Can one really estimate nonstationary GARCH models?

CHRISTIAN FRANCQ
Université Lille 3, EQUIPPE-GREMARS
BP 149, 59653 Villeneuve d'Ascq cedex, France
E-mail: christian.francq@univ-lille3.fr

JEAN-MICHEL ZAKOÏAN
CREST and Université Lille 3,
15 Boulevard Gabriel Péri, 92245 Malakoff Cedex, France
E-mail: zakoian@ensae.fr

July 8, 2008

Abstract

We consider quasi-maximum likelihood estimation (QMLE) of explosive ARCH(1) models. Our results complement those of Jensen and Rahbek (2004a, 2004b) who considered QMLE under the restriction that the intercept ω is fixed. Consistency and asymptotic normality are established for the ARCH coefficient α. Contrary to the restricted QMLE, this estimator is also consistent in the stationary case. Lack of consistency of the QMLE for ω is illustrated via simulation experiments.

KEYWORDS: ARCH, asymptotic normality, inconsistent estimator, nonstationarity, quasi-maximum likelihood estimation, strong consistency.

1 Introduction

Numerous papers have been devoted to the derivation of asymptotic properties of the quasi-maximum likelihood estimator (QMLE) for GARCH models under mild conditions, see for instance Berkes, Horváth and Kokoszka (2003), Francq and Zakoïan (2004), Hall and Yao (2003), and the comprehensive textbook of Straumann (2005). A common assumption in those references is the strict stationarity of the GARCH process.

In two recent papers, Jensen and Rahbek (henceforth JR) (2004a and 2004b) obtained interesting, and somewhat surprising, results concerning the estimation of GARCH models. They showed that one can consistently estimate the coefficients α and β of a GARCH(1,1) model in the non stationary case. This may be found astonishing if one recalls that, without the strict stationary constraint, GARCH processes are explosive.

Based on these findings, many researchers concluded that the stationarity conditions do not matter for the QMLE of GARCH models. The aim of the present paper is to show that this view is, for a large part, erroneous. For our demonstration, it will be sufficient to consider the ARCH(1) case, as is done in JR (2004a).

In Section 2 we study the *restricted* QMLE on which JR's papers are based. In this approach, the intercept ω is fixed and thus is not estimated. We insist on the fact that this estimator is not the standard QMLE. JR showed consistency in the non stationary case. We show that this estimator is in fact non universal, in the sense that it is not consistent in both stationary and nonstationary situations. Section 3 is devoted to the *unrestricted* QMLE. In this approach, which is the one used by the practitioners, the intercept ω is also estimated. By means of analytical arguments and simulation experiments, we will discuss the issue of consistently estimating *the two coefficients* of the ARCH(1) model without stationarity conditions. The conclusion is given in Section 4, and the proofs are collected in Section 5.

2 Asymptotic properties of the restricted QMLE

Consider the ARCH(1) model, as given by

$$\begin{cases} \epsilon_t = \sqrt{h_t}\eta_t, & t = 1, 2, \dots \\ h_t = \omega_0 + \alpha_0 \epsilon_{t-1}^2 \end{cases} \tag{2.1}$$

under classical assumptions on the noise: the sequence (η_t) is assumed independent and identically distributed (iid) with zero mean and unit variance, and such that $\kappa_\eta = E\eta_1^4 < \infty$. In JR (2004a) the parameter $\omega_0 > 0$ is assumed to be known. In JR (2004b) this assumption is relaxed but the parameter ω_0 is fixed to an arbitrary value τ, and only α_0 is estimated in the ARCH(1) case. For any fixed number $\tau > 0$, JR defined a restricted QMLE $\hat{\alpha}_n^c(\tau)$ of α_0 by minimizing the objective function

$$L_n(\alpha) = \frac{1}{n} \sum_{t=1}^{n} \frac{\epsilon_t^2}{\sigma_t^2(\alpha)} + \log \sigma_t^2(\alpha),$$

where $\sigma_t^2(\alpha) = \tau + \alpha\epsilon_{t-1}^2$, and an initial value is introduced for ϵ_0^2 (for instance $\epsilon_0^2 = 0$).

The necessary and sufficient condition for the existence of a strictly stationary solution to (2.1) is $E\log(\alpha_0\eta_1^2) < 0$. Under the assumption

$$\alpha_0 > \exp\left\{-E\log\eta_1^2\right\}, \tag{2.2}$$

$h_t \to \infty$ almost surely as $t \to \infty$, and JR showed that for all $\tau > 0$ there is a fixed open neighborhood U of α_0 such that, with probability tending to one as $n \to \infty$, $L_n(\alpha)$ has a unique minimum point $\hat{\alpha}_n^c(\tau)$ in U. Furthermore

$$\hat{\alpha}_n^c(\tau) \quad \text{is consistent in probability}$$

and asymptotically normal:

$$\sqrt{n}\left(\hat{\alpha}_n^c(\tau) - \alpha_0\right) \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)\alpha_0^2\right\}, \qquad \text{as } n \to \infty. \tag{2.3}$$

Two remarks are in order.

1. The estimator $\hat{\alpha}_n^c(\tau)$ is not the standard QMLE of α_0. This estimator will be studied in the next section.

2. The previous results are not informative about the asymptotic behavior of $\hat{\alpha}_n^c(\tau)$ under the strict stationarity condition

$$\alpha_0 < \exp\left\{-E\log\eta_1^2\right\}. \tag{2.4}$$

Actually the following lemma shows that $\hat{\alpha}_n^c(\tau)$ is in general not consistent under (2.4).

Lemma 2.1 *Let (ϵ_t) be a stationary ARCH(1) with parameters ω_0 and α_0. Assume that $E\epsilon_t^4 < \infty$. Then, if $\tau \neq \omega_0$*

$$\hat{\alpha}_n^c(\tau) \textit{ does not converge in probability to } \alpha_0.$$

In the next section we consider the standard QMLE, which is the commonly used estimator for GARCH models.

3 Asymptotic properties of the unrestricted QMLE

The QMLE of the ARCH(1) model (2.1), is defined as a measurable solution of

$$(\hat{\omega}_n, \hat{\alpha}_n) = \arg\min_{\theta\in\Theta} \frac{1}{n}\sum_{t=1}^{n} \ell_t(\theta), \quad \ell_t(\theta) = \frac{\epsilon_t^2}{\sigma_t^2(\theta)} + \log\sigma_t^2(\theta), \tag{3.1}$$

where $\theta = (\omega, \alpha)$, Θ is a compact subset of $(0, \infty)^2$, and $\sigma_t^2(\theta) = \omega + \alpha\epsilon_{t-1}^2$ for $t = 1, \ldots, n$ (with an initial value for ϵ_0^2). We will use the next result establishing the rate of the almost sure convergence of ϵ_t^2 to infinity under the condition (2.2).

Lemma 3.1 *Let the ARCH(1) defined by (2.1), with initial condition $\epsilon_0^2 \geq 0$. Then, if (2.2) holds,*

$$\frac{1}{h_n} = o(\rho^n) \qquad \text{and} \qquad \frac{1}{\epsilon_n^2} = o(\rho^n)$$

almost surely as $n \to \infty$ for any constant ρ such that

$$1 > \rho > \exp\left\{-E \log \eta_1^2\right\} / \alpha_0. \tag{3.2}$$

This lemma allows to obtain the strong consistency and asymptotic normality of the QMLE of α_0.

Theorem 3.1 *Under the assumptions of Lemma 3.1, and if $\theta_0 = (\omega_0, \alpha_0) \in \Theta$, the QMLE defined in (3.1) satisfies*

$$\hat{\alpha}_n \to \alpha_0 \qquad a.s. \tag{3.3}$$

and, if θ_0 belongs to the interior of Θ,

$$\sqrt{n}\,(\hat{\alpha}_n - \alpha_0) \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)\alpha_0^2\right\}, \qquad as\ n \to \infty. \tag{3.4}$$

It is worth noting from (2.3) and (3.4) that the asymptotic distribution of the restricted and unrestricted QMLE of α_0 coincide in the non-stationary case. However, an important difference is that the restricted estimator is generally not consistent in the stationary case, as proved in Lemma 2.1. On the contrary, the aforementioned references on GARCH estimation have established the consistency and asymptotic normality of the QMLE in the stationary situation.

It is also important to note that these results do not give any insight on the asymptotic behavior of the QMLE of ω_0. A few remarks and numerical illustrations are in order concerning this issue.

In the proof of Theorem 3.1 it is shown that the score vector satisfies

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial}{\partial\theta} \ell_t(\theta_0) \xrightarrow{d} \mathcal{N}\left\{0, J = (\kappa_\eta - 1)\begin{pmatrix} 0 & 0 \\ 0 & \alpha_0^{-1} \end{pmatrix}\right\}. \tag{3.5}$$

The form of the asymptotic covariance matrix J of the score vector shows that, for n sufficiently large and almost surely, the variation of the log-likelihood $n^{-1/2}\sum_{t=1}^{n} \log \ell_t(\theta)$ is negligible when θ varies between (ω_0, α_0) and $(\omega_0 + h, \alpha_0)$ for small h.

Note that a score vector with a degenerate asymptotic variance J can arise when a central limit theorem with a non standard rate of convergence applies. This is for instance the case in regressions with trends, or in unit root and cointegration models. In such situations, the rate of convergence of the QMLE is obtained by finding a diagonal matrix Λ_n such that the asymptotic distribution of $\Lambda_n \sum_{t=1}^n \frac{\partial}{\partial\theta}\ell_t(\theta_0)$ is not degenerated. If, for instance, $\Lambda_n = \mathrm{diag}(n^{-1}, n^{-1/2})$ then the second component of QMLE is expected to converge at the standard rate $\sqrt{n}$, and the first one at the faster rate n. The situation here is completely different. In the proof of Theorem 3.1 it is shown that $\frac{\partial}{\partial\omega}\ell_t(\theta_0) = O_P(\rho^t)$ with $|\rho| < 1$ (see Equation (5.9) below). The equation (3.5) can thus be extended as

$$\Lambda_n \sum_{t=1}^n \frac{\partial}{\partial\theta}\ell_t(\theta_0) \xrightarrow{d} \mathcal{N}\left\{0, (\kappa_\eta - 1)\begin{pmatrix} 0 & 0 \\ 0 & \alpha_0^{-1} \end{pmatrix}\right\}, \quad \Lambda_n = \begin{pmatrix} \lambda_n & 0 \\ 0 & n^{-1/2} \end{pmatrix}$$

for any sequence λ_n tending to zero as $n \to \infty$. It means that the log-likelihood is completely flat in the direction where α_0 is fixed and ω_0 varies. Thus there is little hope concerning the existence of any consistent estimator of ω_0.

This leads to think that the QMLE of ω_0 may be inconsistent without the strict stationarity condition. Figure 3 presents some numerical evidence on the performance of the QMLE in finite samples through a simulation study. In all experiments, we use the sample size $n = 200$ and $n = 4000$ with 100 replications. The data of the top panel are generated from the second-order stationary ARCH(1) model (2.1) with the true parameter $\theta_0 = (1, 0.95)$. The data of the middle panel are generated from the strict stationary ARCH(1) model with $\theta_0 = (1, 1.5)$ and infinite variance. In those two panels the results are very similar, confirming that the second-order stationarity condition is not necessary for the use of the QMLE. The bottom panel, obtained for the explosive ARCH(1) model with $\theta_0 = (1, 4)$, confirms the asymptotic results for the QMLE of α_0. It also illustrates the impossibility to estimate parameter ω_0 with a reasonable accuracy under the nonstationarity condition (2.2). The results even worsen when the sample size increases.

4 Conclusion

The results obtained by JR are very interesting from a theoretical point of view, because they showed that strict stationarity is not compulsory for the estimation of ARCH coefficients. However, there is now a tendency among practitioners, and also theoreticians, to

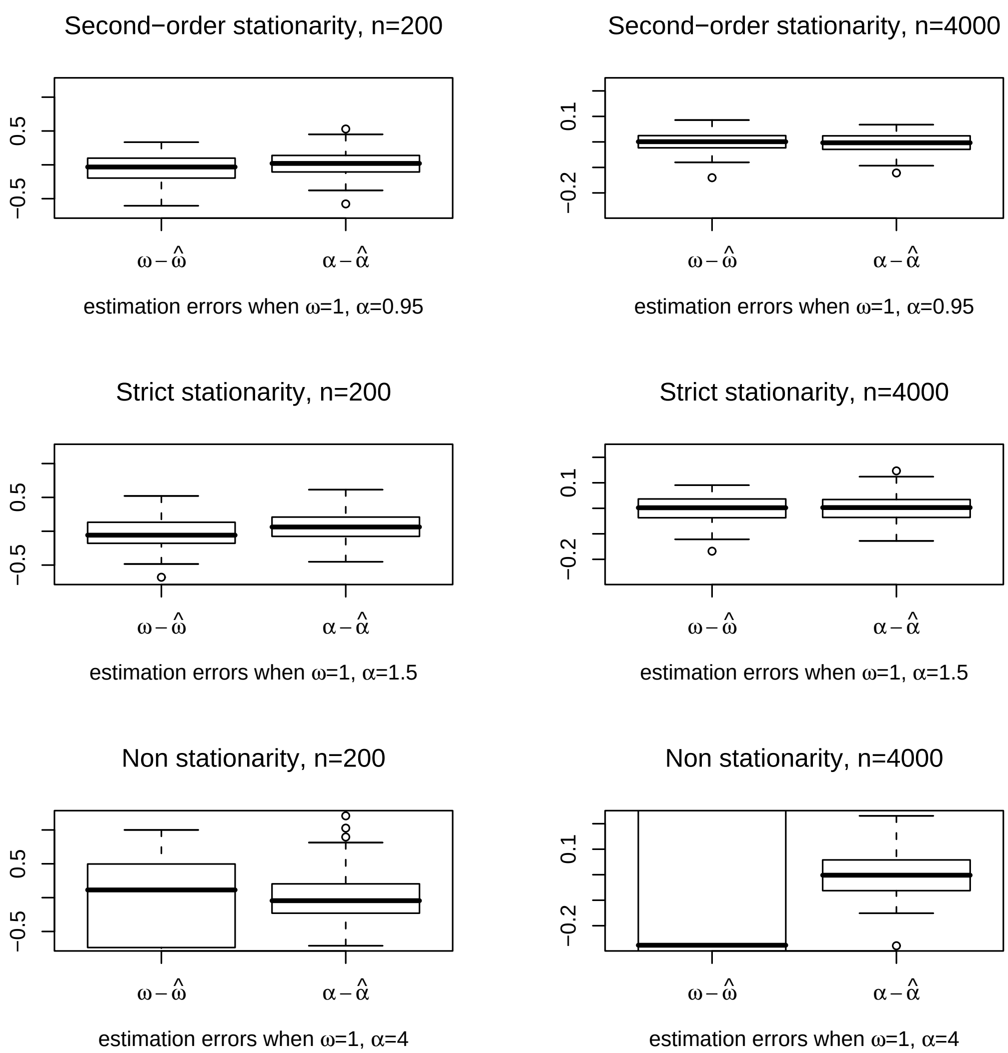

Figure 1: Boxplots of estimation errors for the QMLE of the parameters ω_0 and α_0 of an ARCH(1), with $\eta_t \sim \mathcal{N}(0, 1)$.

believe that the quasi-maximum likelihood estimator for GARCH is consistent and asymptotically normal without any stationarity constraint. In this Note we showed that

i) the restricted QMLE, in which an arbitrary value is attributed to the intercept, is not consistent in the stationary case.

ii) the QMLE of α_0 is indeed strongly consistent and asymptotically normal in the explosive case.

iii) no asymptotic result holds for the parameter ω_0. The flat shape of the log-likelihood along the direction where ω_0 varies, as well as numerical experiments, lead to think that the intercept cannot be estimated without the strict stationarity condition.

To conclude, one can say that the response to the question raised in the title is clearly negative.

5 Proofs

Proof of Lemma 2.1. The ergodic theorem entails that, almost surely,

$$L_n(\alpha) \to L(\alpha) = E\left\{\frac{\omega_0 + \alpha_0\epsilon_{t-1}^2}{\tau + \alpha\epsilon_{t-1}^2} + \log\left(\tau + \alpha\epsilon_{t-1}^2\right)\right\}$$

as $n \to \infty$. The dominated convergence theorem implies that

$$\begin{aligned} L'(\alpha) &= E\frac{\partial}{\partial\alpha}\left\{\frac{\omega_0 + \alpha_0\epsilon_{t-1}^2}{\tau + \alpha\epsilon_{t-1}^2} + \log\left(\tau + \alpha\epsilon_{t-1}^2\right)\right\} \\ &= E\left\{\frac{\epsilon_{t-1}^2}{\left(\tau + \alpha\epsilon_{t-1}^2\right)^2}\{(\tau - \omega_0) + (\alpha - \alpha_0)\epsilon_{t-1}^2\}\right\}. \end{aligned}$$

First suppose that $\tau < \omega_0$. Then $L'(\alpha) < 0$ for $\alpha \le \alpha_0$. The intermediate values theorem shows that the function $L(\cdot)$ has a minimum at a point $\alpha^* > \alpha_0$ and that $L(\alpha^*) < L(\alpha_0)$. Now suppose that $\tau > \omega_0$. Then $L'(\alpha) > 0$ for $\alpha \ge \alpha_0$. This shows that $L(\cdot)$ has a minimum at a point $\alpha^* \in [0, \alpha_0[$, with $L(\alpha^*) < L(\alpha_0)$. Thus, we have shown that for any $\tau \ne \omega_0$, the function $L(\cdot)$ has a minimum at a point $\alpha^* \ne \alpha_0$ and $L(\alpha^*) < L(\alpha_0)$.

A Taylor expansion of $L_n(\cdot)$ yields

$$L_n\left\{\hat{\alpha}_n^c(\tau)\right\} = L_n(\alpha_0) + L_n'(\tilde{\alpha}_n)\{\hat{\alpha}_n^c(\tau) - \alpha_0\} \tag{5.1}$$

where $\tilde{\alpha}_n$ is between $\hat{\alpha}_n^c(\tau)$ and α_0. Note that since $E\epsilon_t^4 < \infty$, almost surely,

$$\limsup_{n\to\infty}\sup_{\alpha}\left|L_n'(\alpha)\right| \le \lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\left(1 + \frac{\epsilon_t^2}{\tau}\right)\frac{\epsilon_{t-1}^2}{\tau} < \infty.$$

Now suppose that

$$\hat{\alpha}_n^c(\tau) \to \alpha_0, \quad \text{in probability as } n \to \infty. \tag{5.2}$$

Then, it follows from (5.1) that

$$L_n\left\{\hat{\alpha}_n^c(\tau)\right\} \to L(\alpha_0), \quad \text{in probability as } n \to \infty.$$

Then, taking the limit in probability in the following inequality

$$L_n\{\hat{\alpha}_n^c(\tau)\} \leq L_n(\alpha^*)$$

we find that

$$L(\alpha_0) \leq L(\alpha^*),$$

which is in contradiction with the definition of $\alpha^* \neq \alpha_0$. Thus (5.2) cannot be true.

□

Proof of Lemma 3.1. We have

$$\begin{aligned}
\rho^n h_n &= \rho^n \omega_0 \left\{1 + \sum_{t=1}^{n-1} \alpha_0^t \eta_{n-1}^2 \ldots \eta_{n-t}^2\right\} + \rho^n \alpha_0^n \eta_{n-1}^2 \ldots \eta_1^2 \epsilon_0^2 \\
&\geq \rho^n \omega_0 \prod_{t=1}^{n-1} \alpha_0 \eta_t^2.
\end{aligned} \tag{5.3}$$

Thus

$$\begin{aligned}
\liminf_{n\to\infty} \frac{1}{n} \log \rho^n h_n &\geq \lim_{n\to\infty} \frac{1}{n} \left\{\log \rho\omega_0 + \sum_{t=1}^{n-1} \log \rho\alpha_0 \eta_t^2\right\} \\
&= E \log \rho\alpha_0 \eta_1^2 > 0,
\end{aligned}$$

using (3.2) for the last inequality. It follows that $\log \rho^n h_n$, and hence $\rho^n h_n$, tends to $+\infty$ almost surely as $n \to \infty$. For any real-valued function f, let $f^+(x) = \max\{f(x), 0\}$ and $f^-(x) = \max\{-f(x), 0\}$, so that $f(x) = f^+(x) - f^-(x)$. Since $E \log^+ \eta_1^2 \leq E\eta_1^2 = 1$, we have $E|\log \eta_1^2| = \infty$ if and only if $E \log \eta_1^2 = -\infty$. Thus (2.2) implies $E|\log \eta_1^2| < \infty$, which entails that $\log \eta_n^2/n \to 0$ almost surely as $n \to \infty$. Therefore, using (5.3), $\liminf_{n\to\infty} n^{-1} \log \rho^n \eta_n^2 h_n \geq E \log \rho\alpha_0\eta_1^2 > 0$, and $\rho^n \epsilon_n^2 = \rho^n \eta_n^2 h_n \to +\infty$ almost surely by already given arguments.

□

Proof of (3.3). Note that $(\hat{\omega}_n, \hat{\alpha}_n) = \arg\min_{\theta\in\Theta} Q_n(\theta)$, where

$$Q_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}\left\{\ell_t(\theta) - \ell_t(\theta_0)\right\}.$$

We have

$$\begin{aligned} Q_n(\theta) &= \frac{1}{n}\sum_{t=1}^{n}\eta_t^2\left\{\frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} - 1\right\} + \log\frac{\sigma_t^2(\theta)}{\sigma_t^2(\theta_0)} \\ &= \frac{1}{n}\sum_{t=1}^{n}\eta_t^2\frac{(\omega_0-\omega)+(\alpha_0-\alpha)\epsilon_{t-1}^2}{\omega+\alpha\epsilon_{t-1}^2} + \log\frac{\omega+\alpha\epsilon_{t-1}^2}{\omega_0+\alpha_0\epsilon_{t-1}^2}. \end{aligned}$$

For any $\theta\in\Theta$, we have $\alpha\neq 0$. Letting

$$O_n(\alpha) = \frac{1}{n}\sum_{t=1}^{n}\eta_t^2\frac{(\alpha_0-\alpha)}{\alpha} + \log\frac{\alpha}{\alpha_0}$$

and

$$d_t = \frac{\alpha(\omega_0-\omega)-\omega(\alpha_0-\alpha)}{\alpha(\omega+\alpha\epsilon_{t-1}^2)},$$

we have

$$Q_n(\theta) - O_n(\alpha) = \frac{1}{n}\sum_{t=1}^{n}\eta_t^2 d_{t-1} + \frac{1}{n}\sum_{t=1}^{n}\log\frac{(\omega+\alpha\epsilon_{t-1}^2)\alpha_0}{(\omega_0+\alpha_0\epsilon_{t-1}^2)\alpha} \to 0 \quad \text{a.s.}$$

since, by Lemma 3.1, $\epsilon_t^2 \to \infty$ almost surely as $t\to\infty$. Moreover this convergence is uniform on the compact set Θ:

$$\lim_{n\to\infty}\sup_{\theta\in\Theta}|Q_n(\theta) - O_n(\alpha)| = 0 \quad \text{a.s.} \tag{5.4}$$

Let α_0^- and α_0^+ denote two constants such that $0 < \alpha_0^- < \alpha_0 < \alpha_0^+$. Introducing $\hat{\sigma}_\eta^2 = n^{-1}\sum_{t=1}^{n}\eta_t^2$, the solution of

$$\alpha_n^* = \arg\min_{\alpha} O_n(\alpha)$$

is $\alpha_n^* = \alpha_0\hat{\sigma}_\eta^2$. This solution belongs to the interval (α_0^-, α_0^+) for sufficiently large n. Thus

$$\alpha_n^{**} = \arg\min_{\alpha\notin(\alpha_0^-,\alpha_0^+)} O_n(\alpha) \in \{\alpha_0^-, \alpha_0^+\}$$

and

$$\lim_{n\to\infty} O_n(\alpha_n^{**}) = \min\left\{\lim_{n\to\infty} O_n(\alpha_0^-), \lim_{n\to\infty} O_n(\alpha_0^+)\right\} > 0.$$

This result and (5.4) show that almost surely

$$\lim_{n\to\infty}\min_{\theta\in\Theta,\,\alpha\notin(\alpha_0^-,\alpha_0^+)} Q_n(\theta) > 0.$$

Since $\min_\theta Q_n(\theta) \le Q_n(\theta_0) = 0$, it follows that

$$\lim_{n\to\infty} \arg\min_{\theta\in\Theta} Q_n(\theta) \in (0,\infty) \times (\alpha_0^-, \alpha_0^+).$$

Because the interval (α_0^-, α_0^+) containing α_0 can be chosen arbitrarily small, we get the convergence in (3.3).

$\square$

The following result will be used to establish the asymptotic normality of the QMLE of α_0.

Lemma 5.1 *Under the assumptions of Theorem 3.1, we have*

$$\begin{aligned}
\sum_{t=1}^{\infty} \sup_{\theta\in\Theta} \left| \frac{\partial}{\partial\omega} \ell_t(\theta) \right| &< \infty \quad a.s., && (5.5)\\
\sum_{t=1}^{\infty} \sup_{\theta\in\Theta} \left\| \frac{\partial^2}{\partial\omega\partial\theta} \ell_t(\theta) \right\| &< \infty \quad a.s., && (5.6)\\
\sup_{\theta\in\Theta} \left| \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial\alpha^2} \ell_t(\omega,\alpha_0) - \frac{1}{\alpha_0^2} \right| &= o(1) \quad a.s., && (5.7)\\
\frac{1}{n} \sum_{t=1}^{n} \sup_{\theta\in\Theta} \left| \frac{\partial^3}{\partial\alpha^3} \ell_t(\theta) \right| &= O(1) \quad a.s., && (5.8)
\end{aligned}$$

Proof. Using Lemma 3.1, there exist a real random variable K and a constant $\rho \in (0,1)$ independent of θ and t such that

$$\begin{aligned}
\left| \frac{\partial}{\partial\omega} \ell_t(\theta) \right| &= \left| \frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial\omega} \left(1 - \frac{\epsilon_t^2}{\sigma_t^2(\theta)} \right) \right| \\
&= \left| \frac{-(\omega_0 + \alpha_0\epsilon_{t-1}^2)\eta_t^2}{(\omega + \alpha\epsilon_{t-1}^2)^2} + \frac{1}{\omega + \alpha\epsilon_{t-1}^2} \right| \le K\rho^t(\eta_t^2 + 1). \qquad (5.9)
\end{aligned}$$

Since $\sum_{t=1}^{\infty} K\rho^t(\eta_t^2+1)$ has a finite expectation, it is almost surely finite. Thus (5.5) is proved, and (5.6) can be obtained by the same arguments. We have

$$\begin{aligned}
\frac{\partial^2 \ell_t(\omega,\alpha_0)}{\partial\alpha^2} - \frac{1}{\alpha_0^2} &= \left\{ 2\frac{(\omega_0 + \alpha_0\epsilon_{t-1}^2)\eta_t^2}{\omega + \alpha_0\epsilon_{t-1}^2} - 1 \right\} \frac{\epsilon_{t-1}^4}{(\omega + \alpha_0\epsilon_{t-1}^2)^2} - \frac{1}{\alpha_0^2} \\
&= \left(2\eta_t^2 - 1\right) \frac{\epsilon_{t-1}^4}{(\omega + \alpha_0\epsilon_{t-1}^2)^2} - \frac{1}{\alpha_0^2} + r_{1,t} \\
&= 2\left(\eta_t^2 - 1\right) \frac{1}{\alpha_0^2} + r_{1,t} + r_{2,t}
\end{aligned}$$

where

$$\sup_{\theta\in\Theta} |r_{1,t}| = \sup_{\theta\in\Theta} \left| \frac{2(\omega_0 - \omega)\eta_t^2}{(\omega + \alpha_0\epsilon_{t-1}^2)} \frac{\epsilon_{t-1}^4}{(\omega + \alpha_0\epsilon_{t-1}^2)^2} \right| = o(1) \quad \text{a.s.}$$

and

$$\begin{aligned}
\sup_{\theta\in\Theta}|r_{2,t}| &= \sup_{\theta\in\Theta}\left|(2\eta_t^2-1)\left\{\frac{\epsilon_{t-1}^4}{(\omega+\alpha_0\epsilon_{t-1}^2)^2}-\frac{1}{\alpha_0^2}\right\}\right| \\
&= \sup_{\theta\in\Theta}\left|(2\eta_t^2-1)\left\{\frac{\omega^2+2\alpha_0\epsilon_{t-1}^2}{\alpha_0^2(\omega+\alpha_0\epsilon_{t-1}^2)^2}\right\}\right| = o(1) \quad \text{a.s.}
\end{aligned}$$

as $t\to\infty$. Therefore (5.7) is established. To prove (5.8), it suffices to remark that

$$\begin{aligned}
\left|\frac{\partial^3}{\partial\alpha^3}\ell_t(\theta)\right| &= \left|\left\{2-6\frac{(\omega_0+\alpha_0\epsilon_{t-1}^2)\eta_t^2}{\omega+\alpha\epsilon_{t-1}^2}\right\}\left(\frac{\epsilon_{t-1}^2}{\omega+\alpha\epsilon_{t-1}^2}\right)^3\right| \\
&\leq \left\{2+6\left(\frac{\omega_0}{\omega}+\frac{\alpha_0}{\alpha}\right)\eta_t^2\right\}\frac{1}{\alpha^3}.
\end{aligned}$$

□

Proof of (3.4). Notice that we cannot use that fact that the derivative of the criterion cancels at $\hat\theta_n=(\hat\omega_n,\hat\alpha_n)$ since we only have the convergence of $\hat\alpha_n$ to α_0. Thus the minimum could lie on the boundary of Θ, even asymptotically. However, the partial derivative with respect to α is asymptotically equal to zero at the minimum since $\hat\alpha_n\to\alpha_0$ and $(\omega_0,\alpha_0)\in\overset{\circ}{\Theta}$. Hence, an expansion of the criterion derivative gives

$$\begin{pmatrix} \frac{1}{\sqrt{n}}\sum_{t=1}^n\frac{\partial}{\partial\omega}\ell_t(\hat\theta_n) \\ 0 \end{pmatrix} = \frac{1}{\sqrt{n}}\sum_{t=1}^n\frac{\partial}{\partial\theta}\ell_t(\theta_0)+J_n\sqrt{n}(\hat\theta_n-\theta_0) \tag{5.10}$$

where J_n is a 2×2 matrix whose elements have the form

$$J_n(i,j)=\frac{1}{n}\sum_{t=1}^n\frac{\partial^2}{\partial\theta_j\partial\theta_j}\ell_t(\theta_{i,j}^*)$$

where $\theta_{i,j}^*=(\omega_{i,j}^*,\alpha_{i,j}^*)$ is between $\hat\theta_n$ and θ_0. By Lemma 3.1 and from the Lindeberg central limit theorem for martingale differences we have

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{t=1}^n\frac{\partial}{\partial\alpha}\ell_t(\theta_0) &= \frac{1}{\sqrt{n}}\sum_{t=1}^n(1-\eta_t^2)\frac{\epsilon_{t-1}^2}{\omega_0+\alpha_0\epsilon_{t-1}^2} \\
&= \frac{1}{\sqrt{n}}\sum_{t=1}^n(1-\eta_t^2)\frac{1}{\alpha_0}+o_P(1) \\
&\xrightarrow{d} \mathcal{N}\left(0,\frac{\kappa_\eta-1}{\alpha_0^2}\right).
\end{aligned} \tag{5.11}$$

By (5.6), in Lemma 5.1, and the compactness of Θ we have

$$J_n(2,1)\sqrt{n}(\hat\omega_n-\omega_0)\leq\sum_{t=1}^\infty\sup_{\theta\in\Theta}\left\|\frac{\partial^2}{\partial\omega\partial\theta}\ell_t(\theta)\right\|\frac{1}{\sqrt{n}}(\hat\omega_n-\omega_0)\to 0 \quad \text{a.s.} \tag{5.12}$$

An expansion of the function

$$\alpha \mapsto \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial \alpha^2} \ell_t(\omega_{2,2}^*, \alpha)$$

gives

$$J_n(2,2) = \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial \alpha^2} \ell_t(\omega_{2,2}^*, \alpha_0) + \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^3}{\partial \alpha^3} \ell_t(\omega_{2,2}^*, \alpha^*)(\alpha_{2,2}^* - \alpha_0)$$

where α^* is between $\alpha_{2,2}^*$ and α_0. Using (5.7), (5.8) and (3.3) we get

$$J_n(2,2) \to \frac{1}{\alpha_0^2} \quad \text{a.s.} \tag{5.13}$$

The conclusion follows, by considering the second component in (5.10) and from (5.11), (5.12) and (5.13).

$\square$

References

Berkes I., Horváth L. and P.S. Kokoszka (2003) GARCH processes: structure and estimation. *Bernoulli* 9, 201–227.

Billingsley P. (1995) *Probability and Measure*. John Wiley, New York.

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

Hall P. and Q. Yao (2003) Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica* 71, 285–317.

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.

Nelson D.B. (1990) Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory* 6, 318–334.

Straumann, D. (2005) *Estimation in conditionally heteroscedastic time series models.* Lecture Notes in Statistics, Springer Berlin Heidelberg.