Limits for the Gaussian QMLE in the Non-Stationary GARCH(1,1) Mod

Stelios Arvanitis & Alexandros Louka
Stable Limits for the Gaussian QMLE in the Non-Stationary GARCH(1,1) Model

Stelios Arvanitis* Advocates Louka

March 26, 2017

Address: Dept. of Economics, AUEB, Patision str. 76, Athens P.O. Box 10434, GREECE.

Abstract

We derive the limit theory of the Gaussian QMLE in the non-stationary GARCH(1,1) model when the squared innovation process lies in the domain of attraction of a stable law. Analogously to the stationary case, when the stability parameter lies in (1, 2], we find regularly varying rates and stable limits for the QMLE of the ARCH and GARCH parameters.

Keywords: Martingale Limit Theorem, Domain of Attraction, Stable Distribution, Slowly Varying Sequence, Non-Stationarity, Gaussian QMLE, Regularly Varying Rate.

JEL: C32.

1 Introduction

We derive the limit theory of the Gaussian QMLE in the non-stationary GARCH(1,1) model when the squared innovation process lies in the domain of attraction (DoA) of a $p$-stable law for $p \in (1, 2]$. Our interest stems from the empirical fact that distributions of financial asset returns exhibit fat tail behavior. This renders plausible the consideration of heavy-tailed distributions for the innovation process of GARCH-type models in financial applications. In the stationary versions of such cases, $\sqrt{n}$-consistency and possibly asymptotic normality can break down for the Gaussian QMLE (see for example Hall and Yao (2003); Mikosch and Straumann (2006); Arvanitis and Louka (2017)). Hence the question of whether this holds under non-stationarity arises naturally, and can be important for the determination of the asymptotic validity of inferential procedures based on the QMLE.

For the non-stationary GARCH(1,1), when the innovations fourth moments exist (hence $p = 2$), Jensen and Rahbek (2004a) and Francq and Zakoïan (2012) establish standard limit theories for the ARCH and GARCH parameters QMLE. In the non-stationary ARCH(1) case...

*Corresponding Author-Email address: stelios@aueb.gr.
Arvanitis and Louka (2016), extend Jensen and Rahbek (2004b) by allowing \( p \in (1, 2] \). They derive a Martingale Limit Theorem (MLT) for an appropriate martingale transform, and use it to establish slower than \( \sqrt{n} \)-rates and stable weak limits for the estimator of the ARCH parameter. In the GARCH(1,1) case this is not applicable since it requires a.s. convergence for the transform’s scaling process. Similarly, the MLT of Mikosch and Straumann (2006) is not immediately applicable since it requires mixing, while it is in any case not very informative on the form of the rate and the parameters of the limiting distribution.

In order to tackle those difficulties we partially extend the MLT in Arvanitis and Louka (2017) by allowing the approximation of the scaling process by a stationary and ergodic sequence. Hence we obtain as a byproduct an MLT for martingale transforms with non-stationary scaling sequences and stable limits (see the Supplement). Using the relevant approximation results of Francq and Zakoian (2012), we apply this to the non-stationary GARCH(1,1) case for \( p \in (1, 2] \), and thus derive regularly varying rates and (multivariate) stable limits for the QMLE of the ARCH and GARCH parameters. In the following section, we present our framework, derive and comment the results. We conclude with some questions for further research.

2 Framework and Results

The GARCH(1,1) process is defined by

\[
\begin{align*}
\sigma_t^2 &= \omega_0 + \alpha_0 y_{t-1}^2 + \beta_0 \sigma_{t-1}^2, \\
y_t &= \sigma_t z_t, \\
n &\in \mathbb{N}^*,
\end{align*}
\]

(1)

with initial value \( \sigma_0^2 \geq 0 \), a.s., where \( \omega_0 > 0, \alpha_0 > 0, \beta_0 \geq 0 \), and \( (z_t)_{t \in \mathbb{N}} \) is an iid sequence with \( \mathbb{E}[z_1] = 0, \mathbb{E}[z_t^2] = 1 \) and \( \mathbb{P}[z_t^2 = 1] < 1 \). Given \( (y_t)_{t=0,\ldots,n} \) from (1), we are interested in the limit theory of the Gaussian QMLE in the scope of the assumption framework that follows. The estimator, say \( \hat{\theta}_n \), of \( \theta_0 = (\omega_0, \alpha_0, \beta_0)' \) minimizes \( c_n(\theta) = \frac{1}{n} \sum_{t=1}^n \ell_t(\theta) \), with respect to \( \theta \equiv (\omega, \alpha, \beta)' \) on \( \Theta \), which is a compact subset of \( (0, \infty)^3 \), and \( \theta_0 \in \Theta \). There, \( \ell_t(\theta) = \frac{y_t^2}{h_t(\theta)} + \log h_t(\theta) \) and \( h_t(\theta) = \omega + \alpha y_{t-1}^2 + \beta h_{t-1}(\theta) \), with arbitrary \( h_0 \geq 0 \), a.s. In what follows all limits are considered as \( n \to +\infty \), \( \Rightarrow \) denotes convergence in distribution, and \( \|\cdot\|_p \) denotes the \( L_p \) norm.

Let \( S_p(s, c, \gamma) \) be the (univariate) stable distribution with parameters \( p, s, c, \gamma \) denoting stability, skewness, scale and location respectively (see Ibragimov (1971)). When \( p = 2 \), then \( s = 0 \) and \( S_2(0, c, \gamma) = N(\gamma, c) \). In what follows, multivariate stable distributions can be characterized by collections of univariate ones via appropriate projections (see Arjun K. Gupta (1994)).

Assumption 1. The distribution of \( z_t^2 - 1 \) lies in the DoA of some \( S_p(1, c, 0) \) with \( p \in (1, 2] \).

Notice that since the support of \( z_t^2 - 1 \) is a subset of \( [-1, +\infty) \), Assumption 1 directly implies that \( s = 1 \) for the attractor when \( p < 2 \) (see Ibragimov (1971) for the relation of \( s \) with the tails of the attracted distribution). Furthermore, the assumption is
equivalent to that (see Ibragimov (1971); Aaronson and Denker (1998)) the partial sums of \( (z_t^2 - 1)_{t \in \mathbb{N}} \) when appropriately translated, and scaled by \( \frac{1}{n^{1/\alpha} r_n^{1/\alpha}} \), converge in distribution to \( S_p(1, c, 0) \). \( (r_n)_{n \in \mathbb{N}} \) forms a slowly varying sequence (see Bingham et al. (1989)) defined by
\[
(nr_n)^{-1/p} = \inf \left\{ x > 0 : P \left( z_t^2 - 1 > x^{-1} \right) = 1/n \right\}.
\]
When (and only when) \( p < 2 \), or \( p = 2 \) and \( r_n \) diverges, then \( \mathbb{E} [z_t^4] = +\infty \). Hence this generalizes the usual assumption that \( \mathbb{E} [z_t^4] < +\infty \), which corresponds to \( p = 2 \) and \( r_n \to +\infty \).

**Theorem 1.** For the process defined in (1), let Assumptions 1 and 2.(A) hold and \( \mathbb{P} (z_1 = 0) = 0 \). If \( \Lambda_{\theta_0} > 0 \), \( (\alpha_n - \alpha_0, \beta_n - \beta_0)^T \) converges a.s. to zero. Moreover, if \( \theta_0 \in \Theta^* \),
\[
n^{-1/p} r_n^{-1/2} (\alpha_n - \alpha_0, \beta_n - \beta_0)^T \Rightarrow J_{\theta_0}^{-1} z_{\theta_0}.
\]

Here,
\[
J_{\theta_0} = \begin{pmatrix} \frac{1}{\sigma_n^2} & \frac{\mu_1}{\alpha_0 \sigma_n^2 (1 - \mu_1)} \\ \frac{\alpha_0 \sigma_n (1 - \mu_1)}{\mu_1} & \frac{\alpha_0 (1 - \mu_1)}{(1 + \mu_1) \mu_2} \end{pmatrix},
\]
with \( \mu_i \equiv \mathbb{E} (\beta_i / Z_1)^i \), \( i = 1, 2 \), \( z_{\theta_0} \) follows a bivariate \( p \)-stable distribution characterized by
\[
\lambda^T z_{\theta_0} \sim S_p (s_{\lambda}, c_{\lambda}, 0), \text{ for any } \lambda \in \mathbb{R}^2 \setminus \{0\},
\]
where \( s_{\lambda} \equiv \mathbb{E} [\lambda^T U_1 | \text{sgn}(\lambda^T U_1)] / \mathbb{E} [\lambda^T U_1^2] \), and \( c_{\lambda} \equiv c \mathbb{E} [\lambda^T U_1^p] \).

If \( \Lambda_{\theta_0} = 0 \), and Assumption 2.(B) holds, \( (\alpha_n - \alpha_0, \beta_n - \beta_0)^T \) converges to zero in probability. If moreover, \( \theta_0 \in \Theta^* \), and Assumption 2.(C) holds, (4) also holds.
For the above, consistency follows exactly as in Francq and Zakoïan (2012). For the limiting behavior of the relevant part of the score at $\theta_0$ notice that it has the form of the martingale transform $\sum_{t=1}^{n} (z_t^2 - 1) \sigma_t^{-2} h_t' (\theta_0)$, where $h_t'$ is the gradient of $h_t$ w.r.t. $\theta^* \equiv (\alpha, \beta)$. Given the approximation results of Francq and Zakoïan (2012) and Jensen and Rahbek (2004a) for the scaling process $(\sigma_t^{-2} h_t' (\theta_0))_{t \in \mathbb{N}}$ by $\left(\frac{1}{\alpha_0}, \sum_{j=1}^{\infty} \beta_j^{-1} \prod_{k=1}^{j} 1_{z_{t-k}}\right)^T_{t \in \mathbb{N}}$, this conforms to the MLT for martingale transforms we derive in the Supplement using the “Principle of Conditioning” of Jakubowski (1986). The remaining steps are handled as in Francq and Zakoïan (2012) under the evident modifications for the rates (see the Supplement).

As in Francq and Zakoïan (2012) the estimator for $\omega_0$ is inconsistent due to lack of asymptotic identification, something that is idiosyncratic to non-stationarity. The results would remain the same if the likelihood was maximized w.r.t. $\theta^*$ for arbitrary initial values $\omega, h_0$ when $\Lambda_{\theta_0} > 0$, and for the true ones $\omega_0, \sigma_0^2$ when $\Lambda_{\theta_0} = 0$. Then Assumption (2).(C) could be avoided, exactly as in Jensen and Rahbek (2004a).

The regularly varying rate is $n^{p-1} r_n^{-\frac{1}{p}}$ which is slower than the usual $\sqrt{n}$ whenever $p < 2$, or $p = 2$ and $r_n$ diverges. The limiting distribution is bivariate stable with stability parameter $p$ and spectral measure (see Mikosch and Straumann (2006)) characterized by linear transformations. The limiting marginals are $p$-stable as linear combinations of the bivariate $p$-stable distribution of $z_{\theta_0}$. The marginals of $z_{\theta_0}$ have symmetry parameters equal to 1 due to positivity. For example the first element of $z_{\theta_0}$ follows the

$$S_p \left( 1, \frac{c}{\alpha_0 p}, 0 \right).$$

When $p = 2$ and $r_n$ converges (necessarily to $\mathbb{E} [z_1^4] - 1$) we recover the results of Jensen and Rahbek (2004a); Francq and Zakoïan (2012), i.e. $\sqrt{n}$ rate and $N \left( 0, J_{\theta_0}^{-1} \right)$ limit. However, we still obtain asymptotic normality, when $p = 2$ and $r_n$ diverges. For example if $\sqrt{2} z_1 \sim t_4$ then simple calculations show that

$$\sqrt{\frac{n}{\log n}} (\alpha_n - \alpha_0, \beta_n - \beta_0)^T \Rightarrow N \left( 0, \frac{3}{2} J_{\theta_0}^{-1} \right).$$

The results of Theorem 1 can be extended to $p = 1$ as long as the condition $\mathbb{E} [z_1^2] = 1$ is retained. This case would involve a diverging translating sequence for the partial sums and a bivariate 1-stable limiting distribution with non-zero location (in the spirit of Arvanitis and Louka (2016, 2017)). We do not engage to such derivations for economy of space.

Notice that we obtain the same rates, and weak limits with the same stability parameters as in the stationary case (see Arvanitis and Louka (2017)), yet with different spectral measures due to the differing limiting behavior of the scaling sequence $(\sigma_t^{-2} h_t' (\theta_0))_{t \in \mathbb{N}}$ between the stationary and the non-stationary frameworks.

Compared to the $\mathbb{E} [z_1^4] < +\infty$ case, the results above can render non-robust, inferential procedures based on the QMLE when those are designed via the standard limit theory.
However there exist examples where this is not true. As an illustration consider the hypothesis structure of weak stationarity

\[ H_0 : \alpha_0 + \beta_0 < 1 \quad \text{against} \quad H_1 : \alpha_0 + \beta_0 \geq 1, \]

with statistic

\[ T_n \equiv \sqrt{n} \frac{(\alpha_n + \beta_n - 1)}{\sqrt{\sum_{t=1}^{n} \frac{y_t^2}{h_t^2(\theta_n)} - 1}} \left[ \frac{1}{n} \sum_{t=1}^{n} \left( \begin{array}{c} h_t'(\theta_n) \left( h_t'(\theta_n) \right)^T \end{array} \right) \right]^{-1} \left( \begin{array}{c} 0 \\ 1 \\ 1 \end{array} \right), \]

and rejection region \( R_a \equiv \{ T_n > \Phi (1 - a) \} \), for \( a \in (0, 1) \), where now \( h_t \) denotes the gradient of \( h_t \) w.r.t. \( \theta \). Notice that under the alternative there may exist \( \theta_0 \) corresponding to \( \Lambda_{\theta_0} \geq 0 \). Under the relevant framework, and if, Assumption 1 holds for \( p = 2 \), \( \mathbb{E} [z_1^2] \) slowly diverges to infinity, and \( \theta_0 \in \Theta^* \), using the results above, the ones in Arvanitis and Louka (2017) or Hall and Yao (2003), Theorem 2.3 of Francq and Zakoïan (2012) and the Generalized LLN (Theorem 2) in Section VII.7 of Feller (1971), it can be shown that \( T_n \) is self-normalized under the null, and the procedure above is asymptotically exact and (whenever \( \alpha_0 + \beta_0 > 1 \)) consistent. Using Francq and Zakoïan (2012), this means that it is robust in all cases where \( p = 2 \). This is not true when \( p < 2 \), whence it is possible that modifications of \( T_n \) can be used in the spirit of Arvanitis and Louka (2016) and rejection regions based on parametric bootstrap in the spirit of Hall and Yao (2003) in order to obtain asymptotic exactness and consistency.

3 Further Research

In a recent note in Economics Letters, Pedersen and Rahbek (2016) study the limit theory of the MLE for the non stationary GARCH(1,1) when \( z_1 \sim t_{v_0}, \ v_0 > 2 \). They derive \( \sqrt{n} \) rates and asymptotic normality for the estimator of \((\alpha_0, \beta_0, v_0)^T\). Notice that this conforms to Assumption 1 for \( p = \frac{v_0}{2} \). Hence when \( 2 < v_0 \leq 4 \), the MLE has faster rate and is relatively asymptotically efficient compared to the Gaussian QMLE. However when 1 holds, yet \( z_1 \) is not \( t \)-distributed the student-t QMLE may be inconsistent (see Berkes et al. (2004)). One possible line of future research concerns the study of the existence of an indirect estimator, that combines the use of the Gaussian and student-t QMLE, and has standard limit theory in the framework of Assumption 1 and in the spirit of Fan et al. (2014).

Finally, an obvious issue of further research concerns the derivation of the limit theory of the Gaussian QMLE in the non-stationary versions of other models of conditional heteroskedasticity.

Acknowledgement. This research was funded by the Research Centre of the Athens University of Economics and Business, in the framework of “Research Funding at AUEB for Excellence and Extroversion”.
References


