Econometric Theory, **18**, 2002, 119–139. Printed in the United States of America.

THE INVARIANCE PRINCIPLE FOR LINEAR PROCESSES WITH APPLICATIONS

QIYING WANG, YAN-XIA LIN, AND CHANDRA M. GULATI
University of Wollongong

Let X_t be a linear process defined by $X_t = \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k}$, where $\{\psi_k, k \geq 0\}$ is a sequence of real numbers and $\{\epsilon_k, k = 0, \pm 1, \pm 2, \ldots\}$ is a sequence of random variables. Two basic results, on the invariance principle of the partial sum process of the X_t converging to a standard Wiener process on $[0,1]$, are presented in this paper. In the first result, we assume that the innovations ϵ_k are independent and identically distributed random variables but do not restrict $\sum_{k=0}^{\infty} |\psi_k| < \infty$. We note that, for the partial sum process of the X_t converging to a standard Wiener process, the condition $\sum_{k=0}^{\infty} |\psi_k| < \infty$ or stronger conditions are commonly used in previous research. The second result is for the situation where the innovations ϵ_k form a martingale difference sequence. For this result, the commonly used assumption of equal variance of the innovations ϵ_k is weakened. We apply these general results to unit root testing. It turns out that the limit distributions of the Dickey–Fuller test statistic and Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) test statistic still hold for the more general models under very weak conditions.

1. INTRODUCTION

Let $\{X_t, t \geq 1\}$ be a sequence of random variables such that $EX_t = 0$. Let

$$S_n = \sum_{t=1}^{n} X_t \quad \text{and} \quad \sigma_n^2 = \mathrm{Var}(S_n).$$

We denote by $\Rightarrow$ the weak convergence of probability measures in $D[0,1]$, where $D[0,1]$ is the space of all right continuous real-valued functions having finite left limits on $[0,1]$ endowed with the sup norm. Under appropriate conditions, it is well known that

$$\frac{S_{[nt]}}{\sigma_n} \Rightarrow W(t), \quad 0 \leq t \leq 1, \tag{1}$$

where $W(t)$ is a standard Wiener process on $[0,1]$ and $[nt]$ denotes the integer part of the nt. The result of form (1) is commonly called the invariance princi-

The authors thank a referee, Professor Tanaka, Professor Phillips, and Professor Rayner for their valuable comments, which have led to this much improved version of the paper. Address correspondence to: Qiying Wang, School of Mathematics and Statistics, University of Wollongong, NSW, 2522, Australia; e-mail: qiying@uow.edu.au.

© 2002 Cambridge University Press 0266-4666/02 $9.50

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

ple or the functional limit theorem. It is quite useful in characterizing the limit distribution of various statistics arising from the inference in economic time series. To elaborate, let us consider a stochastic process generated according to

$$y_t = \alpha y_{t-1} + X_t, \qquad t = 1,2,\ldots, \tag{2}$$

where y_0 is a constant with probability one or has a certain specified distribution. Denote the ordinary least squares (OLS) estimator of α by $\hat{\alpha}_n = \sum_{t=1}^n y_t y_{t-1} / \sum_{t=1}^n y_{t-1}^2$. To test $\alpha = 1$ against $\alpha < 1$, a key step is to derive the limit distribution of the well-known DF (Dickey–Fuller) test statistic (Dickey and Fuller, 1979):

$$n(\hat{\alpha}_n - 1) = \left\{ n^{-1} \sum_{t=1}^n y_{t-1}(y_t - y_{t-1}) \right\} \Big/ \left\{ n^{-2} \sum_{t=1}^n y_{t-1}^2 \right\}. \tag{3}$$

As shown by Phillips (1987), in null hypothesis $\alpha = 1$, the asymptotic properties of the DF test statistic relied heavily on the invariance principle of the form (1).

In past decades, under different assumptions on X_t, there are many articles that discuss the invariance principle of the form (1). Here, we cite two basic textbooks, Billingsley (1968) and Hall and Heyde (1980), for the collections of related articles for independent random variables and martingale difference sequences; the review paper for mixing sequence given by Peligrad (1986); and also Peligrad's recent work (Peligrad, 1998). For more general mixingale sequences, we refer to Mcleish (1975, 1977) and Truong-van (1995).

In this paper, we restrict our attention to linear processes, an important case in economic time series. In what follows, we always assume that

$$X_t = \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k}, \tag{4}$$

where $\{\psi_k, k \geq 0\}$ is a sequence of real numbers and innovations $\epsilon_k, k = 0, \pm 1, \pm 2, \ldots$, are random variables specialized later.

On the invariance principle of the form (1) for linear processes, this paper establishes two basic results. In the first result, we assume that the innovations ϵ_k are independent and identically distributed (i.i.d.) random variables, but the condition $\sum_{k=0}^{\infty} |\psi_k| < \infty$ (or stronger conditions), commonly used in previous research given by Hannan (1979), Stadtmüller and Trautner (1985), and Phillips and Solo (1992) (also see Tanaka, 1996) and also by Yokoyama (1995), is weakened. Only finite second moments for ϵ_k are required in this paper. It gives an essential improvement of the previous similar results given by Davydov (1970). The second result is for the situation where the innovations ϵ_k form a martingale difference sequence. In this result, the commonly used assumption, the innovations ϵ_k having the same variance, is weakened. This will be of interest to researchers from the viewpoint of practice.

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

We give the statements of main theorems and detailed remarks on the previous results in the next section. In Section 3, the applications to the Dickey–Fuller test statistic and Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) test statistic are discussed. We find these important statistics still have similar limit distributions for the more general models under quite weak conditions. In Section 4, some general conclusions are drawn. Finally in Section 5, we give the proofs of the main theorems.

2. MAIN RESULTS AND REMARKS

For brevity, we denote $\lim_{n\to\infty} a_n/b_n \to 1$ by $a_n \sim b_n$, and A, with or without subscript, is for positive constant.

In Theorem 2.1, which follows, we assume that the innovations ϵ_k are i.i.d. random variables but, to cover some interesting cases, the ψ_k are rather general. Write, for $j = 1,2,3,\ldots,$

$$v_j = \sum_{k=0}^{j-1} \psi_k \quad \text{and} \quad s_n^2 = \sum_{j=1}^{n} v_j^2.$$

THEOREM 2.1. *Let $\epsilon_k, k = 0,\pm1,\pm2,\ldots,$ be i.i.d. random variables with $E\epsilon_0 = 0$ and $E\epsilon_0^2 = 1$. Assume that $\psi_0 \neq 0$,*

$$\frac{1}{s_n} \max_{1\le j\le n} |v_j| \to 0 \quad \textit{and} \quad \sum_{j=0}^{n} \left(\sum_{k=j}^{\infty} \psi_k^2 \right)^{1/2} = o(s_n). \tag{5}$$

Under these assumptions, we have that

$$\frac{1}{s_n} \sum_{j=1}^{k_n(t)} X_j \Rightarrow W(t), \qquad 0 \le t \le 1, \tag{6}$$

where $k_n(t) = \sup\{m : s_m^2 \le t s_n^2\}$.

In particular, if $\psi_k = k^{-1}l(k)$, where $l(0)/0 \equiv 1$ and the positive function $l(k)$ is slowly varying at infinity satisfying $\sum_{k=1}^{\infty} k^{-1}l(k) = \infty$, then

$$\frac{1}{\sqrt{n}\sum_{k=1}^{n} k^{-1}l(k)} \sum_{j=1}^{k_n(t)} X_j \Rightarrow W(t), \qquad 0 \le t \le 1, \tag{7}$$

where $k_n(t)$ is defined as in (6).

If $0 < |\sum_{k=0}^{\infty} \psi_k| < \infty$ and $\sum_{k=1}^{\infty} k\psi_k^2 < \infty$ or $\sum_{k=0}^{\infty} |\psi_k| < \infty$ and $\sum_{k=0}^{\infty} \psi_k \neq 0$, then

$$\frac{1}{\sigma_n} \sum_{j=1}^{[nt]} X_j \Rightarrow W(t), \qquad 0 \le t \le 1, \tag{8}$$

where $\sigma_n^2 = n(\sum_{k=0}^{\infty} \psi_k)^2$.

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

Remark 2.1. It follows from Hall (1992, p. 118) that

$$\sigma_n^2 = \mathrm{Var}\left(\sum_{j=1}^{n} X_j\right) \sim n\left(\sum_{k=1}^{n} k^{-1}l(k)\right)^2,$$

provided $\psi_k = k^{-1}l(k)$ and $\sum_{k=1}^{\infty} k^{-1}l(k) = \infty$. Hence, we can replace $\sqrt{n}\sum_{k=1}^{n} k^{-1}l(k)$ by σ_n in (7). It is unclear whether or not s_n in (6) can be replaced by σ_n.

Remark 2.2. Let $\psi_k = k^{-\alpha}$, where $\frac{1}{2} < \alpha < 1$. It is easy to show that the second condition of (5) fails to hold. In this case, we also know that $(1/\sigma_n)\sum_{j=1}^{k_n(t)} X_j$ fails to converge to $W(t)$. In fact, by applying Liu (1998) (also see Marinucci and Robinson, 1998), $(1/\sigma_n)\sum_{j=1}^{[nt]} X_j$ converges to a fractional Brownian motion with $d = 1 - \alpha$. Therefore, to make the partial sum process of the X_j converge to a standard Wiener process, the condition (5) is close to the necessary condition.

Remark 2.3. The conditions given in this theorem are different from those given by Davydov (1970). Specifically, Theorem 2.1 abolishes the condition $E\epsilon_0^4 < \infty$, which is an essential improvement of Davydov's result for the moment condition.

In the next theorem, the i.i.d. assumption for the innovations ϵ_k is weakened to being a martingale difference sequence. In this case, an excellent result is given by Hannan (1979), where it is required that $E\epsilon_k^2 = \sigma^2$ and $\lim_{n\to\infty} E(\epsilon_k^2|\mathcal{F}_{k-n}) = \sigma^2$, a.s. ($\mathcal{F}_k$ is defined as in Theorem 2.2, which follows) for all k. In Theorem 2.2, these conditions are moderated. Our Corollary 2.1, which follows Theorem 2.2, also improves Theorem 3.15 in Phillips and Solo (1992), where the authors assumed $\sum_{k=0}^{\infty} k|\psi_k| < \infty$ and $\{\epsilon_k\}$ is a s.u.i. (strongly uniformly integrable; the definition can be found in Billingsley, 1968, p. 32) martingale difference sequence.

THEOREM 2.2. *Let ψ_k satisfy*

$$b_0 \equiv \sum_{k=0}^{\infty} \psi_k \neq 0, \qquad \sum_{k=0}^{\infty} |\psi_k| < \infty, \qquad \textit{and} \quad \sum_{k=1}^{\infty} k\psi_k^2 < \infty.$$

Let ϵ_k be random variables such that

$$E(\epsilon_k|\mathcal{F}_{k-1}) = 0, \quad a.s.\ k = 0, \pm 1, \pm 2, \ldots,$$

where $\mathcal{F}_k$ is the σ-field generated by $\{\epsilon_j, j \le k\}$. If

$$\sup_{n\ge 1} \frac{1}{n}\sum_{k=-n}^{n} E\epsilon_k^2 < \infty, \qquad \inf_{n\ge 1} \frac{1}{n}\sum_{k=1}^{n} E\epsilon_k^2 > 0; \tag{9}$$

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

and as $n \to \infty$,

$$\frac{1}{n}\sum_{k=1}^{n}(\epsilon_k^2 - E\epsilon_k^2) \to 0, \quad \textit{in probability}; \tag{10}$$

$$\frac{1}{n}\sum_{k=-n}^{n} E\epsilon_k^2 I_{(|\epsilon_k| \geq \delta\sqrt{n})} \to 0, \quad \textit{for any } \delta > 0, \tag{11}$$

then

$$\frac{1}{\sigma_n^*}\sum_{j=1}^{k_n^*(t)} X_j \Rightarrow W(t), \qquad 0 \leq t \leq 1, \tag{12}$$

where $\sigma_n^{*2} = b_0^2 \sum_{k=1}^{n} E\epsilon_k^2$ *and* $k_n^*(t) = \sup\{j : \sum_{k=1}^{j} E\epsilon_k^2 \leq t \sum_{k=1}^{n} E\epsilon_k^2\}$.

From Theorem 2.2, we obtain the following corollary.

COROLLARY 2.1. *If conditions (9)–(11) in Theorem 2.2 are replaced by one of the following conditions (a)–(c), then (12) still holds.*

(a) $\{\epsilon_k^2\}$ *is uniformly integrable and* $E(\epsilon_k^2|\mathcal{F}_{k-1}) = \sigma^2 > 0$ *for all* $k \geq 1$*;*
(b) $\{\epsilon_k^2\}$ *is s.u.i. and* $(1/n)\sum_{k=1}^{n} E(\epsilon_k^2|\mathcal{F}_{k-1}) \to \sigma^2 > 0$*, in probability;*
(c) $E(\sup_k \epsilon_k^2) < \infty$ *and* $(1/n)\sum_{k=1}^{n} E(\epsilon_k^2|\mathcal{F}_{k-1}) \to \sigma^2 > 0$*, in probability.*

Proof. If condition (a) holds, then $E\epsilon_k^2 = E(\epsilon_k^2|\mathcal{F}_{k-1}) = \sigma^2$. Condition (12) follows immediately from Theorem 2.2 by using Lemma 5.5 (from Section 5).

If condition (b) holds, it follows from Lemma 5.5 that

$$\frac{1}{n}\sum_{k=-n}^{n} E\epsilon_k^2 I_{(|\epsilon_k| \geq \delta\sqrt{n})} \to 0 \quad \text{and} \quad \frac{1}{n}\sum_{k=1}^{n}\epsilon_k^2 \to \sigma^2, \quad \text{in probability.} \tag{13}$$

On the other hand, it is known that $\{(1/n)\sum_{k=1}^{n}\epsilon_k^2\}$ is s.u.i. if $\{\epsilon_k^2\}$ is s.u.i. (Chow and Teicher, 1988, p. 102). This fact, together with the second relation of (13), implies that (Chow and Teicher, 1988, p. 100)

$$\frac{1}{n}\sum_{k=1}^{n} E\epsilon_k^2 \to \sigma^2. \tag{14}$$

In terms of (13) and (14), it is easy to check that all conditions in Theorem 2.2 are satisfied and hence (12) holds.

Finally, if condition (c) holds, (12) follows obviously because $E(\sup_k \epsilon_k^2) < \infty$ implies that $\{\epsilon_k^2\}$ is s.u.i. ■

3. APPLICATIONS

In this section, we discuss the applications of this paper to time series. At first, we assume that the process $\{y_t\}$ is generated by (2) with $\alpha = 1$. Phillips (1987) investigated the limit behavor of the DF test statistic $n(\hat{\alpha}_n - 1)$ de-

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

fined by (3) provided $\{X_t\}$ is a strong mixing sequence with appropriate mixing conditions. Here, we assume that $\{X_t\}$ satisfies (4), i.e., $\{X_t\}$ forms a linear process. Under quite general conditions for ψ_k and ϵ_k in (4), it is shown that the DF test statistic $n(\hat{\alpha}_n - 1)$ has a similar distribution as that in Phillips and Xiao (1998), where the authors obtained the limit distribution of $n(\hat{\alpha}_n - 1)$ provided $\sum_{k=0}^{\infty} k^{1/2}|\psi_k| < \infty$.

THEOREM 3.1. *Let* $\epsilon_k, k = 0, \pm 1, \pm 2, \ldots,$ *be i.i.d. random variables with* $E\epsilon_0 = 0$ *and* $E\epsilon_0^2 = \sigma^2$. *If* $0 < |\sum_{k=0}^{\infty} \psi_k| < \infty$ *and* $\sum_{k=1}^{\infty} k\psi_k^2 < \infty$ *or* $\sum_{k=0}^{\infty}|\psi_k| < \infty$ *and* $\sum_{k=0}^{\infty} \psi_k \neq 0$, *then as* $n \to \infty$,

(a) $(1/n^2)\sum_{t=1}^{n} y_{t-1}^2 \Rightarrow \sigma^2 b_0^2 \int_0^1 W(r)^2 dr$;
(b) $(1/n)\sum_{t=1}^{n} y_{t-1}(y_t - y_{t-1}) \Rightarrow (\sigma^2 b_0^2/2)(W(1)^2 - \gamma)$;
(c) $n(\hat{\alpha}_n - 1) \Rightarrow (\frac{1}{2})(W(1)^2 - \gamma)/\int_0^1 W(r)^2 dr$;
(d) $\hat{\alpha}_n \to 1$, *in probability;*
(e) $t_\alpha \Rightarrow (\frac{1}{2}\gamma^{-1/2})(W(1)^2 - \gamma)/\{\int_0^1 W(r)^2 dr\}^{1/2}$,

where

$$b_0 = \sum_{k=0}^{\infty} \psi_k, \qquad \gamma = \sum_{k=0}^{\infty} \psi_k^2 / b_0^2,$$

$$\hat{\alpha}_n = \sum_{t=1}^{n} y_t y_{t-1} \Big/ \sum_{t=1}^{n} y_{t-1}^2, \qquad \delta_n^2 = \frac{1}{n}\sum_{t=1}^{n} (y_t - \hat{\alpha}_n y_{t-1})^2, \qquad \textit{and}$$

$$t_\alpha = \left(\sum_{t=1}^{n} y_{t-1}^2\right)^{1/2} (\hat{\alpha}_n - 1)/\delta_n.$$

As in Phillips (1987), the proof of Theorem 3.1 may be obtained by applying Theorem 2.1. The details are omitted.

The limit distribution given in Theorem 3.1 depends on the unknown parameter

$$\gamma = \sum_{k=0}^{\infty} \psi_k^2 \Big/ \left(\sum_{k=0}^{\infty} \psi_k\right)^2.$$

As in Phillips (1987, p. 285), we can construct an estimate of γ as follows:

$$\hat{\gamma} = \hat{\sigma}_n^{*2}/\hat{\sigma}_n^2, \quad \text{where } \hat{\sigma}_n^{*2} = \frac{1}{n}\sum_{t=1}^{n} X_t^2$$

and $\hat{\sigma}_n^2 = (1/n)\sum_{t=1}^{n} X_t^2 + (2/n)\sum_{r=1}^{l_n}\sum_{t=r+1}^{n} X_t X_{t-r}$. Here and subsequently, $\{l_n, n \geq 1\}$ denotes a sequence of positive real numbers. The following theorem shows that $\hat{\gamma}$ is a consistent estimate of γ for any l_n satisfying $l_n = o(n)$ and $l_n \to \infty$.

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

THEOREM 3.2. *Let* $\epsilon_k, k = 0, \pm 1, \pm 2, \ldots$, *be i.i.d. random variables with* $E\epsilon_0 = 0$ *and* $E\epsilon_0^2 = \sigma^2$.

(a) If $\sum_{k=0}^{\infty} \psi_k^2 < \infty$, *then* $\hat{\sigma}_n^{*2}/\sigma^2 \to \sum_{k=0}^{\infty} \psi_k^2$, *a.s.*
(b) If $\sum_{k=0}^{\infty} |\psi_k| < \infty$, *then for any* l_n *satisfying* $l_n = o(n)$ *and* $l_n \to \infty$, $\hat{\sigma}_n^2/\sigma^2 \to (\sum_{k=0}^{\infty} \psi_k)^2$, *in probability.*

Proof. Noting that $\{X_t, t \geq 1\}$ is a stationary ergodic sequence and $EX_1^2 = \sigma^2 \sum_{k=0}^{\infty} \psi_k^2 < \infty$, it follows from the stationary ergodic theorem (Stout, 1974, p. 181) that $\hat{\sigma}_n^{*2} \to \sigma^2 \sum_{k=0}^{\infty} \psi_k^2$, a.s. This proves part (a).

It is well known that (Brockwell and Davis, 1987, p. 212)

$$\frac{1}{n} E\left(\sum_{t=1}^{n} X_t\right)^2 = \frac{1}{n} \sum_{t=1}^{n} EX_t^2 + \frac{2}{n} \sum_{r=1}^{n-1} \sum_{t=r+1}^{n} EX_t X_{t-r} \to \sigma^2 \left(\sum_{k=0}^{\infty} \psi_k\right)^2. \tag{15}$$

This fact, together with part (a), implies that, to prove $\hat{\sigma}_n^2/\sigma^2 \to (\sum_{k=0}^{\infty} \psi_k)^2$ in probability, it suffices to show that

$$\frac{1}{n} \sum_{r=1}^{l_n} \sum_{t=r+1}^{n} (X_t X_{t-r} - EX_t X_{t-r}) \to 0, \quad \text{in probability;} \tag{16}$$

$$\frac{1}{n} \sum_{r=l_n+1}^{n-1} \sum_{t=r+1}^{n} EX_t X_{t-r} \to 0. \tag{17}$$

The proofs of (16) and (17) appear in the Appendix. ■

If $\sum_{k=0}^{\infty} \psi_k = \infty$, the results differ from those in Theorem 3.1. In this case, we find that the limit distribution of the DF test statistic $n(\hat{\alpha}_n - 1)$ is free from the unknown parameters but t_α diverges to ∞ in probability. Explicitly, we obtain the following theorem.

THEOREM 3.3. *Let* $\epsilon_k, k = 0, \pm 1, \pm 2, \ldots$, *be i.i.d. random variables with* $E\epsilon_0 = 0$ *and* $E\epsilon_0^2 = \sigma^2$. *If* $\psi_k = k^{-1} l(k)$, *where* $l(0)/0 \equiv 1$ *and positive function* $l(k)$ *is slowly varying at infinity satisfying* $\sum_{k=1}^{\infty} k^{-1} l(k) = \infty$, *then*

(a) $(nv_n)^{-2} \sum_{t=1}^{n} y_{t-1}^2 \Rightarrow \sigma^2 \int_0^1 W(r)^2 dr$;
(b) $(\sqrt{n} v_n)^{-2} \sum_{t=1}^{n} y_{t-1}(y_t - y_{t-1}) \Rightarrow (\sigma^2/2) W(1)^2$;
(c) $n(\hat{\alpha}_n - 1) \Rightarrow (\frac{1}{2}) W(1)^2 / \int_0^1 W(r)^2 dr$;
(d) $t_\alpha \to \infty$, *in probability,*

where $v_n = \sum_{k=1}^{n} k^{-1} l(k)$, $\hat{\alpha}_n$, *and* t_α *are defined as in Theorem 3.1.*

Proof. Recall

$$s_n^2 = \sum_{j=1}^{n} v_j^2, \qquad k_n(t) = \sup\{m : s_m^2 \leq t s_n^2\}$$

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

and the process $\{y_t\}$ is defined by (2). If $\alpha = 1$, then $y_t = \sum_{j=1}^{t} X_j$ (without loss of generality, here and subsequently, we assume $y_0 = 0$), and hence

$$y_{t-1}^2 = \frac{s_n^2}{v_t^2} \int_{s_{t-1}^2/s_n^2}^{s_t^2/s_n^2} \left(\sum_{j=1}^{k_n(r)} X_j \right)^2 dr.$$

Therefore, we obtain that

$$\begin{aligned} \sum_{t=1}^{n} y_{t-1}^2 &= \sum_{t=1}^{n} \frac{v_t^2}{v_n^2} y_{t-1}^2 + \sum_{t=1}^{n} \left(\frac{1}{v_t^2} - \frac{1}{v_n^2} \right) v_t^2 y_{t-1}^2 \\ &= \frac{s_n^2}{v_n^2} \sum_{t=1}^{n} \int_{s_{t-1}^2/s_n^2}^{s_t^2/s_n^2} \left(\sum_{j=1}^{k_n(r)} X_j \right)^2 dr + R_n, \quad \text{say,} \\ &= \frac{s_n^4}{v_n^2} \int_0^1 \left(\frac{1}{s_n} \sum_{j=1}^{k_n(r)} X_j \right)^2 dr + R_n. \end{aligned}$$

It follows from Theorem 2.1 and the continuous mapping theorem (see Billingsley, 1968, Sect. 5) that

$$\int_0^1 \left(\frac{1}{s_n} \sum_{j=1}^{k_n(r)} X_j \right)^2 dr \Rightarrow \sigma^2 \int_0^1 W(r)^2 dr.$$

This fact, together with Theorem 1.4.1 given by Billingsley (1968, p. 25), implies that part (a) follows if

$$s_n^2 = \sum_{j=1}^{n} v_j^2 \sim n v_n^2 \quad \text{and} \quad \frac{1}{n^2 v_n^2} R_n \to 0, \quad \text{in probability.} \tag{18}$$

Because $v_n = \sum_{k=1}^{n} k^{-1} l(k)$ is still a slowly varying function, the first relation of (18) follows from Bingham, Goldie, and Teugels (1987, p. 26).

By noting that $Ey_t^2 \sim t v_t^2$ (recalling Remark 2.1), we have that

$$\frac{1}{n^2 v_n^2} \sum_{t=1}^{n} Ey_{t-1}^2 \sim \frac{1}{2} \quad \text{and} \quad \frac{1}{n^2 v_n^4} \sum_{t=1}^{n} v_t^2 Ey_{t-1}^2 \sim \frac{1}{2}$$

by using the slowly varying properties of v_t. Hence, by noting $v_t \uparrow$, as $t \uparrow \infty$, it follows that

$$\frac{1}{n^2 v_n^2} E|R_n| = \frac{1}{n^2 v_n^2} \sum_{t=1}^{n} Ey_{t-1}^2 - \frac{1}{n^2 v_n^4} \sum_{t=1}^{n} v_t^2 Ey_{t-1}^2 \to 0.$$

The second relation of (18) follows from Markov's inequality. The proof of part (a) is complete.

The proof of part (b) follows directly from Theorem 2.1 and part (a) of Theorem 3.2 by noting $\sum_{t=1}^{n} y_{t-1}(y_t - y_{t-1}) = \frac{1}{2} y_n^2 - \sum_{t=1}^{n} X_t^2$ (see Phillips, 1987, Appendix).

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

The proof of part (c) is straightforward by applying parts (a) and (b) and the continuous mapping theorem.

To prove part (d), we rewrite

$$\delta_n^2 = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{\alpha}_n y_{t-1})^2$$

$$= \frac{1}{n}\sum_{t=1}^{n} X_t^2 - \frac{2(\hat{\alpha}_n - 1)}{n}\sum_{t=1}^{n} y_{t-1}X_t + \frac{(\hat{\alpha}_n - 1)^2}{n}\sum_{t=1}^{n} y_{t-1}^2.$$

Because $n^{-\delta}v_n \to 0$, for any $\delta > 0$ (see Feller, 1971, p. 277), and $v_n \to \infty$, it follows from parts (a)–(c) that for $\forall\, \epsilon > 0$, as $n \to \infty$,

$$P\left(\left|\frac{\hat{\alpha}_n - 1}{n}\sum_{t=1}^{n} y_{t-1}X_t\right| \geq \epsilon\right) \leq P(n|\hat{\alpha}_n - 1| \geq \epsilon v_n) + P\left(\frac{1}{nv_n^2}\left|\sum_{t=1}^{n} y_{t-1}X_t\right| \geq nv_n^{-3}\right) \to 0, \qquad \textbf{(19)}$$

$$P\left(\frac{(\hat{\alpha}_n - 1)^2}{n}\sum_{t=1}^{n} y_{t-1}^2 \geq \epsilon\right) \leq P(n|\hat{\alpha}_n - 1| \geq \epsilon v_n) + P\left(\frac{1}{n^2 v_n^2}\sum_{t=1}^{n} y_{t-1}^2 \geq nv_n^{-4}\right) \to 0. \qquad \textbf{(20)}$$

In terms of (19), (20), and part (a) of Theorem 3.2, we have that

$$\delta_n^2 \to \sigma^2 \sum_{k=0}^{\infty} \psi_k^2 < \infty, \quad \text{in probability.}$$

Therefore, part (d) follows easily by applying parts (a) and (c). The proof of Theorem 3.3 is complete. ■

Next we discuss another application of the present results. Let us consider the model

$$y_t = \psi + r_t + z_t, \qquad t = 1,2,\ldots,n. \qquad \textbf{(21)}$$

Here ψ is a constant, z_t is a stationary error, and r_t is a random walk:

$$r_t = r_{t-1} + u_t \quad \text{with } r_0 = 0, \qquad \textbf{(22)}$$

where the u_t are i.i.d. random variables with $Eu_t = 0$ and $Eu_t^2 = \sigma_u^2$. To test $\sigma_u^2 = 0$, i.e., to test whether the data generating process is stationary, the commonly used statistic (known as the KPSS test statistic) is

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

$$\hat{\eta}_u = n^{-2}\sum_{t=1}^{n} S_t^2/s^2(l_n), \quad \text{where } S_t = \sum_{j=1}^{t} e_j, \tag{23}$$

$$s^2(l_n) = \frac{1}{n}\sum_{t=1}^{n} e_t^2 + \frac{2}{n}\sum_{s=1}^{l_n}\sum_{t=s+1}^{n} e_t e_{t-s},$$

and $e_t = y_t - (1/n)\sum_{t=1}^{n} y_t$ is the residual from the regression of y on intercept ψ. Kwiatkowski, Phillips, Schmidt, and Shin (1992) discussed the asymptotic distribution of the $\hat{\eta}_u$ provided z_t satisfies the (strong mixing) regularity conditions given by Phillips and Perron (1988, p. 336) or the linear process conditions given by Phillips and Solo (1992, Theorems 3.4 and 3.15). One of Phillips and Solo's conditions is $\sum_{k=0}^{\infty} k^{1/2}|\psi_k| < \infty$. In this paper, we only require that $\sum_{k=0}^{\infty}|\psi_k| < \infty$. In particular, we only need l_n satisfying $l_n = o(n)$ and $l_n \to \infty$, which, in practice, provides more choice for $s^2(l_n)$. Therefore, our result is an extension of theirs.

THEOREM 3.4. *Let $\epsilon_k, k = 0,\pm1,\pm2,\ldots$, be i.i.d. random variables with $E\epsilon_0 = 0$ and $E\epsilon_0^2 = \sigma^2$. Assume that the data generating process is given by (21) with*

$$z_t = X_t = \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k}.$$

If $\sum_{k=0}^{\infty}|\psi_k| < \infty$ and $\sum_{k=0}^{\infty}\psi_k \neq 0$, then for any l_n satisfying $l_n = o(n)$ and $l_n \to \infty$,

$$\hat{\eta}_u \Rightarrow \int_0^1 V(r)^2 dr, \quad \textit{where } V(r) = W(r) - rW(1). \tag{24}$$

Proof. Under the hypothesis $\sigma_u^2 = 0$, it is well known that $e_t = X_t - (1/n)\sum_{t=1}^{n} X_t$. By applying Theorem 2.1, we have that, for any $0 \le r \le 1$,

$$\frac{1}{\sigma_n} S_{[nr]} = \frac{1}{\sigma_n}\sum_{t=1}^{[nr]} X_t - \frac{[nr]}{n\sigma_n}\sum_{t=1}^{n} X_t \Rightarrow W(r) - rW(1) = V(r),$$

where $\sigma_n^2 = n\sigma^2(\sum_{k=0}^{\infty}\psi_k)^2$. Hence, it follows from the continuous mapping theorem that

$$n^{-2}\sum_{t=1}^{n} S_t^2 = \frac{1}{n}\int_0^1 S_{[nr]}^2 dr \Rightarrow \sigma^2\left(\sum_{k=0}^{\infty}\psi_k\right)^2 \int_0^1 V(r)^2 dr. \tag{25}$$

On the other hand, we have that

$$\begin{aligned} s^2(l_n) &= \frac{1}{n}\sum_{t=1}^{n} e_t^2 + \frac{2}{n}\sum_{s=1}^{l_n}\sum_{t=s+1}^{n} e_t e_{t-s} \\ &= \frac{1}{n}\sum_{t=1}^{n} X_t^2 + \frac{2}{n}\sum_{s=1}^{l_n}\sum_{t=s+1}^{n} X_t X_{t-s} + R_{1n}, \end{aligned} \tag{26}$$

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

where, after a simple calculation,

$$|R_{1n}| \le \frac{4l_n}{n^2}\left(\sum_{j=1}^{n} X_j\right)^2 + \frac{2}{n^2}\left|\sum_{s=1}^{l_n}\sum_{t=s+1}^{n}(X_t + X_{t-s})\right|\left|\sum_{j=1}^{n} X_j\right|$$

$$\le \frac{Cl_n}{n^2}\left(\sum_{j=1}^{n} X_j\right)^2.$$

By noting (15), Markov's inequality implies that for any $l_n = o(n)$, $|R_{1n}| \to 0$ in probability. Therefore, by using part (b) of Theorem 3.2, we obtain that for any l_n satisfying $l_n = o(n)$ and $l_n \to \infty$,

$$s^2(l_n) \to \sigma^2\left(\sum_{k=0}^{\infty}\psi_k\right)^2, \quad \text{in probability.} \tag{27}$$

Thus, (24) follows immediately from (25) and (27). The proof of Theorem 3.4 is complete. ■

Remark 3.1. In terms of Theorem 2.2, Theorems 3.1, 3.3, and 3.4 still hold for stationary ergodic martingale difference sequence provided $\sum_{k=0}^{\infty}|\psi_k| < \infty$ and $\sum_{k=0}^{\infty}\psi_k \neq 0$. We omitted the details here.

4. CONCLUSION

This paper derives two basic results on the invariance principle for the partial sum process of a linear process. The first result assumes that the innovations are i.i.d. random variables, but absolute summability of coefficients for the linear process (i.e., $\sum_{k=0}^{\infty}|\psi_k| < \infty$) is weakened. This relaxation of conditions is interesting because some linear processes do not have absolutely summable coefficients. Especially, a linear process with $\psi_k = k^{-1}l(k)$ where $\sum_{k=0}^{\infty}k^{-1}l(k) = \infty$ is important because it is expressed by neither a finite-order autoregressive moving average process nor a fractional process. The second result is for the situation where the innovations form a martingale difference sequence. For this result, the commonly used assumption of equal variance is removed. This is of interest to researchers from a practical point of view. We apply these general results to unit root testing and stationarity testing. It turns out the limit distributions of the Dickey–Fuller test statistic and KPSS test statistic still hold for the more general models under very weak conditions. This paper also shows that the "long-run variance," σ^2, can be consistently estimated by a nonparametric method with a lag-truncation parameter l_n of $o(n)$. In previous research, it was usually assumed to be of $o(n^{1/2})$. This provides more choice for the estimation of σ^2, and it is theoretically interesting.

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

5. PROOFS OF MAIN RESULTS

5.1. Some Preliminary Lemmas

In this section, we provide some lemmas that will be needed in the proofs of the main results. Some of these lemmas are also interesting in their own right.

LEMMA 5.1. *Let $\{\eta_k, k \geq 0\}$ be a sequence of arbitrary random variables and $\{b_i, i \geq 0\}$ a sequence of real numbers. Assume that*

$$\psi_0^2 + \sum_{k=1}^{\infty} k\psi_k^2 < \infty, \qquad \sup_{n\geq 1} \frac{1}{n} \sum_{i=0}^{n} b_i^2 < \infty$$

and there exists a positive constant A such that

$$E\left(\sum_{k=0}^{\infty} \psi_{j+k}\eta_k\right)^2 \leq A \sum_{k=0}^{\infty} \psi_{j+k}^2 b_k^2, \quad \text{for } j \geq 0. \tag{28}$$

Then, as $n \to \infty$,

$$\frac{1}{\sqrt{n}} \max_{0\leq m\leq n} \left|\sum_{j=0}^{m}\sum_{k=0}^{\infty} \psi_{j+k}\eta_k\right| \to 0, \quad \text{in probability.} \tag{29}$$

Proof. By using $E|Y| \leq (EY^2)^{1/2}$ for any random variable Y, it follows from (28) that

$$E \max_{0\leq m\leq n} \left|\sum_{j=0}^{m}\sum_{k=0}^{\infty} \psi_{j+k}\eta_k\right| \leq \sum_{j=0}^{n} E\left|\sum_{k=0}^{\infty} \psi_{j+k}\eta_k\right| \leq A\sum_{j=0}^{n}\left(\sum_{k=0}^{\infty} \psi_{j+k}^2 b_k^2\right)^{1/2}. \tag{30}$$

Put $\alpha_j = \sum_{k=0}^{\infty} \psi_{j+k}^2 b_k^2$. For any $0 \leq l \leq m$, we have that

$$\begin{aligned}\sum_{j=l}^{m} \alpha_j &= \sum_{j=l}^{m}\sum_{k=0}^{\infty} \psi_{j+k}^2 b_k^2 \\ &= \sum_{k=l}^{\infty} \psi_k^2 \sum_{j=l}^{\min\{k,m\}} b_{k-j}^2 \leq \left(\sup_{k\geq 1} \frac{1}{k}\sum_{i=0}^{k} b_i^2\right)\sum_{k=l}^{\infty}(k+1)\psi_k^2.\end{aligned}$$

This inequality implies that

$$\begin{aligned}\left(\sum_{j=0}^{n} \alpha_j^{1/2}\right)^2 &= \left\{\left(\sum_{j=0}^{[\sqrt{n}]} + \sum_{j=[\sqrt{n}]+1}^{n}\right)\alpha_j^{1/2}\right\}^2 \\ &\leq \sqrt{2}\left\{\left(\sum_{j=0}^{[\sqrt{n}]} \alpha_j^{1/2}\right)^2 + \left(\sum_{j=[\sqrt{n}]}^{n} \alpha_j^{1/2}\right)^2\right\} \\ &\leq 2\left([\sqrt{n}]\sum_{j=0}^{[\sqrt{n}]} \alpha_j + n\sum_{j=[\sqrt{n}]}^{n} \alpha_j\right) \\ &\leq 2\left(\sup_{k\geq 1}\frac{1}{k}\sum_{i=0}^{k} b_i^2\right)\left([\sqrt{n}]\sum_{k=0}^{\infty}(k+1)\psi_k^2 + n\sum_{k=[\sqrt{n}]}^{\infty}(k+1)\psi_k^2\right). \\ &= o(n).\end{aligned}$$

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

Now (29) follows from Markov's inequality, (30), and the bound established earlier. The proof of Lemma 5.1 is complete. ■

LEMMA 5.2. *Let* $\{\eta_n, \mathcal{F}_n, s \le n \le t\}$ *be a martingale difference sequence. Then there exists a constant K such that for any constant sequence* α_k,

$$E \max_{s\le n\le t}\left(\sum_{k=s}^{n} \alpha_k \eta_k\right)^2 \le K \sum_{k=s}^{t} \alpha_k^2 E\eta_k^2. \tag{31}$$

Proof. Apply Doob and Burkholder's inequality (see, e.g., Hall and Heyde, 1980, pp. 15 and 23, respectively). ■

LEMMA 5.3. *Let* $\{\eta_n, \mathcal{F}_n, -\infty < n < \infty\}$ *be a martingale difference sequence satisfying* $\sup_{k\ge 1}(1/k)\sum_{i=-k}^{k} E\eta_i^2 < \infty$. *Assume that*

$$\sum_{k=0}^{\infty} |\psi_k| < \infty \quad \text{and} \quad \sum_{k=1}^{\infty} k\psi_k^2 < \infty.$$

Then, for any $\delta > 0$,

$$\lim_{l\to\infty} \limsup_{n\to\infty} P\left\{\max_{1\le m\le n}\left|\sum_{j=1}^{m} u_j^{(l)}\right| \ge \delta\sqrt{n}\right\} = 0, \tag{32}$$

where $u_j^{(l)} = \sum_{k=l+1}^{\infty} \psi_k \eta_{j-k}$ *and* $l \ge -1$.

Proof. We first note that $\sum_{k=0}^{\infty} \psi_k \eta_{j-k} < \infty$, a.s., for every fixed $j \ge 1$; i.e., $u_j^{(l)}$ is well defined. In fact, by applying Lemma 5.2, there exists a constant K such that for any $j \le m \le m'$,

$$\begin{aligned} E \max_{m\le n\le m'}\left(\sum_{k=m}^{n} \psi_k \eta_{j-k}\right)^2 &\le K \sum_{k=m}^{m'} \psi_k^2 E\eta_{j-k}^2 \\ &\le K\left(\sup_{k\ge 1} \frac{1}{k}\sum_{i=-k}^{k} E\eta_i^2\right)\sum_{k=m}^{\infty}(k+1)\psi_k^2. \end{aligned} \tag{33}$$

From (33) and Markov's inequality, it follows that for any $\delta > 0$, as $n \to \infty$,

$$P\left(\sup_{i\ge 1}\left|\sum_{k=n}^{n+i} \psi_k \eta_{j-k}\right| \ge \delta\right) \le 2K\delta^{-2}\left(\sup_{k\ge 1}\frac{1}{k}\sum_{i=-k}^{k} E\eta_i^2\right)\sum_{k=n}^{\infty} k\psi_k^2 \to 0.$$

So we conclude by the Cauchy criterion that $\sum_{k=0}^{n} \psi_k \eta_{j-k}$ converges almost surely; i.e., for every fixed $j \ge 1$, $\sum_{k=0}^{\infty} \psi_k \eta_{j-k} < \infty$, a.s.

In terms of $\sum_{k=0}^{\infty} \psi_k \eta_{j-k} < \infty$, a.s., it is easy to show (let $\sum_{j=1}^{l} = 0$ for $l \le 0$) that for $m \ge l + 1$,

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

$$\sum_{j=1}^{m} u_j^{(l)} = \sum_{j=1}^{m} \sum_{k=-\infty}^{j-(l+1)} \psi_{j-k}\eta_k$$

$$= \left(\sum_{j=l+2}^{m} \sum_{k=1}^{j-(l+1)} + \sum_{j=1}^{m} \sum_{k=-\infty}^{0} - \sum_{j=1}^{l} \sum_{k=j-l}^{0} \right) \psi_{j-k}\eta_k$$

$$= \sum_{k=1}^{m-(l+1)} \eta_k \sum_{j=l+1}^{m-k} \psi_j + \sum_{j=1}^{m} \sum_{k=0}^{\infty} \psi_{j+k}\eta_{-k} - \sum_{j=1}^{l} \sum_{k=0}^{l-j} \psi_{j+k}\eta_{-k}$$

$$= \Delta_{m1}^{(l)} + \Delta_{m2}^{(l)} + \Delta_{m3}^{(l)}, \quad \text{say,} \tag{34}$$

where $\Delta_{m3}^{(l)} \equiv 0$ for $l = 0$ and -1. Now (32) follows if for any $\delta > 0$,

$$\lim_{l\to\infty} \limsup_{n\to\infty} P\left\{ \max_{1\le m\le n} |\Delta_{mt}^{(l)}| \ge \delta\sqrt{n} \right\} = 0, \qquad t = 1,2,3. \tag{35}$$

For every fixed $j \ge 0$, it follows from the Fatou lemma and Lemma 5.2 that

$$E\left(\sum_{k=0}^{\infty} \psi_{j+k}\eta_{-k} \right)^2 = E \lim_{n\to\infty} \left(\sum_{k=j}^{n} \psi_k \eta_{j-k} \right)^2$$

$$\le \lim_{n\to\infty} E\left(\sum_{k=j}^{n} \psi_k \eta_{j-k} \right)^2 \le K \sum_{k=j}^{\infty} \psi_k^2 E\eta_{j-k}^2$$

$$= K \sum_{k=0}^{\infty} \psi_{j+k}^2 E\eta_{-k}^2.$$

By applying Lemma 5.1 (choosing $b_k^2 = E\eta_{-k}^2$), (35) holds for $t = 2$.

To prove (35) for $t = 1$, put $S_k = \sum_{i=1}^{k} \eta_i$ and $S_0 = 0$. We obtain that

$$\Delta_{m1}^{(l)} = \sum_{k=1}^{m-(l+1)} (S_k - S_{k-1}) \sum_{j=l+1}^{m-k} \psi_j = \sum_{k=1}^{m-(l+1)} \psi_{m-k} S_k$$

and hence,

$$\max_{1\le m\le n} |\Delta_{m1}^{(l)}| \le \sum_{k=l+1}^{n} |\psi_k| \max_{1\le m\le n} |S_m|.$$

Again, it follows from Markov's inequality and Lemma 5.2 that

$$P\left\{ \max_{1\le m\le n} |\Delta_{m1}^{(l)}| \ge \delta\sqrt{n} \right\} \le \delta^{-2} n^{-1} \left(\sum_{k=l+1}^{n} |\psi_k| \right)^2 E \max_{1\le m\le n} S_m^2$$

$$\le K\delta^{-2} \left(\sum_{k=l+1}^{\infty} |\psi_k| \right)^2 \sup_{k\ge 1} \frac{1}{k} \sum_{i=0}^{k} E\eta_i^2.$$

Because $\sum_{k=0}^{\infty} |\psi_k| < \infty$, we conclude that (35) holds for $t = 1$.

That (35) holds for $t = 3$ is obvious and omitted. The proof of Lemma 5.3 is complete. ■

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

LEMMA 5.4. *Let $\{\eta_k, k = 0, \pm 1, \pm 2, \ldots\}$ be a sequence of arbitrary random variables. Assume that, as $n \to \infty$, positive constant series $d_n \to \infty$ and*

$$\frac{1}{d_n^2} \sum_{k=-n}^{n} E\eta_k^2 I_{(|\eta_k| \geq \delta d_n)} \to 0, \quad \text{for any } \delta > 0.$$

Then,

$$\frac{1}{d_n} \max_{-n \leq k \leq n} |\eta_k| \to 0, \quad \text{in probability.}$$

Proof. By applying (36) in Phillips and Solo (1992), we obtain that

$$P\left(\frac{1}{d_n} \max_{-n \leq k \leq n} |\eta_k| \geq \delta\right) = P\left(\frac{1}{d_n^2} \sum_{k=-n}^{n} \eta_k^2 I_{(|\eta_k| \geq \delta d_n)} \geq \delta^2\right).$$

Using Markov's inequality, the result follows. ■

LEMMA 5.5. *Let $\{\eta_k, k = 0, \pm 1, \pm 2, \ldots\}$ be a sequence of arbitrary random variables. If $\{\eta_k^2\}$ is uniformly integrable,*

(i) for any $\delta > 0$, $(1/n)\sum_{k=-n}^{n} E\eta_k^2 I_{(|\eta_k| \geq \delta\sqrt{n})} \to 0$;
(ii) $(1/\sqrt{n})\max_{-n \leq k \leq n} |\eta_k| \to 0$, in probability;
(iii) $(1/n)\sum_{k=1}^{n} (\eta_k^2 - E(\eta_k^2 | \mathcal{F}_{k-1}^*)) \to 0$, in probability, where $\mathcal{F}_k^*$ is the σ-field generated by $\{\eta_j, j \leq k\}$.

By definition of uniform integrability, it follows that $\sup_k E\eta_k^2 I_{(|\eta_k| \geq \delta_n)} \to 0$, for any $\delta_n \to \infty$. Therefore, the proof of Lemma 5.5 is straightforward, and details are omitted.

5.2. Proofs of Results

In this section, we provide the proofs of the main results.

Proof of Theorem 2.1. According to (34) (for $l = -1$), for any $0 \leq t \leq 1$,

$$\sum_{j=1}^{k_n(t)} X_j = \sum_{k=1}^{k_n(t)} \epsilon_k \sum_{j=0}^{k_n(t)-k} \psi_j + \sum_{j=1}^{k_n(t)} \sum_{k=0}^{\infty} \psi_{j+k} \epsilon_{-k}.$$

Similar to (30), we have that

$$E \sup_{0 \leq t \leq 1} \left| \sum_{j=1}^{k_n(t)} \sum_{k=0}^{\infty} \psi_{j+k} \epsilon_{-k} \right| \leq A \sum_{j=1}^{n} \left(\sum_{k=j}^{\infty} \psi_k^2 \right)^{1/2} = o(s_n).$$

This, together with Markov's inequality, implies that

$$\frac{1}{s_n} \sup_{0 \leq t \leq 1} \left| \sum_{j=1}^{k_n(t)} \sum_{k=0}^{\infty} \psi_{j+k} \epsilon_{-k} \right| \to 0, \quad \text{in probability.}$$

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

On the other hand, it is well known (noting that the ϵ_k are i.i.d. random variables) that for any $0 \le t \le 1$,

$$\sum_{k=1}^{k_n(t)} \epsilon_k \sum_{j=0}^{k_n(t)-k} \psi_j \stackrel{d}{=} \sum_{k=1}^{k_n(t)} \epsilon_k \sum_{j=0}^{k-1} \psi_j, \tag{36}$$

where $\stackrel{d}{=}$ denotes the same in distribution. Therefore, by applying Theorem 1.4.1 given in Billingsley (1968, p. 25), (6) follows if

$$\frac{1}{s_n} \sum_{k=1}^{k_n(t)} \epsilon_k \sum_{j=0}^{k-1} \psi_j \Rightarrow W(t), \qquad 0 \le t \le 1. \tag{37}$$

Recall that $\upsilon_k = \sum_{j=0}^{k-1} \psi_j$. Because $\max_{1\le k\le n} |\upsilon_k|/s_n \to 0$, we see that for any $\delta > 0$,

$$\frac{1}{s_n^2} \sum_{k=1}^{n} \upsilon_k^2 E\epsilon_k^2 I_{(|\upsilon_k \epsilon_k| \ge \delta s_n)} \le E\epsilon_1^2 I_{\left(|\epsilon_1| \ge \delta s_n / \max\limits_{1\le k\le n} |\upsilon_k|\right)} \to 0. \tag{38}$$

It follows from Lemma 5.4 that

$$\frac{1}{s_n} \max_{1\le k\le n} |\upsilon_k \epsilon_k| \to 0, \quad \text{in probability.} \tag{39}$$

In terms of (38) and (39), (37) follows from Prokhorov's theorem (see Rao, 1984, p. 343). This completes the proof of (6).

If $\psi_k = k^{-1} l(k)$, where positive function $l(k)$ is slowly varying at infinity, it is easy to check that $\sum_{k=0}^{n} k^{-1} l(k)$ still is slowly varying at infinity. When $\sum_{k=0}^{\infty} k^{-1} l(k) = \infty$, we obtain (see Bingham et al., 1987, p. 26) that

$$s_n^2 = \sum_{j=1}^{n} \left(\sum_{k=0}^{j-1} k^{-1} l(k) \right)^2 \sim n \left(\sum_{k=1}^{n} k^{-1} l(k) \right)^2,$$

$$\sum_{j=0}^{n} \left(\sum_{k=j}^{\infty} k^{-2} l^2(k) \right)^{1/2} \sim \sum_{j=1}^{n} j^{-1/2} l(j) \sim 2n^{1/2} l(n) = o(s_n).$$

Hence (7) follows from (6).

If $0 < |\sum_{k=0}^{\infty} \psi_k| < \infty$ and $\sum_{k=1}^{\infty} k\psi_k^2 < \infty$, by applying Lemma 5.1 and the similar method of the proof used in (6), it suffices to show that

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} \epsilon_k \sum_{j=0}^{k-1} \psi_j \Rightarrow \left(\sum_{j=0}^{\infty} \psi_j \right) W(t), \qquad 0 \le t \le 1. \tag{40}$$

By noting

$$\sum_{k=1}^{[nt]} \epsilon_k \sum_{j=0}^{k-1} \psi_j = \left(\sum_{j=0}^{\infty} \psi_j \right) \sum_{k=1}^{[nt]} \epsilon_k - \sum_{k=1}^{[nt]} \epsilon_k \sum_{j=k}^{\infty} \psi_j,$$

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

(40) follows from Donsker's theorem (see Billingsley, 1968, p. 137), and, as $n \to \infty$,

$$P\left(\sup_{0\le t\le 1} \left| \sum_{k=1}^{[nt]} \epsilon_k \sum_{j=k}^{\infty} \psi_j \right| \ge \delta\sqrt{n} \right) \le C(\delta^2 n)^{-1} E \left(\sum_{k=1}^{n} \epsilon_k \sum_{j=k}^{\infty} \psi_j \right)^2$$

$$\le \frac{C_1}{n} \sum_{k=1}^{n} \left(\sum_{j=k}^{\infty} \psi_j \right)^2 \to 0,$$

where we use the estimate $\sum_{j=k}^{\infty} \psi_j \to 0$ as $k \to \infty$.

If $\sum_{k=0}^{\infty} |\psi_k| < \infty$ and $\sum_{k=1}^{\infty} \psi_k \ne 0$, the result follows from Hannan (1979). The proof of Theorem 2.1 is complete. ■

Proof of Theorem 2.2. Generally speaking, (36) fails to hold for martingale differences. To prove Theorem 2.2, we need a new method.

For every fixed $l \ge 1$, put

$$Z_{1j}^{(l)} = \sum_{k=0}^{l} \psi_k \epsilon_{j-k} \quad \text{and} \quad Z_{2j}^{(l)} = \sum_{k=l+1}^{\infty} \psi_k \epsilon_{j-k}.$$

From Fuller (1996, p. 320), we obtain that for any $m \ge 1$,

$$\sum_{j=1}^{m} Z_{1j}^{(l)} = \sum_{j=1}^{m} \sum_{k=0}^{l} \psi_k \epsilon_{j-k}$$

$$= \sum_{k=0}^{l} \psi_k \sum_{j=1}^{m} \epsilon_j + \sum_{s=1}^{l} \epsilon_{1-s} \sum_{j=s}^{l} \psi_j - \sum_{s=0}^{l-1} \epsilon_{m-s} \sum_{j=s+1}^{l} \psi_j$$

$$= \sum_{k=0}^{l} \psi_k \sum_{j=1}^{m} \epsilon_j + R(m,l), \quad \text{say.}$$

Therefore, it follows that for every fixed $l \ge 1$,

$$\frac{1}{\sigma_n^*} \sum_{j=1}^{k_n^*(t)} X_j = \left(\frac{1}{\sigma_n^*} \sum_{k=0}^{l} \psi_k \right) \sum_{j=1}^{k_n^*(t)} \epsilon_j + \frac{1}{\sigma_n^*} R(k_n^*(t), l) + \frac{1}{\sigma_n^*} \sum_{j=1}^{k_n^*(t)} Z_{2j}^{(l)}. \tag{41}$$

Noting $\sum_{k=0}^{l} \psi_k \to b_0$, as $l \to \infty$, and existing positive constants A_1 and A_2 such that $A_1 n \le \sigma_n^{*2} \le A_2 n$, by applying Theorem 1.4.1 given in Billingsley (1968, p. 25), we only need to show for any $\delta > 0$,

$$\lim_{l\to\infty} \limsup_{n\to\infty} P\left\{ \sup_{0\le t\le 1} \left| \sum_{j=1}^{k_n^*(t)} Z_{2j}^{(l)} \right| \ge \delta\sqrt{n} \right\} = 0; \tag{42}$$

$$\limsup_{n\to\infty} P\left\{ \sup_{0\le t\le 1} |R(k_n^*(t), l)| \ge \delta\sqrt{n} \right\} = 0, \tag{43}$$

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

for every fixed $l \geq 1$; and

$$\frac{1}{s_n^*} \sum_{j=1}^{k_n^*(t)} \epsilon_j \Rightarrow W(t), \qquad 0 \leq t \leq 1, \tag{44}$$

where $s_n^{*2} = \sum_{j=1}^{n} E\epsilon_j^2$.

In fact, (42) follows from Lemma 5.3 because $\{\epsilon_k, \mathcal{F}_n, -\infty < n < \infty\}$ is a martingale difference sequence.

In terms of Lemma 5.4, we have that

$$\frac{1}{\sqrt{n}} \max_{-n \leq j \leq n} |\epsilon_j| \to 0, \quad \text{in probability.} \tag{45}$$

By (45), (43) holds because $\sum_{k=0}^{\infty} |\psi_k| < \infty$ and hence

$$\frac{1}{\sqrt{n}} \sup_{0 \leq t \leq 1} |R(k_n^*(t), l)| \leq \frac{1}{\sqrt{n}} \max_{-l \leq j \leq n} |\epsilon_j| \sum_{s=0}^{l} \left(\sum_{j=s}^{l} |\psi_j| + \sum_{j=s+1}^{l} |\psi_j| \right)$$

$$\to 0, \quad \text{in probability.}$$

Finally, (44) follows Brown (1971) (see also Tanaka, 1996, p. 80) by using (10) and (11). The proof of Theorem 2.2 is complete. ■

REFERENCES

Brown, B. (1971) Martingale central limit theorems. *Annals of Mathematical Statistics* 42, 59–66.

Billingsley, P. (1968) *Convergence of Probability Measure*. New York: Wiley.

Bingham, N.H., C.M. Goldie, & J.L. Teugels (1987) *Regular Variation*. Cambridge, UK: Cambridge University Press.

Brockwell, P.J. & R.A. Davis (1987) *Time Series: Theory and Methods*. New York: Springer-Verlag.

Chow, Y.S. & H. Teicher (1988) *Probability Theory*, 2nd ed. New York: Springer-Verlag.

Davydov, Yu.A. (1970) The invariance principle for stationary processes. *Theory of Probability and Its Applications* 15, 487–498.

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

Feller, W. (1971) *An Introduction to the Theory of Probability and Its Applications*, 2nd ed. New York: Wiley.

Fuller, W.A. (1996) *Introduction to Statistical Time Series*, 2nd ed. New York: Wiley.

Hall, P. (1992) Convergence rates in the central limit theorem for means of autoregressive and moving average sequences. *Stochastic Processes and Their Applications* 43, 115–131.

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

Hannan, E.J. (1979) The central limit theorem for time series regression. *Stochastic Processes and Their Applications* 9, 281–289.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, & Y. Shin (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics* 54, 159–178.

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

Liu, M. (1998) Asymptotics of nonstationary fractional integrated series. *Econometric Theory* 14, 641–662.
Marinucci, D. & P.M. Robinson (1998) Alternative Forms of Fractional Brownian Motion. Discussion paper 354, London School of Economics and Political Science.
Mcleish, D.L. (1975) Invariance principles for dependent variables. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 32, 165–178.
Mcleish, D.L. (1977) On the invariance principle for nonstationary mixingale. *Annals of Probability* 5, 616–621.
Peligrad, M. (1986) Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables. *Progress in Probability and Statistics* 11, 193–223.
Peligrad, M. (1998) Maximum of partial sums and an invariance principle for a class of weak dependent random variables. *Proceedings of the American Mathematical Society* 126, 1181–1189.
Phillips, P.C.B. (1987) Time series regression with a unit root. *Econometrica* 55, 277–302.
Phillips, P.C.B. & P. Perron (1988) Testing for a unit root in time series regression. *Biometrika* 75, 335–346.
Phillips, P.C.B. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.
Phillips, P.C.B. & Z. Xiao (1998) A Primer on Unit Root Testing. Manuscript.
Rao, M.M. (1984) *Probability Theory with Applications*. A series of monographs and textbooks, Z.W. Birnbaum & E. Lukacs (eds.). New York: Academic Press.
Stadtmüller, U. & R. Trautner (1985) Asymptotic behaviour of discrete linear processes. *Journal of Time Series Analysis* 6(2), 97–108.
Stout, W.F. (1974) *Almost Sure Convergence*. New York: Academic Press.
Tanaka, K. (1996) *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*. New York: Wiley.
Truong-van, B. (1995) Invariance principles for semi-stationary sequence of linear processes and applications to ARMA process. *Stochastic Processes and Their Applications* 58, 155–172.
Yokoyama, R. (1995) On the central limit theorem and law of the iterated logarithm for stationary processes with applications to linear processes. *Stochastic Processes and Their Applications* 59, 343–351.

APPENDIX

Proof of (16). Recalling (4), we have that (noting that $E\epsilon_j = 0$ for all j)

$$A_n \equiv \frac{1}{n} E \left| \sum_{r=1}^{l_n} \sum_{t=r+1}^{n} (X_t X_{t-r} - E X_t X_{t-r}) \right|$$

$$= \frac{1}{n} E \left| \sum_{k=0}^{\infty} \sum_{s=0}^{\infty} \psi_k \psi_s \sum_{r=1}^{l_n} \sum_{t=r+1}^{n} (\epsilon_{t-k}\epsilon_{t-r-s} - E\epsilon_{t-k}\epsilon_{t-r-s}) \right|$$

$$\leq \frac{1}{n} \sum_{k=0}^{\infty} \sum_{s=0}^{\infty} |\psi_k \psi_s| E \left| \sum_{t=\lambda}^{n} (\epsilon_{t-k}^2 - E\epsilon_{t-k}^2) \right|$$

$$+ \frac{1}{n} \sum_{k=0}^{\infty} \sum_{s=0}^{\infty} |\psi_k \psi_s| E \left| \sum_{\substack{r=1 \\ r \neq k-s}}^{l_n} \sum_{t=r+1}^{n} \epsilon_{t-k}\epsilon_{t-r-s} \right|,$$

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

where $\lambda = \max\{2, k-s+1\}$. By Markov's inequality, it suffices to show that $A_n \to 0$, as $n \to \infty$. This follows from

$$B_n \equiv \sup_{k,\, s\geq 0} \frac{1}{n} E \left| \sum_{\substack{r\neq k-s \\ r=1}}^{l_n} \sum_{t=r+1}^{n} \epsilon_{t-k}\epsilon_{t-r-s} \right| \to 0, \tag{A.1}$$

where $l_n = o(n)$; and

$$C_n \equiv \sup_{k,\, s\geq 0} \frac{1}{n} E \left| \sum_{t=\lambda}^{n} (\epsilon_{t-k}^2 - E\epsilon_{t-k}^2) \right| \to 0, \tag{A.2}$$

where $\lambda = \max\{2, k-s+1\}$.

Because ϵ_k are i.i.d. random variables with $E\epsilon_0 = 0$ and $E\epsilon_0^2 < \infty$, it follows that

$$E\left(\sum_{\substack{r\neq k-s \\ r=1,\, l_n}} \sum_{t=r+1}^{n} \epsilon_{t-k}\epsilon_{t-r-s} \right)^2 = \sum_{\substack{r\neq k-s \\ r=1}}^{l_n} \sum_{t=r+1}^{n} E\epsilon_{t-k}^2\epsilon_{t-r-s}^2 \leq n l_n (E\epsilon_0^2)^2.$$

Hence, (A.1) follows from, as $n \to \infty$,

$$B_n \leq \frac{1}{n} \sup_{k,\, s\geq 0} \left(E\left(\sum_{\substack{r\neq k-s \\ r=1}}^{l_n} \sum_{t=r+1}^{n} \epsilon_{t-k}\epsilon_{t-r-s} \right)^2 \right)^{1/2}$$

$$\leq (l_n/n)^{1/2} (E\epsilon_0^2)^2 \to 0.$$

To prove (A.2), for every j, let

$$\epsilon_{1,j}^* = \epsilon_j^2 I_{(|\epsilon_j| \leq n^{1/4})} - E\epsilon_j^2 I_{(|\epsilon_j| \leq n^{1/4})} \quad \text{and} \quad \epsilon_{2,j}^* = \epsilon_j^2 I_{(|\epsilon_j| > n^{1/4})} - E\epsilon_j^2 I_{(|\epsilon_j| > n^{1/4})}.$$

After some algebra, we obtain

$$I_{n,k,s} \equiv E\left(\sum_{t=\lambda}^{n} \epsilon_{1,t-k}^* \right)^4 \leq A\{n^2 (E\epsilon_0^4 I_{(|\epsilon_0| \leq n^{1/4})})^2 + nE\epsilon_0^8 I_{(|\epsilon_0| \leq n^{1/4})}\}. \tag{A.3}$$

The relation (A.3) implies that, as $n \to \infty$,

$$H_{n1} \equiv \sup_{k,\, s\geq 0} \frac{1}{n} E \left| \sum_{t=\lambda}^{n} \epsilon_{1,t-k}^* \right|$$

$$\leq \frac{1}{n} \sup_{k,\, s\geq 0} (I_{n,k,s})^{1/4} \leq A\{n^{-1/4} (E\epsilon_0^2)^2 + n^{-3/8} E\epsilon_0^2\} \to 0,$$

where we use the following estimate: $E|X| \leq (EX^4)^{1/4}$ for any X. Therefore, it follows that, as $n \to \infty$,

$$C_n = \sup_{k,\, s\geq 0} \frac{1}{n} E \left| \sum_{t=\lambda}^{n} (\epsilon_{1,t-k}^* + \epsilon_{2,t-k}^*) \right| \leq H_{n1} + 2E\epsilon_0^2 I_{(|\epsilon_0| > n^{1/4})} \to 0.$$

The proof of (16) is complete. ■

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press

Proof of (17). Because $E\epsilon_j\epsilon_k = 0$ for $j \neq k$, we have that

$$\sum_{r=l_n+1}^{n-1}\sum_{t=r+1}^{n} EX_tX_{t-r} = \sum_{k=0}^{\infty}\sum_{s=0}^{\infty}\psi_k\psi_s\sum_{r=l_n+1}^{n}\sum_{t=r+1}^{n} E\epsilon_{t-k}\epsilon_{t-r-s}$$

$$= \sum_{s=0}^{\infty}\sum_{k=s+l_n}^{\infty}\psi_k\psi_s\sum_{t=k-s+1}^{n} E\epsilon_{t-k}^2.$$

Therefore, as $n \to \infty$,

$$\frac{1}{n}\left|\sum_{r=l_n+1}^{n-1}\sum_{t=r+1}^{n} EX_tX_{t-r}\right| \leq E\epsilon_1^2\left(\sum_{k=l_n}^{\infty}|\psi_k|\right)\left(\sum_{s=0}^{\infty}|\psi_s|\right) \to 0.$$

The proof of (17) is complete. ■

https://doi.org/10.1017/S0266466602181072 Published online by Cambridge University Press