Quasi Maximum Likelihood Estimation for Non-Stationary TGARCH(1, 1) Models

Hui Wang[1,3] * Jiazhu Pan [2]

[1]School of Finance, Central University of Finance and Economics, China.

[2]Department of Mathematics and Statistics, University of Strathclyde, UK .

[3]Corresponding author:Hui Wang, email: xiaohuipk@163.com

Abstract. The threshold GARCH (TGARCH) models have been very useful for analyzing asymmetric volatilities arising from financial time series. Most research on TGARCH has been directed to the stationary case. This paper studies the estimation of non-stationary first order TGARCH models. Gaussian quasi-maximum likelihood estimation (G-QMLE) and normal mixture quasi-maximum likelihood estimation (NM-QMLE) for non-stationary TGARCH models are proposed. We show that the proposed estimators are consistent and asymptotically normal under mild regular conditions. The impact of relative tail heaviness of the innovations distribution and quasi-likelihood distributions on the asymptotic efficiency has been thoroughly discussed.

Key words . QMLE;normal mixture; consistency; asymptotic normality; efficiency

1 Introduction

Since the seminal papers by Engle (1982) and Bollerslev (1986), GARCH models have been proved particularly valuable in modelling time varying volatility. Most literature on inference of GARCH models is based on Gaussian quasi-maximum likelihood estimation (G-QMLE) due to its simplicity. Regarding to the asymptotic inference of the G-QMLE for stationary GARCH models, the consistency and asymptotic normality have been established under different conditions, see Berkes et al. (2003), Hall and Yao (2003), and Francq and Zakoïan (2004) etc.. However, gain in robustness comes with efficiency loss, which means that the variance of the estimates fails to reach the Cramér-Rao bound, see González-Rivera and Drost (1999). On other hand, many authors pointed out strong evidence against the normality assumption through empirical studies, see for instance Mikosch and Stărică (2000). Based on a general quasi-likelihood distribution which may be very heavy-tailed, Berks and Horváth (2004) proposed a class of QMLE for stationary GARCH models and compared the efficiency of QMLE based on different quasi-likelihood distribution assumptions. In the statistics literature, mixtures of distributions have been widely used in modeling of heavy-tailed distributions, see for instance Zhang et al. (2006). Specially, to capture the skewness and heavy-tail of the innovations Lee and Lee (2009) proposed the normal mixture QMLE (NM-QMLE), which is obtained from the normal

*Partially supported by a grant from the project 211 of the Central University of Finance and Economics, the CUFE Young Scholar Innovation Fund, and the National Natural Science Foundation of China (grant 11101448).

mixture quasi-likelihood, and demonstrated that the NM-QMLE is consistent and asymptotically normal.

Nonstationarity in the volatility process has been well documented for macroeconomic and financial time series data, see Loretan and Phillips (1994) and Hwang et al. (2010). Jensen and Rahbek (2004 a, 2004 b) are the first to consider the asymptotic theory of the G-QMLE for non-stationary ARCH/GARCH(1,1) models. The further studies for inference of non-stationary GARCH models include Linton et al. (2010) and Francq and Zakoïan (2012). Among the asymmetric GARCH models, threshold GARCH (TGARCH) model is one of the most popular models in the literature, see Li and Li (1996), Pan et al.(2008) and Hwang et al. (2010). Note that nonstationary TGARCH models capture the nonstationarity and asymmetry of the volatility of time series data simultaneously. This motivates us to study the estimation problem of nonstationary TGARCH models.

In this paper, we propose G-QMLE and NM-QMLE for nonstationry TGARCH(1, 1) model and show that the G-QMLE and NM-QMLE are both consistent and asymptotically normal under some regular conditions. We find that the NM-QMLE is more efficient than G-QMLE when the distribution of innovations is more heavy-tailed than Gaussian distribution through a simulation.

The rest of this paper is organized as follows. In Section 2, G-QMLE of nonstationary TGARCH(1, 1) models is considered. Section 3 presents G-QMLE of nonstationary TGARCH(1, 1) models. The asymptotic efficiencies are discussed in Section 4.

2 The model and the Gaussian-QMLE

The TGARCH(1, 1) model is defined by

$$X_t = \sigma_t \varepsilon_t \quad \text{and} \quad \sigma_t^2 = \omega + \alpha^+ (X_{t-1}^+)^2 + \alpha^- (X_{t-1}^-)^2 + \beta \sigma_{t-1}^2, \tag{2.1}$$

where $\omega > 0$, $\alpha^+ \geq 0$, $\alpha^- \geq 0$, $\beta \geq 0$ are unknown parameters, and $\{\varepsilon_t\}$ is a sequence of independent and identically distributed (iid) random variables with $E\varepsilon_t = 0$ and $E\varepsilon_t^2 = 1$, such that ε_t is independent of $\{X_{t-k}, k \geq 1\}$ for all t. According to Pan et al. (2008), there exists a unique strictly stationary and ergodic solution to model (2.1) if and only if $E \log \left[\alpha^+ (\varepsilon_{t-1}^+)^2 + \alpha^- (\varepsilon_{t-1}^-)^2 + \beta\right] < 0$. The initial value of X_t is assumed to be X_0 and the unobserved σ_0^2 is parameterized by η_0. The parameter of model (2.1) is then $\phi = (\alpha^+, \alpha^-, \beta, \omega, \eta)'$ with true value $\phi_0 = (\alpha_0^+, \alpha_0^-, \beta_0, \omega_0, \eta_0)'$. Let $\varphi = (\alpha^+, \alpha^-, \beta)'$ and $\psi = (\omega, \eta)'$ with the true value $\varphi_0 = (\alpha_0^+, \alpha_0^-, \beta_0)'$ and $\psi_0 = (\omega_0, \eta_0)'$ respectively. Define

$$\sigma_t^2(\phi) = \omega + \alpha^+ (X_{t-1}^+)^2 + \alpha^- (X_{t-1}^-)^2 + \beta \sigma_{t-1}^2(\phi) \tag{2.2}$$

with $\sigma_0^2(\phi) = \eta$ and $\sigma_t^2(\phi_0) = \sigma_t^2$.

The G-QMLE $\hat{\phi}_{1n}$ is defined as a maximizer of Gaussian quasi log-likelihood function, equivalently $\hat{\phi}_{1n} = \arg\min_\phi l_n^G(\phi)$, where

$$l_n^G(\phi) = \frac{1}{n} \sum_{t=1}^{n} \Big[\log \sigma_t^2(\phi) + \frac{X_t^2}{\sigma_t^2(\phi)} \Big], \tag{2.3}$$

and $\sigma_t^2(\phi)$ is defined by (2.2). Denote $l_n^G(\phi)|_{\phi=(\varphi',(\psi^0)')'}$ by $l_n^G(\varphi)$ and $l_n^G(\phi)|_{\phi=(\varphi',\psi_*')'}$ by $l_{n*}^G(\varphi)$, where $l_n^G(\phi)$ is defined in (2.3), and $\psi_* = (\omega_*, \eta_*)'$ is some fixed value of ψ.

We will carry out the discussion under the following basic assumptions.

A1. $\gamma_0 = E\log\left[\alpha_0^+(\varepsilon_{t-1}^+)^2 + \alpha_0^-(\varepsilon_{t-1}^-)^2 + \beta_0\right] \geq 0$

A2. $E\varepsilon_t^2 = 1$ and neither ε_t^+ nor ε_t^- is constant.

A3. $E\varepsilon_t^4 < \infty$.

Theorem 1. *Suppose assumptions A1-A3 hold. Then it follows that*

(i) There exists a fixed open neighborhood $N(\varphi_0)$ of φ_0 such that $l_n^G(\varphi)$ has a unique minimum $\hat\varphi_{1n}$ in $N(\varphi_0)$ with probability tending to one as $n \to \infty$ and $\hat\varphi_{1n}$ is consistent. Furthermore, $\hat\varphi_{1n}$ is asymptotically normal

$$\sqrt{n}(\hat\varphi_{1n} - \varphi_0) \xrightarrow{\mathcal{L}} N\Big(0, (E\varepsilon_t^4 - 1)J^{-1}\Big),$$

where $J = ED_tD_t'$ and $D_t = (D_{1t}, D_{2t}, D_{3t})'$ with

$$D_{1t} = \sum_{j=1}^{+\infty} \beta_0^{j-1}(\varepsilon_{t-j}^+)^2 \prod_{k=1}^{j} \frac{1}{\alpha_0^+(\varepsilon_{t-k}^+)^2 + \alpha_0^-(\varepsilon_{t-k}^-)^2 + \beta_0},$$

$$D_{2t} = \sum_{j=1}^{+\infty} \beta_0^{j-1}(\varepsilon_{t-j}^-)^2 \prod_{k=1}^{j} \frac{1}{\alpha_0^+(\varepsilon_{t-k}^+)^2 + \alpha_0^-(\varepsilon_{t-k}^-)^2 + \beta_0},$$

$$D_{3t} = \sum_{j=1}^{+\infty} \beta_0^{j-1} \prod_{k=1}^{j} \frac{1}{\alpha_0^+(\varepsilon_{t-k}^+)^2 + \alpha_0^-(\varepsilon_{t-k}^-)^2 + \beta_0}.$$

(ii) If γ_0 in assumption A1 is strictly positive, the results in (i) hold for $l_{n*}^G(\varphi)$.

3 The NM-QMLE

Compared with Gaussian distribution, normal mixture distribution is more appropriate for modeling heavy-tailed and skewed data. In this subsection we establish the asymptotic properties of NM-QMLE for nonstationary TGARCH(1, 1) models. The s component normal mixture (NM) density is of the form

$$g_\vartheta(y) = \sum_{k=1}^{s} p_k f(y; \mu_k, \varrho_k), \tag{3.1}$$

where $\vartheta = (p_1, \cdots, p_{s-1}, \mu_1, \cdots, \mu_{s-1}, \varrho_1, \cdots, \varrho_s)'$ and $f(y; \mu_k, \varrho_k) = \frac{1}{\sqrt{2\pi}\varrho_k} \exp\Big\{-\frac{(y-\mu_k)^2}{2\varrho_k^2}\Big\}$ satisfying

$$\sum_{k=1}^{s} p_k = 1, \quad \sum_{k=1}^{s} p_k\mu_k = 0 \quad \textit{and} \quad \sum_{k=1}^{s} p_k(\mu_k^2 + \varrho_k^2) < \infty. \tag{3.2}$$

In general, the s component normal mixture distribution is not identifiable, so we need the same identification condition as in Lee and Lee (2009). Furthermore, we assume $\mathcal{G}$ is nondegenerate, that is, any s-component normal mixture density in $\mathcal{G}$ can not be represented as a mixture with the number of components less than s.

Conditionally on initial values X_0, η_0, the normal mixture quasi-likelihood is given by

$$L_n^{NM}(\vartheta, \phi) = \prod_{t=1}^{n} \Big\{ \sum_{k=1}^{s} p_k \frac{1}{\sqrt{2\pi \varrho_k^2 \sigma_t^2(\phi)}} \exp\Big\{ - \frac{(X_t - \mu_k \sigma_t(\phi))^2}{2\varrho_k^2 \sigma_t^2(\phi)} \Big\}\Big\}. \tag{3.3}$$

A natural idea is to obtain an estimator of ϕ by maximizing $L_n^{NM}(\vartheta, \phi)$, and nuisance parameters ϑ are also estimated at the same time. Since the density function g of ε_t may be not in $\mathcal{G}$, then what does the true value ϑ_0 of ϑ mean? One may hope ϑ_0 can minimize the discrepancy between the true innovation density g and the quasi likelihood normal mixture density in the sense of Kullback-Leibler Information Distance (KLID). Thus, we define the true value $\vartheta_0 = (p_{10}, \cdots, p_{(s-1)0}, \mu_{10}, \cdots, \mu_{(s-1)0}, \varrho_{10}, \cdots, \varrho_{s0})'$ as follows,

$$\vartheta_0 = \big\{\vartheta \in \tilde{\Theta} : d(g, g_\vartheta) = \min_{t \in \tilde{\Theta}} \ d(g, g_t)\big\}, \tag{3.4}$$

where $d(g, g_t) = \int g(x)\big(\log g(x) - \log g_t(x)\big)dx$ is the KLID between g and g_t. Note that ϑ_0 here only depends on the KLID of the two densities under consideration. Once ϑ_0 is given, namely the true innovation distribution g is known, the NM-QMLE $\hat{\phi}_{2n}$ is defined by maximizing normal mixture quasi likelihood with parameter ϑ_0. Put $\theta = (\vartheta', \phi')'$ with true value $\theta_0 = ({\vartheta_0}', {\phi_0}')'$ and

$$l_n^{NM}(\theta) = l_n^{NM}(\vartheta, \phi) =: -n^{-1} \log L_n^{NM}(\vartheta, \phi) = n^{-1} \sum_{t=1}^{n} W_t(\theta), \tag{3.5}$$

where

$$W_t(\theta) = W_t(\vartheta, \phi) = -\log \Big\{ \frac{1}{\sigma_t(\phi)} g_\vartheta\Big(\frac{X_t}{\sigma_t(\phi)}\Big)\Big\} \tag{3.6}$$

and $L^{NM}(\vartheta, \phi)$ is defined in (3.3). Namely, we define

$$\hat{\phi}_{2n} = \arg\min l_n^{NM}(\vartheta_0, \phi). \tag{3.7}$$

For $-\infty < y < \infty$ and $t > 0$, let

$$h_\vartheta(y, t) = \log\big(t g_\vartheta(yt)\big), \quad h_{1\vartheta}(y, t) = \frac{\partial h_\vartheta(y, t)}{\partial t} = \frac{1}{t} + \frac{y}{g_\vartheta(yt)} \frac{\partial g_\vartheta(yt)}{\partial y}, \tag{3.8}$$

$$h_{2\vartheta}(y, t) = \frac{\partial^2 h_\vartheta((y, t)}{\partial t^2} = -\frac{1}{t^2} - \frac{y^2}{g_\vartheta^2(yt)} \Big[\frac{\partial g_\vartheta(yt)}{\partial y}\Big]^2 + \frac{y^2}{g_\vartheta(yt)} \frac{\partial^2 g_\vartheta(yt)}{\partial y^2} \tag{3.9}$$

In order to obtain asymptotic properties of $\hat{\phi}_{2n}$, we need the following regularity conditions:

A4. Θ is compact and θ_0 lies in the interior of Θ;

A5. $Eh_{2\vartheta_0}(\varepsilon_0, 1) \neq 0$;

A6. $E\varepsilon_t^6 < \infty$.

We are now ready to state our asymptotic results for the NM-QMLE $\hat{\phi}_{2n}$.

Theorem 2. *Suppose $\mathcal{G}$ is identifiable and nondegenerate. If assumptions A1-A6 hold, then it follows that*

(i) There exists a fixed open neighborhood $N(\varphi_0)$ of φ_0 such that $l_n^{NM}(\varphi)$ has a unique minimum $\hat{\varphi}_{2n}$ in $N(\varphi_0)$ with probability tending to one as $n \to \infty$, and $\hat{\varphi}_{2n}$ is consistent. Furthermore, $\hat{\varphi}_{2n}$ is asymptotically normal,

$$\sqrt{n}(\hat{\varphi}_{2n} - \varphi_0) \xrightarrow{\mathcal{L}} N\big(0, 4\tau^2 J^{-1}\big).$$

Here $\tau^2 = Eh_{1\vartheta_0}^2(\varepsilon_0, 1)/E\big(h_{2\vartheta_0}(\varepsilon_0, 1)\big)^2$ and $J = ED_tD_t'$.

(ii) If γ_0 in assumption A1 is strictly positive, the results in (i) hold for $l_{n*}^{NM}(\varphi)$.

If ϑ_0 is unknown, the NM-QMLE of θ_0 is then defined by

$$\hat{\theta}_n^{NM} =: (\hat{\vartheta}_n', \hat{\phi}_{3n}')' = \arg\min l_n^{NM}(\vartheta, \phi) \tag{3.10}$$

where $l_n^{NM}(\vartheta, \phi)$ is defined in (3.5). Next, we discuss the asymptotic properties of the NM-QMLE $\hat{\theta}_n^{NM}$.The following condition is needed.

A5'. H_2 (The limit of the second derivative of $l_n^{NM}(\theta_0)$) is a positive definite matrix.

The following theorem gives asymptotic properties of the NM-QMLE $\hat{\theta}_n^{NM}$.

Theorem 3. *Suppose $\mathcal{G}$ is identifiable and nondegenerate. If assumptions A1, A2, A4, A5$'$ and A6 hold, it follows that*

(i) There exists a fixed open neighborhood $N(\zeta_0)$ of ζ_0 such that $l_n^{NM}(\zeta)$ has a unique minimum $\hat{\zeta}_n^{NM}$ in $N(\zeta_0)$ with probability tending to one as $n \to \infty$ and $\hat{\zeta}_n^{NM}$ is consistent. Furthermore, $\hat{\zeta}_n^{NM}$ is asymptotically normal

$$\sqrt{n}(\hat{\zeta}_n^{NM} - \zeta_0) \xrightarrow{\mathcal{L}} N\big(0, H_2^{-1}H_1H_2^{-1}\big).$$

Here H_1 is the asymptotic variance of $\sqrt{n}\frac{\partial l_n^{NM}(\theta_0)}{\partial \zeta}$ and H_2 is the same as assumption A5'.

(ii) If γ_0 in assumption A1 is strictly positive, the results in (i) hold for $l_{n*}^{NM}(\zeta)$.

4 The asymptotic efficiencies of G-QMLE and NM-QMLE

From Theorem 2, we have that the ratio of the asymptotic variances between $\hat{\phi}_{2n}$ and $\hat{\phi}_{1n}$ are $ER = 4\tau^2/(E\varepsilon_t^4 - 1)$. Therefore, as long as the ER is smaller than 1, the NM-QMLE is more efficient than the G-QMLE. However, it is not easy to get the theoretical value of ER. Our simulation results below indicate that the heavier the true density of the innovations ε_t, the more efficient the NM-QMLE. we present numerical evidence on the performance of asymptotic efficiencies of the proposed G-QMLE and NM-QMLE through simulation studies. The data are generated from the non-stationary TGARCH(1,1) model (2.1) with the true parameter $\phi_0 = (0.001, 0.1, 0.3, 1, 0.01)'$. In all experiments, we use the sample size $n = 1000$ with 2000 replications. For the distribution of innovation ε_t, we consider the standard Gaussian distribution, the skewed normal-mixture distribution with $\theta_0 = (0.2, 1, 2, 1)'$ and the generalized error distribution with shape parameter 3. Table 1 gives the ratio of the asymptotic variances of NM-QMLE and G-QMLE. As expected, the NM-QMLE gets more efficient than G-QMLE when the distribution of ε_t becomes more skewed and more heavy-tailed.

Table 1: The variance ratio

Distribution	(a)	(b)	(c)
ER	1.0849	0.7775	0.7694

References

[1] Berkes, I., Horváth, L. (2004). The efficiency of the estimators of the parameters in GARCH processes. *Annals of Statistics*. 32, 633-655.

[2] Berkes, I., Horváth, L. and Kokoszka, P. (2003). GARCH processes: structure and estimation. *Bernoulli*. 9(2), 201-227.

[3] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*. 31, 307-327.

[4] Francq, C. and Zakoïan, J.M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* . 10, 605C637.

[5] Francq, C. and Zakoïan, J.M. (2012). Strict staitonarity testing and estimation of explosive and stationary generalized autoregressive conditional heteroscedasticity models. *Econometrica* . 80, 821C861.

[6] González-Rivera, G. and Drost, F.C. (1999). Efficiency comparisons of maximum-likelihood-based estimators in GARCH models. *Journal of Econometrics*, 93(1), 93-111.

[7] Hwang, S. Y., Baek, J.S., Park, J.A. and Choi, M.S. (2010). Explosive volatilities for threshold-GARCH processes generated by asymmetric innovations. *Statistics and probability Letters* 80, 26-33.

[8] Jensen, S. T. and Rahbek, A. (2004 a). Asymptotic normality of the QMLE estimator of ARCH in the nonstatinary case. *Econometrica* 72, 641-646.

[9] Jensen, S. T. and Rahbek, A. (2004 b). Asymptotic inference for nonstationary GARCH. *Econometric Theory* 20, 1203-1226.

[10] Lee, T.-W. and Lee, S. (2009). Normal mixture quasi-maximum likelihood estimator for GARCH models. *Scandinavian Journal of Statistics*, 36,157-170.

[11] Linton, O., Pan, J.Z. and Wang, H. (2010). Estimation for a non-stationary semi-strong GARCH(1,1) model with heavy-tailed errors. *Econometric Theory.* 26(1), 1-28.

[12] Loretan, M. and Phillips, P. C. B. (1994). Testing the covariance stationarity of heavy-tailed time series. *Journal of Empirical Finance*, 1, 211-248.

[13] Pan, J. Z., Wang, H. and Tong, H. (2008). Estimation and tests for power-transformed and threshold GARCH models. *Journal of Econometrics*, 142, 352-378.