LIMIT THEORY FOR MODERATE DEVIATIONS FROM A UNIT ROOT

BY

PETER C. B. PHILLIPS AND TASSOS MAGDALINOS

COWLES FOUNDATION PAPER NO. 1172

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

2006

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

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Econometrics 136 (2007) 115–130
JOURNAL OF Econometrics
www.elsevier.com/locate/jeconom

Limit theory for moderate deviations from a unit root

Peter C.B. Phillips[a,b,*], Tassos Magdalinos[c]

[a]Cowles Foundation for Research in Economics, Yale University, USA
[b]University of Auckland & University of York, UK
[c]University of York, UK

Accepted 19 August 2005
Available online 6 October 2005

Abstract

An asymptotic theory is given for autoregressive time series with a root of the form $\rho_n = 1 + c/k_n$, which represents moderate deviations from unity when $(k_n)_{n\in\mathbb{N}}$ is a deterministic sequence increasing to infinity at a rate slower than n, so that $k_n = o(n)$ as $n \to \infty$. For $c<0$, the results provide a $\sqrt{nk_n}$ rate of convergence and asymptotic normality for the first order serial correlation, partially bridging the $\sqrt{n}$ and n convergence rates for the stationary ($k_n = 1$) and conventional local to unity ($k_n = n$) cases. For $c>0$, the serial correlation coefficient is shown to have a $k_n\rho_n^n$ convergence rate and a Cauchy limit distribution without assuming Gaussian errors, so an invariance principle applies when $\rho_n>1$. This result links moderate deviation asymptotics to earlier results on the explosive autoregression proved under Gaussian errors for $k_n = 1$, where the convergence rate of the serial correlation coefficient is $(1+c)^n$ and no invariance principle applies.

JEL classification: C22

Keywords: Central limit theory; Explosive autoregression; Local to unity; Moderate deviations; Unit root distribution

*Corresponding author. Cowles Foundation for Research in Economics, Yale University, USA.
E-mail address: peter.phillips@yale.edu (P.C.B. Phillips).

0304-4076/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2005.08.002

1. Introduction

Regression asymptotics with roots at or near unity have played an important role in time series econometrics over the last two decades. The limit theory makes extensive use of functional laws of partial sums to Brownian motion, functional laws of weighted partial sums to linear diffusions and weak convergence of discrete martingales to stochastic integrals. Almost all this theory involves time series with autoregressive roots that are at unity (or on the unit circle) or roots that are local to unity in the sense that they have the form $\rho = 1 + c/n$, where n is the sample size. In the latter case, the situation of primary importance occurs when $c < 0$, so that $\rho < 1$ and the local asymptotics therefore seek to characterize alternatives to a unit root that lie in the stationary region. The asymptotic theory turns out to be similar whether $c = 0$ or $c < 0$, and the same rate of convergence in terms of the sample size n applies in both cases. These results have been useful in power evaluations and in confidence interval construction.

To characterize greater deviations from unity we can allow the parameter c to be large and negative or even consider limits as $c \to -\infty$ (Phillips, 1987; Chan and Wei, 1987). While such analysis has proved insightful, it does not resolve all difficulties of the discontinuities of unit root asymptotics. In particular, it does not effectively bridge the very different convergence rates of the stationary and unit root cases.

The present paper takes another approach and provides an asymptotic theory for time series with an autoregressive root of the form $\rho_n = 1 + c/k_n$, where $(k_n)_{n\in\mathbb{N}}$ is a deterministic sequence increasing to infinity at a rate slower than n. Such roots represent moderate deviations from unity in the sense that they belong to larger neighborhoods of one than conventional local to unity roots. An interesting family of moderate deviations from unity roots occurs when we consider $k_n = n^\alpha$ or $\rho_n = 1 + c/n^\alpha$, where the exponent α lies on $(0, 1)$. The boundary value as $\alpha \to 1$ includes the conventional local to unity case, whereas the boundary value as $\alpha \to 0$ includes the stationary or explosive AR(1) process, depending on the sign of c.

The paper provides limit results for a standardized version of such time series, for various sample moments in both the near-stationary ($c < 0$) and the near-explosive ($c > 0$) cases, and for the serial correlation coefficient. When there are near-stationary moderate deviations from unity, the centered first order serial correlation coefficient $\hat{\rho}_n - \rho_n$ is shown to have a $\sqrt{nk_n}$ rate of convergence and a limit normal distribution. In the special case $k_n = n^\alpha$, this rate of convergence becomes $n^{(1+\alpha)/2}$, bridging the $\sqrt{n}$ and n asymptotics of the stationary ($\alpha = 0$) and conventional local to unity ($\alpha = 1$) autoregressions. For near-explosive moderate deviations from unity, the rate of convergence of $\hat{\rho}_n - \rho_n$ is $k_n \rho_n^{-n}$, or $n^\alpha \rho_n^n$ when $k_n = n^\alpha$, which increases with α from $O(n)$ when $\alpha \to 1$ to $O((1 + c)^n)$ when $\alpha \to 0$, thereby bridging the asymptotics of local to unity and explosive autoregressions. An interesting feature of the moderate deviation explosive case is that the limit distribution theory is Cauchy even for non-Gaussian errors. This result differs from conventional theory for the explosive case where the limit distribution is dependent on the distribution of the errors and no invariance principle applies (Anderson, 1959).

After these results were obtained, we learnt of some independent, related work by Park (2003) on weak unit root asymptotics. Park considers autoregressive processes with a root that can be written in the form $\rho = 1 - m/n$ where $m, n \to \infty$. This (weak unit root) setup is analogous to our formulation (see (1) below) of moderate deviations from unity of the form $\rho_n = 1 + c/k_n$. However, the weak unit root specification considers only the stationary side of unity. Using different methods and among some other results, Park shows a rate of convergence of $n/\sqrt{m}$ and asymptotic normality for the serial correlation coefficient in autoregressions with independent identically distributed errors when $(1/m) + (m/n) \to 0$. Theorem 3.2(c) of the present paper also establishes asymptotic normality of the serial correlation coefficient with a rate of convergence $\sqrt{nk_n}$, which corresponds to $n/\sqrt{m}$, on the stationary side of unity ($c<0$). As discussed above, this paper further provides a limit theory for the explosive side of unity ($c>0$).

Some subsequent work on the model (1) considered here has been done by Giraitis and Phillips (2006). These authors consider only the near-stationary case and establish a normal limit theory that is uniform over the autoregressive coefficient $\rho = \rho_n \in [0, 1)$ satisfying $(1 - \rho_n)n \to \infty$ under martingale difference errors. In the present paper, independent identically distributed errors are assumed, which facilitates the use of a Lindeberg–Feller CLT in the near-explosive case.

2. The moderate deviations from unity model

Consider the time series

$$y_t = \rho_n y_{t-1} + u_t, \quad t = 1, \dots, n; \qquad \rho_n = 1 + \frac{c}{k_n} \tag{1}$$

initialized at some $y_0 = o_p(\sqrt{k_n})$ independent of $\sigma(u_1, \dots, u_n)$, where $(k_n)_{n\in\mathbb{N}}$ is a sequence increasing to ∞ such that $k_n = o(n)$ as $n \to \infty$ and u_t is a sequence of independent and identically distributed random variables with $\mathrm{E}u_1 = 0$ and $\mathrm{E}u_1^2 = \sigma^2 < \infty$. For the near-stationary case, $c<0$, we impose a further moment condition, $\mathrm{E}|u_1|^{2+\delta} < \infty$ for some $\delta > 0$.

Since the autoregressive parameter ρ_n is a sequence of the sample size n, the moderate deviations from unity process defined in (1) is, strictly speaking, a triangular array $\{y_{nt} : 1 \leqslant t \leqslant n, n \in \mathbb{N}\}$. Similarly, the initial condition is, in general, a process $(y_{n0})_{n\in\mathbb{N}}$ satisfying the conditions following (1) for each $n \in \mathbb{N}$. In what follows, we employ the abbreviated notation y_t and y_0 for the sake of notational simplicity.

Our approach to developing a limit theory for statistics arising from model (1) is related to standard methods used for stationary and explosive autoregressions.[1] For the near-stationary case, we obtain a law of large numbers for $\sum_{t=1}^n y_{t-1}^2$ under suitable normalization. A martingale central limit theorem is established for a normalized version of $\sum_{t=1}^n y_{t-1}u_t$, giving rise to a Gaussian asymptotic distribution for the normalized and centered least squares estimator $\hat{\rho}_n$.

For the near-explosive case, the standard joint convergence result for $\sum_{t=1}^n y_{t-1}^2$ and $\sum_{t=1}^n y_{t-1}u_t$ applies. Unlike the explosive case $|\rho|>1$, however, we are able to establish a Lindeberg type central limit theorem for a suitably normalized version of $(\sum_{t=1}^n y_{t-1}^2, \sum_{t=1}^n y_{t-1}u_t)$, giving rise to a Cauchy limit theory for the serial correlation coefficient without making any distributional assumptions on the innovation sequence $(u_t)_{1\leqslant t\leqslant n}$.

[1]The original version of this paper used a blocking method and a combination of functional limit theory and standard central limit theory and laws of large numbers in the near-stationary case and martingale limit theory in the near-explosive case.

3. Limit theory for the near-stationary case

This section develops the asymptotic properties of the serial correlation coefficient

$$\hat{\rho}_n - \rho_n = \frac{\sum_{t=1}^{n} y_{t-1}u_t}{\sum_{t=1}^{n} y_{t-1}^2}, \tag{2}$$

when $\rho_n = 1 + (c/k_n)$ and $c<0$. Our approach is to derive a law of large numbers and a central limit theorem, respectively, for the denominator and numerator of (2).

We start by considering the sample variance $\sum_{t=1}^{n} y_{t-1}^2$. The following result allows an analysis of the asymptotic behavior of $\sum_{t=1}^{n} y_{t-1}^2$ analogous to the stationary case $|\rho|<1$.

Lemma 3.1. *For model* (1) *with* $c<0$ *we have, for each* $s \in [0,1]$, *as* $n \to \infty$

(a) $(1/n)y_{\lfloor ns \rfloor}^2 \longrightarrow_{\mathrm{p}} 0$,
(b) $(1/n)\sum_{t=1}^{\lfloor ns \rfloor} y_{t-1}u_t \longrightarrow_{\mathrm{p}} 0$.

By squaring (1) and summing over $t \in \{1, \ldots, \lfloor ns \rfloor\}$ we obtain

$$\begin{aligned}(1-\rho_n^2)\frac{1}{n}\sum_{t=1}^{\lfloor ns \rfloor} y_{t-1}^2 &= \frac{1}{n}y_0^2 - \frac{1}{n}y_{\lfloor ns \rfloor}^2 + \frac{1}{n}\sum_{t=1}^{\lfloor ns \rfloor} u_t^2 + 2\rho_n \frac{1}{n}\sum_{t=1}^{\lfloor ns \rfloor} y_{t-1}u_t \\ &= \frac{1}{n}\sum_{t=1}^{\lfloor ns \rfloor} u_t^2 + \mathrm{o_p}(1) = \sigma^2 s + \mathrm{o_p}(1),\end{aligned}$$

as $n \to \infty$ by Lemma 3.1 and the weak law of large numbers. Using the asymptotic equivalence $1-\rho_n^2 = -2c/k_n[1+\mathrm{O}(k_n^{-1})]$, we obtain

$$\frac{1}{nk_n}\sum_{t=1}^{\lfloor ns \rfloor} y_{t-1}^2 \longrightarrow_{\mathrm{p}} \frac{\sigma^2}{-2c}s \quad \text{as } n \to \infty, \tag{3}$$

for each $s \in [0,1]$. Letting $s=1$, (3) yields the probability limit of the sample variance:

$$\frac{1}{nk_n}\sum_{t=1}^{n} y_{t-1}^2 \longrightarrow_{\mathrm{p}} \frac{\sigma^2}{-2c} \quad \text{as } n \to \infty. \tag{4}$$

The limiting distribution of the sample covariance can be obtained by using the fact that, as in the case of stationary asymptotics, the standardized sample variance has a constant (non-random) probability limit. Defining $\xi_{nt} = (nk_n)^{-1/2}y_{t-1}u_t$, ξ_{nt} is a martingale difference array with respect to the filtration $\mathscr{F}_{nt} = \sigma(y_0, u_1, \ldots, u_t)$ for $1 \leqslant t \leqslant n$ and $n \in \mathbb{N}$. The conditional variance of the martingale array $\sum_{t=1}^{n} \xi_{nt}$ is given by

$$\begin{aligned}\sum_{t=1}^{n} E_{\mathscr{F}_{nt-1}}(\xi_{nt}^2) &= \frac{1}{nk_n}\sum_{t=1}^{n} E_{\mathscr{F}_{nt-1}}(y_{t-1}^2u_t^2) = \frac{1}{nk_n}\sum_{t=1}^{n} y_{t-1}^2 E_{\mathscr{F}_{nt-1}}(u_t^2) \\ &= \sigma^2\frac{1}{nk_n}\sum_{t=1}^{n} y_{t-1}^2 = \frac{\sigma^4}{-2c} + \mathrm{o_p}(1),\end{aligned}$$

by (4), since y_{t-1} is $\mathscr{F}_{nt-1}$ measurable and u_t is independent from $\mathscr{F}_{nt-1}$. By virtue of the Lindeberg condition

$$\sum_{t=1}^{n} E_{\mathscr{F}_{nt-1}}(\xi_{nt}^2 \mathbf{1}\{|\xi_{nt}|>\eta\}) = o_p(1), \quad \eta>0 \tag{5}$$

established in the Appendix, the martingale central limit theorem (e.g. Pollard, 1984, Theorem VIII.1) yields

$$\frac{1}{\sqrt{nk_n}}\sum_{t=1}^{n} y_{t-1}u_t \implies \mathrm{N}\left(0, \frac{\sigma^4}{-2c}\right). \tag{6}$$

Finally, the asymptotic distribution of the centered least squares estimator $\hat{\rho}_n - \rho_n = \sum_{t=1}^{n} y_{t-1}u_t / \sum_{t=1}^{n} y_{t-1}^2$ can be derived by combining (4) and (6):

$$\sqrt{nk_n}(\hat{\rho}_n - \rho_n) \implies \mathrm{N}(0, -2c) \quad \text{as } n \to \infty.$$

We collect these results together as follows.

Theorem 3.2. *For model* (1) *with* $\rho_n = 1 + c/k_n$ *and* $c<0$ *the following limits apply as* $n \to \infty$:

(a) $(1/nk_n)\sum_{t=1}^{n} y_t^2 \longrightarrow_p \sigma^2/-2c$,
(b) $(1/\sqrt{nk_n})\sum_{t=1}^{n} y_{t-1}u_t \Longrightarrow \mathrm{N}(0, \sigma^4/-2c)$,
(c) $\sqrt{nk_n}(\hat{\rho}_n - \rho_n) \Longrightarrow \mathrm{N}(0, -2c)$,

where $(k_n)_{n\in\mathbb{N}}$ *is any sequence increasing to* ∞ *with* $k_n = o(n)$.

Remarks 3.3. (i) The convergence rates of Theorem 3.2 provide a bridge between those that apply for unit root (or local to unity) processes and those that apply under stationarity. This is best understood by letting $k_n = n^\alpha$ for some $\alpha \in (0,1)$, i.e. considering moderate deviations from unity roots of the form $\rho_n = 1 + c/n^\alpha$. Using this parametrization, ρ_n approaches the boundary with the stationary region when $\alpha \to 0$ and the boundary with local to unity roots when $\alpha \to 1$. Part (c) of Theorem 3.2 yields a convergence rate $n^{(1/2)+(\alpha/2)}$ for the serial correlation coefficient, which for $\alpha \in (0,1)$ covers the interval $(n^{1/2}, n)$, providing a link between the $\sqrt{n}$ and n asymptotics of stationary and local to unity autoregressions. While the parametrization $\rho_n = 1 + c/n^\alpha$ is very intuitive, the results of Theorem 3.2 are more general, allowing for arbitrarily large neighborhoods of unity, with ρ_n approaching 1 slower than any polynomial rate, such as $k_n = \log n$.

(ii) Results (a)–(c) match the standard stationary limit theory for fixed $|\rho|<1$. In particular,

$$n^{-1}\sum_{t=1}^{n} y_t^2 \longrightarrow_p \frac{\sigma^2}{1-\rho^2},$$

$$n^{-(1/2)}\sum_{t=1}^{n} y_{t-1}u_t \implies \mathrm{N}\left(0, \frac{\sigma^4}{1-\rho^2}\right),$$

$$\sqrt{n}(\hat{\rho}_n - \rho) \implies \mathrm{N}(0, 1-\rho^2).$$

A heuristic argument for the correspondence is that upon replacing ρ by $1+c/k_n$ in each of the above results, a simple rescaling of the first order approximation delivers (a)–(c) of Theorem 3.2. Thus, for the serial correlation coefficient $\hat{\rho}_n$, substituting $1-\rho^2 = -(2c/k_n)[1+o(1)]$ into the limit distribution of $\sqrt{n}(\hat{\rho}_n - \rho)$ gives the asymptotic approximation

$$\sqrt{n}(\hat{\rho}_n - \rho) \sim_{\mathrm{d}} \mathrm{N}\left(0, -\frac{2c}{k_n}\right) \quad \text{or} \quad \sqrt{nk_n}(\hat{\rho}_n - \rho) \sim_{\mathrm{d}} \mathrm{N}(0, -2c),$$

just as in part (c) of the theorem.

(iii) Under an additional symmetry assumption on the distribution of the innovation errors, a straightforward application of Theorem 3.2 produces a limit theory for time series with moderate deviations from a negative unit root. Suppose that the process $(z_t)_{1\leqslant t\leqslant n}$ is defined by

$$z_t = \theta_n z_{t-1} + v_t = \theta_n^t z_0 + \sum_{j=1}^{t} \theta_n^{t-j} v_j, \quad \theta_n = -1 + \frac{c}{k_n}, \ c>0,$$

where $(v_t)_{1\leqslant t\leqslant n}$ is a sequence of i.i.d. $(0,\sigma^2)$ random variables with $\mathrm{E}|v_1|^{2+\delta} < \infty$ for some $\delta > 0$ and distribution function $F_{v_1}(\cdot)$ on $\mathbb{R}$ satisfying the symmetry condition $F_{v_1}(x) = 1 - \mathrm{P}(v_1 < -x)$, and $z_0 = o_p(\sqrt{k_n})$ is independent of $\sigma(v_1, \ldots, v_n)$. Writing z_t in the form

$$z_t = (-1)^t \left\{ \left(1 - \frac{c}{k_n}\right)^t z_0 + \sum_{j=1}^{t} \left(1 - \frac{c}{k_n}\right)^{t-j} (-1)^{-j} v_j \right\}, \tag{7}$$

and letting $\varepsilon_t = (-1)^t v_t$, $x_t = (-1)^t z_t$, we obtain that ε_t is an i.i.d. F_{v_1} sequence and $x_0 = z_0$. Consequently, $(x_t)_{1\leqslant t\leqslant n}$ is a moderate deviations from unity process with root $\rho_n = 1 - c/k_n$, $c>0$. Since

$$\sum_{t=1}^{n} z_{t-1} v_t = \sum_{t=1}^{n} x_{t-1}\varepsilon_t \quad \text{and} \quad \sum_{t=1}^{n} z_{t-1}^2 = \sum_{t=1}^{n} x_{t-1}^2,$$

part (c) of Theorem 3.2 yields

$$\sqrt{nk_n}(\hat{\theta}_n - \theta_n) = \sqrt{nk_n}(\hat{\rho}_n - \rho_n) \implies \mathrm{N}(0, 2c).$$

4. Limit theory for the near-explosive case

This section considers the limit behavior of the serial correlation coefficient $\hat{\rho}_n - \rho_n$ when $\rho_n = 1 + c/k_n$ and $c>0$. In this case the normalized sample variance $\sum_{t=1}^{n} y_{t-1}^2$ no longer converges in probability to a constant as $n \to \infty$. We show that, under appropriate normalization, the random vector $(\sum_{t=1}^{n} y_{t-1}u_t, \sum_{t=1}^{n} y_{t-1}^2)$ converges weakly to (XY, Y^2), where X and Y are independent $\mathrm{N}(0, \sigma^2/2c)$ random variables.

We begin the analysis of the near-explosive case by presenting two useful preliminary results proved in the Appendix. Lemma 4.1 implies that the normalized sample covariance $\sum_{t=1}^{n} y_{t-1}u_t$ can be approximated by the product of the stochastic sequences

$$X_n := \frac{1}{\sqrt{k_n}} \sum_{t=1}^{n} \rho_n^{-(n-t)-1} u_t \quad \text{and} \quad Y_n := \frac{1}{\sqrt{k_n}} \sum_{j=1}^{n} \rho_n^{-j} u_j, \tag{8}$$

whereas Lemma 4.2 discusses the asymptotic behavior of the approximating sequences $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$.

Lemma 4.1. *For each* $\rho_n = 1 + c/k_n$ *with* $c>0$ *and* $k_n = \mathrm{o}(n)$,

$$\frac{\rho_n^{-n}}{k_n}\sum_{t=1}^{n}\sum_{j=t}^{n}\rho_n^{t-j-1}u_ju_t \longrightarrow_{L_1} 0 \quad \text{as } n\to\infty.$$

Lemma 4.2. *For each* $c>0$, *the sequences* $(X_n)_{n\in\mathbb{N}}$ *and* $(Y_n)_{n\in\mathbb{N}}$ *defined in* (8) *satisfy*

$$(X_n, Y_n) \Longrightarrow (X, Y) \quad \text{as } n\to\infty,$$

where X *and* Y *are independent* $\mathrm{N}(0,\sigma^2/2c)$ *random variables.*

The asymptotic behavior of the sample variance can be derived by expanding $\sum_{t=1}^{n} y_{t-1}^2$ as in the near-stationary case and using Lemma 4.2. By squaring (1) and summing over $t\in\{1,\ldots,n\}$ we obtain

$$\begin{aligned}\frac{\rho_n^{-2n}}{k_n^2}\sum_{t=1}^{n}y_{t-1}^2 &= \frac{1}{k_n^2(\rho_n^2-1)}\left\{\rho_n^{-2n}(y_n^2-y_0^2) - 2\rho_n^{-2n+1}\sum_{t=1}^{n}y_{t-1}u_t - \rho_n^{-2n}\sum_{t=1}^{n}u_t^2\right\}\\ &= \frac{1}{k_n(\rho_n^2-1)}\left\{\frac{\rho_n^{-2n}}{k_n}y_n^2 - \frac{2\rho_n^{-2n+1}}{k_n}\sum_{t=1}^{n}y_{t-1}u_t - \frac{\rho_n^{-2n}}{k_n}\sum_{t=1}^{n}u_t^2\right\} + \mathrm{o_p}(\rho_n^{-2n}),\end{aligned}$$

since $y_0 = \mathrm{o_p}(k_n^{1/2})$. Now $k_n(\rho_n^2-1)\to 2c$ as $n\to\infty$ and

$$\frac{\rho_n^{-2n}}{k_n}\sum_{t=1}^{n}u_t^2 = \mathrm{O_p}\left(\frac{n}{k_n}\rho_n^{-2n}\right) = \mathrm{o_p}(1),$$

by the law of large numbers and Proposition A.1. Also

$$\begin{aligned}\frac{\rho_n^{-2n+1}}{k_n}\sum_{t=1}^{n}y_{t-1}u_t &= \frac{y_0}{\sqrt{k_n}}\frac{\rho_n^{-n}}{\sqrt{k_n}}\sum_{t=1}^{n}\rho_n^{-(n-t)}u_t + \frac{\rho_n^{-2n+1}}{k_n}\sum_{t=1}^{n}\left(\sum_{j=1}^{t-1}\rho_n^{t-1-j}u_j\right)u_t\\ &= \frac{\rho_n^{-2n+1}}{k_n}\sum_{t=1}^{n}\left(\sum_{j=1}^{t-1}\rho_n^{t-1-j}u_j\right)u_t + \mathrm{o_p}(\rho_n^{-n})\end{aligned}$$

by Lemma 4.2. Moreover, independence of the sequence $(u_t)_{1\leqslant t\leqslant n}$ yields

$$\begin{aligned}\mathrm{E}\left[\frac{\rho_n^{-2n+1}}{k_n}\sum_{t=1}^{n}\left(\sum_{j=1}^{t-1}\rho_n^{t-1-j}u_j\right)u_t\right]^2 &= \frac{\sigma^4\rho_n^{-4n}}{k_n^2}\sum_{t=1}^{n}\sum_{j=1}^{t-1}\rho_n^{2(t-j-1)}\\ &= \frac{\sigma^4\rho_n^{-4n}}{k_n^2(\rho_n^2-1)}\left[\sum_{t=1}^{n}\rho_n^{2(t-1)} - n\right] = \mathrm{O}(\rho_n^{-2n}),\end{aligned}$$

implying that $(\rho_n^{-2n+1}/k_n)\sum_{t=1}^n y_{t-1}u_t = o_p(1)$. Thus, the sample variance becomes

$$\begin{aligned}\frac{\rho_n^{-2n}}{k_n^2}\sum_{t=1}^n y_{t-1}^2 &= \frac{1}{k_n(\rho_n^2-1)}\left(\frac{\rho_n^{-n}}{\sqrt{k_n}}y_n\right)^2 + o_p(1)\\ &= \frac{1}{k_n(\rho_n^2-1)}\left(\frac{y_0}{\sqrt{k_n}}+\frac{1}{\sqrt{k_n}}\sum_{j=1}^n \rho_n^{-j}u_j\right)^2 + o_p(1)\\ &= \frac{1}{k_n(\rho_n^2-1)}\left(\frac{1}{\sqrt{k_n}}\sum_{j=1}^n \rho_n^{-j}u_j\right)^2 + o_p(1)\\ &= \frac{1}{2c}Y_n^2 + o_p(1).\end{aligned}$$

Lemma 4.2 now implies that the limiting distribution of the sample variance is given by

$$\frac{\rho_n^{-2n}}{k_n^2}\sum_{t=1}^n y_{t-1}^2 \implies \frac{1}{2c}Y^2, \quad Y \equiv N\left(0,\frac{\sigma^2}{2c}\right), \tag{9}$$

as $n\to\infty$, where Y is the weak limit of Y_n derived in Lemma 4.2.

The asymptotic behavior of the sample covariance can be determined by approximating a suitably normalized version of $\sum_{t=1}^n y_{t-1}u_t$ by the product X_nY_n. We start by writing the sample covariance as

$$\frac{\rho_n^{-n}}{k_n}\sum_{t=1}^n y_{t-1}u_t = \frac{y_0}{\sqrt{k_n}}\frac{1}{\sqrt{k_n}}\sum_{t=1}^n \rho_n^{-(n-t+1)}u_t + \frac{\rho_n^{-n}}{k_n}\sum_{t=1}^n\sum_{j=1}^{t-1}\rho_n^{t-j-1}u_ju_t.$$

The term containing the initial condition tends to 0 in probability, since $y_0/\sqrt{k_n} = o_p(1)$ and

$$\frac{1}{\sqrt{k_n}}\sum_{t=1}^n \rho_n^{-(n-t+1)}u_t = \frac{1}{\sqrt{k_n}}\sum_{j=1}^n \rho_n^{-j}u_{n+1-j} =_d \frac{1}{\sqrt{k_n}}\sum_{j=1}^n \rho_n^{-j}u_j = O_p(1),$$

by Lemma 4.2. Thus, using Lemma 4.1 we obtain

$$\begin{aligned}\frac{\rho_n^{-n}}{k_n}\sum_{t=1}^n y_{t-1}u_t &= \frac{\rho_n^{-n}}{k_n}\sum_{t=1}^n\sum_{j=1}^{t-1}\rho_n^{t-j-1}u_ju_t + o_p(1)\\ &= \frac{\rho_n^{-n}}{k_n}\sum_{t=1}^n\sum_{j=1}^{n}\rho_n^{t-j-1}u_ju_t + \frac{\rho_n^{-n}}{k_n}\sum_{t=1}^n\sum_{j=t}^{n}\rho_n^{t-j-1}u_ju_t + o_p(1)\\ &= \frac{\rho_n^{-n}}{k_n}\sum_{t=1}^n\sum_{j=1}^{n}\rho_n^{t-j-1}u_ju_t + o_p(1)\\ &= \left(\frac{1}{\sqrt{k_n}}\sum_{t=1}^n \rho_n^{-(n-t)-1}u_t\right)\left(\frac{1}{\sqrt{k_n}}\sum_{j=1}^n \rho_n^{-j}u_j\right) + o_p(1)\\ &= X_nY_n + o_p(1).\end{aligned}$$

Since convergence of X_n and Y_n applies jointly, Lemma 4.2 and the continuous mapping theorem imply that the asymptotic distribution of the sample covariance is given by

$$\frac{\rho_n^{-n}}{k_n}\sum_{t=1}^{n} y_{t-1}u_t \implies XY \quad X, Y \equiv \mathrm{N}\left(0, \frac{\sigma^2}{2c}\right). \tag{10}$$

The asymptotic behavior of the serial correlation coefficient in the near-explosive case is an easy consequence of (9), (10) and the fact that the limiting random variables X and Y are independent.

Theorem 4.3. *For model* (1) *with* $\rho_n = 1 + c/k_n$, $c>0$ *the following limits apply as* $n \to \infty$:

(a) $((\rho_n^{-n}/k_n)\sum_{t=1}^{n} y_{t-1}u_t, (\rho_n^{-2n}/k_n^2)\sum_{t=1}^{n} y_{t-1}^2) \Longrightarrow (XY, Y^2)$,
(b) $(k_n\rho_n^n/2c)(\hat{\rho}_n - \rho_n) \Longrightarrow C$,

where X *and* Y *are independent* $\mathrm{N}(0, \sigma^2/2c)$ *random variables and* C *is a standard Cauchy variate.*

Remarks 4.4. (i) Theorem 4.3 relates to earlier work (White, 1958; Anderson, 1959; Basawa and Brockwell, 1984) on the explosive Gaussian AR(1) process. For a Gaussian first order autoregressive process with fixed $|\rho|>1$ and $y_0 = 0$, White showed that

$$\frac{\rho^n}{\rho^2 - 1}(\hat{\rho}_n - \rho) \implies C \quad \text{as } n \to \infty. \tag{11}$$

Replacing ρ by $\rho_n = 1 + c/k_n$, we obtain $\rho^2 - 1 = (2c/k_n)[1 + \mathrm{o}(1)]$. Hence, the normalizations in Theorem 4.3 and (11) are asymptotically equivalent as $n \to \infty$. Anderson (1959) showed that $(\rho^n/(\rho^2 - 1))(\hat{\rho}_n - \rho)$ has a limit distribution that depends on the distribution of the errors u_t when $\rho>1$ and that no central limit theory or invariance principle is applicable.

(ii) The limit theory in this section for the moderate deviations case is not restricted to Gaussian processes. In particular, the Cauchy limit result of part (b) applies for $\rho_n = 1 + c/k_n$ and innovations u_t with finite second moment. This essential difference between the moderately explosive processes of Theorem 4.3 and explosive autoregressions is caused by the different rates of convergence that apply in the two cases. Letting $k_n = n^\alpha$ for some $\alpha \in (0, 1)$, the asymptotic behavior of the serial correlation coefficient in both the pure and the moderately explosive case is determined by the stochastic sequences

$$X_n = \frac{1}{n^{\alpha/2}}\sum_{t=1}^{n} \rho_n^{-(n-t)-1}u_t,$$

$$Y_n = \frac{1}{n^{\alpha/2}}\sum_{j=1}^{n} \rho_n^{-j}u_j.$$

In the explosive case where $\alpha = 0$, both X_n and Y_n converge a.s. by the martingale convergence theorem and joint convergence of (X_n, Y_n) still applies. However, unlike Lemma 4.2, the limit of (X_n, Y_n) is not necessarily a Gaussian random vector. In this explosive case, there is no contribution from the normalization $n^{-\alpha/2}$ as $\alpha = 0$ and a Lindeberg condition does not generally hold, in contrast to the moderately explosive case where $\alpha>0$.

(iii) The limit theory for near-explosive moderate deviations from unity is invariant to the initial condition y_0 being any fixed constant value or random process of smaller asymptotic order than $k_n^{1/2}$. This property is not shared by explosive autoregressions where y_0 does influence the limit theory, as shown by Anderson (1959).

(iv) Theorem 4.3 produces a limit theory for near-explosive time series with moderate deviations from a negative unit root under an additional symmetry assumption on the distribution of the innovation errors. Proceeding as in the near-stationary case, we define the process $(z_t)_{1 \leqslant t \leqslant n}$ by

$$z_t = \theta_n z_{t-1} + v_t, \quad \theta_n = -1 + \frac{c}{k_n}, \; c < 0,$$

where $z_0 = o_p(k_n^{1/2})$ and $(v_t)_{1 \leqslant t \leqslant n}$ is a sequence of i.i.d. $(0, \sigma^2)$ random variables with distribution function F_{v_1} satisfying the symmetry condition of Remark 3.3(iii). Letting $\varepsilon_t = (-1)^t v_t$, $x_t = (-1)^t z_t$ in (7) we obtain that ε_t is an i.i.d. F_{v_1} sequence and x_t is a near-explosive moderate deviations from unity process with root $\rho_n = 1 - c/k_n$, $c < 0$. Since

$$\sum_{t=1}^{n} z_{t-1} v_t = \sum_{t=1}^{n} x_{t-1} \varepsilon_t, \quad \sum_{t=1}^{n} z_{t-1}^2 = \sum_{t=1}^{n} x_{t-1}^2.$$

Theorem 4.3 yields

$$\frac{k_n(-\theta_n)^n}{-2c}(\hat{\theta}_n - \theta_n) = \frac{k_n \rho_n^n}{-2c}(\hat{\rho}_n - \rho_n) \implies C.$$

5. Discussion

We have seen in Sections 3 and 4 that the study of moderate deviations from a unit root provides a useful insight into the discontinuities between stationary, unit root and explosive autoregressions. By investigating the asymptotic behavior of the intermediate cases, Theorems 3.2 and 4.3 describe explicitly the transition from stationary to unit root and explosive asymptotics, with normalizations varying continuously on $(n^{1/2}, n) \cup (n, \rho^n)$, where ρ is an explosive root $\rho = 1 + c$ with $c > 0$. This section addresses the issue of whether the bridging of the asymptotics between these different types of autoregression is complete at the boundary cases. Using the convenient parametrization $\rho_n = 1 + c/n^\alpha$ of Remark 3.3(i), we investigate whether taking the limit as $\alpha \to 0, 1$ produces not only the correct rates of convergence, but also the correct limiting distributions. As the following discussion will show, this is only true for the boundary case $\alpha \to 0$.

The convergence rates of Theorem 3.2 bridge those for unit root or local to unity processes and those that apply under stationarity. Part (c) with $k_n = n^\alpha$ becomes

$$n^{1/2+\alpha/2}(\hat{\rho}_n - \rho_n) \implies \mathrm{N}(0, -2c), \tag{12}$$

with a convergence rate ranging over $(n^{1/2}, n)$ for $\alpha \in (0, 1)$. However, when $\alpha \to 0$, (12) becomes $\sqrt{n}(\hat{\rho}_n - \rho_n) \Longrightarrow \mathrm{N}(0, -2c)$, whereas the correct stationary result when $\rho = 1 + c$ is $\sqrt{n}(\hat{\rho}_n - \rho) \Longrightarrow \mathrm{N}(0, -2c - c^2)$. Thus, part (c) of Theorem 3.2 as it stands overestimates the variance of $\hat{\rho}_n$ in the boundary case where $\alpha = 0$. Continuity at this boundary can be achieved for parts (a)–(c) through replacement of c by $c + c^2/2n^\alpha$ (or more generally by $c + c^2/2k_n$) without affecting the asymptotic results of Theorem 3.2. For the boundary $\alpha = 1$, a law of large numbers no longer applies for the normalized sample variance so the

Lindeberg type argument leading to the Gaussian limit of part (c) is not valid. The asymptotic behavior of autoregressions at this boundary is determined by the invariance principle $n^{-1/2}y_{\lfloor np\rfloor}\Longrightarrow J_c(p)$ on $D[0,1]$—see Phillips (1987)—which gives rise to the non-Gaussian local to unity limit result

$$n(\hat{\rho}_n-\rho_n)\Longrightarrow\frac{\int_0^1 J_c(s)\,\mathrm{d}B(s)}{\int_0^1 J_c(s)^2\,\mathrm{d}s},\tag{13}$$

where $B(\cdot)$ is standard Brownian motion and $J_c(s)=\int_0^s \mathrm{e}^{c(s-r)}\,\mathrm{d}B(r)$ is an Ornstein–Uhlenbeck process.

When $c>0$, part (b) of Theorem 4.3 becomes $n^{\alpha}\rho_n^{-n}(\hat{\rho}_n-\rho_n)\Longrightarrow C$ when $k_n=n^{\alpha}$, with a convergence rate that takes values on (n,ρ^n) as α ranges from 1 to 0. Since $\rho=1+c$ is the autoregressive root of an explosive AR(1) process when $\alpha=0$, there is a discontinuity due to the discrepancy between $1-\rho_n^2=-2c/n^{\alpha}+\mathrm{O}(n^{-2\alpha})$ when $\alpha\in(0,1)$ and $1-\rho^2=2c+c^2$ when $\alpha=0$. As in the near-stationary case, continuity can be achieved through replacement of c by $c+c^2/2n^{\alpha}$ (or by $c+c^2/2k_n$) without affecting Theorem 4.3. However, when $\alpha=1$, part (b) of Theorem 4.3 is replaced by the local to unity limit theory (13). Thus, continuity is achieved at the outside boundaries with the stationary and explosive case asymptotics, but not at the inside boundaries with the conventional local to unity asymptotics.

Finally, we give a brief summary of how departures from i.i.d. errors affect the limit theory. When $c>0$, the signal of the moderately explosive regressor y_{t-1} is strong enough to allow for serially correlated errors without affecting the Cauchy limit theory of Theorem 4.3. In subsequent work, Phillips and Magdalinos (2005) extend Theorem 4.3 for moderately explosive processes with linear process errors $u_t=\sum_{j=0}^{\infty}c_j\varepsilon_{t-j}$, where ε_t is a sequence of i.i.d. $(0,\sigma^2)$ random variables and c_j is a sequence of constants satisfying $\sum_{j=1}^{\infty}j|c_j|<\infty$. Relaxing the independence assumption in the near-stationary case is more complicated. Giraitis and Phillips (2006) prove a version of Theorem 3.2 for martingale difference errors with constant conditional variance $\mathrm{E}_{\mathscr{F}_{nt-1}}(u_t^2)=\sigma^2$ for all t. Serial correlation in the errors, however, induces an asymptotic bias for $\hat{\rho}_n$ and contributes to the variance of the Gaussian limiting distribution. For linear process innovations $u_t=\sum_{j=0}^{\infty}c_j\varepsilon_{t-j}$ as above with $\mathrm{E}\varepsilon_1^4<\infty$, Phillips and Magdalinos (2005) derive an expression for the asymptotic bias of $\hat{\rho}_n$ and provide formulae for the asymptotic variance of the normalized and centered serial correlation coefficient.

Notation

$:=$	definitional equality
$\mathrm{E}_{\mathscr{F}}(\cdot)$	conditional expectation $\mathrm{E}(\cdot\vert\mathscr{F})$
$\mathrm{P}_{\mathscr{F}}(\cdot)$	conditional expectation $\mathrm{P}(\cdot\vert\mathscr{F})$
$\mathscr{F}_{nt}$	$:=\sigma(y_0,u_1,\ldots,u_t)$
$\mathbf{1}\{\cdot\}$	indicator function
$\mathrm{o_p}(1)$	tends to zero in probability
$\mathrm{o_{a.s.}}(1)$	tends to zero almost surely
$\lfloor\cdot\rfloor$	integer part
$\longrightarrow_{\mathrm{a.s.}}$	almost sure convergence
$\longrightarrow_{\mathrm{p}}$	convergence in probability
$\longrightarrow_{L_{\mathrm{p}}}$	convergence in L_{p} norm
$\Longrightarrow$	weak convergence

$\equiv, =_{\mathrm{d}}$	distributional equivalence
$\sim_{\mathrm{d}}$	asymptotically distributed as

Acknowledgements

The authors thank Peter Robinson and three referees for helpful comments and suggestions. A first draft of the paper was written in April, 2003. Phillips thanks the NSF for partial research support under Grant nos. SES-00-92509 and SES 042254. Magdalinos thanks the EPSRC and the Onassis Foundation for scholarship support.

Appendix A. Technical appendix and proofs

Proposition A.1.

(a) *For each* $c<0$, $\rho_n^n = \mathrm{o}(k_n/n)$ *as* $n \to \infty$.
(b) *For each* $c>0$, $\rho_n^{-n} = \mathrm{o}(k_n/n)$ *as* $n \to \infty$.

Proof. For part (a), using the asymptotic equivalence $\log(1+x) = x + \mathrm{O}(x^2)$ as $x \to 0$, we obtain, as $n \to \infty$

$$\begin{aligned}
\log\left(\frac{n}{k_n}\rho_n^n\right) &= n \log \rho_n + \log \frac{n}{k_n} \\
&= n \log\left(1 + \frac{c}{k_n}\right) + \log \frac{n}{k_n} \\
&= n\left[\frac{c}{k_n} + \mathrm{O}\left(\frac{1}{k_n^2}\right)\right] + \log \frac{n}{k_n} \\
&= \frac{cn}{k_n}\left[1 + \frac{1}{c}\frac{\log(n/k_n)}{n/k_n} + \mathrm{O}\left(\frac{1}{k_n}\right)\right] \\
&= \frac{cn}{k_n}[1 + \mathrm{o}(1)].
\end{aligned}$$

Thus, since $cn/k_n \to -\infty$, $(n/k_n)\rho_n^n = \exp\{(cn/k_n)[1+\mathrm{o}(1)]\} = \mathrm{o}(1)$ as $n \to \infty$ and part (a) follows. An identical argument establishes part (b) for $c>0$. □

Proof of Lemma 3.1. For part (a), we have for each $s \in [0,1]$

$$\begin{aligned}
\frac{1}{n}y_{\lfloor ns\rfloor}^2 &= \frac{1}{n}\left(\rho_n^{\lfloor ns\rfloor}y_0 + \sum_{j=1}^{\lfloor ns\rfloor}\rho_n^{\lfloor ns\rfloor - j}u_j\right)^2 \leqslant \frac{2}{n}\rho_n^{2\lfloor ns\rfloor}y_0^2 + \frac{2}{n}\left(\sum_{j=1}^{\lfloor ns\rfloor}\rho_n^{\lfloor ns\rfloor - j}u_j\right)^2 \\
&= \frac{2}{n}\left(\sum_{j=1}^{\lfloor ns\rfloor}\rho_n^{\lfloor ns\rfloor - j}u_j\right)^2 + \mathrm{o_p}\left(\frac{k_n}{n}\right)
\end{aligned}$$

since $y_0 = \mathrm{o_p}(\sqrt{k_n})$. Also, $n^{-1}(\sum_{j=1}^{\lfloor ns\rfloor}\rho_n^{\lfloor ns\rfloor - j}u_j)^2 \to 0$ in L_1 since

$$\frac{1}{n}\mathrm{E}\left(\sum_{j=1}^{\lfloor ns\rfloor}\rho_n^{\lfloor ns\rfloor - j}u_j\right)^2 = \frac{\sigma^2}{n}\sum_{j=1}^{\lfloor ns\rfloor}\rho_n^{2(\lfloor ns\rfloor - j)} = \mathrm{O}\left(\frac{k_n}{n}\right)$$

for each $s \in [0,1]$. For part (b), we have

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{\lfloor ns \rfloor} y_{t-1}u_t &= \frac{1}{n}\sum_{t=1}^{\lfloor ns \rfloor}\left(\rho_n^{t-1}y_0 + \sum_{j=1}^{t-1}\rho_n^{t-1-j}u_j\right)u_t \\ &= \frac{y_0}{\sqrt{k_n}}\frac{\sqrt{k_n}}{n}\sum_{t=1}^{\lfloor ns \rfloor}\rho_n^{t-1}u_t + \frac{1}{n}\sum_{t=1}^{\lfloor ns \rfloor}\left(\sum_{j=1}^{t-1}\rho_n^{t-1-j}u_j\right)u_t \\ &= \frac{1}{n}\sum_{t=1}^{\lfloor ns \rfloor}\left(\sum_{j=1}^{t-1}\rho_n^{t-1-j}u_j\right)u_t + o_p(1)\end{aligned}$$

since $E((\sqrt{k_n}/n)\sum_{t=1}^{\lfloor ns \rfloor}\rho_n^{t-1}u_t)^2 = O(k_n/n)$. Also, since $(u_t)_{1\leqslant t\leqslant n}$ is an independent sequence we obtain

$$\begin{aligned}E\left[\frac{1}{n}\sum_{t=1}^{\lfloor ns \rfloor}\left(\sum_{j=1}^{t-1}\rho_n^{t-1-j}u_j\right)u_t\right]^2 &= \frac{1}{n^2}\sum_{t=1}^{\lfloor ns \rfloor}E\left[\left(\sum_{j=1}^{t-1}\rho_n^{t-1-j}u_j\right)^2\right]Eu_t^2 \\ &= \frac{\sigma^4}{n^2}\sum_{t=1}^{\lfloor ns \rfloor}\sum_{j=1}^{t-1}\rho_n^{2(t-1-j)} \\ &= \frac{\sigma^4}{n^2}(1-\rho_n^2)\left\{\lfloor ns \rfloor - \sum_{t=1}^{\lfloor ns \rfloor}\rho_n^{2(t-1)}\right\} = O\left(\frac{k_n}{n}\right)\end{aligned}$$

as $n \to \infty$ for each $s \in [0,1]$. This completes the proof of the lemma. □

Proof of (5). We can write, for each $\eta > 0$

$$\begin{aligned}&\sum_{t=1}^{n} E_{\mathscr{F}_{nt-1}}(\xi_{nt}^2\mathbf{1}\{|\xi_{nt}|>\eta\}) \\ &\quad = \frac{1}{nk_n}\sum_{t=1}^{n} y_{t-1}^2 E_{\mathscr{F}_{nt-1}}\left(u_t^2\mathbf{1}\left\{|y_{t-1}u_t|>\eta\sqrt{nk_n}\right\}\right) \\ &\quad \leqslant \max_{1\leqslant t\leqslant n} E_{\mathscr{F}_{nt-1}}\left(u_t^2\mathbf{1}\left\{|y_{t-1}u_t|>\eta\sqrt{nk_n}\right\}\right)\frac{1}{nk_n}\sum_{t=1}^{n} y_{t-1}^2.\end{aligned}$$

By (4), $(nk_n)^{-1}\sum_{t=1}^{n} y_{t-1}^2 = O_p(1)$, so (5) will follow from

$$\max_{1\leqslant t\leqslant n} E_{\mathscr{F}_{nt-1}}\left(u_t^2\mathbf{1}\left\{|y_{t-1}u_t|>\eta\sqrt{nk_n}\right\}\right) \longrightarrow_p 0 \quad \text{as } n\to\infty \tag{14}$$

for each $\eta > 0$. Applying the Hölder and Chebyshev inequalities, we obtain, for some $\delta > 0$

$$\begin{aligned}&E_{\mathscr{F}_{nt-1}}(u_t^2\mathbf{1}\{|y_{t-1}u_t|>\eta\sqrt{nk_n}\}) \\ &\quad \leqslant E_{\mathscr{F}_{nt-1}}^{2/(2+\delta)}(|u_t|^{2+\delta})P_{\mathscr{F}_{nt-1}}^{\delta/(2+\delta)}\left\{|y_{t-1}u_t|>\eta\sqrt{nk_n}\right\} \\ &\quad \leqslant (E|u_1|^{2+\delta})^{2/(2+\delta)}\left(\frac{E_{\mathscr{F}_{nt-1}}\{y_{t-1}^2u_t^2\}}{\eta^2 nk_n}\right)^{\delta/(2+\delta)} \\ &\quad = (E|u_1|^{2+\delta})^{2/(2+\delta)}\left(\frac{\sigma^2}{\eta^2}\right)^{\delta/(2+\delta)}\left(\frac{y_{t-1}^2}{nk_n}\right)^{\delta/(2+\delta)},\end{aligned}$$

for each $t \in \{1, \dots, n\}$. Thus, since $\mathrm{E}|u_1|^{2+\delta} < \infty$,

$$\max_{1 \leqslant t \leqslant n} \frac{y_{t-1}^2}{nk_n} \longrightarrow_{\mathrm{p}} 0 \quad \text{as } n \to \infty \tag{15}$$

is sufficient for (14). To show (15) we employ an argument similar to that in Aldous (1978).[2] For $m \in \{1, \dots, n\}$, define the sets

$$B_{n,m} := \bigcap_{j=1}^{m} \left\{ \omega : \left| \frac{1}{nk_n} \sum_{t=1}^{\lfloor n(j/m) \rfloor} y_{t-1}^2(\omega) - \frac{j}{m} \frac{\sigma^2}{-2c} \right| \leqslant \frac{1}{m} \right\}.$$

For each m, (3) implies that $\mathrm{P}(B_{n,m}) \to 1$ as $n \to \infty$. Next, note that

$$\max_{1 \leqslant t \leqslant n} \frac{y_{t-1}^2}{nk_n} \leqslant \frac{1}{nk_n} \sup_{s \in [0,1]} \sum_{t=\lfloor ns \rfloor + 1}^{\lfloor n(s+(1/m)) \rfloor} y_{t-1}^2.$$

For given $s \in [0,1]$ choose $j \in \{1, \dots, m\}$ so that $s \in [(j-1)/m, j/m]$. Then, for each $s \in [0,1]$, $\omega \in B_{n,m}$ implies

$$\begin{aligned}
\frac{1}{nk_n} \sum_{t=\lfloor ns \rfloor + 1}^{\lfloor n(s+1/m) \rfloor} y_{t-1}^2 &\leqslant \frac{1}{nk_n} \sum_{t=\lfloor n(j-1)/m \rfloor + 1}^{\lfloor n(j+1)/m \rfloor} y_{t-1}^2 \\
&= \left(\frac{1}{nk_n} \sum_{t=1}^{\lfloor n(j+1)/m \rfloor} y_{t-1}^2 - \frac{j+1}{m} \frac{\sigma^2}{-2c} \right) \\
&\quad - \left(\frac{1}{nk_n} \sum_{t=1}^{\lfloor n(j-1)/m \rfloor} y_{t-1}^2 - \frac{j-1}{m} \frac{\sigma^2}{-2c} \right) + \frac{\sigma^2}{-2c} \frac{2}{m} \\
&\leqslant \frac{2}{m} + \frac{\sigma^2}{-2c} \frac{2}{m} = \left(1 + \frac{\sigma^2}{-2c} \right) \frac{2}{m}.
\end{aligned}$$

Thus, for any $m \in \mathbb{N}$

$$\lim_{n \to \infty} \mathrm{P} \left\{ \max_{1 \leqslant t \leqslant n} \frac{y_{t-1}^2}{nk_n} \leqslant \left(1 + \frac{\sigma^2}{-2c} \right) \frac{2}{m} \right\} \geqslant \lim_{n \to \infty} \mathrm{P}(B_{n,m}) = 1,$$

showing (15). □

Proof of Lemma 4.1. We can write

$$\frac{\rho_n^{-n}}{k_n} \sum_{t=1}^{n} \sum_{j=t}^{n} \rho_n^{t-j-1} u_j u_t = \frac{\rho_n^{-n-1}}{k_n} \sum_{t=1}^{n} u_t^2 + \frac{\rho_n^{-n}}{k_n} \sum_{t=1}^{n} \left(\sum_{j=t+1}^{n} \rho_n^{t-j-1} u_j \right) u_t. \tag{16}$$

The first term on the right of (16) converges to 0 in L_1 since

$$\frac{\rho_n^{-n-1}}{k_n} \sum_{t=1}^{n} \mathrm{E} u_t^2 = \mathrm{O}\left(\rho_n^{-n} \frac{n}{k_n} \right) = \mathrm{o}(1) \quad \text{as } n \to \infty,$$

[2] We are grateful to a referee for bringing this argument to our attention.

by Proposition A.1(b). For the second term on the right of (16), note that the sequence $(\sum_{j=t+1}^{n}\rho_n^{t-j-1}u_j)u_t$ is uncorrelated for $t\in\{1,\dots,n\}$, which implies that

$$\begin{aligned}\mathrm{E}\left[\frac{\rho_n^{-n}}{k_n}\sum_{t=1}^{n}\left(\sum_{j=t+1}^{n}\rho_n^{t-j-1}u_j\right)u_t\right]^2 &= \frac{\sigma^2\rho_n^{-2n}}{k_n^2}\sum_{t=1}^{n}\mathrm{E}\left(\sum_{j=t+1}^{n}\rho_n^{t-j-1}u_j\right)^2\\ &= \frac{\sigma^4\rho_n^{-2n}}{k_n^2}\sum_{t=1}^{n}\sum_{j=t+1}^{n}\rho_n^{2(t-j-1)}\\ &= \frac{\sigma^4\rho_n^{-2n-2}}{k_n^2}\frac{n(1-\rho_n^{-2n})}{\rho_n^2-1}\\ &= \mathrm{O}\left(\frac{\rho_n^{-2n}}{k_n^2}nk_n\right)=\mathrm{O}\left(\rho_n^{-2n}\frac{n}{k_n}\right)=\mathrm{o}(1),\end{aligned}$$

as $n\to\infty$, by Proposition A.1(b). □

Proof of Lemma 4.2. By the Cramér–Wold device (e.g. Kallenberg, 2002, Corollary 5.5), it is sufficient to show that

$$aX_n+bY_n \implies aX+bY \quad \text{for all } a,b\in\mathbb{R}, \tag{17}$$

where X and Y are independent $\mathrm{N}(0,\sigma^2/2c)$ random variables. If Z is an $\mathrm{N}(0,(a^2+b^2)\sigma^2/2c)$ random variable, $aX+bY=_dZ$, so $aX_n+bY_n\Longrightarrow Z$ for all $a,b\in\mathbb{R}$ is sufficient for (17). We can write $aX_n+bY_n=\sum_{j=1}^{n}\zeta_{nj}$, where the array

$$\zeta_{nj}:=\frac{1}{\sqrt{k_n}}\{a\rho_n^{-j}+b\rho_n^{-(n-j)-1}\}u_j, \quad 1\leqslant j\leqslant n,$$

consists of independent non-identically distributed random variables. Hence, weak convergence to a Gaussian random variable can be derived as a consequence of the Lindeberg–Feller central limit theorem (see e.g. Kallenberg, 2002, Theorem 5.12). The variance of $\sum_{j=1}^{n}\zeta_{nj}$ is given by

$$\begin{aligned}\mathrm{E}\left(\sum_{j=1}^{n}\zeta_{nj}\right)^2 &= \frac{\sigma^2}{k_n}\sum_{j=1}^{n}\{a\rho_n^{-j}+b\rho_n^{-(n-j)-1}\}^2\\ &= \frac{\sigma^2}{k_n}\left\{a^2\sum_{j=1}^{n}\rho_n^{-2j}+b^2\sum_{j=1}^{n}\rho_n^{-2(n-j)-2}+2ab\rho_n^{-n-1}n\right\}\\ &= \sigma^2\left\{\frac{a^2}{k_n}\sum_{j=1}^{n}\rho_n^{-2j}+\frac{b^2}{k_n}\sum_{j=1}^{n}\rho_n^{-2(n-j)-2}\right\}+\mathrm{O}\left(\rho_n^{-n}\frac{n}{k_n}\right)\\ &= \frac{(a^2+b^2)\sigma^2}{2c}+\mathrm{o}(1) \quad \text{as } n\to\infty,\end{aligned}$$

by Proposition A.1. Next, note that

$$\frac{2}{k_n}\sum_{j=1}^{n}[a^2\rho_n^{-2j}+b^2\rho_n^{-2(n-j)-2}]\to\frac{a^2+b^2}{c} \quad \text{as } n\to\infty, \tag{18}$$

implying that the left side of (18) is uniformly bounded by a constant $K\in(0,\infty)$. By using the inequality $(x+y)^2\leqslant 2(x^2+y^2)$ the Lindeberg condition can be written, for each $\eta>0$

and $a, b \in \mathbb{R}$, as

$$\begin{aligned}
&\sum_{j=1}^{n} \mathrm{E}(\zeta_{nj}^2 \mathbf{1}\{|\zeta_{nj}| > \eta\}) \\
&\quad = \frac{1}{k_n} \sum_{j=1}^{n} [a\rho_n^{-j} + b\rho_n^{-(n-j)-1}]^2 \mathrm{E}(u_j^2 \mathbf{1}\{|a\rho_n^{-j} + b\rho_n^{-(n-j)-1}||u_j| > \eta\sqrt{k_n}\}) \\
&\quad \leqslant \frac{2}{k_n} \sum_{j=1}^{n} [a^2\rho_n^{-2j} + b^2\rho_n^{-2(n-j)-2}] \mathrm{E}[u_j^2 \mathbf{1}\{(a\rho_n^{-j} + b\rho_n^{-(n-j)-1})^2 u_j^2 > \eta^2 k_n\}] \\
&\quad \leqslant K \max_{1 \leqslant j \leqslant n} \mathrm{E}[u_1^2 \mathbf{1}\{(a\rho_n^{-j} + b\rho_n^{-(n-j)-1})^2 u_1^2 > \eta^2 k_n\}] \\
&\quad \leqslant K \max_{1 \leqslant j \leqslant n} \mathrm{E}[u_1^2 \mathbf{1}\{2(a^2\rho_n^{-2j} + b^2\rho_n^{-2(n-j)-2}) u_1^2 > \eta^2 k_n\}] \\
&\quad \leqslant K\mathrm{E}[u_1^2 \mathbf{1}\{2(a^2 + b^2) u_1^2 > \eta^2 k_n\}] \\
&\quad = K\mathrm{E}\left[u_1^2 \mathbf{1}\left\{u_1^2 > \frac{\eta^2}{2(a^2 + b^2)} k_n\right\}\right] \\
&\quad = \mathrm{o}(1) \quad \text{as } n \to \infty,
\end{aligned}$$

by integrability of u_1^2. Thus, $aX_n + bY_n \Longrightarrow Z$ for all $a, b \in \mathbb{R}$, establishing (17). □

References

Aldous, D., 1978. Stopping times and tightness. Annals of Probability 6, 335–340.

Anderson, T.W., 1959. On asymptotic distributions of estimates of parameters of stochastic difference equations. Annals of Mathematical Statistics 30, 676–687.

Basawa, I.V., Brockwell, P.J., 1984. Asymptotic conditional inference for regular nonergodic models with an application to autoregressive processes. Annals of Statistics 12, 161–171.

Chan, N.H., Wei, C.Z., 1987. Asymptotic inference for nearly nonstationary AR(1) processes. Annals of Statistics 15, 1050–1063.

Giraitis, L., Phillips, P.C.B., 2006. Uniform limit theory for stationary autoregression. Journal of Time Series Analysis, forthcoming.

Kallenberg, O., 2002. Foundations of Modern Probability. Springer, Berlin.

Park, J.Y., 2003. Weak unit roots. Mimeographed, 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., Magdalinos, T., 2005. Limit theory for moderate deviations from a unit root under weak dependence. Cowles Foundation Discussion Paper No. 1471, Yale University.

Pollard, D., 1984. Convergence of Stochastic Processes. Springer, Berlin.

White, J.S., 1958. The limiting distribution of the serial correlation coefficient in the explosive case. Annals of Mathematical Statistics 29, 1188–1197.