

Properties and Estimation of GARCH(1,1) Model

Petra Posedel¹

Abstract

We study in depth the properties of the GARCH(1,1) model and the assumptions on the parameter space under which the process is stationary. In particular, we prove ergodicity and strong stationarity for the conditional variance (squared volatility) of the process. We show under which conditions higher order moments of the GARCH(1,1) process exist and conclude that GARCH processes are heavy-tailed. We investigate the sampling behavior of the quasi-maximum likelihood estimator of the Gaussian GARCH(1,1) model. A bounded conditional fourth moment of the rescaled variable (the ratio of the disturbance to the conditional standard deviation) is sufficient for the result. Consistent estimation and asymptotic normality are demonstrated, as well as consistent estimation of the asymptotic covariance matrix.

1 Introduction

Financial markets react nervously to political disorders, economic crises, wars or natural disasters. In such *stress periods* prices of financial assets tend to fluctuate very much. Statistically speaking, it means that the conditional variance for the given past

$$\text{Var}(X_t | X_{t-1}, X_{t-2}, \dots)$$

is not constant over time and the process X_t is conditionally heteroskedastic. Econometricians usually say that *volatility*

$$\sigma_t = \sqrt{\text{Var}(X_t | X_{t-1}, X_{t-2}, \dots)}$$

changes over time. Understanding the nature of such time dependence is very important for many macroeconomic and financial applications, e.g. irreversible investments, option pricing, asset pricing etc. Models of conditional heteroskedasticity for time series have a very important role in today's financial risk management and its attempts to make financial decisions on the basis of the observed price asset data P_t in discrete time. Prices P_t are believed to be nonstationary so they are usually transformed in the so-called *log returns*

$$X_t = \log P_t - \log P_{t-1}.$$

Log returns are supposed to be stationary, at least in periods of time that are not too long. Very often in the past it was suggested that (X_t) represents a sequence of independent identically distributed random variable, in other words, that log returns evolve

¹Faculty of Economics, University of Zagreb, Zagreb, Croatia

like a random walk. Samuelson suggested modelling speculative prices in the continuous time with the geometric Brownian motion. Discretization of that model leads to a random walk with independent identically distributed Gaussian increments of log return prices in discrete time. This hypothesis was rejected in the early sixties. Empirical studies based on the log return time series data of some US stocks showed the following observations, the so-called stylized facts of financial data:

- serial dependence are present in the data
- volatility changes over time
- distribution of the data is heavy-tailed, asymmetric and therefore not Gaussian.

These observations clearly show that a random walk with Gaussian increments is not a very realistic model for financial data. It took some time before R. Engle found a discrete model that described very well the previously mentioned stylized facts of financial data, but it was also relatively simple and stationary so the inference was possible. Engle called his model *autoregressive conditionally heteroskedastic*- ARCH, because the conditional variance (squared volatility) is not constant over time and shows autoregressive structure. Some years later, T. Bollerslev generalized the model by introducing *generalized autoregressive conditionally heteroskedastic* - GARCH model. The properties of GARCH models are not easy to determine.

2 GARCH(1,1) process

Definition 2.1 Let (Z_n) be a sequence of i.i.d. random variables such that $Z_t \sim N(0, 1)$. (X_t) is called the **generalized autoregressive conditionally heteroskedastic** or GARCH(q, p) process if

$$X_t = \sigma_t Z_t, \quad t \in \mathbb{Z} \quad (2.1)$$

where (σ_t) is a nonnegative process such that

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_q X_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2, \quad t \in \mathbb{Z} \quad (2.2)$$

and

$$\alpha_0 > 0, \quad \alpha_i \geq 0 \quad i = 1, \dots, q \quad \beta_i \geq 0 \quad i = 1, \dots, p. \quad (2.3)$$

The conditions on parameters ensure strong positivity of the conditional variance (2.2).

If we write the equation (2.2) in terms of the *lag-operator* B we get

$$\sigma_t^2 = \alpha_0 + \alpha(B) X_t^2 + \beta(B) \sigma_t^2, \quad (2.4)$$

where

$$\alpha(B) = \alpha_1 B + \alpha_2 B^2 + \dots + \alpha_q B^q$$

and

$$\beta(B) = \beta_1 B + \beta_2 B^2 + \dots + \beta_p B^p. \quad (2.5)$$

If the roots of the characteristic equation, i.e.

$$1 - \beta_1x - \beta_2x^2 - \dots - \beta_px^p = 0$$

lie outside the unit circle and the process (X_t) is stationary, then we can write (2.2) as

$$\begin{aligned} \sigma_t^2 &= \frac{\alpha_0}{1 - \beta(1)} + \frac{\alpha(B)}{1 - \beta(B)} X_t^2 \\ &= \alpha_0^* + \sum_{i=1}^{\infty} \delta_i X_{t-i}^2 \end{aligned} \tag{2.6}$$

where $\alpha_0^* = \frac{\alpha_0}{1 - \beta(1)}$, and δ_i are coefficients of B^i in the expansion of $\alpha(B)[1 - \beta(B)]^{-1}$. Note that the expression (2.6) tells us that the GARCH(q,p) process is an ARCH process of infinite order with a fractional structure of the coefficients.

From (2.1) it is obvious that the GARCH(1,1) process is stationary if the process (σ_t^2) is stationary. So if we want to study the properties and higher order moments of GARCH(1,1) process it is enough to do so for the process (σ_t^2) .

The following theorem gives us the main result for stochastic difference equations that we are going to use in order to establish the stationarity of the process (σ_t^2) .

Theorem 2.2 *Let (Y_t) be the stochastic process defined by*

$$Y_t = A_t + B_t Y_{t-1}, \quad t \in \mathbb{N}, \tag{2.7}$$

or explicitly

$$Y_t = Y_0 \prod_{j=1}^t B_j + \sum_{m=1}^t A_m \prod_{j=m+1}^t B_j, \quad t \in \mathbb{N}. \tag{2.8}$$

Suppose that Y_0 is independent of the i.i.d. sequence $((A_t, B_t))_{t \geq 1}$. Assume that

$$E \ln^+ |A| < \infty \quad \text{and} \quad -\infty \leq E \ln |B| < 0. \tag{2.9}$$

Then

(a) $Y_t \xrightarrow{\mathcal{D}} Y$ for some random variable Y such that it satisfies the identity in law

$$Y = A + BY, \tag{2.10}$$

where Y and (A, B) are independent.

(b) Equation (2.10) has a solution, unique in distribution, which is given by

$$Y \stackrel{\mathcal{D}}{=} \sum_{m=1}^{\infty} A_m \prod_{j=1}^{m-1} B_j. \tag{2.11}$$

The right hand side of (2.11) converges absolutely with probability 1.

(c) If we choose $Y_0 \stackrel{\mathcal{D}}{=} Y$ as in (2.11), then the process $(Y_t)_{t \geq 0}$ is strictly stationary.

Now assume the moment conditions

$$E|A|^p < \infty \quad \text{and} \quad E|B|^p < 1 \quad \text{for some } p \in [1, \infty).$$

(d) Then $E|Y|^p < \infty$, and the series in (2.11) converges in p th mean.

(e) If $E|Y_0|^p < \infty$, then (Y_t) converges to Y in p th mean, and in particular

$$E|Y_t|^p \rightarrow E|Y|^p \quad \text{as } t \rightarrow \infty.$$

(f) The moments EY^m are uniquely determined by the equations

$$EY^m = \sum_{k=0}^m \binom{m}{k} E(B^k A^{m-k}) EY^k, \quad m = 1, \dots, \lfloor p \rfloor \quad (2.12)$$

where $\lfloor p \rfloor$ denotes the floor function.

In the next theorem we present the stationarity of the conditional variance process (σ_t^2) .

Theorem 2.3 Let (σ_t^2) be the conditional variance of GARCH(1,1) process defined with (2.1) and (2.2). Additionally, assume that

$$E[\ln(\alpha_1 Z_0^2 + \beta_1)] < 0 \quad (2.13)$$

and that σ_0^2 is independent from (Z_t) . Then it holds

(a) the process (σ_t^2) is strictly stationary if

$$\sigma_0^2 \stackrel{\mathcal{D}}{=} \alpha_0 \sum_{m=1}^{\infty} \prod_{j=1}^{m-1} (\beta_1 + \alpha_1 Z_{j-1}^2) \quad (2.14)$$

and the series (2.14) converges absolutely with probability 1.

(b) Assume that (σ_t^2) is strictly stationary and let $\sigma = \sigma_0^2$, $Z = Z_1$. Let $E(\beta_1 + \alpha_1 Z^2)^p < 1$ for some $p \in [1, \infty)$. Then $E(\sigma^2)^m < \infty$ for some $1 \leq m \leq \lfloor p \rfloor$. For such integer m it holds

$$\begin{aligned} E[\sigma^{2m}] &= [1 - E(\beta_1 + \alpha_1 Z^2)^m]^{-1} \sum_{k=0}^{m-1} \binom{m}{k} E(\alpha_1 Z^2 + \beta_1)^k \alpha_0^{m-k} \times \\ &\times E[\sigma^{2k}] < \infty. \end{aligned} \quad (2.15)$$

Proof: From (2.2) we have

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

or

$$\sigma_t^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta_1) \sigma_{t-1}^2$$

that represents a stochastic difference equation

$$Y_t = A_t + B_t Y_{t-1},$$

where $Y_t = \sigma_t^2$, $A_t = \alpha_0$ and $B_t = \alpha_1 Z_{t-1}^2 + \beta_1$. From the assumptions of the theorem we have that $E \ln^+ |A| < \infty$ and $E \ln |B| = E [\ln (\beta_1 + \alpha_1 Z_{t-1}^2)] < 0$. So, from Theorem 2.2 we have that (σ_t^2) is strictly stationary with unique marginal distribution given by (2.14) and this shows the first statement of the theorem. Additionally, suppose that $E(\beta_1 + \alpha_1 Z^2)^p < 1$. In that case we have $E|B|^p = E(\beta_1 + \alpha_1 Z^2)^p < 1$ for some $p \in [1, \infty)$ so from part (f) of Theorem 2.2 it follows (2.15). ■

Example 2.4 Let (X_t) be GARCH(1,1) process. Let

$$\mu(\alpha_1, \beta_1, p) = E(\alpha_1 Z^2 + \beta_1)^p, \quad p \in [1, \infty).$$

In that case, it follows from Theorem 2.3 that a necessary condition for the existence of the stationary moment of order $2m$, $1 \leq m \leq p$, of a GARCH(1,1) process is given by

$$\mu(\alpha_1, \beta_1, p) < 1.$$

In the special case of $m = 2$ it follows that the stationary fourth moment of the GARCH(1,1) process exists if

$$\mu(\alpha_1, \beta_1, 2) = \sum_{j=0}^2 \binom{2}{j} \alpha_1^j \beta_1^{2-j} < 1,$$

that is equivalent to

$$\beta_1^2 + 2\alpha_1\beta_1 + 3\alpha_1^2 < 1.$$

From the recursive formula given in the Theorem 2.2 in the case of $m = 1$ and $m = 2$ we obtain

$$E(X_t^2) = E(Z_t^2) \cdot E(\sigma_t^2) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$$

and

$$\begin{aligned} E(X_t^4) &= E(Z_t^4) \cdot E(\sigma_t^4) \\ &= 3[\alpha_0^2 + 2E(X_t^2)\alpha_0(\alpha_1 + \beta_1)] \cdot [1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2]^{-1} \\ &= 3\left[\alpha_0^2 + 2\frac{\alpha_0^2}{1 - \alpha_1 - \beta_1}(\alpha_1 + \beta_1)\right] \cdot [1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2]^{-1} \\ &= 3\alpha_0^2\left[1 + 2\frac{\alpha_1 + \beta_1}{1 - \alpha_1 - \beta_1}\right] \cdot [1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2]^{-1} \\ &= 3\alpha_0^2(1 + \alpha_1 + \beta_1)[(1 - \alpha_1 - \beta_1)(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)]^{-1}. \end{aligned}$$

Since the marginal kurtosis is given by

$$k = \frac{E(X_t^4)}{[E(X_t^2)]^2},$$

from the previous calculus it immediately follows that

$$k = \frac{3(1 + \alpha_1 + \beta_1)(1 - \alpha_1 - \beta_1)}{1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2}.$$

A little calculus shows

$$\begin{aligned} 3\text{Var}(\sigma_t^2) &= E(X_t^4) - 3[E(X_t^2)]^2 \\ &= \frac{3\alpha_0^2(1 + \alpha_1 + \beta_1)}{(1 - \alpha_1 - \beta_1)(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)} - 3\left[\frac{\alpha_0}{1 - \alpha_1 - \beta_1}\right]^2 \\ &= \frac{3\alpha_0^2}{(1 - \alpha_1 - \beta_1)^2} \cdot \frac{2\alpha_1^2}{(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)}. \end{aligned} \quad (2.16)$$

Since from the assumptions we have that $\alpha_0 > 0$, $1 - \alpha_1 - \beta_1 > 0$ and $1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2 < 1$, it follows that all the factors in (2.16) are positive so we conclude that the GARCH(1,1) process has the so-called *leptokurtic distribution*.

3 Estimation of the GARCH(1,1) model

Although in this section we assume that (Z_t) are i.i.d. sequence of random variables, the results we shall present can also be shown for the (Z_t) strictly stationary and ergodic sequence of random variables. In that case, the assumptions for the process (Z_t) are little modified but the main part of the calculus we present here also holds for not such strong assumptions.

3.1 Description of the model and the quasi-likelihood function

Suppose we observe the sequence (Y_t) such that

$$Y_t = C_0 + \varepsilon_{0t}, \quad t = 1, \dots, n,$$

where we assume that (ε_{0t}) is GARCH(1,1) process, exactly

$$\varepsilon_{0t} = Z_t\sigma_{0t}, \quad \mathcal{F}_t = \sigma(\{\varepsilon_{0s}, s \leq t\}),$$

where (Z_t) is a sequence of i.i.d. random variables and

$$\sigma_{0t}^2 = \omega_0(1 - \beta_0) + \alpha_0\varepsilon_{0t-1}^2 + \beta_0\sigma_{0t-1}^2 \quad \text{a.s.} \quad (3.1)$$

From Theorem 2.2 we have that the strict stationary solution of (3.1) is given by

$$\sigma_{0t}^2 = \omega_0 + \alpha_0 \sum_{k=0}^{\infty} \beta_0^k \varepsilon_{0t-1-k}^2 \quad \text{a.s.}$$

if it holds $E[\ln(\beta_0 + \alpha_0 Z^2)] < 0$. The process is described with the vector of parameters

$$\theta_0 = (C_0, \omega_0, \alpha_0, \beta_0).$$

The model for the unknown parameters $\theta = (C, \omega, \alpha, \beta)'$ is given by

$$Y_t = C + \varepsilon_t, \quad t = 1, \dots, n,$$

and

$$\sigma_t^2(\theta) = \omega(1 - \beta) + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2(\theta), \quad t = 2, \dots, n$$

and with the initial condition $\sigma_1^2(\theta) = \omega$. With that kind of notation we have the following expression for the process of conditional variance:

$$\sigma_t^2 = \omega + \alpha \sum_{k=0}^{t-2} \beta^k \varepsilon_{t-1-k}^2.$$

Let us define the compact space

$$\begin{aligned} \Theta &= \{ \theta : C_l \leq C \leq C_d, 0 < \omega_l \leq \omega \leq \omega_d, 0 < \alpha_l \leq \alpha \leq \alpha_d, \\ &\quad 0 < \beta_l \leq \beta \leq \beta_d < 1 \} \\ &\subseteq \{ \theta : E[\ln(\beta + \alpha Z^2)] < 0 \}. \end{aligned}$$

Additionally, assume that $\theta_0 \in \Theta$ so it immediately follows that $\alpha_0 > 0$ and $\beta_0 > 0$. Inference for GARCH(1,1) process usually assumes that (Z_t) are i.i.d. random variables such that $Z_t \sim N(0, 1)$ so the likelihood function is easy to determine. Assuming that the likelihood function is Gaussian, the log-likelihood function is of the form (ignoring constants)

$$L_T(\theta) = \frac{1}{2T} \sum_{t=1}^T l_t(\theta), \quad \text{where } l_t(\theta) = - \left(\ln \sigma_t^2(\theta) + \frac{\varepsilon_t^2}{\sigma_t^2(\theta)} \right).$$

Since the likelihood function does not need to be Gaussian, in other words, the process (Z_t) does not need to be the Gaussian white noise, L_T is called the **quasi-likelihood function**.

3.2 Consistency of the quasi-maximum likelihood estimator

Although a finite data set is available in practice, this is not enough to determine good properties of an estimator. We shall see in this section how useful results can be obtained taking into consideration the strictly stationary model for the conditional variance that we have previously defined. We shall note it in the following way

$$\sigma_{ut}^2(\theta) = \omega + \alpha \sum_{k=0}^{\infty} \beta^k \varepsilon_{t-1-k}^2, \quad \varepsilon_t = Y_t - C,$$

to avoid confusion with the original conditional variance process (σ_t^2) . In that case the quasi-likelihood function is given by

$$L_{uT}(\theta) = \frac{1}{2T} \sum_{t=1}^T l_{ut}(\theta), \quad \text{where } l_{ut}(\theta) = - \left(\ln \sigma_{ut}^2(\theta) + \frac{\varepsilon_t^2}{\sigma_{ut}^2(\theta)} \right).$$

Additionally, we are going to show that the stationary and the non-stationary model are not "far away" in some sense. So, all the calculus is done using the stationary model and then connecting the two models.

Let us define

$$\sigma_{\varepsilon t}^2(\theta) = \omega + \alpha \sum_{k=0}^{\infty} \beta^k \varepsilon_{0t-1-k}^2.$$

The process (σ_{ut}^2) is a strictly stationary model of the conditional variance which assumes an infinite history of the observed data. The process $(\sigma_{\varepsilon t}^2)$ is in fact identical to the process (σ_{ut}^2) except that it is expressed as a function of the true innovations (ε_{0t}) instead of the residuals (ε_t) .

We suppose that the following conditions on the process (Z_t) hold:

- (1) (Z_t) is a sequence of i.i.d. random variables such that $EZ_t = 0$;
- (2) Z_t^2 is nondegenerate;
- (3) for some $\delta > 0$ exists $S_\delta < \infty$ such that $E[Z_t^{2+\delta}] \leq S_\delta < \infty$;
- (4) $E[\ln(\beta_0 + \alpha_0 Z_t^2)] < 0$;
- (5) θ_0 is in the interior of Θ ;
- (6) if for some t holds

$$\sigma_{0t}^2 = c_0 + \sum_{k=1}^{\infty} c_k \varepsilon_{t-k}^2 \quad \text{i} \quad \sigma_{0t}^2 = c_0^* + \sum_{k=1}^{\infty} c_k^* \varepsilon_{t-k}^2$$

then $c_i = c_i^*$ for every $1 \leq i < \infty$.

We call the conditions (1) – (6) **elementary conditions**.

The proof for the following result for the case of the general GARCH(q, p) process can be found in [5].

Proposition 3.3 *If the elementary conditions hold, there are not two different vectors $(\omega, \alpha, \beta, C)$ and $(\omega^*, \alpha^*, \beta^*, C^*)$ such that*

$$\sigma_{0t}^2 = \omega^* + \alpha^*(Y_{t-1} - C^*)^2 + \beta^* \sigma_{0t-1}^2$$

and

$$\sigma_{0t}^2 = \omega + \alpha(Y_{t-1} - C)^2 + \beta \sigma_{0t-1}^2.$$

The following lemma would be very helpful for the results we shall provide. The proof can be found in [10].

Lemma 3.4 *Uniformly on θ*

$$B^{-1} \sigma_{\varepsilon t}^2(\theta) \leq \sigma_{ut}^2(\theta) \leq B \sigma_{\varepsilon t}^2(\theta) \quad \text{a.s.}$$

where

$$B = 1 + 2(1 - \beta_d)^{-\frac{1}{2}}(C_d - C_l) \times \max\left(\frac{\alpha_d}{\omega_l}, 1\right) + \frac{\alpha_d}{\omega_l(1 - \beta_d)}(C_d - C_l)^2.$$

Although we are not going to discuss the rational moments of the process (σ_{0t}^2) , we will still mention that, under the elementary conditions, there exists $0 < p < 1$ such that

$$E(\sigma_{0t}^2)^p < \infty. \tag{3.2}$$

The proof for such a result can be found in [13], Theorem 4.

The following lemma gives us the basic properties of the process (σ_{ut}^2) and the likelihood function (l_{ut}) .

Lemma 3.5 *If the elementary conditions hold*

- (i) *The process $(\sigma_{ut}^2(\theta))$ is strictly stationary and ergodic;*
- (ii) *The process $(l_{ut}(\theta))$ and the processes of its first and second derivatives with respect to θ are strictly stationary and ergodic for every θ in Θ ;*
- (iii) *For some $0 < p < 1$ and for every $\theta \in \Theta$ it holds*

$$E|\sigma_{ut}^2(\theta)|^p \leq H_p < \infty.$$

Proof: The statement (1) follows from Theorem 2.3.

Since $l_{ut}(\theta) = -\left(\ln \sigma_{ut}^2(\theta) + \frac{\varepsilon_t^2}{\sigma_{ut}^2(\theta)}\right)$, and

$$\frac{\partial l_{ut}}{\partial \omega} = \left(\frac{\varepsilon_t^2}{\sigma_{ut}^2} - 1\right) \frac{\partial \sigma_{ut}^2(\theta)}{\partial \omega} \frac{1}{\sigma_{ut}^2(\theta)}, \tag{3.3}$$

$$\frac{\partial l_{ut}}{\partial \alpha} = \left(\frac{\varepsilon_t^2}{\sigma_{ut}^2} - 1\right) \frac{\partial \sigma_{ut}^2(\theta)}{\partial \alpha} \frac{1}{\sigma_{ut}^2(\theta)}, \tag{3.4}$$

$$\frac{\partial l_{ut}}{\partial C} = \left(\frac{\varepsilon_t^2}{\sigma_{ut}^2} - 1\right) \frac{\partial \sigma_{ut}^2(\theta)}{\partial C} \frac{1}{\sigma_{ut}^2(\theta)} - 2 \frac{\varepsilon_t}{\sigma_{ut}^2(\theta)} \tag{3.5}$$

and

$$\frac{\partial l_{ut}}{\partial \beta} = \left(\frac{\varepsilon_t^2}{\sigma_{ut}^2} - 1\right) \frac{\partial \sigma_{ut}^2(\theta)}{\partial \beta} \frac{1}{\sigma_{ut}^2(\theta)}, \tag{3.6}$$

where

$$\frac{\partial \sigma_{ut}^2}{\partial \omega} = 1 + \beta \frac{\partial \sigma_{ut-1}^2}{\partial \omega}, \tag{3.7}$$

$$\frac{\partial \sigma_{ut}^2}{\partial \alpha} = \varepsilon_{t-1}^2 + \beta \frac{\partial \sigma_{ut-1}^2}{\partial \alpha}, \tag{3.8}$$

$$\frac{\partial \sigma_{ut}^2}{\partial C} = -2\alpha \varepsilon_{t-1} + \beta \frac{\partial \sigma_{ut-1}^2}{\partial C} \tag{3.9}$$

and

$$\frac{\partial \sigma_{ut}^2}{\partial \beta} = \sigma_{ut-1}^2 + \beta \frac{\partial \sigma_{ut-1}^2}{\partial \beta}, \tag{3.10}$$

it follows that the process $l_{ut}(\theta)$ and processes of its first and second derivatives are measurable functions of strictly stationary and ergodic process (ε_t) and so they are also strictly

stationary and ergodic. Finally, let $0 < p < 1$ from (3.2). Then it follows from Lemma 3.4

$$\begin{aligned} E(\sigma_{ut}^2(\theta))^p &\leq B^p E(\sigma_{\varepsilon t}^2(\theta))^p \\ &= B^p E\left(\omega + \alpha \sum_{k=0}^{\infty} \beta^k \varepsilon_{0t-1-k}^2\right)^p \\ &\leq B^p \left[\omega^p + \alpha^p \sum_{k=0}^{\infty} \beta^{kp} E(\varepsilon_{0t-k-1}^{2p})\right]. \end{aligned}$$

Since $\varepsilon_{0t-1-k}^2 \leq \alpha_0^{-1} \sigma_{0t}^2$ for every k , using (3.2) it follows

$$\begin{aligned} E(\sigma_{ut}^2)^p &\leq B^p \left[\omega^p + \alpha^p \sum_{k=0}^{\infty} \beta^{kp} \frac{1}{\alpha_0^p} E(\sigma_{0t}^{2p})\right] \\ &= B^p \left[\omega^p + \frac{\alpha^p}{\alpha_0^p} E(\sigma_{0t}^{2p}) \frac{1}{1 - \beta^p}\right] \\ &\leq B^p \left[\omega_d^p + \frac{\alpha_d^p}{\alpha_0^p} E(\sigma_{0t}^{2p}) \frac{1}{1 - \beta_d^p}\right] \equiv H_p < \infty. \end{aligned}$$

■

Some nontrivial calculus give us the following result.

Lemma 3.6 *Under the elementary conditions it holds*

$$\sup_{\theta \in \Theta} |L_{uT}(\theta) - L_T(\theta)| \longrightarrow 0 \quad \text{a.s. when } T \rightarrow \infty.$$

Finally, we want to find additional constraints for the expression $\frac{\sigma_{0t}^2}{\sigma_{ut}^2}$ and its inverse uniformly on $\theta \in \Theta$. We will do so by splitting the parameter space. Let $\mathcal{R}_l = \mathcal{R}(K_l^{-1} \alpha_l) < 1$ where $\mathcal{R}(\psi) = \frac{2 + \psi \mathcal{P}}{2 + \psi} < 1$, for $\psi > 0$, $\mathcal{P} = 1 - \left[\frac{1}{2^{\frac{2+\delta}{\delta}} S_{\delta}^{\frac{2}{\delta}}}\right] \in (0, 1)$

and S_{δ} define in the elementary conditions and $K_l = \frac{\omega_d}{\omega_0} + \frac{\alpha_d}{\alpha_0} < \infty$. Let η_l and η_d be positive constants such that

$$\eta_l < \beta_0(1 - \mathcal{R}_l^{\frac{1}{12}}) \quad \text{and} \quad \eta_d < \beta_0(1 - \mathcal{R}_0^{\frac{1}{12}}),$$

where $\mathcal{R}_0 = \mathcal{R}(\alpha_0) < 1$. For $1 \leq r \leq 12^2$ define constants

$$\beta_{rl} = \beta_0 \mathcal{R}_l^{\frac{1}{r}} + \eta_l < \beta_0 \quad \text{and} \quad \beta_{rd} = \frac{\beta_0 - \eta_d}{\mathcal{R}_0^{\frac{1}{r}}} > \beta_0,$$

subspaces

$$\Theta_l^r = \{\theta \in \Theta : \beta_{rd} \leq \beta \leq \beta_0\} \quad \text{and} \quad \Theta_d^r = \{\theta \in \Theta : \beta_0 \leq \beta \leq \beta_{rd}\}$$

²We will need r to be 12 in Lemma 4.2. Our aim is to find the minimal r so that all the statements presented bellow hold for every $\theta \in \Theta_r$.

and $\Theta_r = \Theta_l^r \cup \Theta_d^r$. The values η_l and η_d will depend on constants \mathcal{R}_l and \mathcal{R}_0 which are functions of the parameter space Θ .

Observe that we can choose $\Theta = \Theta_{r_{\max}} \subseteq \Theta_r$, for all $1 \leq r \leq 12$. Now we are able to present the result about the convergence in probability of the unconditional likelihood process.

Lemma 3.7 *Under the elementary conditions for every $\theta \in \Theta_1$ it holds:*

$$(1) E\left(\frac{\varepsilon_t^2}{\sigma_{ut}^2(\theta)}\right) \leq H_1 \equiv \frac{(C_d - C_l)^2}{\omega_l} + BH_c \text{ where } H_c = \frac{\omega_0}{\omega_l} + \frac{\alpha_0}{\alpha_l \eta_l} < \infty.$$

In this case it holds

$$(2) L_{uT}(\theta) \xrightarrow{P} L(\theta) \text{ when } T \rightarrow \infty, \text{ where } L(\theta) = E\left(\frac{l_{ut}(\theta)}{2}\right).$$

Proof: It is straightforward to show that $\left\|\frac{\sigma_{0t}^2}{\sigma_{\varepsilon t}^2}(\theta)\right\|_r \leq H_c$. Hence, using Lemma 3.4 and $g = C_0 - C$ we have the following

$$\begin{aligned} E\left(\frac{\varepsilon_t^2}{\sigma_{ut}^2}\right) &= E\left(\frac{(\varepsilon_{0t} + g)^2}{\sigma_{ut}^2}\right) \\ &= BE\left[\frac{\varepsilon_{0t}^2}{\sigma_{\varepsilon t}^2}\right] + 2gE\left[\frac{1}{\sigma_{ut}^2}E(\varepsilon_{0t}|\mathcal{F}_{t-1})\right] + E\left[\frac{g^2}{\sigma_{ut}^2}\right] \\ &\leq BE\left[\frac{\varepsilon_{0t}^2}{\sigma_{\varepsilon t}^2}\right] + \frac{g^2}{\omega_l} = BE\left[\frac{\sigma_{0t}^2}{\sigma_{\varepsilon t}^2}\right] + \frac{g^2}{\omega_l} \\ &\leq B\left\|\frac{\sigma_{0t}^2}{\sigma_{\varepsilon t}^2}\right\|_1 + \frac{g^2}{\omega_l} \end{aligned}$$

so

$$\begin{aligned} E\left(\frac{\varepsilon_t^2}{\sigma_{ut}^2}\right) &\leq BH_c + \frac{g^2}{\omega_l} \\ &\leq BH_c + \frac{(C_d - C_l)^2}{\omega_l} \equiv H_1 \end{aligned}$$

that proves the first statement. Additionally, we have

$$\begin{aligned} E|l_{ut}(\theta)| &= E\left|\ln(\sigma_{ut}^2(\theta)) + \frac{\varepsilon_t^2}{\sigma_{ut}^2(\theta)}\right| \\ &\leq E|\ln(\sigma_{ut}^2(\theta))| + E\left(\frac{\varepsilon_t^2}{\sigma_{ut}^2(\theta)}\right). \end{aligned}$$

But, for $x \geq 1$ and $0 < p < 1$ it holds the inequality $\ln x < \frac{1}{p}x^p$, so we have

$$\begin{aligned} E|\ln \sigma_{ut}^2(\theta)| &\leq |\ln \omega_l| + E\left|\ln \frac{\sigma_{ut}^2(\theta)}{\omega_l}\right| \\ &\leq |\ln \omega_l| + \frac{1}{p}E\left[\left(\frac{\sigma_{ut}^2(\theta)}{\omega_l}\right)^p\right] \\ &= |\ln \omega_l| + \frac{1}{p\omega_l}E[\sigma_{ut}^{2p}(\theta)] \end{aligned}$$

since $\frac{\sigma_{ut}^2(\theta)}{\omega_l} \geq 1$. Finally, using Lemma 3.5 we have

$$E|l_{ut}(\theta)| < \infty.$$

Since $(l_{ut}(\theta))$ is strictly stationary and ergodic, it follows

$$L_{uT}(\theta) = \frac{1}{2T} \sum_{t=1}^T l_{ut}(\theta) \xrightarrow{P} \frac{1}{2}E[l_{ut}(\theta)] = L(\theta), \quad \forall \theta \in \Theta_1.$$

■

The convergence in probability that we have presented in Lemma 3.7 is not a sufficient condition for the consistency of the quasi-maximum likelihood estimator. It is necessary that the convergence we have previously obtained holds uniformly. In order to obtain that, it is sufficient to find an upper bound for the score vector of the log-likelihood function $\nabla l_{ut}(\theta)$ uniformly on θ . The details regarding the explicit forms of the upper bounds can be found in [10].

Let

$$|A| = (\text{tr}(AA'))^{\frac{1}{2}} \quad \text{and} \quad \|A\|_r = (E|A|^r)^{\frac{1}{r}}$$

be the Euclidean norm of a matrix or a vector and the L^r norm of a random matrix or a vector respectively.

Now we are going to present the local consistency of the quasi-maximum likelihood estimator. Let us define

$$\hat{\theta}_T = \arg \max_{\theta \in \Theta_3} L_T(\theta).$$

$\hat{\theta}_T$ is the parameter value that maximizes the likelihood function on the set $\Theta_3 \subset \Theta$.

Theorem 3.8 Under *the* elementary conditions

$$\hat{\theta}_T \xrightarrow{P} \theta_0 \quad \text{when} \quad T \rightarrow \infty.$$

4 Asymptotic normality of the quasi-maximum likelihood estimator

In this section we present the asymptotic distribution of the quasi-maximum likelihood estimator (QMLE). In order to do so, we need stronger conditions on the process (Z_t) than the *elementary conditions* we have given in the previous section. In fact, we pretend that the fourth moment of the random variable Z_t is finite. We are going to call the following condition **additional condition**.

$$E(Z_0^4) \leq \mathcal{K} < \infty.$$

We do not present the proof for the following results as this would require long and non-trivial calculus.

Lemma 4.1 *Under the elementary conditions and under additional condition it holds*

- (i) $E|\nabla l_{ut}(\theta)\nabla l_{ut}(\theta)'| < \infty$, for every $\theta \in \Theta_{12}$;
- (ii) $\frac{1}{\sqrt{T}} \sum_{t=1}^T \nabla l_t(\theta_0) \xrightarrow{\mathcal{D}} N(0, A_0)$, where $A_0 = E(\nabla l_{ut}(\theta_0)\nabla l_{ut}(\theta_0)')$.

Let

$$B_T(\theta) = -\frac{1}{T} \sum_{t=1}^T \nabla^2 l_t(\theta) \quad \text{and} \quad B(\theta) = -E\nabla^2 l_{ut}(\theta).$$

Lemma 4.2 *Suppose the elementary conditions and the additional condition to hold. Then*

- (i) $E \sup_{\theta \in \Theta_{12}} |\nabla^2 l_{ut}(\theta)| < \infty$;
- (ii) For $i = 1, 2, 3, 4$, $E \sup_{\theta \in \Theta_{12}} \left| \frac{\partial}{\partial \theta_i} \nabla^2 l_{ut}(\theta) \right| < \infty$, where θ_i is the i -th element of θ ;
- (iii) $\sup_{\theta \in \Theta_{12}} |B_T(\theta) - B(\theta)| \xrightarrow{P} 0$ and $B(\theta)$ is a continuous function on Θ_{12} .

The following result presents one of the classical results in asymptotic analysis and it will be the basic tool for our further considerations. The details regarding the proof can be found in [9, p. 185].

Theorem 4.3 *Let (X_T) be a sequence of random $(m \times n)$ matrices and let (Y_T) be a sequence of random $(n \times 1)$ vectors such that $X_T \xrightarrow{P} C$ and $Y_T \xrightarrow{\mathcal{D}} Y \sim N(\mu, \Omega)$ when $T \rightarrow \infty$. Then the limiting distribution of $(X_T Y_T)$ is the same as that of CY ; that is*

$$X_T Y_T \xrightarrow{\mathcal{D}} N(C\mu, C\Omega C') \quad \text{when} \quad T \rightarrow \infty.$$

The following result assures that B_0 is a regular matrix.

Lemma 4.4 *Suppose that the joint distribution of $(\varepsilon_t, \varepsilon_t^2, \sigma_{ut}^2)$ is nondegenerate. Then for every $\theta \in \Theta$ the matrix*

$$E \left[\frac{\partial \sigma_{ut}^2}{\partial \theta} \frac{\partial \sigma_{ut}^2}{\partial \theta'} \sigma_{ut}^{-4} \right]$$

is positive definite.

Finally, we have all the necessary results for studying the asymptotic behavior of the parameter estimator. In fact, using the results presented above, the following theorem can be proved.

Theorem 4.5 *Suppose the elementary conditions and the additional condition to hold. Then*

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{\mathcal{D}} N(0, V_0),$$

where $V_0 = B_0^{-1}A_0B_0^{-1}$, $B_0 = B(\theta_0) = -E(\nabla^2 l_{ut}(\theta_0))$ and A_0 is defined in Lemma 4.1.

Notice that $A_0 = -\frac{1}{2}(EZ_0^4 - 1)B_0$. So, in the case in which (Z_t) is a sequence of random variables such that $Z_t \sim N(0, 1)$ we would have $EZ_0^4 - 1 = 2$ and $A_0 = -B_0$.

Let $\hat{B}_T = B_T(\hat{\theta}_T)$. In the case of maximum likelihood estimator, \hat{B}_T would be the standard estimator of the covariance matrix. But in a more general case of quasi-maximum likelihood estimator, the asymptotic covariance matrix is $B_0^{-1}A_0B_0^{-1}$ according to Theorem 4.5. Since this is not equal to B_0^{-1} , \hat{B}_T would not be a consistent estimator of that value.

Let us define

$$A_T(\theta) = \frac{1}{T} \sum_{t=1}^T \nabla l_t(\theta) \nabla l_t(\theta)'$$

and

$$\hat{A}_T = A_T(\hat{\theta}_T) \quad \text{and} \quad A(\theta) = E \nabla l_{ut}(\theta) \nabla l_{ut}(\theta)'$$

The following result presents the consistency of the covariance matrix estimator.

Lemma 4.6 *Suppose the elementary conditions and the additional condition to hold. Then*

$$(i) \sup_{\theta \in \Theta_{12}} |A_T(\theta) - A(\theta)| \xrightarrow{P} 0 \text{ and } A(\theta) \text{ is continuous on } \Theta_{12};$$

$$(ii) \hat{V}_T = \hat{B}_T^{-1} \hat{A}_T \hat{B}_T^{-1} \xrightarrow{P} B_0^{-1} A_0 B_0^{-1}.$$

Lemma 4.6 completes our characterization of classical properties of the QMLE for GARCH(1, 1) model. We show that the covariance matrix estimator is consistent for the asymptotic variance of the parameter estimator.

References

- [1] Amemiya, T. (1985): *Advanced Econometrics*. Cambridge: Harvard University Press.
- [2] Anderson, T.W. (1971): *The Statistical Analysis of Time Series*. New York: Wiley.
- [3] Basrak, B., Davis R.A., and Mikosch, T. (2002): Regular variation of GARCH processes. *Stochastic. Process. Appl.*, **99**, 95-115.
- [4] Basrak, B., Davis R.A., and Mikosch, T. (2002): A characterization of multivariate regular variation. *Ann. Appl. Probab.*, **12**, 908-920.
- [5] Berkes, I., Horváth, L., and Kokoszka, P. (2003): GARCH process: structure and estimation. *Bernoulli*, **9**, 201-227.
- [6] Billingsley, P. (1968): *Convergence of Probability Measures*. New York: Wiley.
- [7] Billingsley, P. (1979): *Probability and Measure*. New York: Wiley.
- [8] Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997): *Modelling Extremal Events*. Berlin: Springer-Verlag.
- [9] Hamilton, J.D. (1994): *Time Series Analysis*. Princeton: Princeton University Press.
- [10] Lee, S.W. and Hansen, B.E. (1994): Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. *Econometric Theory*, **10**, 29-52.
- [11] Mikosch, T. and Stărică, C. (1999): Change of structure in financial time series, long range dependence and the GARCH model. *IWI Preprint*, **5**.
- [12] Mikosch, T. and Straumann, D. (2003): *Stable limits of martingale transforms with application to the estimation of GARCH parameters*. Working Paper, no. 189.
- [13] Nelson, D.B. (1990): Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory*, **6**, 318-334.
- [14] Straumann, D. (2003): *Estimation in Conditionally Heteroscedastic Time Series Models*, Ph.D. Thesis, University of Copenhagen.
- [15] Williams, D. (1991): *Probability with Martingales*. Cambridge: Cambridge University Press.