96-F-19

On Parameter Estimation of Unit Root Processes with Missing Observations

by

Haruhisa Nishino
Yoshihiro Yajima
Faculty of Economics
University of Tokyo

October 1996

ON PARAMETER ESTIMATION OF UNIT ROOT PROCESSES WITH MISSING OBSERVATIONS *

Haruhisa Nishino

Faculty of Economics, University of Tokyo †

Yoshihiro Yajima

Faculty of Economics,University of Tokyo ‡

October 8, 1996

*The research of the first author was supported by JSPS Research Fellowships for Young Scientists. The research of the second author was partly supported by the Seimei-kai Foundation.

†Hongo 7-3-1, Bunkyo-ku Tokyo JAPAN 113, nishino@e.u-tokyo.ac.jp

‡Hongo 7-3-1, Bunkyo-ku Tokyo JAPAN 113, yajima@e.u-tokyo.ac.jp

Abstract

Limiting distribution of estimators for parameters of an AR model with unit roots has been considered by many authors. On the other hand missing observations in time series data often arise. In this paper we consider two estimators for the parameter of an AR(1) model with a unit root in the presence of missing observations. One is the Yule-Walker estimator which is originally used to estimate parameters of a stationary process. The other is the least-square estimator obtained by using all of the pairs among the data observed consecutively. We derive the limiting distributions of these estimators. The result is a generalization of that given by White (1958), Fuller (1976) and Dicky and Fuller (1979) for the case of complete sampling. As an application, we adopt these estimators to test statistics of unit root tests and their performances are investigated by computational experiments.

1 INTRODUCTION AND MODEL

Many authors have been concerned with a nonstationary process with unit roots in the analysis of economic time series. And limiting distributions of parameters of the process have been considered for a long time. On the other hand missing observations arise from a variety of causes, such as machinery disorder, clerical error or financial markets being closed on holidays or weekends. In this paper we shall consider two estimators of the parameter of an AR(1) model with a unit root in the presence of missing observations and derive the limiting distributions of these estimators. As an application we apply them to a unit root test and their performance is investigated by computational experiments.

We consider the autoregressive model

$$X_k = \rho X_{k-1} + \varepsilon_k, \qquad k = 1, ..., n,$$

where $X_0 = 0$, $\rho = 1$ and $\{\varepsilon_k\}$ is a sequence of independently and identically distributed random variables with mean 0, variance σ^2 and finite fourth-order moment. These assumptions on $\{\varepsilon_k\}$ are used throughout this paper. This model is the prototype of a general AR model with unit roots. And an extension of the results derived here to a general AR model with unit roots will be discussed in a subsequent paper.

While Parzen (1963) introduced the time series model with missing observations as a specific case of an amplitude modulated stationary process. Following him, we express observed data $\{Y_1, Y_2, ..., Y_n\}$ by

$$\begin{cases} X_k &= \rho X_{k-1} + \varepsilon_k, \\ Y_k &= a_k X_k, \end{cases} \tag{1}$$

where $\{a_k, k = 1, 2, ...\}$ represents the state of observation,

$$\begin{cases} a_k = 1 & \text{observed,} \\ a_k = 0 & \text{missing.} \end{cases} \tag{2}$$

Typical examples of $\{a_k\}$ are a stochastic Markov process and a periodically deterministic case. We assume that $\{\varepsilon_k\}$ and $\{a_k\}$ are independent if $\{a_k\}$ is a stochastic process. If $\{X_k\}$ is a stationary process, it starts from the infinite past and $|\rho| < 1$. While we assume throughout this paper that the initial value $X_0 = 0$ and $\rho = 1$ as mentioned before.

Now we propose two estimators of ρ. One is originally proposed by Parzen (1963) for a stationary process and investigated its asymptotic properties under various assumptions on $\{X_k\}$ and $\{a_k\}$ by Dunsmuir and Robinson (1981). Denote this estimator by

$$\hat{\rho} = \frac{\sum_{k=1}^{n-1} Y_k Y_{k+1} / \sum_{k=1}^{n-1} a_k a_{k+1}}{\sum_{k=1}^{n} Y_k^2 / \sum_{k=1}^{n} a_k}. \tag{3}$$

This estimator is the ratio of estimators of autocovariances. If $a_k \equiv 1$, that is, the data is observed completely, $\hat{\rho}$ is identical to the Yule-Walker estimator. We apply $\hat{\rho}$ to estimate ρ.

Recently Takeuchi(1995) suggested another estimator. We defined it by

$$\tilde{\rho} = \frac{\sum_{k=1}^{n-1} Y_k Y_{k+1}}{\sum_{k=1}^{n-1} a_{k+1} Y_k^2} \tag{4}$$

Noting that $\tilde{\rho} = \frac{\sum_{k=1}^{n-1} a_k a_{k+1} X_k X_{k+1}}{\sum_{k=1}^{n-1} a_k a_{k+1} X_k^2}$, we see that $\tilde{\rho}$ is the least-square estimator based all of the pairs of data observed consecutively. Shin and Sarker(1995) also used $\tilde{\rho}$ as the starting value for Newton-Raphson estimation method for a stationary AR(1) model. If the data is observed completely, $\tilde{\rho}$ is equal to an ordinary least-square estimator,

$$\hat{\rho}_{OLS} = \frac{\sum_{k=1}^{n-1} X_k X_{k+1}}{\sum_{k=1}^{n-1} X_k^2}. \tag{5}$$

In section 2, first we prepare some limit theorems to derive the main results. Next we apply them to obtain the limiting distributions of $\hat{\rho}$ and $\tilde{\rho}$. Our main results are following. In the case of complete sampling, $\hat{\rho}_{OLS}$ is an n-consistent estimator and the limiting distribution derived by White (1958) and Fuller (1976) is

$$n(\hat{\rho}_{OLS} - 1) \xrightarrow{D} \frac{\int_0^1 BdB}{\int_0^1 B^2 dt}, \tag{6}$$

where $(B(t); 0 \le t \le 1)$ is a standard Brownian motion and $\xrightarrow{D}$ implies convergence in distribution as the sample size $n \to \infty$.

While $\hat{\rho}$ is not n-consistent but $\sqrt{n}$-consistent when $\{a_k\}$ follows a stochastic Markov process and the limiting distribution of $\sqrt{n}(\hat{\rho} - 1)$ is a functional of Brownian motion being different from (6). In contrast $\hat{\rho}$ is still an n-consistent estimator in a periodically deterministic case. The limiting distribution of $n(\hat{\rho} - 1)$ does not exist as $n \to \infty$. But if we put $n - 1 = n^*M + r$ $(r = 0, 1, ..., M - 1)$ with period M, and let n^* go to infinite, r being fixed, the limiting distribution of $(n^*M + r + 1)(\hat{\rho} - 1)$ exists and depends on r. On the other hand $\tilde{\rho}$ is always an n-consistent estimator. But its mean square error is larger than that of (6).

In section 3 we shall reinforce the theoretical results of these estimators by computational experiments. Also we use these estimators as test statistics of unit root tests and compare their performance with the test statistics proposed by Shin and Sarker (1996). We give additional comments in section 4. Finally Appendix includes the proofs of lemmas and theorems.

2 LIMITING DISTRIBUTIONS

First we prepare some lemmas.

We obtain the following lemma by modifying slightly Phillips and Durlauf (1986) Theorem2.1.

Lemma 1 *Suppose that* $\{\xi_k, k=1,2,...\}$ *is a sequence of* R^d*-valued random variables and* $\{\alpha(i;\xi)\}$ *are strong mixing coefficients of* $\{\xi_k, k=1,2,...\}$. *Let* $W_n(t) = \frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]}\xi_k$. *If*

(i) $E\xi_k = 0$ *for all k,*

(ii) $lim_{n\to\infty} E[\frac{1}{n}(\sum_{k=1}^{n}\xi_k)(\sum_{k=1}^{n}\xi_k)'] = \Sigma$, *a positive definite matrix,*

(iii) $sup_k\, E|\xi_{jk}|^\beta < \infty$ *for some* $\beta\,(2 \le \beta \le \infty)$ *and all* $j = 1, ..., d$,

(iv) $\sum_{m=1}^{\infty}\alpha(i;\xi)^{1-\frac{2}{\beta}} < \infty$, *for some* $\beta\,(2 < \beta \le \infty)$,

then $W_n \xrightarrow{D} W^d$, *where* W^d *is a d-dimensional Brownian motion with a covariance matrix* Σ.

And we extend Chan and Wei (1988)'s result.

Lemma 2 *Let* $\{\varepsilon_k, k=1,2,..\}$, $\{\zeta_k, k=1,2,...\}$ *be two sequences of random variables. Let* $U_n(t) = \frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]}\varepsilon_k$, $V_n(t) = \frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]}\zeta_k$. *Then the following two statements hold.*

(i) [Chan and Wei (1988) Theorem 2.4 (ii)]: *Assume that there are increasing σ-fields* $\mathcal{F}_k$ *such that* $(\varepsilon_k, \zeta_k)'$ *is a sequence of martingale differences with respect to* $\mathcal{F}_k$. *Moreover,*

$$E({\varepsilon_k}^2 + {\zeta_k}^2|\mathcal{F}_{k-1}) \le c \quad a.s.\, for\ some\ constant\ c > 0 \tag{7}$$

and

$$(U_n, V_n) \xrightarrow{D} (W_1, W_2), \tag{8}$$

where $(W_1, W_2)'$is a two-dimensional Brownian motion with respect to an increasing sequence of σ-fields $\mathcal{G}_t$. Then

$$\left(U_n, V_n, \frac{1}{\sqrt{n}} \sum_{k=1}^{n-1} U_n\left(\frac{k}{n}\right) \zeta_{k+1}\right) \xrightarrow{D} \left(W_1, W_2, \int_0^1 W_1 dW_2\right). \tag{9}$$

(ii) Assume that the condition of (i) is satisfied and

$$E({\varepsilon_k}^2|\mathcal{F}_{k-1}) = \sigma^2,\ E({\varepsilon_k}^3|\mathcal{F}_{k-1}) = \nu,\ E({\varepsilon_k}^4|\mathcal{F}_{k-1}) = \tau \quad a.s.. \tag{10}$$

Then

$$\left(U_n, V_n, \frac{1}{\sqrt{n}} \sum_{k=1}^{n-1} U_n\left(\frac{k}{n}\right)^2 \zeta_{k+1}\right) \xrightarrow{D} \left(W_1, W_2, \int_0^1 {W_1}^2 dW_2\right). \tag{11}$$

Next we derive the following theorem on $\tilde{\rho}$.

Theorem 1 *Let $\xi_k = (\varepsilon_k, a_{k-1}a_k\varepsilon_k)'$ be a sequence of martingale differences with respect to increasing σ-fields $\mathcal{F}_k$ and satisfy the conditions of Lemma 1 and (7) of Lemma 2 (ii). Assume that $\frac{1}{n^2}\sum_{k=1}^{n-1} a_k a_{k+1} X_k^2 = v\frac{1}{n^2}\sum_{k=1}^{n-1} X_k^2 + o_p(1)$. Then*

$$n(\tilde{\rho} - 1) \xrightarrow{D} \frac{\sigma_{12}\int_0^1 BdB + \sqrt{\sigma_{11}\sigma_{22} - \sigma_{12}^2}\int_0^1 Bd\tilde{B}}{v\sigma_{11}\int_0^1 B^2 dt}, \tag{12}$$

where $(B, \tilde{B})'$is a two-dimensional standard Brownian motion and σ_{ij} is the (i,j)-th component of the covariance matrix $\boldsymbol{\Sigma}$ of Lemma 1 (ii).

If $a_k \equiv 1$, that is, in the case of complete sampling, $v = 1$, $\sigma_{ij} = 1 (i, j = 1, 2)$ and hence this limiting distribution is identical to (6).

Now we prepare a lemma to show the following example.

Lemma 3 *Let* $\{a_k, k = 1,..\}$ *and* $\{\varepsilon_k, k = 1,..\}$ *be sequences of random variables on independent probability spaces* $(\Omega_a, \mathcal{F}_a, P_a)$ *and* $(\Omega_e, \mathcal{F}_e, P_e)$ *respectively. Suppose that* $\{\varepsilon_k, k = 1,..\}$ *is a sequence of independently and identically distributed random variables. Let random variable* ξ_k *be* $\xi_k = (\varepsilon_k, a_k a_{k-1}\varepsilon_k)$. *Moreover let the strong mixing coefficients of* $\{\xi_k, k = 1,..\}$ *and* $\{a_k, k = 1,..\}$ *be* $\alpha(i;\xi)$ *and* $\alpha(i;a)$ *respectively. Then*

$$\alpha(i;\xi) \leq \alpha(i-1;a). \tag{13}$$

Example 1 (A Markov Process). $\{a_k, k = 1, 2, ...\}$ is a stochastic Markov process. Let p_1 and p_0 be $p_1 = P(a_k = 1|a_{k-1} = 1)$ and $p_0 = P(a_k = 1|a_{k-1} = 0)$. If $p_0 = p_1$, this process reduces to a sequence of Bernoulli trials. And we find that the stationary distribution $\pi = P(a_k = 1)$ is $\frac{p_0}{1-p_1+p_0}$.

Setting $\psi = p_1 - p_0$ and $u_k = a_k - E[a_k|a_i, i \leq k-1] = a_k - [p_0 + (p_1 - p_0)a_{k-1}]$, we can obtain the representation $a_k = \psi a_{k-1} + u_k$. Then the argument of a stationary Markov process (Billingsley (1968) p167-8) gives

$$\alpha(i;a) \leq 2|\psi|^i. \tag{14}$$

By (14) and Lemma 3, $\{\xi_k\}$ satisfies Lemma 1 (iv). Also the other conditions of Theorem 1 are satisfied if we put $\mathcal{F}_k = \sigma(a_i, \varepsilon_i; 1 \leq i \leq k)$.

Consequently we can apply Theorem 1 to this example. Then the covariance matrix of Lemma 1 is

$$\mathbf{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \pi p_1 \\ \pi p_1 & \pi p_1 \end{pmatrix}.$$

And v in Theorem 1 equals πp_1. Hence we have

$$n(\tilde{\rho}-1) \xrightarrow{D} \frac{\int_0^1 BdB + r\int_0^1 Bd\tilde{B}}{\int_0^1 B^2 dt}, \tag{15}$$

where $r = \sqrt{\frac{1-\pi p_1}{\pi p_1}}$.

Example 2 (Periodically Deterministic Case). $\{a_k, k=1,2,...\}$ is periodically deterministic. Let M be the period and hence $a_k = a_{k+M}$ for any k. And let $L = \sum_{k=1}^{M} a_k$, the number of observed data in one period and $K = \sum_{k=1}^{M} a_k a_{k+1}$.

Then the conditions of Theorem 1 are satisfied. The covariance matrix of Lemma 1 is

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \frac{K}{M} \\ \frac{K}{M} & \frac{K}{M} \end{pmatrix},$$

and v in Theorem1 equals $\frac{K}{M}$. Hence we have

$$n(\tilde{\rho}-1) \xrightarrow{D} \frac{\int_0^1 BdB + r\int_0^1 Bd\tilde{B}}{\int_0^1 B^2 dt}, \tag{16}$$

where $r = \sqrt{\frac{M-K}{K}}$.

Next we consider the limiting distribution of $\hat{\rho}$. First we derive the limiting distribution, when $\{a_k\}$ follows a stochastic Markov process. In contrast to complete sampling, $\hat{\rho}$ is not n-consistent but $\sqrt{n}$-consistent.

Theorem 2 *Suppose that $\{a_k, k=1,2,...\}$ is a stochastic Markov process. Then*

$$\sqrt{n}(\hat{\rho}-1) \xrightarrow{D} \frac{\int_0^1 \left\{B^2 - \left(\int_0^1 B^2 dt\right)\right\} d\tilde{B}}{\int_0^1 B^2 dt} \sqrt{\frac{(1-p_1)}{p_1}\frac{(1-p_1+p_0)}{p_0}}, \tag{17}$$

where $\{(B(t), \tilde{B}(t)); 0 \le t \le 1\}$ is a two-dimensional standard Brownian motion.

It should be remarked that the limiting distribution is degenerated if the data is observed completely and therefore $p_0 = p_1 = 1$ holds.

Next we give asymptotic properties of $\hat{\rho}$ in a periodically deterministic case. $n(\hat{\rho}-1)$ does not have a limiting distribution. Nevertheless if we put $n-1 = n^*M + r$ ($r = 0, 1, ..., M-1$) with period M and let n^* go to infinite, r being fixed, the limiting distribution of $(n^*M + r + 1)(\hat{\rho}-1)$ exists and depends on r.

Theorem 3 *Suppose that $\{a_k, k = 1, 2, ...\}$ is periodically deterministic. Then*

$$n(\hat{\rho}-1) \xrightarrow{\mathcal{D}} \frac{Q_r\left(\int_0^1 W_1^2 dt - W_1(1)^2\right) + \int_0^1 W_1 dW_2 + R}{\frac{LK}{M}\int_0^1 W_1^2 dt}, \tag{18}$$

where $Q_r = K\sum_{m=1}^{r+1} a_m - L\sum_{m=1}^{r} a_m a_{m+1}$, $R = \frac{1}{M}\sum_{m=1}^{M}(La_m a_{m+1} - Ka_m)m$, and $(W_1, W_2)'$ is a Brownian motion with a covariance matrix Σ whose $(i-j)$th-components are $\sigma_{11} = 1$, $\sigma_{12} = \sigma_{21} = R + \frac{LK}{M}$ and $\sigma_{22} = \frac{1}{M}\sum_{j=1}^{M}\left[2\sum_{j=m}^{m}(La_m a_{m+1} - Ka_m) + La_j a_{j-1}\right]^2$.

We have that $L = K = M$, $Q_r = M$, $R = 0$, $\sigma_{12} = M$ and $\sigma_{22} = M^2$, in the case of complete sampling. Then $(W_1, W_2) = (B, MB)$, where $\{B\}$ is a standard Brownian motion. The limiting distribution of $n(\tilde{\rho}-1)$ is

$$\frac{\int_0^1 B^2 dt - B(1)^2 + \int_0^1 B dB}{\int_0^1 B^2 dt}.$$

3 COMPUTATIONAL EXPERIMENTS

In this section we shall reinforce the results in the previous section by computational experiments. First we compare the finite sample distributions with the limiting distributions. Second as an application, we also consider testing for a unit root.

To generate random numbers, we used "ran2.c" of *Numerical Recipes in C* [Press et al (1988)]. And to generate a Brownian motion, we used an expansion of Brownian motion,

$$B(t) = \sum_{k=0}^{\infty} \frac{2\sqrt{2}}{(2k+1)\pi} sin\left(k+\frac{1}{2}\right)\pi t Z_k,$$

where Z_k are independently and identically normal variables $N(0,1)$. [See Chan and Wei (1988)]. Then this expansion and term by term integration give

$$\int_0^1 B^2 dt = \sum_{k=0}^{\infty} \frac{4}{(2k+1)^2\pi^2} Z_k^2,$$

and

$$\int_0^1 B^4 dt = \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\sum_{k_3=0}^{\infty}\sum_{k_4=0}^{\infty} \frac{8 Z_{k_1} Z_{k_2} Z_{k_3} Z_{k_4}}{\prod_{i=1}^{4}[(2k_i+1)\pi]} A(k_1.k_2,k_3,k_4),$$

where

$$\begin{aligned} &A(k_1,k_2,k_3,k_4) \\ &= I(k_1+k_2-k_3-k_4=0) + I(k_1-k_2-k_3+k_4=0) + I(k_1-k_2+k_3-k_4=0) \\ &\quad -I(k_1+k_2+k_3-k_4+1=0) - I(k_1+k_2-k_3+k_4+1=0) \\ &\quad -I(k_1-k_2+k_3+k_4+1=0) - I(-k_1+k_2+k_3+k_4+1=0), \end{aligned}$$

and $I(\cdot)$ means an indicator function. And we note that

$$\int_0^1 BdB = \frac{1}{2}(B(1)^2 - 1)$$

by Ito's rule, [see Karatzas and Shreve (1991)] and

$$\int_0^1 Bd\tilde{B} \overset{D}{\sim} Z\sqrt{\int_0^1 B^2 dt},$$

where $\overset{D}{\sim}$ means equivalence in distribution, $(B, \tilde{B})$ is a standard Brownian motion and Z is a normal variable $N(0,1)$ independent of B, since $\{B\}$ and $\{\tilde{B}\}$ are independent processes.

When $\{a_k\}$ follows a stochastic Markov process, the limiting distributions $\tilde{\rho}$ and $\hat{\rho}$ are quite different. Here we consider $p_0 = p_1 = p = 0.95$, that is, the case that $\{a_k; k = 1, 2, ...\}$ is a sequence of Bernoulli trials. The limiting distribution of $n(\tilde{\rho} - 1)$ is given by Example 1 with $r = \sqrt{\frac{1-\pi p_1}{\pi p_1}} = \sqrt{\frac{1-0.95}{0.95}}$. Figure 1 gives the comparison of the finite sample distribution with the limiting distribution. The sample size is $n = 100$. The number of replications is $5,000$. We can see that the finite sample distribution is close to the limiting distribution.

Next we compare the finite sample distribution of $\sqrt{n}(\hat{\rho} - 1)$ with its limiting distribution. This limiting distribution is

$$\frac{\int_0^1 \left\{B^2 - \left(\int_0^1 B^2 dt\right)\right\} d\tilde{B}}{\int_0^1 B^2 dt} \frac{\sqrt{1-p}}{p} \overset{D}{\sim} \frac{\sqrt{\int_0^1 B^4 dt - \left(\int_0^1 B^2 dt\right)^2}}{\int_0^1 W_1^2 dt} Z \frac{\sqrt{1-p}}{p},$$

where $\{(B, \tilde{B}); 0 \leq t \leq 1\}$is a two-dimensional standard Brownian motion and Z is a standard normal variable independent of B. This distribution is a mixed normal. Figure 2 compares the finite sample distribution with the limiting distribution when the sample size are $n = 100$, 500 and $100,000$. The number of replications is 5,000. Then Figure 2 shows that $\sqrt{n}(\hat{\rho} - 1)$ converges very slowly as n increases.

On the other hand, in a periodically deterministic case, we adopt A-B sampling in which A consecutive values of X_k are observed, B missed, A observed and so on. This sampling scheme was used in Parzen(1965), Dunsmuir and Robinson (1981), Shin and Sarker (1996). We give quantiles of $n(\tilde{\rho} - 1)$ and $n(\hat{\rho} - 1)$, to compare the finite sample distribution with the limiting distribution. These quantiles are applied to a unit root test later. Theorem 3 shows that the limiting distribution of $(n^*M + r + 1)(\hat{\rho} - 1)$ exists as n^* goes to infinity, r being fixed. Therefore we consider $n = 51$, 100 and 499. Then r is equal

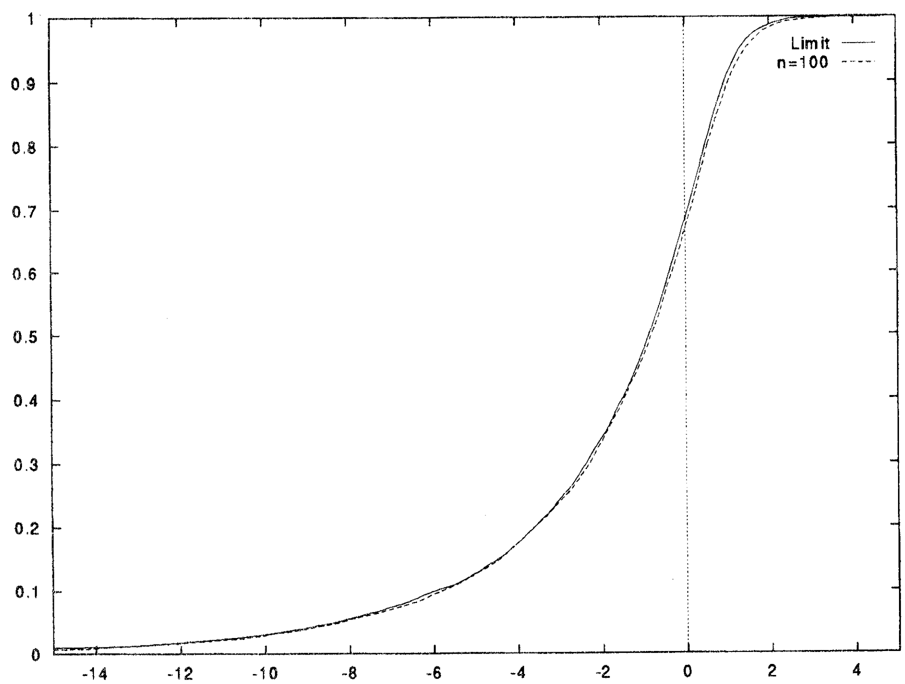

Figure 1: Comparison of distribution of finite sample distribution with that of limiting distribution of $n(\tilde{\rho} - 1)$.

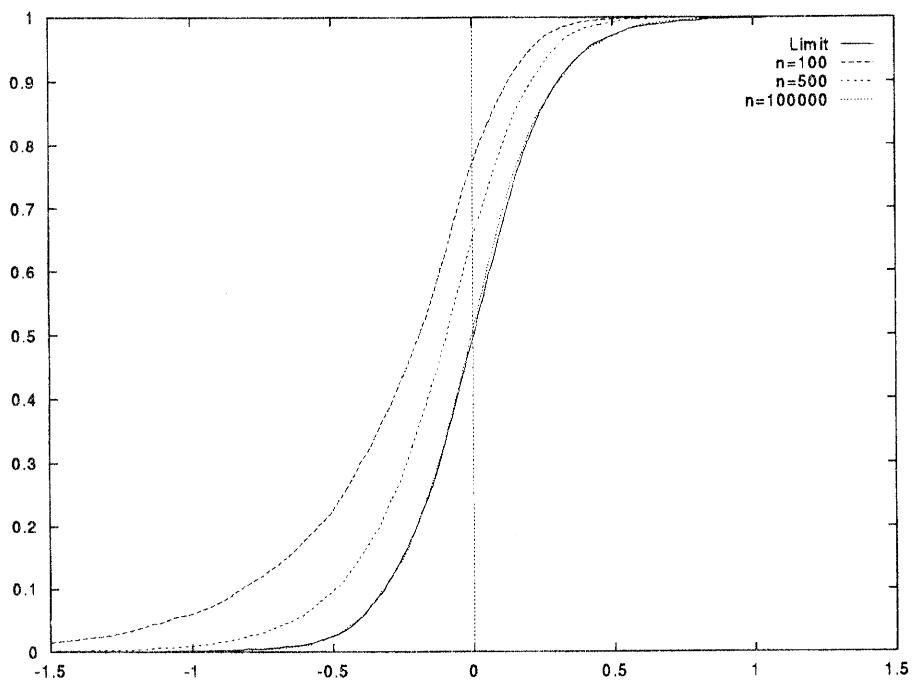

Figure 2: Comparison of distribution of finite sample distribution with that of limiting distribution of $\sqrt{n}(\hat{\rho}-1)$.

Table 1: Empirical Cumulative Distributions of $n(\tilde{\rho}-1)$.

	n	.01	.05	.10	.20	.50	.80	.90	.95	.99
A=6,B=1										
	51	-13.85	-8.38	-5.93	-3.62	-0.77	0.72	1.33	1.85	3.01
	100	-14.20	-8.45	-5.99	-3.65	-0.78	0.71	1.30	1.79	2.89
	499	-14.76	-8.65	-6.12	-3.70	-0.79	0.69	1.27	1.75	2.80
	Limit	-14.86	-8.66	-6.12	-3.71	-0.80	0.68	1.26	1.74	2.77
A=5,B=2										
	51	-14.77	-8.77	-6.21	-3.78	-0.72	0.94	1.65	2.31	3.78
	100	-14.94	-8.85	-6.30	-3.82	-0.74	0.90	1.59	2.21	3.62
	499	-15.59	-9.12	-6.44	-3.86	-0.76	0.87	1.56	2.16	3.47
	Limit	-15.75	-9.15	-6.44	-3.87	-0.77	0.86	1.55	2.14	3.42
A=4,B=3										
	51	-15.80	-9.39	-6.66	-3.98	-0.69	1.25	2.17	3.04	5.07
	100	-16.35	-9.58	-6.77	-4.05	-0.71	1.20	2.09	2.91	4.84
	499	-16.85	-9.83	-6.92	-4.12	-0.74	1.15	2.02	2.80	4.57
	Limit	-17.07	-9.87	-6.93	-4.13	-0.75	1.14	2.00	2.77	4.50

to 1 in all cases. Table 1-2 gives the result of $\tilde{\rho}$ and $\hat{\rho}$. The numbers of replications are $100,000$ in the case of finite sample size, while the numbers of replications are $200,000$ in the case of limiting distributions. These tables show that the quantiles of the finite sample distribution is close to those of the limiting distributions as n increases. We also observe that the left tails are longer as n increases and that the tails are fatter as B is larger. And $\hat{\rho}$ has a fatter tail than $\tilde{\rho}$ does.

And we compare the limiting distributions of $\tilde{\rho}$ and $\hat{\rho}$ in complete sampling and various

Table 2: Empirical Cumulative Distributions of $n(\hat{\rho}-1)$.

	n	p .01	.05	.10	.20	.50	.80	.90	.95	.99
A=6,B=1										
	51	-15.88	-10.43	-8.01	-5.62	-2.57	-0.90	-0.28	0.15	0.88
	100	-16.60	-10.76	-8.22	-5.71	-2.60	-0.92	-0.31	0.12	0.85
	499	-17.28	-11.02	-8.41	-5.86	-2.66	-0.95	-0.34	0.10	0.84
	Limit	-17.39	-11.09	-8.45	-5.88	-2.67	-0.96	-0.34	0.10	0.84
A=5,B=2										
	51	-16.74	-10.95	-8.41	-5.92	-2.71	-0.92	-0.26	0.23	1.02
	100	-17.33	-11.21	-8.54	-5.96	-2.73	-0.93	-0.27	0.22	1.02
	499	-17.86	-11.42	-8.72	-6.09	-2.77	-0.95	-0.28	0.20	1.01
	Limit	-17.89	-11.46	-8.75	-6.11	-2.77	-0.95	-0.28	0.20	1.00
A=4,B=3										
	51	-17.62	-11.34	-8.75	-6.14	-2.81	-0.94	-0.22	0.31	1.15
	100	-18.09	-11.62	-8.89	-6.23	-2.84	-0.94	-0.22	0.30	1.17
	499	-18.36	-11.79	-9.02	-6.32	-2.88	-0.95	-0.22	0.30	1.17
	Limit	-18.42	-11.84	-9.06	-6.34	-2.88	-0.95	-0.22	0.30	1.17

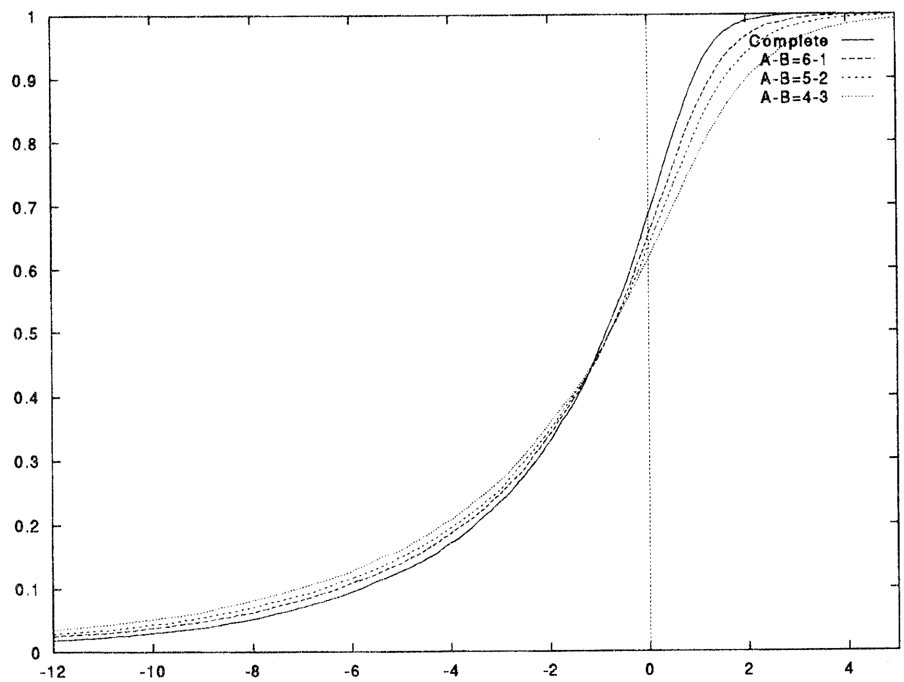

Figure 3: Limiting distributions of $n(\tilde{\rho}-1)$ with complete sampling, A-B=6-1. A-B=5-2 and A-B=4-3.

A-B sampling. In Figure 3-4, the numbers of replications are 10,000. Figure 3 gives the limiting distributions of $\tilde{\rho}$ with complete sampling and A-B sampling (A-B = 6-1, 5-2 and 4-3). Figure 3 shows that the limiting distributions of $n(\tilde{\rho}-1)$ with A-B sampling converge to that for complete sampling as B is smaller. On the other hand Figure 4 gives the result of $\hat{\rho}$. In contrast to $\tilde{\rho}$, the limiting distribution of $n(\hat{\rho}-1)$ with A-B sampling converges slowly to that for complete sampling as B is smaller.

The main purpose of this paper is to derive asymptotic properties of $\hat{\rho}$ and $\tilde{\rho}$ if they are used to estimate the autoregressive parameter in the case that the true underlying process is a nonstationary AR(1) process with a unit root.

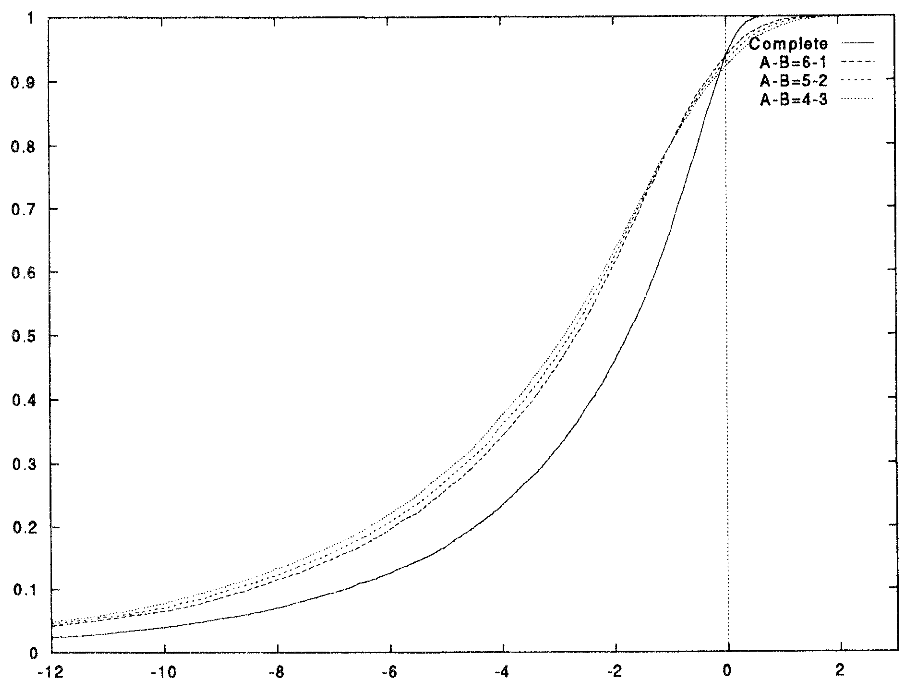

Figure 4: Limiting distributions of $n(\hat{\rho}-1)$ with complete sampling, A-B=6-1. A-B=5-2 and A-B=4-3.

While a unit root test is an important issue in economic time series analysis. Hence as an application we use $\hat{\rho}$ and $\tilde{\rho}$ to be test statistics for unit root tests. Recently Toda and Mckenzie (1994) and Shin and Sarker (1996) proposed testing procedures for a unit root test of time series data with missing observations. Their test statistics have the same limiting distribution as that given by Dicky and Fuller (1979) for the case of complete sampling. Among them we take up $\hat{\rho}_{SS}$ defined by

$$\hat{\rho}_{SS} = \frac{\sum_{k=1}^{n-1} z_k z_{k+1}}{\sum_{k=1}^{n-1} z_k^2},$$

where $\{X_k\}$ is observed at times $k_1, k_2, ..., k_l$ and $z_k = X_{k_{i-1}}$ for $k_{i-1} \leq k < k_i$, $1 \leq k \leq n$. $\hat{\rho}_{SS}$ is proposed by Shin and Sarker (1996) and is the dominating term of the one-step Newton-Raphson estimator for log Gaussian likelihood conditional on observed data.

We compare the performance of $\hat{\rho}$ and $\tilde{\rho}$ with that of $\hat{\rho}_{SS}$ by computational experiments. Toda and Mckenzie (1994) and Shin and Sarker (1996) cover only the case that the sampling intervals, $k_{i+1} - k_i$ $(i = 1, 2, ..)$ are bounded almost surely as the sample size $n \to \infty$. Hence their results cannot apply to the case that $\{a_k\}$ is a stochastic Markov process and then we consider only A-B sampling. We use $n(\hat{\rho}_{SS} - 1)$, $n(\hat{\rho} - 1)$ and $n(\tilde{\rho} - 1)$ as test statistics. We use the previous quantile as a critical value 5% for $\tilde{\rho}$ and $\hat{\rho}$ and use a critical value in Fuller (1976) of the finite sample distribution for $\hat{\rho}_{SS}$ respectively. Then we calculate the empirical power. Table 3 gives empirical power with the sample size $n = 51$, 100 and 499 and the numbers of replications are 10,000. A-B are 6-1, 5-2 and 4-3.

Table 3: Comparison of Empirical Power of $n(\hat{\rho}_{SS} - 1)$, $n(\tilde{\rho} - 1)$ and $n(\hat{\rho} - 1)$.

	A=6 B=1			A=5 B=2			A=4 B=3		
ρ	$\hat{\rho}_{SS}$	$\tilde{\rho}$	$\hat{\rho}$	$\hat{\rho}_{SS}$	$\tilde{\rho}$	$\hat{\rho}$	$\hat{\rho}_{SS}$	$\tilde{\rho}$	$\hat{\rho}$
n=51									
0.50	1.000	0.998	0.993	0.999	0.990	0.983	0.996	0.972	0.967
0.55	0.999	0.993	0.983	0.998	0.980	0.965	0.988	0.950	0.940
0.60	0.998	0.983	0.962	0.994	0.964	0.938	0.976	0.923	0.909
0.65	0.992	0.961	0.913	0.982	0.931	0.882	0.952	0.877	0.846
0.70	0.970	0.917	0.838	0.947	0.868	0.798	0.900	0.808	0.761
0.75	0.910	0.835	0.723	0.875	0.786	0.680	0.815	0.714	0.653
0.80	0.773	0.702	0.567	0.730	0.647	0.523	0.659	0.584	0.505
0.85	0.548	0.499	0.383	0.516	0.471	0.365	0.466	0.431	0.353
0.90	0.317	0.297	0.229	0.302	0.284	0.217	0.271	0.271	0.220
0.95	0.149	0.147	0.121	0.141	0.150	0.122	0.129	0.143	0.123
1.00	0.050	0.050	0.051	0.048	0.052	0.051	0.046	0.052	0.050
n=100									
0.50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.55	1.000	1.000	1.000	1.000	1,000	1.000	1.000	1.000	1.000
0.60	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.999
0.65	1.000	1.000	1.000	1.000	0.999	0.999	1.000	0.993	0.998
0.70	1.000	1.000	0.998	1.000	0.994	0.997	1.000	0.983	0.991
0.75	1.000	0.996	0.990	1.000	0.987	0.980	0.999	0.956	0.965
0.80	0.998	0.980	0.948	0.996	0.954	0.924	0.991	0.903	0.899
0.85	0.969	0.906	0.814	0.959	0.854	0.775	0.932	0.780	0.746
0.90	0.759	0.673	0.538	0.734	0.633	0.511	0.699	0.572	0.487
0.95	0.316	0.302	0.224	0.306	0.285	0.214	0.290	0.266	0.210
1.00	0.048	0.050	0.047	0.047	0.053	0.045	0.045	0.050	0.048
n=499									
0.90	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000
0.91	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000
0.92	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	1.000
0.93	1.000	1.000	0.999	1.000	0.999	0.999	1.000	0.987	0.998
0.94	1.000	1.000	0.999	1.000	0.995	0.997	1.000	0.977	0.994
0.95	1.000	0.996	0.989	1.000	0.982	0.986	1.000	0.948	0.978
0.96	0.998	0.975	0.949	0.998	0.945	0.935	0.998	0.890	0.917
0.97	0.966	0.900	0.813	0.964	0.845	0.786	0.962	0.775	0.762
0.98	0.748	0.664	0.532	0.744	0.616	0.504	0.738	0.560	0.482
0.99	0.319	0.295	0.220	0.318	0.284	0.210	0.314	0.265	0.202
1.00	0.052	0.050	0.049	0.051	0.050	0.047	0.052	0.051	0.049

Table 4: Size and Power of $n(\hat{\rho}_{SS}-1)$, $n(\tilde{\rho}-1)$ and $n(\hat{\rho}-1)$.

	A=6 B=1			A=5 B=2			A=4 B=3		
ρ	$\hat{\rho}_{SS}$	$\tilde{\rho}$	$\hat{\rho}$	$\hat{\rho}_{SS}$	$\tilde{\rho}$	$\hat{\rho}$	$\hat{\rho}_{SS}$	$\tilde{\rho}$	$\hat{\rho}$
n=51									
0.50	1.000	0.997	0.990	0.999	0.988	0.978	0.993	0.966	0.961
0.55	0.999	0.991	0.975	0.997	0.976	0.957	0.984	0.941	0.932
0.60	0.997	0.980	0.948	0.992	0.957	0.924	0.969	0.909	0.894
0.65	0.988	0.953	0.891	0.976	0.918	0.858	0.936	0.858	0.826
0.70	0.962	0.903	0.798	0.932	0.849	0.765	0.877	0.781	0.735
0.75	0.889	0.812	0.674	0.847	0.757	0.642	0.779	0.680	0.620
0.80	0.739	0.669	0.509	0.694	0.610	0.483	0.618	0.545	0.474
0.85	0.510	0.463	0.336	0.475	0.434	0.334	0.426	0.392	0.328
0.90	0.288	0.270	0.193	0.270	0.254	0.192	0.243	0.238	0.201
0.95	0.133	0.131	0.100	0.124	0.130	0.104	0.114	0.121	0.109
1.00	0.044	0.045	0.040	0.041	0.046	0.043	0.039	0.044	0.044
n=100									
0.50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.55	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.60	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.999
0.65	1.000	1.000	1.000	1.000	0.999	0.999	1.000	0.992	0.998
0.70	1.000	0.999	0.997	1.000	0.994	0.995	1.000	0.980	0.989
0.75	1.000	0.995	0.987	1.000	0.984	0.976	0.999	0.949	0.962
0.80	0.997	0.978	0.938	0.996	0.947	0.916	0.989	0.890	0.892
0.85	0.964	0.892	0.789	0.952	0.835	0.756	0.924	0.756	0.734
0.90	0.741	0.649	0.508	0.716	0.602	0.485	0.681	0.538	0.472
0.95	0.302	0.279	0.204	0.291	0.258	0.199	0.275	0.239	0.200
1.00	0.046	0.046	0.041	0.045	0.046	0.042	0.042	0.043	0.046
n=499									
0.90	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000
0.91	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000
0.92	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	1.000
0.93	1.000	1.000	0.999	1.000	0.999	0.999	1.000	0.987	0.998
0.94	1.000	1.000	0.999	1.000	0.995	0.997	1.000	0.977	0.994
0.95	1.000	0.996	0.989	1.000	0.982	0.986	1.000	0.947	0.977
0.96	0.998	0.975	0.947	0.998	0.944	0.934	0.998	0.890	0.915
0.97	0.964	0.899	0.808	0.961	0.843	0.784	0.958	0.774	0.759
0.98	0.741	0.663	0.527	0.735	0.613	0.502	0.730	0.556	0.479
0.99	0.312	0.294	0.217	0.311	0.281	0.208	0.307	0.263	0.201
1.00	0.051	0.050	0.048	0.049	0.049	0.047	0.050	0.050	0.048

Moreover we also use the quantile of limiting distribution as a critical value for $\tilde{\rho}$ and $\hat{\rho}$. Similarly we use a critical value in Fuller (1976) of the limiting distribution for $\hat{\rho}_{SS}$. Table 4 gives empirical power with the sample size $n = 51$, 100 and 499 and the numbers of replications are 10,000. A-B are 6-1, 5-2 and 4-3. The size and the power are smaller than in Table 3, because the critical values of Table 4 are smaller than those of Table 3.

Table 3-4 show that $\hat{\rho}_{SS}$ mostly performs better than $\tilde{\rho}$ and $\hat{\rho}$. And there is not a clear rating between $\tilde{\rho}$ and $\hat{\rho}$. However when $\{X_k\}$ is a stationary process. $\hat{\rho}_{SS}$ is not a consistent estimator and converges almost surely to

$$\frac{(A-1)\rho + B + \rho^{B+1}}{A+B} \qquad as\ n \to \infty.$$

Consequently the bias is positive since

$$\begin{aligned} \frac{(A-1)\rho + B + \rho^{B+1}}{A+B} - \rho &= \frac{\rho^{B+1} + B - (B+1)\rho}{A+B} \\ &= \frac{(1-\rho)[B - \rho(1+\rho+\dots+\rho^{B-1})]}{A+B}, \end{aligned}$$

and becomes larger as ρ is closer to zero and A is smaller. On the other hand $\tilde{\rho}$ and $\hat{\rho}$ are still consistent estimators of ρ.

4 ADDITIONAL COMMENTS

(1) Throughout this paper we assumed that $\{\varepsilon_k\}$ and $\{a_k\}$ are independent stochastic sequences if $\{a_k\}$ is stochastic one. However Theorem 1 can be applied to some cases that $\{\varepsilon_k\}$ and $\{a_k\}$ depends on each other.

For example let $\{a_k\}$ be defined by

$$\begin{cases} a_k = 1 & |\varepsilon_k| \leq C, \\ a_k = 0 & |\varepsilon_k| > C, \end{cases}$$

with some constant C. This sampling scheme means that if $\{\varepsilon_k\}$ is an outlier, X_k is not observed. And we assume that $\{\varepsilon_k\}$ is independently and identically distributed and has a symmetric distribution with mean 0, variance σ^2 and finite fourth-order moment. Then

$$E(a_k a_{k-1}\varepsilon_k|\mathcal{F}_{k-1}) = a_{k-1}E(\varepsilon_k I(|\varepsilon_k| \leq C) = 0 \quad a.s., \tag{19}$$

where $\mathcal{F}_k$ is the σ-field generated by $\{\varepsilon_1, \varepsilon_2, ..., \varepsilon_k\}$ and $I(\cdot)$ implies a indicator function. Hence $\{\varepsilon_k, a_k a_{k-1}\varepsilon_k\}$ is a sequence of martingale differences with respect to $\mathcal{F}_k$. And it is easily shown that the covariance matrix of $(W_1, W_2)'$ in Theorem 1 is

$$\mathbf{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \tau p \\ \tau p & \tau p \end{pmatrix}$$

and

$$\frac{1}{n^2}\sum_{k=1}^{n-1} a_k a_{k+1} X_k^2 = \frac{p^2}{n^2}\sum_{k=1}^{n-1} X_k^2 + o_p(1),$$

where $\tau = \frac{1}{\sigma^2}E({\varepsilon_k}^2 I(|\varepsilon_k| \leq C)$ and $p = Pr(|\varepsilon_k| \leq C)$. Then

$$n(\tilde{\rho} - 1) \xrightarrow{D} \frac{\tau \int_0^1 BdB + \sqrt{\tau(1-\tau p)/p}\int_0^1 Bd\tilde{B}}{p\int_0^1 B^2 dt}$$

holds.

(2) We consider asymptotic properties of $\tilde{\rho}$ under the assumption that $\{X_k\}$ is a non-stationary process. But $\tilde{\rho}$ is also an useful estimator of an autocorrelation function of a stationary process. Now we assume that $|\rho| < 1$. Then

$$\sqrt{n}(\tilde{\rho} - \rho) = \frac{\frac{1}{\sqrt{n}}\sum_{k=1}^{n-1} X_k \varepsilon_{k+1} a_k a_{k+1}}{\frac{1}{n}\sum_{k=1}^{n-1} a_k a_{k+1} X_k^2}.$$

If $\{a_k; k = 1, 2, ..\}$ is a sequence of Bernoulli trials, then we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n-1} a_k a_{k+1} X_k^2 = EX_1^2 p^2$$

and

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n-1} X_k \varepsilon_{k+1} a_k a_{k+1} \xrightarrow{D} N(0, (EX_1^2)^2 p^2 (1-\rho^2)),$$

where $p = Pr(a_k = 1)$. Hence

$$\sqrt{n}(\tilde{\rho} - \rho) \xrightarrow{D} N(0, \frac{1-\rho^2}{p^2}).$$

On the other hand asymptotic variance of $\sqrt{n}(\hat{\rho} - \rho)$ is $\frac{1+\rho^2(1-2p)}{p^2}$ [See Dunsmuir and Robinson (1981)]. Consequently $\tilde{\rho}$ is asymptotically more efficient than $\hat{\rho}$. In order to estimate the autocorrelation of a stationary process at lag l, we can generalize the definition of $\hat{\rho}$ and $\tilde{\rho}$ by

$$\hat{\rho}(l) = \frac{\sum_{k=1}^{n-l} Y_k Y_{k+l} / \sum_{k=1}^{n-l} a_k a_{k+l}}{\sum_{k=1}^{n} Y_k^2 / \sum_{k=1}^{n} a_k}$$

and

$$\tilde{\rho}(l) = \frac{\sum_{k=1}^{n-l} Y_k Y_{k+l}}{\sum_{k=1}^{n-l} a_{k+l} Y_k^2}$$

respectively. Asymptotic properties of $\hat{\rho}$ and $\tilde{\rho}$ are investigated by Yajima and Nishino (1996).

References

[1] Billngsley, P. (1968) *Convergence of Probability Measure.* New York: John Wiley.

[2] Chow, Y.S. and Teicher, H. (1988) *Probability Theory 2nd.* New York: Springer

[3] Chan, N.H. and Wei, C.Z. (1988) Limiting distributions of least squares estimates of unstable autoregressive processes. *Annals of Statistics* 16, 367-401.

[4] Dickey, D.A. and Fuller, W.A. (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427-431.

[5] Dunsmuir, W. and Robinson, P.M. (1981) Asymptotic theory for time series containing missing and amplitude modulated observations. *Sankhyā, Ser A*, 43, 260-281.

[6] Fuller, W.A. (1976) *Introduction to Statistical Time Series.* New York: Wiley.

[7] Herrndorf, N. (1984) A functional central limit theorem for weakly dependent sequences of random variables. *Annals of Probability* 12, 141-153.

[8] Karatzas, I. and Shreve, S. E. (1991) *Brownian Motion and Stochastic Calculus 2nd.* New York: Springer.

[9] McLeish, D.L. (1975) Invariance principles for dependent variables. *Z. Wahrsheinlichkeitstheorie und Verw. Gebiete* 32, 165-178.

[10] Parzen.E. (1963) On spectral analysis with missing observations and amplitude modulation. *Sankhyā, Ser A* 25, 383-392.

[11] Phillips, P.C.B. and Durlauf, S.N. (1986) Multiple time series regressions with integrated processes. *Review of Economic Studies* 53, 473-495.

[12] Shin, D.W. and Sarker, S.(1995) Estimation for AR(1) models with nonconsecutively observed samples. *Sankhyā Ser A*, 57, 287-298.

[13] Shin, D.W. and Sarker, S.(1996) Testing for a unit root using irregularly observed data. *Journal of Time Series Analysis* 17, 309-320.

[14] Takeuchi, K. (1995) A comment on "Recent Development of Economic Data Analysis "at the 63th Annual Meeting of Japan Statistical Society.

[15] Toda, H.Y. and Mckenzie, C.R. (1994) LM tests for unit roots in the presence of missing observations. *Discussion Paper 94-10 Faculty of Economics Osaka University.*

[16] White, J.S. (1958) The limiting distribution of the serial correlation coefficient in the explosive case. *Annals of Mathematical Statistics* 29, 1188-1197.

[17] Press, W.H at el (1988) *Numerical Recipes in C.* New York: Cambridge University Press.

[18] Yajima, Y and Nishino, H (1996) On estimation of autocorrelation functions for stationary processes with missing observations (forthcoming).

APPENDIX

A *Proof of Lemma 1*

Let $X_n(t)$ be $\frac{1}{\sqrt{n}}\sum_{s=1}^{[nt]}\Sigma^{-\frac{1}{2}}\xi_s$. Then since the tightness of $\{X_n(t)\}$ is proved in the same way as the proof of Phillips and Durlauf (1986) Theorem 2.1, it suffices to show that the finite dimensional distributions of $X_n(t)$ converge weakly as $n \to \infty$ to those of a d-dimensional standard Brownian motion $B(t)$.

We consider arbitrary linear combination $Y_n(t) = \lambda' X_n(t) = \frac{1}{\sqrt{n}} \sum_{s=1}^{[nt]} v_s(\text{say})$ with $\lambda'\lambda = 1$. $\{v_s\}$ satisfy the condition of Lemma 1 for $d = 1$ and by univariate invariance principle of Herrndorf (1984) Corollary 1 we obtain $Y_n(t) \xrightarrow{D} V(t)$ as $n \to \infty$, where $V(t)$ is a 1-dimensional Brownian motion. Writing $V(t) = \lambda' B(t)$, we obtain $\lambda' X_n(t) \xrightarrow{D} \lambda' B(t)$ as $n \to \infty$ for arbitrary λ with $\lambda'\lambda = 1$. By the Cramer-Wold device we show that the finite dimensional distributions of $X_n(t)$ converge weakly as $n \to \infty$ to those of $B(t)$. ∎

B *Proof of Lemma 2*

See Chan and Wei (1988) Theorem 2.4 (ii) for the proof of (i).

We only show the outline of the proof of (ii), since the proof is similar to that of Chan and Wei (1988) Theorem 2.4 (ii). By the Skorokhod representation theorem, there is another probability space $\hat{\Omega}$ and $D[0,1]$-valued random variables U^n, V^n such that

$$\|(U^n, V^n) - (W_1, W_2)\|_\infty \quad \to \quad 0 \qquad a.s. \tag{20}$$

and

$$(U^n, V^n) \quad \overset{D}{\sim} \quad (U_n, V_n) \quad . \tag{21}$$

Let $\overset{D}{\sim}$ denote equivalent in distribution. Defining

$$G^n = \sum_{k=1}^{n-1} U^n \left(\frac{k}{n}\right)^2 \left(V^n\left(\frac{k+1}{n}\right) - V^n\left(\frac{k}{n}\right)\right)$$

and

$$G_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n-1} U_n \left(\frac{k}{n}\right)^2 \zeta_{k+1},$$

we have

$$(U^n, V^n, G^n) \quad \overset{D}{\sim} \quad (U_n, V_n, G_n) \quad . \tag{22}$$

In order to show(11), it suffices to show that

$$G^n \quad \xrightarrow{p} \quad \int_0^1 {W_1}^2 dW_2. \tag{23}$$

By (20) and Egorov's theorem, for given $\epsilon > 0$, there is an event $\hat{\Omega}_\epsilon \subset \hat{\Omega}$ such that $P(\hat{\Omega}_\epsilon) \geq 1 - \epsilon$ and

$$sup\{\|(U^n(\omega), V^n(\omega)) - (W_1(\omega), W_2(\omega))\|_\infty : \omega \in \hat{\Omega}_\epsilon\} = \delta_n \quad \rightarrow \quad 0. \tag{24}$$

Note that δ_n is a sequence of constants. We can choose integers $N(n) \rightarrow \infty$ such that

$$N(n){\delta_n}^2 \rightarrow 0 \quad and \quad N(n)/n \rightarrow 0. \tag{25}$$

For each n, we can further choose a partition $\{t_0, ..., t_{N(n)}\}$ of [0,1] such that

$$0 = t_0 < t_1(n) = \frac{n_1^*}{n} < t_2(n) = \frac{n_2^*}{n} < \quad ... \quad < t_{N(n)}(n) = \frac{n_{N(n)}^*}{n} = 1, \tag{26}$$

$$max\{|t_{i+1} - t_i| : 0 \leq i \leq N(n) - 1\} \quad = \quad o(1). \tag{27}$$

We first show that

$$G^n = \quad \sum_{k=1}^{N(n)} U^n(t_{k-1})^2 (V^n(t_k) - V^n(t_{k-1})) + o_p(1). \tag{28}$$

We set that

$$J_n = \quad G^n - \sum_{k=1}^{N(n)} U^n(t_{k-1})^2 (V^n(t_k) - V^n(t_{k-1})).$$

Using the fact that (ε_k, ζ_k) are martingale differences and (10), we have

$$\begin{aligned}
E(J_n)^2 &= E\left[\sum_{k=1}^{N(n)}\sum_{i=n_{k-1}^*}^{n_k^*-1}\left(U_n\left(\frac{i}{n}\right)^2 - U_n(t_{k-1})^2\right)\left(V_n\left(\frac{i+1}{n}\right) - V_n\left(\frac{i}{n}\right)\right)\right]^2 \\
&= \sum_{k=1}^{N(n)}\sum_{i=n_{k-1}^*}^{n_k^*-1} E\left[\left(U_n\left(\frac{i}{n}\right)^2 - U_n(t_{k-1})^2\right)\left(V_n\left(\frac{i+1}{n}\right) - V_n\left(\frac{i}{n}\right)\right)\right]^2 \\
&\leq \frac{c}{n}\sum_{k=1}^{N(n)}\sum_{i=n_{k-1}^*}^{n_k^*-1} E\left[\left(U_n\left(\frac{i}{n}\right)^2 - U_n(t_{k-1})^2\right)\right]^2 \\
&= \frac{c}{n}\sum_{k=1}^{N(n)}\sum_{i=n_{k-1}^*}^{n_k^*-1} E\left[\left(U_n\left(\frac{i}{n}\right) - U_n(t_{k-1})\right)^4 + 4U_n(t_{k-1})^2\left(U_n\left(\frac{i}{n}\right) - U_n(t_{k-1})\right)^2\right] \\
&\leq \frac{c}{n^3}\sum_{k=1}^{N(n)}\sum_{i=n_{k-1}^*}^{n_k^*-1}\left[(i-n_{k-1}^*)\tau + 3(i-n_{k-1}^*)^2\sigma^4 + 4(i-n_{k-1}^*)n_{k-1}^*\sigma^4\right] \\
&\leq \frac{c}{n^3}\sum_{k=1}^{N(n)}\left[\frac{1}{2}(n_k^*-n_{k-1}^*)^2\tau + (n_k^*-n_{k-1}^*)^3\sigma^4 + 2(n_k^*-n_{k-1}^*)^2 n_{k-1}^*\sigma^4\right] \\
&\leq \frac{c\tau}{2n} max_k|t_k - t_{k-1}|\sum_{k=1}^{N(n)}\left(\frac{n_k^*}{n} - \frac{n_{k-1}^*}{n}\right) + c\sigma^4 max_k|t_k - t_{k-1}|^2\sum_{k=1}^{N(n)}\left(\frac{n_k^*}{n} - \frac{n_{k-1}^*}{n}\right) \\
&\quad +2c\sigma^4 max_k|t_k - t_{k-1}|\sum_{k=1}^{N(n)}\left(\frac{n_k^*}{n} - \frac{n_{k-1}^*}{n}\right) = o(1). \qquad (29)
\end{aligned}$$

The last identity of (29) is given by (27) and $\sum_{k=1}^{N(n)}\left(\frac{n_k^*}{n} - \frac{n_{k-1}^*}{n}\right) = 1$. Thus (28) is shown.

Next we show that

$$I_{\hat{\Omega}_\epsilon}\sum_{k=1}^{N(n)} U^n(t_{k-1})^2\left(V^n(t_k) - V^n(t_{k-1})\right) = I_{\hat{\Omega}_\epsilon}\sum_{k=1}^{N(n)} W_1(t_{k-1})^2\left(V^n(t_k) - V^n(t_{k-1})\right) + o_p(1). (30)$$

We obtain the following inequality

$$\begin{aligned}
&|I_{\hat{\Omega}_\epsilon}\sum_{k=1}^{N(n)}\left(U^n(t_{k-1})^2 - W_1(t_{k-1})^2\right)\left(V^n(t_k) - V^n(t_{k-1})\right)|^2 \\
&\leq 2|I_{\hat{\Omega}_\epsilon}\sum_{k=1}^{N(n)}\left(U^n(t_{k-1})^2 - W_1(t_{k-1})^2\right)^2\left(V^n(t_k) - V^n(t_{k-1})\right)|^2 \\
&\quad +2|I_{\hat{\Omega}_\epsilon}\sum_{k=1}^{N(n)} 2\left(U^n(t_{k-1}) - W_1(t_{k-1})\right)U^n(t_{k-1})\left(V^n(t_k) - V^n(t_{k-1})\right)|^2
\end{aligned}$$

$$\leq \; 2\,N(n)\delta_n{}^4 \sum_{k=1}^{N(n)} (V^n(t_k) - V^n(t_{k-1}))^2 + 8\,N(n)\delta_n{}^2 \sum_{k=1}^{N(n)} U^n\,(t_{k-1})\,(V^n(t_k) - V^n(t_{k-1}))\,.$$

Hence

$$E|I_{\hat{\Omega}_\epsilon} \sum_{k=1}^{N(n)} \left(U^n\,(t_{k-1})^2 - W_1\,(t_{k-1})^2\right) (V^n\,(t_k) - V^n\,(t_{k-1}))\,|^2$$

$$\leq \; 2\,N(n)\delta_n{}^4 c \sum_{k=1}^{N(n)} (t_k - t_{k-1}) + 8\,N(n)\delta_n{}^2\sigma^2 c \sum_{k=1}^{N(n)} t_{k-1}(t_k - t_{k-1}) \;=\; o(1).$$

By the Chebyshev inequality, (30) is proved.

Finally we have

$$\begin{aligned}
& I_{\hat{\Omega}_\epsilon} \sum_{k=1}^{N(n)} W_1(t_{k-1})^2\,(V^n\,(t_k) - V^n\,(t_{k-1})) \\
&= \; -I_{\hat{\Omega}_\epsilon} \sum_{k=1}^{N(n)} V^n(t_k)\left(W_1(t_k)^2 - W_1(t_{k-1})^2\right) + I_{\hat{\Omega}_\epsilon} W_1(1)^2 V^n(1) \\
&= \; -I_{\hat{\Omega}_\epsilon} \sum_{k=1}^{N(n)} W_2(t_k)\left(W_1(t_k)^2 - W_1(t_{k-1})^2\right) + I_{\hat{\Omega}_\epsilon} W_1(1)^2 W_2(1) + o_p(1) \\
&= \; I_{\hat{\Omega}_\epsilon} \sum_{k=1}^{N(n)} W_1(t_{k-1})^2\,(W_2(t_k) - W_2(t_{k-1})) + o_p(1) \\
&= \; I_{\hat{\Omega}_\epsilon} \int_0^1 W_1(t)^2 dW_2(t) + o_p(1). \qquad (31)
\end{aligned}$$

Summation by parts gives the first and third identities, the second identity is shown by a similar method to (30) and the last identity is given by the fact that

$$\begin{aligned}
& E\left[\sum_{k=1}^{N(n)} \int_{t_{k-1}}^{t_k} \left\{W_1(t_{k-1})^2 - W_1(t)^2\right\} dW_2(t)\right]^2 \\
&= \; \sum_{k=1}^{N(n)} E\left[\int_{t_{k-1}}^{t_k} \left\{W_1(t_{k-1})^2 - W_1(t)^2\right\} dW_2(t)\right]^2 \\
&\leq \; \sum_{k=1}^{N(n)} 2E\left[\int_{t_{k-1}}^{t_k} \{W_1(t_{k-1}) - W_1(t)\}^2\, dW_2(t)\right]^2 \\
&\quad + \sum_{k=1}^{N(n)} 8E\left[W_1(t_{k-1}) \int_{t_{k-1}}^{t_k} \{W_1(t_{k-1}) - W_1(t)\}\, dW_2(t)\right]^2
\end{aligned}$$

$$\leq \quad 6c\sigma^4 \sum_{k=1}^{N(n)} \int_{t_{k-1}}^{t_k} (t - t_{k-1})^2 dt + 8c\sigma^4 \sum_{k=1}^{N(n)} t_{k-1} \int_{t_{k-1}}^{t_k} (t - t_{k-1}) dt \quad = \quad o(1) \quad .$$

The last inequality is given by Karatzas and Shreve (1991).

Combining (28), (30) and (31), we show (23) and therefore proved Lemma 2 (ii). ∎

C *Proof of Theorem 1*

First we have

$$n(\tilde{\rho} - 1) = \frac{\frac{1}{n}\sum_{k=1}^{n-1} X_k \varepsilon_{k+1} a_k a_{k+1}}{v\frac{1}{n^2}\sum_{k=1}^{n} X_k^2 + o_p(1)}. \tag{32}$$

Defining $U_n(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]} \varepsilon_i$ and $V_n(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]} \varepsilon_i a_{i-1} a_i$, we have that (U_n, V_n) converge in distribution (W_1, W_2) by Lemma 1. Hence the assertion (8) of Lemma 2 (ii) is satisfied. Then by Lemma 2 (i) and continuous mapping theorem, we have

$$\frac{1}{n}\sum_{k=1}^{n-1} X_k \varepsilon_{k+1} a_k a_{k+1} \xrightarrow{D} \int_0^1 W_1 dW_2,$$

and

$$\frac{1}{n^2}\sum_{k=1}^{n} X_k^2 \xrightarrow{D} \int_0^1 W_1^2 dt.$$

Orthogonalizing (W_1, W_2), we define B and $\tilde{B}$ by

$$\begin{aligned} B &= \frac{1}{\sqrt{\sigma_{11}}} W_1, \\ \tilde{B} &= \sqrt{\frac{\sigma_{11}}{\sigma_{11}\sigma_{22} - \sigma_{12}^2}} \left(W_2 - \frac{\sigma_{12}}{\sigma_{11}} W_1 \right). \end{aligned}$$

Then $(B, \tilde{B})$ is a two-dimensional standard Brownian motion. Consequently we have

$$n(\tilde{\rho} - 1) \xrightarrow{D} \frac{\sigma_{12}\int_0^1 B dB + \sqrt{\sigma_{11}\sigma_{22} - \sigma_{12}^2}\int_0^1 B d\tilde{B}}{v\sigma_{11}\int_0^1 B^2 dt}. \tag{33}$$

∎

D *Proof of Lemma 3*

Herrndorf (1984) defined the coefficients of strong mixing by

$$\alpha(i;\xi) = \sup_n \alpha_n(i;\xi), \qquad i = 1, 2, ...,$$

where

$$\alpha_n(i;\xi) = sup\{|P(A \cap B) - P(A)P(B)| : A \in \sigma(\xi_k, 1 \leq k \leq l), B \in \sigma(\xi_k, l + i \leq k \leq n), \\ 1 \leq l \leq n - i\}.$$

Similarly we define $\alpha(i;a)$ and $\alpha_n(i;a)$ by replacing $\{\xi_k\}$ by $\{a_k\}$ respectively. Let

$$\mathcal{A}_A = \{\text{all finite unions of disjoint rectangles } A_a \times A_e \\ \text{with } A_a \in \sigma(a_k, 1 \leq k \leq l) \text{ and } A_e \in \sigma(\varepsilon_k, 1 \leq k \leq l)\}, \tag{34}$$

and

$$\mathcal{F}_A = \sigma(\mathcal{A}_A) \quad (\text{the } \sigma - \text{field generated by } \mathcal{A}_A), \tag{35}$$

noting that $\mathcal{A}_A$ is a algebra. Similarly define $\mathcal{A}_B$ and $\mathcal{F}_B$ for sequences $\{a_k, l+i-1 \leq k \leq n\}$ and $\{\varepsilon_k, l + i \leq k \leq n\}$.

First we shall show

$$|P(A \cap B) - P(A)P(B)| \leq \alpha_n(i;a) \qquad \textit{for every } A \in \mathcal{A}_A,\ B \in \mathcal{A}_B. \tag{36}$$

For any $A \in \mathcal{A}_A$, we can choose $\{A_j = A_{aj} \times A_{ej};\ 1 \leq j \leq r_A,\ A_{aj} \in \sigma(a_k; 1 \leq k \leq l)$ and $\mathrm{A_{ej}} \in \sigma(\varepsilon_\mathrm{k}; 1 \leq \mathrm{k} \leq 1)\}$ such that

$$A = \bigcup_{j=1}^{r_A} A_j, \\ A_{ei} \cap A_{ej} = \phi \; for \quad i \neq j. \tag{37}$$

We can choose in the following way. By (34) we have

$$A = \bigcup_{j=1}^{K} A'_j, \quad where\ A'_j = A'_{aj} \times A'_{ej},$$

$$A'_i \cap A'_j = \phi\ for \quad i \neq j.$$

Then we have

$$\begin{aligned} A &= \bigcup_{p=1}^{K} \bigcup_{i} \left\{ \left(\bigcup_{h=i_1,\dots,i_p} A'_{ah} \right) \times \left[\left(\bigcap_{h=i_1,\dots,i_p} A'_{eh} \right) \cap \left(\bigcap_{h=i_{p+1},\dots,i_K} A'^c_{eh} \right) \right] \right\} \\ &= \bigcup_{j=1}^{r_A} A_j, \end{aligned}$$

where $\cup_i$ means the union of all of the sets satisfying

$$\begin{aligned} \{i_1, \dots, i_p\} \cup \{i_{p+1}, \dots, i_K\} &= \{1, \dots, K\}, \\ \{i_1, \dots, i_p\} \cap \{i_{p+1}, \dots, i_K\} &= \phi, \end{aligned}$$

and $r_A = 2^K - 1$ and A_j implies a product set

$$\left(\bigcup_{h=i_1,\dots,i_p} A'_{ah} \right) \times \left[\left(\bigcap_{h=i_1,\dots,i_p} A'_{eh} \right) \cap \left(\bigcap_{h=i_{p+1},\dots,i_K} A'^c_{eh} \right) \right].$$

Then A_j satisfies (37).

Similarly we can choose for any $B \in \mathcal{A}_B$, $\{B_j = B_{aj} \times B_{ej};\ 1 \leq j \leq r_B,\ B_{aj} \in \sigma(a_k; l+i-1 \leq k \leq n)$ and $B_{ej} \in \sigma(\varepsilon_k; l+i \leq k \leq n)\}$ such that

$$B = \bigcup_{j=1}^{r_B} B_j,$$

$$B_{ei} \cap B_{ej} = \phi\ for \quad i \neq j.$$

If we note that $\sum_{j=1}^{r_A} P(A_{ej}) = P(\cup_{j=1}^{r_A} A_{ej}) \leq 1$, $\sum_{j=1}^{r_B} P(B_{ej}) = P(\cup_{j=1}^{r_B} B_{ej}) \leq 1$ and $\{\varepsilon_k\}$ are independent, we have

$$|P(A \cap B) - P(A)P(B)| = |\sum_{j=1}^{r_A} \sum_{k=1}^{r_B} [P(A_j \cap B_k) - P(A_j)P(B_k)]|$$

$$
\begin{aligned}
&= \left|\sum_{j=1}^{r_A}\sum_{k=1}^{r_B}[P(A_{aj}\cap B_{ak}) - P(A_{aj})P(B_{ak})]P(A_{ej})P(B_{ek})\right| \\
&\le \sum_{j=1}^{r_A}\sum_{k=1}^{r_B}|P(A_{aj}\cap B_{ak}) - P(A_{aj})P(B_{ak})|P(A_{ej})P(B_{ek}) \\
&\le \alpha_n(i-1;a)\left[\sum_{j=1}^{r_A}P(A_{ej})\right]\left[\sum_{k=1}^{r_B}P(B_{ek})\right] \quad \le \quad \alpha_n(i-1;a).
\end{aligned}
$$

Hence (36) is proved.

Second we shall show

$$
|P(A\cap B) - P(A)P(B)| \le \alpha_n(i-1;a) \qquad \textit{for every } A \in \mathcal{F}_{\mathcal{A}},\ B \in \mathcal{F}_{\mathcal{B}}. \tag{38}
$$

By Chow and Teicher (1988) Theorem 1.5.3, for any $\epsilon > 0$ (fixed), for every $A \in \mathcal{F}_{\mathcal{A}}$, $B \in \mathcal{F}_{\mathcal{B}}$, there exist $A'_\epsilon \in \mathcal{A}_{\mathcal{A}}$, $B'_\epsilon \in \mathcal{A}_{\mathcal{B}}$, such that $P(A\Delta A'_\epsilon) < \frac{\epsilon}{4}$, $P(B\Delta B'_\epsilon) < \frac{\epsilon}{4}$, where Δ means symmetric difference. Then

$$
\begin{aligned}
|P(A\cap B) - P(A)P(B)| &\le |P(A'_\epsilon\cap B'_\epsilon) - P(A'_\epsilon)P(B'_\epsilon)| + 2P(A\Delta A'_\epsilon) + 2P(B\Delta B'_\epsilon) \\
&< |P(A'_\epsilon\cap B'_\epsilon) - P(A'_\epsilon)P(B'_\epsilon)| + \epsilon.
\end{aligned}
$$

By this inequality and (36), we show that for any $\epsilon > 0$, for every $A \in \mathcal{F}_{\mathcal{A}}$, $B \in \mathcal{F}_{\mathcal{B}}$,

$$
|P(A\cap B) - P(A)P(B)| < \alpha_n(i-1;a) + \epsilon. \tag{39}
$$

Hence (38) is proved.

Last, noting that $\sigma(\xi_k, 1 \le k \le l) \subset \mathcal{F}_{\mathcal{A}}$ and $\sigma(\xi_k, l+i \le k \le n) \subset \mathcal{F}_{\mathcal{B}}$, we have

$$
\alpha_n(i;\xi) \le sup\{|P(A\cap B) - P(A)P(B)|; A \in \mathcal{F}_{\mathcal{A}},\ B \in \mathcal{F}_{\mathcal{B}}, 1 \le l \le n-i\} \ \le\ \alpha_n(i-1;a). \tag{40}
$$

Hence

$$
\alpha(i;\xi) \quad \le \quad \alpha(i-1;a). \tag{41}
$$

∎

E Proof of Theorem 2

Let $\mathcal{F}_k = \sigma(a_1, \varepsilon_1, ..., a_k, \varepsilon_k)$ and $u_k = a_k - E(a_k|\mathcal{F}_{k-1}) = a_k - \{p_0 + (p_1 - p_0)a_{k-1}\}$. Then we have

$$n(\hat{\rho} - 1) = \frac{\frac{1}{n\sqrt{n}} \sum_{k=1}^{n-1} X_k^2 a_k u_{k+1} \frac{1}{n} \sum_{k=1}^{n} a_k}{\frac{1}{n^2} \sum_{k=1}^{n} a_k X_k^2 \frac{1}{n} \sum_{k=1}^{n-1} a_k a_{k+1}} - \frac{\frac{1}{\sqrt{n}} \sum_{k=1}^{n-1} a_k u_{k+1}}{\frac{1}{n} \sum_{k=1}^{n-1} a_k a_{k+1}}. \tag{42}$$

And

$$\frac{1}{n} \sum_{k=1}^{n} a_k \stackrel{a.s.}{\rightarrow} \pi \quad \mathit{and} \quad \frac{1}{n} \sum_{k=1}^{n-1} a_k a_{k+1} \stackrel{a.s.}{\rightarrow} \pi p_1, \tag{43}$$

since $\{a_k\}$ is a strong-mixing process. From

$$E\left[\frac{1}{n^2} \sum_{k=1}^{n} (a_k - \pi) X_k^2\right]^2 = o(1)$$

and Chebyshev's inequality, we have $\frac{1}{n^2} \sum_{k=1}^{n} a_k X_k^2 = \pi \frac{1}{n^2} \sum_{k=1}^{n} X_k^2 + o_p(1)$.

Let $\xi = \sqrt{E(a_{k-1}u_k)^2} = \sqrt{(1-p_1)p_1\pi}$ and $\zeta_k = \frac{a_{k-1}u_k}{\xi}$. Then $(\varepsilon_k, \zeta_k)'$ is a sequence of martingale differences with respect to $\mathcal{F}_k$. Let $U_n(t) = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{[nt]} \varepsilon_k$, $V_n(t) = \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} \zeta_k$. By Lemma 1 we have $(U_n, V_n) \stackrel{D}{\rightarrow} (W_1, W_2)$. Therefore we obtain by Lemma 2 (ii) and continuous mapping theorem

$$\begin{pmatrix} \frac{1}{n^2} \sum_{k=1}^{n} X_k^2 \\ \frac{1}{n\sqrt{n}} \sum_{k=1}^{n-1} X_k^2 a_k u_{k+1} \\ \frac{1}{\sqrt{n}} \sum_{k=1}^{n-1} a_k u_{k+1} \end{pmatrix} = \begin{pmatrix} \sigma^2 \int_0^1 U_n(t)^2 dt \\ \sigma^2 \xi \frac{1}{\sqrt{n}} \sum_{k=1}^{n-1} U_n(\frac{k}{n}) \zeta_{k+1} \\ \xi V_n(1) + o_p(1) \end{pmatrix} \stackrel{D}{\rightarrow} \begin{pmatrix} \sigma^2 \int_0^1 W_1^2 dt \\ \sigma^2 \xi \int_0^1 W_1^2 dW_2 \\ \xi \int_0^1 dW_2 \end{pmatrix}. \tag{44}$$

Substituting (43) and (44) into (42), we completed the proof. ∎

F *Proof of Theorem 3*

By definition we have

$$n(\hat{\rho}-1)=\frac{\frac{1}{n}\sum_{k=1}^{n-1}Y_kY_{k+1}\frac{1}{n}\sum_{k=1}^{n}a_k-\frac{1}{n}\sum_{k=1}^{n}Y_k^2\frac{1}{n}\sum_{k=1}^{n-1}a_ka_{k+1}}{\frac{1}{n^2}\sum_{k=1}^{n}Y_k^2\frac{1}{n}\sum_{k=1}^{n-1}a_ka_{k+1}}. \tag{45}$$

First, we shall give some relations used in this proof. Noting that $E\left[\frac{1}{n^2}\sum_{k=1}^{n}\left(a_k-\frac{L}{M}\right)X_k^2\right]^2=o(1)$, we obtain by Chebyshev inequality,

$$\frac{1}{n^2}\sum_{k=1}^{n}Y_k^2=\frac{L}{M}\frac{1}{n^2}\sum_{k=1}^{n}X_k^2+o_p(1). \tag{46}$$

In the same way, we have

$$\frac{1}{n^2}\sum_{k=1}^{n-1}Y_kY_{k+1}=\frac{K}{M}\frac{1}{n^2}\sum_{k=1}^{n-1}X_kX_{k+1}+o_p(1). \tag{47}$$

Moreover we have

$$\begin{aligned}\frac{1}{n^2}\sum_{k=1}^{n}X_k^2 &= \frac{1}{n^2}\sum_{k=1}^{n^*M}X_k^2+o_p(1)\\ &= \frac{1}{n^2}\sum_{i=0}^{n*-1}\sum_{m=1}^{M}\left(X_{iM}+\sum_{j=iM+1}^{iM+m}\varepsilon_j\right)^2+o_p(1)\\ &= \frac{M}{n^2}\sum_{i=1}^{n*-1}X_{iM}^2+o_p(1).\end{aligned} \tag{48}$$

Similarly

$$\frac{1}{n}X_n^2=\frac{1}{n}X_{n^*M}^2+o_p(1). \tag{49}$$

And we note that

$$\sum_{k=1}^{n}a_k=n^*L+\sum_{m=1}^{r+1}a_m, \tag{50}$$

and

$$\sum_{k=1}^{n-1}a_ka_{k-1}=n^*K+\sum_{m=1}^{r}a_ma_{m+1}. \tag{51}$$

Then from (46), (48) and (51), the denominator of (45) is

$$\left(\frac{L}{M}\frac{1}{n^2}\sum_{k=1}^{n}X_k^2+o_p(1)\right)\left[\frac{1}{n}\left(n^*K+\sum_{m=1}^{r}a_ma_{m+1}\right)\right]$$

$$=\frac{LK}{M}\frac{n^*}{n^3}\sum_{k=1}^{n}X_k^2+o_p(1)=LK\frac{n^*}{n^3}\sum_{i=1}^{n*-1}X_{iM}^2+o_p(1). \quad (52)$$

Next, by (50) and (51), the numerator of (45) is

$$\frac{1}{n}\sum_{k=1}^{n-1}Y_kY_{k+1}\frac{1}{n}\sum_{k=1}^{n}a_k-\frac{1}{n}\sum_{k=1}^{n}Y_k^2\frac{1}{n}\sum_{k=1}^{n-1}a_ka_{k+1}$$

$$= \frac{n^*}{n^2}\left[\sum_{k=1}^{n-1}\left(La_ka_{k+1}X_kX_{k+1}-Ka_kX_k^2\right)-Ka_nX_n^2\right]$$

$$+\frac{1}{n^2}\sum_{k=1}^{n-1}Y_kY_{k+1}\sum_{m=1}^{r+1}a_m-\frac{1}{n^2}\sum_{k=1}^{n-1}Y_k^2\sum_{m=1}^{r}a_ma_{m+1}$$

$$= \frac{n^*}{n^2}\left\{\sum_{k=1}^{n-1}\left[(La_ka_{k+1}-Ka_k)X_k^2+La_ka_{k+1}X_k\varepsilon_{k+1}\right]\right\}-\frac{n^*}{n^2}Ka_{r+1}X_n^2$$

$$+\frac{1}{n^2}\sum_{k=1}^{n}X_k^2\left(\frac{K}{M}\sum_{m=1}^{r+1}a_m-\frac{L}{M}\sum_{m=1}^{r}a_ma_{m+1}\right)+o_p(1). \quad (53)$$

This last equality is obtained by (46), (47) and $\frac{1}{n^2}X_n^2=o_p(1)$. If we put

$$Q_r=K\sum_{m=1}^{r+1}a_m-L\sum_{m=1}^{r}a_ma_{m+1}, \quad (54)$$

by (48) and (49), the right hand side of (53) is equal to

$$\frac{n^*}{n^2}\left\{\sum_{k=1}^{n-1}\left[(La_ka_{k+1}-Ka_k)X_k^2+La_ka_{k+1}X_k\varepsilon_{k+1}\right]\right\}$$

$$-\frac{n^*}{n^2}Ka_{r+1}X_{n^*M}^2+Q_r\frac{1}{n^2}\sum_{i=1}^{n*-1}X_{iM}^2+o_p(1). \quad (55)$$

Now we evaluate the first term of (55). First we have the following relations,

$$\sum_{m=1}^{M}[La_ma_{m+1}-Ka_m]X_{iM}^2=0, \quad (56)$$

$$\frac{n^*}{n^2}\sum_{i=0}^{n^*-1}\sum_{m=1}^{M}\left[(La_ma_{m+1}-Ka_m)(\sum_{j=iM+1}^{iM+m}\varepsilon_j)^2\right]$$

$$= \frac{n^*}{n^2} \sum_{i=0}^{n^*-1} \sum_{m=1}^{M} (La_m a_{m+1} - Ka_m) m\sigma^2 + o_p(1), \tag{57}$$

$$\frac{n^*}{n^2} \sum_{k=1}^{n-1} La_k a_{k+1} X_k \varepsilon_{k+1} = \frac{n^*}{n^2} \sum_{i=0}^{n^*-1} \sum_{m=1}^{M} La_m a_{m+1} X_{iM} \varepsilon_{iM+m} + o_p(1), \tag{58}$$

and

$$\frac{n^*}{n^2} \sum_{k=n^*M+1}^{n^*M+r} (La_k a_{k+1} - Ka_k) X_k^2 = \frac{n^*}{n^2} X_{n^*M}^2 \sum_{m=1}^{r} (La_m a_{m+1} - Ka_m) + o_p(1). \tag{59}$$

By (56), (57), (58) and (59), the first term of (55) is

$$\begin{aligned}
&\frac{n^*}{n^2} \left\{ \sum_{k=1}^{n-1} \left[(La_k a_{k+1} - Ka_k) X_k^2 + La_k a_{k+1} X_k \varepsilon_{k+1} \right] \right\} \\
&= \frac{n^*}{n^2} \sum_{i=0}^{n^*-1} \sum_{m=1}^{M} \left\{ (La_m a_{m+1} - Ka_m)[2X_{iM} \sum_{j=iM+1}^{iM+m} \varepsilon_j + (\sum_{j=iM+1}^{iM+m} \varepsilon_j)^2] \right\} \\
&\quad + \frac{n^*}{n^2} \sum_{k=n^*M+1}^{n^*M+r} (La_m a_{m+1} - Ka_m) X_k^2 + \frac{n^*}{n^2} \sum_{k=1}^{n-1} La_k a_{k+1} X_k \varepsilon_{k+1} \\
&= \frac{n^*}{n^2} \sum_{i=0}^{n^*-1} \sum_{m=1}^{M} \left[2X_{iM} (La_m a_{m+1} - Ka_m) \sum_{j=iM+1}^{iM+m} \varepsilon_j + La_{m-1} a_m X_{iM} \varepsilon_{iM+m} \right] \\
&\quad + \frac{n^*}{n^2} \sum_{i=0}^{n^*-1} \sum_{m=1}^{M} (La_m a_{m+1} - Ka_m) m\sigma^2 + \frac{n^*}{n^2} X_{n^*M}^2 \sum_{m=1}^{r} (La_m a_{m+1} - Ka_m) + o_p(1).
\end{aligned} \tag{60}$$

If we put

$$v_{i+1} = \sum_{m=1}^{M} [2(La_m a_{m+1} - Ka_m) \sum_{j=iM+1}^{iM+m} \varepsilon_j + La_m a_{m-1} \varepsilon_{iM+m}] \tag{61}$$

and

$$R = \frac{1}{M} \sum_{m=1}^{M} (La_m a_{m+1} - Ka_m) m, \tag{62}$$

(60) is equal to

$$\frac{n^*}{n^2} \sum_{i=0}^{n^*-1} X_{iM} v_{i+1} + \frac{n^{*2}}{n^2} R\sigma^2 M + \frac{n^*}{n^2} X_{n^*M}^2 \sum_{m=1}^{r} (La_m a_{m+1} - Ka_m) + o_p(1). \tag{63}$$

Now, substituting (63) into (55), we can rewrite the numerator of (45) as

$$
\begin{aligned}
&\frac{n^*}{n^2}\sum_{i=0}^{n^*-1} X_{iM}\, v_{i+1} + \frac{n^{*2}}{n^2}R\sigma^2 M + Q_r\left[\frac{1}{n^2}\sum_{i=1}^{n^*-1} X_{iM}^2 - \frac{n^*}{n^2}X_{n^*M}^2\right] + o_p(1) \\
&= \frac{n^{*2}}{n^2}\left[\frac{1}{n^*}\sum_{i=0}^{n^*-1} X_{iM}\, v_{i+1} + R\sigma^2 M + Q_r\left(\frac{1}{n^{*2}}\sum_{i=1}^{n^*-1} X_{iM}^2 - \frac{X_{n^*M}^2}{n^*}\right)\right] + o_p(1). \quad (64)
\end{aligned}
$$

Then substituting (52) and (64) into (45), we have

$$
n(\hat{\rho}-1) = \frac{\frac{1}{n^*}\sum_{i=0}^{n^*-1} X_{iM}\, v_{i+1} + R\sigma^2 M + Q_r\left(\frac{1}{n^{*2}}\sum_{i=1}^{n^*-1} X_{iM}^2 - \frac{1}{n^*}X_{n^*M}^2\right) + o_p(1)}{\frac{n^*}{n}LK\frac{1}{n^{*2}}\sum_{i=1}^{n^*-1} X_{iM}^2 + o_p(1)}. \quad (65)
$$

Finally we set that $u_i = \sum_{m=1}^{M}\varepsilon_{(i-1)M+m}$ and $\mathcal{F}_i = \sigma(\varepsilon_1,...,\varepsilon_M,...,\varepsilon_{iM})$. Then $(u_i, v_i)'$ is a sequence of martingale differences with respect to $\mathcal{F}_i$. Defining $U_{n^*}(t) = \frac{1}{\sqrt{n^*M\sigma^2}}\sum_{i=1}^{[n^*t]} u_i$ and $V_{n^*}(t) = \frac{1}{\sqrt{n^*M\sigma^2}}\sum_{i=1}^{[n^*t]} v_i$, we show that (U_{n^*}, V_{n^*}) converge in distribution (W_1, W_2) by Lemma 1, where the (i,j)-components of the covariance matrix of (W_1, W_2) are $\sigma_{11} = 1$, $\sigma_{22} = E\left[\sum_{m=1}^{M} v_m\right]^2 =$ and $\sigma_{12} = \sigma_{21} = E\left[\sum_{m=1}^{M} u_m \sum_{m=1}^{M} v_m\right] = 2\,R + \frac{LK}{M}$. Then using Lemma 2 (i) and continuous mapping theorem, we can see that

$$
\begin{pmatrix} \frac{1}{n^{*2}}\sum_{i=1}^{n^*-1} X_{iM}^2 \\ \frac{1}{n^*}\sum_{i=1}^{n^*-1} X_{iM}\, v_{i+1} \\ \frac{1}{n^*}X_{n^*M}^2 \end{pmatrix} \xrightarrow{D} M\sigma^2 \begin{pmatrix} \int_0^1 W_1^2 dt \\ \int_0^1 W_1^2 dW_2 \\ W_1(1)^2 \end{pmatrix} \quad as\ n^* \to \infty. \quad (66)
$$

From (65) and (66), the proof is completed. ∎