Communications for Statistical Applications and Methods
2013, Vol. 20, No. 5, 395–404
DOI: http://dx.doi.org/10.5351/CSAM.2013.20.5.395

Skewness of Gaussian Mixture Absolute Value GARCH(1, 1) Model

Taewook Lee[1,a]

[a]Department of Statistics, Hankuk University of Foreign Studies

Abstract

This paper studies the skewness of the absolute value GARCH(1, 1) models with Gaussian mixture innovations (Gaussian mixture AVGARCH(1, 1) models). The maximum estimated-likelihood estimator (MELE) employed (a two- step estimation method in order to estimate the skewness of Gaussian mixture AVGARCH(1, 1) models. Through the real data analysis, the adequacy of adopting Gaussian mixture innovations is exhibited in reflecting the skewness of two major Korean stock indices.

Keywords: AVGARCH model, Gaussian mixture, skewness, Taylor effect.

1. Introduction

Many researchers have studied the dependence of financial time series data by analyzing the autocorrelation function (ACF) of absolute and squared returns. One of the stylized facts in the pattern of ACF is the Taylor effect discovered by Taylor (1988), which implies that the ACF of absolute returns tends to be larger than those of squared returns. Generally, for any given integer value k, $\rho_\delta(k) := \mathrm{Corr}(|r_{t-k}|^\delta, |r_t|^\delta), \delta > 0$, is maximized for δ close to unity for returns, r_t. We refer to Ding *et al.* (1993) and Ding and Granger (1996) for the generalized Taylor effect. In order to reflect the Taylor effect, some researchers such as He and Teräsvirta (1999), Gonçalves *et al.* (2009) and Haas (2009), consider the absolute value GARCH(1, 1) (AVGARCH(1, 1)) models under the assumption that the innovations have a symmetric density. However, the degree of the conditional kurtosis of the innovations is crucial for the appearance of the Taylor effect (*cf.* Haas, 2009). Most studies on the Taylor effect have been done under symmetric innovation distributions with high kurtosis such as generalized exponential and student-t distributions. However, such a symmetric assumption should be carefully investigated since financial time series data is frequently reported to be left-skewed by many empirical studies (*cf.* Haas *et al.*, 2004; Lee *et al.*, 2009; Lee and Lee, 2009); subsequently, symmetric AVGARCH(1,1) models might be inconsistent with skewed financial time series data.

In this paper, we derive the expression for the skewness of AVGARCH(1, 1) models under the assumption that innovations follow Gaussian mixture distributions. We refer to this model as Gaussian mixture AVGARCH(1, 1) models. This is because the Gaussian mixture approach is a functional tool to model leptokurtic and skewed distributions (*cf.* McLachlan and Peel, 2000). In order to estimate the skewness of Gaussian mixture AVGARCH(1, 1) models, one can simply adopt the usual conditional maximum likelihood estimator (CMLE) and EM-like algorithm as in Lee and Lee (2009). However, the EM-like algorithm for computing CMLE is highly time consuming and does not completely guarantee its convergence (*cf.* Lee and Lee, 2009). Therefore, we employ the maximum

[1] Professor, Department of Statistics, Hankuk University of Foreign Studies, Yongin, Kyonggi 449-791, Korea.
E-mail: twlee@hufs.ac.kr

estimated-likelihood estimator (MELE), which is a type of the two-step estimation method of Lee and Lee (2012). Through the real data analysis, the adequacy of adopting Gaussian mixture innovations is exhibited in reflecting the skewness of two major Korean stock indices, KOSPI200 (Korea Stock Price Index 200) and KOSDAQ (Korea Securities Dealers Automated Quotation).

This paper is organized as follows. In Section 2, we introduce the MELE and present the expression of the skewness of Gaussian mixture AVGARCH(1, 1) models. In Section 3, we perform a real data analysis to illustrate our findings. In Section 4, we provide the proof of the theorems presented in Section 2.

2. Main Results

Consider the absolute value GARCH(1, 1) model:

$$\begin{aligned} X_t &= \epsilon_t \sigma_t(\theta), \\ \sigma_t(\theta) &= \omega + \alpha |X_{t-1}| + \beta \sigma_{t-1}(\theta), \quad t \in Z, \end{aligned} \tag{2.1}$$

where $\{\epsilon_t\}$ is a sequence of iid r.v.'s with zero mean and unit variance. We further assume that ϵ_t follows a two component Gaussian mixture density of the form

$$f_\eta(y) = \pi f(y; \mu_1, \sigma_1) + (1 - \pi) f(y; \mu_2, \sigma_2),$$

where $f(y; \mu_i, \sigma_i) := 1/\{\sqrt{2\pi}\sigma_i\} \exp[-1/2\{(y - \mu_i)/\sigma_i\}^2]$ for $i = 1, 2$, where

$$\begin{aligned} E[\epsilon_t] &= \pi\mu_1 + (1 - \pi)\mu_2 = 0, \\ E\left[\epsilon_t^2\right] &= \pi\left(\mu_1^2 + \sigma_1^2\right) + (1 - \pi)\left(\mu_2^2 + \sigma_2^2\right) = 1. \end{aligned}$$

The parameter vector is $\varphi = (\theta^T, \eta^T)^T$, where $\theta = (\omega, \alpha, \beta)^T$ and $\eta = (\pi, \mu_1, \sigma_1)^T$ and the true parameter vector is $\varphi_0 = (\theta_0^T, \eta_0^T)^T$, where $\theta_0 = (\omega_0, \alpha_0, \beta_0)^T$ and $\eta_0 = (\pi_0, \mu_{10}, \sigma_{10})^T$. The parameter space of φ is $\Phi = \Phi_1 \times \Phi_2$, where $\Phi_1 \subset (0, \infty) \times [0, \infty)^2$ and $\Phi_2 \subset [0, 1] \times R \times (0, \infty)$.

2.1. Estimation of Gaussian mixture AVGARCH model

In this section, we study the MELE, which is a two-step estimated-likelihood method to estimate φ_0. First, conditional on the initial r.v.'s X_0 and $\tilde{\sigma}_0$, the residuals are obtained by using the Gaussian QMLE $\hat{\theta}_n = (\omega_n, \alpha_n, \beta_n)^T$ proposed by Pan *et al.* (2008) as follows:

$$\tilde{\epsilon}_t := \frac{X_t}{\tilde{\sigma}_t\left(\hat{\theta}_n\right)}, \quad t = 1, \ldots, n, \tag{2.2}$$

where $\tilde{\sigma}_t(\hat{\theta}_n)$ are defined recursively using

$$\tilde{\sigma}_t\left(\hat{\theta}_n\right) := \hat{\omega} + \hat{\alpha}|X_{t-1}| + \hat{\beta}\tilde{\sigma}_{t-1}\left(\hat{\theta}_n\right). \tag{2.3}$$

Initial r.v.'s for X_0 and $\tilde{\sigma}_0$ are often chosen as X_1 (*cf.* Francq and Zakoïan, 2004, p.608).

Then, we obtain the estimator for η_0 based on the residuals as $\tilde{\eta}_n := \arg\max_{\eta \in \Theta_2} \tilde{l}_n(\eta)$, where $\tilde{l}_n(\eta) := 1/n \sum_{t=1}^n \log f_\eta(\tilde{\epsilon}_t)$. It is worth noting that any $\sqrt{n}$-consistent estimator can be used to obtain the residuals. By plugging $\tilde{\eta}_n$ into the conditional likelihood $\tilde{l}_n(\theta, \eta)$ defined as

$$\tilde{l}_n(\theta, \eta) := \frac{1}{n} \log \tilde{L}_n(\theta, \eta) = \frac{1}{n} \sum_{t=1}^n \tilde{W}_t(\theta, \eta) \tag{2.4}$$

and $\tilde{W}_t(\theta, \eta) := \log\{(1/\sigma_t(\theta))f_\eta(X_t/\sigma_t(\theta))\}$, the estimated quasi-likelihood and the MELE for θ are obtained as follows:

$$\tilde{l}_n(\theta, \tilde{\eta}_n) := \frac{1}{n}\sum_{t=1}^{n} \tilde{W}_t(\theta, \tilde{\eta}_n) \tag{2.5}$$

and $\hat{\theta}_n^e := \arg\max_{\theta\in\Theta_1} \tilde{l}_n(\theta, \tilde{\eta}_n)$, respectively, where

$$\tilde{W}_t(\theta, \tilde{\eta}_n) := \log\left\{\frac{1}{\tilde{\sigma}_t(\theta)} f_{\tilde{\eta}_n}\left(\frac{X_t}{\tilde{\sigma}_t(\theta)}\right)\right\}.$$

Throughout this paper, it is assumed that all r.v.'s are defined on a probability space $(\Lambda, \mathcal{F}, P)$. The spectral radius of square matrix A is denoted by $\rho(A)$. Let $\mathcal{A}_\theta(z) = \alpha z$ and $\mathcal{B}_\theta(z) = 1 - \beta z$. In order to obtain the asymptotic properties of $\tilde{\eta}_n$ and $\hat{\theta}_n^e$, we consider the following regularity conditions:

(A1) φ_0 is an interior point of Φ and Φ is compact

(A2) $E[\log(\beta_0 + \alpha_0|\epsilon_t|] < 0$ and $\beta < 1$ for each $\theta \in \Phi_1$

(A3) If $\beta_0 \neq 0$, $\mathcal{A}_{\theta_0}(1) = \alpha_0 \neq 0$

Given below are the asymptotic properties for $\tilde{\eta}_n$ and $\hat{\theta}_n^e$. The proofs are given in Section 4.

Theorem 1. *Suppose that (A1)–(A3) hold. Then, $\tilde{\eta}_n \to \eta_0$ a.s. as $n \to \infty$.*

Theorem 2. *Suppose that (A1)–(A3) hold. Then, $\hat{\theta}_n^e \to \theta_0$ a.s. as $n \to \infty$.*

Remark 1. For the strictly stationarity of AVGARCH models, the top Lyapunov exponent γ_{θ_0} at θ_0 is assumed to be strictly negative. However, the top Lyapunov exponent can be explicitly expressed for AVGARCH(1,1) models. It can be seen that $\gamma_{\theta_0} = E[\log(\beta_0 + \alpha_0|\epsilon_t|]$ by following the arguments (for example) in Section 2 of Lee and Noh (2013). Thus, in (A2), the strictly stationary condition $\gamma_{\theta_0} < 0$ is replaced by $E[\log(\beta_0 + \alpha_0|\epsilon_t|] < 0$.

2.2. Skewness of Gaussian mixture AVGARCH model

In this section, we derive the skewness of Gaussian mixture AVGARCH(1, 1) models. In the following, we denote the skewness of X_t by γ^X, which is called the overall skewness. First, one can see from elementary calculations that

$$E[\sigma_t] = \frac{\omega}{1 - \alpha E|\epsilon_t| - \beta},$$

$$E\left[\sigma_t^2\right] = \frac{\omega^2 + 2\omega\alpha E|\epsilon_t| E[\sigma_t] + 2\omega\beta E[\sigma_t]}{1 - \beta^2 - \alpha^2 - 2\alpha\beta E|\epsilon_t|},$$

and

$$E\left[\sigma_t^3\right] = \frac{\omega^3 + 3\omega^2\alpha E|\epsilon_t| E[\sigma_t] + 3\omega\alpha^2 E\left[\sigma_t^2\right] + 3\omega^2\beta E[\sigma_t] + 6\omega\alpha E|\epsilon_t| E\left[\sigma_t^2\right] + 3\omega\beta^2 E\left[\sigma_t^2\right]}{1 - \alpha^3 E|\epsilon_t|^3 - 3\alpha^2\beta - 3\alpha\beta^2 E|\epsilon_t| - \beta^3}.$$

Table 1: Descriptive statistics for log returns of KOSPI200 and KOSDAQ indices

	KOSPI200	KOSDAQ
Mean	0.0006	0.0001
Standard deviation	0.0158	0.0164
Maximum	0.1154	0.1086
Minimum	−0.1090	−0.1103
Skewness	−0.4131	−1.0778
Kurtosis	8.1377	10.4598
Jarque-Bera	2256.5740	5024.5250
(P-value)	(0.0000)	(0.0000)

Table 2: QMLE estimates for AVGARCH(1, 1) model

Series	ω	α	β
KOSPI200	0.0015	0.1644	0.7744
KOSDAQ	0.0013	0.2426	0.7336

Table 3: MELE estimates for Gaussian mixture AVGARCH(1, 1) model

Series	ω	α	β	π	μ_1	μ_2	σ_1	σ_2
KOSPI200	0.0015	0.1558	0.7845	0.4742	−0.1894	0.1708	1.2834	0.5928
KOSDAQ	0.0012	0.2267	0.7517	0.2394	−0.7355	0.2314	1.4274	0.6703

Table 4: Estimates of skewness for Gaussian mixture AVGARCH(1, 1) model

Series	Overall skewness	Sample skewness
KOSPI200	−0.3815	−0.4131
KOSDAQ	−1.2743	−1.0778

Then, using $E[X_t^2] = E[\epsilon_t^2]E[\sigma_t^2] = E[\sigma_t^2]$ and $E[X_t^3] = E[\epsilon_t^3]E[\sigma_t^3]$, we get

$$\gamma_1^X = \frac{E\left[X_t^3\right]}{\left(E\left[X_t^2\right]\right)^{\frac{3}{2}}} = \frac{E\left[\epsilon_t^3\right]E\left[\sigma_t^3\right]}{\left(E\left[\sigma_t^2\right]\right)^{\frac{3}{2}}}. \tag{2.6}$$

By replacing the parameters in (2.6) with the MELE, $\tilde{\eta}_n$ and $\hat{\theta}_n^e$, in Section 2.1, we can obtain the estimator $\hat{\gamma}^X$ for the overall skewness γ^X. This estimator will be used for the real data analysis in Section 3.

3. Real Data Analysis

In this section, we applied the proposed estimation method and calculated the overall skewness for KOSPI200 (Korea Stock Price Index 200) and KOSDAQ(Korea Securities Dealers Automated Quotation) indices from December 13, 2002 to December 30, 2010. Both the KOSPI 200 and KOSDAQ indices contained 2000 observations. Table 1 provides descriptive statistics for the log returns of KOSPI 200 and KOSDAQ indices. The skewness and kurtosis estimates show that the log returns series are negatively skewed and leptokurtic. Moreover, it is seen from the Jarque-Bera normality test that the log return series are not normally distributed.

The Gaussian quasi-maximum likelihood (QMLE) and maximum estimated-likelihood estimates (MELE) are given in Table 2 and Table 3, respectively. The estimated overall skewness of Gaussian mixture AVGARCH(1, 1) models for KOSPI200 and KOSDAQ data are obtained by applying the results in Section 2.2 and the estimates given in Table 3 and shown in Table 4, together with the

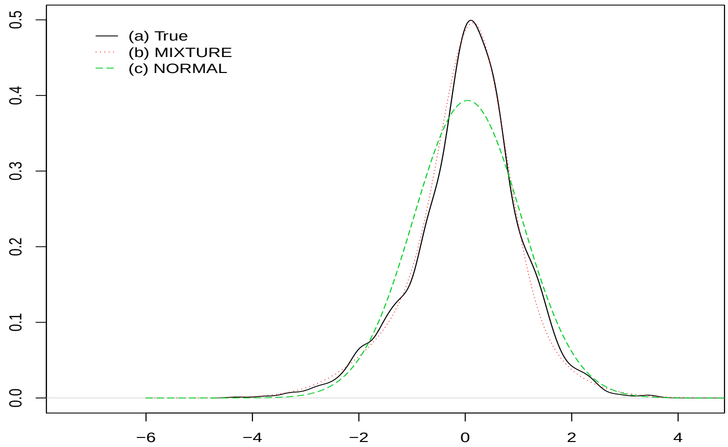

Figure 1: *Densities of KOSPI200 residuals; (a) Kernel; (b) Mixture; (c) Normal*

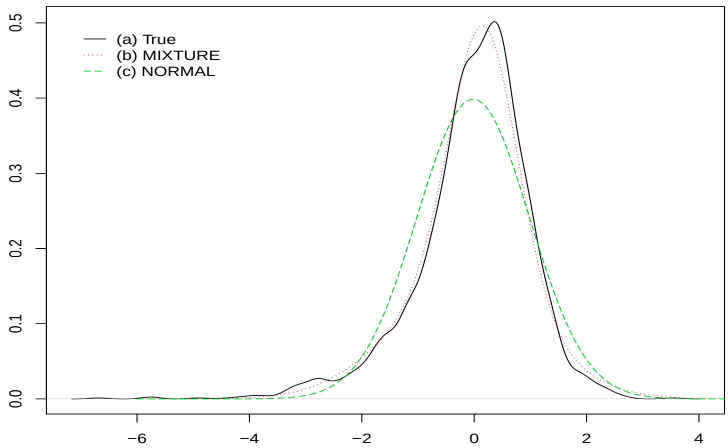

Figure 2: *Densities of KOSDAQ residuals; (a) Kernel; (b) Mixture; (c) Normal*

sample skewness. Figure 1 and Figure 2 show the shape of the estimated kernel, standard normal and Gaussian mixture densities of the AVGARCH(1, 1) residuals. From the Tables and Figures, we have several important findings. First, it is observed that the deviation between the QMLE and MELE of KOSDAQ index is more distinctive than the KOSPI200 index. This may be because the skewness of KOSDAQ index to the left is more distinctive as seen in Table 4 and Figure 1 and Figure 2. Next, the estimated overall skewness and the sample skewness in Table 3 are found to be very close to each other. This may imply that Gaussian mixture AVGARCH(1, 1) models are successful to capture the skewness of KOSPI200 and KOSDAQ data. Finally, the estimated densities in Figure 1 and Figure 2 indicate that Gaussian mixture density is more appropriate to capture the left-skewed distributions of AVGARCH(1, 1) residuals than the standard normal distribution. The findings strongly suggest that Gaussian mixture models for AVGARCH(1, 1) innovations are more adequate to reflect the left-skewness of KOSPI200 and KOSDAQ data that provide a better fit to skewed financial time series data.

4. Proofs

In this section, we provide the proofs for the theorems in Section 2. The following lemma is from Ha and Lee (2011).

Lemma 1. *Under the condition* **(A1)***,*

$$\sup_{\eta\in\Omega}\left|\frac{\partial}{\partial y}\log f_\eta(y)\right| \le C\left(|y|+1\right), \tag{4.1}$$

$$\sup_{\eta\in\Omega}\left|\frac{\partial}{\partial \eta}\log f_\eta(y)\right| \le C\left(|y|+1\right). \tag{4.2}$$

Lemma 2. *Let σ_t and $\tilde{\sigma}_t$ be the symbols in (2.1) and (2.3), respectively. Suppose that the conditions (A1) and (A2) hold. Then, there exists $\rho\in(0,1)$ such that, for all $t\ge 1$ and $0\le i\le 1$,*

$$\sup_{\theta\in\Phi_1}|\sigma_t(\theta)-\tilde{\sigma}_t(\theta)| \le C\rho^t. \tag{4.3}$$

Proof: It can be similarly done as in the proof of Theorem 2.1 of Francq and Zakoïan (2004). □

Below, two lemmas are introduced. The first lemma is from Corollary 1 in Lee and Lee (2011) and the second lemma is done in Theorem 3 of Leroux (1992).

Lemma 3. *Let $\hat{\theta}_n$ be the Gaussian QMLE in (2.2). Suppose that the conditions (A1)–(A3) hold. Then, we have, $\hat{\theta}_n\to\theta_0$ a.s., as $n\to\infty$.*

Lemma 4. *Let $\hat{\eta}_n := \arg\max_{\eta\in\Omega} l_n^*(\eta)$, where $l_n^*(\eta) := 1/n\sum_{t=1}^n \log f_\eta(\epsilon_t)$. Then, $\hat{\eta}_n\to\eta_0$ a.s., as $n\to\infty$.*

Lemma 5. *Let $\tilde{\epsilon}_t$ be the symbol in (2.2). Suppose that the conditions (A1)–(A3) hold. Then, we have, as $n\to\infty$,*

$$\frac{1}{n}\sum_{t=1}^n|\tilde{\epsilon}_t-\epsilon_t| \to 0 \quad a.s.. \tag{4.4}$$

$$\frac{1}{n}\sum_{t=1}^n|\epsilon_t||\tilde{\epsilon}_t-\epsilon_t| \to 0 \quad a.s.. \tag{4.5}$$

$$\frac{1}{n}\sum_{t=1}^n|\tilde{\epsilon}_t||\tilde{\epsilon}_t-\epsilon_t| \to 0 \quad a.s.. \tag{4.6}$$

Proof: Since (4.5) and (4.6) can be similarly proved as (4.7), we only deal with (4.4). First, we have

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^n|\tilde{\epsilon}_t-\epsilon_t| &= \frac{1}{n}\sum_{t=1}^n\left|\frac{X_t}{\tilde{\sigma}_t\left(\hat{\theta}_n\right)}-\epsilon_t\right| \le \frac{1}{n}\sum_{t=1}^n|\epsilon_t|\left|\frac{\tilde{\sigma}_t\left(\hat{\theta}_n\right)-\sigma_t(\theta_0)}{\tilde{\sigma}_t\left(\hat{\theta}_n\right)}\right| \\
&\le C\frac{1}{n}\sum_{t=1}^n|\epsilon_t|\left|\frac{\tilde{\sigma}_t\left(\hat{\theta}_n\right)-\tilde{\sigma}_t(\theta_0)}{\tilde{\sigma}_t\left(\hat{\theta}_n\right)}\right| + C\frac{1}{n}\sum_{t=1}^n|\epsilon_t|\left|\tilde{\sigma}_t(\theta_0)-\sigma_t(\theta_0)\right| \\
&\le C\frac{1}{n}\sum_{t=1}^n|\epsilon_t|\left|\frac{\tilde{\sigma}_t\left(\hat{\theta}_n\right)-\tilde{\sigma}_t(\theta_0)}{\tilde{\sigma}_t\left(\hat{\theta}_n\right)}\right| + C\frac{1}{n}\sum_{t=1}^n|\epsilon_t|\rho^t.
\end{aligned}$$

Then, it suffices to show that

$$\frac{1}{n}\sum_{t=1}^{n}|\epsilon_t|\left|\frac{\tilde{\sigma}_t\left(\hat{\theta}_n\right)-\tilde{\sigma}_t(\theta_0)}{\tilde{\sigma}_t\left(\hat{\theta}_n\right)}\right| \to 0 \quad a.s. \tag{4.7}$$

$$\frac{1}{n}\sum_{t=1}^{n}|\epsilon_t|\rho^t \to 0 \quad a.s.. \tag{4.8}$$

Now, we have

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}|\epsilon_t|\left|\frac{\tilde{\sigma}_t\left(\hat{\theta}_n\right)-\tilde{\sigma}_t(\theta_0)}{\tilde{\sigma}_t\left(\hat{\theta}_n\right)}\right| &\le \frac{1}{n}\sum_{t=1}^{n}|\epsilon_t|\left|\frac{\hat{\omega}_n-\omega_0}{\tilde{\sigma}_t\left(\hat{\theta}_n\right)}\right| + \frac{1}{n}\sum_{t=1}^{n}|\epsilon_t|\left|\frac{X_{t-1}}{\tilde{\sigma}_t\left(\hat{\theta}_n\right)}\right||\hat{\alpha}_n-\alpha_0| \\
&\quad + \frac{1}{n}\sum_{t=1}^{n}|\epsilon_t|\left|\frac{\tilde{\sigma}_{t-1}\left(\hat{\theta}_n\right)}{\tilde{\sigma}_t\left(\hat{\theta}_n\right)}\right|\left|\hat{\beta}_n-\beta_0\right| + \frac{1}{n}\sum_{t=1}^{n}|\epsilon_t|\beta_0\left|\frac{\tilde{\sigma}_{t-1}\left(\hat{\theta}_n\right)-\tilde{\sigma}_{t-1}(\theta_0)}{\tilde{\sigma}_t\left(\hat{\theta}_n\right)}\right| \\
&\le C\,|\hat{\omega}_n-\omega_0|\,\frac{1}{n}\sum_{t=1}^{n}|\epsilon_t| + C\frac{1}{\hat{\alpha}_n}\,|\hat{\alpha}_n-\alpha_0|\,\frac{1}{n}\sum_{t=1}^{n}|\epsilon_t| \\
&\quad + C\frac{1}{\hat{\beta}_n}\left|\hat{\beta}_n-\beta_0\right|\frac{1}{n}\sum_{t=1}^{n}|\epsilon_t| + C\frac{1}{\hat{\beta}_n}\frac{1}{n}\sum_{t=1}^{n}|\epsilon_t|\left|\frac{\tilde{\sigma}_{t-1}\left(\hat{\theta}_n\right)-\tilde{\sigma}_{t-1}(\theta_0)}{\tilde{\sigma}_{t-1}\left(\hat{\theta}_n\right)}\right| \\
&= R_{1n} + R_{2n}\frac{1}{n}\sum_{t=2}^{n}|\epsilon_t|\left|\frac{\tilde{\sigma}_{t-1}\left(\hat{\theta}_n\right)-\tilde{\sigma}_{t-1}(\theta_0)}{\tilde{\sigma}_{t-1}\left(\hat{\theta}_n\right)}\right| \\
&\;\;\vdots \\
&= R'_{1n} + R'_{2n}\frac{1}{n}\,|\epsilon_t|\left|\frac{\tilde{\sigma}_1\left(\hat{\theta}_n\right)-\tilde{\sigma}_1(\theta_0)}{\tilde{\sigma}_1\left(\hat{\theta}_n\right)}\right| \\
&\le R'_{1n} + R'_{2n}\frac{1}{n}\,|\epsilon_t|\left\{|\hat{\omega}_n-\omega_0| + |X_0|\,|\hat{\alpha}_n-\alpha_0|\right\},
\end{aligned}$$

where $R_{1n} \to 0$ *a.s.*, $R'_{1n} \to 0$ *a.s.*, $R_{2n} \to C$ *a.s.* and $R'_{2n} \to C$ *a.s.* and thus, (4.7) is established. Next, (4.8) is obtained due to the Cesàro lemma and the fact that $|\epsilon_t|\rho^t \to 0$ almost surely. Hence, the lemma is verified. □

Proof of Theorem 1: The theorem can be easily proven by using Lemma 1–Lemma 5 as the proof of Theorem 1 of Ha and Lee (2011). The details are omitted for brevity. □

Next, the following lemma is necessary to prove Theorem 2.

Lemma 6. *Suppose that the conditions (A1) and (A2) hold. Then, as $n \to \infty$,*

$$n^{-1}\sum_{t=1}^{n}\sup_{\theta\in\Phi_1}\left|\log\left(\frac{\sigma_t(\theta)}{\tilde{\sigma}_t(\theta)}\right)\right| \;\to\; 0 \;\; a.s.$$

Proof: We simply have

$$\left|\log\left(\frac{\sigma_t(\theta)}{\tilde{\sigma}_t(\theta)}\right)\right| \le \max\left\{\frac{1}{\sigma_t(\theta)}, \frac{1}{\tilde{\sigma}_t(\theta)}\right\} |\sigma_t(\theta) - \tilde{\sigma}_t(\theta)|$$

and thus,

$$\begin{aligned} n^{-1}\sum_{t=1}^{n}\sup_{\theta\in\Phi_1}\left|\log\left(\frac{\sigma_t(\theta)}{\tilde{\sigma}_t(\theta)}\right)\right| &\le n^{-1}\sum_{t=1}^{n}\sup_{\theta\in\Phi_1}\left\{\max\left\{\frac{1}{\sigma_t(\theta)}, \frac{1}{\tilde{\sigma}_t(\theta)}\right\}|\sigma_t(\theta) - \tilde{\sigma}_t(\theta)|\right\} \\ &\le \left\{\sup_{\theta\in\Phi_1}\frac{1}{\omega}\right\} n^{-1}\sum_{t=1}^{n}\sup_{\theta\in\Phi_1}|\sigma_t(\theta) - \tilde{\sigma}_t(\theta)| \\ &\le Cn^{-1}\sum_{t=1}^{n}\rho^t, \end{aligned}$$

where the last inequality follows from (A1) and Lemma 2. Since $\rho \in (0, 1)$ due to (A1) and the second part of (A2), the lemma is established. □

Let $\tilde{l}_n(\theta, \tilde{\eta}_n^*)$ be the symbol in (2.5) and define $l_n(\theta, \tilde{\eta}_n^*) := 1/n \sum_{t=1}^{n} W_t(\theta, \tilde{\eta}_n^*)$ and

$$W_t(\theta, \tilde{\eta}_n^*) := \log\left\{\frac{1}{\sigma_t(\theta)} f_{\tilde{\eta}_n^*}\left(\frac{X_t}{\sigma_t(\theta)}\right)\right\}.$$

Lemma 7. *Suppose that the conditions (A1) and (A2). Then, as $n \to \infty$,*

$$\lim_{n\to\infty}\sup_{\theta\in\Phi_1}\left|\tilde{l}_n(\theta, \tilde{\eta}_n^*) - l_n(\theta, \tilde{\eta}_n^*)\right| = 0 \ \ a.s..$$

Proof: Note that we have

$$\begin{aligned} \left|\tilde{l}_n(\theta, \tilde{\eta}_n^*) - l_n(\theta, \tilde{\eta}_n^*)\right| &\le \frac{1}{n}\sum_{t=1}^{n}\left|\log\left\{\frac{1}{\tilde{\sigma}_t(\theta)} f_{\tilde{\eta}_n^*}\left(\frac{X_t}{\tilde{\sigma}_t(\theta)}\right)\right\} - \log\left\{\frac{1}{\sigma_t(\theta)} f_{\tilde{\eta}_n^*}\left(\frac{X_t}{\sigma_t(\theta)}\right)\right\}\right| \\ &\le \frac{1}{n}\sum_{t=1}^{n}\left|\log\left\{\frac{\sigma_t(\theta)}{\tilde{\sigma}_t(\theta)}\right\}\right| + \frac{1}{n}\sum_{t=1}^{n}\left|\log\left\{f_{\tilde{\eta}_n^*}\left(\frac{X_t}{\tilde{\sigma}_t(\theta)}\right)\right\} - \log\left\{f_{\tilde{\eta}_n^*}\left(\frac{X_t}{\sigma_t(\theta)}\right)\right\}\right|. \end{aligned}$$

Therefore, it suffices to show that

$$n^{-1}\sum_{t=1}^{n}\sup_{\theta\in\Phi_1}\left|\log\left(\frac{\sigma_t(\theta)}{\tilde{\sigma}_t(\theta)}\right)\right| \to 0 \ \text{a.s.} \tag{4.9}$$

$$n^{-1}\sum_{t=1}^{n}\sup_{\theta\in\Phi_1}\left|\log\left\{f_{\tilde{\eta}_n^*}\left(\frac{X_t}{\tilde{\sigma}_t(\theta)}\right)\right\} - \log\left\{f_{\tilde{\eta}_n^*}\left(\frac{X_t}{\sigma_t(\theta)}\right)\right\}\right| \to 0 \ \text{a.s.} \tag{4.10}$$

First, we prove (4.9). We simply have

$$\left|\log\left(\frac{\sigma_t(\theta)}{\tilde{\sigma}_t(\theta)}\right)\right| \le \max\left\{\frac{1}{\sigma_t(\theta)}, \frac{1}{\tilde{\sigma}_t(\theta)}\right\} |\sigma_t(\theta) - \tilde{\sigma}_t(\theta)|$$

and thus,

$$n^{-1}\sum_{t=1}^{n}\sup_{\theta\in\Phi_1}\left|\log\left(\frac{\sigma_t(\theta)}{\tilde{\sigma}_t(\theta)}\right)\right| \le n^{-1}\sum_{t=1}^{n}\sup_{\theta\in\Phi_1}\left\{\max\left\{\frac{1}{\sigma_t(\theta)},\frac{1}{\tilde{\sigma}_t(\theta)}\right\}|\sigma_t(\theta)-\tilde{\sigma}_t(\theta)|\right\}$$
$$\le \left\{\sup_{\theta\in\Phi_1}\frac{1}{\omega}\right\} n^{-1}\sum_{t=1}^{n}\sup_{\theta\in\Phi_1}|\sigma_t(\theta)-\tilde{\sigma}_t(\theta)|$$
$$\le Cn^{-1}\sum_{t=1}^{n}\rho^t,$$

where the last inequality follows from (A1) and Lemma 2. Since $\rho \in (0,1)$ due to (A1) and the second part of (A2), (4.9) is established.

Next, we deal with (4.10). Using Lemma 1 we get,

$$\left|\log\left\{f_{\tilde{\eta}_n^*}\left(\frac{X_t}{\tilde{\sigma}_t(\theta)}\right)\right\}-\log\left\{f_{\tilde{\eta}_n^*}\left(\frac{X_t}{\sigma_t(\theta)}\right)\right\}\right| \le \left|\frac{X_t}{\tilde{\sigma}_t(\theta)}-\frac{X_t}{\sigma_t(\theta)}\right|\left(\max\left\{\frac{|X_t|}{\tilde{\sigma}_t(\theta)},\frac{|X_t|}{\sigma_t(\theta)}\right\}+1\right)$$
$$\le C\,|\tilde{\sigma}_t(\theta)-\sigma_t(\theta)|\,(|X_t|+1)^2\,,$$

and therefore,

$$\frac{1}{n}\sum_{t=1}^{n}\sup_{\theta\in\Phi_1}\left|\log\left\{f_{\tilde{\eta}_n^*}\left(\frac{X_t}{\tilde{\sigma}_t(\theta)}\right)\right\}-\log\left\{f_{\tilde{\eta}_n^*}\left(\frac{X_t}{\sigma_t(\theta)}\right)\right\}\right| \le C\frac{1}{n}\sum_{t=1}^{n}\left\{\sup_{\theta\in\Phi_1}|\tilde{\sigma}_t(\theta)-\sigma_t(\theta)|\right\}(|X_t|+1)^2\,,$$
$$\le C\frac{1}{n}\sum_{t=1}^{n}\rho^t\,(|X_t|+1)^2\,.$$

Then, (4.10) is easily proved similarly as (4.8). □

The following two lemmas can be similarly proved as in the proof of Lemma 10–11 of Ha and Lee (2011).

Lemma 8. *Suppose that (A1)–(A3) hold. Let* $W_t(\theta,\eta) := \log\{(1/\sigma_t(\theta))f_\eta(X_t/\sigma_t(\theta))\}$. *Then,*

$$E\left(\sup_{(\theta,\eta)\in\Phi} W_t(\theta,\eta)\right) < \infty$$

and for any $(\theta,\eta)\neq(\theta_0,\eta_0)$*,*

$$E\left[W_t(\theta,\eta)\right] < E\left[W_t(\theta_0,\eta_0)\right].$$

Lemma 9. *Suppose that (A1)–(A3) hold. Any* $\theta\,(\neq\theta_0)$ *has a neighborhood* $N(\theta)$ *such that*

$$\lim_{n\to\infty}\sup_{\theta'\in N(\theta)}\tilde{l}_n(\theta,\tilde{\eta}_n^*) < EW_t(\theta_0,\eta_0)\ \ a.s..$$

Proof of Theorem 2: Similar to the proof of Theorem 2 of Ha and Lee (2011), the consistency of the MELE can be obtained by using Theorem 1 and Lemmas 6–9. The details are omitted for brevity. □

References

Ding, Z. and Granger, C. W. J. (1996). Modeling volatility persistence of speculative returns: A new approach, *Journal of Econometrics*, **73**, 185–215.

Ding, Z., Granger, C. W. J. and Engle, R. F. (1993). A long memory property of stock market returns and a new model, *Journal of Empirical Finance*, **1**, 83–106.

Francq, C. and Zakoïan, J. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes, *Bernoulli*, **10**, 605–637.

Gonçalves, E., Leite, J. and Mendes-Lopez, N. (2009). A mathematical approach to detect the Taylor property in TARCH processes, *IEEE Signal Processing Letters*, **79**, 602–610.

Ha, J. and Lee, T. (2011). NM-QELE for ARMA-GARCH Models with non-Gaussian innovations, *Statistics and Probability Letters*, **81**, 694–703.

Haas, M. (2009). Persistence in volatility, conditional kurtosis, and the Taylor property in absolute value GARCH processes, *IEEE Signal Processing Letters*, **79**, 1674–1683.

Haas, M., Mittnik, S. and Paolella, M. S. (2004). Mixed normal conditional heteroskedasticity, *Journal of Financial Econometrics*, **2**, 211–250.

He, C. and Teräsvirta, T. (1999). Properties of moments of a family of GARCH processes, *Journal of Econometrics*, **92**, 173–192.

Lee, S. and Lee, T. (2012). Inference for Box-Cox transformed threshold GARCH models with nuisance parameters, *Scandinavian Journal of Statistics*, **39**, 568–589.

Lee, S. and Noh, J. (2013). An empirical study on explosive volatility test with possibly nonstationary GARCH(1, 1) models, *Communications for Statistical Applications and Methods*, **20**, 207–215.

Lee, S., Park, S. and Lee, T. (2009). A note on the Jarque-Bera Normality Test for GARCH Innovations, *Journal of the Korean Statistical Society*, **39**, 93–102.

Lee, T. and Lee, S. (2009). Normal mixture quasi-maximum Likelihood estimator for GARCH Models, *Scandinavian Journal of Statistics*, **36**, 157–170.

Leroux, B. G. (1992). Consistent estimation of a mixing distribution, *The Annals of Statistics*, **20**, 1350–1360.

McLachlan, G. and Peel, D. (2000). *Finite Mixture Models*, Wiley, New York.

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

Taylor, S. J. (1988). *Modelling Financial Time Series*, John Wiley and Sons, Chichester.

Received June 26, 2013; Revised July 29, 2013; Accepted August 7, 2013