

QUANTILE INFERENCE FOR NEAR-INTEGRATED AUTOREGRESSIVE TIME SERIES WITH INFINITE VARIANCE

Ngai Hang Chan, Liang Peng and Yongcheng Qi

*The Chinese University of Hong Kong, Georgia Institute of Technology
and University of Minnesota Duluth*

Abstract: The limiting distribution of the quantile estimate for the autoregressive coefficient of a near-integrated first order autoregressive model with infinite variance errors is derived. Since the limiting distribution depends on the unknown density function of the errors, an empirical likelihood ratio statistic is proposed from which confidence intervals can be constructed for the near unit root model without knowing the density function. Numerical simulations are conducted to compare the performance of the empirical likelihood method and the least squares procedure. It is found that the empirical likelihood method outperforms the least squares procedure in general.

Key words and phrases: Empirical likelihood method, infinite variance, near unit root, quantile estimate.

1. Introduction

Consider the first-order autoregressive (AR(1)) model

$$Y_{t,n} = \alpha_n Y_{t-1,n} + \epsilon_t, \quad t = 1, \dots, n, \quad (1.1)$$

where $\alpha_n = \alpha_n(\varphi) = 1 - \varphi/n$, φ is a real number, $Y_{0,n} = 0$ for all n , and $\{\epsilon_t\}$ is a sequence of i.i.d. random variables. This is called a near-integrated AR(1) model. If $\varphi = 0$, (1.1) reduces to the traditional unit root model.

There is an extensive literature on unit root estimation and testing for the case $\varphi = 0$. For example, Dickey and Fuller (1979) and Chan and Tran (1989) studied the least squares estimation of a nonstationary AR(1) model when $E(\epsilon_t^2) < \infty$ and $E(\epsilon_t^2) = \infty$, respectively; Callegari, Cappuccio and Lubian (2003) extended the results in Chan and Tran (1989) to the case where a drift exists; Knight (1989, 1991) studied M estimation and least absolute deviations estimate for the case $E(\epsilon_t^2) = \infty$; Hecce (1996) studied least absolute deviations estimate for the case $E(\epsilon_t^2) < \infty$; Shin and So (1999) proposed Cauchy M -estimate when ϵ_t may have infinite variance; a sequential estimate was proposed by Lai and Siegmund (1983) for the case $E(\epsilon_t^2) < \infty$; rank tests of unit root were investigated

by Hasan and Koenker (1997) for the case $E(\epsilon_t^2) < \infty$, and by Hasan (2001) for the case $E(\epsilon_t^2) = \infty$; Ling and Li (1998) studied the maximum likelihood procedure for nonstationary time series with GARCH errors; Ahn, Fotopoulos and He (2001) compared different testing procedures for the unit root with infinite variance errors; tests based on the bootstrap were proposed by Horváth and Kokoszka (2003).

Although the study of the unit root AR(1) model has been actively pursued by statisticians and econometricians alike, a related question that needs to be addressed is what happens to the limiting distribution of the test statistics when the autoregressive parameter α_n is close to one? Equivalently, for finite sample theory or analysis of tests under local alternatives such as $\alpha_n = 1 - \varphi/n$, what kind of approximation should be used for the distribution of the test statistics? To answer this question, Chan and Wei (1987) proposed the triangular array framework (1.1), which they called the nearly nonstationary AR(1) model, and established the limiting distributions of the least squares statistics for α_n under the assumption that the conditional variance of ϵ_t is finite. Similar results can also be found in Phillips (1987). Further studies of least squares estimate of a near-integrated AR(1) model and related issues are given in Chan (1988, 1990) and the references therein.

In addition to the popular least squares procedure, another important technique examined by econometricians is the so-called quantile regression, which has been receiving considerable attention in the literature since the seminal work of Koenker and Basset (1978). Recent advances can be found, for example, in Koenker and Xiao (2002) and Ling and McAleer (2004). Most of the research, however, has focused on the case where $E(\epsilon_t^2) < \infty$. In this paper, we investigate quantile regression estimation for (1.1) when $E(\epsilon_t^2) = \infty$. This model is different from the models studied in Knight (1989, 1991), where the median of ϵ_t is assumed to be zero and $\varphi = 0$. Since quantile estimates are more robust, as will be seen in Section 4, they offer better inferential procedures than the least squares method when the underlying series is heavy-tailed.

The paper is organized as follows. In Section 2, the asymptotic limit of the quantile regression estimate is derived. An empirical likelihood method for constructing a confidence interval for the parameter α_n is proposed in Section 3. A simulation study is presented in Section 4. Proofs are given in Section 5.

2. Quantile Regression Estimation

Herein, when no confusion arises, write $Y_{t,n}$ as Y_t in (1.1). Let $\alpha(\tau) = \alpha_n$ and $\beta(\tau)$ denote the τ -th quantile of ϵ_t . Let $\rho_\tau(u) = u(\tau - I(u < 0))$, $\theta(\tau) = (\beta(\tau), \alpha(\tau))^T$, and $X_t = (1, Y_{t-1})^T$. If we let $Q_t(\tau|t-1)$ denote the τ -th conditional

quantile of Y_t conditional on Y_{t-1} , then $Q_t(\tau|t-1) = X_t^\top \theta(\tau)$. Hence, it follows from Koenker and Bassett (1978) that the quantile regression estimate is

$$\hat{\theta}(\tau) = \operatorname{argmin}_{\theta(\tau)} \sum_{t=1}^n \rho_\tau(Y_t - X_t^\top \theta(\tau)). \quad (2.1)$$

Unless otherwise stated, all limits are taken as n tends to infinity. We assume the following conditions throughout.

Condition 1. There exists $a_n > 0$ such that $a_n^{-1} \sum_{t=1}^n \epsilon_t$ converges in distribution to a stable law with index $\eta \in (0, 2)$;

Condition 2. The distribution function F of ϵ_t has a continuous Lebesgue density f , which is positive on $\{u : 0 < F(u) < 1\}$.

The main result is the following.

Theorem 1. *Assume (1.1) holds with $\alpha_n = \alpha_n(\varphi)$, and Conditions 1 and 2 hold. Then*

$$D_n(\hat{\theta}(\tau) - \theta_0(\tau)) \xrightarrow{d} \frac{1}{f(\beta_0(\tau))} \Sigma^{-1} \left(W(\tau, 1), \int_0^1 S_1(s) dW(\tau, s) \right)^\top, \quad (2.2)$$

where $D_n = \operatorname{diag}(\sqrt{n}, a_n \sqrt{n})$, $\theta_0(\tau) = (\beta_0(\tau), \alpha_{n,0})^\top$ denotes the true value of $\theta(\tau)$, $\Sigma = \int_0^1 (1, S_1(s))^\top (1, S_1(s)) ds$, and $W(\tau, s)$ and $S_1(s)$ are two independent stochastic processes whose specific forms are given in Lemma 2 in Section 5. In particular,

$$a_n \sqrt{n}(\hat{\alpha}(\tau) - \alpha_{n,0}) \xrightarrow{d} \frac{1}{f(\beta_0(\tau))} \frac{\int_0^1 S_1(s) dW(\tau, s) - W(\tau, 1) \int_0^1 S_1(s) ds}{\int_0^1 S_1^2(s) ds - (\int_0^1 S_1(s) ds)^2}, \quad (2.3)$$

$$\left\{ \sum_{t=1}^n Y_{t-1}^2 - \left(\sum_{t=1}^n Y_{t-1} \right)^2 \right\}^{\frac{1}{2}} (\hat{\alpha}(\tau) - \alpha_{n,0}) \xrightarrow{d} N\left(0, \frac{\tau(1-\tau)}{f^2(\beta_0(\tau))}\right). \quad (2.4)$$

3. Empirical Likelihood

Although the limiting distribution given in (2.4) is normal, its form involves the unknown density f of the error distribution. As a result, a confidence interval for α_n based on the normal approximation requires an estimate of the unknown density function f . To circumvent this difficulty, we propose to use the empirical likelihood method. Chuang and Chan (2002) recently apply the method to unit root AR models with finite variance errors. In this paper, a profile empirical likelihood method of Qin and Lawless (1994) is adopted to study the near unit

root AR(1) model with infinite variance errors. We refer to Owen (2001) for a comprehensive description of the empirical likelihood method.

Let $p = (p_1, \dots, p_n)$ be a probability vector, i.e., $\sum_{i=1}^n p_i = 1$ and $p_i \geq 0$ for $i = 1, \dots, n$. The empirical likelihood is defined as

$$L(\theta(\tau)) = \sup \left\{ \prod_{t=1}^n p_t : \sum_{t=1}^n p_t = 1, \sum_{t=1}^n p_t X_t \psi_\tau(Y_t - X_t^\top \theta(\tau)) = 0 \right\},$$

where $\psi_\tau(u) = \tau - I(u < 0)$.

By the method of Lagrange multipliers, we have

$$p_t = \frac{1}{n} \{1 + \lambda^\top X_t \psi_\tau(Y_t - X_t^\top \theta(\tau))\}^{-1}, t = 1, \dots, n, \quad (3.1)$$

where $\lambda = (\lambda_1, \lambda_2)^\top$ satisfies

$$\frac{1}{n} \sum_{t=1}^n \frac{X_t \psi_\tau(Y_t - X_t^\top \theta(\tau))}{1 + \lambda^\top X_t \psi_\tau(Y_t - X_t^\top \theta(\tau))} = 0. \quad (3.2)$$

The empirical likelihood ratio is

$$l(\theta(\tau)) = 2 \sum_{t=1}^n \log \{1 + \lambda^\top X_t \psi_\tau(Y_t - X_t^\top \theta(\tau))\}$$

and the profile empirical likelihood ratio is $l_p(\alpha(\tau)) = \min_{\theta(\tau)} l(\theta(\tau))$.

Theorem 2. *Under the conditions of Theorem 1, $l_p(\alpha_{n,0}) - \min_{\theta(\tau)} l(\theta(\tau)) \xrightarrow{d} \chi^2(1)$.*

Based on this result, a confidence interval for $\alpha_{n,0}$ with significance level γ is

$$I_\gamma = \{\alpha : l_p(\alpha) - \min_{\theta(\tau)} l(\theta(\tau)) \leq u_\gamma\}, \quad (3.3)$$

where u_γ is the 100 γ %-level quantile of $\chi^2(1)$. This gives a confidence interval for $\alpha_{n,0}$ with an asymptotic coverage probability γ , as shown in the next corollary.

Corollary 1. *Assume the conditions of Theorem 1 hold. Then, as $n \rightarrow \infty$, $P(\alpha_{n,0} \in I_\gamma) \rightarrow \gamma$.*

4. Simulations

Finite sample performances in terms of standard deviations and coverage probabilities of the quantile estimate and the least squares estimate are compared in this section. The least squares estimate $\tilde{\alpha}_n = \sum_{i=1}^n Y_i Y_{i-1} / \sum_{i=1}^n Y_i^2$ was studied by Chan (1990).

First, we compared the quantile estimate at $\tau = 0.5$ with the least squares estimate. We drew 1,000 samples of size $n = 50$ and 500 from (1.1) with ϵ_t a standard Cauchy. We considered the cases $\varphi = 0, 1, 5, 10, 20, 50, 100$. The sample averages of the estimates and the corresponding standard deviations of $\hat{\alpha}_n$ and $\tilde{\alpha}_n$ are given in Table 1. It shows that the quantile estimate performs better than the least squares estimate both in terms of precision and standard errors. This improvement is more prominent for moderate sample size, such as $n = 50$.

Table 1. Estimates of $\hat{\alpha}_n$ and $\tilde{\alpha}_n$. Numbers in bracket are the corresponding standard deviations.

$n = 50$							
α_n	1	0.98	0.9	0.8	0.6	0	-1
$\hat{\alpha}_n$	0.985 (0.040)	0.966 (0.036)	0.889 (0.042)	0.790 (0.044)	0.591 (0.046)	-0.003 (0.054)	-0.997 (0.014)
$\tilde{\alpha}_n$	0.929 (0.098)	0.918 (0.100)	0.852 (0.102)	0.760 (0.104)	0.571 (0.102)	-0.006 (0.093)	-0.934 (0.092)
$n = 500$							
α_n	1	0.998	0.99	0.98	0.96	0.9	0.8
$\hat{\alpha}_n$	0.9998 (0.001)	0.9978 (0.001)	0.9898 (0.001)	0.9798 (0.001)	0.9599 (0.001)	0.8999 (0.002)	0.7998 (0.003)
$\tilde{\alpha}_n$	0.9923 (0.014)	0.9911 (0.014)	0.9842 (0.016)	0.9744 (0.018)	0.9546 (0.022)	0.8947 (0.028)	0.7950 (0.032)

Second, we compared the confidence intervals in terms of coverage accuracy based on the empirical likelihood method in Section 3 with those based on the limit of least squares estimate. Let $g(\phi)$ denote the limit of $n(\tilde{\alpha}_n - \alpha_n)$ and z_γ denote the $100\gamma\%$ -level quantile of $g(\phi)$. Values of z_γ can be found in Chan (1990). A γ level confidence interval based on the above approximation is

$$I_\gamma^* = \left(\tilde{\alpha}_n - \frac{z_{1-\frac{\gamma}{2}}}{n}, \tilde{\alpha}_n - \frac{z_{\frac{\gamma}{2}}}{n} \right). \tag{4.1}$$

We drew 1,000 samples of size $n = 500$ from (1.1) with ϵ_t a standard Cauchy. Again, we considered $\varphi = 0, 1, 5, 10, 20, 50, 100$ with the true coverage probabilities $\gamma = 0.90$ and 0.95 . Table 2 shows the coverage probabilities for I_γ using the empirical likelihood given in (3.3) with I_γ^* using the least squares given in (4.1). Once again, the empirical likelihood method clearly outperforms the least squares approximation method.

In conclusion, these simulations furnish strong evidence that the proposed quantile estimate procedure provides a reliable alternative for the least squares procedure when conducting statistical inference for near-integrated AR(1) models. Furthermore, the improvement exhibited by the empirical procedure is more striking for series with moderate lengths.

Table 2. Coverage probabilities with $n = 500$.

α_n	1	0.998	0.99	0.98	0.96	0.9	0.8
$I_{0.90}$	0.922	0.913	0.929	0.927	0.929	0.930	0.919
$I_{0.90}^*$	0.865	0.877	0.887	0.885	0.861	0.855	0.808
$I_{0.95}$	0.954	0.952	0.967	0.963	0.963	0.965	0.957
$I_{0.95}^*$	0.932	0.930	0.940	0.940	0.927	0.924	0.937

5. Proofs

For convenience, introduce the following notation:

$$S_n(s) = a_n^{-1} \sum_{t=1}^{[ns]} \epsilon_t, \quad T_n(s) = a_n^{-1} Y_{[ns]}, \quad W_n(\tau, s) = n^{-\frac{1}{2}} \sum_{t=1}^{[ns]} \psi_\tau(\epsilon_t - \beta_0(\tau))$$

$$V_n = a_n^{-1} n^{-\frac{1}{2}} \sum_{t=1}^n Y_{t-1} \psi_\tau(\epsilon_t - \beta_0(\tau)).$$

Lemma 1. *Under the conditions of Theorem 1,*

$$(S_n(s_1), W_n(\tau, s_2)) \xrightarrow{d} (S(s_1), W(\tau, s_2)) \quad (5.1)$$

in $D\{[0, 1]^2\}$, where $S(\cdot)$ and $W(\cdot, \cdot)$ are independent, $S(s)$ is a stable process with index η , and $W(\tau, s)$ is a rescaled Brownian bridge for fixed s , and a Brownian motion with variance $\tau(1 - \tau)$ for fixed τ . Thus, for each fixed pair (τ, s) , $W(\tau, s) \sim N(0, \tau(1 - \tau)s)$.

Proof. Equation (5.1) follows from the arguments in Resnick and Greenwood (1979).

Lemma 2. *Under the conditions of Theorem 1, we have*

$$T_n(\cdot) = a_n^{-1} Y_{[n\cdot]} \xrightarrow{d} S_1(\cdot) \quad (5.2)$$

in $D[0, 1]$ and

$$V_n \xrightarrow{d} \int_0^1 S_1(s) dW(\tau, s), \quad (5.3)$$

where $S_1(t) = S(t) - \varphi \int_0^t e^{\varphi(s-t)} S(s) ds$, and $W(\cdot, \cdot)$ and $S(\cdot)$ are defined in Lemma 1.

Remark 1. Note that $S_1(t)$ defined above is the analog of the Ornstein-Uhlenbeck process in the finite variance case.

Proof of Lemma 2.

$$\begin{aligned}
 T_n(s) &= a_n^{-1} \sum_{j=1}^{[ns]} \epsilon_j + \sum_{i=1}^{[ns]} (\alpha_n^{[ns]-i+1} - \alpha_n^{[ns]-i}) a_n^{-1} \sum_{j=1}^i \epsilon_j \\
 &= a_n^{-1} \sum_{j=1}^{[ns]} \epsilon_j - \frac{\varphi}{n} \sum_{i=1}^{[ns]} \alpha_n^{[ns]-i} a_n^{-1} \sum_{j=1}^i \epsilon_j \\
 &= S_n(s) - \varphi \int_0^{n^{-1}([ns]+1)} \left(1 - \frac{\varphi}{n}\right)^{[ns]-[nu]} S_n(u) du.
 \end{aligned}$$

Therefore, (5.2) follows directly from Lemma 1. The rest of the proof follows along the lines of the proof of Lemma 1 in Knight (1989) by noting that

$$a_n^{-1} n^{-\frac{1}{2}} \sum_{t=1}^n Y_{t-1} \psi_\tau(\epsilon_t - \beta_0(\tau)) = \int_0^1 T_n(s-) dW_n(s).$$

Proof of Theorem 1. Put $\nu = (\nu_1, \nu_2)^\top = D_n(\beta(\tau) - \beta_0(\tau), \alpha(\tau) - \alpha_n)^\top$ and

$$Z_n(\nu) = \sum_{t=1}^n \{\rho_\tau(\epsilon_t - \beta_0(\tau) - \nu^\top D_n^{-1} X_t) - \rho_\tau(\epsilon_t - \beta_0(\tau))\}.$$

Since $\rho_\tau(u - v) - \rho_\tau(u) = -v\psi_\tau(u) + (u - v)\{I(0 > u > v) - I(0 < u < v)\}$, we have

$$\begin{aligned}
 Z_n(\nu) &= - \sum_{t=1}^n \{\nu^\top D_n^{-1} X_t \psi_\tau(\epsilon_t - \beta_0(\tau))\} \\
 &\quad + \sum_{t=1}^n \{\epsilon_t - \beta_0(\tau) - \nu^\top D_n^{-1} X_t\} I(0 > \epsilon_t - \beta_0(\tau) > \nu^\top D_n^{-1} X_t) \\
 &\quad - \sum_{t=1}^n \{\epsilon_t - \beta_0(\tau) - \nu^\top D_n^{-1} X_t\} I(0 < \epsilon_t - \beta_0(\tau) < \nu^\top D_n^{-1} X_t) \\
 &= I_1 + I_2 + I_3, \quad \text{say.} \tag{5.4}
 \end{aligned}$$

It follows from Lemma 2 that

$$I_1 = -\nu^\top (W_n(\tau, 1), V_n)^\top \xrightarrow{d} -\nu^\top (W(\tau, 1), \int_0^1 S_1(s) dW(\tau, s))^\top \tag{5.5}$$

and $\max_{1 \leq t \leq n} |\sqrt{n} \nu^\top D_n^{-1} X_t| \leq |\nu_1| + |\nu_2| \max_{1 \leq t \leq n} |a_n^{-1} Y_{t-1}| \xrightarrow{d} |\nu_1| + |\nu_2| \sup_{0 \leq s \leq 1} |S_1(s)|$. Thus,

$$\max_{1 \leq t \leq n} |\sqrt{n} \nu^\top D_n^{-1} X_t| = O_p(1). \tag{5.6}$$

Adopt the following definitions:

$$\begin{aligned} Z_{tn}(\nu) &= (\nu^T D_n^{-1} X_t - \epsilon_t + \beta_0(\tau)) I(\nu^T D_n^{-1} X_t > \epsilon_t - \beta_0(\tau) > 0) \\ &\quad I(0 < \sqrt{n} \nu^T D_n^{-1} X_t \leq n^{\frac{1}{4}}); \\ \mathcal{F}_t &= \sigma(\epsilon_s; s \leq t); \quad \mu_{tn}(\nu) = E(Z_{tn}(\nu) | \mathcal{F}_{t-1}); \\ A_{tn}(\nu) &= \nu^T D_n^{-1} X_t I(0 < \sqrt{n} \nu^T D_n^{-1} X_t \leq n^{\frac{1}{4}}). \end{aligned}$$

From (5.6) and Lemma 2 it is easy to see that $I_3 = \sum_{t=1}^n Z_{tn}(\nu) + o_p(1)$ and

$$\begin{aligned} \sum_{t=1}^n A_{tn}^2(\nu) &= \sum_{t=1}^n (\nu^T D_n^{-1} X_t)^2 I(\nu^T D_n^{-1} X_t > 0) + o_p(1) \\ &\xrightarrow{d} \int_0^1 \nu^T (1, S(s))^T (1, S(s)) \nu I(0 < \nu^T (1, S(s))^T) ds. \end{aligned}$$

Hence

$$\begin{aligned} \sum_{t=1}^n \mu_{tn}(\nu) &= \sum_{t=1}^n \int_{\beta_0(\tau)}^{\beta_0(\tau) + A_{tn}(\nu)} \{A_{tn}(\nu) + \beta_0(\tau) - x\} f(x) dx \\ &= \sum_{t=1}^n \int_{\beta_0(\tau)}^{\beta_0(\tau) + A_{tn}(\nu)} \left\{ \int_x^{A_{tn}(\nu) + \beta_0(\tau)} ds \right\} f(x) dx \\ &= \sum_{t=1}^n \int_{\beta_0(\tau)}^{\beta_0(\tau) + A_{tn}(\nu)} \left\{ \int_{\beta_0(\tau)}^s f(x) dx \right\} ds \\ &= \sum_{t=1}^n \int_{\beta_0(\tau)}^{\beta_0(\tau) + A_{tn}(\nu)} \{F(s) - F(\beta_0(\tau))\} ds \\ &= \sum_{t=1}^n \int_{\beta_0(\tau)}^{\beta_0(\tau) + A_{tn}(\nu)} (s - \beta_0(\tau)) f(\beta_0(\tau)) (1 + o_p(1)) ds \\ &= \frac{1}{2} \sum_{t=1}^n f(\beta_0(\tau)) A_{tn}^2(\nu) (1 + o_p(1)) \\ &= \frac{1}{2} f(\beta_0(\tau)) \sum_{t=1}^n (\nu^T D_n^{-1} X_t)^2 I(\nu^T D_n^{-1} X_t > 0) + o_p(1). \end{aligned}$$

Since $\max_{1 \leq t \leq n} A_{tn}(\nu) = o_p(1)$,

$$\sum_{t=1}^n E(Z_{tn}^2(\nu) | \mathcal{F}_{t-1}) \leq \left(\max_{1 \leq t \leq n} A_{tn}(\nu) \right) \sum_{t=1}^n \mu_{tn}(\nu) \xrightarrow{p} 0.$$

Thus we have

$$I_3 = \frac{1}{2}f(\beta_0(\tau)) \sum_{t=1}^n (\nu^\top D_n^{-1} X_t)^2 I(\nu^\top D_n^{-1} X_t > 0) + o_p(1),$$

$$I_2 = \frac{1}{2}f(\beta_0(\tau)) \sum_{t=1}^n (\nu^\top D_n^{-1} X_t)^2 I(\nu^\top D_n^{-1} X_t < 0) + o_p(1).$$

Therefore

$$Z_n(\nu) = -\nu^\top (W_n(\tau, 1), V_n)^\top + \frac{1}{2}f(\beta_0(\tau)) \sum_{t=1}^n (\nu^\top D_n^{-1} X_t)^2 + o_p(1)$$

$$\xrightarrow{d} -\nu^\top \left(W(\tau, 1), \int_0^1 S_1(s) dW(\tau, s) \right)^\top + \frac{1}{2}f(\beta_0(\tau)) \nu^\top \Sigma \nu,$$

where $\Sigma = \int_0^1 (1, S_1(s))^\top (1, S_1(s)) ds$.

Since $Z_n(\nu)$ has convex sample paths, the above convergence implies uniform convergence on compact sets. Thus, (2.2) follows from Lemma 2.2 of Davis, Knight and Liu (1992). Equations (2.3) and (2.4) can be checked easily.

Proof of Theorem 2. Let $\bar{\theta}_0(\tau, \nu) = (\beta_0(\tau) + n^{-1/2}\nu, \alpha_{n,0})^\top$, $Z_t(\bar{\theta}_0) = X_t \Psi_\tau(Y_t - X_t^\top \bar{\theta}_0(\tau, \nu))$, and $\lambda = (\lambda_1, \lambda_2)^\top$ be the solution of $(1/n) \sum_{t=1}^n (Z_t(\bar{\theta}_0))/(1 + \lambda^\top Z_t(\bar{\theta}_0)) = 0$. Thus $\bar{\lambda} = (\lambda_1, a_n \lambda_2)^\top$ is the solution of

$$g(\bar{\lambda}) = \frac{1}{n} \sum_{t=1}^n \frac{\bar{Z}_t(\bar{\theta}_0)}{1 + \bar{\lambda}^\top \bar{Z}_t(\bar{\theta}_0)} = 0,$$

where $\bar{Z}_t(\bar{\theta}_0) = \bar{X}_t \Psi_\tau(Y_t - \bar{X}_t^\top \bar{\theta}_0(\tau, \nu))$ and $\bar{X}_t = (1, a_n^{-1} Y_{t-1})^\top$.

Our first step is to prove that

$$\|\bar{\lambda}\| = O_p(n^{-\frac{1}{2}}) \text{ locally uniformly in } \nu. \quad (5.7)$$

Write $\bar{\lambda} = \rho \lambda_0$, where $\rho \geq 0$ and $\|\lambda_0\| = 1$. Since the probabilities p_t appearing in the definition of $l_p(\alpha_{n,0})$ are given by $p_t = n^{-1}(1 + \lambda^\top Z_t(\bar{\theta}_0))^{-1} = n^{-1}(1 + \bar{\lambda}^\top \bar{Z}_t(\bar{\theta}_0))^{-1}$, we have $1 + \bar{\lambda}^\top \bar{Z}_t(\bar{\theta}_0) \geq 0$, i.e.,

$$|1 + \bar{\lambda}^\top \bar{Z}_t(\bar{\theta}_0)|^{-1} = |1 + \rho \lambda_0^\top \bar{Z}_t(\bar{\theta}_0)|^{-1} \geq \left(1 + \rho \max_{1 \leq t \leq n} \|\bar{Z}_t(\bar{\theta}_0)\|\right)^{-1}.$$

Hence

$$\begin{aligned}
0 &= \|g(\bar{\lambda})\| = \|g(\rho\lambda_0)\| \\
&\geq |\lambda_0^\top g(\rho\lambda_0)| \\
&= \frac{1}{n} \left| \lambda_0^\top \left\{ \sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) - \rho \sum_{t=1}^n \frac{\bar{Z}_t(\bar{\theta}_0) \lambda_0^\top \bar{Z}_t(\bar{\theta}_0)}{1 + \rho \lambda_0^\top \bar{Z}_t(\bar{\theta}_0)} \right\} \right| \\
&\geq \frac{\rho}{n} \lambda_0^\top \sum_{t=1}^n \frac{\bar{Z}_t(\bar{\theta}_0) \bar{Z}_t(\bar{\theta}_0)^\top}{1 + \rho \lambda_0^\top \bar{Z}_t(\bar{\theta}_0)} \lambda_0 - \frac{1}{n} \left| \lambda_0^\top \sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) \right| \\
&\geq \frac{\rho}{n} \left\{ 1 + \rho \max_{1 \leq t \leq n} \|\bar{Z}_t(\bar{\theta}_0)\| \right\}^{-1} \lambda_0^\top \sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) \bar{Z}_t(\bar{\theta}_0)^\top \lambda_0 - \frac{1}{n} \left| \lambda_0^\top \sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) \right|.
\end{aligned}$$

That is,

$$\begin{aligned}
&\rho \left\{ \frac{1}{n} \lambda_0^\top \sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) \bar{Z}_t(\bar{\theta}_0)^\top \lambda_0 - \left(\max_{1 \leq t \leq n} \|\bar{Z}_t(\bar{\theta}_0)\| \right) \frac{1}{n} \left| \sum_{t=1}^n \lambda_0^\top \bar{Z}_t(\bar{\theta}_0) \right| \right\} \\
&\leq \frac{1}{n} \left| \sum_{t=1}^n \lambda_0^\top \bar{Z}_t(\bar{\theta}_0) \right|. \tag{5.8}
\end{aligned}$$

Note that

$$\begin{aligned}
\sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) &= \sum_{t=1}^n \bar{X}_t \Psi_{F(\beta_0(\tau) + n^{-\frac{1}{2}}\nu)}(\epsilon_t - \beta_0(\tau) - n^{-\frac{1}{2}}\nu) \\
&\quad + \sum_{t=1}^n \bar{X}_t \{\tau - F(\beta_0(\tau) + n^{-\frac{1}{2}}\nu)\}. \tag{5.9}
\end{aligned}$$

By Lemma 2,

$$\left\| \frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) \right\| = O_p(n^{-\frac{1}{2}}) \text{ locally uniformly in } \nu. \tag{5.10}$$

Write

$$\begin{aligned}
&\frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) \bar{Z}_t(\bar{\theta}_0)^\top \\
&= \frac{1}{n} \sum_{t=1}^n \bar{X}_t \{ \Psi_\tau^2(\epsilon_t - \beta_0(\tau) - n^{-\frac{1}{2}}\nu) - \Psi_\tau^2(\epsilon_t - \beta_0(\tau)) \} \bar{X}_t \\
&\quad + \tau(1 - \tau) \frac{1}{n} \sum_{t=1}^n \bar{X}_t (1, a_n^{-1} Y_{t-1}) + (2\tau - 1) \frac{1}{n} \sum_{t=1}^n \bar{X}_t \Psi_\tau(\epsilon_t - \beta_0(\tau)) \bar{X}_t^\top \\
&= I_1 + I_2 + I_3. \tag{5.11}
\end{aligned}$$

Similar to (5.10) we have, by Lemma 2,

$$I_1 = o_p(1), \quad I_2 \xrightarrow{d} \bar{\Sigma}, \quad I_3 = o_p(1) \quad (5.12)$$

locally uniformly in ν , where $\bar{\Sigma} = \tau(1 - \tau)\Sigma$. Hence

$$\frac{1}{n} \lambda_0^\top \sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) \bar{Z}_t(\bar{\theta}_0)^\top \lambda_0 \xrightarrow{d} \lambda_0^\top \bar{\Sigma} \lambda_0 \quad (5.13)$$

locally uniformly in ν . It is straightforward to show that

$$\max_{1 \leq t \leq n} \|\bar{Z}_t(\bar{\theta}_0)\| = o_p(\sqrt{n}) \text{ locally uniformly in } \nu \text{ a.s.} \quad (5.14)$$

Thus (5.7) follows from (5.8), (5.10), (5.13) and (5.14). Furthermore,

$$\bar{\lambda} = \left(\frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) \bar{Z}_t(\bar{\theta}_0)^\top \right)^{-1} \frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) + o_p(n^{-\frac{1}{2}})$$

and

$$l(\bar{\theta}_0(\tau, \nu)) = n \left[\frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) \right]^\top \left[\frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) \bar{Z}_t(\bar{\theta}_0)^\top \right]^{-1} \left[\frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) \right] + o_p(1) \quad (5.15)$$

locally uniformly in ν . Similarly

$$l(\theta_0(\tau)) = n \left[\frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\theta_0) \right]^\top \left[\frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\theta_0) \bar{Z}_t(\theta_0)^\top \right]^{-1} \left[\frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\theta_0) \right] + o_p(1). \quad (5.16)$$

By Lemmas 1 and 2,

$$\begin{aligned} & \sqrt{n} \left\{ \frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\bar{\theta}_0) - \frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\theta_0) \right\} \\ &= \sqrt{n} \left\{ \frac{1}{n} \sum_{t=1}^n \bar{X}_t \Psi_{F(\beta_0(\tau) + n^{-\frac{1}{2}}\nu)}(Y_t - X_t^\top \bar{\theta}_0(\tau, \nu)) \right. \\ & \quad \left. + \frac{1}{n} \sum_{t=1}^n \bar{X}_t [F(\beta_0(\tau)) - F(\beta_0(\tau) + n^{-\frac{1}{2}}\nu)] - \frac{1}{n} \sum_{t=1}^n \bar{X}_t \Psi_\tau(Y_t - X_t^\top \theta_0(\tau)) \right\} \\ &= \left(W(F(\beta_0(\tau) + n^{-\frac{1}{2}}\nu), 1), \int_0^1 S_1(s) dW(F(\beta_0(\tau) + n^{-\frac{1}{2}}\nu), s) \right)^\top \\ & \quad - F'(\beta_0(\tau)) \nu \frac{1}{n} \sum_{t=1}^n \bar{X}_t - \left(W(\tau, 1), \int_0^1 S_1(s) dW(\tau, s) \right)^\top + o_p(1) \\ &= -F'(\beta_0(\tau)) \nu \left(1, \int_0^1 S_1(s) ds \right)^\top + o_p(1). \end{aligned} \quad (5.17)$$

Hence, by (5.13), (5.15), (5.16) and (5.17),

$$\begin{aligned} & l(\bar{\theta}_0(\tau, \nu)) - l(\theta_0(\tau)) \\ &= [F'(\beta_0(\tau))]^2 \nu^2 \left(1, \int_0^1 S_1(s) ds\right) \Sigma^{-1} \left(1, \int_0^1 S_1(s) ds\right)^{\text{T}} \\ &\quad - 2F'(\beta_0(\tau)) \nu \left(1, \int_0^1 S_1(s) ds\right) \Sigma^{-1} \left(W(\tau, 1), \int_0^1 S_1(s) dW(\tau, s)\right)^{\text{T}} + o_p(1). \end{aligned}$$

By minimizing the above relation with respect to ν , we obtain

$$\begin{aligned} l(\theta_0(\tau)) - l_p(\alpha_{n,0}) &= \left\{ \left(1, \int_0^1 S_1 ds\right) \bar{\Sigma}^{-1} \left(W(\tau, 1), \int_0^1 S_1(s) dW(\tau, s)\right)^{\text{T}} \right\}^2 / \\ &\quad \left\{ \left(1, \int_0^1 S_1(s) ds\right) \bar{\Sigma}^{-1} \left(1, \int_0^1 S_1(s) ds\right)^{\text{T}} \right\} + o_p(1) \\ &= W(\tau, 1)^2 / [\tau(1 - \tau)] + o_p(1). \end{aligned} \quad (5.18)$$

Define $\tilde{\theta}_0(\tau, \nu_1, \nu_2) = (\beta_0(\tau) + n^{-\frac{1}{2}}\nu_1, \alpha_{n,0} + n^{-\frac{1}{2}}\nu_2)^{\text{T}}$. Similar to (5.15), we have

$$l(\tilde{\theta}_0(\tau, \nu_1, \nu_2)) = n \left[\frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\tilde{\theta}_0) \right]^{\text{T}} \left[\frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\tilde{\theta}_0) \bar{Z}_t(\tilde{\theta}_0)^{\text{T}} \right]^{-1} \left[\frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\tilde{\theta}_0) \right] + o_p(1) \quad (5.19)$$

locally uniformly in ν_1 and ν_2 . It follows from Lemmas 1 and 2 that

$$\begin{aligned} & \sqrt{n} \left\{ \frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\tilde{\theta}_0) - \frac{1}{n} \sum_{t=1}^n \bar{Z}_t(\theta_0) \right\} \\ &= \sqrt{n} \left\{ \frac{1}{n} \sum_{t=1}^n \bar{X}_t \Psi_{F(\beta_0(\tau) + n^{-\frac{1}{2}}\nu_1 + a_n^{-1}Y_{t-1}n^{-\frac{1}{2}}\nu_2)} (Y_t - X_t^{\text{T}} \tilde{\theta}_0(\tau, \nu_1, \nu_2)) \right. \\ &\quad \left. + \frac{1}{n} \sum_{t=1}^n \bar{X}_t [F(\beta_0(\tau)) - F(\beta_0(\tau) + n^{-\frac{1}{2}}\nu_1 + a_n^{-1}Y_{t-1}n^{-\frac{1}{2}}\nu_2)] \right. \\ &\quad \left. - \frac{1}{n} \sum_{t=1}^n \bar{X}_t \Psi_{\tau} (Y_t - X_t^{\text{T}} \theta_0(\tau)) \right\} \\ &= -F'(\beta_0(\tau)) \Sigma(\nu_1, \nu_2)^{\text{T}} + o_p(1). \end{aligned} \quad (5.20)$$

Hence, by (5.13), (5.19), (5.16) and (5.20),

$$\begin{aligned} & l(\tilde{\theta}_0(\tau, \nu_1, \nu_2)) - l(\theta_0(\tau)) \\ &= [F'(\beta_0(\tau))]^2 (\nu_1, \nu_2) \Sigma \bar{\Sigma}^{-1} \Sigma(\nu_1, \nu_2)^{\text{T}} \\ &\quad - 2F'(\beta_0(\tau)) (\nu_1, \nu_2) \Sigma \bar{\Sigma}^{-1} \left(W(\tau, 1), \int_0^1 S_1(s) dW(\tau, s)\right)^{\text{T}} + o_p(1). \end{aligned}$$

By minimizing the above equation with respect to ν_1 and ν_2 , we obtain

$$\begin{aligned}
 & l(\theta_0(\tau)) - \min_{\theta(\tau)} l(\theta(\tau)) \\
 &= \left(W(\tau, 1), \int_0^1 S_1(s) dW(\tau, s) \right) \bar{\Sigma}^{-1} \left(W(\tau, 1), \int_0^1 S_1(s) dW(\tau, s) \right)^T + o_p(1) \\
 &= W(\tau, 1)^2 \int_0^1 S_1^2(s) ds - 2W(\tau, 1) \int_0^1 S_1(s) ds \int_0^1 S_1(s) dW(\tau, s) \\
 &\quad + \left\{ \int_0^1 S_1(s) dW(\tau, s) \right\}^2 / \left\{ \tau(1-\tau) \left[\int_0^1 S_1^2(s) ds - \left(\int_0^1 S_1(s) ds \right)^2 \right] \right\} \\
 &\quad + o_p(1). \tag{5.21}
 \end{aligned}$$

Hence, by (5.18), (5.21) and the fact that $S_1(s_1)$ and $W(\tau, s_2)$ are independent processes,

$$\begin{aligned}
 l_p(\alpha_{n,0}) - \min_{\theta(\tau)} l(\theta(\tau)) &= \left\{ \int_0^1 S_1(s) dW(\tau, s) - W(\tau, 1) \int_0^1 S_1(s) ds \right\}^2 / \\
 &\quad \left\{ \tau(1-\tau) \left[\int_0^1 S_1^2(s) ds - \left(\int_0^1 S_1(s) ds \right)^2 \right] \right\} + o_p(1) \\
 &\xrightarrow{d} \chi^2(1).
 \end{aligned}$$

Acknowledgement

This research was supported in part by HKSAR-RGC Grants CUHK 4043/02P and 2060268, and NSF Grant DMS-04-03443. We are grateful to the Editor, an associate editor and a referee for helpful comments.

References

- Ahn, S. K., Fotopoulos, S. B. and He, L. (2001). Unit root tests with infinite variance errors. *Econom. Rev.* **20**, 461-483.
- Callegari, F., Cappuccio, N. and Lubian, D. (2003). Asymptotic inference in time series regressions with a unit root and infinite variance errors. *J. Statist. Plann. Inference* **116**, 277-303.
- Chan, N. H. (1988). The parameter inference for nearly nonstationary time series. *J. Amer. Statist. Assoc.* **83**, 757-862.
- Chan, N. H. (1990). Inference for nearly-integrated time series with infinite variance. *J. Amer. Statist. Assoc.* **85**, 1069-1074.
- Chan, N. H. and Wei, C. Z. (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *Ann. Statist.* **15**, 1050-1063.
- Chan, N. H. and Tran, L. T. (1989). On the first order autoregressive process with infinite variance. *Econometric Theory* **5**, 354-362.
- Chuang, C. and Chan, N. H. (2002). Empirical likelihood for autoregressive models, with applications to unstable time series. *Statist. Sinica* **12**, 387-407.

- Davis, R. A., Knight, K. and Liu, J. (1992). M-estimation for autoregressions with infinite variance. *Stochastic Process. Appl.* **40**, 145-180.
- Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *J. Amer. Statist. Assoc.* **74**, 427-431.
- Hasan, M. N. (2001). Rank tests of unit root hypothesis with infinite variance errors. *J. Econometrics* **104**, 49-65.
- Hasan, M. N. and Koenker, R. W. (1997). Robust rank tests of the unit root hypothesis. *Econometrica* **65**, 133-161.
- Herce, M. A. (1996). Asymptotic theory of LAD estimation in a unit root process with finite variance errors. *Econom. Theory* **12**, 129-153.
- Horváth, L. and Kokoszka, P. (2003). A bootstrap approximation to a unit root test statistic for heavy tailed observations. *Statist. Probab. Lett.* **62**, 163-173.
- Knight, K. (1989). Limit theory for autoregressive-parameter estimates in an infinite-variance random walk. *Canad. J. Statist.* **17**, 261-278.
- Knight, K. (1991). Limit theory for M -estimates in an integrated infinite variance processes. *Econom. Theory* **7**, 200-212.
- Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica* **46**, 33-49.
- Koenker, R. and Xiao, Z. (2002). Inference on the quantile regression processes. *Econometrica* **70**, 1583-1612.
- Lai, T. L. and Siegmund, D. (1983). Fixed accuracy estimation of an autoregressive parameter. *Ann. Statist.* **11**, 478-485.
- Ling, S. and Li, W. K. (1998). Limiting distributions of maximum likelihood estimators for unstable ARMA time series with GARCH errors. *Ann. Statist.* **26**, 84-125.
- Ling, S. and McAleer, M. (2004). Regression quantiles for unstable autoregressive models. *J. Multivariate Anal.* **89**, 304-328.
- Owen, A. (2001). *Empirical Likelihood Method*. Chapman and Hall, London.
- Phillips, P. C. B. (1987). Towards a unified theory for autoregressions. *Biometrika* **74**, 535-547.
- Qin, J. and Lawless, J. (1994). Empirical likelihood and general estimating equations. *Ann. Statist.* **22**, 300-325.
- Resnick, S. and Greenwood, P. (1979). A bivariate stable characterization and domains of attraction. *J. Multivariate Anal.* **9**, 206-221.
- Shin, D. W. and So, B. S. (1999). New test for unit roots in autoregressive processes with possibly infinite variance errors. *Statist. Probab. Lett.* **44**, 387-397.

Department of Statistics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong.

E-mail: nhchan@sta.cuhk.edu.hk

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, U.S.A.

E-mail: peng@math.gatech.edu

Department of Mathematics and Statistics, University of Minnesota Duluth, Campus Center 140, 1117 University Drive, Duluth, MN 55812, U.S.A.

E-mail: yqi@d.umn.edu

(Received July 2004; accepted March 2005)