Ooms, and M. A. Carnero (2007). Periodic Seasonal REG-ARFIMA-GARCH Models for Daily Electricity Spot Prices. *Journal of the American Statistical Association* 102, 16–27.
- Ljung, G. and G. Box (1979). On a Measure of Lack of Fit in Time Series Models. *Biometrika* 66, 265–270.
- McLeod, A. (1978). On the Distribution of Residual Autocorrelations in Box-Jenkins Method. *Journal of the Royal Statistical Society B* 40, 296–302.
- Milhøj, A. (1987). A Multiplicative Parametrization of ARCH Models. Research Report 101, University of Copenhagen: Institute of Statistics.
- Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica* 59, 347–370.
- Pantula, S. (1986). Modelling the Persistence of Conditional Variance: A Comment. *Econometric Reviews* 5, 71–73.

- Straumann, D. and T. Mikosch (2006). Quasi-Maximum-Likelihood Estimation in Conditionally Heteroscedastic Time Series: A Stochastic Recurrence Equations Approach. *The Annals of Statistics* 34, 2449–2495.
- Sucarrat, G. (2011). AutoSEARCH: An R Package for Automated Financial Modelling.
- Sucarrat, G. and Á. Escribano (2012). Automated Model Selection in Finance: General-to-Specific Modelling of the Mean and Volatility Specifications. *Oxford Bulletin of Economics and Statistics* 74, 716–735.
- Sucarrat, G. and Á. Escribano (2013). Unbiased QML Estimation of Log-GARCH Models in the Presence of Zero Errors. MPRA Paper No. 50699. Online at <http://mpra.ub.uni-muenchen.de/50699/>.
- Sucarrat, G., S. Grønneberg, and Á. Escribano (2013). Estimation and Inference in Univariate and Multivariate Log-GARCH-X Models When the Conditional Density is Unknown. MPRA Paper No. 49344. Online at <http://mpra.ub.uni-muenchen.de/49344/>.
- Wald, A. (1949). Note on the consistency of the maximum likelihood estimate. *Annals of Mathematical Statistics* 20, 595–601.
- White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity. *Econometrica* 48, 817–838.
- Wintenberger, O. (2012). QML estimation of the continuously invertible EGARCH(1,1) model. Unpublished working paper.

## A Asymptotic variance of the standard QMLE

In this Appendix we show that the asymptotic variances of the standard QMLE are equal to the asymptotic variances of the Cex- $\chi^2$  ARMA-QMLE in the log-GARCH(1,1) case.

The asymptotic distribution of the standard gaussian QMLE is equal to  $E(\eta_1^4 - 1)J_Q^{-1}$  where  $J_Q$  is the second-order moment of the vector of the derivatives of  $\ln \sigma_t^2$  with respect to the parameters. To compute  $J_Q$  in the log-GARCH(1,1) case, first recall that

$$E \ln \sigma_1^2 = \frac{\omega_0 + \alpha_0 \mu_0}{1 - (\alpha_0 + \beta_0)}, \quad E \ln \epsilon_1^2 = \nu_0 = \frac{\omega_0 + (1 - \beta_0) \mu_0}{1 - (\alpha_0 + \beta_0)}.$$

Now note that

$$\begin{aligned} \frac{\partial \ln \sigma_t^2}{\partial \omega} &= \frac{1}{1 - \beta_0}, \\ \frac{\partial \ln \sigma_t^2}{\partial \alpha} &= (1 - \beta_0 L)^{-1} \ln \epsilon_{t-1}^2, \\ \frac{\partial \ln \sigma_t^2}{\partial \beta} &= (1 - \beta_0 L)^{-1} \ln \sigma_{t-1}^2. \end{aligned}$$



The first row and column of  $J_Q$  easily follows, and we have

$$\begin{aligned}
J_Q &= \begin{pmatrix} \frac{1}{(1-\beta_0)^2} & \frac{\omega_0+(1-\beta_0)\mu_0}{(1-\beta_0)^2(1-\alpha_0-\beta_0)} & \frac{\omega_0+\alpha_0\mu_0}{(1-\beta_0)^2(1-\alpha_0-\beta_0)} \\ & J_Q(2,2) & J_Q(2,3) \\ & & J_Q(3,3) \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{\mathcal{B}_{\theta_0}^2(1)} & \frac{\nu_0}{\mathcal{B}_{\theta_0}^2(1)} & \frac{(\nu_0-\mu_0)}{\mathcal{B}_{\theta_0}^2(1)} \\ & J_Q(2,2) & J_Q(2,3) \\ & & J_Q(3,3) \end{pmatrix}.
\end{aligned} \tag{A.1}$$

Since  $\ln \epsilon_t^2$  follows an ARMA(1,1)

$$\{1 - (\alpha_0 + \beta_0)L\} \ln \epsilon_t^2 = \omega_0 + (1 - \beta_0)\mu_0 + u_t - \beta_0 u_{t-1},$$

we have the AR(1) equation

$$\{1 - (\alpha_0 + \beta_0)L\} \frac{\partial \ln \sigma_t^2}{\partial \alpha} = \frac{\omega_0}{1 - \beta_0} + \mu_0 + u_{t-1}.$$

The second-order moment of this AR(1) gives

$$\begin{aligned}
J_Q(2,2) &= \frac{\sigma_v^2}{1 - (\alpha_0 + \beta_0)^2} + \left( \frac{\omega_0 + (1 - \beta_0)\mu_0}{(1 - \beta_0)(1 - \alpha_0 - \beta_0)} \right)^2, \\
&= \frac{\sigma_v^2}{1 - (\alpha_0 + \beta_0)^2} + \frac{\nu_0^2}{\mathcal{B}_{\theta_0}^2(1)},
\end{aligned} \tag{A.2}$$

where  $\sigma_v^2 = \text{Var} \ln \eta_1^2$ . Since  $\ln \sigma_t^2$  follows an AR(1) of the form

$$\{1 - (\alpha_0 + \beta_0)L\} \ln \sigma_t^2 = \omega_0 + \alpha_0\mu_0 + \alpha_0 u_{t-1},$$

we also have the AR(2) equation

$$(1 - \beta_0 L) \{1 - (\alpha_0 + \beta_0)L\} \frac{\partial \ln \sigma_t^2}{\partial \beta} = \omega_0 + \alpha_0\mu_0 + \alpha_0 u_{t-2}.$$

The latter equation can be rewritten as

$$\frac{\partial \ln \sigma_t^2}{\partial \beta} = \frac{\omega_0 + \alpha_0\mu_0}{(1 - \beta_0) \{1 - (\alpha_0 + \beta_0)\}} - \sum_{i=0}^{\infty} \beta^{i+1} u_{t-2-i} + \sum_{i=0}^{\infty} (\alpha + \beta)^{i+1} u_{t-2-i}.$$

Using that  $(\omega_0 + \alpha_0\mu_0)/(1 - \alpha_0 - \beta_0) = (\nu_0 - \mu_0)$  it follows that

$$\begin{aligned}
J_Q(2, 3) &= \frac{(\alpha_0 + \beta_0)^2 \sigma_v^2}{1 - (\alpha_0 + \beta_0)^2} - \frac{\beta_0(\alpha_0 + \beta_0) \sigma_v^2}{1 - \beta_0(\alpha_0 + \beta_0)} + \frac{\{\omega_0 + (1 - \beta_0)\mu_0\} \{\omega_0 + \alpha_0\mu_0\}}{(1 - \beta_0)^2(1 - \alpha_0 - \beta_0)^2}, \\
&= \frac{(\alpha_0 + \beta_0)^2 \sigma_v^2}{1 - (\alpha_0 + \beta_0)^2} - \frac{\beta_0(\alpha_0 + \beta_0) \sigma_v^2}{1 - \beta_0(\alpha_0 + \beta_0)} + \frac{\nu_0(\nu_0 - \mu_0)}{\mathcal{B}_{\theta_0}^2(1)}, \\
&= \frac{\alpha_0(\alpha_0 + \beta_0) \sigma_v^2}{(1 - (\alpha_0 + \beta_0)^2)(1 - \beta_0(\alpha_0 + \beta_0))} + \frac{\nu_0(\nu_0 - \mu_0)}{\mathcal{B}_{\theta_0}^2(1)}, \tag{A.3}
\end{aligned}$$

$$\begin{aligned}
J_Q(3, 3) &= \frac{\beta_0^2 \sigma_v^2}{1 - \beta_0^2} + \frac{(\alpha_0 + \beta_0)^2 \sigma_v^2}{1 - (\alpha_0 + \beta_0)^2} - \frac{2\beta_0(\alpha_0 + \beta_0) \sigma_v^2}{1 - \beta_0(\alpha_0 + \beta_0)} + \frac{(\omega_0 + \alpha_0\mu_0)^2}{(1 - \beta_0)^2(1 - \alpha_0 - \beta_0)^2}, \\
&= \frac{\beta_0^2 \sigma_v^2}{1 - \beta_0^2} + \frac{(\alpha_0 + \beta_0)^2 \sigma_v^2}{1 - (\alpha_0 + \beta_0)^2} - \frac{2\beta_0(\alpha_0 + \beta_0) \sigma_v^2}{1 - \beta_0(\alpha_0 + \beta_0)} + \frac{(\nu_0 - \mu_0)^2}{\mathcal{B}_{\theta_0}^2(1)}. \tag{A.4}
\end{aligned}$$

The determinant of  $J_Q$  simplifies to

$$\begin{aligned}
|J_Q| &= \frac{\sigma_v^2}{\mathcal{B}_{\theta_0}^2(1)(1 - (\alpha_0 + \beta_0)^2)} \left[ \frac{\beta_0^2 \sigma_v^2}{(1 - \beta_0^2)} + \frac{(\alpha_0 + \beta_0)^2 \sigma_v^2}{1 - (\alpha_0 + \beta_0)^2} - \frac{2\beta_0(\alpha_0 + \beta_0) \sigma_v^2}{1 - \beta_0(\alpha_0 + \beta_0)} \right. \\
&\quad \left. - \frac{\alpha_0^2(\alpha_0 + \beta_0)^2 \sigma_v^2}{(1 - (\alpha_0 + \beta_0)^2)(1 - \beta_0(\alpha_0 + \beta_0))^2} \right]. \tag{A.5}
\end{aligned}$$

Next, multiplying the term  $(\alpha_0 + \beta_0^2)/(1 - (\alpha_0 + \beta_0)^2)$  by  $[1 - \beta_0(\alpha_0 + \beta_0)]^2/[1 - \beta_0(\alpha_0 + \beta_0)]^2$  yields

$$\begin{aligned}
|J_Q| &= \frac{(\sigma_v^2)^2}{\mathcal{B}_{\theta_0}^2(1)(1 - (\alpha_0 + \beta_0)^2)} \left[ \frac{\beta_0^2}{(1 - \beta_0^2)} - \frac{2\beta_0(\alpha_0 + \beta_0)}{1 - \beta_0(\alpha_0 + \beta_0)} + \frac{(\alpha_0 + \beta_0)^2(1 - \beta_0^2)}{(1 - \beta_0(\alpha_0 + \beta_0))^2} \right] \\
&= \frac{(\sigma_v^2)^2}{\mathcal{B}_{\theta_0}^2(1)(1 - (\alpha_0 + \beta_0)^2)} \left[ \frac{\alpha_0^2}{(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0))^2} \right]. \tag{A.6}
\end{aligned}$$

The asymptotic variance of  $\hat{\omega}_0$  is  $E(\eta_t^4 - 1)$  times

$$J_Q^{-1}(1, 1) = \frac{1}{|J_Q|} \left[ \left( \frac{\sigma_v^2}{1 - (\alpha_0 + \beta_0)^2} + \frac{\nu_0^2}{\mathcal{B}_{\theta_0}^2(1)} \right) \cdot \left( \frac{\beta_0^2 \sigma_v^2}{(1 - \beta_0^2)} + \frac{(\alpha_0 + \beta_0)^2 \sigma_v^2}{1 - (\alpha_0 + \beta_0)^2} - \frac{2\beta_0(\alpha_0 + \beta_0)\sigma_v^2}{1 - \beta_0(\alpha_0 + \beta_0)} + \frac{(\nu_0 - \mu_0)^2}{\mathcal{B}_{\theta_0}^2(1)} \right) - \left( \frac{\alpha_0(\alpha_0 + \beta_0)\sigma_v^2}{(1 - (\alpha_0 + \beta_0)^2)(1 - \beta_0(\alpha_0 + \beta_0))} + \frac{\nu_0(\nu_0 - \mu_0)}{\mathcal{B}_{\theta_0}^2(1)} \right)^2 \right] \quad (\text{A.7})$$

$$= \frac{\beta_0^2 \mathcal{B}_{\theta_0}^2(1)(1 - \beta_0(\alpha_0 + \beta_0))^2}{\alpha_0^2} + \frac{(\alpha_0 + \beta_0)^2 \mathcal{B}_{\theta_0}^2(1)(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0))^2}{\alpha_0^2(1 - (\alpha_0 + \beta_0)^2)} - \frac{2\beta_0(\alpha_0 + \beta_0)\mathcal{B}_{\theta_0}^2(1)(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0))}{\alpha_0^2} - \frac{(\alpha_0 + \beta_0)^2 \mathcal{B}_{\theta_0}^2(1)(1 - \beta_0^2)}{1 - (\alpha_0 + \beta_0)^2} + \frac{\nu_0^2(1 - \beta_0^2(\alpha_0 + \beta_0)^2)}{\sigma_v^2} - \frac{2(\alpha_0 + \beta_0)\nu_0(\nu_0 - \mu_0)(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0)^2)}{\alpha_0 \sigma_v^2} + \frac{(\nu_0 - \mu_0)^2(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0))^2}{\alpha_0^2 \sigma_v^2} \quad (\text{A.8})$$

$$= \mathcal{B}_{\theta_0}^2(1) + \frac{\nu_0^2(1 - \beta_0^2(\alpha_0 + \beta_0)^2)}{\sigma_v^2} + \frac{2(\alpha_0 + \beta_0)\nu_0(\mu_0 - \nu_0)(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0)^2)}{\alpha_0 \sigma_v^2} + \frac{(\mu_0 - \nu_0)^2(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0))^2}{\alpha_0^2 \sigma_v^2} \quad (\text{A.9})$$

$$= \mathcal{B}_{\theta_0}^2(1) + \gamma' \Sigma_u^{-1} \gamma. \quad (\text{A.10})$$

Using the simplification-step in the computation of the determinant gives

$$J_Q^{-1}(2, 2) = \frac{1}{|J_Q|} \left[ \frac{1}{\mathcal{B}_{\theta_0}^2(1)} \left( \frac{\beta_0^2 \sigma_v^2}{(1 - \beta_0^2)} + \frac{(\alpha_0 + \beta_0)^2 \sigma_v^2}{1 - (\alpha_0 + \beta_0)^2} - \frac{2\beta_0(\alpha_0 + \beta_0)\sigma_v^2}{1 - \beta_0(\alpha_0 + \beta_0)} + \frac{(\nu_0 - \mu_0)^2}{\mathcal{B}_{\theta_0}^2(1)} \right) - \frac{(\nu_0 - \mu_0)^2}{[\mathcal{B}_{\theta_0}^2(1)]^2} \right] \quad (\text{A.11})$$

$$= \frac{\sigma_v^2}{|J_Q|} \left[ \frac{1}{\mathcal{B}_{\theta_0}^2(1)} \left( \frac{\alpha_0^2}{(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0))^2} - \frac{(\alpha_0 + \beta_0)^2(1 - \beta_0^2)}{(1 - \beta_0(\alpha_0 + \beta_0))^2} \right) + \frac{(\alpha_0 + \beta_0)^2}{1 - (\alpha_0 + \beta_0)^2} \right] - \frac{(\nu_0 - \mu_0)^2}{[\mathcal{B}_{\theta_0}^2(1)]^2} \quad (\text{A.12})$$

$$= \frac{\sigma_v^2}{|J_Q|} \left[ \frac{1}{\mathcal{B}_{\theta_0}^2(1)} \left( \frac{\alpha_0^2(1 - (\alpha_0 + \beta_0)^2) + \alpha_0^2(\alpha_0 + \beta_0)^2(1 - \beta_0^2)}{(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0))^2(1 - (\alpha_0 + \beta_0)^2)} \right) \right] \quad (\text{A.13})$$

$$= \frac{1 - \beta_0^2(\alpha_0 + \beta_0)^2}{\sigma_v^2}. \quad (\text{A.14})$$

Finally,

$$J_Q^{-1}(3, 3) = \frac{1}{|J_Q|} \left[ \frac{1}{\mathcal{B}_{\theta_0}^2(1)} \left( \frac{\sigma_v^2}{1 - (\alpha_0 + \beta_0)^2} + \frac{\nu_0^2}{\mathcal{B}_{\theta_0}^2(1)} \right) - \frac{\nu_0^2}{[\mathcal{B}_{\theta_0}^2(1)]^2} \right] \quad (\text{A.15})$$

$$= \frac{(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0))^2}{\sigma_v^2 \alpha_0^2}. \quad (\text{A.16})$$

## B Asymptotic variance of the Gaussian ARMA-QMLE

In this Appendix we derive the asymptotic variance-covariance matrix of the Gaussian ARMA-QMLE for the log-GARCH(1,1) case. We do this for the  $\alpha_0$  and  $\beta_0$  parameters only, since the expression for the  $\hat{\omega}_0$  is currently not available.

For the Gaussian QMLE of an ARMA(1,1) model the asymptotic  $2 \times 2$  covariance matrix of the AR and MA parameters  $\phi_1$  and  $\theta_1$  is given by

$$C^A = \frac{1 + \phi_1 \theta_1}{(\phi_1 + \theta_1)^2} \cdot \begin{pmatrix} (1 - \phi_1^2)(1 + \phi_1 \theta_1) & -(1 - \theta_1^2)(1 - \phi_1^2) \\ -(1 - \theta_1^2)(1 - \phi_1^2) & (1 - \theta_1^2)(1 + \phi_1 \theta_1) \end{pmatrix}, \quad (\text{B.1})$$

see for example [Brockwell and Davis \(2006, pp. 259-260\)](#). Using that  $\theta_1 = -\beta_0$  and  $\phi_1 = \alpha_0 + \beta_0$  in the ARMA(1,1) representation of the log-GARCH(1,1), we obtain the following asymptotic  $2 \times 2$  covariance matrix for  $\alpha_0$  and  $\beta_0$ :

$$\begin{pmatrix} C^A(1, 1) + C^A(2, 2) + 2C^A(1, 2) & -C^A(1, 2) - C^A(2, 2) \\ -C^A(1, 2) - C^A(2, 2) & C^A(2, 2) \end{pmatrix} \quad (\text{B.2})$$

$$= \begin{pmatrix} 1 - \beta_0^2(\alpha_0 + \beta_0)^2 & -\frac{(\alpha_0 + \beta_0)(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0))}{\alpha_0} \\ -\frac{(\alpha_0 + \beta_0)(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0))}{\alpha_0} & \frac{(1 - \beta_0^2)(1 - \beta_0(\alpha_0 + \beta_0))^2}{\alpha_0^2} \end{pmatrix}. \quad (\text{B.3})$$

Table 1: Asymptotic efficiency comparisons

Density	Cex- $\chi^2$ /Gaussian ARMA				Cex- $\chi^2$ /standard QMLE		
	$\omega_0$	$\alpha_0$	$\beta_0$	$\mu_0$	$\omega_0$	$\alpha_0$	$\beta_0$
$N(0, 1)$ :	–	0.405	0.405	1.000	1.000	1.000	1.000
$t(7)$ :	–	0.758	0.758	1.000	1.000	1.000	1.000
$t(7, 0.8)$ :	–	0.860	0.860	1.000	1.000	1.000	1.000
$t(5)$ :	–	1.335	1.335	1.000	1.000	1.000	1.000
$t(5, 0.7)$ :	–	2.073	2.073	1.000	1.000	1.000	1.000

Cex- $\chi^2$ /Gaussian ARMA, the ratio of the Cex- $\chi^2$  ARMA-QMLE over the Gaussian ARMA-QMLE asymptotic variances. Cex- $\chi^2$ /standard QMLE, the ratio of the Cex- $\chi^2$  ARMA-QMLE over the standard QMLE asymptotic variances.  $N(0, 1)$ ,  $\eta_t$  is standard normal.  $t(df, skew)$ ,  $\eta_t$  is distributed as a standardised skew  $t$  with  $df$  degrees of freedom and  $skew > 0$ . Symmetry obtains when  $skew = 1$ , and left-skewness (right-skewness) obtains when  $skew < 1$  ( $skew > 1$ ). The skewing method used is that of [Fernández and Steel \(1998\)](#)

Table 2: Finite sample ( $n = 1000$ ) efficiency comparisons

Density	DGP ( $\omega_0, \alpha_0, \beta_0$ )	Cex- $\chi^2$ /Gaussian ARMA				Cex- $\chi^2$ /standard QMLE		
		$\omega_0$	$\alpha_0$	$\beta_0$	$\mu_0$	$\omega_0$	$\alpha_0$	$\beta_0$
$N(0, 1)$ :	0, 0.3, 0.1	0.549	0.393	0.381	1.029	0.996	0.903	0.906
	0, 0.1, 0.8	0.113	0.377	0.166	0.800	0.981	0.966	0.950
	0, 0.05, 0.94	0.005	0.247	0.015	0.307	0.754	1.002	0.914
$t(5)$ :	0, 0.3, 0.1	1.020	1.157	1.211	0.822	1.042	1.040	1.013
	0, 0.1, 0.8	0.441	0.925	0.546	0.734	1.117	0.913	0.920
	0, 0.05, 0.94	0.086	0.810	0.186	0.334	1.504	1.264	1.616
$t(5, 0.7)$ :	0, 0.3, 0.1	1.150	1.370	1.511	1.005	1.021	0.800	0.939
	0, 0.1, 0.8	0.981	1.720	1.033	0.964	1.395	1.217	1.160
	0, 0.05, 0.94	0.063	1.248	0.159	0.351	0.329	0.748	0.775

Cex- $\chi^2$ /Gaussian ARMA, the ratio of the Cex- $\chi^2$  ARMA-QMLE over the Gaussian ARMA-QMLE empirical variances. Cex- $\chi^2$ /standard QMLE, the ratio of the Cex- $\chi^2$  ARMA-QMLE over the standard QMLE empirical variances.  $N(0, 1)$ ,  $\eta_t$  is standard normal.  $t(df, skew)$ ,  $\eta_t$  is distributed as a standardised skew  $t$  with  $df$  degrees of freedom and  $skew > 0$ . Symmetry obtains when  $skew = 1$ , and left-skewness (right-skewness) obtains when  $skew < 1$  ( $skew > 1$ ). The skewing method used is that of [Fernández and Steel \(1998\)](#)

Table 3: Estimation results of a simplified AR(12) model of 12-month Euro-area inflation

	Estimate	St.Err.	t-ratio	P-value
Intercept	0.382	0.102	3.740	0.000
AR(1)	1.073	0.060	17.923	0.000
AR(3)	-0.293	0.117	-2.498	0.014
AR(4)	0.335	0.140	2.399	0.018
AR(5)	-0.379	0.116	-3.254	0.001
AR(6)	0.278	0.104	2.664	0.009
AR(8)	-0.304	0.114	-2.662	0.009
AR(9)	0.214	0.094	2.272	0.025
AR(12)	-0.114	0.040	-2.865	0.005

$$AR_{12}(\hat{\epsilon}_t) : 18.35 \quad AR_{13}(\hat{\epsilon}_t) : 20.55$$

$$[p\text{-val.}] \quad [0.11] \quad [p\text{-val.}] \quad [0.08]$$

$$ARCH_{12}(\hat{\epsilon}_t) : 25.51 \quad ARCH_{13}(\hat{\epsilon}_t) : 25.60$$

$$[p\text{-val.}] \quad [0.01] \quad [p\text{-val.}] \quad [0.02]$$

The estimated model is  $y_t = b_0 + \sum_{k=1}^{12} b_k y_{t-k} + \epsilon_t$ ,  $t = 1, 2, \dots, 126$ . St.Err., standard error of the [White \(1980\)](#) type. P-value, the  $p$ -values from two-sided tests.  $AR_{12}(\hat{\epsilon}_t)$  and  $AR_{13}(\hat{\epsilon}_t)$ , [Ljung and Box \(1979\)](#) tests for 12th. and 13th. order autocorrelation in  $\hat{\epsilon}_t$ , respectively.  $ARCH_{12}(\hat{\epsilon}_t)$  and  $ARCH_{13}(\hat{\epsilon}_t)$ , [Ljung and Box \(1979\)](#) tests for 12th. and 13th. order autocorrelation in  $\hat{\epsilon}_t^2$ , respectively.

Table 4: Estimation results of a log-GARCH(1,12) model of 12-month Euro-area inflation volatility

	Estimate	St.Err.	t-ratio	P-value
Log-intercept( $\omega_0$ )	-1.052	–	–	–
ARCH(1)	0.130	0.060	2.186	0.031
ARCH(2)	0.021	0.112	0.183	0.855
ARCH(3)	-0.092	0.076	-1.223	0.224
ARCH(4)	-0.016	0.071	-0.221	0.825
ARCH(5)	0.198	0.079	2.513	0.013
ARCH(6)	0.054	0.137	0.394	0.695
ARCH(7)	-0.165	0.109	-1.516	0.132
ARCH(8)	0.068	0.100	0.678	0.499
ARCH(9)	0.002	0.077	0.027	0.979
ARCH(10)	-0.103	0.078	-1.324	0.188
ARCH(11)	0.049	0.103	0.481	0.632
ARCH(12)	0.036	0.060	0.607	0.545
GARCH(1)	0.383	0.700	0.548	0.585
$\mu_0$	-1.362	0.126	-10.786	0.000

$$AR_{12}(\hat{\eta}_t) : 13.41 \quad AR_{13}(\hat{\eta}_t) : 14.29$$

$[p\text{-val.}] \quad [0.34] \quad [p\text{-val.}] \quad [0.35]$

$$ARCH_{12}(\hat{\eta}_t) : 6.26 \quad ARCH_{13}(\hat{\eta}_t) : 6.54$$

$[p\text{-val.}] \quad [0.90] \quad [p\text{-val.}] \quad [0.92]$

The estimated model is  $\epsilon_t = \sigma_t \eta_t$ ,  $\ln \sigma_t = \omega_0 + \sum_{i=1}^{12} \alpha_{0i} \ln \epsilon_{t-i} + \beta_{01} \ln \sigma_{t-1}^2$ ,  $t = 1, 2, \dots, 126$ , where  $\epsilon_t$  is the error term of the dynamic model in Table 3. P-value, the  $p$ -values from two-sided tests.  $AR_{12}(\hat{\eta}_t)$  and  $AR_{13}(\hat{\eta}_t)$ , [Ljung and Box \(1979\)](#) tests for 12th. and 13th. order autocorrelation in  $\hat{\eta}_t$ , respectively.  $ARCH_{12}(\hat{\eta}_t)$  and  $ARCH_{13}(\hat{\eta}_t)$ , [Ljung and Box \(1979\)](#) tests for 12th. and 13th. order autocorrelation in  $\hat{\eta}_t^2$ , respectively.

Table 5: Estimation results of a simplified log-GARCH(1,12) model of 12-month Euro-area inflation volatility

	Estimate	St.Err.	t-ratio	P-value
Log-intercept ( $\omega_0$ )	-0.829	–	–	–
ARCH(1)	0.126	0.052	2.404	0.018
ARCH(3)	-0.112	0.043	-2.591	0.011
ARCH(5)	0.206	0.054	3.788	0.000
ARCH(7)	-0.109	0.052	-2.102	0.038
GARCH(1)	0.556	0.244	2.274	0.025
$\mu_0$	-1.394	0.126	-11.055	0.000

$$AR_{12}(\hat{\eta}_t) : 13.84 \quad AR_{13}(\hat{\eta}_t) : 15.17$$

$_{[p-val.]} \quad_{[0.31]} \quad_{[p-val.]} \quad_{[0.30]}$

$$ARCH_{12}(\hat{\eta}_t) : 7.83 \quad ARCH_{13}(\hat{\eta}_t) : 8.04$$

$_{[p-val.]} \quad_{[0.80]} \quad_{[p-val.]} \quad_{[0.84]}$

The estimated model is  $\epsilon_t = \sigma_t \eta_t$ ,  $\ln \sigma_t = \omega_0 + \sum_{i \in \{1,3,5,7\}} \alpha_{0i} \ln \epsilon_{t-i}^2 + \beta_{01} \ln \sigma_{t-1}^2$ ,  $t = 1, 2, \dots, 126$ , where  $\epsilon_t$  is the error term of the dynamic model in Table 3. P-value, the  $p$ -values from two-sided tests.  $AR_{12}(\hat{\eta}_t)$  and  $AR_{13}(\hat{\eta}_t)$ , [Ljung and Box \(1979\)](#) tests for 12th. and 13th. order autocorrelation in  $\hat{\eta}_t$ , respectively.  $ARCH_{12}(\hat{\eta}_t)$  and  $ARCH_{13}(\hat{\eta}_t)$ , [Ljung and Box \(1979\)](#) tests for 12th. and 13th. order autocorrelation in  $\hat{\eta}_t^2$ , respectively.



### Monthly Euro-area prices and inflation

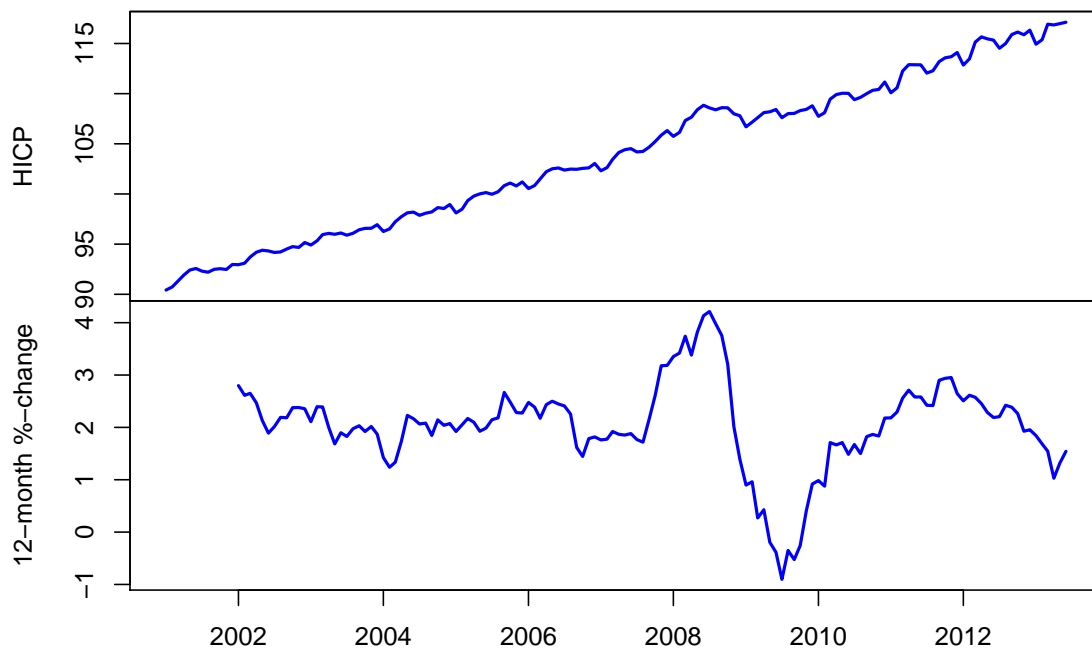


Figure 1: Monthly Euro-area prices (the HICP index) and 12-month inflation from January 2001 to June 2013. Datasource: European Central Bank