ASYMPTOTICS FOR STATIONARY VERY NEARLY UNIT ROOT PROCESSES

By

Donald W. K. Andrews Patrik Guggenberger

March 2007

COWLES FOUNDATION DISCUSSION PAPER NO. 1607

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Asymptotics for Stationary Very Nearly Unit Root Processes

Donald W. K. Andrews
Cowles Foundation
Yale University

Patrik Guggenberger
Department of Economics
UCLA

October 2006

Abstract

This paper considers a mean zero stationary first-order autoregressive (AR) model. It is shown that the least squares estimator and t statistic have Cauchy and standard normal asymptotic distributions, respectively, when the AR parameter ρ_n is very near to one in the sense that $1 - \rho_n = o(n^{-1})$.

1 Introduction

A recent paper by Giraitis and Phillips (2006), also see Park (2002) and Phillips and Magdalinos (2007), establishes the asymptotic distribution of the least squares (LS) estimator $\widehat{\rho}_n$ in a stationary first-order AR model without intercept when the AR parameter ρ_n deviates from unity by more than $O(n^{-1})$, i.e., $n(1 - \rho_n) \to \infty$. The result is $(1 - \rho_n^2)^{-1/2} n^{1/2} (\widehat{\rho}_n - \rho_n) \to_d N(0, 1)$. That is, provided ρ_n is not too close to unity the LS estimator has a standard normal distribution. The LS t statistic also has a standard normal distribution.

In addition, results in the literature can be used to obtain the asymptotic distribution of the LS estimator in a stationary AR model when ρ_n deviates from unity by $O(n^{-1})$, but not $o(n^{-1})$—the so-called near unit root case—e.g., see Elliott (1999), Elliott and Stock (2001), and Müller and Elliott (2003). In this case, $n(\widehat{\rho}_n - \rho_n)$ and the LS t statistic have distributions that are functions of an Ornstein-Uhlenbeck process plus an independent normal random variable that arises due to the stationary initial condition. (Bobkowski (1983), Cavanagh (1985), Chan and Wei (1987), and Phillips (1987) consider the AR model with an initial condition that is not stationary. In this case, the independent normal random variable does not appear in the limit distribution.)

In this paper, we consider the case of a stationary AR model with AR parameter $\rho_n < 1$ that is "very nearly" unity in the sense that ρ_n deviates from unity by $o(n^{-1})$. We show that the LS estimator has a Cauchy distribution and the LS t statistic has a standard normal distribution. The rate of convergence of the LS estimator is arbitrarily fast in the sense that any rate can be obtained by letting ρ_n approach one sufficiently fast. These asymptotic results hold because the initial condition dominates the asymptotics. In a model with an estimated intercept or intercept and time trend, the asymptotics are substantially different because the estimation of an intercept eliminates the effect of the initial condition when ρ_n is very nearly a unit root.

The results just described have implications for unit root tests in an AR model with no intercept. The same asymptotic results for the LS estimator and t statistic (as described in the previous paragraph) hold when the initial condition is determined by an AR parameter ρ_n that is very nearly unity and the AR parameter in the model is exactly unity. Because the LS estimator converges to one at a rate faster than $1/n$, the usual LS-estimator-based unit root test under-rejects the null hypothesis of a unit root asymptotically when the true root is unity and the initial condition is very nearly a unit root. In addition, because the α quantile of the standard normal distribution is larger than that of the LS t statistic "unit root distribution," the same is true for the usual LS t-statistic-based unit root test. Hence, both of these unit root tests are robust to the initial condition being very nearly a unit root distribution. These results are related to results of Phillips (2006) for the unit root model with an initial condition that is determined by a unit root process that starts at a time $t_n < 0$, where $t_n \to -\infty$ as $n \to \infty$.

Finite-sample numerical results (not reported here) indicate that the asymptotic results established here only hold for ρ being *extremely* close to one.

Below, we denote convergence in distribution, convergence in probability, and weak convergence as $n \to \infty$ by " $\to_d$ ", " $\to_p$ ", and " $\Rightarrow$," respectively.

2 Results

We consider a (strictly) stationary mean zero first-order autoregressive model:

$$Y_{n,i} = \rho_n Y_{n,i-1} + U_i, \text{ for } i = 1, ..., n, \tag{1}$$

where $\rho_n \in (-1, 1)$ is a nonrandom scalar and the innovations $\{U_i : i = 0, \pm 1, ...\}$ and initial condition $Y_{n,0}$ satisfy the following assumptions.

Assumption I. $\{U_i : i = 0, \pm 1, ...\}$ are i.i.d. with mean zero and variance $\sigma_U^2 \in (0, \infty)$.

Assumption S. $Y_{n,0} = \sum_{j=0}^{\infty} \rho_n^j U_{-j}$.

The sum in Assumption S converges almost surely, e.g., see Brockwell and Davis (1987, Prop. 3.1.1).

Under Assumption S, we have

$$var(Y_{n,0}) = \sigma_U^2/(1 - \rho_n^2). \tag{2}$$

If ρ_n is local to unity in the sense that $\rho_n = 1 - h_n/n$ for $0 < h_n \to h \in (0, \infty)$, then (2) implies that $var(Y_{n,0}^2)$ is $O(n)$ (and not $o(n)$). In the near unit root literature it is often assumed that $Y_{n,0}$ has a distribution that does not depend on n and thus $var(Y_{n,0}^2) = O(1)$, e.g., see Chan and Wei (1987) and Phillips (1987). This yields a triangular array model with random variables $\{Y_{n,i} : 0 \leq i \leq n\}$ that are not stationary in each row. Also, it eliminates the impact of the initial condition $Y_{n,0}$ on the asymptotic theory. There are some near unit root papers, however, that consider a model with stationary initial condition as in the model considered here, e.g., see Elliott (1999), Elliott and Stock (2001), and Müller and Elliott (2003). In these papers, the initial condition has an impact on the asymptotic theory in the AR model.

The least squares (LS) estimator of ρ_n, $\widehat{\rho}_n$, and the studentized t statistic, $T_n(\rho)$, are defined by

$$\widehat{\rho}_n = \frac{\sum_{i=1}^n Y_{n,i-1} Y_{n,i}}{\sum_{i=1}^n Y_{n,i-1}^2} \text{ and } T_n(\rho) = \frac{n^{1/2}(\widehat{\rho}_n - \rho)}{\widehat{\sigma}_n}, \tag{3}$$

where $\widehat{\sigma}_n$ is the usual LS standard deviation estimator. That is, $\widehat{\sigma}_n^2 = \widehat{\sigma}_{Un}^2(n^{-1}\sum_{i=1}^n Y_{n,i-1}^2)^{-1}$ and $\widehat{\sigma}_{Un}^2$ is the sum of squared residuals divided by $n-1$.

The main result of this note is the following.

Theorem 1 *Suppose Assumptions* I *and* S *hold and* $\rho_n \in (-1,1)$ *is such that* $\rho_n = 1 - h_n/n$ *and* $h_n \to 0$ *as* $n \to \infty$. *Then,*

$$(2h_n)^{-1/2} n(\widehat{\rho}_n - \rho_n) \to_d C \text{ and } T_n(\rho_n) \to_d Z,$$

where C *is a Cauchy random variable and* Z *is a standard normal random variable.*

Comments. 1. Theorem 1 shows that the rate of convergence of the LS estimator to the true AR parameter is arbitrarily fast. That is, any rate can be obtained by having ρ_n converge to one (equivalently, h_n converge to zero) sufficiently fast. This occurs because the signal from the regressor $Y_{n,i-1}$ can be made arbitrarily strong by having ρ_n converge to one very fast, whereas the noise in the innovation $U_{n,i}$ is not affected by ρ_n.

2. The intuition behind the result in Theorem 1 is that when $h_n \to 0$ the AR parameter ρ_n is so close to one that the initial condition $Y_{n,0}$ is the realization of the process that is almost a unit root process—$Y_{n,0} = \rho_n Y_{n,-1} + U_0$ for $\rho_n = 1 - o(n^{-1})$, where $Y_{n,-1} = \sum_{j=0}^{\infty} \rho_n^j U_{-j-1}$—and it dominates the behavior of $Y_{n,i}$ for all $i = 0, ..., n$. In particular, $(2h_n)^{1/2} n^{-1/2} Y_{n,[nr]}/\sigma_U \Rightarrow Z$ for a standard normal random variable Z that does not depend on r for $r \in [0,1]$. In contrast, if Assumption S is replaced by $Y_{n,0} = o_p(n)$, then $n^{-1/2} Y_{n,[nr]} \Rightarrow \sigma_U W$ for a Brownian motion W on $[0,1]$.

3. The results of Theorem 1 still hold if $\rho_n = 1$ in (1), but ρ_n in Assumption S satisfies the assumptions of Theorem 1. That is, the LS estimator and t statistic when the model is a unit root model with a very nearly unit root initial condition have Cauchy and normal distributions. The proof just requires minor changes from that of Theorem 1.

For comparative purposes, we now consider the case in which $\rho_n = 1 - h_n/n$ and $h_n \to h \in (0,\infty]$. The result for $h \in (0,\infty)$ is closely related to results in Elliott (1999), Elliott and Stock (2001), and Müller and Elliott (2003), although they do not consider the no-intercept model. The result for $h = \infty$ is due to Giraitis and Phillips (2006).

For a Brownian motion W on $[0,1]$ and an independent standard normal random variable Z, define the Ornstein-Uhlenbeck process $I_h(r)$ and the process $I_h^*(r)$ for $r \in [0,1]$ by

$$\begin{aligned} I_h(r) &= \int_0^r \exp(-(r-s)h)dW(s) \text{ and} \\ I_h^*(r) &= I_h(r) + (2h)^{-1/2}\exp(-hr)Z \text{ for } h > 0. \end{aligned} \tag{4}$$

Proposition 2 *Suppose Assumptions* I *and* S *hold and* $\rho_n \in (-1,1)$ *is such that* $\rho_n = 1 - h_n/n$ *and* $h_n \to h \in (0,\infty]$ *as* $n \to \infty$. *Then,*

(a) *for* $h \in (0,\infty)$,

$$n(\widehat{\rho}_n - \rho_n) \to_d [\int_0^1 I_h^*(r)dW(r)]/[\int_0^1 I_h^*(r)^2 dr] \text{ and}$$

$$T_n(\rho_n) \to_d [\int_0^1 I_h^*(r)dW(r)]/[\int_0^1 I_h^*(r)^2 dr]^{1/2}.$$

(b) *for* $h = \infty$,

$$(1-\rho_n^2)^{-1/2} n^{1/2}(\widehat{\rho}_n - \rho_n) \to_d Z \text{ and } T_n(\rho_n) \to_d Z.$$

Comment. The a.s. limit as $h \to 0$ of $(2h)^{-1/2}$ times the first limit random variable in Proposition 2(a) yields a random variable whose distribution is Cauchy, which corresponds to the first asymptotic distribution in Theorem 1. The a.s. limit as $h \to 0$ of the second limit random variable in Proposition 2(a) yields a random variable whose distribution is standard normal, which corresponds to the second asymptotic distribution in Theorem 1.

3 Proofs

In the integral expressions below, we often leave out the lower and upper limit zero and one, the argument r, and dr to simplify notation. For example, $\int_0^1 I_h(r)^2 dr$ is written as $\int I_h^2$. For simplicity, in the proofs, we drop the subscript n on $Y_{n,i}$.

The proofs of Theorem 1 and Proposition 2 use the following lemmas.

Lemma 3 *Suppose Assumptions* I *and* S *hold and* $\rho_n \in (-1, 1)$ *is such that* $\rho_n = 1 - h_n/n$ *and* $h_n \to h \in [0, \infty)$ *as* $n \to \infty$. *Then,*

$$(2h_n)^{1/2} n^{-1/2} Y_{n,0}/\sigma_U \to_d Z \sim N(0, 1).$$

Define $h_n^* > 0$ by $\rho_n = \exp(-h_n^*/n)$. By a mean-value expansion of $\exp(-h_n^*/n)$, we have $h_n^*/h_n \to 1$ if $h_n = O(1)$, where $\rho_n = 1 - h_n/n$ (see the proof of Lemma 3). The next lemma shows that Lemma 1 in Phillips (1987) continues to hold under our slightly more general assumption that $\rho_n = \exp(-h_n^*/n)$, where h_n^* may depend on n, rather than the sequence $\rho_n = \exp(-h/n)$ used in Phillips (1987).

By recursive substitution, we have

$$\begin{aligned} Y_{n,i} &= \widetilde{Y}_{n,i} + \exp(-h_n^* i/n) Y_{n,0}, \text{ where} \\ \widetilde{Y}_{n,i} &= \textstyle\sum_{j=1}^{i} \exp(-h_n^*(i-j)/n) U_j. \end{aligned} \quad (5)$$

Under Assumption I, it is standard that the innovations satisfy a functional central limit theorem:

$$S_n \Rightarrow W, \text{ where } S_n(r) = n^{-1/2} \textstyle\sum_{i=1}^{[nr]} U_i/\sigma_U \text{ for } r \in [0, 1] \quad (6)$$

and W is a standard Brownian motion. (The same result holds with martingale difference sequences $\{U_i : i = 0, \pm 1, ...\}$ and the results in this paper could be generalized correspondingly.)

Lemma 4 *Suppose Assumption* I *holds and* $\rho_n \in (-1, 1)$ *satisfies* $\rho_n = 1 - h_n/n$, *where* $h_n \to h \in [0, \infty)$. *Then, the following results hold jointly,*

(a) $n^{-1/2} \widetilde{Y}_{n,[nr]} \Rightarrow \sigma_U I_h(r)$ *for* $r \in [0, 1]$,

(b) $n^{-3/2} \sum_{i=1}^{n} \widetilde{Y}_{n,i-1} \Rightarrow \sigma_U \int I_h$,

(c) $n^{-2} \sum_{i=1}^{n} \widetilde{Y}_{n,i-1}^2 \Rightarrow \sigma_U^2 \int I_h^2$,

(d) $n^{-1} \sum_{i=1}^{n} \widetilde{Y}_{n,i-1} U_i \Rightarrow \sigma_U^2 \int I_h(r) dW(r)$, *and*

(e) $\widehat{\sigma}_{Un}^2 \to_p \sigma_U^2$.

Lemmas 3 and 4 and some calculations show that when $h_n \to 0$ the initial condition component of $Y_{n,i}$ in (5) dominates in the asymptotics for the components of the LS estimator. The following Lemma provides the results.

Lemma 5 *Suppose Assumptions* I *and* S *hold and* $\rho_n \in (-1, 1)$ *satisfies* $\rho_n = 1 - h_n/n$, *where* $h_n \to 0$. *Let* Z *and* Z^* *be independent standard normal random variables. Then, the following results hold jointly,*

(a) $(2h_n)^{1/2}n^{-3/2}\sum_{i=1}^n Y_{n,i-1} \to_d \sigma_U Z$,
(b) $2h_n n^{-2}\sum_{i=1}^n Y_{n,i-1}^2 \to_d \sigma_U^2 Z^2$, *and*
(c) $(2h_n)^{1/2}n^{-1}\sum_{i=1}^n Y_{i-1}U_i \to_d \sigma_U^2 ZZ^*$.

Proof of Theorem 1. Lemma 5(c) and (d) and the continuous mapping theorem (CMT) yield

$$(2h_n)^{-1/2}n(\widehat{\rho}_n - \rho_n) = \frac{(2h_n)^{1/2}n^{-1}\sum_{i=1}^n Y_{i-1}U_i}{2h_n n^{-2}\sum_{i=1}^n Y_{i-1}^2} \to_d \frac{\sigma_U^2 ZZ^*}{\sigma_U^2 Z^2} = \frac{Z^*}{Z}. \tag{7}$$

Given that Z^*/Z is a ratio of two independent standard normal random variables, the limit distribution is Cauchy. Furthermore, by Lemma 5(c) and (d) and Lemma 4(e), we have

$$\begin{aligned} T_n(\rho_n) &= \frac{\widehat{\rho}_n - \rho_n}{(\sum_{i=1}^n Y_{i-1}^2)^{-1/2}\widehat{\sigma}_{Un}} = \frac{(2h_n)^{1/2}n^{-1}\sum_{i=1}^n Y_{i-1}U_i}{(2h_n n^{-2}\sum_{i=1}^n Y_{i-1}^2)^{1/2}\widehat{\sigma}_{Un}} \\ &\to_d \frac{\sigma_U^2 ZZ^*}{(\sigma_U^2 Z^2)^{1/2}\sigma_U} = sgn(Z)Z^*. \end{aligned} \tag{8}$$

By independence of Z and Z^*, the conditional distribution of $sgn(Z)Z^*$ given $sgn(Z) = \pm 1$ is $N(0,1)$ and, hence, $sgn(Z)Z^*$ is $N(0,1)$ unconditionally. $\square$

Proof of Lemma 3. As in the text, define h_n^* by $\rho_n = \exp(-h_n^*/n)$. We have $\rho_n = 1 - h_n/n$ and $h_n = O(1)$ implies that $\rho_n \to 1$. Hence, $\exp(-h_n^*/n) = \rho_n \to 1$ and $h_n^* = o(n)$. By a mean-value expansion of $\exp(-h_n^*/n)$ about 0,

$$0 = \rho_n - \rho_n = \exp(-h_n^*/n) - (1 - h_n/n) = h_n/n - \exp(-h_n^{**}/n)h_n^*/n, \tag{9}$$

where $h_n^{**} = o(n)$ given that $h_n^* = o(n)$. Hence, $h_n - (1 + o(1))h_n^* = 0$, $h_n^*/h_n \to 1$, and it suffices to prove the result with h_n^* in place of h_n.

Let $\{m_n : n \geq 1\}$ be a sequence such that $m_n h_n^*/n \to \infty$. By Assumption S, we can write $(2h_n^*/n)^{1/2}Y_0/\sigma_U = A_{1n} + A_{2n}$ for $A_{1n} = (2h_n^*/n)^{1/2}\sum_{j=0}^{m_n}\rho^j U_{-j}/\sigma_U$ and $A_{2n} = (2h_n^*/n)^{1/2}\sum_{j=m_n+1}^{\infty}\rho^j U_{-j}/\sigma_U$. Note that $EA_{2n} = 0$ and

$$\begin{aligned} var(A_{2n}) &= (2h_n^*/n)\textstyle\sum_{j=m_n+1}^{\infty}\rho^{2j} \\ &= (2h_n^*/n)\rho^{2(m_n+1)}/(1-\rho^2) \\ &= (2h_n^*/n)\rho^{2(m_n+1)}/((2h_n^*/n)(1+o(1))) \\ &= O(\exp(-2(m_n+1)h_n^*/n)) \\ &= o(1), \end{aligned} \tag{10}$$

where the third equality holds because $\rho^2 = \exp(-2h_n^*/n) = 1 - (2h_n^*/n)(1 + o(1))$ by a mean value expansion and the last equality holds because $m_n h_n^*/n \to \infty$ by assumption. Therefore, $A_{2n} \to_p 0$.

The result now follows from $A_{1n} \to_d Z$ which holds by the CLT given in Corollary 3.1 in Hall and Heyde (1980) for their $X_{n,i}$ being equal to $(2h_n^*/n)^{1/2}\rho^i U_{-i}/\sigma_U$. Without loss of generality, suppose $\sigma_U = 1$. To apply their Corollary 3.1 we have to verify their (3.21), a Lindeberg condition, and a conditional variance condition. By independence of $\{U_i : i = 0, \pm 1, ...\}$, (3.21) in Hall and Heyde (1980) holds automatically and conditioning on $\mathcal{F}_{n,i-1}$ is superfluous. To check the remaining two conditions, note first that $\sum_{i=0}^{m_n} EX_{n,i}^2 = 2h_n^*\sum_{i=0}^{m_n}\rho^{2i}/n \to 1$ which holds because $\sum_{i=0}^{m_n}\rho^{2i} = (1-\rho^{2(m_n+1)})/(1-\rho^2)$, $\rho^{2(m_n+1)} = \exp(-2h_n^*(m_n+1)/n) \to 0$, and

$$n(1-\rho^2) = n(1-\rho)(1+\rho) = h_n(1+\rho) \to 2h. \tag{11}$$

In addition, for $\varepsilon > 0$,

$$\sum_{i=0}^{m_n} EX_{n,i}^2 I(|X_{n,i}| > \varepsilon) \le (2h_n^*/n)(\sum_{i=0}^{m_n} \rho^{2i}) E(U_0^2 I(2h_n^* U_0^2/n > \varepsilon^2)) = O(1)o(1), \quad (12)$$

where the inequality uses the identical distributions of U_{-j} and the equality uses the result above that $(2h_n^*/n)\sum_{i=0}^{m_n} \rho^{2i} \to 1$ and the dominated convergence theorem. □

Proof of Lemma 4. The proof of parts (a)-(d) follows from the proof of Lemma 1 in Phillips (1987) by using (i) the functional central limit theorem in (6) and (ii) an application of the extended CMT, see Theorem 1.11.1 in van der Vaart and Wellner (1996), rather than the CMT used in Phillips (1987). The extended CMT is needed because the continuous function depends on n. For illustration, we prove part (a). By (5), we have

$$\begin{aligned} n^{-1/2}\widetilde{Y}_{[nr]}/\sigma_U &= \sum_{j=1}^{[nr]} \exp(-h_n^*([nr]-j)/n) U_j/(n^{1/2}\sigma_U) \\ &= \sum_{j=1}^{[nr]} \exp(-h_n^*([nr]-j)/n) \int_{(j-1)/n}^{j/n} dS_n(s) \\ &= \sum_{j=1}^{[nr]} \int_{(j-1)/n}^{j/n} \exp(-h_n^*(r-s)) dS_n(s) + o_p(1) \\ &= \int_0^r \exp(-h_n^*(r-s)) dS_n(s) + o_p(1) \\ &= S_n(r) + h_n^* \int_0^r \exp(-h_n(r-s)) S_n(s) ds + o_p(1) \\ &\Rightarrow W(r) + h\int_0^r \exp(-h(r-s)) W(s) ds \\ &= I_h(r), \end{aligned} \quad (13)$$

where the second to last equality uses integration by parts, the convergence statement uses (6) and the extended CMT. The function $g_n : D_n \to E$ in van der Vaart and Wellner (1996) is given by $g_n(x)(r) = h_n^* \int_0^r \exp(-h_n^*(r-s)) x ds$, where $D_n = D[0,1]$ is the (not separable) metric space of CADLAG functions on the interval $[0,1]$ equipped with the uniform metric and $E = C[0,1]$ is the set of continuous functions on the interval $[0,1]$ also equipped with the uniform metric. Their set D_0 is also chosen as $D[0,1]$. If $x_n \to x$ in $D[0,1]$ then $g_n(x_n) \to g(x)$ in $C[0,1]$ because the function $h_n^* \exp(-h_n^*(r-s))$ converges uniformly (in $r \in [0,1]$) to $h\exp(-h(r-s))$ and any function in $D[0,1]$ is bounded.

To prove part (e), we write

$$\widehat{\sigma}_{Un}^2 = (\widehat{\rho}-\rho)^2 \sum_{i=1}^n Y_{i-1}^2/(n-1) + 2(\widehat{\rho}-\rho)\sum_{i=1}^n Y_{i-1}U_i/(n-1) + \sum_{i=1}^n U_i^2/(n-1). \quad (14)$$

The first two summands are $O_p(n^{-1})$ by (7) and Lemma 5(c) and (d). The third summand is $\sigma_U^2 + o_p(1)$ by the law of large numbers. □

Proof of Lemma 5. By a mean value expansion,

$$\begin{aligned} \max_{1\le j\le 2n} |1-\rho^j| &= \max_{1\le j\le 2n} |1-\exp(-h_n^* j/n)| \\ &= \max_{1\le j\le 2n} |1-(1-h_n^* j\exp(m_j)/n)| \\ &\le 2h_n^* \max_{1\le j\le 2n} |\exp(m_j)| = o(1), \end{aligned} \quad (15)$$

for $0 \le |m_j| \le h_n^* j/n \le 2h_n^*$, where the last equality in (15) holds because $h_n^* \to 0$.

To prove part (a), by (5) we have

$$\begin{aligned} &(2h_n)^{1/2} n^{-3/2} \sum_{i=1}^n Y_{i-1}/\sigma_U \\ &= (2h_n)^{1/2} n^{-3/2} \sum_{i=1}^n \widetilde{Y}_{i-1}/\sigma_U + ((2h_n/n)^{1/2} Y_0/\sigma_U)\sum_{i=1}^n \rho^{i-1}/n \\ &\to_d Z \end{aligned} \quad (16)$$

because the first summand is $o_p(1)$ by Lemma 4(b), $\sum_{i=1}^n \rho^{i-1}/n \to 1$ by (15), and $(2h_n/n)^{1/2} Y_0/\sigma_U \to_d Z$ by Lemma 3.

For part (b), note that by (5),

$$\begin{aligned} 2h_n n^{-2} \textstyle\sum_{i=1}^n Y_{i-1}^2/\sigma_U^2 &= 2h_n n^{-2} \textstyle\sum_{i=1}^n (\widetilde{Y}_{i-1} + \rho^{i-1} Y_0)^2/\sigma_U^2 \\ &= B_{1n} + B_{2n} + B_{3n}, \end{aligned} \tag{17}$$

where $B_{1n} = 2h_n n^{-2} \sum_{i=1}^n \widetilde{Y}_{i-1}^2/\sigma_U^2$, $B_{2n} = 4h_n n^{-2} \sum_{i=1}^n \widetilde{Y}_{i-1}\rho^{i-1} Y_0/\sigma_U^2$, and $B_{3n} = (2h_n n^{-1} Y_0^2/\sigma_U^2) n^{-1} \sum_{i=1}^n \rho^{2(i-1)}$. Lemma 3 implies $B_{3n} \to_d Z^2$ because $n^{-1}\sum_{i=1}^n \rho^{2(i-1)} \to 1$ by (15). Note that $|B_{1n}| \leq 2h_n \sup_{1 \leq i \leq n} |n^{-1/2}\widetilde{Y}_{i-1}/\sigma_U|^2 = h_n O_p(1) = o_p(1)$, where the first equality holds by Lemma 4(a) and the CMT. Finally, by the Cauchy-Schwarz inequality, $|B_{2n}| \leq 2B_{1n}^{1/2} B_{3n}^{1/2} = o_p(1)O_p(1) = o_p(1)$.

To prove part (c), we decompose

$$(2h_n)^{1/2} n^{-1} \textstyle\sum_{i=1}^n Y_{i-1} U_i/\sigma_U^2 = C_{1n} + C_{2n}, \tag{18}$$

where $C_{1n} = (2h_n)^{1/2} n^{-1} \sum_{i=1}^n \widetilde{Y}_{i-1} U_i/\sigma_U^2$ and $C_{2n} = ((2h_n/n)^{1/2} Y_0/\sigma_U) n^{-1/2} \sum_{i=1}^n \rho^{i-1} U_i/\sigma_U$. By Lemma 4(d) and $h_n \to 0$, $C_{1n} = o(1)O_p(1) = o_p(1)$. For C_{2n}, note that by Lemma 3, $(2h_n/n)^{1/2} Y_0/\sigma_U \to_d Z$ and by Assumptions I and S this random variable is independent of $n^{-1/2}\sum_{i=1}^n \rho^{i-1} U_i/\sigma_U$. As in the proof of Lemma 3, an application of Corollary 3.1 in Hall and Heyde (1980) shows that the latter sum converges in distribution to $Z^* \sim N(0,1)$. Note that (15) implies that for $X_{ni} = n^{-1/2}\rho^{i-1} U_i/\sigma_U$ we have $\sum_{i=1}^n EX_{ni}^2 = \sum_{i=1}^n (\rho^2)^{i-1}/n \to 1$. The Lindeberg condition is verified as in (12). From the calculations above, it is clear that the convergence in parts (a)-(c) holds jointly. □

The proof of Proposition 2 uses the following result that follows from Lemmas 3 and 4. Part (a) also can be found in equation (3) of Elliott and Stock (2001).

Corollary 6 *Suppose Assumptions* I *and* S *hold and* $\rho_n \in (-1,1)$ *satisfies* $\rho_n = 1 - h_n/n$, *where* $h_n \to h \in (0,\infty)$. *Then, the following limits hold jointly:*
(a) $n^{-1/2} Y_{n,[nr]} \Rightarrow \sigma_U I_h^*(r)$,
(b) $n^{-3/2}\sum_{i=1}^n Y_{n,i-1} \Rightarrow \sigma_U \int I_h^*$,
(c) $n^{-2}\sum_{i=1}^n Y_{n,i-1}^2 \Rightarrow \sigma_U^2 \int I_h^{*2}$, *and*
(d) $n^{-1}\sum_{i=1}^n Y_{n,i-1} U_i \Rightarrow \sigma_U^2 \int I_h^*(r) dW(r)$.

Proof of Corollary 6. Part (a) follows by

$$\begin{aligned} n^{-1/2} Y_{[nr]}/\sigma_U &= n^{-1/2}\widetilde{Y}_{[nr]}/\sigma_U + n^{-1/2}\exp(-h_n[nr]/n) Y_0/\sigma_U \\ &\Rightarrow I_h(r) + (2h)^{-1/2}\exp(-rh) Z, \end{aligned} \tag{19}$$

where the equality holds by (5), and the convergence holds by Lemma 4(a), Lemma 3, and $\exp(-h_n[nr]/n) \to \exp(-rh)$ uniformly in $r \in [0,1]$. By (5), Z and the Brownian motion W are clearly independent. Parts (b)-(d) are now proved exactly as in Lemma 1 in Phillips (1987). □

Proof of Proposition 2. The result of part (a) (where $h \in (0,\infty)$) follows directly from parts (c) and (d) of Corollary 6 and Lemma 4(e).

For part (b) (where $h = \infty$), it follows from (2) that $EY_{n0}^2 = o(n)$ and thus Assumption A.2 in the Corrigendum to Giraitis and Phillips (2006) holds. The result follows from their Theorem 2.1 and Lemmas 2.1 and 2.2. □

References

BOBKOWSKI, M. J. (1983) Hypothesis testing in nonstationary time series. Unpublished Ph. D. Thesis, University of Wisconsin, Madison.

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

CAVANAGH, C. (1985) Roots local to unity. Unpublished manuscript, Department of Economics, Harvard University.

CHAN, N. H. AND WEI, C. Z. (1987) Asymptotic inference for nearly nonstationary AR(1) processes. *Annals of Statistics* 15, 1050-1063.

ELLIOTT, G. (1999) Efficient tests for a unit root when the initial observation is drawn from its unconditional distribution. *International Economic Review* 40, 767-783.

ELLIOTT, G. and STOCK, J. H. (2001) Confidence intervals for autoregressive coefficients near one. *Journal of Econometrics* 103, 155-181.

GIRAITIS, L. and PHILLIPS, P. C. B. (2006) Uniform limit theory for stationary autoregression. *Journal of Time Series Analysis* 27, 51-60. Corrigendum, forthcoming.

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

MÜLLER, U. and ELLIOTT, G. (2003) Tests for unit roots and the initial condition. *Econometrica* 71, 1269-1286.

PARK, J. Y. (2002) Weak unit roots. Unpublished working paper, Department of Economics, Rice University.

PHILLIPS, P. C. B. (1987) Towards a unified asymptotic theory for autoregression. *Biometrika* 74, 535-547.

PHILLIPS, P. C. B. (2006) When the tail wags the unit root limit distribution. Unpublished manuscript, Cowles Foundation, Yale University.

PHILLIPS, P. C. B. and MAGDALINOS, T. (2007) Limit theory for moderate deviations from a unit root. *Journal of Econometrics* 136, 115-130.

VAN DER VAART, A. W. and WELLNER, J. A. (1996) *Weak Convergence and Empirical Processes.* New York: Springer-Verlag.