

Article

A Time-Varying Coefficient Double Threshold GARCH Model with Explanatory Variables

Tongwei Zhang ¹, Lianyan Fu ^{2,*} , Dehui Wang ^{2,*}  and Zhuoxi Yu ²¹ School of Economics, Liaoning University, Shenyang 110036, China² School of Mathematics and Statistics, Liaoning University, Shenyang 110036, China

* Correspondence: fulianyan@lnu.edu.cn (L.F.); wangdehui@lnu.edu.cn (D.W.)

Abstract: In this article, we consider the nonparametric inference for the time-varying coefficient double-threshold generalized autoregressive conditional heteroscedastic models. The quasi-maximum exponential likelihood estimators (QMELEs) of the model's parameters and the asymptotic properties of the estimators are obtained. The simulation study implies that the distribution of the estimators is asymptotically normal. A real data application to stock returns is given. Both the simulations and real data example imply that the model and the QMELE are proper, compatible and accurately fit the financial time series data of the Nikkei 225.

Keywords: time-varying coefficient; threshold model; QMELE; asymptotic properties; financial time series; explanatory variables

MSC: 62P20; 62F99



Citation: Zhang, T.; Fu, L.; Wang, D.; Yu, Z. A Time-Varying Coefficient Double Threshold GARCH Model with Explanatory Variables. *Axioms* **2023**, *12*, 476. <https://doi.org/10.3390/axioms12050476>

Academic Editor: Valery Y. Glizer

Received: 25 March 2023

Revised: 5 May 2023

Accepted: 11 May 2023

Published: 15 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The random coefficient autoregressive (RCA) model proposed by [1] is a nonlinear time series models that is widely used when analyzing data generated on economics and finance. The RCA model incorporates a random variable in its autoregressive coefficient part, and thus makes the model able to capture the random influential factors. Refs. [2–5] studied various problems of inference for non-linear time-series models, including random coefficient models. Subsequently, ref. [6] proposed the generalized random coefficient autoregressive (GRCA) model. Refs. [7,8] studied properties of the RCA model with order 1 (the RCA(1) model). Ref. [9] studied the RCA model with GARCH innovations and its properties.

When analyzing the economic and financial time series, researchers found that the data frequently exhibit apparent evidence of asymmetry. For example, when the stock price is rising (in a bull market), the return yield (volatility) is different from the relevant yield when the price is falling (in a bear market). Refs. [10,11] proposed the threshold autoregressive model (TAR) to capture the asymmetric features shown in the financial time series. By using the threshold variable, the TAR model set up several different regimes with different parameters to describe asymmetric phenomena. Since [12] proposed the autoregressive conditional heteroscedastic (ARCH) model and [13] extended the ARCH models to the generalized autoregressive conditional heteroscedastic (GARCH) models, ref. [14] implied that the asymmetric volatility is exhibited in financial series data obtained from the stock market. Ref. [15] extended the TAR model to the model with conditional heteroscedasticity, which is the self-exciting threshold autoregressive ARCH (SETAR-ARCH) model. Then, ref. [16] proposed the double-threshold autoregressive conditional heteroscedastic (DTARCH) model, which captures asymmetry in the conditional mean and conditional variance at the same time, and used their model to analyze the Hang Seng Index return series in Hong Kong. Ref. [17] extended the DTARCH model to the DTGARCH model and used it to investigate the French Franc/Deutschmark exchange rates. Based on these

achievements, several researchers conducted further studies of applications of this model in finance and economics, e.g., [18–22].

In the literature related to the RCA model, the random coefficient is usually affected by an independent and identically distributed random variable. In [1], the random coefficient part is influenced by a white noise sequence. However, in practical application, the random coefficients may depend on some observed data and external factors. For instance, when analyzing the stock data, the return of yield at time t should be influenced by the one at time $t - 1$ and some factors external to the stock market. Ref. [23] proposed the smooth transition autoregressive (STAR) model, and the autoregressive parameters of STAR model incorporated the influence of the observations. Then, ref. [24] used the smooth transition autoregressive (STAR) model to analyze the data of industrial production in West Germany. Since then, STAR models have widely been used in economics and finance. When analyzing the economic and financial data, the models are not only affected by observations but also some external factors. Therefore, incorporating the observations and explanatory variables in the model and influencing the model through the random coefficient part are considered. Refs. [25–27] proposed several integer-valued autoregressive models, in which the random coefficients were influenced by observations or explanatory variables. Motivated by the above literature, which shows different approaches of constructing time series models, we incorporate observations and explanatory variables into the autoregressive coefficient part, and investigate a time-varying coefficient double-threshold generalized autoregressive conditional heteroscedastic (TVCDT-GARCH) model.

Ref. [28] studied the estimation of the random coefficient AR(1) processes based on the least squares estimator. Ref. [8] studied the model parameter estimation of the RCA(1) model, and [9] discussed the RCA models with GARCH innovations and its parameter-estimation problem. Ref. [29] found that some parameters of the nonstationary RCA model cannot be estimated using quasi-maximum likelihood estimation (QMLE); then, ref. [30] presented the quasi-likelihood estimation for the stationary and nonstationary process of RCA(1) model. In general, when solving the problem related to the model parameter estimation of the G/ARCH type model, the QMLE is an approach that is frequently used and is discussed by several researchers, such as [31–34]. However, QMLE requires a finite fourth moment for the error term. Due to the heavy-tailed features of financial time series, this requirement is not easily satisfied in many practical applications, especially in economics and finance. Ref. [35] studied the models with heavy-tailed errors and showed that QMLE is not asymptotically normal. Then, ref. [36] proposed the quasi-maximum exponential likelihood estimators (QMELE) for the ARMA-GARCH model under a fractional moment condition of the error term without specification of the distribution of the error term. Consequently, QMELE exhibits superior performance when dealing with heavy-tailed data in practical applications. These considerations motivate us to study the QMELE of the TVCDT-GARCH model.

As financial data usually involve non-normality and fat tails, the double exponential distribution is better than the normal distribution for characterizing heavy-tailed data. Compared with traditional QMLE, the QMELE is a better approach to estimate model parameters. The time-varying coefficients of our model incorporate not only the information of observations but also the information of external factors. Hence, by using QMELE, we investigate inference for TVCDT-GARCH models, which provides practitioners with an attractive modeling alternative for fitting financial data.

The rest of this paper is organized as follows. In Section 2, we discuss the model, parameter estimation and the asymptotic results. Section 3 shows the results from some simulation studies. In Section 4, we use our results to analyze the real data of the Nikkei 225. Section 5 provides our conclusions and some extensions for further research. We put the proofs in Appendix A.

2. Model and Estimation

Ref. [1] proposed the classic random coefficient autoregressive (RCA) model as the following:

$$y_t = (\varphi + b_t)y_{t-1} + \epsilon_t, \quad t \in \mathbb{N}.$$

where (b_t, ϵ_t) are independent and identically distributed pairs of random variables and satisfy

$$E(b_t, \epsilon_t) = (0, 0), \quad \text{cov}(b_t, \epsilon_t) = \begin{pmatrix} \sigma_b^2 & 0 \\ 0 & \sigma_\epsilon^2 \end{pmatrix}.$$

Although the RCA models are widely used in biology, economics and finance, and studied in the literature, there are some drawbacks that can be improved. For instance, when analyzing stock data, the return of yield at time t should be influenced by the one at time $t - 1$ and some external factors; however, in the RCA model, the random coefficient is only affected by the random variable b_t . Based on the discussion above, we consider the following time-varying coefficient double threshold GARCH (TVCDT-GARCH) model:

$$y_t = \begin{cases} \sum_{i=1}^p \phi_{i,t}^{(1)} y_{t-i} + \epsilon_t, & y_{t-d} \leq r, \\ \sum_{i=1}^p \phi_{i,t}^{(2)} y_{t-i} + \epsilon_t, & y_{t-d} > r, \end{cases} \quad \epsilon_t = \eta_t \sqrt{h_t}, \quad (1)$$

$$h_t = \begin{cases} \omega_1 + \sum_{i=1}^q \alpha_{1i} \epsilon_{1t-i}^2 + \sum_{i=1}^s \beta_{1i} h_{1t-i}, & y_{t-d} \leq r, \\ \omega_2 + \sum_{i=1}^q \alpha_{2i} \epsilon_{2t-i}^2 + \sum_{i=1}^s \beta_{2i} h_{2t-i}, & y_{t-d} > r, \end{cases}$$

where

$$\ln \frac{\phi_{i,t}^{(j)}}{1 - \phi_{i,t}^{(j)}} = c_{j,i} y_{t-1} + d_j z_t,$$

where z_t represents observations of explanatory variables, $\{z_t\}$ is a stationary sequence, $j = \{1, 2\}$, $i = \{1, 2, \dots, n\}$, $d \leq p, q, s$ and d are positive integers, $\{\eta_t\}$ are independent and identically distributed noise with $E|\eta_t| = 1$, and r is a threshold parameter. Let $\theta = (c', \delta', r')'$ be the model parameters with true value $\theta_0 = (c'_0, \delta'_0, r'_0)'$, where $\delta = (\omega_1, \alpha_{11}, \dots, \alpha_{1q}, \beta_{11}, \dots, \beta_{1s}, \omega_2, \alpha_{21}, \dots, \alpha_{2q}, \beta_{21}, \dots, \beta_{2s})'$ and $c = (c_{11}, \dots, c_{1p}, d_1, c_{21}, \dots, c_{2p}, d_2)$. The parameter space is $\Theta = \Theta_c \times \Theta_\delta \times \Theta_r$, where $\Theta_c \subset R^{2(p+1)}$, $\Theta_\delta \subset R_0^{2(q+s+1)}$, $\Theta_r \subset R$, $R = (-\infty, \infty)$, $R_0 = [0, \infty)$. For the sake of representation, let $\lambda = (c', \delta')'$ denote the parameters in our model, except the threshold parameter r ; then, $\Theta_\lambda = \Theta_c \times \Theta_\delta$. Assume that Θ_c and Θ_δ are compact, and θ_0 is an interior point of Θ . We consider the following function:

$$L_n(\theta) = \sum_{t=1}^n l_t(\theta), \text{ with } l_t(\theta) = \left\{ \log \sqrt{h_{1t}(\theta)} + \frac{|\epsilon_{1t}|}{\sqrt{h_{1t}(\theta)}} \right\} I_{1t} + \left\{ \log \sqrt{h_{2t}(\theta)} + \frac{|\epsilon_{2t}|}{\sqrt{h_{2t}(\theta)}} \right\} I_{2t},$$

where $I_{1t} = I(y_{t-d} \leq r)$ and $I_{2t} = I(y_{t-d} > r)$. We look for $\hat{\theta}_n$, the minimizer of $L_n(\theta)$ on Θ , where

$$\hat{\theta}_n = \arg \min_{\Theta} L_n(\theta).$$

$\hat{\theta}_n$ is called the QMELE of θ_0 . To estimate $\hat{\theta}_n$, we have to fix the threshold parameter r , then minimize $L_n(\theta)$ on Θ_λ and finally derive its minimizer $\hat{\lambda}_n(r)$. In what follows, we assume that r is already known.

In the circumstances that r is unknown, we can use several methods to estimate the threshold parameter, such as the single grid search (SGS) algorithm (e.g., [37–39]) and the nested sub-sample search (NeSS) algorithm ([40]), etc. Researchers then can choose a proper method to locate the threshold parameter. Using the optimization algorithms to optimize the log-likelihood function is an alternative procedure to obtain the threshold parameter. Whether this estimator is proper or not requires judgement. We use the SGS method to estimate the threshold parameter r in our estimation process. When r is fixed, we can calculate the AIC of the model, and the optimal r is the one with the minimum model AIC. Using the SGS method, we search for r over the range of the 20th to the 80th percentile of the sample data with a search step size of 0.01 (1%), and then we obtain the estimated threshold $\hat{r}$ that corresponds to the minimum model AIC.

Remark 1. The least absolute deviation (LAD) estimation can also relax the requirement of the moment condition. However, we do not use the LAD method to estimate the parameters of our model. Remark 2.2 in [41] gives a detailed description, and thus we will not consider the LAD method henceforward.

Next, we discuss the asymptotic properties of our estimators. Let $A_j(z) = \sum_{i=1}^q \alpha_{ji} z^i$, and $B_j(z) = 1 - \sum_{i=1}^s \beta_{ji} z^i$. We need the following assumptions.

Assumption 1. $\{y_t\}$ satisfying Equation (1) is strictly stationary and ergodic. The parameter space Θ is compact. $\omega_j > 0$, $\alpha_{ji} > 0$, $\beta_{ji} > 0$, $\sum_{i=1}^q \alpha_{ji} + \sum_{i=1}^s \beta_{ji} < 1$ for $i = 1, 2, \dots$, and $j = 1, 2$.

Assumption 2. For each $\theta \in \Theta$, $A_j(z)$ and $B_j(z)$ have no common root, $A_j(1) \neq 0$, $\alpha_q + \beta_s \neq 0$, and we let $A_j = (\alpha_{j1}, \dots, \alpha_{jq})'$, $A_j \neq A_{j'}$ if $j \neq j'$. A similar condition holds for $B_j = (\beta_{j1}, \dots, \beta_{js})'$, $\Phi_j = (\phi_{j1}, \dots, \phi_{jp})'$ and ω_j .

Assumption 3. η_t^2 has a nondegenerate distribution with $E\eta_t^2 < \infty$. $\{\eta_t\}$ is independent and identically distributed with median 0, $E|\eta_t| = 1$; its density function $f(x)$ is continuous and satisfies $f(0) > 0$, $\sup_{x \in \mathbb{R}} f(x) < \infty$, and $\sup_{x \in \mathbb{R}} |f'(x)| < \infty$.

Assumption 4. $E[\xi_{\rho t-1}^3] < \infty$ for any $\rho \in (0, 1)$, where $\xi_{\rho t} = 1 + \sum_{i=1}^{\infty} \rho^i |y_{t-i}|$.

Assumption 5. $\sup_{(y,z) \in \mathbb{R}^2} |c_{j,i}y + d_jz| < \infty$, $j = \{1, 2\}$, $i = \{1, 2, \dots, p\}$, and $\{z_t\}$ is a stationary sequence and independent of $\{\epsilon_t\}$.

Remark 2. For the time-varying coefficients part, we notice that $0 < \phi_{i,t}^{(j)} = \frac{\exp(c_{j,i}y_{t-1} + d_jz_t)}{\exp(c_{j,i}y_{t-1} + d_jz_t) + 1} < 1$, $j = \{1, 2\}$, $i = \{1, 2, \dots, p\}$. With Assumptions 1–5 and similarly to Proposition 1.1 in [42], the parameters of the autoregressive part fit the strictly stationary and ergodic solution condition. Our model is a two-regime threshold model that can be seen as a special case of [43] that then combines the corresponding results of Theorem 1 in [43] with the restrictions on the parameters of the variance part: $\omega_j > 0$, $\alpha_{ji} > 0$, $\beta_{ji} > 0$, $\sum_{i=1}^q \alpha_{ji} + \sum_{i=1}^s \beta_{ji} < 1$. We can guarantee the stationarity of our model.

For the sake of clarity, let $\hat{\theta}_{sn}$ denote the minimizer of $L_n(\theta)$ on Θ_λ for fixed r . The following theorem gives the strong consistency of $\hat{\theta}_{sn}$.

Theorem 1. Suppose η_t has a zero median with $E|\eta_t| = 1$. If Assumptions 1–5 hold, then $\hat{\theta}_{sn} \rightarrow \theta_0$ a.s., as $n \rightarrow \infty$.

To investigate the rate of convergence of $\hat{\theta}_{sn}$, we rewrite the log-likelihood function as follows: $H_n(u, r) \equiv L_n(\lambda_0 + u, r) - L_n(\lambda_0, r)$, where $u \in \{u = (u'_1, u'_2)' : u + \lambda_0 \in \Theta_\lambda\}$ ($u'_i = (u_{i1}, u_{i2}, \dots, u_{i(p+q+s+2)})', i = 1, 2$). Let $\hat{u}_n(r) = \hat{\lambda}(r) - \lambda_0$; then, $\hat{u}_n(r)$ is the minimizer of $H_n(u, r)$ for fixed r . Note that

$$H_n(u, r) = H_n(u, r_0) + H_n(u, r) - H_n(u, r_0),$$

and

$$H_n(u, r_0) = L_n(\lambda_0 + u, r_0) - L_n(\lambda_0, r_0) = \sum_{i=1}^2 \tilde{L}_{in}(u_i, r_0),$$

where

$$\begin{aligned} \tilde{L}_{in}(u_i, r_0) &= \sum_{t=1}^n \left\{ \log \sqrt{h_{it}(\theta_0 + u)} + \frac{|\epsilon_{it}(\gamma_0 + u_i)|}{\sqrt{h_{it}(\theta_0 + u)}} \right. \\ &\quad \left. - \log \sqrt{h_{it}(\theta_0)} - \frac{|\epsilon_{it}(\gamma_0)|}{\sqrt{h_{it}(\theta_0)}} \right\} I_{it}. \end{aligned}$$

Now we divide $\tilde{L}_{in}(u_i, r_0)$ into three parts:

$$\tilde{L}_{in}(u_i, r_0) = \sum_{t=1}^n A_{it}(u_i, r_0) + \sum_{t=1}^n B_{it}(u_i, r_0) + \sum_{t=1}^n C_{it}(u_i, r_0),$$

where

$$\begin{aligned} A_{it}(u_i, r_0) &= \frac{1}{\sqrt{h_{it}(\theta_0)}} [|\epsilon_{it}(\gamma_0 + u_1)| - |\epsilon_{it}(\gamma_0)|] I_{it}, \\ B_{it}(u_i, r_0) &= \left\{ \log \sqrt{h_{it}(\theta_0 + u)} - \log \sqrt{h_{it}(\theta_0)} + \frac{|\epsilon_{it}(\gamma_0)|}{\sqrt{h_{it}(\theta_0 + u)}} - \frac{|\epsilon_{it}(\gamma_0)|}{\sqrt{h_{it}(\theta_0)}} \right\} I_{it}, \\ C_{it}(u_i, r_0) &= \left[\frac{1}{\sqrt{h_{it}(\theta_0 + u)}} - \frac{1}{\sqrt{h_{it}(\theta_0)}} \right] [|\epsilon_{it}(\gamma_0 + u_1)| - |\epsilon_{it}(\gamma_0)|] I_{it}. \end{aligned}$$

Let $\mathcal{F}_{t-1} = \sigma\{\eta_{t-1}, \eta_{t-2}, \dots\}$; using the equality

$$|x - y| - |x| = -y[I(x > 0) - I(x < 0)] + 2 \int_0^y [I(x \leq s) - I(x \leq 0)] ds,$$

we have

$$A_{it}(u_i, d_0) = s_{it}(u_i)[I(\eta_t > 0) - I(\eta_t < 0)] + 2 \int_0^{-s_{it}(u_i)} [I(\eta_t \leq s) - I(\eta_t \leq 0)] ds, \text{ where } s_{it}(u_i) = s_{it}^{(1)}(u_i) + s_{it}^{(2)}(u_i),$$

$$s_{it}^{(1)}(u_i) = \frac{u'_i I_{it}}{\sqrt{h_{it}(\theta_0)}} \frac{\partial \epsilon_{it}(\gamma_0)}{\partial \theta}, \quad s_{it}^{(2)}(u_i) = \frac{u'_i I_{it}}{2\sqrt{h_{it}(\theta_0)}} \frac{\partial^2 \epsilon_{it}(\gamma^*)}{\partial \theta \partial \theta'} u_i,$$

and γ^* lies between γ_0 and $\gamma_0 + u_i$. Let $\xi_{it}(u_i) = \int_0^{-s_{it}(u_i)} [I(\eta_t \leq s) - I(\eta_t \leq 0)] ds$, and note that $I(\eta_t > 0) - I(\eta_t < 0) = \text{sgn}(\eta_t)$; then, we have

$$\sum_{t=1}^n A_{it}(u_i, r_0) = u'_i T_{in}^{(1)} + \Pi_{in}^{(1)}(u_i) + \Pi_{in}^{(2)}(u_i) + \Pi_{in}^{(3)}(u_i),$$

where

$$\begin{aligned} T_{in}^{(1)} &= \sum_{t=1}^n \frac{I_{it}}{\sqrt{h_{it}(\theta_0)}} \frac{\partial \epsilon_{it}(\gamma_0)}{\partial \theta} \text{sgn}(\eta_t), \\ \Pi_{in}^{(1)}(u) &= 2 \sum_{t=1}^n \{ \xi_{it}(u_i) - E[\xi_{it}(u_i) | \mathcal{F}_{t-1}] \}, \\ \Pi_{in}^{(2)}(u) &= 2 \sum_{t=1}^n E[\xi_{it}(u_i) | \mathcal{F}_{t-1}], \text{ and} \\ \Pi_{in}^{(3)}(u) &= \sum_{t=1}^n s_{it}^{(2)}(u_i) \text{sgn}(\eta_t). \end{aligned}$$

By Taylor's expansion, we have

$$\sum_{t=1}^n B_{it}(u_i, r_0) = u' T_{in}^{(2)} + \Pi_{in}^{(4)}(u_i) + \Pi_{in}^{(5)}(u_i),$$

where

$$\begin{aligned} T_{in}^{(2)} &= \sum_{t=1}^n \frac{I_{it}}{2h_{it}(\theta_0)} \frac{\partial h_{it}(\theta_0)}{\partial \theta} (1 - |\eta_t|), \\ \Pi_{in}^{(4)}(u_i) &= u'_i \sum_{t=1}^n \left(\frac{3}{8} \left| \frac{\epsilon_{it}(\gamma_0)}{\sqrt{h_{it}(\theta^*)}} \right| - \frac{1}{4} \right) \frac{I_{it}}{h_{it}^2(\theta^*)} \frac{\partial h_{it}(\theta^*)}{\partial \theta} \frac{\partial h_{it}(\theta^*)}{\partial \theta'} u_i, \\ \Pi_{in}^{(5)}(u_i) &= u'_i \sum_{t=1}^n \left(\frac{1}{4} - \frac{1}{4} \left| \frac{\epsilon_{it}(\gamma_0)}{\sqrt{h_{it}(\theta^*)}} \right| \right) \frac{I_{it}}{h_{it}(\theta^*)} \frac{\partial^2 h_{it}(\theta^*)}{\partial \theta \partial \theta'} u_i, \end{aligned}$$

and θ^* lies between θ_0 and $\theta_0 + u$. We need three lemmas.

Lemma 1. Let $T_{in} = T_{in}^{(1)} + T_{in}^{(2)}$ for $i = 1, 2$. If Assumptions 1–5 hold, then

$$\frac{1}{\sqrt{n}} T_{in} \rightarrow_d N(0, \Omega_i) \quad \text{as } n \rightarrow \infty,$$

where

$$\Omega_i = \begin{pmatrix} E\left(\frac{I_{it}}{h_{it}(\theta_0)} \frac{\partial \epsilon_{it}(\gamma_0)}{\partial \theta} \frac{\partial \epsilon_{it}(\gamma_0)}{\partial \theta'}\right) & \frac{E\eta_t}{2} E\left(\frac{I_{it}}{[h_{it}(\theta_0)]^{\frac{3}{2}}} \frac{\partial \epsilon_{it}(\gamma_0)}{\partial \theta} \frac{\partial h_{it}(\theta_0)}{\partial \theta}\right) \\ \frac{E\eta_t}{2} E\left(\frac{I_{it}}{[h_{it}(\theta_0)]^{\frac{3}{2}}} \frac{\partial \epsilon_{it}(\gamma_0)}{\partial \theta'} \frac{\partial h_{it}(\theta_0)}{\partial \theta'}\right) & \frac{E\eta_t^2 - 1}{4} E\left(\frac{I_{it}}{h_{it}^2(\theta_0)} \frac{\partial h_{it}(\theta_0)}{\partial \theta} \frac{\partial h_{it}(\theta_0)}{\partial \theta'}\right) \end{pmatrix}.$$

Lemma 2. If Assumptions 1–5 hold, then for any sequence of random variables a_n such that $a_n = o_p(1)$, it follows that, for $i = 1, 2$,

$$\begin{aligned} (a) \quad & \Pi_{in}^{(1)}(a_n) = o_p(\sqrt{n}\|a_n\| + n\|a_n\|^2), \\ (b) \quad & \Pi_{in}^{(2)}(a_n) = (\sqrt{n}\|a_n\|)' \Sigma_i^{(1)}(\sqrt{n}\|a_n\|) + o_p(n\|a_n\|^2), \\ (c) \quad & \Pi_{in}^{(3)}(a_n) = o_p(n\|a_n\|^2), \\ (d) \quad & \Pi_{in}^{(4)}(a_n) = (\sqrt{n}\|a_n\|)' \Sigma_i^{(2)}(\sqrt{n}\|a_n\|) + o_p(n\|a_n\|^2), \\ (e) \quad & \Pi_{in}^{(5)}(a_n) = o_p(n\|a_n\|^2), \\ (f) \quad & \sum_{t=1}^n C_{it}(a_i, d_0) = o_p(n\|a_n\|^2), \end{aligned}$$

where

$$\begin{aligned}\Sigma_i^{(1)} &= f(0)E\left(\frac{I_{it}}{h_{it}(\theta_0)}\frac{\partial \epsilon_{it}(\gamma_0)}{\partial \theta}\frac{\partial \epsilon_{it}(\gamma_0)}{\partial \theta'}\right), \\ \Sigma_i^{(2)} &= \frac{1}{8}E\left(\frac{I_{it}}{h_{it}^2(\theta_0)}\frac{\partial h_{it}(\theta_0)}{\partial \theta}\frac{\partial h_{it}(\theta_0)}{\partial \theta'}\right).\end{aligned}$$

Lemma 3. *If Assumptions 1–5 hold, then*

$$\sup_{\|r-r_0\|\leq B/n}|H_n(\hat{u}_n(r), r) - H_n(\hat{u}_n(r), r_0)| = o_p\left(\sup_{\|r-r_0\|\leq B/n}\sqrt{n}\|\hat{u}_n(r)\|\right).$$

Then, we can establish the asymptotic normality of $\hat{\theta}_{sn}$.

Theorem 2. *If Assumptions 1–5 hold, then*

$$(1) \sqrt{n}(\hat{\theta}_{sn} - \theta_0) = O_p(1) \quad \text{as } n \rightarrow \infty,$$

$$(2) \sqrt{n}(\hat{\theta}_{sn} - \theta_0) \xrightarrow{d} N(0, \frac{1}{4}\Sigma_0^{-1}\Omega_0\Sigma_0^{-1}) \quad \text{as } n \rightarrow \infty,$$

where $\Sigma_0 = \text{diag}(\Sigma_1, \Sigma_2)$ and $\Omega_0 = \text{diag}(\Omega_1, \Omega_2)$ with $\Sigma_i = \text{diag}\{\Sigma_i^{(1)}, \Sigma_i^{(2)}\}$, for $i = 1, 2$.

The proofs of Lemmas and Theorems are given in Appendix A.

When we analyze the real data of the Nikkei 225, we calculate the asymptotic standard deviation using $\hat{\chi}_n = \frac{1}{4}\hat{\Sigma}_n^{-1}\hat{\Omega}_n\hat{\Sigma}_n^{-1}$, where $\hat{\Sigma}_n = \text{diag}(\hat{\Sigma}_1, \hat{\Sigma}_2)$ and $\hat{\Omega}_n = \text{diag}(\hat{\Omega}_1, \hat{\Omega}_2)$. For $i = 1, 2$, $\hat{\Sigma}_i$ and $\hat{\Omega}_i$ are estimated by

$$\begin{aligned}\hat{\Sigma}_{in} &= \text{diag}\left\{\hat{f}(0)\hat{\kappa}_{in}^{(1)}, \frac{1}{8}\hat{\kappa}_{in}^{(2)}\right\}, \\ \hat{\Omega}_{in} &= \begin{pmatrix} \hat{\kappa}_{in}^{(1)} & \frac{1}{2}\hat{\mu}\hat{\kappa}_{in}^{(3)} \\ \frac{1}{2}\hat{\mu}\hat{\kappa}_{in}^{(3)'} & \frac{1}{4}(\hat{v} - 1)\hat{\kappa}_{in}^{(2)} \end{pmatrix},\end{aligned}$$

where

$$\begin{aligned}\hat{\kappa}_{in}^{(1)} &= \frac{1}{n}\sum_{t=1}^n \frac{I_{it}}{h_{it}(\hat{\theta}_n)}\frac{\partial \epsilon_{it}(\hat{\gamma}_n)}{\partial \theta}\frac{\partial \epsilon_{it}(\hat{\gamma}_n)}{\partial \theta'}, \\ \hat{\kappa}_{in}^{(2)} &= \frac{1}{n}\sum_{t=1}^n \frac{I_{it}}{h_{it}^2(\hat{\theta}_n)}\frac{\partial h_{it}(\hat{\theta}_n)}{\partial \theta}\frac{\partial h_{it}(\hat{\theta}_n)}{\partial \theta'}, \\ \hat{\kappa}_{in}^{(3)} &= \frac{1}{n}\sum_{t=1}^n \frac{I_{it}}{[h_{it}(\hat{\theta}_n)]^{\frac{3}{2}}}\frac{\partial \epsilon_{it}(\hat{\gamma}_n)}{\partial \theta}\frac{\partial h_{it}(\hat{\theta}_n)}{\partial \theta'}, \\ \hat{\mu} &= \frac{1}{n}\sum_{t=1}^n \hat{\eta}_t, \\ \hat{v} &= \frac{1}{n}\sum_{t=1}^n \hat{\eta}_t^2.\end{aligned}$$

We can obtain $\hat{f}(0)$ from the following function:

$$\hat{f}(0) = \frac{1}{nb_n}\sum_{t=1}^n K\left(\frac{\hat{\eta}_t}{b_n}\right),$$

where $K(x)$ is the kernel function and b_n is the bandwidth. In the simulation study, we choose the normal kernel function and a same bandwidth with [44], $b = 1.06 \cdot sn^{-0.2}$, where s is the sample standard deviation. By replacing the true value θ_0 with our estimator $\hat{\theta}_{sn}$ in

theoretical asymptotic variance, we obtain the consistent asymptotic variance estimation. Researchers can flexibly use the results above to work on the studies they are interested in, such as constructing confidence intervals or test procedure to study some statistical test problems, based on their own requirements.

3. Simulation Studies

In this section, we compare the estimation performance results of QMELE and QMLE under the circumstances that the error terms follow different distributions. The data-generating processes (DGPs) are as follows.

DGP_1 :

$$y_t = \begin{cases} \phi_{1t} \cdot y_{t-1} + \eta_t \sqrt{0.3 + 0.2 \cdot \epsilon_{1t-1}^2 + 0.4 \cdot h_{1t-1}}, & y_{t-1} \leq 1, \\ \phi_{2t} \cdot y_{t-1} + \eta_t \sqrt{0.1 + 0.3 \cdot \epsilon_{2t-1}^2 + 0.5 \cdot h_{2t-1}}, & y_{t-1} > 1, \end{cases}$$

$$\begin{cases} \ln \frac{\phi_{1t}}{1-\phi_{1t}} = 0.8 \cdot y_{t-1}, & y_{t-1} \leq 1, \\ \ln \frac{\phi_{2t}}{1-\phi_{2t}} = 0.3 \cdot y_{t-1}, & y_{t-1} > 1, \end{cases}$$

DGP_2 :

$$y_t = \begin{cases} \phi_{1t} \cdot y_{t-1} + \eta_t \sqrt{0.1 + 0.5 \cdot \epsilon_{1t-1}^2 + 0.3 \cdot h_{1t-1}}, & y_{t-1} \leq 0, \\ \phi_{2t} \cdot y_{t-1} + \eta_t \sqrt{0.1 + 0.2 \cdot \epsilon_{2t-1}^2 + 0.6 \cdot h_{2t-1}}, & y_{t-1} > 0, \end{cases}$$

$$\begin{cases} \ln \frac{\phi_{1t}}{1-\phi_{1t}} = 0.7 \cdot y_{t-1} + 0.2 \cdot z_t, & y_{t-1} \leq 0, \\ \ln \frac{\phi_{2t}}{1-\phi_{2t}} = 0.3 \cdot y_{t-1} + 0.5 \cdot z_t, & y_{t-1} > 0, \end{cases}$$

where $z_t = 0.6z_{t-1} + \varepsilon_t$, $\varepsilon_t \sim N(0, 1)$.

The data of these DGPs can be generated by transforming the time-varying coefficient part into the equation of observations or explanatory variables, and these generating processes do not need the support of any complicated algorithm. We let $\eta_t \sim N(0, \pi/2)$, $Laplace(0, 1)$ and $\frac{\pi}{2\sqrt{3}}t_3$ to ensure $E|\eta_t| = 1$ in these DGPs, and replicate 1000 times in each process. Tables 1 and 2 present the simulation results for DGP_1 and DGP_2 respectively. In these tables, SD and ASD are the standard deviation and asymptotic standard deviation of the estimators, and bias equals the difference between the average of each estimator and the corresponding true value of the parameter. The ASDs of the parameter estimators of QMELE are calculated via $\hat{\Sigma}_{in}$ and $\hat{\Omega}_{in}$, which we demonstrated in Section 2.

From Tables 1 and 2, we can see that when the sample size n increases, the mean of QMELE gets closer to the corresponding true values, regardless of the distribution of η_t . As n increases, all of the bias, SD and ASD decrease. The simulation results imply that when η_t follows the normal distribution, the estimation results of standard QMLE are relatively normal but display greater biases than the corresponding results of QMELE. When η_t follows the Laplace distribution and the T-distribution, the performance of the standard QMLE is not desirable; inversely, QMELE has much better performance.

From the simulation results, we can clearly see that when the error terms follow different distributions, the QMELE are consistent and perform much better than standard QMLE. Especially when the distribution of the error term is not normal, the QMELE provides more precise results and shows obvious advantages.

Table 1. Estimation performance results of DGP_1 (QMELE vs. standard QMLE).

n	Parameters True Value	c_{11} 0.8	ω_1 0.3	α_{11} 0.2	β_{11} 0.4	c_{21} 0.3	ω_2 0.1	α_{21} 0.3	β_{21} 0.5	r 1	c_{11} 0.8	ω_1 0.3	α_{11} 0.2	β_{11} 0.4	c_{21} 0.3	ω_2 0.1	α_{21} 0.3	β_{21} 0.5	r 1	
QMELE											Standard QMLE									
$\eta_t \sim N(0, \frac{\pi}{2})$																				
300	Bias	0.1007	0.0481	0.0039	−0.0457	−0.0236	0.0365	−0.0077	−0.0184	0.0210	−0.0183	0.1870	0.1013	0.0089	−0.0263	0.0828	0.1431	−0.0126	−0.1083	
	SD	0.4195	0.1691	0.1211	0.2341	0.1584	0.1879	0.1231	0.1877	0.1306	0.3206	0.2072	0.1571	0.1966	0.1139	0.2533	0.1643	0.1717	0.1589	
	ASD	0.7650	0.1517	0.1401	0.1895	0.2446	0.1621	0.1020	0.1707	−	0.2697	0.1919	0.1278	0.1654	0.0588	0.2241	0.1436	0.1638	−	
600	Bias	0.0586	0.0395	0.0152	−0.0461	−0.0147	0.0258	−0.0075	−0.0136	0.0294	−0.0408	0.1681	0.1019	0.0171	−0.0134	0.0728	0.1464	−0.0036	−0.1193	
	SD	0.2860	0.1220	0.0784	0.1454	0.0955	0.1199	0.0837	0.1093	0.1049	0.2673	0.1523	0.1185	0.1424	0.0791	0.1588	0.1068	0.1061	0.1491	
	ASD	0.4960	0.1008	0.0682	0.1289	0.1112	0.1026	0.0733	0.1144	−	0.1953	0.1420	0.0965	0.1258	0.0407	0.1416	0.1034	0.1098	−	
900	Bias	0.0564	0.0328	0.0153	−0.0403	−0.0039	0.0219	−0.0008	−0.0176	0.0177	−0.0600	0.1651	0.1013	0.0205	−0.0103	0.0679	0.1444	0.0036	−0.1163	
	SD	0.2456	0.1135	0.0665	0.1201	0.0750	0.0933	0.0626	0.0864	0.0945	0.1840	0.1405	0.0916	0.1132	0.0694	0.1139	0.0867	0.0798	0.1219	
	ASD	0.4346	0.0893	0.0636	0.1130	0.0798	0.0764	0.0596	0.0893	−	0.1569	0.1229	0.0821	0.1050	0.0345	0.1081	0.0838	0.0867	−	
$\eta_t \sim Laplace(0,1)$																				
300	Bias	0.0879	0.0315	0.0132	−0.0071	0.0027	0.0324	−0.0120	−0.0029	0.0238	−0.1411	0.2665	0.1933	0.0543	−0.0094	0.1366	0.2501	−0.0133	−0.2345	
	SD	0.4255	0.3210	0.1480	0.2814	0.2419	0.1818	0.1322	0.1676	0.1953	0.3770	0.3246	0.2838	0.2625	0.1409	0.2843	0.2467	0.1624	0.2522	
	ASD	0.9172	0.2465	0.1407	0.2424	0.1672	0.1560	0.1086	0.1498	−	0.4344	0.3131	0.2361	0.1998	0.0905	0.2697	0.1977	0.1553	−	
600	Bias	0.0266	0.0169	0.0379	0.0037	0.0044	0.0171	−0.0085	−0.0025	0.0140	−0.1510	0.2746	0.1852	0.0438	−0.0050	0.1284	0.2671	−0.0070	−0.2326	
	SD	0.3477	0.1625	0.1254	0.2269	0.1197	0.0923	0.0790	0.0923	0.1361	0.2698	0.2885	0.2609	0.1927	0.1019	0.1769	0.1781	0.1063	0.2089	
	ASD	0.7976	0.1744	0.1417	0.1869	0.1223	0.0803	0.0754	0.0923	−	0.4236	0.3202	0.5972	0.1711	0.0677	0.1664	0.1524	0.1054	−	
900	Bias	0.0503	0.0134	0.0131	0.0007	0.0017	0.0142	−0.0017	−0.0073	0.0145	−0.1730	0.2729	0.1861	0.0388	0.0040	0.1163	0.2783	−0.0051	−0.2459	
	SD	0.3070	0.1306	0.1145	0.1761	0.0893	0.0684	0.0631	0.0674	0.1015	0.2359	0.2300	0.1948	0.1692	0.0856	0.1308	0.1371	0.0798	0.1986	
	ASD	0.4979	0.1286	0.0985	0.1458	0.0854	0.0574	0.0616	0.0719	−	0.2980	0.2428	0.1846	0.1643	0.0551	0.1198	0.1261	0.0810	−	
$\eta_t \sim (\frac{\pi}{2\sqrt{3}})t_3$																				
300	Bias	0.0460	0.0254	0.0361	0.0062	0.0094	0.0353	−0.0110	−0.0175	0.0181	−0.2024	0.2828	0.2106	0.1221	0.0182	0.1552	0.3058	−0.0198	−0.2972	
	SD	0.3874	0.2762	0.2336	0.3134	0.1696	0.1785	0.1516	0.1837	0.1897	0.5492	0.5434	0.4623	0.3884	0.3693	0.3681	0.4019	0.1990	0.4593	
	ASD	0.9868	0.2649	0.3477	0.2615	0.2322	0.1526	0.1204	0.1602	−	0.5655	0.4292	0.3356	0.2399	0.1469	0.3580	0.2790	0.1928	−	
600	Bias	0.0744	0.0221	0.0321	−0.0030	0.0076	0.0197	0.0061	−0.0095	0.0157	−0.2257	0.3067	0.2134	0.1114	0.0139	0.1650	0.3656	−0.0168	−0.3080	
	SD	0.3170	0.1761	0.2035	0.2539	0.1337	0.0953	0.0971	0.1008	0.1807	0.3958	0.3814	0.4269	0.3802	0.2285	0.2714	0.3876	0.1416	0.3693	
	ASD	0.7293	0.2212	0.1850	0.2129	0.1535	0.0796	0.0887	0.1002	−	0.5430	0.4499	0.3654	0.2585	0.1619	0.2369	0.2451	0.1423	−	
900	Bias	0.0520	0.0206	0.0228	−0.0016	0.0017	0.0129	0.0004	−0.0094	0.0061	−0.2230	0.2876	0.2341	0.1072	0.0171	0.1515	0.3606	−0.0065	−0.2269	
	SD	0.2963	0.1592	0.1857	0.1934	0.1109	0.0695	0.0752	0.0756	0.1606	0.3703	0.3629	0.3925	0.2976	0.1730	0.1937	0.3382	0.1224	0.3537	
	ASD	0.6693	0.1832	0.1587	0.1825	0.0965	0.0558	0.0732	0.0921	−	0.5329	0.3752	0.3568	0.2486	0.1112	0.1798	0.2185	0.1162	−	

Table 2. Estimation performance results of DGP_2 (QMELE vs. Standard QMLE).

n	Parameters True Value	c_{11} 0.7	d_1 0.2	ω_1 0.1	a_{11} 0.5	β_{11} 0.3	c_{21} 0.3	d_2 0.5	ω_2 0.1	a_{21} 0.2	β_{21} 0.6	r 0	c_{11} 0.7	d_1 0.2	ω_1 0.1	a_{11} 0.5	β_{11} 0.3	c_{21} 0.3	d_2 0.5	ω_2 0.1	a_{21} 0.2	β_{21} 0.6	r 0	
QMELE													Standard QMLE											
$\eta_t \sim N(0, \frac{\sigma}{2})$																								
300	Bias	0.0359	−0.0137	0.0481	−0.0328	0.0155	−0.0096	0.0320	0.0422	−0.0195	−0.0178	−0.0383	−0.0959	−0.0917	0.1259	0.2274	0.0363	−0.0223	−0.0126	0.1800	0.0783	0.0138	−0.2076	
	SD	0.5409	0.5306	0.0847	0.2338	0.1485	0.2010	0.3838	0.1125	0.0926	0.1354	0.3876	0.2759	0.2987	0.1185	0.2273	0.1475	0.1171	0.2492	0.1677	0.1213	0.1296	0.3679	
	ASD	1.4229	1.1241	0.1953	0.1500	0.2043	0.1847	0.5038	0.0768	0.0908	0.1192	−	0.3105	0.3780	0.3915	0.2390	0.2067	0.0621	0.1651	0.1051	0.1172	0.1092	−	
600	Bias	0.0198	−0.0096	0.0458	−0.0116	0.0102	−0.0006	0.0146	0.0295	−0.0191	−0.0035	−0.0181	−0.1068	−0.0656	0.1197	0.2145	0.0495	−0.0100	−0.0167	0.1522	0.0097	0.0237	−0.1926	
	SD	0.2824	0.3030	0.0940	0.1476	0.1451	0.1057	0.2527	0.0700	0.0667	0.0818	0.3152	0.1887	0.2099	0.0820	0.1782	0.1074	0.0769	0.1600	0.1049	0.0972	0.0845	0.2622	
	ASD	0.5658	0.5857	0.0770	0.1148	0.1335	0.1799	0.3272	0.0526	0.0638	0.0826	−	0.2092	0.2739	0.1385	0.1560	0.1255	0.0389	0.1155	0.0706	0.0833	0.0748	−	
900	Bias	0.0046	−0.0109	0.0426	−0.0187	0.0185	0.0059	0.0153	0.0198	−0.0166	0.0019	0.0009	−0.1023	−0.0685	0.1242	0.2012	0.0554	−0.0090	−0.0188	0.1448	0.0912	0.0307	−0.1842	
	SD	0.3039	0.2515	0.0648	0.1971	0.1064	0.1343	0.2373	0.0542	0.0551	0.0674	0.2467	0.1693	0.1691	0.0824	0.1452	0.1005	0.0593	0.1332	0.0848	0.0834	0.0676	0.2431	
	ASD	0.6637	0.5310	0.0552	0.0906	0.0916	0.2017	0.2595	0.0421	0.0524	0.0671	−	0.1756	0.1883	0.1045	0.1291	0.1011	0.0329	0.0982	0.0574	0.0681	0.0608	−	
$\eta_t \sim Laplace(0,1)$																								
300	Bias	0.0313	−0.0120	0.0486	0.0056	0.0191	0.0014	0.0238	0.0368	−0.0076	−0.0164	−0.0621	−0.1206	−0.1771	0.1924	0.4032	0.0585	0.0047	−0.0641	0.2774	0.1512	0.0214	−0.3335	
	SD	0.3193	0.3893	0.1095	0.2640	0.2000	0.1172	0.2773	0.1165	0.1037	0.1224	0.4687	0.4313	0.4547	0.2527	0.4925	0.2187	0.1716	0.3933	0.2575	0.1904	0.1513	0.5577	
	ASD	1.3226	1.5866	0.3796	0.2624	0.3878	0.1525	0.4386	0.0834	0.1069	0.1185	−	1.1642	1.1136	1.2507	0.5134	0.4134	0.1286	0.3395	0.1632	0.1804	0.1230	−	
600	Bias	0.0017	−0.0226	0.0514	0.0135	0.0135	0.0002	0.0154	0.0266	−0.0064	−0.0123	−0.0289	−0.1515	−0.1625	0.1936	0.3808	0.0636	0.0094	−0.0631	0.2444	0.1841	0.0100	0.4300	
	SD	0.3227	0.3890	0.1052	0.2181	0.1335	0.0716	0.2301	0.0723	0.0705	0.0805	0.3505	0.3311	0.3739	0.1699	0.3582	0.2084	0.1188	0.3481	0.1662	0.1664	0.0978	0.4905	
	ASD	0.8516	0.9170	0.1992	0.1811	0.2317	0.0831	0.2878	0.0570	0.0753	0.0816	−	0.6877	0.9354	1.0392	0.4891	0.6041	0.1211	0.2533	0.1125	0.1332	0.0848	−	
900	Bias	0.0030	−0.0069	0.0504	0.0061	0.0173	0.0056	0.0215	0.0192	−0.0101	−0.0014	−0.0393	−0.1409	−0.1552	0.1997	0.3613	0.0740	0.0146	−0.0530	0.2273	0.1980	0.0318	−0.3594	
	SD	0.2428	0.2748	0.0846	0.1918	0.1305	0.0869	0.1885	0.0550	0.0593	0.0623	0.3283	0.3440	0.3244	0.1704	0.2986	0.1644	0.1552	0.3282	0.1290	0.1459	0.0834	0.4708	
	ASD	0.5947	0.6394	0.1459	0.1683	0.1894	0.1004	0.2974	0.0452	0.0592	0.0651	−	0.8345	1.0395	0.8602	0.3704	0.2656	0.0864	0.2290	0.0915	0.1107	0.0691	−	
$\eta_t \sim (\frac{\pi}{2\sqrt{3}})^{1/3}$																								
300	Bias	0.0433	−0.0023	0.0626	0.0229	0.0030	0.0154	0.0112	0.0438	−0.0093	−0.0229	−0.0537	−0.1922	−0.2268	0.2413	0.4413	0.0825	0.0270	−0.0813	0.3328	0.1905	0.0208	−0.3948	
	SD	0.3766	0.3797	0.1479	0.3230	0.2228	0.2949	0.3053	0.1284	0.1131	0.1362	0.4103	0.5678	0.6324	0.3059	0.6310	0.2967	0.2549	0.5479	0.3934	0.3245	0.1924	0.7518	
	ASD	1.6808	1.3841	0.3948	0.3133	0.4711	0.1528	0.4242	0.0921	0.1236	0.1341	−	2.1121	3.5392	4.8341	1.2046	0.6966	0.3020	0.7381	0.2889	0.3147	0.1813	−	
600	Bias	0.0107	−0.0194	0.0550	0.0115	0.0276	0.0079	−0.0195	0.0282	−0.0095	−0.0107	−0.0467	−0.1745	−0.2191	0.2423	0.4505	0.0855	0.0376	−0.1391	0.3091	0.2309	0.0266	−0.4178	
	SD	0.3062	0.5962	0.1808	1.3040	0.1933	0.1042	0.4318	0.0865	0.0850	0.0948	0.4613	0.4070	0.5125	0.3649	0.6025	0.2615	0.2436	0.5146	0.3219	0.2968	0.1420	0.6121	
	ASD	0.7485	0.8655	0.3999	0.2782	0.2164	0.1242	0.4406	0.0638	0.0879	0.0947	−	1.9790	2.8621	4.0155	1.3184	0.7723	0.5350	0.8205	0.2876	0.2960	0.1757	−	
900	Bias	0.0111	−0.0158	0.0544	0.0121	0.0280	0.0124	0.0082	0.0208	−0.0076	−0.0082	−0.0466	−0.1752	−0.1976	0.2217	0.4025	0.0938	0.0281	−0.1299	0.3051	0.2509	0.0400	−0.3946	
	SD	0.3003	0.3518	0.1512	0.2689	0.1855	0.1521	0.2643	0.0634	0.0707	0.0719	0.3832	0.3332	0.4359	0.2450	0.4416	0.2084	0.1844	0.4231	0.3072	0.2867	0.1272	0.5482	
	ASD	0.8089	0.8741	0.1964	0.2315	0.1794	0.1236	0.3430	0.0511	0.0719	0.0767	−	1.4053	1.8931	4.5213	0.9127	0.5223	0.2029	0.4696	0.1930	0.2297	0.1177	−	

4. Real Data Example

In this section, we provide our analysis of a real data set of the Nikkei 225 in the Japanese stock market. The original data express the closing price on each trading day from 1 May 2014 to 1 April 2019, and there are 1208 valid observations. The explanatory variable is the corresponding daily volume (z_t). Figure 1a presents the the original data of the closing price on each trading day, which we denote as p_t . Figure 1b shows the rescaled 100 times log return, which we denote as $\{y_t\}_{t=1}^{1207}$: $y_t = 100 \cdot \log \frac{p_{t+1}}{p_t}$.

Figure 1c,d present the density plots and the histogram of the log return, respectively. From these two nonparametric plots, we can see that the log return data exhibit the characteristics of non-normality and fat tails.

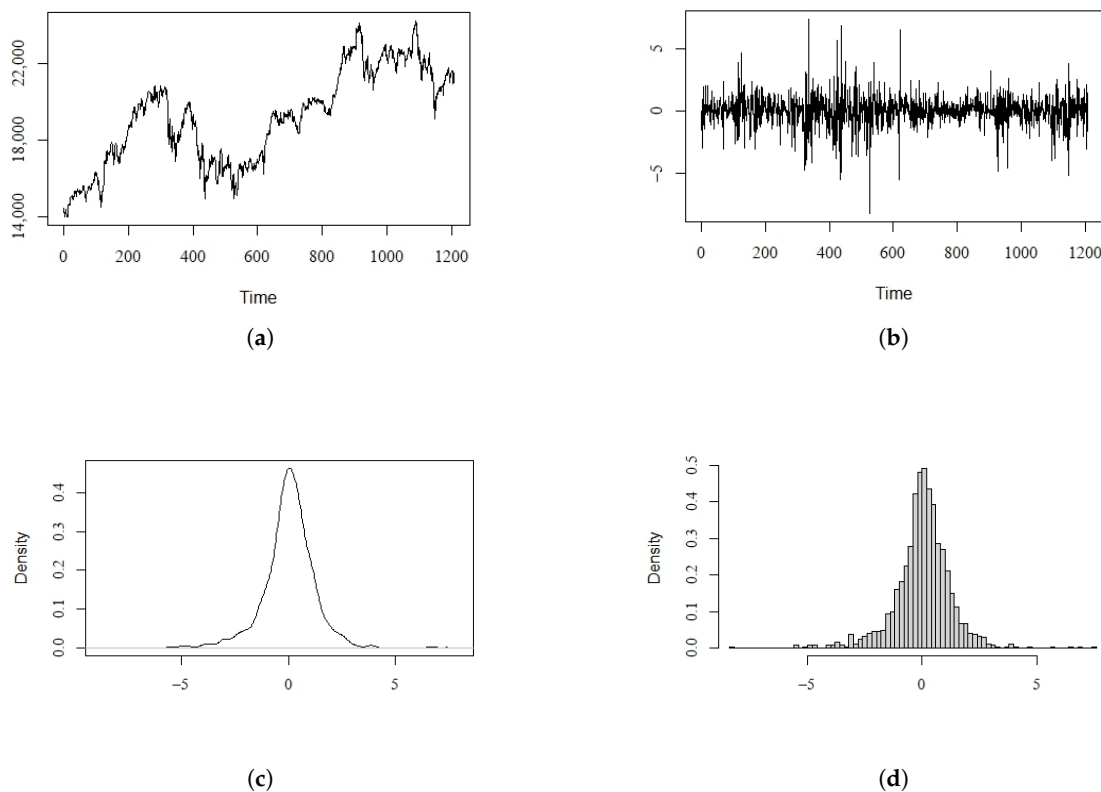


Figure 1. The data of the Nikkei 225. (a) Original closing price. (b) 100 times log return. (c) Density plots of log return. (d) Frequency histogram of log return.

In order to estimate the tail index of $\{y_t^2\}$, we use the Hill's estimator as follows:

$$\hat{H}(k) = \left\{ \frac{1}{k} \sum_{i=1}^k [\log \hat{y}_{n-i}^2 - \log \hat{y}_{n-k}^2] \right\}^{-1},$$

where $k = 1, 2, \dots, 100$, and $\hat{y}_i^2$ is the i th order statistic of $\{y_t^2\}$. From Figure 2, we can clearly see that when $k > 40$, $1 < \hat{H}(k) < 2$. Thus, the tail index of $\{y_t^2\}$ lies between 1 and 2, which means $E y_t^4 = \infty$. Since this result does not satisfy the assumption of the asymptotic normality of QMLE, we use the QMELE to estimate the model parameters. First, we need to determine d and locate the threshold parameter r . Following several studies that analyze the stock return series (Chen et al., 2006 [18]; Yang and Chang, 2008 [19]; Chen

et al. 2016 [22]), we believe that $d = 1$ is an appropriate choice. By using the SGS method, we obtain $\hat{\tau} = 0.0151$ and the fitted model

$$y_t = \begin{cases} \phi_t^{(1)} y_{t-1} + \eta_t \sqrt{0.0086 + 0.1210 \cdot \epsilon_{t-1}^2 + 0.7928 \cdot h_{t-1}}, & y_{t-d} \leq 0.0151, \\ \phi_t^{(2)} y_{t-1} + \eta_t \sqrt{0.000113 + 0.0115 \cdot \epsilon_{t-1}^2 + 0.8302 \cdot h_{t-1}}, & y_{t-d} > 0.0151, \end{cases}$$

$$\begin{cases} \phi_t^{(1)} = 17.2753 \cdot y_{t-1} - 0.4812 \cdot z_t, \\ (5.8739) \quad (0.0001387) \\ \phi_t^{(2)} = -1.8262 \cdot y_{t-1} - 2.1476 \cdot z_t, \\ (0.5202) \quad (0.0006448) \end{cases}$$

where the standard errors are given in parentheses. The estimated value of $E|\hat{\eta}_t|$ is 1.0238, which is very close to 1. Figure 3 shows the density histogram and density plots of $\hat{\eta}_t$. We can see that the distribution of $\hat{\eta}_t$ is asymmetric and heavy tailed and very similar to a double exponential distribution. This evidence determines the choice of using QMELE to fit these Nikkei 225 log return data. Figure 4 shows the diagnostic checking plots for our fitted model. The estimated residuals indicate that the series is stationary. In addition, the histogram plot implies that the standardized residuals are approximately normal, which reveals that our fitted model is suitable. As a comparison, we present the DT-GARCH fitted model of the same data, which is as follows:

$$y_t = \begin{cases} -0.1551 y_{t-1} + \epsilon_t, & y_{t-1} \leq 0.0151, \\ (0.0581) \\ 0.0749 y_{t-1} + \epsilon_t, & y_{t-1} > 0.0151, \end{cases} \quad \epsilon_t = \eta_t \sqrt{h_t},$$

$$h_t = \begin{cases} 0.1008 + 1264 \epsilon_{t-1}^2 + 0.7702 h_{t-1}, & y_{t-1} \leq 0.0151, \\ (0.0169) \quad (0.0239) \quad (0.0498) \\ 0.0135 + 0.0114 \epsilon_{t-1}^2 + 0.8059 h_{t-1}, & y_{t-1} > 0.0151, \\ (0.0106) \quad (0.0148) \quad (0.0321) \end{cases}$$

Table 3 displays the AIC of these two fitted models, from which we can see that the AIC of the TVCDT-GARCH model is smaller than the one of the DT-GARCH model.

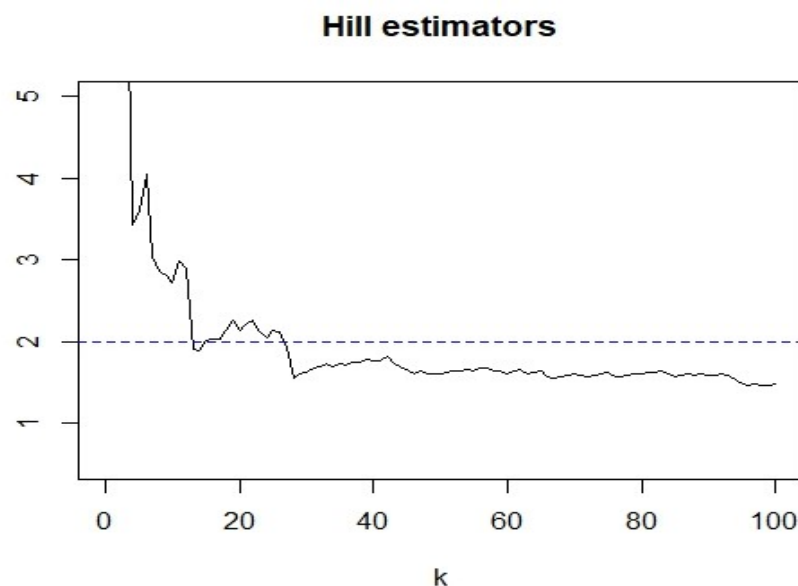


Figure 2. The Hill's estimators of y_t^2 .

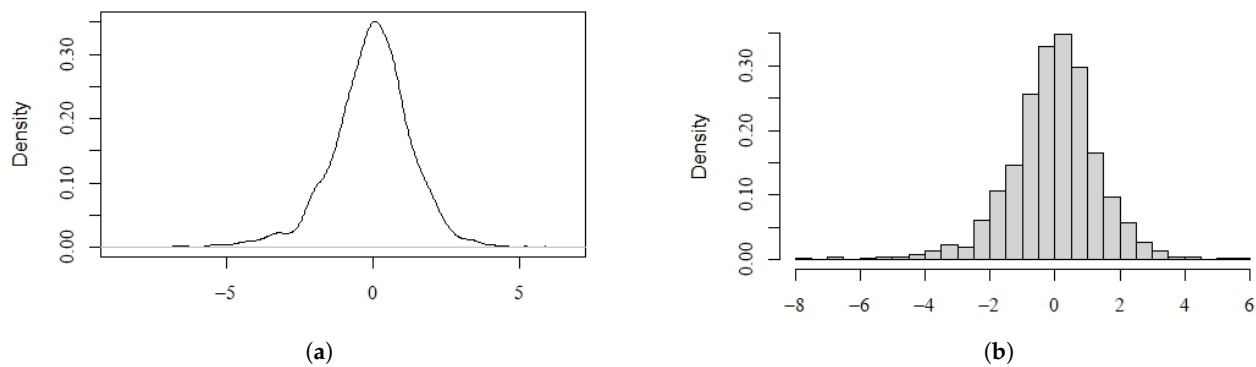


Figure 3. The frequency histogram and density plots of $\hat{\eta}_t$. (a) Density plots. (b) Frequency histogram.

Table 3. AIC of two fitted models.

	TVCDT-GARCH	DT-GARCH
AIC	1922.522	1925.186

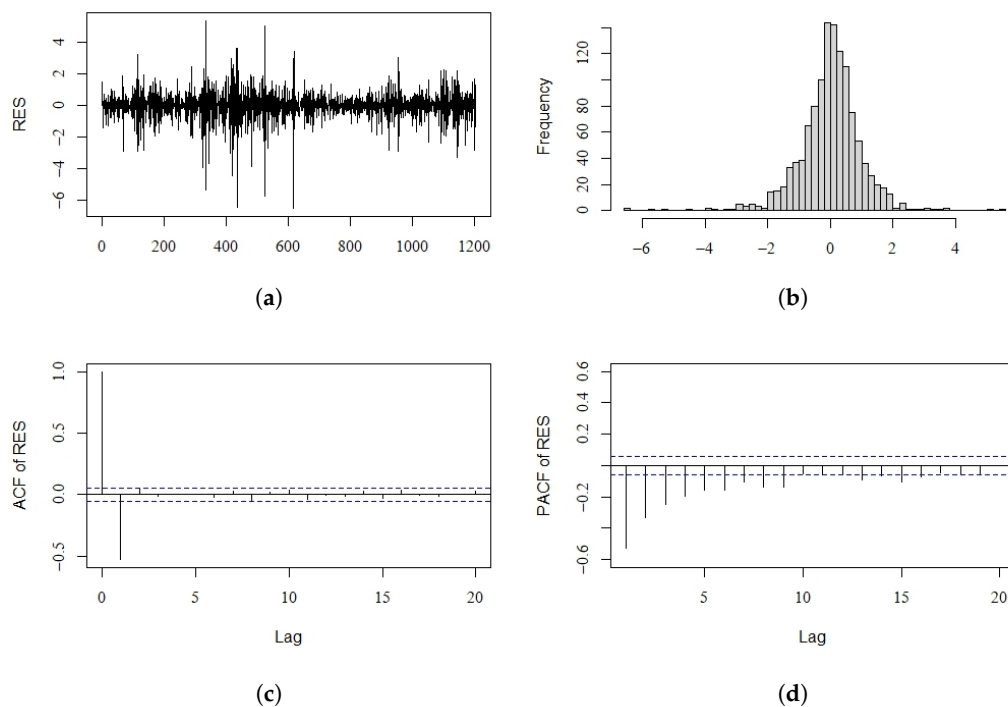


Figure 4. Diagnostic checking plots for the Nikkei 225 (log return) data. (a) Standardized residuals; (b) histogram of standardized residuals; (c) ACF plot of residuals; (d) PACF plot of residuals.

In conclusion, the real data application reveals that the TVCDT-GARCH model is better than the DT-GARCH model, and the QMELE is proper, compatible, and more accurate than the traditional QMLE. This is not surprising, so in the TVCDT-GARCH model, we consider the influence of the explanatory variable (daily volume). Thus, the model contains more information and can fit the real data better than the DT-GARCH model. Furthermore, the distribution of η_t exhibits asymmetric and heavy-tailed features, which is a frequent phenomenon involving financial time series, so we think that using QMELE in parameter-estimation procedures is a pragmatic alternative for practitioners.

5. Conclusions

In this article, we studied the use of the quasi-maximum exponential likelihood estimators (QMELE) to fit the TVCDT-GARCH model. We presented the consistency and the asymptotic normality of the QMELE for the model parameters. The simulation study provided the performance of these estimators. The application of the Nikkei 225 log return series revealed that the TVCDT-GARCH model is appropriate and feasible for analyzing the data, and QMELE provides researchers with a better approach than traditional QMLEs. For further study, the explanatory variables may be developed to higher dimensions, and the models may involve multiple threshold parameters under certain circumstances. These discussions will help us to extend our research in the future.

Author Contributions: Conceptualization, T.Z., L.F., D.W. and Z.Y.; methodology, T.Z. and D.W.; software, T.Z.; validation, T.Z., L.F., D.W. and Z.Y.; formal analysis, T.Z., L.F., D.W. and Z.Y.; investigation, T.Z., L.F., D.W. and Z.Y.; resources, L.F. and D.W.; data curation, T.Z. and D.W.; writing—original draft preparation, T.Z.; writing—review and editing, T.Z., L.F., D.W. and Z.Y.; visualization, T.Z.; supervision, L.F. and D.W.; project administration, L.F., D.W. and Z.Y.; funding acquisition, D.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the National Natural Science Foundation of China (No.12271231, 12001229, 11901053), the National Social Science Foundation of China (No.21BTJ043, No.20BTJ056), and the Social Science Foundation of Liaoning Province (No.L22ZD065).

Data Availability Statement: Data are available on the website: hk.finance.yahoo.com, accessed on 1 January 2023.

Acknowledgments: We gratefully acknowledge the anonymous reviewers and editors for their insightful comments on and suggestions for this paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACF Autocorrelation Function Plot
PACF Partial Autocorrelation Function Plot

Appendix A

We need Lemma A1 and Lemma A2 for our proofs; hence, we list two lemmas below.

Lemma A1. Let $\xi_{pt} = 1 + \sum_{i=0}^{\infty} \rho^i |y_{t-i}|$, where $\rho \in (0, 1)$. If Assumptions (1)–(4) hold, then there exist constants C and $\rho \in (0, 1)$ such that the following holds uniformly in Θ :

- (1) $\sup_{\Theta} |\epsilon_{t-1}(\gamma)| \leq C \xi_{pt-1}$, $\sup_{\Theta} \left\| \frac{\partial \epsilon_t(\gamma)}{\partial \gamma} \right\| \leq C \xi_{pt-1}$, and $\sup_{\Theta} \left\| \frac{\partial^2 \epsilon_t(\gamma)}{\partial \gamma \partial \gamma'} \right\| \leq C \xi_{pt-1}$;
- (2) $\sup_{\Theta} h_t(\theta) \leq C \xi_{pt-1}$;
- (3) $\sup_{\Theta} \left\| \frac{1}{h_t(\theta)} \frac{\partial h_t(\theta)}{\partial \delta} \right\| \leq C \xi_{pt-1}$;
- (4) $\sup_{\Theta} \left\| \frac{1}{\sqrt{h_t(\theta)}} \frac{\partial h_t(\theta)}{\partial \gamma} \right\| \leq C \xi_{pt-1}$.

Lemma A2. For any $\theta^* \in \Theta$, let $B_\eta(\theta^*) = \{\theta \in \Theta : \|\theta - \theta^*\| < \eta\}$ be an open neighborhood of θ^* with radius $\eta > 0$. If Assumptions (1)–(4) hold, then we have the following:

- (a) $E[\sup_{\theta \in \Theta} l_t(\theta)] < \infty$;
- (b) $E[l_t(\theta)]$ has a unique minimum at θ_0 ;
- (c) $E[\sup_{\theta \in B_\eta(\theta^*)} |l_t(\theta) - l_t(\theta^*)|] \rightarrow 0$ as $\eta \rightarrow 0$.

Lemma A1 is taken from Lemma A2 in [45] and Lemma A1 of [36]. Lemma 1 is straightforward and follows from the central limit theorem for a martingale difference sequence. For Lemma 2, by letting $\omega_t \equiv 1$ in Lemmas 2.2 and 2.3 of [36], we can directly obtain Lemma 2.2. The detailed proofs of Lemma A2, Lemma 3, Theorem 1 and Theorem 2 can be found in the appendix of [41], and, hence, we omit the proofs of these lemmas and theorems.

References

- Nicholls, D.F.; Quinn, B.G. *Random Coefficient Autoregressive Models: An Introduction*; Springer: New York, NY, USA, 1982.
- Tjøstheim, D. Estimation in nonlinear time series models. *Stoch. Process. Their Appl.* **1986**, *21*, 251–273. [\[CrossRef\]](#)
- Kim, Y.W.; Basawa, I.V. Empirical Bayes estimation for first order autoregressive processes. *Aust. J. Stat.* **1992**, *34*, 105–114. [\[CrossRef\]](#)
- Hwang, S.Y.; Basawa, I.V. Asymptotic optimal inference for a class of nonlinear time series models. *Stoch. Process. Their Appl.* **1993**, *46*, 91–113. [\[CrossRef\]](#)
- Hwang, S.Y.; Basawa, I.V. Large sample inference for conditional exponential families with applications to nonlinear time series. *J. Stat. Plan. Inference* **1994**, *38*, 141–157. [\[CrossRef\]](#)
- Hwang, S.Y.; Basawa, I.V. Parameter estimation for generalized random coefficient autoregressive processes. *J. Stat. Plan. Inference* **1998**, *68*, 323–337. [\[CrossRef\]](#)
- Aue, A. Strong approximation for RCA(1) time series with applications. *Stat. Probab. Lett.* **2004**, *68*, 369–382. [\[CrossRef\]](#)
- Aue, A.; Horvath, L.; Steinebach, J. Estimation in Random Coefficient Autoregressive Models. *J. Time Ser. Anal.* **2006**, *27*, 61–76. [\[CrossRef\]](#)
- Thavaneswaran, A.; Appadoo, S.S.; Ghahramani, M. RCA models with GARCH innovations. *Appl. Math. Lett.* **2009**, *22*, 110–114. [\[CrossRef\]](#)
- Tong, H. *On a Threshold Model*; Sijthoff & Noordhoff: Amsterdam, The Netherlands, 1978.
- Tong, H.; Lim, K.S. Threshold autoregression, limit cycles and cyclical data. *J. R. Stat. Soc. Ser. B (Methodol.)* **1980**, *42*, 245–268. [\[CrossRef\]](#)
- Engle, R.F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* **1982**, *50*, 987–1007. [\[CrossRef\]](#)
- Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *J. Econom.* **1986**, *31*, 307–327. [\[CrossRef\]](#)
- Karpoff, J.M. The Relation between Price Changes and Trading Volume: A Survey. *J. Financ. Quant. Anal.* **1987**, *22*, 109–126. [\[CrossRef\]](#)
- Tong, H. *Non-Linear Time Series: A Dynamical System Approach*; Oxford University Press: Oxford, UK, 1990.
- Li, C.W.; Li, W.K. On a Double-Threshold Autoregressive Heteroscedastic Time Series Model. *J. Appl. Econometr.* **1996**, *11*, 253–274. [\[CrossRef\]](#)
- Brooks, C. A Double-threshold GARCH Model for the French Franc/Deutschmark exchange rate. *J. Forecast.* **2001**, *20*, 135–143. [\[CrossRef\]](#)
- Chen, C.W.S.; Yang, M.J.; Gerlach, R.; Jim, L.H. The asymmetric reactions of mean and volatility of stock returns to domestic and international information based on a four-regime double-threshold GARCH model. *Phys. Stat. Mech. Appl.* **2006**, *366*, 401–418. [\[CrossRef\]](#)
- Yang, Y.L.; Chang, C.L. A double-threshold GARCH model of stock market and currency shocks on stock returns. *Math. Comput. Simul.* **2008**, *79*, 458–474. [\[CrossRef\]](#)
- Chen, C.W.S.; Gerlach, R. Semi-parametric quantile estimation for double threshold autoregressive models with heteroskedasticity. *Comput. Stat.* **2013**, *28*, 1103–1131. [\[CrossRef\]](#)
- Chen, C.W.S.; Chen, M.; Chen, H. Pairs Trading via Three-Regime Threshold Autoregressive GARCH Models. In Proceedings of the 7th International Conference of the Thailand Econometric Society, TES 2014, Chiang Mai, Thailand, 8–10 January 2014; pp. 127–140.
- Chen, C.W.S.; So, M.K.P.; Chiang, T.C. Evidence of Stock Returns and Abnormal Trading Volume: A Threshold Quantile Regression Approach. *Jpn. Econ. Rev.* **2016**, *67*, 96–124. [\[CrossRef\]](#)
- Luukkonen, R.; Saikkonen, P.; Teräsvirta, T. Testing linearity against smooth transition autoregressive models. *Biometrika* **1988**, *75*, 491–499. [\[CrossRef\]](#)
- Teräsvirta, T. Specification, estimation, and evaluation of smooth transition autoregressive models. *J. Am. Stat. Assoc.* **1994**, *89*, 208–218.
- Zheng, H.T.; Basawa, I.V. First-order observation-driven integer-valued autoregressive processes. *Stat. Probab. Lett.* **2008**, *78*, 1–9. [\[CrossRef\]](#)
- Yu, M.J.; Wang, D.H.; Yang, K. A class of observation-driven random coefficient INAR(1) processes based on negative binomial thinning. *J. Korean Stat. Soc.* **2019**, *48*, 248–264. [\[CrossRef\]](#)
- Yang, K.; Li, H.; Wang, D.H.; Zhang, C.H. Random coefficients integer-valued threshold autoregressive processes driven by logistic regression. *ASTA Adv. Stat. Anal.* **2021**, *105*, 533–557. [\[CrossRef\]](#)

28. Hwang, S.Y.; Basawa, I.V. Explosive Random Coefficient AR(1) Processes and Related Asymptotics for Least Squares Estimation. *J. Time Ser. Anal.* **2005**, *26*, 807–824. [[CrossRef](#)]
29. Berkes, I.; Horvath, L.; Ling, S.Q. Estimation in nonstationary random coefficient autoregressive models. *J. Time Ser. Anal.* **2009**, *30*, 395–416. [[CrossRef](#)]
30. Aue, A.; Horvath, L. Quasi-likelihood estimation in stationary and nonstationary autoregressive models with random coefficients. *Stat. Sin.* **2011**, *21*, 973–999.
31. Weiss, A. Asymptotic theory for ARCH models: Estimation and testing. *Econom. Theory* **1986**, *2*, 107–131. [[CrossRef](#)]
32. Engle, R.F.; Bollerslev, T. Modelling the persistence of conditional variances. *Econom. Rev.* **1986**, *5*, 1–50. [[CrossRef](#)]
33. Berkes, I.; Horvath, L.; Kokoszka, P. GARCH processes: Structure and estimation. *Bernoulli* **2003**, *9*, 201–227. [[CrossRef](#)]
34. Francq, C.; Zokaian, J.M. Maximum Likelihood Estimation of Pure GARCH and ARMA-GARCH Processes. *Bernoulli* **2004**, *10*, 605–637. [[CrossRef](#)]
35. Hall, P.; Yao, Q. Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica* **2003**, *71*, 285–317. [[CrossRef](#)]
36. Zhu, K.; Ling, S.Q. Global self-weighted and local quasi-maximum exponential likelihood estimators for ARMA-GARCH/IGARCH models. *Ann. Stat.* **2011**, *39*, 2131–2163. [[CrossRef](#)]
37. Tsay, R.S. Testing and modeling threshold autoregressive processes. *J. Am. Stat. Assoc.* **1989**, *84*, 231–240. [[CrossRef](#)]
38. Li, D.; Ling, S.Q. On the least squares estimation of multiple-regime threshold autoregressive models. *J. Econom.* **2012**, *167*, 240–253. [[CrossRef](#)]
39. Yu, P. Likelihood estimation and inference in threshold regression. *J. Econom.* **2012**, *167*, 274–294. [[CrossRef](#)]
40. Li, D.; Tong, H. Nested sub-sample search algorithm for estimation of threshold models. *Stat. Sin.* **2016**, *26*, 1543–1554. [[CrossRef](#)]
41. Zhang, T.W.; Wang, D.H.; Yang, K. Quasi-maximum exponential likelihood estimation for double-threshold GARCH models. *Can. J. Stat.* **2021**, *49*, 1152–1178. [[CrossRef](#)]
42. Sheng, D.S.; Wang, D.H.; Kang, Y. A new RCAR(1) model based on explanatory variables and observations. *Commun. Stat.-Theory Methods* **2022**. [[CrossRef](#)]
43. Li, G.D.; Guan, B.; Li, W.K.; Yu, P.L.H. Hysteretic autoregressive time series models. *Biometrika* **2015**, *102*, 717–723. [[CrossRef](#)]
44. Silverman, B.W. *Density Estimation for Statistics and Data Analysis*; Chapman and Hall: New York, NY, USA, 1986.
45. Ling, S.Q. Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH/IGARCH models. *J. Econom.* **2007**, *140*, 849–873. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.