International Journal of Statistics and Management Systems
Vol. 3 No. 1 (January- June, 2018), ISSN: 0973-7395

Tests Based on Empirical Likelihood for an AR(1) Process with ARCH(1) Errors*

CLAUDIA KLÜPPELBERG† **and** LIANG PENG‡

Received: February 25, 2017; Accepted: September 22, 2017

Abstract

For an AR(1) process with ARCH(1) errors, we propose empirical likelihood tests for testing whether the sequence is strictly stationary but has infinite variance, or the sequence is an ARCH(1) sequence or the sequence is an iid sequence. Moreover, an empirical likelihood based confidence interval for the parameter in the AR part is proposed. All of these results do not require more than a finite second moment of the innovations. This includes the case of t-innovations for any degree of freedom larger than 2, which serves as a prominent model for real data.

1. Introduction Consider the following autoregressive model with ARCH(1) errors:

$$X_t = \alpha X_{t-1} + (\beta + \lambda X_{t-1}^2)^{1/2}\epsilon_t, \quad t \in \mathbb{N}, \tag{1}$$

where $\alpha \in \mathbb{R}, \beta > 0, \lambda \geq 0$, $\{\epsilon_t : t \in \mathbb{N}\}$ are independent and identically distributed (iid) random variables with mean zero and variance one, and X_0 is independent of $\{\epsilon_t : t \in \mathbb{N}\}$. Borkovec & Klüppelberg (2001) show the existence and uniqueness of a stationary distribution under some regularity conditions, and prove that the stationary distribution is heavy-tailed. Asymptotic normality of the quasi maximum likelihood estimator for the

---

Key words and phrases: ARCH model, empirical likelihood, stationary, weighted least squares.
AMS 2000 subject classifications. Primary 62M10; secondary 62F03.

†*Mailing Address:* Center for Mathematical Sciences, Munich University of Technology, D-85747 Garching, Germany. *E–mail:* cklu@ma.tum.de.

‡*Mailing Address:* School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, U.S.A. *E–mail:* peng@math.gatech.edu.

parameter vector (α, β, λ) is derived in Ling (2004) under the assumption that $E(\epsilon_t^4) < \infty$. Chan & Peng (2005) study the weighted least absolute deviations estimator and derive its asymptotic normality only assuming that $E(\epsilon_t^2) < \infty$.

Another important issue in economic and financial study is to test the stationarity of a model. It follows from Borkovec & Klüppelberg (2001) that $\{X_t\}$ is geometrically ergodic and has a unique stationary distribution if the following regularity conditions hold:

Condition 1. The noise ϵ_t has a symmetric, positive and continuous Lebesgue density in $(-\infty, \infty)$.

Condition 2. The parameter space is

$$\Theta = \left\{\theta = (\alpha, \beta, \lambda)^{\mathrm{T}} : E(\log|\alpha + \lambda^{1/2}\epsilon_t|) < 0, -\infty < \alpha < \infty, \beta > 0, \lambda \geq 0\right\}.$$

It is clear that $\{X_t\}$ is neither strictly nor weakly stationary when $(\alpha, \lambda) = (\pm 1, 0)$. Ling (2004) employs the Lagrange multiplier test to test the null hypothesis $(\alpha, \lambda) = (\pm 1, 0)$ against the alternative hypothesis $(\alpha, \lambda) \neq (\pm 1, 0)$. However, one can not claim that $\{X_t\}$ is stationary when the above null hypothesis is rejected. On the other hand, Klüppelberg et al. (2002) employ a pseudo-likelihood ratio test to test the null hypothesis $\alpha = 0, \beta > 0, \lambda = 0$ against the alternative hypothesis $\beta > 0, \lambda \geq 0, (\alpha, \lambda) \neq (0, 0)$. Note that both tests require that $E(\epsilon_t^4) < \infty$.

As shown in Remark 5 of Borkovec & Klüppelberg (2001) the strictly stationary distribution has finite second moment if and only if $\alpha^2 + \lambda E\epsilon^2 < 1$. Consequently, for $\alpha = 1$ the process $\{X_t\}$ is strictly but not weakly stationary as the second moment does not exist.

Define $\Theta_1 = \{(\alpha, \beta, \lambda)^{\mathrm{T}} : \alpha = 1\} \cap \Theta$, $\Theta_2 = \{(\alpha, \beta, \lambda)^{\mathrm{T}} : \alpha = 0\} \cap \Theta$ and $\Theta_3 = \{(\alpha, \beta, \lambda)^{\mathrm{T}} : \alpha = 0, \lambda = 0\} \cap \Theta$. In this paper, we propose to apply the empirical likelihood method to test the following three different tests:

$$H_0^{(i)} : \theta \in \Theta_i \quad \text{against} \quad H_1^{(i)} : \theta \in \Theta \setminus \Theta_i$$

for $i = 1, 2, 3$. We remark that $H_0^{(1)}, H_0^{(2)}$ and $H_0^{(3)}$ imply that $\{X_t\}$ is strictly stationary but not weakly stationary, is an ARCH(1) sequence and is an iid sequence, respectively.

The empirical likelihood method as a non-parametric robust statistical method has many advantages in comparison to parametric likelihood methods, see Owen (2001). Recently, Chuang & Chan (2002) applied the empirical likelihood method to unit root AR models with finite variance errors, and Chan et al. (2005) apply the empirical likelihood method to near-integrated AR models with infinite variance errors.

We organize this paper as follows. In Section 2, the empirical likelihood tests are proposed. Moreover, an empirical likelihood based confidence interval for α is given. A simulation study supports our theory in Section 3. All proofs are postponed to the appendix.

2. Empirical Likelihood Method Throughout we assume that the median of ϵ_t^2 is m, which is unknown. Rewrite model (1) as

$$(X_t - \alpha X_{t-1})^2 - (\beta m + \lambda m X_{t-1}^2) = (\beta m + \lambda m X_{t-1}^2)(\frac{\epsilon_t^2}{m} - 1). \tag{2}$$

When m is assumed to be known and equal to one, Chan & Peng (2005) propose the following weighted least absolute deviations estimator for $\theta^* = (\alpha^*, \beta^*, \lambda^*)^{\mathrm{T}} = (2\alpha, \beta m, \alpha^2 - \lambda m)^{\mathrm{T}}$, which is defined as

$$\hat{\theta}^* = (\hat{\alpha}^*, \hat{\beta}^*, \hat{\lambda}^*)^{\mathrm{T}} = \arg \min_{(\alpha^*,\beta^*,\lambda^*)} \sum_{t=1}^{n} \frac{1}{1+X_{t-1}^2} |X_t^2 - \alpha^* X_t X_{t-1} - \beta^* + \lambda^* X_{t-1}^2|. \tag{3}$$

Here we propose to employ the empirical likelihood method to the above weighted least absolute deviations with unknown m as follows.

Let $p = (p_1, \ldots, p_n)$ be a probability vector, i.e., $\sum_{i=1}^n p_i = 1$ and $p_i \geq 0$ for $i = 1, \ldots, n$. Put $Y_t(\theta^*) = X_t^2 - \alpha^* X_t X_{t-1} - \beta^* + \lambda^* X_{t-1}^2$ and $Z_t = (-\frac{X_t X_{t-1}}{1+X_{t-1}^2}, -\frac{1}{1+X_{t-1}^2}, \frac{X_{t-1}^2}{1+X_{t-1}^2})^{\mathrm{T}}$ for $t = 1, \ldots, n$. Then the empirical likelihood is defined as

$$L(\theta^*) = \sup\left\{\prod_{t=1}^n p_t : \sum_{t=1}^n p_t = 1, p_t \geq 0, \sum_{t=1}^n p_t Z_t sgn(Y_t(\theta^*)) = 0\right\},$$

where $sgn(x)$ equals 1 if $x \geq 0$, and -1 if $x < 0$. By the method of Lagrange multipliers, we have

$$p_t = \frac{1}{n}\{1 + \gamma^{\mathrm{T}} Z_t sgn(Y_t(\theta^*))\}^{-1}, \quad t = 1, \ldots, n, \tag{4}$$

where $\gamma = (\gamma_1, \gamma_2, \gamma_3)^{\mathrm{T}}$ satisfies

$$g(\gamma) = \frac{1}{n}\sum_{t=1}^n \frac{Z_t sgn(Y_t(\theta^*))}{1 + \gamma^{\mathrm{T}} Z_t sgn(Y_t(\theta^*))} = 0. \tag{5}$$

The empirical likelihood ratio is defined as

$$l(\theta^*) = 2\sum_{t=1}^n \log\{1 + \gamma^{\mathrm{T}} Z_t sgn(Y_t(\theta^*))\}.$$

Our main results are as follows.

Theorem 1. Suppose model (1) holds with Conditions 1 and 2. Then

$$l_p(\alpha_0) = \arg \min_{\theta^*=(2\alpha_0,\beta^*,\lambda^*)^{\mathrm{T}}} l(\theta^*) - \arg \min_{\theta^*} l(\theta^*) \xrightarrow{d} \chi^2(1),$$

where α_0 denotes the true value of α. Therefore, an asymptotic confidence interval for α_0 with significance level 100a% is

$$I_a = \{\alpha : l_p(\alpha) \le u_a\},$$

where u_a denotes the 100a%-level quantile of $\chi^2(1)$.

Theorem 2. Suppose model (1) holds with Conditions 1 and 2. Then,
(i) under $H_0^{(1)}$, we have

$$T_1 = \arg \min_{\theta^*=(2,\beta^*,\lambda^*)^{\mathrm{T}}} l(\theta^*) - \arg \min_{\theta^*} l(\theta^*) \xrightarrow{d} \chi^2(1);$$

(ii) under $H_0^{(2)}$, we have

$$T_2 = \arg \min_{\theta^*=(0,\beta^*,\lambda^*)^{\mathrm{T}}} l(\theta^*) - \arg \min_{\theta^*} l(\theta^*) \xrightarrow{d} \chi^2(1);$$

(iii) under $H_0^{(3)}$, we have

$$T_3 = \arg \min_{\theta^*=(0,\beta^*,0)^{\mathrm{T}}} l(\theta^*) - \arg \min_{\theta^*} l(\theta^*) \xrightarrow{d} \chi^2(2).$$

Remark 1. Klüppelberg et al. (2002) employ the pseudo likelihood ratio test to test $H_0^{(3)}$, but obtained a different limiting distribution from that given in case (iii) of Theorem 2. The reason is that $\theta^* = (0, \beta^*, 0)^{\mathrm{T}}$ is not at the boundary of the parameter set of θ^* although $\theta = (0, \beta, 0)^{\mathrm{T}}$ is indeed at the boundary of the parameter set of θ. Moreover, the limit in Klüppelberg et al. (2002) involves the fourth moment of ϵ_t.

3. Numerical Studies We investigate the finite sample behaviors of our tests based on empirical likelihood for the case, where the fourth moment of the innovations is infinite. Since other methods like the pseudo likelihood ratio test in Klüppelberg et al. (2002) and the Lagrange multiplier test in Ling (2004) require finite fourth moment, we concentrate on noise variables with infinite fourth moment.

We draw 1 000 random samples with size $n = 1000$ from model (1) with $\alpha = 1 - \delta/n$, $\beta = 1$, $\lambda = 0.5$ for testing $H_0^{(1)}$, $\alpha = \delta/n$, $\beta = 1$, $\lambda = 0.5$ for testing $H_0^{(2)}$, and $\alpha = \delta/n$, $\lambda = \delta/n$, $\beta = 1$ for testing $H_0^{(3)}$. We consider $\delta = 0, 1, 10, 50, 100, 500$, and ϵ_t having a standardized $t(3)$ or $t(4)$ distribution such that $E(\epsilon_t^2) = 1$. For the significance level 0.05, we compute the empirical sizes and powers based on our empirical likelihood method, see Table 1. We conclude from Table 1 that the sizes of the tests based on empirical likelihood are reasonably close to the nominal level 0.05 and the powers show that these tests are powerful.

$H_0^{(1)} : \alpha = 1$						
δ	0	1	10	50	100	500
$t(3)$	0.043	0.041	0.052	0.226	0.644	0.999
$t(4)$	0.036	0.033	0.053	0.223	0.690	1.000
$H_0^{(2)} : \alpha = 0$						
δ	0	1	10	50	100	500
$t(3)$	0.051	0.043	0.053	0.125	0.358	0.998
$t(4)$	0.035	0.049	0.047	0.120	0.389	1.000
$H_0^{(3)} : \alpha = 0, \lambda = 0$						
δ	0	1	10	50	100	500
$t(3)$	0.041	0.039	0.043	0.161	0.565	1.000
$t(4)$	0.044	0.034	0.051	0.163	0.609	1.000

Table 1: Empirical sizes and powers of the tests based on empirical likelihood at the significance level 0.05.

4. Appendix: Proofs

Proof of Theorem 1. Define $v = (v_1, v_2, v_3)^{\mathrm{T}}$, $v_1 = n^{1/2}(\alpha^* - \alpha_0^*)$, $v_2 = n^{1/2}(\beta^* - \beta_0^*)$, and $v_3 = n^{1/2}(\lambda^* - \lambda_0^*)$, where $\theta_0^* = (\alpha_0^*, \beta_0^*, \lambda_0^*)^T$ denotes the true value of θ^*. Our first step is to prove that

$$||\gamma|| = O_p(n^{-1/2}) \quad \text{locally uniformly in } v, \tag{6}$$

where $||\cdot||$ denotes the Euclidean norm. Write $\gamma = \rho\gamma_0$, where $\rho \geq 0$ and $||\gamma_0|| = 1$. By (4), we have

$$1 + \gamma^{\mathrm{T}} Z_t sgn(Y_t(\theta^*)) > 0,$$

i.e.,

$$(1+\gamma^{\mathrm{T}} Z_t sgn(Y_t(\theta^*)))^{-1} = (1+\rho\gamma_0^{\mathrm{T}} Z_t sgn(Y_t(\theta^*)))^{-1} \geq \{1+\rho \max_{1\leq t\leq n} ||Z_t sgn(Y_t(\theta^*))||\}^{-1}.$$

Hence

$$\begin{aligned}
0 &= ||g(\gamma)|| = ||g(\rho\gamma_0)|| \\
&\geq |\gamma_0^{\mathrm{T}} g(\rho\gamma_0)| \\
&= \frac{1}{n}\left|\gamma_0^{\mathrm{T}}\{\sum_{t=1}^{n} Z_t sgn(Y_t(\theta^*)) - \rho\sum_{t=1}^{n}\frac{Z_t sgn(Y_t(\theta^*))\gamma_0^{\mathrm{T}} Z_t sgn(Y_t(\theta^*))}{1+\rho\gamma_0^{\mathrm{T}} Z_t sgn(Y_t(\theta^*))}\}\right| \\
&\geq \frac{\rho}{n}\gamma_0^{\mathrm{T}}\sum_{t=1}^{n}\frac{Z_t Z_t^{\mathrm{T}}}{1+\rho\gamma_0^{\mathrm{T}} Z_t sgn(Y_t(\theta^*))}\gamma_0 - \frac{1}{n}\left|\gamma_0^{\mathrm{T}}\sum_{t=1}^{n} Z_t sgn(Y_t(\theta^*))\right| \\
&\geq \frac{\rho}{n}\{1+\rho\max_{1\leq t\leq n}||Z_t sgn(Y_t(\theta^*))||\}^{-1}\gamma_0^{\mathrm{T}}\sum_{t=1}^{n} Z_t Z_t^{\mathrm{T}}\gamma_0 - \frac{1}{n}\left|\gamma_0^{\mathrm{T}}\sum_{t=1}^{n} Z_t sgn(Y_t(\theta^*))\right|.
\end{aligned}$$

That is,

$$\begin{aligned}
&\rho\left\{\frac{1}{n}\gamma_0^{\mathrm{T}}\sum\nolimits_{t=1}^{n} Z_t Z_t^{\mathrm{T}}\gamma_0 - (\max\nolimits_{1\leq t\leq n}||Z_t sgn(Y_t(\theta^*))||)\frac{1}{n}\left|\sum\nolimits_{t=1}^{n}\gamma_0^{\mathrm{T}} Z_t sgn(Y_t(\theta^*))\right|\right\} \\
&\leq \frac{1}{n}\left|\sum\nolimits_{t=1}^{n}\gamma_0^{\mathrm{T}} Z_t sgn(Y_t(\theta^*))\right|.
\end{aligned} \tag{7}$$

Recall that $v_1 = n^{1/2}(\alpha^* - \alpha_0^*)$ and define $\Delta_{t-1} = \beta_0 + \lambda_0 X_{t-1}^2$. Then,

$$Y_t(\theta^*) = \left\{\Delta_{t-1}^{1/2}\epsilon_t - \frac{v_1 X_{t-1}}{2n^{1/2}}\right\}^2 - \beta m - \lambda m X_{t-1}^2,$$

and, denoting by m_0 the true median, we have

$$\begin{aligned}
& sgn(Y_t(\theta^*)) - sgn(Y_t(\theta_0^*)) \\
=\ & 2\{I(Y_t(\theta_0^*) < 0) - I(Y_t(\theta^*) < 0)\} \\
=\ & 2\Big\{I(-\sqrt{m_0} < \epsilon_t < \sqrt{m_0}) \\
& -I\Big(\frac{-(\beta m + \lambda m X_{t-1}^2)^{1/2} + 2^{-1}n^{-1/2}v_1 X_{t-1}}{\Delta_{t-1}^{1/2}} < \epsilon_t \\
& < \frac{(\beta m + \lambda m X_{t-1}^2)^{1/2} + 2^{-1}n^{-1/2}v_1 X_{t-1}}{\Delta_{t-1}^{1/2}}\Big)\Big\}.
\end{aligned}$$

Let F denote the distribution function of ϵ_t and put

$$S_{t-1} = 1 + X_{t-1}^2 \quad \text{and} \quad h(c,d) = E\{\epsilon_t I(c < \epsilon_t < d)\}.$$

Then, we can write

$$\begin{aligned}
& \tfrac{1}{\sqrt{n}} \textstyle\sum_{t=1}^n \tfrac{X_t X_{t-1}}{S_{t-1}} \Big\{ sgn(Y_t(\theta^*)) - sgn(Y_t(\theta_0^*)) \Big\} \\
= \; & \tfrac{2}{\sqrt{n}} \textstyle\sum_{t=1}^n \tfrac{\alpha_0 X_{t-1}^2}{S_{t-1}} \Big\{ I(-\sqrt{m_0} < \epsilon_t < \sqrt{m_0}) - F(\sqrt{m_0}) + F(-\sqrt{m_0}) \Big\} \\
& - \tfrac{2}{\sqrt{n}} \textstyle\sum_{t=1}^n \tfrac{\alpha_0 X_{t-1}^2}{S_{t-1}} \Big\{ I\Big(\tfrac{-(\beta m + \lambda m X_{t-1}^2)^{1/2} + 2^{-1} n^{-1/2} v_1 X_{t-1}}{\Delta_{t-1}^{1/2}} < \epsilon_t \\
& < \tfrac{(\beta m + \lambda m X_{t-1}^2)^{1/2} + 2^{-1} n^{-1/2} v_1 X_{t-1}}{\Delta_{t-1}^{1/2}} \Big) \\
& - F\Big(\frac{(\beta m + \lambda m X_{t-1}^2)^{1/2} + 2^{-1} n^{-1/2} v_1 X_{t-1}}{\Delta_{t-1}^{1/2}} \Big) \\
& + F\Big(\frac{-(\beta m + \lambda m X_{t-1}^2)^{1/2} + 2^{-1} n^{-1/2} v_1 X_{t-1}}{\Delta_{t-1}^{1/2}} \Big) \Big\} \\
& + \tfrac{2}{\sqrt{n}} \textstyle\sum_{t=1}^n \tfrac{X_{t-1} \Delta_{t-1}^{1/2}}{S_{t-1}} \Big\{ \epsilon_t I(-\sqrt{m_0} < \epsilon_t < \sqrt{m_0}\Big) - h(-\sqrt{m_0}, \sqrt{m_0}) \Big\} \\
& - \tfrac{2}{\sqrt{n}} \textstyle\sum_{t=1}^n \tfrac{X_{t-1} \Delta_{t-1}^{1/2}}{S_{t-1}} \Big\{ \epsilon_t I\Big(\tfrac{-(\beta m + \lambda m X_{t-1}^2)^{1/2} + 2^{-1} n^{-1/2} v_1 X_{t-1}}{\Delta_{t-1}^{1/2}} < \epsilon_t \\
& < \tfrac{(\beta m + \lambda m X_{t-1}^2)^{1/2} + 2^{-1} n^{-1/2} v_1 X_{t-1}}{\Delta_{t-1}^{1/2}} \Big) \\
& - h\Big(\tfrac{-(\beta m + \lambda m X_{t-1}^2)^{1/2} + 2^{-1} n^{-1/2} v_1 X_{t-1}}{\Delta_{t-1}^{1/2}}, \\
& \tfrac{(\beta m + \lambda m X_{t-1}^2)^{1/2} + 2^{-1} n^{-1/2} v_1 X_{t-1}}{\Delta_{t-1}^{1/2}} \Big) \Big\} \\
& + \frac{2}{\sqrt{n}} \textstyle\sum_{t=1}^n \tfrac{\alpha_0 X_{t-1}^2}{S_{t-1}} \Big\{ F(\sqrt{m_0}) - F(-\sqrt{m_0}) \\
& - F\Big(\tfrac{(\beta m + \lambda m X_{t-1}^2)^{1/2} + 2^{-1} n^{-1/2} v_1 X_{t-1}}{\Delta_{t-1}^{1/2}} \Big) \\
& + F\Big(\frac{-(\beta m + \lambda m X_{t-1}^2)^{1/2} + 2^{-1} n^{-1/2} v_1 X_{t-1}}{\Delta_{t-1}^{1/2}} \Big) \Big\} \\
& + \tfrac{2}{\sqrt{n}} \textstyle\sum_{t=1}^n \tfrac{X_{t-1} \Delta_{t-1}^{1/2}}{S_{t-1}} \Big\{ h(-\sqrt{m_0}, \sqrt{m_0}) \\
& - h\Big(\frac{-(\beta m + \lambda m X_{t-1}^2)^{1/2} + 2^{-1} n^{-1/2} v_1 X_{t-1}}{\Delta_{t-1}^{1/2}}, \\
& \frac{(\beta m + \lambda m X_{t-1}^2)^{1/2} + 2^{-1} n^{-1/2} v_1 X_{t-1}}{\Delta_{t-1}^{1/2}} \Big) \Big\} \\
=: \; & I_1 + \cdots + I_6.
\end{aligned}$$

By Corollary 3.1 of Hall & Heyde (1980), we can show that

$$|I_1 + I_2| = o_p(1) \quad \text{and} \quad |I_3 + I_4| = o_p(1) \tag{8}$$

locally uniformly in v. Hence it follows from (8), Condition 1, and the ergodicity result in Borkovec & Klüppelberg (1998) that

$$\begin{aligned} & \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{X_tX_{t-1}}{S_{t-1}}\{sgn(Y_t(\theta^*))-sgn(Y_t(\theta_0^*))\} \\ = & \{-2\alpha_0^2 f(\sqrt{m_0})E(\frac{X_1^4}{S_1\Delta_1})-2f(\sqrt{m_0})\sqrt{m_0}E(\frac{X_1^2}{S_1})\}v_1 \\ & -2\alpha_0 f(\sqrt{m_0})E(\frac{X_1^2}{S_1\Delta_1})v_2+2\alpha_0 f(\sqrt{m_0})E(\frac{X_1^4}{S_1\Delta_1})v_3+o_p(1) \end{aligned} \tag{9}$$

locally uniformly in v. Similarly,

$$\begin{aligned} & \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{1}{S_{t-1}}\{sgn(Y_t(\theta^*))-sgn(Y_t(\theta_0^*))\} \\ = & -2\alpha_0 f(\sqrt{m_0})E(\frac{X_1^2}{S_1\Delta_1})v_1-2f(\sqrt{m_0})E(\frac{1}{S_1\Delta_1})v_2 \\ & +2f(\sqrt{m_0})E(\frac{X_1^2}{S_1\Delta_1})v_3+o_p(1) \end{aligned} \tag{10}$$

and

$$\begin{aligned} & \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{X_{t-1}^2}{S_{t-1}}\{sgn(Y_t(\theta^*))-sgn(Y_t(\theta_0^*))\} \\ = & -2\alpha_0 f(\sqrt{m_0})E(\frac{X_1^4}{S_1\Delta_1})v_1-2f(\sqrt{m_0})E(\frac{X_1^2}{S_1\Delta_1})v_2 \\ & +2f(\sqrt{m_0})E(\frac{X_1^4}{S_1\Delta_1})v_3+o_p(1) \end{aligned} \tag{11}$$

locally uniformly in v. Thus, by (9) - (11),

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}Z_t\{sgn(Y_t(\theta^*))-sgn(Y_t(\theta_0^*))\}=\Sigma_1 v+o_p(1) \tag{12}$$

locally uniformly in v, where

$$\Sigma_1=2f(\sqrt{m_0})\begin{pmatrix} \alpha_0^2E(\frac{X_1^4}{S_1\Delta_1})+\sqrt{m_0}E(\frac{X_1^2}{S_1}) & \alpha_0E(\frac{X_1^2}{S_1\Delta_1}) & -\alpha_0E(\frac{X_1^4}{S_1\Delta_1}) \\ \alpha_0E(\frac{X_1^2}{S_1\Delta_1}) & E(\frac{1}{S_1\Delta_1}) & -E(\frac{X_1^2}{S_1\Delta_1}) \\ -\alpha_0E(\frac{X_1^4}{S_1\Delta_1}) & -E(\frac{X_1^2}{S_1\Delta_1}) & E(\frac{X_1^4}{S_1\Delta_1}) \end{pmatrix}.$$

Then it follows from (12) and the proof of Theorem 1 in Chan & Peng (2005) that

$$||\frac{1}{n}\sum_{t=1}^{n}Z_t sgn(Y_t(\theta^*))||=O_p(n^{-1/2}) \quad \text{locally uniformly in} \quad v \tag{13}$$

and

$$\frac{1}{n}\gamma_0^{\mathrm{T}}\sum_{t=1}^{n} Z_t Z_t^{\mathrm{T}}\gamma_0 \xrightarrow{p} \gamma_0^{\mathrm{T}}\Sigma_2\gamma_0, \tag{14}$$

where

$$\Sigma_2 = \begin{pmatrix} \alpha_0^2 E(X_1^4/S_1^2) + \beta_0 E(X_1^2/S_1^2) + \lambda_0 E(X_1^4/S_1^2) & \alpha_0 E(X_1^2/S_1^2) & -\alpha_0 E(X_1^4/S_1^2) \\ \alpha_0 E(X_1^2/S_1^2) & E(1/S_1^2) & -E(X_1^2/S_1^2) \\ -\alpha_0 E(X_1^4/S_1^2) & -E(X_1^2/S_1^2) & E(X_1^4/S_1^2) \end{pmatrix}.$$

Apply conditional expectation arguments to the proof of Lemma 3 of Owen (1990), we have

$$\max_{1\le t\le n} ||Z_t sgn(Y_t(\theta^*))|| = o(n^{1/2}) \quad \text{locally uniformly in} \quad v \tag{15}$$

with probability one as $n \to \infty$. Hence, (6) follows from (7) - (15). Furthermore

$$\gamma = \Big\{\frac{1}{n}\sum_{t=1}^{n} Z_t Z_t^{\mathrm{T}}\Big\}^{-1}\frac{1}{n}\sum_{t=1}^{n} Z_t sgn(Y_t(\theta^*)) + o_p(n^{-1/2})$$

and

$$l(\theta^*) = n\Big\{\frac{1}{n}\sum_{t=1}^{n} Z_t sgn(Y_t(\theta^*))\Big\}^{\mathrm{T}}\Big\{\frac{1}{n}\sum_{t=1}^{n} Z_t Z_t^{\mathrm{T}}\Big\}^{-1}\Big\{\frac{1}{n}\sum_{t=1}^{n} Z_t sgn(Y_t(\theta^*))\Big\} + o_p(1) \tag{16}$$

locally uniformly in v. Similarly we can show that

$$l(\theta_0^*) = n\Big\{\frac{1}{n}\sum_{t=1}^{n} Z_t sgn(Y_t(\theta_0^*))\Big\}^{\mathrm{T}}\Big\{\frac{1}{n}\sum_{t=1}^{n} Z_t Z_t^{\mathrm{T}}\Big\}^{-1}\Big\{\frac{1}{n}\sum_{t=1}^{n} Z_t sgn(Y_t(\theta_0^*))\Big\} + o_p(1) \tag{17}$$

locally uniformly in v. Using (12), (14), (16) and (17), we have

$$l(\theta^*) - l(\theta_0^*) = v^{\mathrm{T}}\Sigma_1^{\mathrm{T}}\Sigma_2^{-1}\Sigma_1 v + 2v^{\mathrm{T}}\Sigma_1^{\mathrm{T}}\Sigma_2^{-1}\Big\{\frac{1}{\sqrt{n}}\sum_{t=1}^{n} Z_t sgn(Y_t(\theta_0^*))\Big\} + o_p(1) \tag{18}$$

locally uniformly in v. By minimizing the above equation with respect to v, we obtain

$$\begin{aligned} & l(\theta_0^*) - \arg\min_{\theta^*} l(\theta^*) \\ & = \Big\{\frac{1}{\sqrt{n}}\sum\nolimits_{t=1}^{n} Z_t sgn(Y_t(\theta_0^*))\Big\}^{\mathrm{T}}\Sigma_2^{-1}\Big\{\frac{1}{\sqrt{n}}\sum\nolimits_{t=1}^{n} Z_t sgn(Y_t(\theta_0^*))\Big\} + o_p(1). \end{aligned} \tag{19}$$

Set $\bar{Z}_t = (-\frac{1}{1+X_{t-1}^2}, \frac{X_{t-1}^2}{1+X_{t-1}^2})^{\mathrm{T}}$ for $t = 1, \ldots, n$. Using the same arguments as in proving (12) and (18), we can show that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} \bar{Z}_t\{sgn(Y_t((\alpha_0^*, \beta^*, \gamma^*)^{\mathrm{T}})) - sgn(Y_t(\theta_0^*))\} = \Sigma_3(v_2, v_3)^{\mathrm{T}} + o_p(1)$$

and

$$\begin{aligned} & l((\alpha_0^*, \beta^*, \gamma^*)^{\mathrm{T}}) - l(\theta_0^*) \\ = \quad & (v_2, v_3)\, \Sigma_3^{\mathrm{T}}\, \Sigma_4^{-1}\, \Sigma_3\, (v_2, v_3)^{\mathrm{T}} + 2(v_2, v_3)\, \Sigma_3^{\mathrm{T}}\, \Sigma_4^{-1} \left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \bar{Z}_t sgn(Y(\theta_0^*)) \right\} + o_p(1) \end{aligned}$$

locally uniformly in v_2 and v_3, where

$$\Sigma_3 = 2f(\sqrt{m_0}) \begin{pmatrix} E(\frac{1}{S_1\Delta_1}) & -E(\frac{X_1^2}{S_1\Delta_1}) \\ -E(\frac{X_1^2}{S_1\Delta_1}) & E(\frac{X_1^4}{S_1\Delta_1}) \end{pmatrix}$$

and

$$\Sigma_4 = \begin{pmatrix} E(1/S_1^2) & -E(X_1^2/S_1^2) \\ -E(X_1^2/S_1^2) & E(X_1^4/S_1^2) \end{pmatrix}.$$

By minimizing the above equation with respect to v_2 and v_3, we obtain

$$\begin{aligned} & l(\theta_0^*) - \arg\min_{\theta^*=(\alpha_0^*, \beta^*, \gamma^*)^{\mathrm{T}}} l(\theta^*) \\ & = \left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \bar{Z}_t sgn(Y_t(\theta_0^*)) \right\}^{\mathrm{T}} \Sigma_4^{-1} \left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \bar{Z}_t sgn(Y_t(\theta_0^*)) \right\} + o_p(1). \end{aligned} \tag{20}$$

Hence, the theorem follows from (14), (19) and (20).

Proof of Theorem 2. Cases (i) and (ii) follow from Theorem 1 immediately, and case (iii) can be shown in a way similar to Theorem 1.

Acknowledgment

We thank two reviewers for their helpful comments. Peng's research was supported by NSF grant DMS-04-03443 and a Humboldt Research Fellowship.

References

[1] Borkovec, M. and Klüppelberg, C. (2001). The tail of the stationary distribution of an autoregressive process with ARCH(1) errors. *Ann. Appl. Prob.* **11** 1220–41.

[2] Chan, N.G. & Peng, L. (2005). Weighted least absolute deviations estimation for an AR(1) process with ARCH(1) errors. *Biometrika* **92** 477–487.

[3] Chan, N.G., Peng, L. & Qi, Y. (2005). Quantile inference for near-integrated autoregressive time series with infinite variance. *Statistica Sinica*, to appear.

[4] Chuang, C. & Chan, N.H. (2002). Empirical likelihood for autoregressive models, with applications to unstable time series. *Statist. Sinica* **12** 387–407.

[5] Kluppelberg, C., Maller, R.A., Van de Vyver, M. & wee, D. (2002). Testing for reduction to random walk in autoregressive conditional heteroskedasticity models. *Econometr. J.* **5** 387–416.

[6] Hall, P. & Heyde, C.C. (1980). *Martingale Limit Theory and Its Applications.* New York: Academic Press.

[7] Ling, S. (2004). Estimation and testing stationarity for double autoregressive models. *J. R. Statist. Soc.* B **66** 63–78.

[8] Owen, A. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18(1)**, 90 - 120.

[9] Owen, A. (2001). *Empirical Likelihood Method.* Chapman & Hall, London.