

RESEARCH

Open Access



Empirical likelihood inference for threshold autoregressive conditional heteroscedasticity model

Cuixin Peng^{1*}  and Zhiwen Zhao²

*Correspondence:

pengcuixin@126.com

¹School of Foreign Languages, Jilin Normal University, Siping, 136000, P.R. China

Full list of author information is available at the end of the article

Abstract

This paper considers the parameter estimation problem of a first-order threshold autoregressive conditional heteroscedasticity model by using the empirical likelihood method. We obtain the empirical likelihood ratio statistic based on the estimating equation of the least squares estimation and construct the confidence region for the model parameters. Simulation studies indicate that the empirical likelihood method outperforms the normal approximation-based method in terms of coverage probability.

MSC: Primary 62M10; secondary 91B62

Keywords: Empirical likelihood; Threshold autoregressive model; Conditional heteroscedasticity; Confidence region; Least squares method

1 Introduction

Consider the following first-order threshold autoregressive conditional heteroscedasticity model:

$$X_t = \theta_1 X_{t-1}^+ + \theta_2 X_{t-1}^- + \varepsilon_t, \quad (1)$$

where $X_t^+ = \max(X_t, 0)$, $X_t^- = \min(X_t, 0)$, $\varepsilon_t = \sqrt{h_t} e_t$, $h_t = \alpha_0 + \alpha_1 (\varepsilon_{t-1}^+)^2 + \alpha_2 (\varepsilon_{t-1}^-)^2$, e_t is a sequence of independent and identically distributed random variables satisfying $Ee_t = 0$ and $\text{Var}(e_t) = 1$. θ_1 , θ_2 , α_0 , α_1 , and α_2 are the model parameters with $\alpha_0 > 0$, $0 \leq \alpha_j < 1$, $j = 1, 2$.

When $\theta_1 = \theta_2$, model (1) becomes the usual autoregressive model whose innovation is a conditional heteroscedasticity process. Threshold autoregressive model is a nonlinear time series model. Because the threshold autoregressive model can explain nonlinear features such as asymmetry and limit cycles, it is widely used in time series modeling (see Tong [1]). Petrucci and Woolford [2] first defined model (1) and investigated its properties and the parameter estimation problems, but they assumed that the error sequence is a sequence of independent and identically distributed random variables. Brockwell et al. [3] and Hwang and Basawa [4] further generalized the model coefficients to be random

© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

variables. Hwang and Woo [5] first considered the parameter estimation problems when the error sequence is a conditional heteroscedasticity process and proposed to use the conditional least squares method to estimate the model parameters. In this paper, we use the empirical likelihood method to estimate the model parameters.

Similar to the parametric likelihood, Owen [6–8] introduced empirical likelihood method. It is a nonparametric likelihood method which establishes a likelihood function through placing positive probability on every one of the observed data values, but often makes no assumptions on the data-generating mechanism. Empirical likelihood has many advantages compared with the normal approximation method. For example, the limiting distribution of empirical likelihood ratio statistic is a chi-squared distribution. Therefore, we need not estimate the asymptotic variance when we construct the confidence region. Moreover, the confidence region is completely decided by the data themselves because we make no assumptions on the probability distribution of the data. These attract the attention of statisticians to make inference for all kinds of statistical models using the empirical likelihood method, such as linear regressive model [9–13], generalized linear models [14–17], and partially linear models [18–21]. In recent years, empirical likelihood method is also applied to make statistical inference about time series models, such as autoregressive model [22–24], random coefficient autoregressive model [25–28], and integer-valued autoregressive model [29–32].

In this paper, we obtain the limiting distribution of empirical log-likelihood ratio statistic and construct the confidence region for the parameters in model (1) by using the empirical likelihood method. Some simulation studies indicate that the empirical likelihood method has a higher coverage probability compared with the normal approximation-based method.

This paper is organized as follows. In Sect. 2, we present the main methods and results. Some simulation results and real data analysis are given in Sect. 3. Section 4 is concerned with the proofs of the main results. Moreover, the symbols “ \xrightarrow{d} ” and “ \xrightarrow{p} ” denote convergence in distribution and convergence in probability, respectively. $O_p(1)$ means a term which is bounded in probability. $o_p(1)$ means a term which converges to zero in probability. “Almost surely” and “independent identical distributed” are denoted by “a.s.” and “i.i.d.,” respectively.

2 Methods and main results

For model (1), Hwang and Woo [5] obtained the least square estimation of the model parameters and its limiting properties. Now we use the empirical likelihood method to estimate the model parameters. Before giving the main results, we assume that the following conditions are true:

- (A1) Probability density function $f(\cdot)$ of e_t has its support on $(-\infty, +\infty)$. $\theta_{\max} + \sqrt{\alpha_{\max}} < 1$, where $\theta_{\max} = \max\{|\theta_1|, |\theta_2|\}$, $\alpha_{\max} = \max\{\alpha_1, \alpha_2\}$.
- (A2) $E(X_t^6) < \infty$.

According to Theorem 1 in Hwang and Woo [5], if (A1) holds, then $\{X_t, t \geq 1\}$ is geometrically ergodic, and the sequence $\{X_t\}$ has a unique stationary distribution.

Hwang and Woo [5] used the least square method to estimate the model parameters. Let $\theta = (\theta_1, \theta_2)^T$. Based on the observation data $\{X_0, X_1, \dots, X_n\}$, the least square estimation θ^* of θ can be obtained by minimizing $Q(\theta) = \sum_{t=1}^n (X_t - E(X_t | \mathcal{F}_{t-1}))^2 = \sum_{t=1}^n (X_t - \theta_1 X_{t-1}^+ -$

$\theta_2 X_{t-1}^-)^2$ with respect to θ . Solving

$$\frac{\partial Q(\theta)}{\partial \theta} = -2 \sum_{t=1}^n (X_t - \theta_1 X_{t-1}^+ - \theta_2 X_{t-1}^-) \begin{pmatrix} X_{t-1}^+ \\ X_{t-1}^- \end{pmatrix} = 0 \tag{2}$$

for θ , we know that

$$\theta^* = \begin{pmatrix} \sum_{t=1}^n X_t X_{t-1}^+ / \sum_{t=1}^n (X_{t-1}^+)^2 \\ \sum_{t=1}^n X_t X_{t-1}^- / \sum_{t=1}^n (X_{t-1}^-)^2 \end{pmatrix}.$$

Let $\mathcal{X}_t = (X_{t-1}^+, X_{t-1}^-)^\tau$. Then the estimating equation (2) can be written as

$$\sum_{t=1}^n (X_t - \mathcal{X}_t^\tau \theta) \mathcal{X}_t = 0. \tag{3}$$

Further, let $H_t(\theta) = (X_t - \mathcal{X}_t^\tau \theta) \mathcal{X}_t$. By (3), we can obtain the following empirical likelihood ratio statistic:

$$L(\theta) = \max \left\{ \prod_{t=1}^n n p_t : \sum_{t=1}^n p_t H_t(\theta) = 0, p_t \geq 0, \sum_{t=1}^n p_t = 1 \right\}. \tag{4}$$

By using the Lagrange multiplier method, it is easy to know that

$$p_t = \frac{1}{n} \frac{1}{1 + b^\tau(\theta) H_t(\theta)}, \tag{5}$$

where the Lagrange multiplier $b(\theta)$ satisfies

$$\frac{1}{n} \sum_{t=1}^n \frac{H_t(\theta)}{1 + b^\tau(\theta) H_t(\theta)} = 0. \tag{6}$$

Therefore, we have

$$-2 \log(L(\theta)) = 2 \sum_{t=1}^n \log(1 + b^\tau(\theta) H_t(\theta)). \tag{7}$$

The following theorem indicates that the limiting distribution of $-2 \log(L(\theta))$ is a chi-squared distribution.

Theorem 2.1 *If (A1) and (A2) hold, then when $n \rightarrow \infty$,*

$$-2 \log(L(\theta)) \xrightarrow{d} \chi^2(2), \tag{8}$$

where $\chi^2(2)$ is the chi-squared distribution with two degrees of freedom.

Using the above theorem, we can construct the empirical likelihood ratio confidence region for the parameter θ . For $0 < \delta < 1$, the $100(1 - \delta)\%$ asymptotic confidence region for the parameter θ is

$$\mathcal{C}\{\delta\} = \{\theta : L(\theta) \leq \chi_\delta^2(2)\}, \tag{9}$$

where $\chi_\delta^2(2)$ is the upper δ quantile of chi-squared distribution with two degrees of freedom.

3 Simulation studies

In this section, we carry out some simulation studies to compare the performances of our empirical likelihood (EL) method with the least square (LS) method proposed by Hwang and Woo [5] through random simulation. Consider the simulation results of model (1) in the following error sequence:

Sequence I: $\{e_t\}$ is a sequence of independent and identically distributed (i.i.d.) standard normal distribution $N(0, 1)$ random variables.

Sequence II: $\{e_t\}$ is an independent noise sequence with ϵ -contamination distribution, and the distribution function of $\{e_t\}$ is

$$F_{e_t}(x) = \epsilon \Phi\left(\frac{x}{\sigma_1}\right) + (1 - \epsilon)\Phi\left(\frac{x}{\sigma_2}\right),$$

where $\sigma_i > 0$ ($i = 1, 2$), ϵ is a fixed constant satisfying $0 < \epsilon < 1$ and $\Phi(x)$ is the distribution function of the standard normal random variable.

Sequence III: $\{e_t\}$ is a sequence of independent and identically distributed (i.i.d.) mixing random variable sequence, and the distribution function of $\{e_t\}$ is

$$F_{e_t}(x) = \epsilon \Phi\left(\frac{x}{\sigma}\right) + (1 - \epsilon)T_k(x),$$

where $\sigma > 0$, $0 < \epsilon < 1$, $T_k(x)$ is the distribution function of T distribution with k degrees of freedom, $\Phi(x)$ is the distribution function of standard normal random variable.

We calculate the coverage probabilities of the empirical likelihood and the least square methods for different model parameters. The nominal confidence level $1 - \delta$ is chosen to be 0.90. All simulation studies are based on 1000 repetitions, and the sample sizes considered in these simulations are $n = 100, 300$, and 500. The simulation results for sequence I are presented in Table 1. For sequence II, we simulate $(\epsilon, \sigma_1, \sigma_2) = (0.9, 1, 3)$ and $(\epsilon, \sigma_1, \sigma_2) = (0.75, 1, \sqrt{7})$, and the simulation results are presented in Table 2 and Table 3, respectively. For sequence III, we simulate $(\epsilon, \sigma, k) = (0.2, 1, 6)$ and $(\epsilon, \sigma, k) = (0.5, 1, 3)$, and the simulation results are presented in Table 4 and Table 5, respectively. The first figures in parentheses are the simulation results obtained by the empirical likelihood method, and the second figures are the simulation results obtained by the least square method.

From the simulation results in Tables 1–5, it can be seen that, for different error distribution, the confidence region constructed by the empirical likelihood method has a higher coverage probabilities for different parameters, sample sizes, pollution levels, and pollution distributions. Moreover, the confidence region constructed by the empirical likelihood method is closer to the confidence level 0.90. This shows that the empirical likelihood method is more robust than the least square method.

4 Real data analysis

In this section, we use our method to fit student teacher ratio (number of teachers = 1) data in Chinese universities, which are provided by the website of China National Bureau of Statistics (<http://data.stats.gov.cn/easyquery.htm?cn=C01&zb=A060E01&sj=2019>). Student teacher ratio is an important index to measure the level of universities. There are

Table 1 The simulation results for sequence I

$(\theta_1, \theta_2, \alpha_0, \alpha_1, \alpha_2)$	$n = 100$	$n = 300$	$n = 500$
(0.1, 0.1, 1, 0.1, 0.1)	(0.9520, 0.3140)	(0.9410, 0.4710)	(0.9330, 0.6950)
(0.1, 0.1, 3, 0.3, 0.3)	(0.9570, 0.3170)	(0.9440, 0.4890)	(0.9360, 0.6870)
(0.1, 0.1, 5, 0.5, 0.5)	(0.9520, 0.3170)	(0.9450, 0.4890)	(0.9330, 0.6880)
(0.1, 0.1, 7, 0.7, 0.7)	(0.9530, 0.3000)	(0.9350, 0.4850)	(0.9260, 0.7110)
(-0.1, -0.1, 1, 0.1, 0.1)	(0.9630, 0.3210)	(0.9330, 0.4870)	(0.9350, 0.7020)
(-0.1, -0.1, 3, 0.3, 0.3)	(0.9360, 0.3110)	(0.9510, 0.5020)	(0.9290, 0.6860)
(-0.1, -0.1, 5, 0.5, 0.5)	(0.9420, 0.3330)	(0.9370, 0.5100)	(0.9510, 0.7310)
(-0.1, -0.1, 7, 0.7, 0.7)	(0.9420, 0.3200)	(0.9480, 0.5250)	(0.9220, 0.6890)
(0.2, 0.2, 1, 0.1, 0.1)	(0.9580, 0.3290)	(0.9430, 0.5140)	(0.9330, 0.7120)
(0.2, 0.2, 3, 0.3, 0.3)	(0.9480, 0.3490)	(0.9310, 0.4920)	(0.9380, 0.7110)
(0.2, 0.2, 5, 0.5, 0.5)	(0.9540, 0.3540)	(0.9420, 0.4890)	(0.9290, 0.6900)
(-0.2, -0.2, 1, 0.1, 0.1)	(0.9530, 0.3290)	(0.9570, 0.5080)	(0.9220, 0.6860)
(-0.2, -0.2, 3, 0.3, 0.3)	(0.9590, 0.3640)	(0.9440, 0.5210)	(0.9350, 0.7020)
(-0.2, -0.2, 5, 0.5, 0.5)	(0.9550, 0.3400)	(0.9350, 0.5000)	(0.9280, 0.7130)
(0.3, 0.3, 1, 0.1, 0.1)	(0.9400, 0.3420)	(0.9440, 0.5250)	(0.9290, 0.6910)
(0.3, 0.3, 3, 0.3, 0.3)	(0.9530, 0.3470)	(0.9310, 0.5170)	(0.9390, 0.7440)
(-0.3, -0.3, 1, 0.1, 0.1)	(0.9410, 0.3610)	(0.9380, 0.5090)	(0.9360, 0.7050)
(-0.3, -0.3, 3, 0.3, 0.3)	(0.9460, 0.3840)	(0.9330, 0.5300)	(0.9320, 0.7010)
(0.4, 0.4, 1, 0.1, 0.1)	(0.9740, 0.3100)	(0.9530, 0.4120)	(0.9430, 0.7210)
(0.4, 0.4, 1, 0.3, 0.3)	(0.9580, 0.3700)	(0.9440, 0.5260)	(0.9370, 0.7090)
(-0.4, -0.4, 1, 0.1, 0.1)	(0.9540, 0.3850)	(0.9290, 0.7360)	(0.9120, 0.6950)
(-0.4, -0.4, 1, 0.3, 0.3)	(0.9530, 0.3870)	(0.9320, 0.7260)	(0.9310, 0.7170)
(0.5, 0.5, 1, 0.1, 0.1)	(0.9580, 0.4140)	(0.9380, 0.7350)	(0.9330, 0.7270)
(-0.5, -0.5, 1, 0.1, 0.1)	(0.9510, 0.3950)	(0.9420, 0.7280)	(0.9290, 0.7350)

Table 2 The simulation results for sequence II $(\epsilon, \sigma_1, \sigma_2) = (0.9, 1, 3)$

$(\theta_1, \theta_2, \alpha_0, \alpha_1, \alpha_2)$	$n = 100$	$n = 300$	$n = 500$
(0.1, 0.1, 1, 0.1, 0.1)	0.9770 (0.4330)	0.9510 (0.8740)	0.9490 (0.8620)
(0.1, 0.1, 3, 0.3, 0.3)	0.9820 (0.4590)	0.9370 (0.8530)	0.9410 (0.8500)
(0.1, 0.1, 5, 0.5, 0.5)	0.9740 (0.4300)	0.9540 (0.8560)	0.9580 (0.8740)
(0.1, 0.1, 7, 0.7, 0.7)	0.9760 (0.4340)	0.9370 (0.8660)	0.9480 (0.8540)
(-0.1, -0.1, 1, 0.1, 0.1)	0.9750 (0.4170)	0.9500 (0.8700)	0.9380 (0.8570)
(-0.1, -0.1, 3, 0.3, 0.3)	0.9820 (0.4320)	0.9570 (0.8610)	0.9620 (0.8700)
(-0.1, -0.1, 5, 0.5, 0.5)	0.9800 (0.4230)	0.9410 (0.8730)	0.9470 (0.8790)
(-0.1, -0.1, 7, 0.7, 0.7)	0.9730 (0.4380)	0.9580 (0.8660)	0.9450 (0.8430)
(0.2, 0.2, 1, 0.1, 0.1)	0.9820 (0.458)	0.9540 (0.8690)	0.9340 (0.8500)
(0.2, 0.2, 3, 0.3, 0.3)	0.9680 (0.4320)	0.9430 (0.8400)	0.9520 (0.8600)
(0.2, 0.2, 5, 0.5, 0.5)	0.9670 (0.4240)	0.9480 (0.8490)	0.9390 (0.8610)
(-0.2, -0.2, 1, 0.1, 0.1)	0.9730 (0.4510)	0.9440 (0.8520)	0.9410 (0.8660)
(-0.2, -0.2, 3, 0.3, 0.3)	0.9710 (0.4460)	0.9470 (0.8680)	0.9390 (0.8590)
(-0.2, -0.2, 5, 0.5, 0.5)	0.9820 (0.4510)	0.9450 (0.8700)	0.9480 (0.8490)
(0.3, 0.3, 1, 0.1, 0.1)	0.9790 (0.4490)	0.9530 (0.8440)	0.9670 (0.8690)
(0.3, 0.3, 3, 0.3, 0.3)	0.9810 (0.4440)	0.9540 (0.8440)	0.9470 (0.8490)
(-0.3, -0.3, 1, 0.1, 0.1)	0.9790 (0.4450)	0.9410 (0.8630)	0.9610 (0.8570)
(-0.3, -0.3, 3, 0.3, 0.3)	0.9790 (0.4320)	0.9550 (0.8560)	0.9400 (0.8350)
(0.4, 0.4, 1, 0.1, 0.1)	0.9780 (0.4420)	0.9320 (0.8310)	0.9390 (0.8320)
(0.4, 0.4, 1, 0.3, 0.3)	0.9790 (0.4090)	0.9460 (0.8370)	0.9440 (0.8610)
(-0.4, -0.4, 1, 0.1, 0.1)	0.9820 (0.4420)	0.9500 (0.8340)	0.9400 (0.8510)
(-0.4, -0.4, 1, 0.3, 0.3)	0.9820 (0.4170)	0.9540 (0.8490)	0.9480 (0.8260)
(0.5, 0.5, 1, 0.1, 0.1)	0.9840 (0.4220)	0.9480 (0.8430)	0.9490 (0.8320)
(-0.5, -0.5, 1, 0.1, 0.1)	0.9760 (0.4470)	0.9520 (0.8450)	0.9490 (0.8360)

70 available observations, which are denoted by X_1, X_2, \dots, X_{70} . The observations represent the yearly counts of student teacher ratio in China over the period from 1949 to 2018. Let $Y_t = X_t - X_{t-1}$. The plot of sample path, autocorrelation function (ACF), and partial autocorrelation function (PACF) for the series $\{Y_t\}$ are given in Figs. 1, 2, and 3, respectively.

Table 3 The simulation results for sequence II $(\epsilon, \sigma_1, \sigma_2) = (0.75, 1, \sqrt{7})$

$(\theta_1, \theta_2, \alpha_0, \alpha_1, \alpha_2)$	$n = 100$	$n = 300$	$n = 500$
(0.1, 0.1, 1, 0.1, 0.1)	0.9680 (0.3430)	0.9390 (0.7960)	0.9400 (0.8040)
(0.1, 0.1, 3, 0.3, 0.3)	0.9780 (0.3620)	0.9250 (0.7850)	0.9330 (0.7900)
(0.1, 0.1, 5, 0.5, 0.5)	0.9730 (0.3750)	0.9320 (0.8040)	0.9510 (0.7970)
(0.1, 0.1, 7, 0.7, 0.7)	0.9680 (0.3750)	0.9380 (0.7980)	0.9410 (0.7850)
(-0.1, -0.1, 1, 0.1, 0.1)	0.9690 (0.3640)	0.9380 (0.7750)	0.9340 (0.8000)
(-0.1, -0.1, 3, 0.3, 0.3)	0.9800 (0.3580)	0.9310 (0.7910)	0.9280 (0.7820)
(-0.1, -0.1, 5, 0.5, 0.5)	0.9690 (0.3480)	0.9240 (0.8050)	0.9350 (0.7860)
(-0.1, -0.1, 7, 0.7, 0.7)	0.9740 (0.3400)	0.9420 (0.7770)	0.9370 (0.7880)
(0.2, 0.2, 1, 0.1, 0.1)	0.9700 (0.3940)	0.9310 (0.7890)	0.9400 (0.7970)
(0.2, 0.2, 3, 0.3, 0.3)	0.9740 (0.3700)	0.9360 (0.7960)	0.9330 (0.7800)
(0.2, 0.2, 5, 0.5, 0.5)	0.9710 (0.3890)	0.9370 (0.7820)	0.9460 (0.7870)
(-0.2, -0.2, 1, 0.1, 0.1)	0.9700 (0.3890)	0.9370 (0.7820)	0.9230 (0.7760)
(-0.2, -0.2, 3, 0.3, 0.3)	0.9630 (0.3940)	0.9420 (0.8110)	0.9380 (0.7840)
(-0.2, -0.2, 5, 0.5, 0.5)	0.9650 (0.3780)	0.9270 (0.7810)	0.9280 (0.7740)
(0.3, 0.3, 1, 0.1, 0.1)	0.9660 (0.4020)	0.9190 (0.7720)	0.9470 (0.8060)
(0.3, 0.3, 3, 0.3, 0.3)	0.9680 (0.3680)	0.9330 (0.7940)	0.9450 (0.8080)
(-0.3, -0.3, 1, 0.1, 0.1)	0.9710 (0.4050)	0.9270 (0.7720)	0.9350 (0.7900)
(-0.3, -0.3, 3, 0.3, 0.3)	0.9820 (0.3950)	0.9440 (0.7900)	0.9350 (0.7850)
(0.4, 0.4, 1, 0.1, 0.1)	0.9670 (0.3610)	0.9420 (0.7940)	0.9420 (0.7750)
(0.4, 0.4, 1, 0.3, 0.3)	0.9690 (0.3740)	0.9360 (0.7960)	0.9470 (0.8010)
(-0.4, -0.4, 1, 0.1, 0.1)	0.9740 (0.4120)	0.9250 (0.7790)	0.9330 (0.7930)
(-0.4, -0.4, 1, 0.3, 0.3)	0.9760 (0.3970)	0.9440 (0.8030)	0.9260 (0.7890)
(0.5, 0.5, 1, 0.1, 0.1)	0.9800 (0.4140)	0.9310 (0.8010)	0.9320 (0.7960)
(-0.5, -0.5, 1, 0.1, 0.1)	0.9820 (0.4290)	0.9370 (0.7910)	0.9390 (0.7930)

Table 4 The simulation results for sequence III $(\epsilon, \sigma, k) = (0.2, 1, 6)$

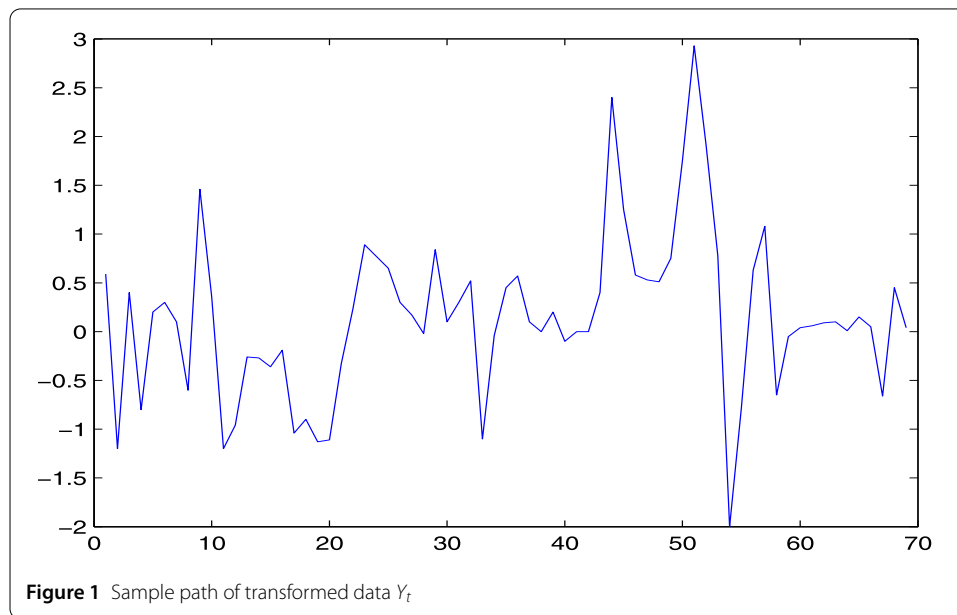
$(\theta_1, \theta_2, \alpha_0, \alpha_1, \alpha_2)$	$n = 100$	$n = 300$	$n = 500$
(0.1, 0.1, 1, 0.1, 0.1)	0.9560 (0.3180)	0.9400 (0.7260)	0.9490 (0.7200)
(0.1, 0.1, 3, 0.3, 0.3)	0.9500 (0.3130)	0.9460 (0.7090)	0.9380 (0.7000)
(0.1, 0.1, 5, 0.5, 0.5)	0.9510 (0.3440)	0.9320 (0.7330)	0.9420 (0.7180)
(0.1, 0.1, 7, 0.7, 0.7)	0.9570 (0.3270)	0.9400 (0.7110)	0.9270 (0.7230)
(-0.1, -0.1, 1, 0.1, 0.1)	0.9510 (0.3270)	0.9340 (0.7110)	0.9220 (0.6960)
(-0.1, -0.1, 3, 0.3, 0.3)	0.9580 (0.3640)	0.9200 (0.7030)	0.9350 (0.7330)
(-0.1, -0.1, 5, 0.5, 0.5)	0.9460 (0.3460)	0.9260 (0.7120)	0.9270 (0.7050)
(-0.1, -0.1, 7, 0.7, 0.7)	0.9640 (0.3350)	0.9120 (0.6850)	0.9390 (0.7320)
(0.2, 0.2, 1, 0.1, 0.1)	0.9630 (0.3550)	0.9310 (0.7180)	0.9330 (0.7140)
(0.2, 0.2, 3, 0.3, 0.3)	0.9520 (0.3470)	0.9290 (0.7310)	0.9440 (0.7400)
(0.2, 0.2, 5, 0.5, 0.5)	0.9390 (0.3330)	0.9360 (0.7250)	0.9500 (0.7210)
(-0.2, -0.2, 1, 0.1, 0.1)	0.9600 (0.3430)	0.9470 (0.7390)	0.9440 (0.7260)
(-0.2, -0.2, 3, 0.3, 0.3)	0.9560 (0.3290)	0.9480 (0.7220)	0.9370 (0.7020)
(-0.2, -0.2, 5, 0.5, 0.5)	0.9540 (0.3490)	0.9390 (0.7460)	0.9380 (0.7250)
(0.3, 0.3, 1, 0.1, 0.1)	0.9600 (0.3650)	0.9370 (0.7270)	0.9220 (0.7320)
(0.3, 0.3, 3, 0.3, 0.3)	0.9570 (0.3550)	0.9410 (0.7260)	0.9490 (0.7210)
(-0.3, -0.3, 1, 0.1, 0.1)	0.9610 (0.3520)	0.9410 (0.7320)	0.9390 (0.7150)
(-0.3, -0.3, 3, 0.3, 0.3)	0.9500 (0.3740)	0.9390 (0.7500)	0.9380 (0.7280)
(0.4, 0.4, 1, 0.1, 0.1)	0.9620 (0.3690)	0.9380 (0.7290)	0.9480 (0.7220)
(0.4, 0.4, 1, 0.3, 0.3)	0.9600 (0.3710)	0.9380 (0.7290)	0.9480 (0.7220)
(-0.4, -0.4, 1, 0.1, 0.1)	0.9610 (0.3810)	0.9420 (0.7370)	0.9310 (0.7290)
(-0.4, -0.4, 1, 0.3, 0.3)	0.9400 (0.3930)	0.9370 (0.7470)	0.9340 (0.7330)
(0.5, 0.5, 1, 0.1, 0.1)	0.9670 (0.3960)	0.9430 (0.7420)	0.9240 (0.7340)
(-0.5, -0.5, 1, 0.1, 0.1)	0.9670 (0.3950)	0.9150 (0.7170)	0.9230 (0.7330)

The corresponding plots of sample autocorrelation function (ACF) and partial autocorrelation function (PACF) indicate an AR(1)-like autocorrelation structure.

In what follows, based on the observation data $\{Y_t\}$, we give the figure of the empirical likelihood ratio confidence region when the confidence level is 0.95 (see Fig. 4). After a

Table 5 The simulation results for sequence III $(\epsilon, \sigma, k) = (0.5, 1, 3)$

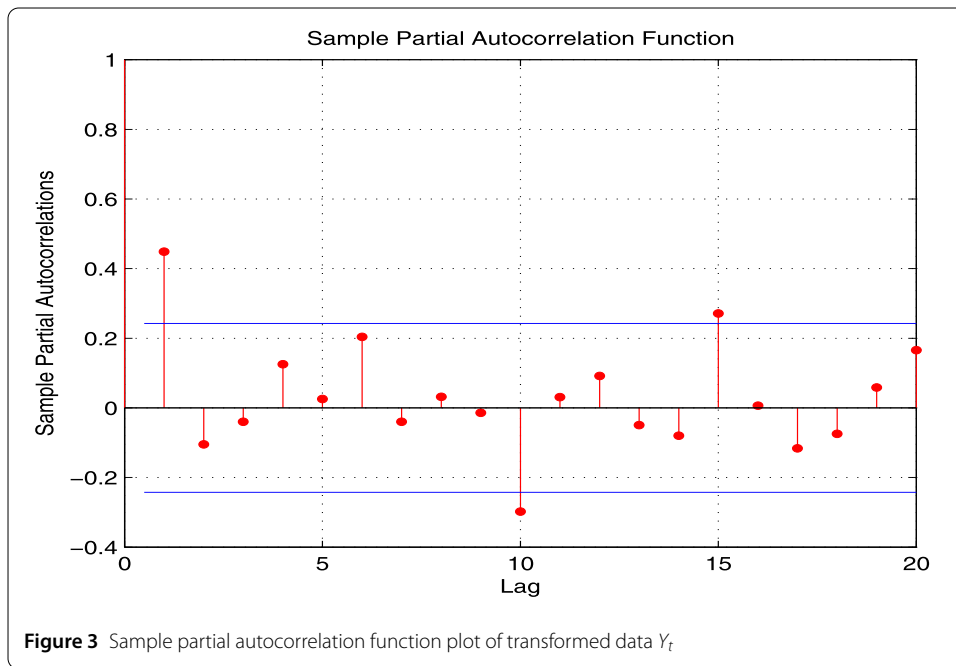
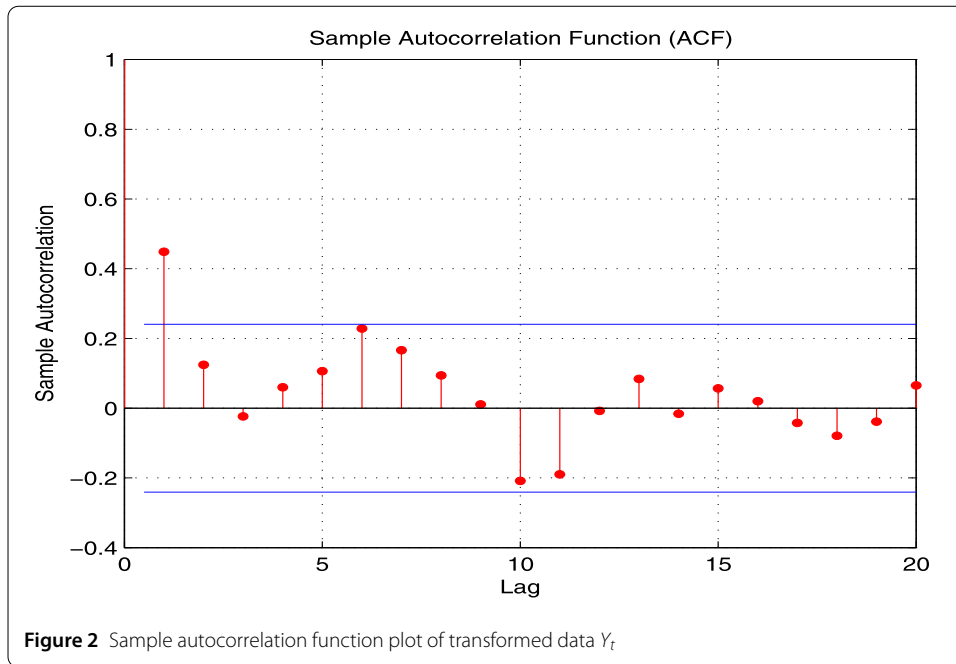
$(\theta_1, \theta_2, \alpha_0, \alpha_1, \alpha_2)$	$n = 100$	$n = 300$	$n = 500$
(0.1, 0.1, 1, 0.1, 0.1)	0.9650 (0.3300)	0.9410 (0.7330)	0.9510 (0.7660)
(0.1, 0.1, 3, 0.3, 0.3)	0.9560 (0.3260)	0.9410 (0.7390)	0.9280 (0.7440)
(0.1, 0.1, 5, 0.5, 0.5)	0.9720 (0.3320)	0.9310 (0.7280)	0.9310 (0.7140)
(0.1, 0.1, 7, 0.7, 0.7)	0.9560 (0.3050)	0.9490 (0.7280)	0.9210 (0.7130)
(-0.1, -0.1, 1, 0.1, 0.1)	0.9610 (0.3020)	0.9260 (0.7240)	0.9350 (0.7530)
(-0.1, -0.1, 3, 0.3, 0.3)	0.9710 (0.3300)	0.9390 (0.7310)	0.9430 (0.7560)
(-0.1, -0.1, 5, 0.5, 0.5)	0.9540 (0.3220)	0.9350 (0.7230)	0.9410 (0.7460)
(-0.1, -0.1, 7, 0.7, 0.7)	0.9690 (0.3390)	0.9360 (0.7070)	0.9330 (0.7200)
(0.2, 0.2, 1, 0.1, 0.1)	0.9530 (0.3210)	0.9460 (0.7350)	0.9240 (0.7140)
(0.2, 0.2, 3, 0.3, 0.3)	0.9670 (0.3140)	0.9400 (0.7260)	0.9290 (0.7550)
(0.2, 0.2, 5, 0.5, 0.5)	0.9670 (0.3610)	0.9390 (0.7420)	0.9510 (0.7660)
(-0.2, -0.2, 1, 0.1, 0.1)	0.9560 (0.3490)	0.9390 (0.7280)	0.9400 (0.7490)
(-0.2, -0.2, 3, 0.3, 0.3)	0.9660 (0.3570)	0.9360 (0.7090)	0.9350 (0.7180)
(-0.2, -0.2, 5, 0.5, 0.5)	0.9550 (0.3290)	0.9380 (0.7300)	0.9190 (0.7180)
(0.3, 0.3, 1, 0.1, 0.1)	0.9640 (0.3400)	0.9380 (0.7320)	0.9330 (0.7460)
(0.3, 0.3, 3, 0.3, 0.3)	0.9670 (0.3780)	0.9340 (0.7320)	0.9350 (0.7560)
(-0.3, -0.3, 1, 0.1, 0.1)	0.9640 (0.3600)	0.9400 (0.7470)	0.9500 (0.7450)
(-0.3, -0.3, 3, 0.3, 0.3)	0.9700 (0.3710)	0.9380 (0.5650)	0.9280 (0.7440)
(0.4, 0.4, 1, 0.1, 0.1)	0.9420 (0.7570)	0.9150 (0.5650)	0.9290 (0.7640)
(0.4, 0.4, 1, 0.3, 0.3)	0.9570 (0.3810)	0.9350 (0.7380)	0.9440 (0.7430)
(-0.4, -0.4, 1, 0.1, 0.1)	0.9610 (0.4130)	0.9540 (0.7430)	0.9360 (0.7470)
(-0.4, -0.4, 1, 0.3, 0.3)	0.9600 (0.3850)	0.9400 (0.7440)	0.9400 (0.7690)
(0.5, 0.5, 1, 0.1, 0.1)	0.9540 (0.3940)	0.9320 (0.754)	0.9250 (0.7640)
(-0.5, -0.5, 1, 0.1, 0.1)	0.9590 (0.4230)	0.9450 (0.7550)	0.9500 (0.7760)



simple calculation, we know that the least square estimation $\theta^* = (0.1616, 0.6191)$, and it is denoted by * in Fig. 4. From Fig. 4, we can see that the least square estimation θ^* is in the empirical likelihood ratio confidence region. Moreover, the empirical likelihood ratio confidence region is relatively small although the confidence level is 0.95.

5 Proofs

In order to establish Theorem 2.1, we first prove the following lemmas.

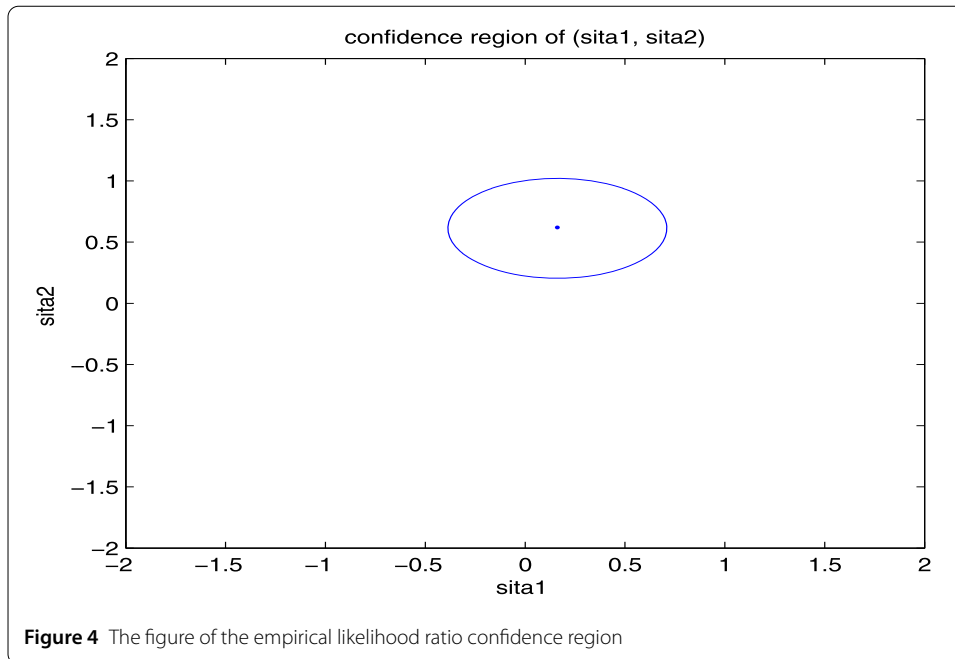


Lemma 5.1 *If (A1) and (A2) hold, then*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n H_t(\theta) \xrightarrow{d} N(0, D) \quad \text{as } n \rightarrow \infty, \tag{10}$$

where

$$D = \begin{pmatrix} E(\varepsilon_1^2(X_0^+)^2) & 0 \\ 0 & E(\varepsilon_1^2(X_0^-)^2) \end{pmatrix}.$$



Proof Note that

$$\begin{aligned}
 \sqrt{n}(\theta^* - \theta) &= \sqrt{n} \left(\left(\sum_{t=1}^n \mathcal{X}_t \mathcal{X}_t^\tau \right)^{-1} \sum_{t=1}^n X_t \mathcal{X}_t - \theta \right) \\
 &= \sqrt{n} \left(\left(\sum_{t=1}^n \mathcal{X}_t \mathcal{X}_t^\tau \right)^{-1} \sum_{t=1}^n X_t \mathcal{X}_t - \left(\sum_{t=1}^n \mathcal{X}_t \mathcal{X}_t^\tau \right)^{-1} \sum_{t=1}^n \mathcal{X}_t \mathcal{X}_t^\tau \theta \right) \\
 &= \sqrt{n} \left(\sum_{t=1}^n \mathcal{X}_t \mathcal{X}_t^\tau \right)^{-1} \left(\sum_{t=1}^n X_t \mathcal{X}_t - \sum_{t=1}^n \mathcal{X}_t \mathcal{X}_t^\tau \theta \right) \\
 &= \sqrt{n} \left(\sum_{t=1}^n \mathcal{X}_t \mathcal{X}_t^\tau \right)^{-1} \left(\sum_{t=1}^n \mathcal{X}_t (X_t - \mathcal{X}_t^\tau \theta) \right) \\
 &= \left(\frac{1}{n} \sum_{t=1}^n \mathcal{X}_t \mathcal{X}_t^\tau \right)^{-1} \frac{1}{\sqrt{n}} \left(\sum_{t=1}^n \mathcal{X}_t (X_t - \mathcal{X}_t^\tau \theta) \right).
 \end{aligned}$$

Therefore we have

$$\begin{aligned}
 \frac{1}{\sqrt{n}} \sum_{t=1}^n H_t(\theta) &= \frac{1}{\sqrt{n}} \sum_{t=1}^n (X_t - \mathcal{X}_t^\tau \theta) \mathcal{X}_t \\
 &= \frac{1}{n} \sum_{t=1}^n \mathcal{X}_t \mathcal{X}_t^\tau \left(\frac{1}{n} \sum_{t=1}^n \mathcal{X}_t \mathcal{X}_t^\tau \right)^{-1} \frac{1}{\sqrt{n}} \left(\sum_{t=1}^n \mathcal{X}_t (X_t - \mathcal{X}_t^\tau \theta) \right) \\
 &= \frac{1}{n} \sum_{t=1}^n \mathcal{X}_t \mathcal{X}_t^\tau \sqrt{n}(\theta^* - \theta).
 \end{aligned} \tag{11}$$

By the ergodic theorem, we have

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n \mathcal{X}_t \mathcal{X}_t^\tau &\xrightarrow{\text{a.s.}} E(\mathcal{X}_t \mathcal{X}_t^\tau) \\ &= E \left(\begin{pmatrix} X_{t-1}^+ \\ X_{t-1}^- \end{pmatrix} \begin{pmatrix} X_{t-1}^+ & X_{t-1}^- \end{pmatrix} \right) \\ &= E \begin{pmatrix} (X_{t-1}^+)^2 & 0 \\ 0 & (X_{t-1}^-)^2 \end{pmatrix} \\ &= \begin{pmatrix} E(X_{t-1}^+)^2 & 0 \\ 0 & E(X_{t-1}^-)^2 \end{pmatrix} \\ &\triangleq W. \end{aligned}$$

According to the result of Lemma 1 in Hwang and Woo [5], we have

$$\sqrt{n}(\theta^* - \theta) \xrightarrow{d} N(0, W^{-1}DW^{-1}) \quad \text{as } n \rightarrow \infty. \tag{12}$$

Combining with (11), we know that Lemma 5.1 holds. □

Lemma 5.2 *If (A1) and (A2) hold, then*

$$\frac{1}{n} \sum_{t=1}^n H_t(\theta)H_t^\tau(\theta) \xrightarrow{p} D \quad \text{as } n \rightarrow \infty. \tag{13}$$

Proof Notice that

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n H_t(\theta)H_t^\tau(\theta) &= \frac{1}{n} \sum_{t=1}^n (X_t - \mathcal{X}_t^\tau \theta)^2 \mathcal{X}_t \mathcal{X}_t^\tau \\ &= \frac{1}{n} \sum_{t=1}^n (X_t - \mathcal{X}_t^\tau \theta)^2 \begin{pmatrix} (X_{t-1}^+)^2 & 0 \\ 0 & (X_{t-1}^-)^2 \end{pmatrix} \\ &= \frac{1}{n} \sum_{t=1}^n \begin{pmatrix} (X_t - \mathcal{X}_t^\tau \theta)^2 (X_{t-1}^+)^2 & 0 \\ 0 & (X_t - \mathcal{X}_t^\tau \theta)^2 (X_{t-1}^-)^2 \end{pmatrix} \\ &= \frac{1}{n} \sum_{t=1}^n \begin{pmatrix} \varepsilon_t^2 (X_{t-1}^+)^2 & 0 \\ 0 & \varepsilon_t^2 (X_{t-1}^-)^2 \end{pmatrix}. \end{aligned}$$

Therefore, according to the ergodic theorem, Lemma 5.2 is established. □

Lemma 5.3 *If (A1) and (A2) hold, then*

$$\max_{1 \leq t \leq n} \|H_t(\theta)\| = o_p(n^{\frac{1}{2}}) \quad \text{as } n \rightarrow \infty. \tag{14}$$

Proof Based on assumption (A2), we know that $E(H_t(\theta)H_t^\tau(\theta)) < \infty$, which implies that

$$\sum_{n=1}^{\infty} P(H_t(\theta)H_t^\tau(\theta) > n) < \infty.$$

Further by assumption (A1), we know that the model is stationary, from which we can conclude that

$$\sum_{n=1}^{\infty} P(H_n(\theta)H_n^{\tau}(\theta) > n) < \infty.$$

By using the Borel–Cantelli lemma, we know that $\|H_t(\theta)\| > n^{\frac{1}{2}}$ holds for finite n , which implies that $\max_{1 \leq t \leq n} \|H_t(\theta)\| > n^{\frac{1}{2}}$ holds for finite n . Similarly, we can obtain that for $\forall \varepsilon > 0$, $\max_{1 \leq t \leq n} \|H_t(\theta)\| > \varepsilon n^{\frac{1}{2}}$ holds for finite n . Thus, Lemma 5.3 is established. \square

Lemma 5.4 *If (A1) and (A2) hold, then*

$$\|b\| = O_p(n^{-\frac{1}{2}}). \tag{15}$$

Proof Let $b = \|b\|_{\zeta}$. By (6), we have

$$\begin{aligned} 0 &= \frac{1}{n} \sum_{t=1}^n \frac{\zeta^{\tau} H_t(\theta)}{1 + b^{\tau}(\theta)H_t(\theta)} \\ &= \zeta^{\tau} \frac{1}{n} \sum_{t=1}^n H_t(\theta) \left(1 - \frac{b^{\tau}(\theta)H_t(\theta)}{1 + b^{\tau}(\theta)H_t(\theta)} \right) \\ &= \zeta^{\tau} \frac{1}{n} \sum_{t=1}^n H_t(\theta) - \zeta^{\tau} \frac{1}{n} \sum_{t=1}^n \frac{H_t(\theta)H_t^{\tau}(\theta)b(\theta)}{1 + b^{\tau}(\theta)H_t(\theta)} \\ &= \zeta^{\tau} \frac{1}{n} \sum_{t=1}^n H_t(\theta) - \zeta^{\tau} \frac{1}{n} \sum_{t=1}^n \frac{H_t(\theta)H_t^{\tau}(\theta)_{\zeta} \|b(\theta)\|}{1 + b^{\tau}(\theta)H_t(\theta)} \\ &= \zeta^{\tau} \frac{1}{n} \sum_{t=1}^n H_t(\theta) - \|b(\theta)\|_{\zeta} \zeta^{\tau} \frac{1}{n} \sum_{t=1}^n \frac{H_t(\theta)H_t^{\tau}(\theta)}{1 + b^{\tau}(\theta)H_t(\theta)}_{\zeta}. \end{aligned}$$

Hence we have

$$\zeta^{\tau} \frac{1}{n} \sum_{t=1}^n H_t(\theta) = \|b(\theta)\|_{\zeta} \zeta^{\tau} \tilde{D}_n_{\zeta}, \tag{16}$$

where

$$\tilde{D}_n = \frac{1}{n} \sum_{t=1}^n \frac{H_t(\theta)H_t^{\tau}(\theta)}{1 + b^{\tau}(\theta)H_t(\theta)}.$$

Let $D_n = \frac{1}{n} \sum_{t=1}^n H_t(\theta)H_t^{\tau}(\theta)$. From (5), we can see that $1 + b^{\tau}(\theta)H_t(\theta) > 0$. Thus we have

$$\begin{aligned} \|b(\theta)\|_{\zeta} \zeta^{\tau} D_n_{\zeta} &\leq \|b(\theta)\|_{\zeta} \zeta^{\tau} \frac{1}{n} \sum_{t=1}^n \frac{H_t(\theta)H_t^{\tau}(\theta)}{1 + b^{\tau}(\theta)H_t(\theta)}_{\zeta} \left(1 + \max_{1 \leq t \leq n} b^{\tau}(\theta)H_t(\theta) \right) \\ &\leq \|b(\theta)\|_{\zeta} \zeta^{\tau} \tilde{D}_n_{\zeta} \left(1 + \max_{1 \leq t \leq n} b^{\tau}(\theta)H_t(\theta) \right) \\ &\leq \|b(\theta)\|_{\zeta} \zeta^{\tau} \tilde{D}_n_{\zeta} \left(1 + \|b(\theta)\|_{\zeta} \max_{1 \leq t \leq n} \|H_t(\theta)\| \right) \end{aligned}$$

$$= \varsigma^\tau \frac{1}{n} \sum_{t=1}^n H_t(\theta) \left(1 + \|b(\theta)\| \max_{1 \leq t \leq n} \|H_t(\theta)\| \right), \tag{17}$$

which implies that

$$\|b(\theta)\| \left(\varsigma^\tau D_n \varsigma - \max_{1 \leq t \leq n} \|H_t(\theta)\| \varsigma^\tau \frac{1}{n} \sum_{t=1}^n H_t(\theta) \right) \leq \varsigma^\tau \frac{1}{n} \sum_{t=1}^n H_t(\theta). \tag{18}$$

According to Lemma 5.1, we have

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n H_t(\theta) = O_p(1), \tag{19}$$

which implies that

$$\varsigma^\tau \frac{1}{n} \sum_{t=1}^n H_t(\theta) = O_p(n^{-\frac{1}{2}}). \tag{20}$$

Further, by Lemma 5.3, we obtain

$$\begin{aligned} \max_{1 \leq t \leq n} \|H_t(\theta)\| \varsigma^\tau \frac{1}{n} \sum_{t=1}^n H_t(\theta) &= \frac{1}{\sqrt{n}} \varsigma^\tau \max_{1 \leq t \leq n} \|H_t(\theta)\| \frac{1}{\sqrt{n}} \sum_{t=1}^n H_t(\theta) \\ &= \frac{1}{\sqrt{n}} o_p(n^{\frac{1}{2}}) O_p(1) \\ &= o_p(1). \end{aligned} \tag{21}$$

Note that D is a positive definite matrix. Thus we have

$$\varsigma^\tau D_n \varsigma \xrightarrow{p} \varsigma^\tau D \varsigma > 0 \tag{22}$$

and

$$\sigma_{\min} + o_p(1) \leq \varsigma^\tau D_n \varsigma \leq \sigma_{\max} + o_p(1), \tag{23}$$

where σ_{\max} and σ_{\min} are the smallest and the largest eigenvalue of D , respectively. Combining with (18)–(23), we can obtain that

$$\|b(\theta)\| (\varsigma^\tau D_n \varsigma + o_p(1)) = O_p(n^{-\frac{1}{2}}). \tag{24}$$

Combined with (22), we know that $\|b(\theta)\| = O_p(n^{-\frac{1}{2}})$. Lemma 5.4 is established. \square

Lemma 5.5 *If (A1) and (A2) hold, then*

$$b(\theta) = \left(\sum_{t=1}^n H_t(\theta) H_t^T(\theta) \right)^{-1} \sum_{t=1}^n H_t(\theta) + B_n, \tag{25}$$

where

$$B_n = \left(\sum_{t=1}^n H_t(\theta)H_t^\tau(\theta) \right)^{-1} \sum_{t=1}^n \frac{H_t(\theta)(b^\tau(\theta)H_t(\theta))^2}{1 + b^\tau(\theta)H_t(\theta)}, \tag{26}$$

and

$$\|B_n\| = o_p(n^{-\frac{1}{2}}). \tag{27}$$

Proof By (6), we have

$$\begin{aligned} 0 &= \frac{1}{n} \sum_{t=1}^n \frac{H_t(\theta)}{1 + b^\tau(\theta)H_t(\theta)} \\ &= \frac{1}{n} \sum_{t=1}^n H_t(\theta) \left(1 - b^\tau(\theta)H_t(\theta) + \frac{(b^\tau(\theta)H_t(\theta))^2}{1 + b^\tau(\theta)H_t(\theta)} \right) \\ &= \frac{1}{n} \sum_{t=1}^n H_t(\theta) - \frac{1}{n} \sum_{t=1}^n H_t(\theta)H_t^\tau(\theta)b(\theta) + \frac{1}{n} \sum_{t=1}^n \frac{H_t(\theta)(b^\tau(\theta)H_t(\theta))^2}{1 + b^\tau(\theta)H_t(\theta)}. \end{aligned}$$

Thus, (25) can be established.

In what follows, we consider (27). Note that

$$\begin{aligned} \|B_n\| &= \left\| \left(\sum_{t=1}^n H_t(\theta)H_t^\tau(\theta) \right)^{-1} \sum_{t=1}^n \frac{H_t(\theta)(b^\tau(\theta)H_t(\theta))^2}{1 + b^\tau(\theta)H_t(\theta)} \right\| \\ &= \left\| \left(\frac{1}{n} \sum_{t=1}^n H_t(\theta)H_t^\tau(\theta) \right)^{-1} \frac{1}{n} \sum_{t=1}^n \frac{H_t(\theta)(b^\tau(\theta)H_t(\theta))^2}{1 + b^\tau(\theta)H_t(\theta)} \right\| \\ &\leq \left\| \left(\frac{1}{n} \sum_{t=1}^n H_t(\theta)H_t^\tau(\theta) \right)^{-1} \right\| \left\| \frac{1}{n} \sum_{t=1}^n \frac{H_t(\theta)(b^\tau(\theta)H_t(\theta))^2}{1 + b^\tau(\theta)H_t(\theta)} \right\| \\ &\leq \left\| \left(\frac{1}{n} \sum_{t=1}^n H_t(\theta)H_t^\tau(\theta) \right)^{-1} \right\| \|b^\tau(\theta)\|^2 \frac{1}{n} \sum_{t=1}^n \frac{\|H_t(\theta)\|^3}{\|1 + b^\tau(\theta)H_t(\theta)\|} \\ &\leq O_p(1)O_p(n^{-\frac{1}{2}})O_p(n^{-\frac{1}{2}})O_p(1) \\ &= o_p(n^{-\frac{1}{2}}). \end{aligned} \tag{28}$$

So (27) holds. □

Lemma 5.6 *If (A1) and (A2) hold, then*

$$\begin{aligned} -2\log(L(\theta)) &= \left(\sum_{t=1}^n H_t(\theta) \right)^\tau \left(\sum_{t=1}^n H_t(\theta)H_t^\tau(\theta) \right)^{-1} \sum_{t=1}^n H_t(\theta) \\ &\quad - B_n^\tau \sum_{t=1}^n H_t(\theta)H_t^\tau(\theta)B_n + 2 \sum_{t=1}^n \eta_t, \end{aligned} \tag{29}$$

where

$$B_n^\tau \sum_{t=1}^n H_t(\theta)H_t^\tau(\theta)B_n = o_p(1), \tag{30}$$

$$\sum_{t=1}^n \eta_t = o_p(1). \tag{31}$$

Proof We expand

$$\begin{aligned} -2\log(L(\theta)) &= 2 \sum_{t=1}^n \log(1 + b^\tau(\theta)H_t(\theta)) \\ &= 2 \sum_{t=1}^n b^\tau(\theta)H_t(\theta) - \sum_{t=1}^n (b^\tau(\theta)H_t(\theta))^2 + 2 \sum_{t=1}^n \eta_t \\ &= 2 \left(\left(\sum_{t=1}^n H_t(\theta)H_t^\tau(\theta) \right)^{-1} \sum_{t=1}^n H_t(\theta) + B_n \right)^\tau \sum_{t=1}^n H_t(\theta) \\ &\quad - \left(\left(\sum_{t=1}^n H_t(\theta)H_t^\tau(\theta) \right)^{-1} \sum_{t=1}^n H_t(\theta) + B_n \right)^\tau \left(\sum_{t=1}^n H_t(\theta)H_t^\tau(\theta) \right) \\ &\quad \times \left(\left(\sum_{t=1}^n H_t(\theta)H_t^\tau(\theta) \right)^{-1} \sum_{t=1}^n H_t(\theta) + B_n \right) \\ &\quad + 2 \sum_{t=1}^n \eta_t. \end{aligned} \tag{32}$$

After a simple algebraic operation, we know that (29) holds.

Next, we consider (30). Note that

$$\begin{aligned} \|B_n^\tau \sum_{t=1}^n H_t(\theta)H_t^\tau(\theta)B_n\| &\leq \|B_n^\tau\| \left\| \sum_{t=1}^n H_t(\theta)H_t^\tau(\theta) \right\| \|B_n\| \\ &= o_p(n^{-\frac{1}{2}})o_p(n^{-\frac{1}{2}})O_p(n) = o_p(1), \end{aligned} \tag{33}$$

which implies that (30) holds.

Last, we consider (31). For this, we first prove that there exists a finite real number $Q > 0$ such that

$$P(|\eta_t| \leq Q|b^\tau(\theta)H_t(\theta)|^3, 1 \leq t \leq n) \rightarrow 1 \quad \text{as } n \rightarrow \infty. \tag{34}$$

Consider the third-order Taylor expansion of $\log(1 + x)$ at $x = 0$:

$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \varpi(x),$$

where $\frac{\varpi(x)}{x^3} \rightarrow 0$ as $x \rightarrow 0$. Therefore, there exists $\iota > 0$ such that, for any $|x| < \iota$, $|\frac{\varpi(x)}{x^3}| < \frac{1}{6}$. In addition, note that

$$\max_{1 \leq t \leq n} \|b^\tau(\theta)H_t(\theta)\| = o_p(1).$$

Therefore, we have

$$\lim_{n \rightarrow \infty} P \left\{ \max_{1 \leq t \leq n} |b^\tau(\theta)H_t(\theta)|^3 < \iota^3 \right\} = 1.$$

Let $A_n = \{\omega : \max_{1 \leq t \leq n} |b^\tau(\theta)H_t(\theta)|^3 < \iota^3\}$. It is easy to prove that, for any $\omega \in A_n$ and $1 \leq t \leq n$,

$$\frac{|\eta_t|}{|b^\tau(\theta)H_t(\theta)|^3} = \frac{|\frac{(b^\tau(\theta)H_t(\theta))^3}{3} + \varpi(b^\tau(\theta)H_t(\theta))|}{|b^\tau(\theta)H_t(\theta)|^3} \leq \frac{1}{3} + \frac{1}{6} = \frac{1}{2}.$$

Thus we have

$$P(|\eta_t| \leq Q|b^\tau(\theta)H_t(\theta)|^3, 1 \leq t \leq n) \rightarrow 1 \quad \text{as } n \rightarrow \infty,$$

where $Q = \frac{1}{2}$. This implies that

$$\begin{aligned} \left\| \sum_{t=1}^n \eta_t \right\| &\leq Q \|b(\theta)\|^3 \sum_{t=1}^n \|H_t(\theta)\|^3 \\ &\leq O_p(n^{-\frac{3}{2}}) o_p(n^{\frac{3}{2}}) \\ &= o_p(1). \end{aligned}$$

□

Proof of Theorem 2.1 By Lemma 5.6, we can conclude that $-2\log(L(\theta))$ and $(\sum_{t=1}^n H_t(\theta))^\tau (\sum_{t=1}^n H_t(\theta)H_t^\tau(\theta))^{-1} \sum_{t=1}^n H_t(\theta)$ have the same limit distribution. By Lemma 5.1 and Lemma 5.2, we can conclude that

$$\sum_{t=1}^n H_t^\tau(\theta) \left(\sum_{t=1}^n H_t(\theta)H_t^\tau(\theta) \right)^{-1} \sum_{t=1}^n H_t(\theta) \xrightarrow{d} \chi^2(2).$$

Hence Theorem 2.1 holds.

□

Acknowledgements

We thank the editor and the anonymous referee for their constructive comments and suggestions, which greatly improved this paper.

Funding

This work is supported by the National Natural Science Foundation of China (No. 11571138, 11671054, 11301137, 11271155, 11371168, J1310022, 11501241). The national social science fund of China (16BTJ020).

Availability of data and materials

The data used to support the findings of this study are available from the corresponding author upon request.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors contributed equally and significantly in this manuscript, and they read and approved the final manuscript.

Author details

¹School of Foreign Languages, Jilin Normal University, Siping, 136000, P.R. China. ²College of Mathematics, Jilin Normal University, Siping, 136000, P.R. China.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 1 August 2020 Accepted: 2 March 2021 Published online: 22 March 2021

References

1. Tong, H.: *Nonlinear Time Series*. Oxford University Press, Oxford (1990)
2. Petruccielli, J.D., Woolford, S.W.: A threshold AR(1) model. *J. Appl. Probab.* **21**, 270–286 (1984)
3. Brockwell, P.J., Liu, J., Tweedie, R.L.: On the existence of stationary threshold autoregressive moving-average processes. *J. Time Ser. Anal.* **13**, 95–107 (1992)
4. Hwang, S.Y., Basawa, I.V.: Large sample inference for conditional exponential families with applications to nonlinear time series. *J. Stat. Plan. Inference* **38**, 141–158 (1994)
5. Hwang, S.Y., Woo, M.J.: Threshold ARCH(1) processes: asymptotic inference. *Stat. Probab. Lett.* **55**, 11–20 (2001)
6. Owen, A.B.: Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75**, 237–249 (1988)
7. Owen, A.B.: Empirical likelihood confidence regions. *Ann. Stat.* **18**, 90–120 (1990)
8. Owen, A.B.: *Empirical Likelihood*. Chapman & Hall, New York (2001)
9. Owen, A.B.: Empirical likelihood for linear models. *Ann. Stat.* **19**, 1725–1747 (1991)
10. Chen, S.X.: Empirical likelihood confidence intervals for linear regression coefficients. *J. Multivar. Anal.* **49**, 24–40 (1994)
11. Xue, L.G.: Empirical likelihood for linear models with missing responses. *J. Multivar. Anal.* **100**, 1353–1366 (2009)
12. Qin, Y.S., Li, Y.H.: Empirical likelihood for linear models under negatively associated errors. *J. Syst. Sci. Complex.* **102**, 153–163 (2011)
13. Qin, J., Lawless, J.: Empirical likelihood and general estimating equations. *Ann. Stat.* **22**, 300–325 (1994)
14. Kolaczyk, E.D.: Empirical likelihood for generalized linear models. *Stat. Sin.* **4**, 199–218 (1994)
15. Chen, S.X., Cui, H.: An extended empirical likelihood for generalized linear models. *Stat. Sin.* **13**, 69–81 (2003)
16. Xue, D., Xue, L.G., Cheng, W.H.: Empirical likelihood for generalized linear models with missing responses. *J. Stat. Plan. Inference* **141**, 2007–2020 (2011)
17. Bai, Y., Fung, W.K., Zhu, Z.: Weighted empirical likelihood for generalized linear models with longitudinal data. *J. Stat. Plan. Inference* **140**, 3446–3456 (2010)
18. Shi, J., Lau, T.S.: Empirical likelihood for partially linear models. *J. Multivar. Anal.* **72**, 132–148 (2000)
19. Li, G.R., Lin, L., Zhu, L.X.: Empirical likelihood for a varying coefficient partially linear model with diverging number of parameters. *J. Multivar. Anal.* **105**, 85–111 (2012)
20. Lu, X.W.: Empirical likelihood for heteroscedastic partially linear models. *J. Multivar. Anal.* **100**, 387–396 (2009)
21. Yan, L., Chen, X.: Empirical likelihood for partly linear models with errors in all variables. *J. Multivar. Anal.* **130**, 275–288 (2014)
22. Li, J.Y., Liang, W., He, S.Y., Wu, X.B.: Empirical likelihood for the smoothed LAD estimator in infinite variance autoregressive models. *Stat. Probab. Lett.* **80**, 1420–1430 (2010)
23. Zhao, Z.W., Wang, D.H.: Empirical likelihood for an autoregressive model with explanatory variables. *Commun. Stat., Theory Methods* **40**, 559–570 (2011)
24. Liu, X.H., Peng, L.: Asymptotic theory and unified confidence region for an autoregressive model. *J. Time Ser. Anal.* **40**, 43–65 (2019)
25. Zhao, Z.W., Wang, D.H.: Statistical inference for generalized random coefficient autoregressive model. *Math. Comput. Model.* **56**, 152–166 (2012)
26. Zhao, Z.W., Wang, D.H., Peng, C.X.: Coefficient constancy test in generalized random coefficient autoregressive model. *Appl. Math. Comput.* **219**, 10283–10292 (2013)
27. Zhao, Z.W., Wang, D.H., Peng, C.X., Zhang, M.L.: Empirical likelihood-based inference for stationary-ergodicity of the generalized random coefficient autoregressive model. *Commun. Stat., Theory Methods* **44**, 2586–2599 (2015)
28. Zhao, Z.W., Li, Y., Peng, C.X., Fu, Z.H.: Empirical likelihood-based inference in generalized random coefficient autoregressive model with conditional moment restrictions. *J. Comput. Appl. Math.* **348**, 146–160 (2019)
29. Zhang, H.X., Wang, D.H., Zhu, F.K.: Empirical likelihood inference for random coefficient INAR(p) process. *J. Time Ser. Anal.* **32**, 195–203 (2011)
30. Peng, C.X., Wang, D.H., Zhao, Z.W.: Empirical likelihood-based inference in Poisson autoregressive model with conditional moment restrictions. *J. Inequal. Appl.* **2015**, 218 (2015)
31. Ding, X., Wang, D.H.: Empirical likelihood inference for INAR(1) model with explanatory variables. *J. Korean Stat. Soc.* **45**, 623–632 (2016)
32. Zhao, Z.W., Wang, D.H., Peng, C.X.: Conditional heteroscedasticity test for Poisson autoregressive model. *Commun. Stat., Theory Methods* **46**, 4437–4448 (2017)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
