ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS
ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
SCHOOL OF ECONOMIC SCIENCES
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΗΣ ΕΠΙΣΤΗΜΗΣ
DEPARTMENT OF ECONOMICS

WORKING PAPER SERIES 05-2017

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

Stelios Arvanitis & Alexandros Louka

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5 / Fax: 210 8238249
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5 / Fax: (+30) 210 8238249 E-mail: econ@aueb.gr / www.aueb.gr

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

Stelios Arvanitis[*] Alexandros 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 tranform's scalling 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 *informable* 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 scalling process by a stationary and ergodic sequence. Hence we obtain as a *byproduct* an MLT for martingale transforms with non-stationary scalling sequences and stable limits (see the Supplement). Using the relevant approximation results of Francq and Zakoïan (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

$$\begin{cases} y_t = \sigma_t z_t \\ \sigma_t^2 = \omega_0 + \alpha_0 y_{t-1}^2 + \beta_0 \sigma_{t-1}^2 \end{cases}, t \in \mathbb{N}^\star \tag{1}$$

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 $\mathbb{E}[z_1] = 0$, $\mathbb{E}[z_1^2] = 1$ and $\mathbb{P}[z_1^2 = 1] < 1$. 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 θ_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 Θ, 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 $n \to +\infty$, $\rightsquigarrow$ denotes convergence in distribution, and $\|\cdot\|_\delta$ denotes the L_δ norm.

Let $S_p(s,c,\gamma)$ be the (univariate) stable distribution with parameters p, s, c, γ 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 $z_1^2 - 1$ lies in the DoA of* some $S_p(1,c,0)$ *with* $p \in (1,2]$.

Notice that since the support of $z_1^2 - 1$ 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

$$(nr_n)^{-1/p}=\inf\left\{x>0:\mathbb{P}\left(z_1^2-1>x^{-1}\right)=1/n\right\}.$$

When (and only when) $p<2$, or $p=2$ and r_n diverges, then $\mathbb{E}\left[z_1^4\right]=+\infty$. Hence this *generalizes* the usual assumption that $\mathbb{E}\left[z_1^4\right]<+\infty$, which corresponds to $p=2$ and $r_n\to\mathbb{E}\left[z_1^4\right]-1$. Furthermore, since $\mathbb{E}\left[z_1^2\right]<+\infty$, when $p=1$ then $r_n\to+\infty$.

Assumption 2. $(\mathbf{A})$ *θ_0 and z_1 satisfy*

$$\Lambda_{\theta_0}\equiv\mathbb{E}\left[\log\left(Z_1\right)\right]\geq 0. \tag{2}$$

When (2) holds with equality, $(\mathbf{B})$ Θ is such that $\beta<\left\|Z_1^{-1}\right\|_\delta^{-1}$ for some $\delta>1$, $(\mathbf{C})$ $\mathbb{E}\left[\left|\log\left(z_1^2\right)\right|\right]<+\infty$, and,

$$\mathbb{E}\left[\frac{1}{1+\sum_{i=1}^{n-1}\prod_{j=1}^{i}Z_j}\right]=o\left(\frac{1}{n^{1/a}r_n^{1/\alpha}}\right), \tag{3}$$

where $Z_t=\alpha_0z_t^2+\beta_0,\quad t=1,\ldots,n-1.$

Assumption (2).$(\mathbf{A})$ is equivalent to non-stationarity (see Nelson (1990)). When $\Lambda_{\theta_0}>0$, σ_n^2 diverges a.s. to infinity and when $\Lambda_{\theta_0}=0$ this holds in probability (see Francq and Zakoïan (2012)). In the latter case, $(\mathbf{B})$-$(\mathbf{C})$ are adaptations of analogous assumptions in Francq and Zakoïan (2012) to the current framework. Notice that (3) is true when $\mathbb{E}\left[\left|\log\left(Z_1\right)\right|^2\right]<+\infty$ (see again Francq and Zakoïan (2012)), which in conjunction with Assumption (1) holds for example when the support of z_1^2 is bounded away from zero.

Our main result follows.

Theorem 1. *For the process defined in (1), let Assumptions 1 and 2.$(\mathbf{A})$ hold and $\mathbb{P}\left(z_1=0\right)=0$. If $\Lambda_{\theta_0}>0$, $(\alpha_n-\alpha_0,\beta_n-\beta_0)^T$* converges a.s. to zero. *Moreover, if $\theta_0\in\Theta^\circ$,*

$$n^{\frac{p-1}{p}}r_n^{-\frac{1}{p}}\left(\alpha_n-\alpha_0,\beta_n-\beta_0\right)^T\rightsquigarrow J_{\theta_0}^{-1}z_{\theta_0}. \tag{4}$$

Here,

$$J_{\theta_0}=\begin{pmatrix}\frac{1}{\alpha_0^2} & \frac{\mu_1}{\alpha_0\beta_0(1-\mu_1)}\\ \frac{\mu_1}{\alpha_0\beta_0(1-\mu_1)} & \frac{(1+\mu_1)\mu_2}{\beta_0^2(1-\mu_1)(1-\mu_2)}\end{pmatrix},$$

with $\mu_i\equiv\mathbb{E}(\beta_0/Z_1)^i,\ i=1,2$, z_{θ_0} follows a bivariate p-stable distribution characterized by

$$\lambda^Tz_{\theta_0}\sim S_p\left(s_\lambda,c_\lambda,0\right),\ \textit{for any}\ \lambda\in\mathbb{R}^2\text{-}\{0\},$$

where $s_\lambda\equiv\frac{\mathbb{E}\left[\left|\lambda^TU_1\right|^p\operatorname{sgn}(\lambda^TU_1)\right]}{\mathbb{E}\left[\left|\lambda^TU_1\right|^p\right]}$, and $c_\lambda\equiv c\mathbb{E}\left[\left|\lambda^TU_1\right|^p\right]$.
If $\Lambda_{\theta_0}=0$, and *Assumption 2.$(\mathbf{B})$* holds, $(\alpha_n-\alpha_0,\beta_n-\beta_0)'$ converges to zero in probability. *If moreover, $\theta_0\in\Theta^\circ$, and Assumption 2.$(\mathbf{C})$* 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 θ_0 notice that it has the form of the martingale transform $\sum_{t=1}^{n}\left(z_t^2-1\right)\sigma_t^{-2}h_t'\left(\theta_0\right)$, where h_t' is the gradient of h_t w.r.t. $\theta^\star \equiv (\alpha,\beta)$. Given the approximation results of Francq and Zakoïan (2012) and Jensen and Rahbek (2004a) for the scalling process $\left(\sigma_t^{-2}h_t'\left(\theta_0\right)\right)_{t\in\mathbb{N}}$ by $\left(\left(\frac{1}{\alpha_0},\sum_{j=1}^{\infty}\beta_0^{j-1}\prod_{k=1}^{j}\frac{1}{Z_{t-k}}\right)^T\right)_{t\in\mathbb{N}}$, 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 ω_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. $\theta^\star$ for arbitrary initial values ω, h_0 when $\Lambda_{\theta_0}>0$, and for the true ones ω_0,σ_0^2 when $\Lambda_{\theta_0}=0$. Then Assumption (2).($\mathbf{C}$) could be avoided, exactly as in Jensen and Rahbek (2004a).

The regularly varying rate is $n^{\frac{p-1}{p}}r_n^{-\frac{1}{p}}$ 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_{θ_0}. The marginals of z_{θ_0} have symmetry parameters equal to 1 due to positivity. For example the first element of z_{θ_0} follows the

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

When $p=2$ and r_n converges (necessarily to $\mathbb{E}\left[z_1^4\right]-1$) 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

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

The results of Theorem 1 can be extended to $p=1$ as long as the condition $\mathbb{E}\left[z_1^2\right]=1$ 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 scalling sequence $\left(\sigma_t^{-2}h_t'\left(\theta_0\right)\right)_{t\in\mathbb{N}}$ between the stationary and the non-stationary frameworks.

Compared to the $\mathbb{E}\left[z_1^4\right]<+\infty$ 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

$$\mathbf{H}_0 : \alpha_0 + \beta_0 < 1 \quad \text{against} \quad \mathbf{H}_1 : \alpha_0 + \beta_0 \geq 1 \ ,$$

with statistic

$$T_n \equiv \frac{\sqrt{n}\left(\alpha_n + \beta_n - 1\right)}{\sqrt{\left(\frac{1}{n}\sum_{t=1}^{n}\frac{y_t^4}{h_t^2(\theta_n)} - 1\right)\begin{pmatrix}0\\1\\1\end{pmatrix}^T \left[\frac{1}{n}\sum_{t=1}^{n}\left(\frac{1}{h_t^2(\theta_n)}h_t^{'}\left(\theta_n\right)\left[h_t^{'}\left(\theta_n\right)\right]^T\right)\right]^{-1}\begin{pmatrix}0\\1\\1\end{pmatrix}}},$$

and rejection region $R_a \equiv \{T_n > \Phi\left(1-a\right)\}$, for $a \in (0,1)$, where *now* $h_t^{'}$ *denotes the gradient of* h_t *w.r.t.* θ. Notice that under the alternative there may exist θ_0 corresponding to $\Lambda_{\theta_0} \geq 0$. Under the relevant framework, and if, Assumption 1 holds for $p = 2$, $\mathbb{E}\left[z_1^4\right]$ slowly diverges to infinity, and $\theta_0 \in \Theta^{\circ}$, 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

Aaronson, J. and M. Denker (1998): "Characteristic functions of random variables attracted to 1-stable laws," *The Annals of Probability*, 26, 399–415.

Arjun K. Gupta, Truc T. Nguyen, W.-B. Z. (1994): "Conditions for Stability of Laws with All Projections Stable," *Sankhya: The Indian Journal of Statistics, Series A (1961-2002)*, 56, 438–443.

Arvanitis, S. and A. Louka (2016): "A Note on the QMLE Limit Theory in the Non-stationary ARCH (1) Model," *Journal of Time Series Econometrics*, 8, 21–39.

——— (2017): "Martingale Transforms with Mixed Stable Limit and the QMLE for Conditionally Heteroskedastic Models," *Working Paper 04-17, Dept. of Economics, AUEB, http://www.econ.aueb.gr/uploadfiles/AllDP042017*.

Berkes, I., L. Horváth, et al. (2004): "The efficiency of the estimators of the parameters in GARCH processes," *The Annals of Statistics*, 32, 633–655.

Bingham, N. H., C. M. Goldie, and J. L. Teugels (1989): *Regular variation*, vol. 27, Cambridge university press.

Fan, J., L. Qi, and D. Xiu (2014): "Quasi-maximum likelihood estimation of GARCH models with heavy-tailed likelihoods," *Journal of Business & Economic Statistics*, 32, 178–191.

Feller, W. (1971): *An Introduction to Probability Theory and Its Applications: Volumes I & II: Second Edition*, John Wiley & Sons.

Francq, C. and J.-M. Zakoïan (2012): "Strict stationarity testing and estimation of explosive and stationary generalized autoregressive conditional heteroscedasticity models," *Econometrica*, 80, 821–861.

Hall, P. and Q. Yao (2003): "Inference in ARCH and GARCH Models with Heavy-Tailed Errors," *Econometrica*, 71, 285–317.

Ibragimov, I. A. (1971): *Independent and stationary sequences of random variables*, Wolters-Noordhoff.

Jakubowski, A. (1986): "Principle of Conditioning in Limit Theorems for Sums of Random Variables," *The Annals of Probability*, 14, 902–915.

Jensen, S. T. and A. Rahbek (2004a): "Asymptotic Inference for Nonstationary GARCH," *Econometric Theory*, 20, 1203–1226.

——— (2004b): "Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case," *Econometrica*, 641–646.

Mikosch, T. and D. Straumann (2006): "Stable Limits of Martingale Transforms with Application to the Estimation of Garch Parameters," *The Annals of Statistics*, 34, 493–522.

Nelson, D. B. (1990): "Stationarity and Persistence in the GARCH(1,1) Model," *Econometric Theory*, 6, 318–334.

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