The Annals of Statistics
1987, Vol. 15, No. 3, 1050–1063

ASYMPTOTIC INFERENCE FOR NEARLY NONSTATIONARY AR(1) PROCESSES

BY N. H. CHAN AND C. Z. WEI[1]

Indiana University and University of Maryland

A first-order autoregressive process, $Y_t = \beta Y_{t-1} + \varepsilon_t$, is said to be nearly nonstationary when β is close to one. The limiting distribution of the least-squares estimate b_n for β is studied when Y_t is nearly nonstationary. By reparameterizing β to be $1 - \gamma/n$, γ being a fixed constant, it is shown that the limiting distribution of $\tau_n = (\sum_{t=1}^n Y_{t-1}^2)^{1/2}(b_n - \beta)$ converges to $\mathscr{L}(\gamma)$ which is a quotient of stochastic integrals of standard Brownian motion. This provides a reasonable alternative to the approximation of the distribution of τ_n proposed by Ahtola and Tiao (1984).

1. Introduction and main results. Consider a first-order autoregressive AR(1) model

$$Y_t = \beta Y_{t-1} + \varepsilon_t, \qquad t = 1, \dots, n. \tag{1.1}$$

Here, Y_t is the observation at time t, ε_t is the random disturbance and β is an unknown parameter. In the sequel, we shall let Y_0 be zero and $\{\varepsilon_t\}$ be a martingale difference sequence with respect to an increasing sequence of σ-fields $\{\mathscr{F}_t\}$ such that

$$\frac{1}{n}\sum_{t=1}^n E\left(\varepsilon_t^2 | \mathscr{F}_{t-1}\right) \to_P 1, \quad \text{as } n \to \infty \tag{1.2}$$

and

$$\forall\, \alpha > 0, \qquad \frac{1}{n}\sum_{t=1}^n E\left(\varepsilon_t^2 I_{(|\varepsilon_t| > n^{1/2}\alpha)} | \mathscr{F}_{t-1}\right) \to_P 0, \quad \text{as } n \to \infty. \tag{1.3}$$

An important example of $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with zero means and finite variances.

The unknown parameter β is customarily estimated by its least-squares estimate $b_n = \sum_{t=1}^n Y_{t-1}Y_t / \sum_{t=1}^n Y_{t-1}^2$. If the ε_t's are normally distributed, b_n is also the maximum-likelihood estimate of β. In recent years, there has been considerable interest in the asymptotic properties of b_n when β is close or equal to one, both in the statistics and in the econometrics literature [cf. Fuller (1976), Dickey and Fuller (1979), Lai and Siegmund (1983), Ahtola and Tiao (1984), Evans and Savin (1981), Anderson (1985) and Phillips (1987)].

Received January 1986; revised December 1986.

[1] Research supported by the National Science Foundation under Grant DMS-8404081.

AMS 1980 *subject classification.* Primary 62M10.

Key words and phrases. Nearly nonstationary, autoregressive process, least squares, stochastic integral.

When $|\beta| < 1$, it is well known that [see, for example, Mann and Wald (1943)]

$$(1.4)\qquad \tau_n = \left(\sum_{\tau=1}^{n} Y_{t-1}^2\right)^{1/2} (b_n - \beta) \to_{\mathscr{D}} N(0,1), \quad \text{as } n \to \infty,$$

where $N(0,1)$ denotes a standard normal random variable and $\to_{\mathscr{D}}$ designates convergence in distribution. For $|\beta| > 1$, Anderson (1959) showed that (1.4) is true when the ε_t's are independently identically distributed. But for general ε_t's, with $|\beta| > 1$, he showed that the limiting distribution of τ_n may not exist. However, when $\beta = 1$, (1.4) is no longer applicable even if the ε_t's are independently normally distributed and it can be shown that [White (1958) and Rao (1978)] as $n \to \infty$,

$$(1.5)\qquad \tau_n \to_{\mathscr{D}} \tau = \tfrac{1}{2}\left(W^2(1) - 1\right) \Big/ \left\{\int_0^1 W^2(t)\, dt\right\}^{1/2},$$

where $\{W(t)\colon 0 \le t \le 1\}$ is a standard Brownian motion process. Observe that $P(\tau \le 0) = P(W^2(1) \le 1) \doteq 0.684$. This indicates that (1.4) may not be a satisfactory approximation when β is close to one and the sample size is moderate. Of course, (1.5) could also be used to approximate the distribution of τ_n when β is close to one, as suggested by Evans and Savin (1981). However, this approximation results in a nonsmooth transformation from a standard normal distribution to a distribution τ which is nonintuitive.

Under a specific sequential sampling scheme, Lai and Siegmund (1983) showed that the convergence in (1.4) is uniform in β for $|\beta| \le 1$. However, most of the time series encountered are collected in a fixed sample size scheme. Ahtola and Tiao (1984), based on fixed sample size considerations, proposed to approximate the distribution of τ_n through a quadratic form decomposition when β is close to but less than one. They expressed this decomposition as a sum of two quantities which in turn were approximated by a normal and an F random variable. However, the approximation by an F random variable was based on some heuristic moment considerations and a rigorous justification remains to be found.

The main purpose of this paper is to provide another alternative to approximate the distribution of τ_n when β is close to one. Our approach is motivated by the following classical Poisson approximation analogue.

Let X be a binomial random variable with parameters n and p. It is well known that for fixed p, $U_n = (X - np)(np(1-p))^{-1/2} \to_{\mathscr{D}} N(0,1)$ as $n \to \infty$. But when $np \sim \lambda$, Poisson's theorem implies that $U_n \to_{\mathscr{D}} G_\lambda$ as $n \to \infty$, where $G_\lambda = \lambda^{-1/2}(Y_\lambda - \lambda)$ and Y_λ is a Poisson random variable with parameter λ. Notice that G_λ possesses the following properties:

(1.6)
- (i) G_λ is a continuous family of distributions in λ;
- (ii) $G_\lambda \to_{\mathscr{D}} 0$ as $\lambda \to 0$; and
- (iii) from the central limit theorem, $G_\lambda \to_{\mathscr{D}} N(0,1)$ as $\lambda \to \infty$.

We expect a reasonable approximation to the distribution of τ_n, for nearly nonstationary AR(1) models, should reveal similar properties. But before proceeding, let us reparameterize (1.1) through a characterization of the closeness of β to unity.

Suppose β lies in a neighborhood of one and has the form $1 - \gamma/a_n$, where γ is a fixed constant and $\{a_n\}$ is a sequence of constants which increases to infinity. Consider the order of $I_n = E(\sum_{t=1}^n Y_{t-1}^2)$ (this is the order of the observed Fisher information about β when the ε_t's are normally distributed). Clearly,

$$E\left(\sum_{t=1}^{n} Y_{t-1}^2\right) \sim \begin{cases} n/(1-\beta^2), & |\beta| < 1, \\ n^2, & \beta = 1. \end{cases} \tag{1.7}$$

As nonstationary behavior of $\{Y_t\}$ becomes dominant when β is near one, a reasonable choice for a_n should retain the order of I_n to be n^2, i.e., corresponding to the nonstationary boundary $\beta = 1$. Putting $\beta = 1 - \gamma/a_n$ into (1.7), this consideration leads to $\beta = \beta_n = 1 - \gamma/n$. Reparameterizing (1.1) according to β_n, we are now ready to state our principal result.

THEOREM 1. *Let $\beta_n = 1 - \gamma/n$. For $t = 1, \dots, n$, suppose $Y_t(n)$ satisfies the reparameterized $AR(1)$ model,*

$$Y_t(n) = \beta_n Y_{t-1}(n) + \varepsilon_t, \qquad Y_0(n) = 0, \text{ for all } n \tag{1.8}$$

and $\{\varepsilon_t\}$ satisfying (1.2) *and* (1.3). *Then, as $n \to \infty$,*

$$\tau_n = \left(\sum_{t=1}^{n} Y_{t-1}^2\right)^{1/2} (b_n - \beta_n) \to_{\mathscr{D}} \mathscr{L}(\gamma),$$

where

$$\mathscr{L}(\gamma) = \int_0^1 (1+bt)^{-1} W(t)\, dW(t) \Big/ \left\{\int_0^1 (1+bt)^{-2} W^2(t)\, dt\right\}^{1/2},$$

$$b = e^{2\gamma} - 1$$

and $\{W(t)\colon 0 \le t \le 1\}$ is a standard Brownian motion.

REMARK. Notice that Theorem 1 is valid only when β_n is close to one. Thus, in practice, we have to require $|\gamma/n|$ to be small so that β_n stays near one.

Observe that $\mathscr{L}(\gamma)$ is a continuous family of distributions in the parameter γ. This corresponds to (i) of (1.6). When $\gamma = 0$, Itô's formula [cf. Arnold (1974), page 90] implies

COROLLARY 1. $\mathscr{L}(0) = \frac{1}{2}(W^2(1) - 1)/\{\int_0^1 W^2(t)\, dt\}^{1/2}$.

Notice that $\mathscr{L}(0) = \tau$ in (1.5), which corresponds to (ii) of (1.6). For (iii) of (1.6), we have

THEOREM 2. $\mathscr{L}(\gamma) \to_{\mathscr{D}} N(0,1)$ *as* $|\gamma| \to \infty$.

Even though Theorem 2 holds for $\gamma \to \infty$ and $\gamma \to -\infty$, the underlying reasoning for this result is quite different between positive and negative γ's. For large positive γ's, β can be thought of much less than one so that the model can be viewed as stationary and asymptotic normality prevails, which corresponds to (iii) in (1.6). On the other hand, negative γ means β is larger than one and our AR(1) model becomes explosive. Theorem 2 shows that asymptotic normality is still legitimate for such nearly explosive AR(1) models and serves as a complement to the result of Anderson (1959).

In recent literature, there has been a discussion concerning the numerical tabulation of the limiting distribution of τ_n. Dickey and Fuller (1979) showed that the limiting distribution of τ_n is a quotient of weighted sums of independent standard normal random variables from which they obtained their numerical tabulations. By means of a fourier series expansion for $W(t)$, Chan and Wei (1985) showed that $\mathscr{L}(0)$ is equivalent to the form obtained by Dickey and Fuller. We believe that similar ideas can be used to obtain a series expansion for $\mathscr{L}(\gamma)$ so that numerical tabulation will be achieved.

The rest of this paper is devoted to proving Theorems 1 and 2.

2. Proof of Theorem 1. We first establish three lemmas. To simplify notation, let $Y_t(n)$ and β_n be written, respectively, as Y_t and β throughout this paper. Let $X_{n,i} = n^{-1/2}\beta^{n-i}\varepsilon_i$. For each n, $\{X_{n,i}\}$ is a martingale difference sequence with respect to the σ-fields $\mathscr{F}_{n,i} = \mathscr{F}_i$. For $t \in [0,1]$, let $X_n(t) = \sum_{i=1}^{[nt]} X_{n,i}$, where $[nt]$ is the greatest integer function of nt. We have the following functional central limit theorem for $X_n(t)$.

LEMMA 2.1. *For $t \in [0,1]$, let $B_\gamma(t) = e^{-2\gamma}(e^{2t\gamma} - 1)/2\gamma$. Then*

$$X_n(t) \to_{\mathscr{D}} W_1(B_\gamma(t)), \quad \text{in } D[0,1] \text{ as } n \to \infty, \tag{2.1}$$

where $W_1(t)$ is a standard Brownian motion.

PROOF. Observe that $B_\gamma(t)$ is a continuous function of t mapping $[0,1]$ onto $[0, c]$, where $c = B_\gamma(1) = (1 - e^{-2\gamma})/2\gamma$. In view of the martingale central limit theorem of Rootzén (1983) (see Theorem A in the Appendix), the proof of (2.2) is completed once we have established the conditional Lindeberg condition and the conditional variance for $X_n(t)$. For the conditional Lindeberg condition, $\forall\ \alpha > 0$, consider

$$\begin{aligned} H_n &= \sum_{i=1}^{n} E\left(X_{n,i}^2 I_{(|X_{n,i}| \geq \alpha)} | \mathscr{F}_{i-1}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} \beta^{2(n-i)} E\left(\varepsilon_i^2 I_{(|\varepsilon_i| \geq n^{1/2}\alpha\beta^{i-n})} | \mathscr{F}_{i-1}\right). \end{aligned}$$

Suppose $\beta = 1 - \gamma/n$ with $\gamma \geq 0$. Then

$$\begin{aligned} H_n &\leq \frac{1}{n}\sum_{i=1}^{n} E\left(\varepsilon_i^2 I_{(|\varepsilon_i| \geq n^{1/2}\alpha)} | \mathscr{F}_{i-1}\right) \\ &= o_p(1), \quad \text{by (1.3)}. \end{aligned}$$

On the other hand, if $\beta = 1 + \gamma/n$ with $\gamma \geq 0$, then

$$H_n \le \frac{e^{2\gamma}}{n} \sum_{i=1}^{n} E\left(\varepsilon_i^2 I_{(|\varepsilon_i| \ge n^{1/2}\alpha e^{-\gamma})} | \mathscr{F}_{i-1}\right) \tag{2.2}$$
$$= o_p(1).$$

For the conditional variance of $X_n(t)$, it is enough to show that for $t \in [0,1]$,

$$\frac{1}{n}\sum_{i=1}^{[nt]} E\left(X_{n,i}^2 | \mathscr{F}_{i-1}\right) = \frac{1}{n}\sum_{i=1}^{[nt]} \beta^{2(n-1)} E\left(\varepsilon_i^2 | \mathscr{F}_{i-1}\right) \tag{2.3}$$
$$\to_P B_\gamma(t), \quad \text{as } n \to \infty.$$

The proof of (2.3) is given in A.2 in the Appendix. Now, we can apply Rootzén's theorem with $\tau_n(t) = [nt]$ and $\tau(t) = B_\gamma(t)$ to obtain

$$X_n(t) \to_{\mathscr{D}} W_1\left(B_\gamma(t)\right), \quad \text{as } n \to \infty. \qquad \square$$

LEMMA 2.2. *Let*

$$R_n = \sum_{i=1}^{n} \beta^{2(i-n)} X_n^2\left(\frac{i}{n}\right)\frac{1}{n} - \int_0^1 e^{2\gamma(1-t)} X_n^2(t)\,dt.$$

Then

$$R_n \to_P 0, \quad \textit{as } n \to \infty.$$

PROOF. For each n such that $|\gamma/n| < 1$, let $d_n = (n/\gamma)\log(1-\gamma/n)$. For fixed γ, $d_n \to -1$ as $n \to \infty$. For $(i-1)/n \le t \le i/n$, $i = 1,\dots,n$, observe that

$$\left|(1-\gamma/n)^{2(i-n)} - e^{2\gamma(1-t)}\right|$$
$$= e^{2\gamma(1-i/n)}\left|e^{(d_n+1)2\gamma(i/n-1)} - e^{2\gamma(i/n-t)}\right| \tag{2.4}$$
$$\le e^{2|\gamma|}\left\{\left|e^{(d_n+1)2\gamma(i/n-1)} - 1\right| + \left|1 - e^{2\gamma(i/n-t)}\right|\right\}.$$

Now,

$$\max_{1\le i\le n}\left|(d_n+1)2\gamma(i/n-1)\right| \le 2|\gamma|\,|d_n+1| \tag{2.5}$$
$$\to 0, \quad \text{as } n \to \infty,$$

$$\max_{1\le i\le n}\ \sup_{(i-1)/n\le t\le i/n}\left|1 - e^{2\gamma(i/n-t)}\right| \le \max\left\{e^{2\gamma/n} - 1, 1 - e^{2\gamma/n}\right\} \tag{2.6}$$
$$\to 0, \quad \text{as } n \to \infty.$$

Thus, it follows from (2.4)–(2.6) that

$$T_n = \max_{1\le i\le n}\left\{\sup_{(i-1)/n\le t\le i/n}\left|(1-\gamma/n)^{2(i-n)} - e^{2\gamma(1-t)}\right|\right\} \tag{2.7}$$
$$= o(1).$$

Observe that by the definition of $X_n(t)$, we have

$$|R_n| = \left| \sum_{i=1}^{n-1} \int_{i/n}^{(i+1)/n} \{\beta^{2(i-n)} - e^{2\gamma(1-t)}\} X_n^2(i/n)\, dt + X_n^2(1)/n \right|$$

$$\le T_n \sum_{i=1}^{n-1} \int_{i/n}^{(i+1)/n} X_n^2(i/n)\, dt + X_n^2(1)/n$$

$$= T_n \int_0^1 X_n^2(t)\, dt + X_n^2(1)/n$$

$$= o_p(1), \quad \text{by Lemma 2.1 and (2.7).}$$

□

LEMMA 2.3. *For $X \in D[0,1]$ with $X \neq 0$, let*

$$h(X) = \left\{\gamma \int_0^1 e^{2\gamma(1-t)} X^2(t)\, dt + \tfrac{1}{2}X^2(1) - \tfrac{1}{2}\right\} \Big/ \left\{\int_0^1 e^{2\gamma(1-t)} X^2(t)\, dt\right\}^{1/2}.$$

Then

$$h\big(W_1(B_\gamma(t))\big) =_{\mathscr{D}} \mathscr{L}(\gamma).$$

PROOF. Let $a = 2\gamma e^{2\gamma}$. Recall that $c = B_\gamma(1)$. Set the variable $s = B_\gamma(t) = e^{-2\gamma}(e^{2t\gamma} - 1)/2\gamma$. By this change of variable, we have

$$\gamma \int_0^1 e^{2\gamma(1-t)} W_1^2\big(B_\gamma(t)\big)\, dt = \gamma \int_0^c \frac{e^{4\gamma}}{(1+as)^2} W_1^2(s)\, ds.$$

Observe that using integration by parts [cf. Arnold (1974), page 93],

$$\int_0^c \frac{e^{4\gamma}}{(1+as)^2} W_1^2(s)\, ds = \gamma e^{4\gamma} \int_0^c \left(-\frac{1}{a}\right) W_1^2(s)\, d\left(\frac{1}{1+as}\right)$$

$$= -\frac{1}{2} e^{2\gamma} W_1^2(c) \frac{1}{1+ac} + \frac{1}{2}\int_0^c \frac{e^{2\gamma}}{1+as}\, d\big(W_1^2(s)\big)$$

$$= -\frac{1}{2} W_1^2(c) + \frac{1}{2}\int_0^c \frac{e^{2\gamma}}{1+as}\big(ds + 2W_1(s)\, dW_1(s)\big),$$

which by Itô's formula [cf. Arnold (1974), page 93] is

$$-\frac{1}{2} W_1^2\big(B_\gamma(1)\big) + \frac{1}{2} + \int_0^c \frac{e^{2\gamma}}{1+as} W_1(s)\, dW_1(s).$$

Hence,

$$h\big(W_1(B_\gamma(t))\big) = \int_0^c \frac{e^{2\gamma}}{1+as} W_1(s)\, dW_1(s) \Big/ \left\{\int_0^c \frac{e^{4\gamma}}{(1+as)^2} W_1^2(s)\, ds\right\}^{1/2}$$

Now let $s = cu$. Then

$$\int_0^c \frac{e^{2\gamma}}{1+as} W_1(s)\, dW_1(s) = \int_0^1 \frac{e^{2\gamma}}{1+acu} W_1(cu)\, dW_1(cu)$$
$$= \int_0^1 \frac{e^{2\gamma}}{1+bu} cW(u)\, dW(u),$$

where $b = ac$ and $W(u) = c^{-1/2}W_1(cu)$ is again a standard Brownian motion. Similarly,

$$\int_0^c \frac{e^{4\gamma}}{(1+as)^2} W_1^2(s)\, ds = \int_0^1 \frac{e^{4\gamma}}{(1+bu)^2} c^2 W^2(u)\, dt. \qquad \square$$

With Lemmas 2.1–2.3, the main idea of the proof of Theorem 1 lies in exploiting the following identity, whose derivation follows directly from squaring (1.8) and summing:

$$(2.8)\quad n^{-1}\sum_{t=1}^n Y_{t-1}\varepsilon_t = \frac{\gamma}{2}\frac{2n-\gamma}{n-\gamma}\sum_{t=1}^n \left(\frac{Y_{t-1}}{n}\right)^2 + \frac{1}{2}\frac{1}{n-\gamma}Y_n^2 - \frac{1}{2}\frac{1}{n-\gamma}\sum_{t=1}^n \varepsilon_t^2.$$

Now, express the preceding quantities in (2.8) in terms of $X_n(t)$, defined in Lemma 2.1. Specifically,

$$n^{-2}\sum_{i=1}^n Y_i^2 = \sum_{i=1}^n n^{-1}\left(\sum_{k=1}^i \beta^{i-k}\varepsilon_k n^{-1/2}\right)^2$$
$$(2.9)\qquad = \sum_{i=1}^n \beta^{2(i-n)} X_n^2\left(\frac{i}{n}\right)\frac{1}{n} - \int_0^1 e^{2\gamma(1-t)}X_n^2(t)\, dt$$
$$+ \int_0^1 e^{2\gamma(1-t)}X_n^2(t)\, dt$$
$$= R_n + \int_0^1 e^{2\gamma(1-t)}X_n^2(t)\, dt.$$

Note that by (1.2) and Theorem B (see the Appendix), we have

$$(2.10)\qquad V_n = \frac{1}{n-\gamma}\sum_{t=1}^n \varepsilon_t^2 - 1$$
$$= o_p(1).$$

Recall

$$\tau_n = \left(\sum_{t=1}^n Y_{t-1}^2\right)^{1/2}(b_n - \beta)$$
$$= \left(\sum_{t=1}^n Y_{t-1}\varepsilon_t\right)\Big/\left(\sum_{t=1}^n Y_{t-1}^2\right)^{1/2}.$$

It follows from (2.8)–(2.10) that

$$\tau_n = \left\{ \frac{\gamma}{2} \frac{2n-\gamma}{n-\gamma} \left(R_n + \int_0^1 e^{2\gamma(1-t)} X_n^2(t)\, dt \right) + \frac{1}{2} \frac{1}{n-\gamma} X_n^2(1) - \frac{1}{2}(V_n + 1) \right\} \Big/ \left\{ R_n + \int_0^1 e^{2\gamma(1-t)} X_n^2(t)\, dt \right\}^{1/2}. \tag{2.11}$$

Since $P\{\int_0^1 e^{2\gamma(1-t)} W_1^2(B_\gamma(t))\, dt > 0\} = 1$, it follows from the continuous mapping theorem [Billingsley (1968), page 30], Lemmas 2.1–2.3, (2.10) and (2.11) that as $n \to \infty$,

$$\begin{aligned} \tau_n &\to_{\mathscr{D}} h\big(W_1\big(B_\gamma(t)\big)\big), \quad \text{with } h \text{ defined in Lemma 2.3,} \\ &=_{\mathscr{D}} \mathscr{L}(\gamma). \end{aligned} \tag{2.12}$$

This completes the proof of Theorem 1. □

3. Proof of Theorem 2. The proof of Theorem 2 is accomplished by using a special case of Theorem 1 of Rootzén (1980) which concerns the limiting distributions of stochastic integrals. We restate this special case in the following lemma.

LEMMA 3.1. *Let $\{W(t)\colon 0 \le t \le 1\}$ be a standard Brownian motion which is measurable with respect to an increasing sequence of σ-fields $\mathscr{F}_t$. Suppose $\{\psi_n(t)\colon t \in [0,1]\}$ is a sequence of random functions which is $\mathscr{F}_t$-measurable. If*

$$\tau_n(1) = \int_0^1 \psi_n^2(s)\, ds \to_P \tau, \quad \text{as } n \to \infty, \tag{3.1}$$

for some random variable τ such that $\tau > 0$ a.s. and

$$\sup_{0 \le t \le 1} \left| \int_0^t \psi_n(s)\, ds \right| \to_p 0, \quad \text{as } n \to \infty, \tag{3.2}$$

then

$$\int_0^1 \psi_n(s)\, dW(s) \Big/ \left\{ \int_0^1 \psi_n^2(s)\, ds \right\}^{1/2} \to_{\mathscr{D}} N(0,1), \quad \text{as } n \to \infty.$$

Let

$$\psi_b(s) = (1+b)^{1/2} W(s)/(1+bs)$$

and

$$\varphi_b(s) = b(\log b)^{-1/2} W(s)/(1+bs).$$

Notice that $b = e^{2\gamma} - 1$. Hence, $b \to -1$ as $\gamma \to -\infty$ and $b \to \infty$ as $\gamma \to \infty$. We have

LEMMA 3.2.

$$\int_0^1 \psi_b^2(s)\, ds \to_P W^2(1), \quad \text{as } b \to -1 \tag{i}$$

and

(ii) $$\int_0^1 \varphi_b^2(s)\,ds \to_P 1, \quad \text{as } b \to \infty.$$

PROOF. Using integration by parts and Itô's formula as in the proof of Lemma 2.3, we obtain

$$\begin{aligned}\int_0^1 \psi_b^2(s)\,ds &= (1+b)\int_0^1 (1+bs)^{-2} W^2(s)\,ds \\ &= (1+b)\int_0^1 \left(-\frac{1}{b}\right) W^2(s)\,d\left(\frac{1}{1+bs}\right) \\ &= \frac{-(1+b)}{b}\left\{ W^2(s)\frac{1}{1+bs}\bigg|_0^1 - \int_0^1 \frac{1}{1+bs}(ds + 2W(s)\,dW(s))\right\} \\ &= -\frac{1}{b}W^2(1) + \frac{1}{b^2}(1+b)\log(1+b) + \frac{2}{b}(1+b)\int_0^1 \frac{W(s)}{1+bs}\,dW(s).\end{aligned} \tag{3.3}$$

Now,

$$E\left(\frac{1}{b}(1+b)\int_0^1 \frac{W(s)}{1+bs}\,dW(s)\right)^2 = \frac{1}{b^2}(1+b)^2\int_0^1 \frac{s}{(1+bs)^2}\,ds$$
$$\to 0, \quad \text{as } b \to -1.$$

It follows from Chebyshev's inequality that

$$\frac{2}{b}(1+b)\int_0^1 \frac{W(s)}{1+bs}\,dW(s) \to_P 0, \quad \text{as } b \to -1.$$

Thus, from (3.3),

$$\int_0^1 \psi_b^2(s)\,ds \to_P W^2(1), \quad \text{as } b \to -1.$$

For (ii), observe that

$$\begin{aligned}\int_0^1 \varphi_b^2(s)\,ds &= b^2(\log b)^{-1}\int_0^1 (1+bs)^{-2} W^2(s)\,ds \\ &= -\frac{b}{\log b}\frac{W^2(1)}{1+b} + \frac{\log(1+b)}{\log b} + 2\frac{b}{\log b}\int_0^1 \frac{W(s)}{1+bs}\,dW(s).\end{aligned} \tag{3.4}$$

Moreover,

$$E\left\{\left(\frac{b}{\log b}\right)\int_0^1 \frac{W(s)}{1+bs}\,dW(s)\right\}^2 = \left(\frac{b}{\log b}\right)^2\int_0^1 \frac{s}{(1+bs)^2}\,ds \to 0, \quad \text{as } b \to \infty.$$

It follows from Chebyshev's inequality and (3.4) that

$$\int_0^1 \varphi_b^2(s)\,ds \to_P 1, \quad \text{as } b \to \infty. \qquad \square$$

In view of this lemma, we have already proved condition (3.1) of Lemma 3.1 for ψ_b and φ_b. Condition (3.2) is the content of

LEMMA 3.3.

(i) $$\sup_{0\le t\le 1}\left|\int_0^t \psi_b(s)\,ds\right| \to_P 0, \quad as\ b \to -1$$

and

(ii) $$\sup_{0\le t\le 1}\left|\int_0^t \varphi_b(s)\,ds\right| \to_P 0, \quad as\ b \to \infty.$$

PROOF. Since $\sup_{0\le t\le 1}|\int_0^t\psi_b(s)\,ds| \le \int_0^1|\psi_b(s)|\,ds$, it is enough to show that $\int_0^1|\psi_b(s)|\,ds \to_P 0$ as $b \to -1$. In view of Markov's inequality, it suffices to show

$$E\int_0^1|\psi_b(s)|\,ds \to 0, \quad \text{as } b \to -1.$$

But

$$\begin{aligned} E\int_0^1|\psi_b(s)|\,ds &\le (1+b)^{1/2}\int_0^1 s^{1/2}(1+bs)^{-1}\,ds \\ &\le (1+b)^{1/2}\int_0^1(1+bs)^{-1}\,ds \\ &\to 0, \quad \text{as } b \to -1. \end{aligned}$$

This establishes (i).

Likewise, (ii) will follow if $E\int_0^1|\varphi_b(s)|\,ds \to 0$ as $b \to \infty$. But this is true since

$$\begin{aligned} E\int_0^1|\varphi_b(s)|\,ds &\le b(\log b)^{-1/2}\int_0^1 \frac{s^{1/2}}{1+bs}\,ds \\ &= 2(b\log b)^{-1/2}\int_0^{b^{1/2}} \frac{v^2}{1+v^2}\,dv, \end{aligned}$$

by letting $v = (bs)^{1/2}$,

$$\begin{aligned} &\le 2(\log b)^{-1/2} \\ &\to 0, \quad \text{as } b \to \infty. \end{aligned}$$ □

Observe that $\mathscr{L}(\gamma) = \int_0^1(1+bs)^{-1}W(s)\,dW(s)/\{\int_0^1(1+bs)^{-2}W^2(s)\,ds\}^{1/2}$ can either be written as $\int_0^1\psi_b(s)\,dW(s)/\{\int_0^1\psi_b^2(s)\,ds\}^{1/2}$ for $\gamma < 0$ or

$$\int_0^1\varphi_b(s)\,dW(s)\bigg/\left\{\int_0^1\varphi_b^2(s)\,ds\right\}^{1/2} \quad \text{for } \gamma > 0.$$

By Lemmas 3.1–3.3, we conclude that $\mathscr{L}(\gamma) \to_{\mathscr{D}} N(0,1)$ as $|\gamma| \to \infty$. This completes the proof of Theorem 2. □

APPENDIX

A.1.

THEOREM A [Rootzén (1983), Theorem 3.5]. *Suppose* $\{X_{n,j}, \mathscr{F}_{n,j}\}$ *is a martingale difference array. Let* $\{\tau_n(t);\ t \in [0,1]\}$ *be a sequence of adapted time* *scales and* $\{\tau(t);\ t \in [0,1]\}$ *a continuous, nonrandom function. If the following holds*:

$$\forall \varepsilon > 0, \qquad \sum_{j=1}^{\tau_n(1)} E\Big(X_{n,j}^2 I_{(|X_{n,j}| > \varepsilon)} | \mathscr{F}_{n,j-1}\Big) \to_P 0, \quad \text{as } n \to \infty$$

and

$$\sum_{j=1}^{\tau_n(t)} E\big(X_{n,j}^2 | \mathscr{F}_{n,j-1}\big) \to_P \tau(t), \quad \text{as } n \to \infty,\ t \in [0,1],$$

then

$$\sum_{j=1}^{\tau_n(t)} X_{n,j} \to_{\mathscr{D}} W(\tau(t)), \quad \text{as } n \to \infty, \text{ in } D[0,1],$$

where $W(t)$ *is a standard Brownian motion.*

REMARK. Under the same assumptions, this result can also be derived from Theorem (3.2) of Helland (1982).

THEOREM B [Hall and Heyde (1980), Theorem 2.23, page 44]. *Let* $\{X_{n,j}, \mathscr{F}_{n,j}\}$ *be a martingale difference array. If*

$$\forall\ \varepsilon > 0, \qquad \sum_{j=1}^{n} E\Big(X_{n,j}^2 I_{(|X_{n,j}| > \varepsilon)} | \mathscr{F}_{n,j-1}\Big) \to_P 0, \quad \text{as } n \to \infty$$

and if $\sum_{j=1}^n E(X_{n,j}^2 | \mathscr{F}_{n,j-1})$ *is tight, then*

$$\left| \sum_{j=1}^{n} E\big(X_{n,j}^2 | \mathscr{F}_{n,j-1}\big) - \sum_{j=1}^{n} X_{n,j}^2 \right| \to_P 0, \quad \text{as } n \to \infty.$$

A.2.

PROOF OF (2.3). Fix $s \in [0,1]$. For any $\delta > 0$, choose $0 = t_0 \le t_1 \le \cdots \le t_k = s$ such that

$$\text{(A.2.1)} \qquad \max_{1 \le i \le k} \left| e^{-2\gamma(1-t_i)} - e^{-2\gamma(1-t_{i-1})} \right| < \delta.$$

It follows from (A.2.1) that

$$\text{(A.2.2)} \qquad \begin{aligned} &\left| \int_0^s e^{-2\gamma(1-t)}\, dt - \sum_{i=1}^{k} e^{-2\gamma(1-t_{i-1})}(t_i - t_{i-1}) \right| \\ &\quad \le \sum_{i=1}^{k} \int_{t_{i-1}}^{t_i} \left| e^{-2\gamma(1-t)} - e^{-2\gamma(1-t_{i-1})} \right| dt < \delta. \end{aligned}$$

Let $I_i = \{l: [nt_{i-1}] < l \le [nt_i]\}$. Then

$$A_n = \frac{1}{n} \sum_{i=1}^{n} \beta^{2(n-i)} E\left(\varepsilon_i^2 | \mathscr{F}_{i-1}\right) - B_\gamma(s)$$

$$= \frac{1}{n} \sum_{i=1}^{k} \sum_{l \in I_i} \beta^{2(n-l)} E\left(\varepsilon_l^2 | \mathscr{F}_{l-1}\right) - B_\gamma(s)$$

$$(A.2.3) \qquad = \left\{ \frac{1}{n} \sum_{i=1}^{k} \sum_{l \in I_i} \left(\beta^{2(n-l)} - \beta^{2(n-[nt_{i-1}])}\right) E\left(\varepsilon_l^2 | \mathscr{F}_{l-1}\right) \right\}$$

$$+ \left\{ \sum_{i=1}^{k} \beta^{2(n-[nt_{i-1}])} \left[\frac{1}{n} \sum_{l \in I_i} E\left(\varepsilon_l^2 | \mathscr{F}_{l-1}\right) - (t_i - t_{i-1}) \right] \right\}$$

$$+ \left\{ \sum_{i=1}^{k} \beta^{2(n-[nt_{i-1}])} (t_i - t_{i-1}) - B_\gamma(s) \right\}$$

$$= I_n + II_n + III_n, \quad \text{say.}$$

Observe that $B_\gamma(s) = \int_0^s e^{-2\gamma(1-t)}\, dt$.

$$(A.2.4) \qquad |III_n| \le \left| \sum_{i=1}^{k} \beta^{2(n-[nt_{i-1}])} (t_i - t_{i-1}) - \sum_{i=1}^{k} e^{-2\gamma(1-t_{i-1})} (t_i - t_{i-1}) \right| + \left| \sum_{i=1}^{k} e^{-2\gamma(1-t_{i-1})} (t_i - t_{i-1}) - \int_0^s e^{-2\gamma(1-t)}\, dt \right|.$$

Since $\beta^{2(n-[nt_{i-1}])} \to e^{-2\gamma(1-t_{i-1})}$ as $n \to \infty$, in view of (A.2.2), there exists an N such that

$$(A.2.5) \qquad |III_n| < 2\delta, \qquad \forall\, n \ge N.$$

Note that (1.2) implies that for all i,

$$\frac{1}{n} \sum_{l \in I_i} E\left(\varepsilon_l^2 | \mathscr{F}_{l-1}\right) \to_P t_i - t_{i-1}, \quad \text{as } n \to \infty.$$

Furthermore,

$$\max_{1 \le i \le k} \left| \beta^{2(n-[nt_{i-1}])} \right| \le e^{2|\gamma|}.$$

Thus,

$$(A.2.6) \qquad |II_n| \le e^{2|\gamma|} \sum_{i=1}^{k} \left| \frac{1}{n} \sum_{l \in I_i} E\left(\varepsilon_l^2 | \mathscr{F}_{l-1}\right) - (t_i - t_{i-1}) \right| = o_p(1).$$

Finally, we have

$$\max_{1\le i\le k}\max_{l\in I_i}\left|\beta^{2(n-1)}-\beta^{2(n-[nt_{i-1}])}\right|$$
$$=\max_{1\le i\le k}\left|\beta^{2(n-[nt_i])}-\beta^{2(n-[nt_{i-1}])}\right|$$
$$\le\max_{1\le i\le k}\left|e^{-2\gamma(1-t_i)}-e^{-2\gamma(1-t_{i-1})}\right|+o(1)$$
$$\le\delta+o(1),\quad\text{by (A.2.1)}.$$

Hence,

$$|I_n|\le(\delta+o(1))\left(\frac{1}{n}\sum_{i=1}^{n}E\left(\varepsilon_i^2|\mathscr{F}_{i-1}\right)\right)$$
(A.2.7)
$$=(\delta+o(1))\left(1+o_p(1)\right),\quad\text{by (1.2)}$$
$$=\delta+o_p(1).$$

So, given $\eta>0$, choose a small δ such that $\delta<\eta/6$. Then, in view of (A.2.3)–(A.2.7), we have for $n\ge N$,

$$P(|A_n|>\eta)\le P(|I_n|>\eta/3)+P(|II_n|>\eta/3)$$
$$+P(|III_n|>\eta/3)$$
$$\le P(\delta>\eta/6)+P\left(|o_p(1)|>\eta/6\right)$$
$$+P\left(|o_p(1)|>\eta/3\right)$$
$$=P\left(|o_p(1)|>\eta/6\right)+P\left(|o_p(1)|>\eta/3\right).$$

Thus,

$$\lim_{n\to\infty}P(|A_n|>\eta)=0.$$

This completes our proof. □

Acknowledgments. The authors wish to thank the referees and the Associate Editor for their constructive comments which helped to improve the presentation of this paper substantially.

NOTE. After submitting this paper, the authors were informed by I. Llatas that M. J. Bobkoski (1983) had obtained a result similar to Theorem 1 under the i.i.d. assumption in his unpublished Ph.D. thesis.

REFERENCES

AHTOLA, J. A. and TIAO, G. C. (1984). Parametric inference for a nearly nonstationary first-order autoregressive model. *Biometrika* **71** 263–272.

ANDERSON, O. D. (1985). Practical examples of ARUMA modelling. In *Time Series Analysis: Theory and Practice* (O. D. Anderson, ed.) **7** 15–24. North-Holland, Amsterdam.

ANDERSON, T. W. (1959). On asymptotic distributions of estimates of parameters of stochastic difference equations. *Ann. Math. Statist.* **30** 676–687.

ARNOLD, L. (1974). *Stochastic Differential Equations: Theory and Practice*. Wiley, New York.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

BOBKOSKI, M. J. (1983). Hypothesis testing in nonstationary time series. Unpublished Ph.D. thesis, Univ. of Wisconsin, Madison.

CHAN, N. H. and WEI, C. Z. (1985). Limiting distributions of least squares estimate of unstable autoregressive processes. Technical Report, Univ. of Maryland.

DICKEY, D. A. and FULLER, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *J. Amer. Statist. Assoc.* **74** 427–431.

EVANS, G. B. A. and SAVIN, N. E. (1981). The calculation of the limiting distribution of the least squares estimator of the parameter in a random walk model. *Ann. Statist.* **9** 1114–1118.

FULLER, W. A. (1976). *Introduction to Statistical Time Series*. Wiley, New York.

HALL, P. and HEYDE, C. C. (1980). *Martingale Limit Theory and Its Application*. Academic, New York.

HELLAND, I. S. (1982). Central Limit theorems for martingales with discrete or continuous time. *Scand. J. Statist.* **9** 79–94.

LAI, T. L. and SIEGMUND, D. (1983). Fixed accuracy estimation of an autoregressive parameter. *Ann. Statist.* **11** 478–485.

MANN, H. B. and WALD, A. (1943). On the statistical treatment of linear stochastic difference equations. *Econometrica* **11** 173–220.

PHILLIPS, P. C. B. (1987). Time series regression with unit roots. *Econometrica* **55** 277–301.

RAO, M. M. (1978). Asymptotic distribution of an estimator of the boundary parameter of an unstable process. *Ann. Statist.* **6** 185–190.

ROOTZÉN, H. (1980). Limit distributions for the error in approximations of stochastic integrals. *Ann. Probab.* **8** 244–251.

ROOTZÉN, H. (1983). Central limit theory for martingales via random change of time. Technical Report 28, Dept. of Statistics, Univ. of North Carolina.

WHITE, J. S. (1958). The limiting distribution of the serial correlation coefficient in the explosive case. *Ann. Math. Statist.* **29** 1188–1197.

DEPARTMENT OF MATHEMATICS
INDIANA UNIVERSITY
BLOOMINGTON, INDIANA 47405

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND 20742