ESTIMATION AND INFERENCE WITH NEAR UNIT ROOTS

By

Peter C. B. Phillips

October 2021

COWLES FOUNDATION DISCUSSION PAPER NO. 2304

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

http://cowles.yale.edu/

Estimation and Inference with Near Unit Roots *

Peter C. B. Phillips
Yale University, University of Auckland,
Singapore Management University & University of Southampton

September 25, 2021

Abstract

New methods are developed for identifying, estimating and performing inference with nonstationary time series that have autoregressive roots near unity. The approach subsumes unit root (UR), local unit root (LUR), mildly integrated (MI) and mildly explosive (ME) specifications in the new model formulation. It is shown how a new parameterization involving a localizing rate sequence that characterizes departures from unity can be consistently estimated in all cases. Simple pivotal limit distributions that enable valid inference about the form and degree of nonstationarity apply for MI and ME specifications and new limit theory holds in UR and LUR cases. Normalizing and variance stabilizing properties of the new parameterization are explored. Simulations are reported that reveal some of the advantages of this alternative formulation of nonstationary time series. A housing market application of the methods is conducted that distinguishes the differing forms of house price behavior in Australian state capital cities over the past decade.

Keywords: Cauchy limit distribution, Local to unity, Localizing rate sequence, Mild integration, Mildly explosive process, Unit root.

JEL classification: C22

1 Introduction

While empirical research makes heavy use of persistent time series asymptotics for modeling nonstationary data it is usually recognized that it is often too restrictive, although certainly convenient, to insist that autoregressive roots be precisely unity. In consequence, much research has been done on time series with local to unit roots (LURs) or near-integrated processes

*This paper is a four decadal sequel to Phillips (1987a). Some preliminary findings were reported in 2011 in a draft paper with a different title (Phillips, 2011) that was never completed. The present paper completes that earlier analysis, studies identification issues, formulates a new localizing rate sequence, and provides limit theory, inferential procedures, simulations, and an empirical application. Computations were performed in Matlab. Support is acknowledged from the NSF under Grant Nos. SES-09 56687 and SES-18 50860, and a Kelly Fellowship at the University of Auckland.

following early work in the 1980s on the development of LUR asymptotics (Chan and Wei, 1987; Phillips, 1987b, 1988) for models with long run autoregressive coefficients of the form $\theta_n = 1 + \frac{c}{n}$ where c is an unknown localizing coefficient and n is the sample size. In the LUR model, the parameter c is identified but is not consistently estimable. Methods of inference concerning c have been suggested (Stock, 1991; Hansen, 1999) and used in applications but have known limitations (Mikusheva, 2007, 2012; Phillips, 2014) and face challenges in extension to practical model settings and multivariate models.

More recent attention has focussed on mildly integrated (MI) and mildly explosive (ME) time series for which the long run autoregressive coefficient has the form $\theta_n = 1 + \frac{c}{n^\alpha}$ where $\alpha \in (0, 1)$ is an unknown localizing rate and c is an unknown constant scale coefficient. Models with this formulation of θ_n offer alternatives closer to the stationary and explosive regions and have opened up new robust estimation possibilities and new options for inference. Such models deliver nonstationary alternatives to the random wandering behavior associated with LUR processes and help to deliver connectivity between stationary and nonstationary asymptotics (Phillips and Magdalinos, 2007a,b, hereafter, PM) and (Giraitis and Phillips, 2006; Phillips et al., 2010), in addition to long memory processes with the (nonstationary) fractional parameter $d = \frac{1}{2}$ (Duffy and Kasparis, 2021). A particular advantage of MI time series is the simple mechanism they provide for constructing endogeneously generated instruments (known as IVX) that validate standard methods of inference in cointegrating and predictive regressions (Phillips and Magdalinos, 2009; Kostakis et al., 2015), thereby overcoming ubiquitous problems of size distortion and non-pivotal inference that are induced by the presence of LUR regressors (Elliott, 1998; Phillips, 2015). In addition, ME time series have opened up new opportunities for estimation and inference concerning explosive phenomena and exuberance in financial and real estate markets, providing methods of real time detection of bubble behavior that have proved useful in practical work on the diagnosis of prevailing market conditions by market participants, banks and regulators (Phillips et al., 2011, 2015a,b).

The present paper contributes to this literature in several ways. First, issues of identification and consistent estimation of the localizing coefficients $\{\alpha, c\}$ in MI and ME models are explored. Contrary to popular thinking it is shown that it is possible to consistently estimate the rate parameter α that controls (in conjunction with the scale parameter) the widths of the mildly integrated and mildly explosive regions as the sample size $n \to \infty$. Consistent estimation of the rate parameter also applies in the LUR case where $\alpha = 1$ although the limit distribution is different. In addition, whereas the localizing parameters $\{\alpha, c\}$ are not separately identified in finite samples, pseudo-identification does hold asymptotically. In particular, unlike LUR models where consistent estimation of the localizing coefficient c is not possible, consistent estimation of pseudo-true values of c, notably $c_* \mp 1$, is attainable in MI and ME systems.[1]

The primary contribution of the paper is to propose an equivalent model formulation in terms of a single localizing rate sequence which is identifiable, consistently estimable, and has

[1]Recent work by Lin and Tu (2020) correctly pointed to the difficulties in the estimation of c but did not observe that pseudo-true values c_* of c are consistently estimable.

pivotal limit theory that enables inference in these models. Some particular advantages of this new formulation stand out. It delivers a simple uni-parameter measure that quantifies departures from unit root and local unit root specifications. This parameter sequence can be estimated at a regularly varying power rate for MI time series and an exponential rate for ME series. Local unit root as well as unit root specifications occur at natural boundary values of the new parameter sequence; and the limit theory for the proposed nonlinear rate estimator belongs to a stable normal class in contrast to the nonstandard limit theory for the autoregressive coefficient estimator. Finally, simulations show good finite sample performance in estimation and inference for this alternative formulation, although performance deteriorates due to a slower logarithmic convergence rate when the localizing rate coefficient approaches unity. Consistency still holds in this case, in contrast to the well-known asymptotic theory for the LUR case in which the localizing coefficient c, as distinct from α, is not consistently estimable.

The paper is organized as follows. Section 2 studies issues of identification, estimation, and inference concerning MI autoregressions, introduces the equivalent uni-parameter sequence representation of these processes, and provides limit theory for estimation of the localizing coefficients. Mildly explosive processes are considered in Section 3. Simulations are reported in Section 4 and an empirical illustration of the methods to housing markets is given in Section 5. Some concluding remarks are given in Section 6. Proofs are in the Appendix.

2 Mildly Integrated Autoregression

2.1 Model and Properties

For simplicity of exposition we consider the prototypical mildly integrated autoregression

$$X_t = \theta_n X_{t-1} + u_t, \quad t = 1, ..., n, \tag{2.1}$$

$$\theta_n = 1 + \frac{c}{n^\alpha}, \quad c < 0, \quad \alpha \in (0,1), \tag{2.2}$$

with initialization $X_0 = o_p(n^\alpha)$ and innovations satisfying

$$u_t \sim_{iid} (0, \sigma^2) \text{ with } \mathbb{E}|u_t|^\nu < \infty \text{ for some } \nu > \frac{2}{\alpha}. \tag{2.3}$$

With some minor modification the methodology given here accommodates systems like (2.1) with weak dependent errors u_t such as those considered in PM(2007b). The extension in the case $c < 0$ is discussed in Section 2.4. In the mildly explosive case where $c > 0$ the limit theory given in Section 3 holds without any modification for weak dependence in the innovations.

Partial sums $S_t := \sum_{i=1}^t u_i$ of u_t satisfy the functional law

$$B_{n^\alpha}(\cdot) := \frac{S_{\lfloor n^\alpha \cdot \rfloor}}{n^{\alpha/2}} = \frac{\sum_{i=1}^{\lfloor n^\alpha \cdot \rfloor} u_i}{n^{\alpha/2}} \rightsquigarrow B(\cdot), \tag{2.4}$$

where $B(\cdot)$ is Brownian motion with variance σ^2. Least squares estimation of θ_n gives $\hat{\theta}_n = \sum_{t=1}^n X_t X_{t-1} / \sum_{t=1}^n X_{t-1}^2$, which carries information about both localizing parameters (α, c) in (2.2). The estimate $\hat{\theta}_n$ is known to be consistent for θ_n with the following limit theory.

Lemma 2.1. *PM(2007a, Theorem 3.2) For model (2.1) with $\theta_n = 1 + c/n^a$, $c < 0$, $\alpha \in (0,1)$ and u_t satisfying (2.3) we have: (i)*

$$n^{\frac{1+\alpha}{2}} \left(\hat{\theta}_n - \theta_n \right) \rightsquigarrow \mathcal{N}(0, -2c) =: \xi_c, \tag{2.5}$$

(ii) $\frac{1}{n^{1+\alpha}} \sum_{t=1}^n X_{t-1}^2 \to_p \frac{\sigma^2}{-2c}$, and (iii) $\frac{1}{n_*^{(1+\alpha)/2}} \sum_{t=1}^n x_{t-1} u_t \rightsquigarrow \mathcal{N}\left(0, \frac{\sigma^4}{-2c}\right)$, as $n \to \infty$.

Result (i) shows that as $c \to 0$ the asymptotic variance tends to zero, matching the fact that the convergence rate rises to n in the unit root case as $\theta_n \to 1$. Similarly, (ii) and (iii) show that the regressor signal $\sum_{t=1}^n X_{t-1}^2$ diverges at a faster rate than $n^{1+\alpha}$ as $c \to 0$. Thus, the value of c and its proximity to zero influence asymptotic behavior in a material way that relates to the localizing rate coefficient α. More directly in terms of the localizing rate parameter α itself, the convergence rate $n^{\frac{1+\alpha}{2}} \to n$ as $\alpha \to 1$. It is therefore evident that the two localizing coefficients (c, α) play joint and related roles in determining both the finite sample and limit behavior of $\hat{\theta}_n$. This interactive role of the unknown parameters (c, α) affects the capacity to identify these parameters.

2.2 Local parameter identification failure

The functional dependence of the autoregressive coefficient $\theta_n = 1 + \frac{c}{n^\alpha} =: \theta_n(c, \alpha)$ on the two localizing parameters (c, α) reveals a fundamental identification uncertainty in the specification. Whereas θ_n is itself identified in finite samples, these two parameters are not separately identified in finite samples even under additional conditions such as the sign of c. This is explained by the fact that the generating mechanism (2.1) implies the moving average representation $X_t = \sum_{j=0}^{t-1} \theta_n^j u_{t-j} + \theta_n^t X_0$ so that the joint distribution of $\{X_j\}_{j=1}^n$ for any given sample size n depends on the pair (c, α) only through θ_n and, hence, only through the ratio $\frac{c}{n^\alpha}$. Given n, the (stationary) local-to-unit root condition $0 < \theta_n < 1$ requires $-n^\alpha < c < 0$ or $0 < |c| < n^\alpha$, so that $-\infty < \log|c| < \alpha \log n$. It follows that the likelihood of $\{X_j\}_{j=1}^n$ is equivalently defined by the following simpler uni-parameter autoregressive parameter specification

$$\theta_n = \theta_n(\gamma_n) := 1 - \frac{1}{n^{\gamma_n}}, \tag{2.6}$$

with

$$\gamma_n = \gamma_n(c, \alpha) := \alpha - \frac{\log|c|}{\log n} \in (0,1), \tag{2.7}$$

because $-1/n^{\gamma_n} = -n^{\frac{\log|c|}{\log n}}/n^\alpha = -|c|/n^\alpha = c/n^\alpha$ for $-n^\alpha < c < -\frac{1}{n^{1-\alpha}}$. The upper and lower limits of c used in the definition of γ_n ensure that $\gamma_n \in (0,1)$. In particular, $\gamma_n \to 1$ as

c tends to the upper limit $-\frac{1}{n^{1-\alpha}}$, and $\gamma_n \to 0$ as c tends to its lower limit $-n^{\alpha}$. For values of $c \geq -\frac{1}{n^{1-\alpha}}$, we have $1/n^{\gamma_n} = |c|/n^{\alpha} \leq 1/n$. Thus, for $c \geq -\frac{1}{n^{1-\alpha}}$ and, as $c \to 0$ from below, the rate parameter $\gamma_n \geq 1$ and the autoregressive coefficient θ_n is local to unity when $c = -1/n^{1-\alpha}$ or closer to unity when $-1/n^{1-\alpha} < c \leq 0$. As will be explained in what follows, the limit theory developed here for the estimation of γ_n accommodates this possibility.

The upshot is the model given by equations (2.1) and (2.2) may equivalently be defined by (2.1) in conjunction with the specification $\theta_n = 1 - \frac{1}{n^{\gamma_n}}$ where $\gamma_n = \gamma_n(c, \alpha) \in (0, 1)$ is given in (2.7). This new formulation of θ_n is a single parameter specification $\theta_n = \theta_n(\gamma_n)$ that involves the rate parameter sequence γ_n. Importantly, given n, the value of $\gamma_n(c, \alpha)$ is determined by the pair (c, α) and we have the following correspondences at the limits of the domain of definition of (c, α) and the key point $c = -1$ in (2.7) where $\gamma_n = \alpha$:

$$\begin{aligned}
&\lim_{c \to -n^{-(1-\alpha)}} (\gamma_n, \theta_n) \to (1, 1 - \frac{1}{n}), \qquad \lim_{c \to -1} (\gamma_n, \theta_n) \to (\alpha, 1 - \frac{1}{n^{\alpha}}), \quad \lim_{c \to -n^{\alpha}} (\gamma_n, \theta_n) \to (0, 0), \\
&\lim_{(c,\alpha) \to (-1,1)} (\gamma_n, \theta_n) \to (1, 1 - \frac{1}{n}), \qquad \lim_{\alpha \to 1} (\gamma_n, \theta_n) \to (1 - \frac{\log |c|}{\log n}, 1 - \frac{|c|}{n}), \\
&\lim_{\alpha \to 0} (\gamma_n, \theta_n) \to \left(-\frac{\log |c|}{\log n}, 1 - |c| \right). \qquad (2.8)
\end{aligned}$$

These relations show that stationary, MI, and LUR models are all captured in the single parameter specification. But the correspondence is evidently not $1:1$. For example, $\gamma_n(-\frac{1}{n^{1-\alpha}}, \alpha) = \gamma_n(-1, 1) = 1$ both yield the same autoregressive coefficient $\theta_n = 1 - \frac{1}{n}$. More generally we have equivalence whenever $\frac{c_1}{n^{\alpha_1}} = \frac{c_2}{n^{\alpha_2}}$, or $c_1 = c_2 n^{\alpha_1 - \alpha_2}$. So if $\alpha_1 = \alpha_2 + b$ then $c_1 = c_2 n^b$ will ensure $\theta_n = 1 + \frac{c_1}{n^{\alpha_1}} = 1 + \frac{c_2 n^b}{n^{\alpha_1}} = 1 + \frac{c_2}{n^{\alpha_2}}$. Then $\theta_n = 1 - \frac{1}{n^{\gamma_n}}$ with $\gamma_n = \alpha_1 - \frac{\log |c_1|}{\log n} = \alpha_2 - \frac{\log |c_2|}{\log n}$ so that $\gamma_n(c_1, \alpha_1) = \gamma_n(c_2, \alpha_2)$, and there is lack of identification in finite samples. Note that in the equivalence $\gamma_n(c_1, \alpha_1) = \gamma_n(c_2, \alpha_2)$, c_1 depends on n. But this is not an issue in finite samples where, for any given n, the allowable range $-n^{\alpha_1} < c_1 < -\frac{1}{n^{1-\alpha_1}}$ for the localizing coefficient c_1 is satisfied and ensures that $\gamma_n \in (0, 1)$ as required for the MI specification of the model. In contrast to this finite sample failure of identification of the pair (c, α), the localizing rate parameter γ_n is identified, just as the autoregressive coefficient θ_n is itself identified.

Attempts to estimate the twin parameter specification (c, α) reveal the presence of the localizing coefficient uncertainty implicit in the dual parameter specification of the MI model (2.1) and (2.2). The impact of the identification uncertainty about (c, α) becomes apparent in the asymptotic theory because separate estimation of the rate parameter α and the localizing coefficient c lead as $n \to \infty$ to the pseudo-parameters $(c_* = -1, \gamma_n)$ in the uni-parameter specification (2.7). Thus, the point of equivalence $\gamma_n = \alpha$ that arises when c takes on the value -1 in (2.7) turns out to be an important pseudo-true limit value for the localizing rate coefficient γ_n in the limit theory. As we proceed to show, it turns out that there is asymptotic identification and consistent estimation of both the pseudo-true value $c_* = -1$ and the specific rate sequence γ_n when $n \to \infty$. Thus, the essential element in the MI model is the implied localizing rate parameter γ_n in the specification $\theta_n = 1 - \frac{1}{n^{\gamma_n}}$. Boundary values for the parameter sequence γ_n are also relevant because when $\gamma_n = 1 - \frac{\log|c|}{\log n}$ the model merges with the LUR class where $\theta_n = 1 - \frac{|c|}{n} = 1 + \frac{c}{n}$ with $c < 0$. In particular: when $\gamma_n = 1$ (i.e., $c = -1$ in $\gamma_n = 1 - \frac{\log|c|}{\log n}$) we have the LUR special case $\theta_n = 1 - \frac{1}{n}$; when $\gamma_n > 1$ we have the LUR coefficient $\theta_n = 1 + \frac{c}{n}$ with $-1 < c < 0$; and $\gamma_n < 1$ captures LUR models with $c < -1$. Further, when $\gamma_n \to \infty$, $\theta_n \to 1$, delivering the UR model. Similar correspondences apply on the right side of the UR model with $\theta_n = 1 + \frac{1}{n^{\gamma_n}} > 1$ and $\gamma_n = 1 - \frac{\log|c|}{\log n}$, as discussed in Section 3. These representations become important in interpreting the results of applied research, as shown later in the empirical illustration.

Limit theory for the autoregressive coefficient estimate $\hat{\theta}_n$ is given in (2.5). On a suitably expanded probability space the convergence (2.5) holds in probability and in this expanded space we can write

$$\hat{\theta}_n = \theta_n + \frac{\xi_c}{n^{\frac{1+\alpha}{2}}} \left\{1 + o_p(1)\right\}. \tag{2.9}$$

It is convenient to work within this expanded space and we often do so subsequently without specific mention. The random component $n^{-\frac{1+\alpha}{2}} \xi_c$ in (2.9) depends on both parameters (c, α). But in view of the equivalent representation (2.7) we may write

$$\frac{\xi_c}{n^{\frac{1+\alpha}{2}}} =_d \mathcal{N}\left(0, \frac{-2c}{n^{1+\alpha}}\right) = \mathcal{N}\left(0, \frac{2}{n^{1+\gamma_n}}\right), \tag{2.10}$$

because $1/n^{\gamma_n} = -c/n^{\alpha}$ and $n^{\frac{1+\alpha}{2}} |c|^{-1/2} = n^{\frac{1+\gamma_n}{2}}$ as shown in (A-1) in the Appendix. The limit theory of Lemma 2.1 may therefore be rewritten in the following simpler form that does not explicitly depend on c although the uni-parameter γ_n implicitly carries the effects of the value of c and its asymptotic behavior when c itself depends on n.

Lemma 2.2. *Under the conditions of Lemma 2.1 and defining the rate parameter sequence $\gamma_n = \gamma_n(c, \alpha) = \alpha - \frac{\log|c|}{\log n} \in (0, 1)$ as in (3.11) we have: (i)*

$$n^{\frac{1+\gamma_n}{2}} \left(\hat{\theta}_n - \theta_n\right) \rightsquigarrow \mathcal{N}(0, 2) =: \xi, \tag{2.11}$$

(ii) $\frac{1}{n^{1+\gamma_n}} \sum_{t=1}^{n} X_{t-1}^2 \to_p \frac{\sigma^2}{2}$, *and (iii)* $\frac{1}{n^{(1+\gamma_n)/2}} \sum_{t=1}^{n} x_{t-1} u_t \rightsquigarrow \mathcal{N}\left(0, \frac{\sigma^4}{2}\right)$, *as* $n \to \infty$.

2.3 Parameter estimation

We now consider methods of estimating the localizing parameters (c, α) and associated uni-parameter sequence γ_n under various conditions concerning the true value of autoregressive parameter sequence θ_n and its proximity to unity.

(a) Estimation of the rate coefficient α

In view of the representation (2.6), define $\hat{A}_n = \hat{\theta}_n - 1$ and construct the nonlinear rate estimator $\hat{\alpha} = -\frac{\log|\hat{A}_n|}{\log n}$, for which the following limit theory holds.

Theorem 2.1. *(i) For model (2.1) and (2.2) with fixed $c<0$, fixed $\alpha\in(0,1)$ and u_t satisfying (2.3) as $n\to\infty$ we have $\hat{\alpha}\to_p\alpha$ and*

$$n^{\frac{1-\alpha}{2}}\log n\left\{\hat{\alpha}-\alpha+\frac{\log|c|}{\log n}\right\}\rightsquigarrow\frac{\xi_c}{-c}=_d\mathcal{N}\left(0,\frac{2}{|c|}\right),\tag{2.12}$$

where ξ_c is given in (2.5).
(ii) For model (2.1) with fixed $\alpha\in(0,1)$, $\theta_n=1+\frac{c_n}{n^\alpha}$ and $c_n=-\frac{1}{n^{1-\alpha-\delta}}$ and $0<\delta\le1-\alpha$ as $n\to\infty$ we have $\hat{\alpha}\to_p 1-\delta$ and

$$n^{\delta/2}\log n\left\{\hat{\alpha}-(1-\delta)\right\}\rightsquigarrow\xi=_d\mathcal{N}(0,2)\,.\tag{2.13}$$

(iii) For model (2.1) with $\alpha=1$, $\theta_n=1+\frac{c}{n}$ and fixed $c\in(-\infty,\infty)$ as $n\to\infty$ we have $\hat{\alpha}\to_p 1$ and

$$(\log n)\left\{\hat{\alpha}-1\right\}\rightsquigarrow-\log|c+\xi_{J_c}|,\tag{2.14}$$

where $\xi_{J_c}=\int_0^1 J_c dW/\int_0^1 J_c^2$, $J_c(r)=\int_0^1 e^{c(r-s)}dW(s)$ is a standard linear diffusion and W is standard Brownian motion.

Result (2.12) shows that the estimator $\hat{\alpha}$ is consistent with convergence rate $O(n^{\frac{1-\alpha}{2}}\log n)$ but with a first order asymptotic bias $-\frac{\log|c|}{\log n}$ and, upon centering and scaling, $\hat{\alpha}$ has the limiting normal distribution $\mathcal{N}\left(0,\frac{2}{|c|}\right)$. When $c=-1$ the asymptotic bias term disappears and the limit distribution is simply $\mathcal{N}(0,2)$. In general, the asymptotic variance depends on the localizing coefficient c and diverges as $\frac{2}{|c|}\to\infty$ when $c\to0$, indicative that the convergence rate changes when c is dependent on the sample size n. In particular, in case (ii) where $c_n=-\frac{1}{n^{1-\alpha-\delta}}\to0$ or even closer to zero[2] as $n\to\infty$, the convergence rate drops to $n^{\delta/2}\log n$ and further approaches $\log n$ when $\delta\to0$. The asymptotics in (i) and (ii) hold when $\alpha\in(0,1)$ and $c<0$ but fail in the unit root case where $c=0$ and $\theta_n=1$ or where $c=c_n\to0$ as fast or faster than as $c_n=-\frac{1}{n^{1-\alpha}}$. In that case the autoregressive coefficient is either $\theta_n=1-\frac{1}{n}$ or closer to unity with $\theta_n=1+o(\frac{1}{n})$. In such near integrated or closer-to-unit root cases, (2.9) fails. Result (2.14) in (iii) then shows that the estimator $\hat{\alpha}$ is consistent for the unit exponent with convergence rate $O(\log n)$, as anticipated from (2.12) when $\alpha\to1$, but with a limit distribution determined by the quantity $\log|c+\xi_{J_c}|$ which involves bias and dependence on the nonstandard LUR limit distribution $\xi_{J_c}=\int_0^1 J_c dW/\int_0^1 J_c^2$. The finite sample behavior of the estimator $\hat{\alpha}$ in these various cases is explored in relation to these asymptotics later in the paper.

From Lemma 2.1 we have $\frac{1}{n^{1+\alpha}}\sum_{t=1}^n X_{t-1}^2\to_p\frac{\sigma^2}{-2c}$, so that $\log\left(\frac{1}{n}\sum_{t=1}^n X_{t-1}^2\right)-\alpha\log n\to_p$

[2] For instance, if $c_n=-\frac{L_n}{n^{1-\alpha}}$ for some slowly varying function for which $L_n\to\infty$ as $n\to\infty$ then a version of (ii) continues to hold but with convergence rate $L_n^{1/2}\log n$. See Remark 2.1 below.

$\log\left(\frac{\sigma^2}{-c}\right)$. Simple transformation then leads to another consistent estimator of α, viz.,

$$\tilde{\alpha} = \frac{\log\left(\frac{1}{n}\sum_{t=1}^{n} X_{t-1}^2\right)}{\log n} \to_p \alpha,$$

for which $(\log n)(\tilde{\alpha} - \alpha) \to_p \log\left(\frac{\sigma^2}{-c}\right)$. So $\tilde{\alpha}$ has a logarithmic convergence rate when $\alpha \in (0,1)$ in contrast to the regularly varying rate of $\hat{\alpha}$. When $\alpha = 1$, $\frac{1}{n^2}\sum_{t=1}^{n} X_{t-1}^2 \rightsquigarrow \int_0^1 J_c^2$ and then

$$\tilde{\alpha} = \frac{\log\left(\frac{1}{n}\sum_{t=1}^{n} X_{t-1}^2\right)}{\log n} \to_p 1,$$

so that $\tilde{\alpha}$ is again consistent with a logarithmic convergence rate in the LUR case but with limit theory $(\log n)(\tilde{\alpha} - 1) \rightsquigarrow \log\left(\int_0^1 J_c^2\right)$, indicative of a random first order bias effect. In cases (i) and (ii) of Theorem 2.1 both estimators $\hat{\alpha}$ and $\tilde{\alpha}$ have limit theory that depends on the unknown localizing coefficient c and both suffer first order asymptotic bias effects. Inference about the rate parameter α using these results for either $\hat{\alpha}$ or $\tilde{\alpha}$ therefore depends on estimation or knowledge of c. As discussed below, this problem is averted by exploiting the pseudo-true value $c_* = -1$ of c in the mildly integrated case and using a uni-parameter representation of θ_n.

Remark 2.1. *If the specification $c_n = \frac{g}{n^{1-\alpha-\delta}}$, with additional localizing constant coefficient $g < 0$, is used in Theorem 2.1 (ii) we again have $\hat{\alpha} \to_p 1 - \delta$ but in place of (2.13) we have the limit theory*

$$n^{\delta/2} \log n \left\{\hat{\alpha} - (1-\delta) + \frac{\log|g|}{\log n}\right\} \rightsquigarrow \frac{\xi_g}{-g} =_d \mathcal{N}\left(0, \frac{2}{|g|}\right), \tag{2.15}$$

analogous to (2.12). Further, if $c_n = -\frac{L_n}{n^{1-\alpha}}$ for some slowly varying (SV) function L_n at infinity, then $\hat{\alpha} \to_p 1$ and

$$L_n^{1/2} \log n \left\{\hat{\alpha} - 1 + \frac{\log L_n}{\log n}\right\} \rightsquigarrow \xi =_d \mathcal{N}(0,2), \tag{2.16}$$

which is shown in (A-10) in the proof of Theorem 2.1(ii)-SV Extension, which is given in the Appendix. The Gaussian limit theory is then maintained up to and including a scale SV factor $L_n \to \infty$ times the $\log n$ rate. So this rate is faster than the logarithmic rate $\log n$ that applies in Theorem 2.1 (iii) when $\alpha = 1$ and $\theta_n = 1 + \frac{c}{n}$ is local to unity. Result (2.16) provides a localizing rate estimator limit theory for the MI case with autoregressive coefficient $\theta_n = 1 + \frac{c}{k_n}$ considered in PM(2007a) where in the present case $k_n = -c\frac{n}{L_n} = o(n)$ so that $\theta_n = 1 - \frac{L_n}{n}$. Observe that even though the bias term $-\frac{\log L_n}{\log n} \to 0$ as $n \to \infty$ in this near unit root case, the bias is nonnegligible asymptotically in the limit distribution given the convergence rate $L_n^{1/2} \log n$ in (2.16).

(b) Estimation of the localizing coefficient c

The asymptotic behavior of the regression signal $\sum_{t=1}^{n} X_{t-1}^2 \sim_a \frac{\sigma^2 n^{1+\alpha}}{-2c}$ suggests the following estimator of the localizing coefficient c

$$\hat{c} = -\frac{n^{1+\hat{\alpha}}\hat{\sigma}^2}{2\sum_{t=1}^{n} X_{t-1}^2} = -\frac{n^{\hat{\alpha}-\alpha}\hat{\sigma}^2}{\frac{2}{n^{1+\alpha}}\sum_{t=1}^{n} X_{t-1}^2}, \tag{2.17}$$

where $\hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n}\hat{u}_t^2 \to_p \sigma^2$ with $\hat{u}_t = X_t - \hat{\theta}_n X_{t-1}$ and $\hat{\alpha} = -\frac{\log|\hat{A}_n|}{\log n}$ is as before. The estimator $\hat{c}$ is consistent for the pseudo-true localizing coefficient $c_* = -1$ in the MI case.

Theorem 2.2. *Under the conditions of Theorem 2.1 and as $n \to \infty$ $\hat{c} \to_p -1$.*

Theorems 2.1 and 2.2 show that, whereas the twin localizing parameters (c, α) in an MI system are not themselves identified in finite samples, consistent estimation is possible for the rate parameter $\hat{\alpha}$ and the pseudo-true value $c_* = -1$. However, there is non-negligible bias in the limit distribution of the rate estimator α and c itself is not consistently estimable. Instead the pseudo-true value $c_* = -1$ is the limiting value of $\hat{c}$ and leads directly to the uni-parameter sequence γ_n for which the identified representation of the MI autoregressive sequence $\theta_n = 1 - \frac{1}{n^{\gamma_n}}$ holds. These asymptotic findings reveal that attempts to estimate consistently the twin localizing parameters (c, α) in an MI system lead, in effect, to consistent estimation of the uni-parameter sequence γ_n in the identified representation $\theta_n = 1 - \frac{1}{n^{\gamma_n}}$, pointing to the advantage of using this representation of mild integration in an autoregression.

(c) Uni-parameter estimation

An alternative approach to rate estimation is to take advantage of the uni-parameter representation of the MI model based on the rate parameter sequence γ_n in (2.7). With this reformulation of the model, we can define the rate estimator

$$\hat{\gamma}_n = -\frac{\log\left|\hat{A}_n\right|}{\log n} = -\frac{\log|\hat{\theta}_n - 1|}{\log n}, \tag{2.18}$$

which has precisely the same form as the estimator $\hat{\alpha}$. But by virtue of the definition of the sequence γ_n in (2.7) $\hat{\gamma}_n$ takes advantage of the presence of the first order asymptotic bias $\frac{\log|c|}{\log n}$ in the estimator $\hat{\alpha}$. In doing so, $\hat{\gamma}_n$ is a natural estimator for the localizing rate sequence γ_n because the uni-parameter formulation of the autoregressive coefficient $\theta_n = 1 - \frac{1}{n^{\gamma_n}}$ gives $A_n = \theta_n - 1 = -\frac{1}{n^{\gamma_n}}$ and so $\frac{\log|A_n|}{\log n} = -\gamma_n$, which leads to the estimator $\hat{\gamma}_n$ in (2.18). The limit theory for $\hat{\gamma}_n$ follows directly from Theorem 2.1 and is formalized in the following result.

Corollary 2.1. *Under the conditions of Theorem 2.1 (i) with $c < 0$, $\alpha \in (0,1)$ and $\gamma_n = \alpha - \frac{\log|c|}{\log n}$, or Theorem 2.1 (ii) with $\theta_n = 1 + \frac{c_n}{n^{\alpha}} = 1 - \frac{1}{n^{\gamma_n}}$, $c_n = -\frac{1}{n^{1-\alpha-\delta}}$, $\gamma_n = 1 - \delta$, and*

$0 < \delta \leq 1 - \alpha$ *as* $n \to \infty$*, we have*

$$n^{\frac{1-\gamma_n}{2}} \log n \left(\hat{\gamma}_n - \gamma_n\right) \rightsquigarrow \xi =_d \mathcal{N}\left(0, 2\right). \tag{2.19}$$

In the SV case discussed in Remark 2.1 where $\theta_n = 1 + \frac{c_n}{n^\alpha} = 1 - \frac{L_n}{n}$ *with* $c_n = -\frac{L_n}{n^{1-\alpha}}$ *and* $L_n \to \infty$ *is slowly varying at infinity, we have, as* $n \to \infty$

$$L_n^{1/2} \log n \left(\hat{\gamma}_n - \gamma_n\right) \rightsquigarrow \xi, \tag{2.20}$$

where $\gamma_n = 1 - \frac{\log L_n}{\log n}$*. Finally, when* $\theta_n = 1 + \frac{c}{n}$ *we have, as* $n \to \infty$*,*

$$(\log n)\left\{\hat{\gamma_n} - 1\right\} \rightsquigarrow -\log\left|c + \xi_{J_c}\right|. \tag{2.21}$$

In (2.19) the Gaussian limit theory of $\hat{\gamma}_n$ has no first order asymptotic bias and has the regularly varying convergence rate $n^{\frac{1-\gamma_n}{2}} \log n$. The asymptotic variance in (2.19) is constant and independent on the sequence γ_n. So $\hat{\gamma}_n$ is a variance stabilizing transformation of $\hat{\theta}_n$, at least up to the rate of convergence.

In (2.20) the limit theory of $\hat{\gamma}_n$ also has no first order bias, is again Gaussian, and involves the convergence rate $L_n^{1/2} \log n$. Observe that when $\gamma_n = 1 - \frac{\log L_n}{\log n}$ it follows that $n^{(1-\gamma_n)/2} = n^{\frac{1}{2}\frac{\log L_n}{\log n}} = L_n^{1/2}$, corresponding to the additional SV factor beyond the $\log n$ rate in (2.20). This linkage means that (2.20) is subsumed within (2.19), which assists inference as explained below. Importantly, result (2.20) also continues to hold for SV functions L_n, such as $L_n = \log n$, for which $\frac{\log L_n}{\log n} \to 0$ and $\gamma_n = 1 - \frac{\log L_n}{\log n} \to 1$ as $n \to \infty$, thereby reaching the lower boundary of the LUR case. In such cases (2.20) implies the Gaussian asymptotic approximation

$$\hat{\gamma}_n \sim_a \mathcal{N}\left(1 - \frac{\log L_n}{\log n}, \frac{2}{L_n \log^2 n}\right) \sim_a \mathcal{N}\left(1, \frac{2}{L_n \log^2 n}\right). \tag{2.22}$$

The last member of (2.22) is a (crude) asymptotic approximation that ignores the relative magnitude of the mean component $\frac{\log L_n}{\log n}$ in relation to the asymptotic standard deviation $\frac{\sqrt(2)}{L_n^{1/2} \log n}$. The comparative merits of the crude Gaussian approximation (2.22) and the LUR approximation (2.21) in the case $\gamma_n = 1$ are explored later in simulations.

Transformations of the usual stationary autoregression limit theory with fixed $\theta < 1$ provide simple heuristics for the limit theory given in (2.19). In particular, using the representation $\theta_n = 1 - \frac{1}{n^{\gamma_n}}$, and applying standard fixed θ asymptotics suggests

$$\sqrt{n}\left(\hat{\theta}_n - \theta_n\right) \Rightarrow \mathcal{N}\left(0, 1 - \theta_n^2\right) = \mathcal{N}\left(0, \frac{2}{n^{\gamma_n}} - \frac{1}{n^{2\gamma_n}}\right).$$

Rescaling gives the asymptotic approximation $n^{(1+\gamma_n)/2}\left(\hat{\theta}_n - \theta_n\right) \sim_a \mathcal{N}\left(0, 2\right)$ for $\theta_n = 1 - \frac{1}{n^{\gamma_n}}$ in the neighbourhood of unity. Transforming $\theta_n \longmapsto \gamma_n = -\frac{\log(1-\theta_n)}{\log n}$ and using the delta method with derivative $\frac{d\theta_n}{d\gamma_n} = \frac{\log n}{n^{\gamma_n}}$ then delivers $n^{(1-\gamma_n)/2} \log n \left(\hat{\gamma}_n - \gamma_n\right) \sim_a \mathcal{N}\left(0, 2\right)$, matching (2.19). This argument becomes rigorous in Corollary 2.1 because the mildly integrated limit theory in Lemma 2.2 validates the asymptotic theory (2.11) for $\theta_n = 1 - \frac{1}{n^{\gamma_n}}$, which leads in turn to (2.19) above.

In the LUR case where $\theta_n = 1 + \frac{c}{n}$, based on the assumption that $\gamma_n = 1$, the nonlinear rate estimator $\hat{\gamma}_n \to_p 1$ and is consistent for the true value $\gamma_n = 1$, unlike the usual estimator of c in the LUR case, viz., $n(\hat{\theta}_n - 1)$. In this LUR case, the convergence rate of $\hat{\gamma}_n$ is logarithmic and the limit distribution is nonstandard, although as shown in the simulations in Fig. 5(b), the distribution of the nonlinear functional (2.21) is much closer in general appearance to a Gaussian distribution than the usual UR and LUR limit distributions. This feature is explained by the form of the limit distribution (2.21). In particular, using the linear diffusion equation $dJ_c = cJ_c dr + dW$, we have the integral representation $\int_0^1 J_c dJ_c = c \int_0^1 J_c^2 dr + \int_0^1 J_c dW$, from which it is easy to see that (2.21) can be written in the alternate form

$$(\log n)\left\{\hat{\gamma}_n - 1\right\} \rightsquigarrow -\log\left|\frac{\int_0^1 J_c dJ_c}{\int_0^1 J_c^2}\right| = -\log|\eta_{J_c}|, \tag{2.23}$$

where $\eta_{J_c} = \int_0^1 J_c dJ_c / \int_0^1 J_c^2$ is the continuous record estimator of the parameter c in the linear diffusion defining J_c – see Phillips (1987a, equation (32)). The limiting representation (2.23) may be interpreted as a transform of a continuous time serial correlation coefficient, having the form of a logarithmic 'variance stabilizing' transform.

In the UR case with $c = 0$ and $\theta_n = 1$ in (2.21), the limit theory of $\hat{\gamma}_n$ is pivotal and given by $(\log n)\left\{\hat{\gamma}_n - 1\right\} \rightsquigarrow -\log|\eta_W|$, with $\eta_W = \int_0^1 W dW / \int_0^1 W^2$ and W standard Brownian motion. This distribution can be used for testing under the null of a unit autoregressive root, with local power function determined by (2.23) with $c \neq 0$. But this approach is equivalent to standard unit root testing because there is a one to one relationship between the estimates $\hat{\theta}_n$ and $\hat{\gamma}_n$ and their asymptotic distributions under the null and the alternative. So, even though the respective asymptotics have different convergence rates and the limit distributions have very different forms, with that of $\hat{\gamma}_n$ being much more bell shaped than the usual UR distribution, they lead to precisely the same inferences. The approach that follows augments this existing testing regime in the pure UR case by using the limit theory for $\hat{\gamma}_n$ in MI and ME cases to construct confidence intervals for $\gamma_n < 1$ that allow for near unit roots on both the left and right sides of unity that approach the LUR boundary, as discussed above.

The limit distributions (2.19) and (2.20) are both $\mathcal{N}(0, 2)$ and do not depend on the localizing coefficient c. But in fact $\xi = \xi_{-1}$, corresponding to the pseudo-true value $c_* = -1$ of c and matching the specification of the localizing scale coefficient -1 in $\theta_n = 1 - \frac{1}{n^{\gamma_n}}$. The limit distribution $\mathcal{N}(0, 2)$ is conducive to inference. In particular, confidence interval (CI) construction for γ_n follows directly by use of the consistent estimate $\hat{\gamma}_n$ for calculation of the asymptotic standard error. Thus, when $\hat{\theta}_n < 1$, an asymptotic $100(1-\lambda)\%$ CI for γ_n can be constructed as

$$\hat{\gamma}_n \pm cv_\lambda \times \frac{\sqrt{2}}{n^{\frac{1-\hat{\gamma}_n}{2}} \log n}, \tag{2.24}$$

where $cv_\lambda = \Phi^{-1}(1-\lambda/2)$, using the standard normal cdf Φ. This confidence interval remains valid even when $\gamma_n = 1 - \frac{\log L_n}{\log n} \to 1$ as $n \to \infty$ because $L_n \to \infty$ and so

$$n^{\frac{1-\hat{\gamma}_n}{2}} \log n = n^{\frac{1-\gamma_n}{2}} \log n \times O_p\left(n^{-\frac{1}{2L_n^{1/2}\log n}}\right) \tag{2.25}$$

$$= n^{\frac{1-\gamma_n}{2}} \log n\left\{1 + O_p\left(\frac{1}{L_n^{1/2}}\right)\right\} \sim_a n^{\frac{1-\gamma_n}{2}} \log n. \tag{2.26}$$

The coverage probability and length of the interval (2.24) are explored later in the simulations. When γ_n is close to unity the intervals can be wide, as is to be expected from the convergence rate of $\hat{\gamma}_n$ and the fact that the asymptotic standard error $\sqrt{2}/(n^{\frac{1-\gamma_n}{2}} \log n)$ tends to zero at a near logarithmic rate in such cases. Nonetheless, useful inferences about γ_n are possible in practice with sample sizes around $n = 100$ when $\hat{\theta}_n < 1$ and considerably smaller sample sizes when $\hat{\theta}_n > 1$, as will be evident in the empirical illustration in Section 5.

(d) Variance stabilizing and normalizing properties of $\hat{\gamma}_n$

The rate estimator $\hat{\gamma}_n = -\frac{\log(1-\hat{\theta}_n)}{\log n}$ bears an interesting relationship to the well-known Fisher z transformation of the sample product-moment correlation coefficient r. Fisher (1921) discovered that, unlike r, the transformation $z(r) = \frac{1}{2}\log\left(\frac{1+r}{1-r}\right) = \tanh^{-1}(r)$ of r is approximately normal with variance that is stable over different values of the population correlation ρ,[3]. The following result shows that the rate estimator $\hat{\gamma}_n$ has asymptotically the same form as Fisher's z transformation, expressed in terms of the serial correlation coefficient $\hat{\theta}_n$ rather than the product-moment correlation.

Corollary 2.2. *Under the conditions of Corollary 2.1, as $n \to \infty$ we have:*

$$n^{\frac{1-\gamma_n}{2}} \log n\,(\hat{\gamma}_n - \gamma_n) = n^{\frac{1-\gamma_n}{2}}\left\{\log\frac{1+\hat{\theta}_n}{1-\hat{\theta}_n} - \log\frac{1+\hat{\theta}_n}{1-\hat{\theta}_n}\right\} + O\left(\frac{1}{n^{\gamma_n}}\right) \rightsquigarrow \mathcal{N}(0,2). \tag{2.27}$$

Note that the scale factor $\frac{1}{2}$ in the usual Fisher formula $\frac{1}{2}\log\frac{1+\hat{\theta}_n}{1-\hat{\theta}_n}$ does not appear here in the serial correlation case (2.27) because the asymptotic variance is 2 not 1. Moreover, the action of variance stabilization is more subtle in the present case. In particular in the case of

[3]If r is the sample product moment correlation of data $(X_i, Y_i)_{i=1}^n$ drawn independently from the same bivariate normal distribution with correlation ρ, then $z = \frac{1}{2}\log\left(\frac{1+r}{1-r}\right)$ is approximately $\mathcal{N}\left(\frac{1}{2}\log\left(\frac{1+\rho}{1-\rho}\right), \frac{1}{n}\right)$ thereby stabilizing the variance - see Fisher (1921); Hotelling (1953). The normalizing properties of the transform have been demonstrated through its skewness reduction and attenuating effects on the first order correction term in its Edgeworth expansion - see Winterbottom (1979); Konishi (1981). Thus, the Fisher transformation is normalizing and stabilizing for the product-moment correlation. Both these properties enhance inference.

fixed $\theta \in (0,1)$, if we consider the transformation $h(\hat{\theta}) = \log \frac{1+\hat{\theta}}{1-\hat{\theta}}$ then $h'(\theta) = \frac{2}{1-\theta^2}$ and

$$\sqrt{n}(h(\hat{\theta}) - h(\theta)) \sim_a h'(\theta)\sqrt{n}(\hat{\theta} - \theta) \rightsquigarrow \mathcal{N}\left(0, h'(\theta)^2(1-\theta^2)\right) = \mathcal{N}\left(0, \frac{4}{1-\theta^2}\right),$$

which is clearly not variance stabilizing. But in the neighborhood of unity with $\theta_n = 1 - \frac{1}{n^{\gamma_n}}$ we have $h'(\theta_n) = \frac{2}{1-\theta_n^2} = n^{\gamma_n}\left(1 + O\left(\frac{1}{n^{\gamma_n}}\right)\right)$. Then

$$\sqrt{n}(h(\hat{\theta_n}) - h(\theta_n)) \sim_a h'(\theta_n)\sqrt{n}(\hat{\theta_n} - \theta_n) \sim_a \mathcal{N}\left(0, \frac{4}{1-\theta_n^2}\right) \sim_a \mathcal{N}\left(0, 2n^{\gamma_n}\right),$$

and, upon rescaling by the factor $n^{-\gamma_n/2}$, we obtain

$$n^{\frac{1-\gamma_n}{2}}(h(\hat{\theta_n}) - h(\theta_n)) \sim_a n^{\frac{1-\gamma_n}{2}} \log n\,(\hat{\gamma}_n - \gamma_n) \rightsquigarrow \mathcal{N}(0,2). \tag{2.28}$$

So the Fisher transformation of the serial correlation coefficient $\hat{\theta}_n$ is variance stabilizing in the near unit root case, at least up to the convergence rate, as is the rate estimator $\hat{\gamma}_n$ from (2.19).

Interestingly for the case of fixed θ, it is known (Jenkins, 1954) that the appropriate variance stabilizing transform of the serial correlation coefficient $\hat{\theta}$ is the angular transform $h(\theta) = \sin^{-1}(\theta)$ not the Fisher transform. Indeed, with derivative $h'(\theta) = (1-\theta^2)^{-1/2}$, direct application of the delta method gives $\sqrt{n}(\sin^{-1}(\hat{\theta}) - \sin^{-1}(\theta)) \rightsquigarrow \mathcal{N}(0,1)$. But in the vicinity of unity the development of the limit theory changes. When $\theta_n = 1 - \frac{1}{n^{\gamma_n}}$ we have $\sin^{-1}(1 - \frac{1}{n^{\gamma_n}}) = \frac{\pi}{2} - \sqrt{2}\frac{1}{n^{\gamma_n/2}} + O(\frac{1}{n^{3\gamma_n/2}})$ as $n \to \infty$. Then $\sqrt{n}(\sin^{-1}(\hat{\theta}_n) - \sin^{-1}(\theta_n)) \sim_a -\sqrt{2n}\left(\frac{1}{n^{\hat{\gamma}_n/2}} - \frac{1}{n^{\gamma_n/2}}\right)$ and further application of the delta method reveals that

$$\sqrt{n}(\sin^{-1}(\hat{\theta}_n) - \sin^{-1}(\theta_n)) \sim_a \frac{n^{(1-\gamma_n)/2}\log n}{\sqrt{2}}(\hat{\gamma}_n - \gamma_n) \rightsquigarrow \mathcal{N}(0,1), \tag{2.29}$$

as $n \to \infty$ in view of the limit theory (2.19) in Corollary 2.1. It follows that the angular transform maintains its variance stabilizing property in the mildly integrated vicinity of unity and is asymptotically equivalent to the rate estimator $\hat{\gamma}_n$, at least up to the respective convergence rates. Thus, in contrast to the fixed θ case, when $\theta_n = 1 - \frac{1}{n^{\gamma_n}}$ is near unity, both the Fisher transformation and the angular transform coincide asymptotically and lead to variance stabilization according to (2.28) and (2.29); and both transformations have the same asymptotic theory as that of the rate coefficient $\hat{\gamma}_n$.

Unlike the product-moment correlation, there is very little literature dealing with normalizing transformations for the serial correlation coefficient. In the fixed $\theta \in (-1,1)$ case Phillips (1977) gave the Edgeworth expansion for the distribution of the serial correlation coefficient to order $O(n^{-1})$ and a subsequent working paper (Phillips, 1979) showed that the Fisher transformation removed the $O(\frac{1}{\sqrt{n}})$ skewness term in the Edgeworth expansion of $\hat{\theta}$. Phillips et al. (2010) developed expansions in the mildly integrated and mildly explosive cases, which smooth transitions to the near-stationary and near-explosive models from the local unit root case. Translating their expansion (Phillips et al., 2010, theorem 1) in the near-stationary case to the present notation with $\theta = 1 - \frac{1}{\gamma_n}$ we have

$$P\left(\frac{n^{(1+\gamma_n)/2}}{\sqrt{2}}\left(\hat{\theta}_n - \theta_n\right) < x\right) = \Phi\left(x\right) + \frac{1}{\sqrt{2}}\frac{1+x^2}{n^{(1-\gamma_n)/2}}\varphi\left(x\right) + O\left(\frac{1}{n^{1-\gamma_n}}\right), \tag{2.30}$$

where $\Phi\left(x\right)$ and $\varphi\left(x\right)$ are the cdf and density of the standard normal distribution. Expression (2.30) gives an asymptotic series expansion in powers of $\frac{1}{n^{(1-\gamma_n)/2}}$ rather than $\frac{1}{\sqrt{n}}$. In contrast to the Edgeworth expansion for fixed $\theta \in (0,1)$, where the correction term on the first order Gaussian asymptotics diverges as $\theta \to 1$,[4] the first order correction in (2.30) remains finite as $\gamma_n \to 1$, leading to a Gram-Charlier series representation of the limit distribution in the LUR model rather than an asymptotic series. Similar properties can be expected for the distribution of the rate estimator $\hat{\gamma}_n$ for which we have the limit theory

$$\frac{n^{(1-\gamma_n)/2}}{\sqrt{2}}\log n\left(\hat{\gamma}_n - \gamma_n\right) \sim_a \frac{n^{(1+\gamma_n)/2}}{\sqrt{2}}\left(\hat{\theta}_n - \theta_n\right) \rightsquigarrow \mathcal{N}\left(0,1\right),$$

corresponding to the fact that $\hat{\gamma}_n \to_p 1$ when $\gamma_n = 1$ and the limit distribution of $\log n\left(\hat{\gamma}_n - 1\right)$ is no longer Gaussian but given in (2.21). A detailed analysis of these expansions and representations is left for future work.

2.4 Weak dependent errors

The above theory extends to the MIR model (2.1)–(2.2) with weak dependent errors under the following condition.

Assumption LP *$u_t = C\left(L\right)\varepsilon_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}$, where $C(1) \neq 0$, $\omega^2 = \sigma^2 C(1)^2$, $\sum_{j=1}^{\infty} j\left|c_j\right| < \infty$ and $\varepsilon_t \sim_{iid} \left(0, \sigma^2\right)$ with $\mathbb{E}\left|\varepsilon_t\right|^{\nu} < \infty$ for some $\nu > \frac{2}{\alpha}$.*

Specifically, we have the following extension of Corollary 2.1.

Theorem 2.3. *(i) For model (2.1) and (2.2) with fixed $c < 0$, fixed $\alpha \in (1/3, 1)$ and u_t satisfying Assumption **LP** as $n \to \infty$ we have $\hat{\gamma}_n - \gamma_n \to_p 0$ and*

$$n^{\frac{1-\gamma_n}{2}}\log n\left\{\hat{\gamma}_n - \gamma_n + \frac{\log\varphi}{\log n}\right\} \rightsquigarrow \xi, \tag{2.31}$$

[4]The Edgeworth expansion in the fixed stationary θ case to $O(\frac{1}{\sqrt{n}})$ has the form

$$P\left(\frac{\sqrt{n}\left(\hat{\theta} - \theta\right)}{\sqrt{1-\theta^2}} < x\right) = \Phi\left(x\right) + \frac{\theta}{\sqrt{1-\theta^2}}\frac{1+x^2}{\sqrt{n}}\varphi\left(x\right) + O\left(\frac{1}{n}\right),$$

whose correction term $\frac{\theta}{\sqrt{1-\theta^2}}\frac{1+x^2}{\sqrt{n}}\varphi\left(x\right)$ on the standard Gaussian cdf $\Phi\left(x\right)$ diverges as $\theta \to 1$. This divergence signals the abrupt discontinuity in the asymptotic theory between the stationary and nonstationary cases of θ. As is evident in (2.30) the passage in the asymptotic theory via the parameter γ_n is far less abrupt as $\gamma_n \to 1$.

where $\xi =_d \mathcal{N}(0,2)$ *and* $\varphi = \frac{\sigma^2}{\omega^2}$.

(ii) Under these conditions, a modified version of case (ii) of Corollary 2.1 where $\theta_n = 1 + \frac{c_n}{n^\alpha} = 1 - \frac{1}{n^{\gamma_n}}$ *with* $c_n = -\frac{1}{n^{1-\alpha-\delta}}$, $\gamma_n = 1-\delta$ *and* $0 < \delta \le 1-\alpha$ *holds in which*

$$n^{\delta/2} \log n \left\{ \hat{\gamma_n} - \gamma_n + \frac{\log \varphi}{\log n} \right\} \rightsquigarrow \xi. \tag{2.32}$$

(iii) In the SV case where $\theta_n = 1 + \frac{c_n}{n^\alpha} = 1 - \frac{L_n}{n}$ *with* $c_n = -\frac{L_n}{n^{1-\alpha}}$ *and* $L_n \to \infty$ *is slowly varying at infinity, again under the above conditions we have as* $n \to \infty$

$$L_n^{1/2} \log n \left(\hat{\gamma_n} - \gamma_n + \frac{\log \varphi}{\log n} \right) \rightsquigarrow \xi, \tag{2.33}$$

where $\gamma_n = 1 - \frac{\log L_n}{\log n}$.

(iv) In the LUR case where $\theta_n = 1 + \frac{c}{n}$

$$(\log n) \{\hat{\gamma_n} - 1\} \rightsquigarrow -\log |c + \xi_{J_c}|. \tag{2.34}$$

The effect of weak dependence in the innovations in the MI model is to induce asymptotic bias in the estimator $\hat{\gamma_n}$. The bias $\frac{\log \varphi}{\log n}$ in (2.31) –(2.33) depends on the extent of the deviation φ from unity and hence the extent of the deviation of the long run variance ω^2 from the variance σ_u^2. There is no bias when the u_t are martingale differences and $\omega^2 = \sigma_u^2$, in which case $\log \varphi = 0$. The condition $\alpha \in (1/3, 1)$ in Theorem 2.3 ensures that the bias in (2.31) takes the simple form shown involving the parameter φ and is a consequence of PM(2007b, Theorem 4.2, equation (24)). The condition can be relaxed but has the advantage in the present context that it leads to a simple bias correction formula.

In fact, correction for the bias in $\hat{\gamma_n}$ in the presence of weak dependence can be achieved by a simple nonparametric serial correlation adjustment, analogous to the corrections employed in unit root tests such as the $\{Z_\alpha, Z_t\}$ tests (Phillips, 1987a). Define $\hat{\varphi} = \hat{\sigma}_u^2/\hat{\omega}^2$ where $\hat{\sigma}_u^2 = \frac{1}{n}\sum_{t=1}^n \hat{u}_t^2$ is the residual variance and $\hat{\omega}^2$ is a consistent HAC estimator of ω^2. The bias corrected estimator is $\tilde{\gamma_n} = \hat{\gamma_n} + \frac{\log \hat{\varphi}}{\log n}$, so that the estimation error is $\tilde{\gamma_n} - \gamma = \hat{\gamma_n} - \gamma + \frac{\log \hat{\varphi}}{\log n}$, for which we have

$$n^{\frac{1-\gamma_n}{2}} \log n \{\tilde{\gamma_n} - \gamma_n\} \rightsquigarrow \xi =_d \mathcal{N}(0,2), \tag{2.35}$$

in place of (2.31). The result holds for any consistent HAC estimator of ω^2 based on standard triangular or quadratic lag kernels, as shown in the proof of Theorem 2.3.

Confidence intervals (CI) for γ_n that are robust to weak dependence may be constructed using the bias corrected estimator $\tilde{\gamma_n}$ in place of $\hat{\gamma_n}$ in the earlier formula (2.24). In particular, when $\hat{\theta}_n < 1$, an asymptotic $100(1-\lambda)\%$ CI for γ_n is

$$\tilde{\gamma_n} \pm cv_\lambda \times \frac{\sqrt{2}}{n^{\frac{1-\tilde{\gamma}_n}{2}} \log n}, \tag{2.36}$$

where the critical value $cv_\lambda = \Phi^{-1}(1-\lambda/2)$ is determined as in (2.24).

3 Mildly Explosive Model

We use the generating mechanism (2.1) with autoregressive coefficient in the mildly explosive region $\theta_n = 1 + \frac{c}{n^\alpha}$, $c > 0$, where $\alpha \in (0,1)$. The limit theory for $\hat{\theta}_n$ was given in PM(2007a) and shown to hold in PM(2007b) with weakly dependent equation errors under standard linear process conditions (Phillips and Solo, 1992).

Lemma 3.1. *PM(2007a, Theorem 3.2) For model (2.1) with $\theta_n = 1 + c/n^\alpha$, $c > 0$, $\alpha \in (0,1)$, and u_t satisfying either (2.3) or the linear process condition **LP**, as $n \to \infty$ we have: (i)*

$$\frac{n^\alpha \theta_n^n}{2c}\left(\hat{\theta}_n - \theta_n\right) \rightsquigarrow \mathbb{C}, \quad \text{as } n \to \infty, \tag{3.1}$$

where $\mathbb{C}$ denotes a standard Cauchy variate; and (ii)

$$\frac{1}{n^\alpha} X_n \rightsquigarrow X(c), \quad \frac{\sum_{t=1}^n X_{t-1}^2}{\theta_n^{2n} n^{2\alpha}} \rightsquigarrow \frac{1}{2c} X(c)^2, \tag{3.2}$$

where $X(c) = \mathcal{N}\left(0, \frac{\omega^2}{2c}\right)$ and ω^2 is the long run variance of u_t or simply the variance when u_t satisfies (2.3).

Proceeding as in the mildly integrated case on a suitably expanded probability space where the convergence (3.1) holds in probability, we have

$$\hat{\theta}_n = \theta_n + \frac{1}{n^\alpha}\frac{2c}{\theta_n^n}\mathbb{C}\left\{1 + o_p(1)\right\} = 1 + \frac{c}{n^\alpha} + \frac{1}{n^\alpha}\frac{2c}{\theta_n^n}\mathbb{C}\left\{1 + o_p(1)\right\}. \tag{3.3}$$

As before, define $\hat{A}_n = \hat{\theta}_n - 1$ and the rate estimator $\hat{\alpha} = -\frac{\log|\hat{A}_n|}{\log n}$. The limit theory for $\hat{\alpha}$ in the mildly explosive case now follows directly from Lemma 3.1.

Theorem 3.1. *(i) For model (2.1) and (2.2) with fixed $c > 0$, under the conditions of Lemma 3.1 we have $\hat{\alpha} \to_p \alpha$ and*

$$\left(1 + \frac{c}{n^\alpha}\right)^n \log n \left\{\hat{\alpha} - \alpha + \frac{\log|c|}{\log n}\right\} \rightsquigarrow \zeta_c =_d 2c\mathbb{C}. \tag{3.4}$$

(ii) For model (2.1) with fixed $\alpha \in (0,1)$, $\theta_n = 1 + \frac{c_n}{n^\alpha}$ and $c_n = \frac{1}{n^{1-\alpha-\delta}}$ and $0 < \delta \leq 1 - \alpha$ as $n \to \infty$ we have $\hat{\alpha} \to_p 1 - \delta$ and

$$\left(1 + \frac{1}{n^{1-\delta}}\right)^n \log n \left\{\hat{\alpha} - (1-\delta)\right\} \rightsquigarrow \zeta =_d 2\mathbb{C}. \tag{3.5}$$

Noting that $\left(1+\frac{c}{n^\alpha}\right)^n = e^{cn^{1-\alpha}}\left(1-O(n^{1-2\alpha})\right) \sim_a e^{cn^{1-\alpha}}$ when $\alpha > \frac{1}{2}$, we can write (3.4) in this case as

$$e^{cn^{1-\alpha}} \log n \left\{ \hat{\alpha} - \alpha + \frac{\log |c|}{\log n} \right\} \rightsquigarrow 2c\mathbb{C}, \tag{3.6}$$

and (3.5) as

$$e^{n^\delta} \log n \left\{ \hat{\alpha} - (1-\delta) \right\} \rightsquigarrow 2\mathbb{C}, \tag{3.7}$$

when $\delta \leq 1-\alpha < 1/2$. Further, an extension of (ii) analogous to the MI case holds when $\theta_n = 1 + \frac{L_n}{n}$ differs from unity by a slowly varying function at infinity L_n. In this case we have $c_n = \frac{L_n}{n^{1-\alpha}} = \frac{n^{\frac{\log L_n}{\log n}}}{n^{1-\alpha}}$ so that

$$\theta_n = 1 + \frac{c_n}{n^\alpha} = 1 + \frac{1}{n^{1-\frac{\log L_n}{\log n}}} = 1 + \frac{L_n}{n}, \tag{3.8}$$

and the following limit theory holds when $L_n \to \infty$ as $n \to \infty$

$$\left(1 + \frac{1}{n^{1-\frac{\log L_n}{\log n}}}\right)^n \log n \left(\hat{\alpha} - \left\{ 1 - \frac{\log L_n}{\log n} \right\} \right) \rightsquigarrow 2\mathbb{C}, \tag{3.9}$$

as shown in (A-30) in the proof of the SV extension of Theorem 3.1(ii).

As in the MI case, the limit distributions (3.4) and (3.5) reveal non-negligible bias in the rate estimator $\hat{\alpha}$ of α when $c \neq 1$. Moreover, the pair (c, α) is not identifiable and the localizing scale parameter c is not consistently estimable, just as discussed earlier in the MI case. Instead, the pseudo-true value $c_* = 1$ is identifiable and consistently estimable.

Theorem 3.2. *Under the conditions of Theorem 3.1 and as $n \to \infty$*

$$\tilde{c} = \frac{1}{2} \frac{n^{\hat{\alpha}} X_n^2}{\sum_{t=1}^n X_{t-1}^2} \to_p 1. \tag{3.10}$$

The pseudo-true value $c_* = 1$ is the limiting value of $\tilde{c}$ in the ME case. As before in the MI case, attempts to estimate consistently the twin localizing parameters (c, α) lead, in effect, to consistent estimation of the uni-parameter sequence $\gamma_n = \alpha - \frac{\log c}{\log n}$. It is therefore convenient, as before in the MI case, to reparameterize the ME model so that the autoregressive coefficient is written as $\theta_n = 1 + \frac{1}{n^{\gamma_n}}$. With this parameterization the likelihood of $\{X_t\}_{t=1}^n$ relies on the identified uni-parameter sequence

$$\theta_n = 1 + \frac{1}{n^{\gamma_n}}, \quad \text{with } \gamma_n = \gamma_n(c, \alpha) = \alpha - \frac{\log c}{\log n} \in (0,1), \tag{3.11}$$

as $1/n^{\gamma_n} = n^{\frac{\log |c|}{\log n}}/n^\alpha = c/n^\alpha$ with $\frac{1}{n^{1-\alpha}} < c < n^\alpha$. The inequalities defining the range of c used in the definition of γ_n ensure that $\gamma_n \in (0,1)$ so that θ_n is a mildly explosive coefficient. When $0 \leq c \leq \frac{1}{n^{1-\alpha}}$, the rate parameter $\gamma_n \geq 1$, corresponding to a local-to-unity or closer to unity autoregressive coefficient $\theta_n = 1 + \frac{1}{n^{\gamma_n}} \leq 1 + \frac{1}{n}$.

Estimation of γ_n proceeds as in the MIR case by employing the same form as $\hat{\alpha}$, viz.,

$$\hat{\gamma}_n = -\frac{\log \left|\hat{A}_n\right|}{\log n} = -\frac{\log |\hat{\theta}_n - 1|}{\log n}. \tag{3.12}$$

The limit theory for $\hat{\gamma}_n$ follows the proof of Theorem 3.1(i) and (ii), including the SV extension of (ii) in which $\gamma_n = 1 - \frac{\log L_n}{\log n}$ for $\theta_n = 1 + \frac{L_n}{n}$, as given in the following result.

Corollary 3.1. *(i) Under the conditions of Theorem 3.1 (i) with $c_n > 0$, $\alpha \in (0,1)$ and $\gamma_n = \alpha - \frac{\log |c_n|}{\log n}$ we have, as $n \to \infty$,*

$$\left(1 + \frac{1}{n^{\gamma_n}}\right)^n \log n \left(\hat{\gamma}_n - \gamma_n\right) \rightsquigarrow \zeta =_d 2\mathbb{C}. \tag{3.13}$$

(ii) In the regularly varying case where $\theta_n = 1 + \frac{c_n}{n^\alpha} = 1 + \frac{L_n}{n}$ with $c_n = \frac{L_n}{n^{1-\alpha}}$ and $L_n \to \infty$ is slowly varying at infinity, we have $\gamma_n = 1 - \frac{\log L_n}{\log n}$ and, in place of (3.13),

$$\left(1 + \frac{L_n}{n}\right)^n \log n \left(\hat{\gamma}_n - \gamma_n\right) \rightsquigarrow \zeta =_d 2\mathbb{C}. \tag{3.14}$$

The limit distribution $2\mathbb{C}$ in (3.13) and (3.14) is conducive to inference, using the consistent estimate $\hat{\gamma}_n$ and the quantiles of the Cauchy distribution $\mathbb{C}$ to construct confidence intervals. When $\hat{\theta}_n > 1$, an asymptotic $100(1-\lambda)\%$ CI for γ_n in the ME case based on (3.13) is given by

$$\hat{\gamma}_n \pm cv_{\mathbb{C},\lambda} \times \frac{2}{\left(1 + \frac{1}{n^{\hat{\gamma}_n}}\right)^n \log n}, \tag{3.15}$$

where $cv_{\mathbb{C},\lambda} = \Phi_{\mathbb{C}}^{-1}(1 - \lambda/2)$ is the $1 - \lambda/2$ percentile of the standard Cauchy distribution with cdf $\Phi_{\mathbb{C}}$. The two sided 95% critical value of 12.706 when $\lambda = 0.025$ reflects the heavy tailed nature of the Cauchy distribution and contributes to widening the length of the confidence interval, especially when the rate parameter γ_n is close to unity, in which case the scale factor $\frac{2}{\left(1+\frac{1}{n^{\hat{\gamma}_n}}\right)^n \log n}$ in (3.15) is close to logarithmic. Use of the confidence interval in this ME case is of course conditional on observing $\hat{\theta}_n > 1$, just as the confidence interval in the MI case is conditional on observing $\hat{\theta}_n < 1$.

4 Simulations

This section reports numerical evidence based on $5,000$ replications of the finite sample distributions of the localizing rate estimates $\{\hat{\alpha}, \hat{\gamma}_n\}$ and the localizing coefficient estimates $\{\hat{c}, \tilde{c}\}$. The simulation design employed the model given in (2.1). The observations $\{X_t\}_{t=1}^n$ were generated for both mildly integrated ($c < 0$) and mildly explosive ($c > 0$) cases using standard normally distributed equation errors u_t from an initialization $X_0 \sim \mathcal{N}(0, 1)$. The specific experiments and simulation results for these cases are given in the following two subsections.

4.1 Mildly integrated case

The experiments for Figures 1-5 used the following parameter settings:

$$(i)\ \alpha = 0.85,\ c = -1,\ n \in \{50, 100, 250, 500, 1500\},$$
$$(ii)\ \alpha \in \{0.65, 0.75, 0.85, 0.95\},\ c = -1,\ n = 100,$$
$$(iii)\ \alpha = 0.85,\ c \in \{-0.5, -1.0, -2.0, -5.0\},\ n = 100,$$
$$(iv)\ \gamma_n = \alpha - \frac{\log|c|}{\log n} \in \{0.50, 0.70, 0.85, 1.0\},\ n = 100.$$

Figures 1-3 display kernel estimates of the empirical densities of the estimates $(\hat{\alpha}, \hat{c})$ of the localizing coefficient pair (α, c). Figure 4 shows kernel estimates of the empirical densities of the parameter $\hat{\gamma}_n$ for sample size $n = 100$ and for true values $\gamma_n = \alpha - \frac{\log(|c|)}{n^{\alpha}} \in \{0.50, 0.70, 0.85, 1.0\}$ corresponding to the true values $c \in \{-5.0, -2.0, -1.0, -0.5\}$ and true value $\alpha = 0.85$. For the same true value $\alpha = 0.85$ but with $\gamma_n = 1$, Figure 5 displays kernel estimates of the densities of $\hat{\gamma}_n$ for $n \in \{100, 250, 500, 1500\}$ shown against the asymptotic normal (5(a)) and asymptotic LUR (5(b)) distributions. The results are summarized as follows.

(1) Figure 1(b) shows increasing concentration of the density of $\hat{c}$ around the pseudo-true value $c_* = -1$ as n increases, consonant with the consistency $\hat{c} \to_p c_*$ established in Theorem 2.2. The rate of concentration in the distributions of $\hat{c}$ as n increases with $\alpha = 0.85$ is noticeable but slow, indicative of the convergence rate $O_p(\min\{n^{-(1-\alpha)/2}, n^{-\alpha/2}\})$ shown in (A-14) in the proof of Theorem 2.2.

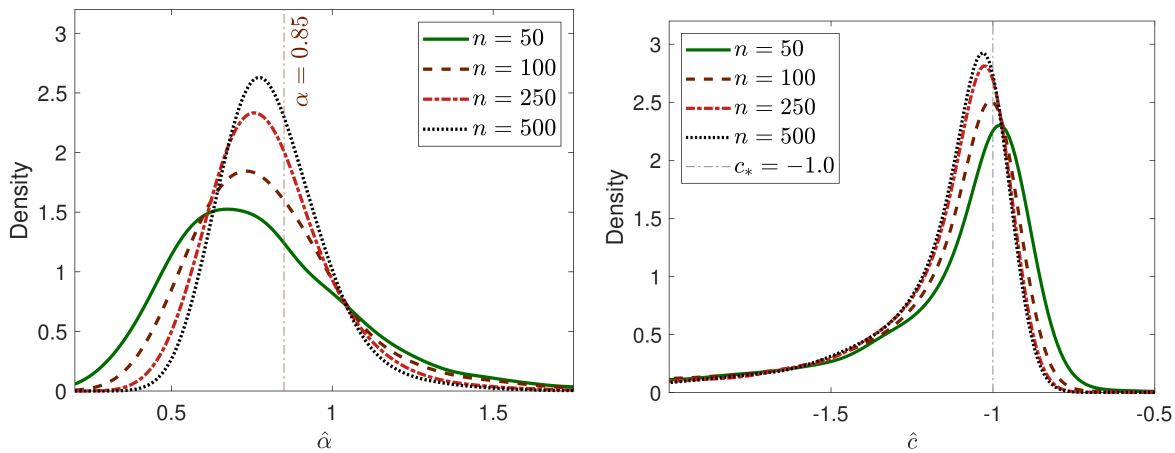
(2) The densities in Figure 1(a) show a similar increasing concentration in the distribution of $\hat{\alpha}$ as n increases, in this case combined with bias reduction. Again, this accords well with the limit theory in Theorem 2.1 where the convergence rate is $O_p(n^{-(1-\alpha)/2}\log n)$ combined with first order downward bias of $-\frac{\log|c|}{\log n}$. The latter is the analogue in the estimation of α of the usual downward bias in the least squares autoregressive coefficient estimate $\hat{\theta}_n$.

(3) Figures 2 (a) and (b) show the effects of varying the rate coefficient α on the distributions of $\hat{\alpha}$ and $\hat{c}$ when the sample size $n = 100$. As expected from Theorem 2.1, the central location of the distribution of $\hat{\alpha}$ shifts to follow the value of α but with clear indication of downward bias in each case, consonant with the known downward bias in the autoregressive coefficient estimate $\hat{\theta}$. On the other hand the central location of $\hat{c}$ is close to the pseudo-true value $c_* = -1$ for all values of α, corroborating the asymptotic theory that $\hat{c} \to_p -1$. The distribution of $\hat{c}$ does show somewhat greater skewness to the right and greater dispersion when $\alpha = 0.95$, no doubt reflecting the well-known skewness and dispersion of the estimate $\hat{\theta}$ in unit root and local to unity cases.

(4) Figures 3 (a) and (b) show the effects of the localizing coefficient value c on the distributions of $\hat{\alpha}$ and $\hat{c}$, with fixed $\alpha = 0.85$ and sample size $n = 100$. The impact on the distribution of $\hat{\alpha}$, seen in Figure 3 (a), is to shift the central location in accord with the changing value of $\alpha = 1 + \frac{c}{n^{\alpha}}$ for $c \in \{-0.5, -1, -2, -5\}$. The densities show greater concentration for larger

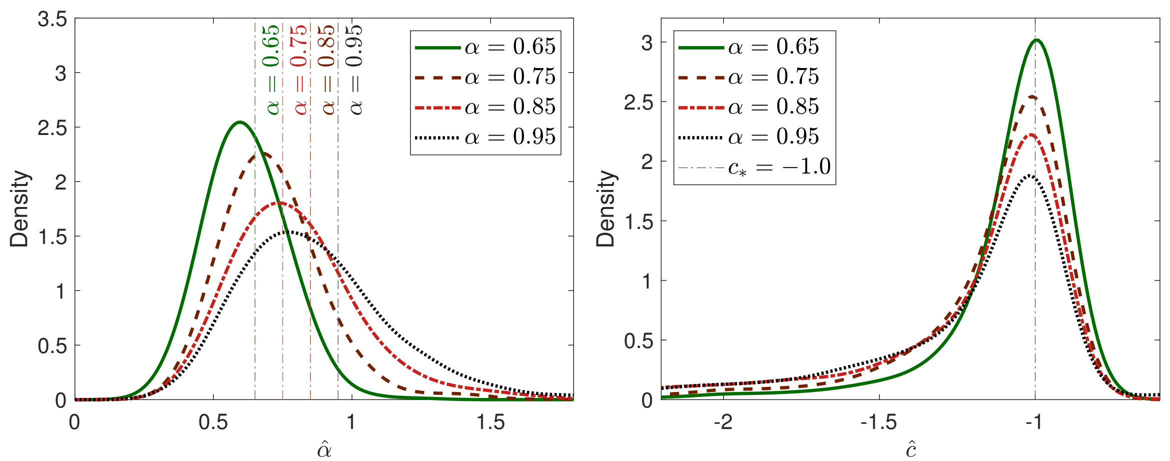
$|c|$ and again reflect the wider dispersion and skewness associated with the near-unit root case that applies when $|c|$ is small. For all these values of c the densities of $\hat{c}$ displayed in Figure 3 (b) show remarkable concentration about the pseudo-true limit value $c_* = -1$.

(5) Figure 4 (a) shows empirical densities for $n = 100$ of the estimates $\hat{\gamma}_n$ of the parameter γ_n in the uni-parameter MI specification (3.11). In contrast to the distributions of the estimates $\hat{\alpha}$ of the rate coefficient α in the conventional MI model formulation, the densities of $\hat{\gamma}_n$ are much better centered about the true values of γ_n, especially when γ_n takes on small values closer to $\gamma_n = 0.5$. This finite sample finding reflects the asymptotic theory in Corollary 2.1 in which the limit distribution is shown to be Gaussian and centered about the true value of γ_n.



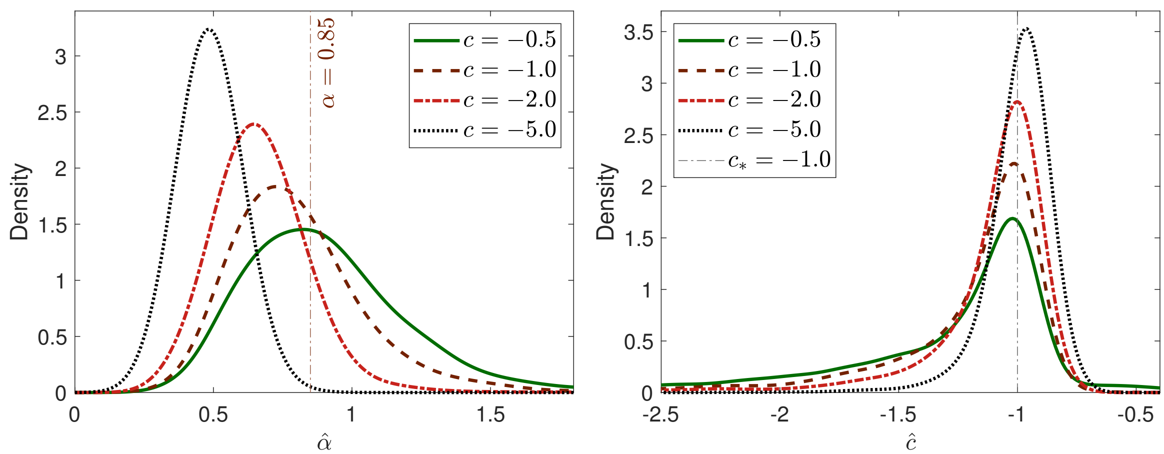
(a) $\hat{\alpha}$ densities (b) $\hat{c}$ densities

Figure 1: Empirical densities of the estimates $\hat{\alpha}$ and $\hat{c}$ for sample sizes $n \in \{50, 100, 250, 500\}$ with true value $\alpha = 0.85$, true $c = -1$ and pseudo-true value $c_* = -1$.



(a) $\hat{\alpha}$ densities (b) $\hat{c}$ densities

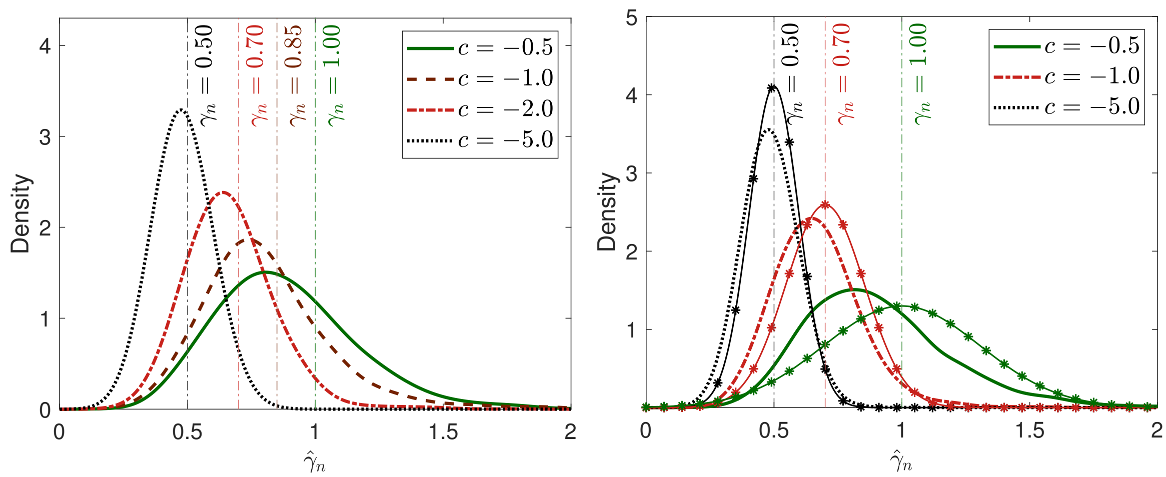
Figure 2: Empirical densities of the estimates $\hat{\alpha}$ and $\hat{c}$ for sample size $n = 100$, for true values of $\alpha \in \{0.65, 0.75, 0.85, 0.95\}$, true $c = -1$, and pseudo-true value $c_* = -1$.



(a) $\hat{\alpha}$ densities

(b) $\hat{c}$ densities

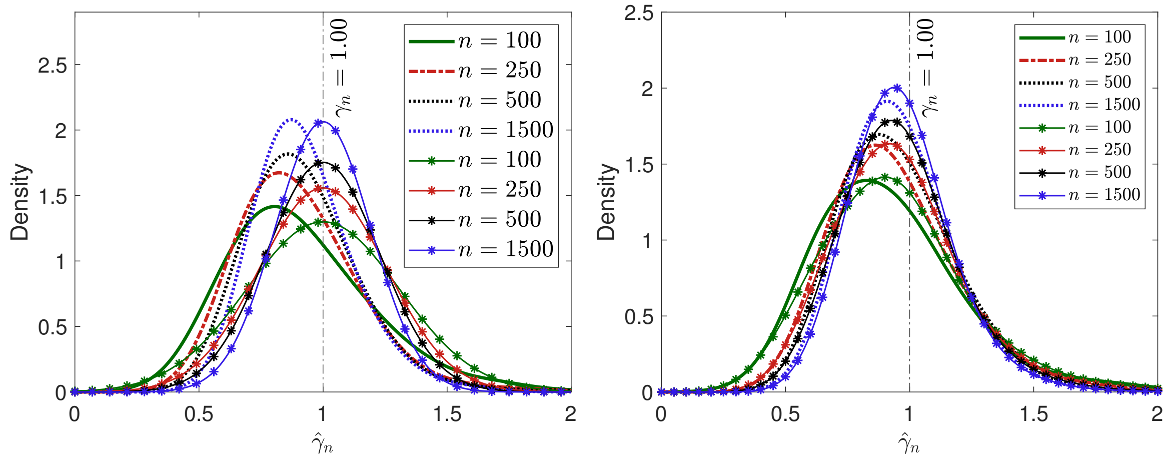
Figure 3: Empirical densities of the estimates $\hat{\alpha}$ and $\hat{c}$ for sample size $n = 100$, for true values $\alpha = 0.85$, true $c \in \{-0.5, -1.0, -2.0, -5.0\}$, and pseudo-true value $c_* = -1$.



(a) $\hat{\gamma}_n$ densities for various c values

(b) $\hat{\gamma}_n$ densities with asymptotic normal approximation

Figure 4: Panel (a): Empirical densities of the estimates $\hat{\gamma}_n$ for sample size $n = 100$ and for true values $\gamma_n = \alpha - \frac{\log|c|}{\log n} \in \{0.50, 0.70, 0.85, 1.0\}$ corresponding to the values $c \in \{-5.0, -2.0, -1.0, -0.5\}$ and fixed $\alpha = 0.85$. Panel (b): Empirical densities of the estimates $\hat{\gamma}_n$ for sample size $n = 100$ as in Panel (a) shown against the asymptotic $\mathcal{N}\left(\gamma_n, \frac{2}{n^{1-\gamma_n}\log^2 n}\right)$ curves displayed with asterisks.



(a) $\hat{\gamma}_n$ densities for $n = 100, 250, 500, 1500$, with asymptotic normal approximation

(b) $\hat{\gamma}_n$ densities for $n = 100, 250, 500, 1500$, with asymptotic transformed LUR approximation

Figure 5: Empirical densities of the estimates $\hat{\gamma}_n$ computed for the pairs $(c, n) \in \{(-0.501, 100), (-0.436, 250), (-0.393, 500), (-0.333, 1500)\}$ and $\alpha = 0.85$, with each pair corresponding to the constant localizing rate coefficient $\gamma_n = 1$. The asymptotic normal approximations (shown in (a)) and the asymptotic transformed LUR approximations (shown in (b)) are given by the curves with asterisks – see the text for details.

(6) Figure 4 (b) shows the same empirical densities of $\hat{\gamma}_n$ as in Panel (a) for $n = 100$ and $c \in -0.5, -1.0, -5.0$ against the asymptotic normal $\mathcal{N}\left(\gamma_n, \frac{2}{n^{1-\gamma_n}\log^2 n}\right)$ displayed with asterisks. For these values of c and n the corresponding values of γ_n are $1.0, 0.70, 0.50$. Evidently, for $\gamma_n = 0.5$ and $\gamma_n = 0.7$ the asymptotic distribution is adequate in terms of location but less so in the unit root case $\gamma_n = 1.0$ with $c = -0.5$. Here, the asymptotic distribution is biased upwards relative to the finite sample distribution when $n = 100$, partly reflecting the slower logarithmic rate of convergence to the limit distribution when $\gamma_n = 1.0$ and partly that in this case the autoregressive coefficient $\theta_n = 1 - \frac{1}{n}$ is in the LUR class and the distribution $\mathcal{N}\left(\gamma_n, \frac{2}{n^{1-\gamma_n}\log^2 n}\right)$ is a crude Gaussian approximation to the LUR limit theory. The comparison of the two approximations is explored more systematically in Figure 5.

(7) Figures 5 (a) and (b) show the empirical densities of $\hat{\gamma}_n$ against two types of asymptotic approximation in the LUR case where the autoregressive coefficient is $\theta_n = 1 - \frac{1}{n}$ and $\gamma_n = 1$. The simulations are performed for values of n increasing from $n = 100$ to $n = 1500$ and for (c, α) pairs where in each case $\alpha = 0.85$ and c rises towards zero according to $c = -1/n^{1-\alpha} = -0.50$ when $n = 100$ up to $c = -0.393$ when $n = 1,500$. In each of these cases, the autoregressive coefficient is $\theta = 1 - \frac{1}{n}$, and the uni-parameter localizing rate coefficient $\gamma_n = 1$ is in the LUR class. The asymptotic normal approximations based on $\mathcal{N}\left(\gamma_n, \frac{2}{n^{1-\gamma_n}\log^2 n}\right) = \mathcal{N}\left(1, \frac{2}{\log^2 n}\right)$ when $\gamma_n = 1$ are displayed by the curves with asterisks in Fig. 5(a). Since $\gamma_n = 1$ the rate of convergence to the asymptotic normal distribution is $\log n$ and this logarithmic rate is evident in the slow convergence of the empirical density plots towards the asymptotic normal

with slowly shrinking bias and variance, corroborating the limit theory.[5] In Fig. 5(a), the asymptotic normal approximations (asterisked) show similar variation to the finite sample distributions of $\hat{\gamma}_n$ but do not capture location as well since for all n the asymptotic distributions are centered on $\gamma_n = 1$ by construction and the finite sample distributions are centered below unity with the gap narrowing as n increases. In Fig. 5(b), the asymptotic transformed LUR approximations (asterisked) are obtained by computing $1 - \log\left|-1+\xi_{J_{-1}}\right| / \log n$ based on Theorem 2.1 (iii) and simulating the distribution of $\xi_{J_{-1}}$, using a sample size of $5,000$ and $25,000$ replications. Evidently the transformed LUR approximations provide substantially improved location estimates and match dispersion well with the finite sample distribution. Notably, both finite sample and LUR approximate densities of the rate estimator $\hat{\gamma}_n$ are bell shaped, in contrast to estimates of the autoregressive coefficient $\hat{\theta}_n$. The nonlinear transformation defining $\hat{\gamma}_n = -\log|\hat{\theta}_n - 1|/\log n$ plays the role of a normalizing transformation, similar to Fisher's $z-$transformation of the autocorrelation coefficient, as discussed earlier.

It is worth drawing attention to the fact that the asymptotic distribution at the local unit root limit where $\theta_n = 1 + \frac{c_n}{n^\alpha} \sim_a 1 - \frac{1}{n}$ is actually normal along the path toward the boundary where the pair $(c_n, n) \to (0, \infty)$ and the autoregressive coefficient $\theta_n \sim_a 1 - \frac{1}{n}$ and uni-parameter rate sequence $\gamma_n \to 1$, while remaining within the mildly integrated class with fixed $\alpha \in (0,1)$. The distributions of the centered and scaled estimates of γ_n along this path belong to a stable normal class in which the asymptotic approximation is Gaussian even though the autoregressive parameter is local to unity in the limit.[6] This normal class is very different from the LUR class where $\theta_n = 1 + \frac{c}{n}$ for which the limit distributions as $n \to \infty$ belong to a non-normal general unit root class involving linear diffusion and Brownian motion processes (J_c, B) which in turn converge to the standard unit root distribution when $c \to 0$, as shown in Phillips (1987b). Similarly, the normal class with $\gamma_n = 1$ and with fixed rate coefficient $\alpha \in (0,1)$, for which the pairs $(c_n, n) \to (0, \infty)$, differs considerably from the LUR class with $\theta_n = 1 + \frac{c}{n}$ considered in Theorem 2.1 (iii) where the rate coefficient $\alpha = 1$ is consistently estimated by $\hat{\alpha}_n$ within this class. However, a common feature of this latter class and the normal class is that they share the same $\log n$ convergence rate, reflecting the increased difficulty in estimating the localizing rate parameter as the unit root is approached. As indicated earlier, the asymptotic density $\mathcal{N}\left(\gamma_n, \frac{2}{n^{1-\gamma_n}\log^2 n}\right) \sim_a \mathcal{N}\left(1, \frac{2}{\log^2 n}\right)$ of γ_n may be considered a crude Gaussian approximation to the non-Gaussian LUR limit theory in the case $\gamma_n \sim_a 1$.

[5] Theorem 2.1 (i) and (ii) fail when $\alpha \in (0,1)$ but $c = c_n \to 0$ as fast as or faster than $c_n = -\frac{1}{n^{1-\alpha}}$. In that case the autoregressive coefficient is either $\theta_n = 1 - \frac{1}{n}$ or closer to unity with $\theta_n = 1 + o(\frac{1}{n})$ and Theorem 2.1 (iii) applies with $\log n$ convergence rate and limit distribution determined by the quantity $\log|c + \xi_{J_c}|$. Nonetheless, as discussed in the paragraph below, the asymptotic normal distribution $\mathcal{N}\left(1, \frac{2}{\log^2 n}\right)$ provides an alternative approximation when $\gamma_n = 1$.

[6] Another stable normal class in the unit root case is the partially aggregated differences estimator (PAE) studied in Han et al. (2011). But in that case, the PAE is an estimator of the autoregressive coefficient θ_n rather than the localizing rate parameters $\hat{\alpha}$ or $\hat{\gamma}_n$.

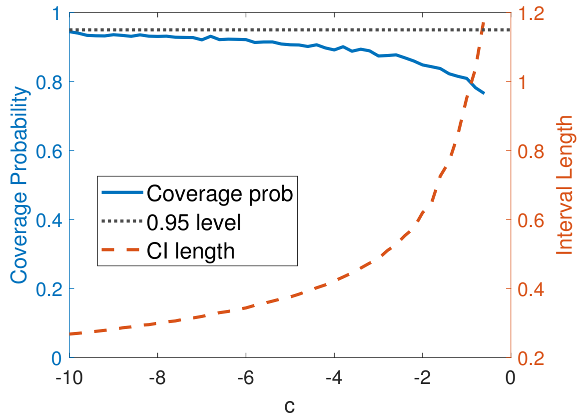


Figure 6: Coverage probabilities (solid blue curve, left axis), 95% confidence level (black dotted line) and confidence interval lengths (dashed sienna curve, right axis) for $\gamma_n \in [0.35, 0.96]$ based on (2.24), corresponding to various values of the localizing coefficient $c \in [-10, -0.6]$ with $\alpha = 0.85$ and $n = 100$.

(8) Figure 6 shows coverage probabilities and confidence interval lengths for inference concerning γ_n based on the asymptotic formula (2.24). The graphics were computed using 5,000 replications with $n = 100$, $\alpha = 0.85$, and values of $c < 0$ at equispaced intervals in the interval $[-10, -0.6]$, leading to a range of γ_n values in $[0.35, 0.96]$ with $\gamma_n = 0.35$ when $c = -10$ and $\gamma_n = 0.96$ when $c = -0.6$. The results show satisfactory coverage in the range $90\% - 95\%$ for values of $c \in [-10, -6]$ with coverage falling to around 75% when $c = -0.6$. Similarly, confidence interval lengths for γ rise from below 0.1 when $c = -10$ to over 1.1 when $c = -0.6$ and $\gamma_n = 0.96$. The lower coverage probabilities and much wider confidence interval lengths when γ_n is close to unity are to be expected, given the near logarithmic convergence rate as γ_n approaches unity.[7] Inference about the value of γ_n is clearly imprecise when the true value is close to unity, matching the poor local power of unit root tests and the difficulty of distinguishing a local unit root from a unit root.

4.2 Mildly explosive case

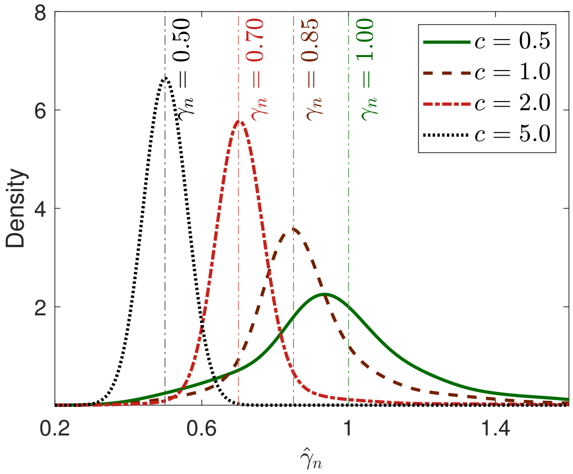
These simulations used experimental designs similar to those in the mildly integrated case but based on model (2.1) with $c > 0$. Figures 7-8 employed settings that give results for various values of the localizing scale coefficient c and the sample size n as follows:

$$(i)\ \ \alpha = 0.85,\ n = 100,\ c \in \{0.5, 1.0, 2.0, 5.0\}, \gamma_n = \alpha - \frac{\log |c|}{\log n} \in \{0.50, 0.70, 0.85, 1.0\};$$

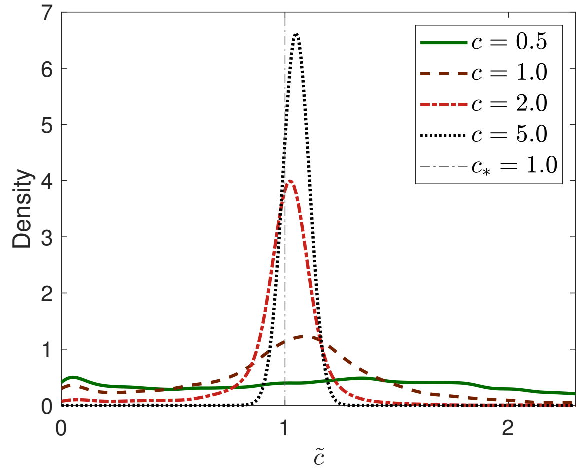
$$(ii)\ \alpha = 0.85,\ c = 1, n \in \{50, 100, 250, 500\}.$$

The results reported concentrate on (i) estimation of the uni-parameter rate sequence γ_n and the localizing scale coefficient $c > 0$; and (ii) inference about γ_n.

[7] When $n = 100$ and $\gamma_n = 0.96$, $n^{(1-\gamma_n)/2} \log n = 5.05$, and $n^{(1-\gamma_n)/2} \log n = 20.57$ when $\gamma_n = 0.35$.



(a) $\hat{\gamma}_n$ densities for various c values

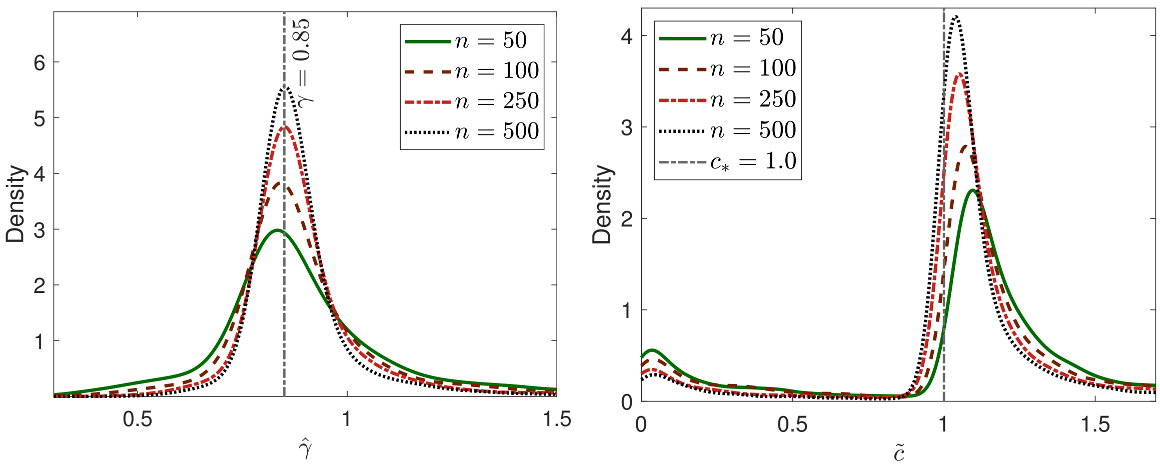


(b) $\tilde{c}_n$ densities for various c values

Figure 7: Panel (a): Empirical densities of the estimates $\hat{\gamma}_n$ for sample size $n = 100$ and for true values $\gamma_n = \alpha - \frac{\log|c|}{\log n} \in \{0.50, 0.70, 0.85, 1.0\}$ corresponding to the true values $c \in \{5.0, 2.0, 1.0, 0.5\}$ and $\alpha = 0.85$. Panel (b): Empirical densities of the estimates $\tilde{c}$ for sample size $n = 100$, $\alpha = 0.85$, and $c \in \{0.5, 1.0, 2.0, 5.0\}$, as in Panel (a).

Figure 7(a) shows that the distributions of $\hat{\gamma}_n$ are generally well centered about the true values in the mildly explosive case when $n = 100$. Dispersion increases as γ_n approaches unity, as expected from the rate of convergence and as occurs in the mildly integrated case (Fig. 4(b)). When $c = 0.5$ we have $\gamma_n = 0.85 - \frac{\log 0.5}{\log 100} = 1.0$ so that the autoregressive coefficient is $\theta = 1 + \frac{1}{n}$ and therefore immediately local to unity. In this case the density of $\hat{\gamma}_n$ is still very close to symmetric and only slightly biased below unity, unlike the density of the OLS estimate of θ in the LUR case or the mildly explosive case when $c = -0.5$ where the downward bias is more substantial (Fig. 4(b)).

In Figure 7(b) the densities of $\tilde{c}$ are shown for $c \in \{0.5, 1.0, 2.0, 5.0\}$, again for $n = 100$. The densities differ considerably between $c = 0.5$ and the higher values of c. When $c = 0.5$ with $n = 100$ we have $\theta = 1 + \frac{0.5}{100^{0.85}} = 1.01 = 1 + \frac{1}{100}$, corresponding to the value of the uni-parameter rate $\gamma_n = 1$. In this case the true value of the pair (c, α) is not identifiable and c is not consistently estimable. This property is reflected in the nearly uniform density of $\tilde{c}$ observed in 7(b) when $c = 0.5$. On the other hand, the uni-parameter γ_n is identifiable in this case and the density of the consistent estimator $\hat{\gamma}_n$ is centered close to unity with a small downward bias (Fig 7(a)). For values of $c > 1$ the densities of $\tilde{c}$ are centered close to the pseudo-true value $c_* = 1$ with progressively less dispersion as c increases, both as predicted by the limit theory.



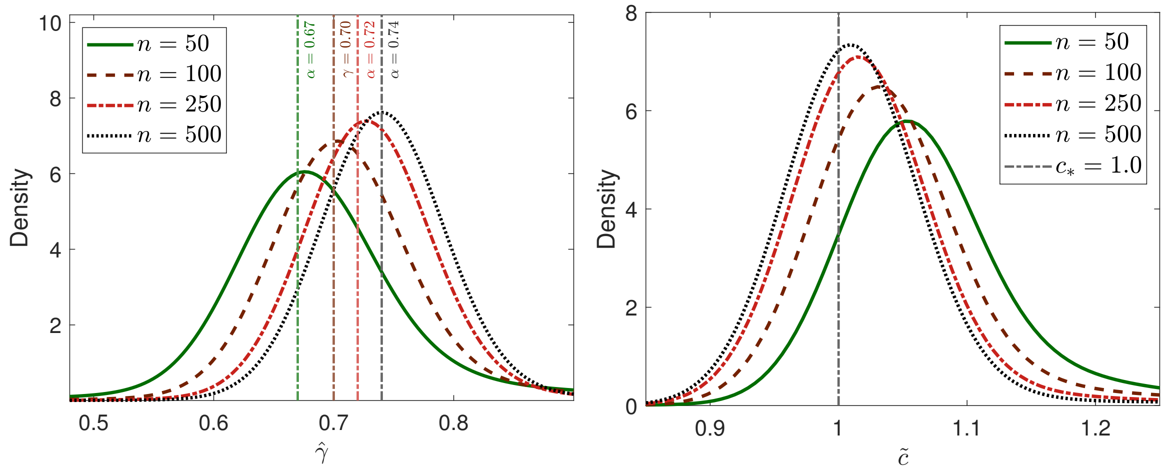
(a) $\hat{\gamma}_n$ densities for various sample sizes n

(b) $\tilde{c}_n$ densities for various sample sizes n

Figure 8: Panel (a): Empirical densities of $\hat{\gamma}_n$ for various sample sizes $n \in \{50, 100, 250, 500\}$ with true value $c = 1$ and $\gamma_n = \alpha = 0.85$. Panel (b): Empirical densities of $\tilde{c}$ for the same sample sizes and same values $c = 1$ and $\gamma = \alpha = 0.85$.

Figures 8(a) and 8(b) show the effects on the distributions of $\hat{\gamma}_n$ and $\tilde{c}$ of raising the sample size when true values of the parameters are $c = 1 = c_*$ and $\alpha = 0.85$. For these fixed parameters, the sample sizes $n \in \{50, 100, 250, 500\}$ lead to the same implied uni-parameter value $\gamma_n = \alpha = 0.85$. Evidently, the distributions of $\hat{\gamma}_n$ are well centered about the true value and show shrinking dispersion as n increases. The distributions of $\tilde{c}$ are located around a dominant primary mode close to the true value $c = 1$ with concentration that increases with n. There is evidence of a small secondary mode close to the origin, which is more evident for the smaller sample sizes $n = 50, 100$, and a small upward bias in the primary mode that diminishes as n increases.

Figure 9(a) shows the distributions of $\hat{\gamma}_n$ for sample sizes $n \in \{50, 100, 250, 500\}$ with true values $c = 2$ and $\alpha = 0.85$, which correspond to the implied values $\gamma_n \in \{0.67, 0.70, 0.72, 0.74\}$. Evidently, the distributions of $\hat{\gamma}_n$ are well centered about the true values in the mildly explosive case for these values of n when the localizing scale coefficient $c = 2$. Figure 9(b) shows the distributions of $\tilde{c}$ when the true value $c = 2$ for the same values of n. For the smaller sample sizes $n = 50, 100$, there is some upward bias from the pseudo-true value $c_* = 1$, but this bias disappears for the larger sample sizes $n = 250, 500$, corroborating the limit theory that $\tilde{c} \to_p 1$ in the mildly explosive case when $c \neq 1$.



(a) $\hat{\gamma}_n$ densities for various sample sizes n (b) $\tilde{c}$ densities for various sample sizes n

Figure 9: Panel (a): Empirical densities of $\hat{\gamma}_n$ for various sample sizes $n \in \{50, 100, 250, 500\}$ with true values $c = 2$, $\alpha = 0.85$, and corresponding values $\gamma_n \in \{0.67, 0.70, 0.72, 0.74\}$. Panel (b): Empirical densities of $\tilde{c}$ for the same sample sizes and same values $c = 2$, $\alpha = 0.85$, and $\gamma_n = 0.70$.

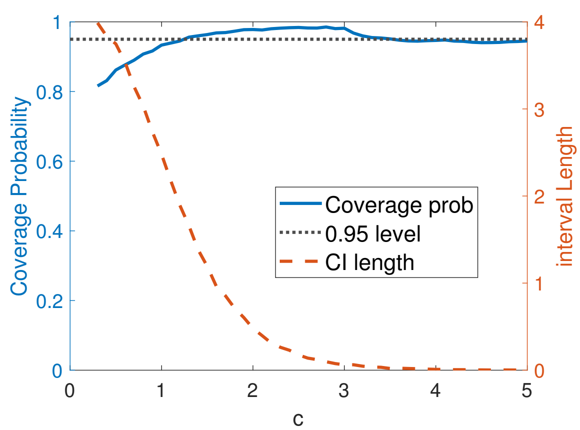


Figure 10: Coverage probabilities (solid blue curve, left axis), 95% confidence level (black dotted line) and confidence interval lengths (dashed sienna curve, right axis) for $\gamma_n \in [1.0, 0.50]$ based on (3.15), corresponding to various values of the localizing coefficient $c \in [0.5, 5.0]$ with $\alpha = 0.85$ and $n = 100$.

Figure 10 shows coverage probabilities and confidence interval lengths for inference concerning γ_n based on the asymptotic formula (3.15) using a nominal (two-sided) asymptotic level of 95% with the Cauchy distribution critical value 12.706.[8] The graphics were computed using 5,000 replications with $n = 100$, $\alpha = 0.85$, and values of $c > 0$ at equispaced intervals in the interval $[0.5, 5.0]$, leading to a range of γ_n values in $[1.0, 0.5]$ with $\gamma_n = 1.0$ when $c = 0.5$ and

[8]The quantile function for percentile p of the standard Cauchy distribution is $\tan(\pi(p - 0.5))$, which gives the critical value 12.706 when $p = 0.975$.

$\gamma_n = 0.5$ when $c = 5.0$ corresponding to autoregressive coefficient values $\theta_n \in [1.01, 1.1]$. The results reveal sharp coverage probabilities at the 95% level for values $c \in [3, 5]$ corresponding to $\theta_n \in [1.06, 1.1]$ with CI lengths at most 0.033 over this range.[9] The sharp outcomes over this range arise from the exponential convergence rate to the limiting Cauchy distribution. For values of $c \in (0.5, 3)$ coverage probability slowly declines to around the 82% level and the interval length increases monotonically towards unity at the point $c = 0.5$ where $\gamma_n = 1$, $\theta_n = 1 + \frac{1}{n} = 1.01$ and a logarithmic convergence rate holds in this boundary local to unity case.

5 Empirics

This section provides an empirical illustration of the paper's methodology to the housing market in Australia. The period since the global financial crisis (GFC) has witnessed rampant house price appreciation in many cities of the developed world. With the onset of the Covid-19 pandemic and the rapid expansion of credit by many monetary authorities in response, house price inflation has accelerated, persistently outpacing income growth and making housing affordability and consumer debt prominent issues for policy makers in many countries. Housing market exuberance has been especially marked in the antipodes, with many Australian and New Zealand cities experiencing in excess of 20% house price appreciation in a single year from March 2020.[10] The methods of this paper are used to make a quantitative assessment of this inflation and, in particular, to provide an empirical measure of its extent using estimates of the parametric rate of exuberance based on the mildly explosive model studied in the paper. Those cities where speculative behavior in the housing market is not identified are studied within the mildly integrated or local to unity framework.

House price exuberance can be defined as explosive or mildly explosive deviations of house prices from underlying market fundamentals. To determine the existence of such deviations or specific episodes of deviation it is necessary to obtain a measure of housing market fundamentals that can be used as a benchmark for the computation of deviations. One method is to select specific fundamentals such as rents or income and employ standardized quantities like price/rent or price/income ratios in conducting the data analysis. Another approach is to employ a reduced form regression method that accounts for the impact of a broad set of fundamental factors that may affect demand and supply pressures in housing markets. This approach avoids the specificity of a single factor fundamental such as that involved in the use of a price/rent or price/income ratio. The reduced form is fitted by IVX regression (Phillips and Magdalinos, 2009; Kostakis et al., 2015) to accommodate endogeneity in the regressors and key

[9] From (3.15) the interval length at $c = 3$ with $n = 100$ and $\alpha = 0.85$ is $\frac{4 \times 12.706}{(1+\frac{3}{100^{0.85}})^{100} \log 100} = 0.033$

[10] For Australia as a country, house price appreciation was 16.4% over the 12 month period to June 2021, placing Australia the 7th highest amoung 55 countries according to global house price indexes https://content.knightfrank.com/research/84/documents/en/global-house-price-index-q2-2021-8422.pdf; the corresponding figure for New Zealand was 25.9% placing New Zealand 2nd highest among the same group of countries.

variables such as the price/rent ratio are decomposed into fundamental and non-fundamental (NF) components. The NF components are the residuals in this regression and are therefore anticipated to have weak dependence or stationary characteristics if there are no other systematic forces at work driving house prices. These residuals may then be used to assess evidence for the presence of explosive or mildly explosive behavior in prices. This is the approach developed in Shi and Phillips (2021) and is employed in the construction of the NF component of the price/rent (P/R) ratio data used here.[11]

Figure 11 provides plots of monthly observations of the NF-P/R data for the eight state capital cities of Australia over the period July 31, 2012 to June 30, 2021. The plots for Sydney, Melbourne, Brisbane, Adelaide, Darwin and Canberra have a hockey stick graph form, each showing evidence of elevating NF-P/R ratios towards the end of the sample period from 2020. The earlier observations before 2020 show no noticeable systematic movements but are clearly highly autoregressive.

The methods of the paper were applied as follows. Empirical regressions with each NF house price/rent series were run giving linear least squares estimates $\hat{\theta}_n$ of θ_n and nonlinear estimates $\hat{\gamma}_n$ of γ_n according to (2.18) and (3.12). Mildly explosive and mildly integrated series were identified according to whether $\hat{\theta}_n \gtrless 1$, the rate parameter γ_n was estimated and confidence intervals were constructed for γ_n based on (3.15) and (2.24). The results are reported in Table 1 and detail both the full period 2012–2021 and the later period 2018–2021. The main findings are summarized as follows.

1. Sydney, Melbourne, Brisbane, Adelaide, Darwin, and Canberra all have autoregressive coefficients $\hat{\theta}_n > 1$ over the more recent period 2018–2021. With the exception of Adelaide, the corresponding rate coefficient estimates $\hat{\gamma}_n$ are all less than unity, signifying mildly explosive behavior for each of these cities. Furthermore, Sydney, Melbourne, Brisbane, and Canberra have 95% confidence intervals for γ_n lying within the $(0, 1)$ interval, confirming mildly explosive behavior in the NF house price/rent ratio at this level of significance. For Darwin, the rate estimate is $\hat{\gamma}_n = 0.8196$ and the confidence interval is wide $(0, 1.9)$, thereby including the explosive LUR and close to unity cases as possible generating mechanisms. For Adelaide, the rate coefficient $\hat{\gamma}_n = 1.1146$ exceeds unity, indicating an explosive close to unity coefficient and the confidence interval for γ_n is also wide, reflecting the slow convergence rate in the LUR case and the short sample size of 36 observations over this recent period.

2. Over the full period 2012–2021, Sydney, Melbourne, Brisbane, and Canberra again have explosive autoregressive coeffcients and $\hat{\gamma}_n$ estimates lying in the $(0, 1)$ interval, indicating that over this longer period the mildly explosive behavior at the end of the period remains

[11] The data employed here are downloadable from the website https://www.housing-fever.com/. The fundamental factors considered include real mortgage interest rates (nominal mortgage rates less inflation expectations), real rents, and real disposal income (proxied by State final demand) for the Australian cities. See Shi and Phillips (2021) and the website https://www.housing-fever.com/ for further details.

evident. For Sydney and Brisbane, the 95% confidence intervals for γ_n continue to fall within the $(0, 1)$ interval so the mildly explosive behavior is sustained at this significance level for the full period. For Melbourne and Canberra the confidence intervals are wider and include both explosive LUR and close to unity cases.

3. Perth and Hobart have autoregressive coefficients $\hat{\theta}_n < 1$ in both the full period and later period. These two cities also have estimated rate coefficients $\hat{\gamma}_n \in (0, 1)$, signifying mildly integrated behavior. The corresponding 95% confidence intervals support this inference although the interval $(0.235, 1.019)$ for Hobart is wider and includes unity, thereby allowing for the possibility of a local to unity coefficient. These cities therefore show no evidence of housing market bubbles. In addition and for the full period, Darwin has autoregressive coefficient $\hat{\theta}_n = 0.922 < 1$ with confidence interval $(0.341, 0.750)$ indicating mildly integrated behavior, so that for Darwin the exuberance observed in the later period is not strong enough to be sustained in estimation and inference over the full sample.

The findings reported above correspond broadly to the results of alternative methods of assessing the presence of exuberance in these Australian city housing markets. In particular, the recursive PSY test procedures developed in Phillips et al. (2015a) provide supportive evidence for exuberance in the latter part of the sample period in Sydney, Melbourne, Brisbane, Canberra and, less so, Adelaide and Darwin.[12] The present findings complement that evidence in two ways. First, the results distinguish the explosive alternative by measuring the departure from the null hypothesis of no exuberance in the NF price/rent ratio by means of the magnitude (lower values signifying greater departures) of estimates of the rate coefficient γ_n and its confidence intervals. Second, the alternative hypothesis in the present work allows for mildly integrated alternatives in addition to unit root and local to unity alternatives (as in the PSY test), again with quantification provided by the magnitude of the estimated rate coefficient $\hat{\gamma}_n$ when $\hat{\theta}_n < 1$. On the other hand, the PSY procedure is designed for real time dating of origination and termination of bubbles as well as detection. This is a feature for which recursive versions of the present estimation and inferential procedures can be developed and explored in later work.

[12] See https://www.housing-fever.com/.

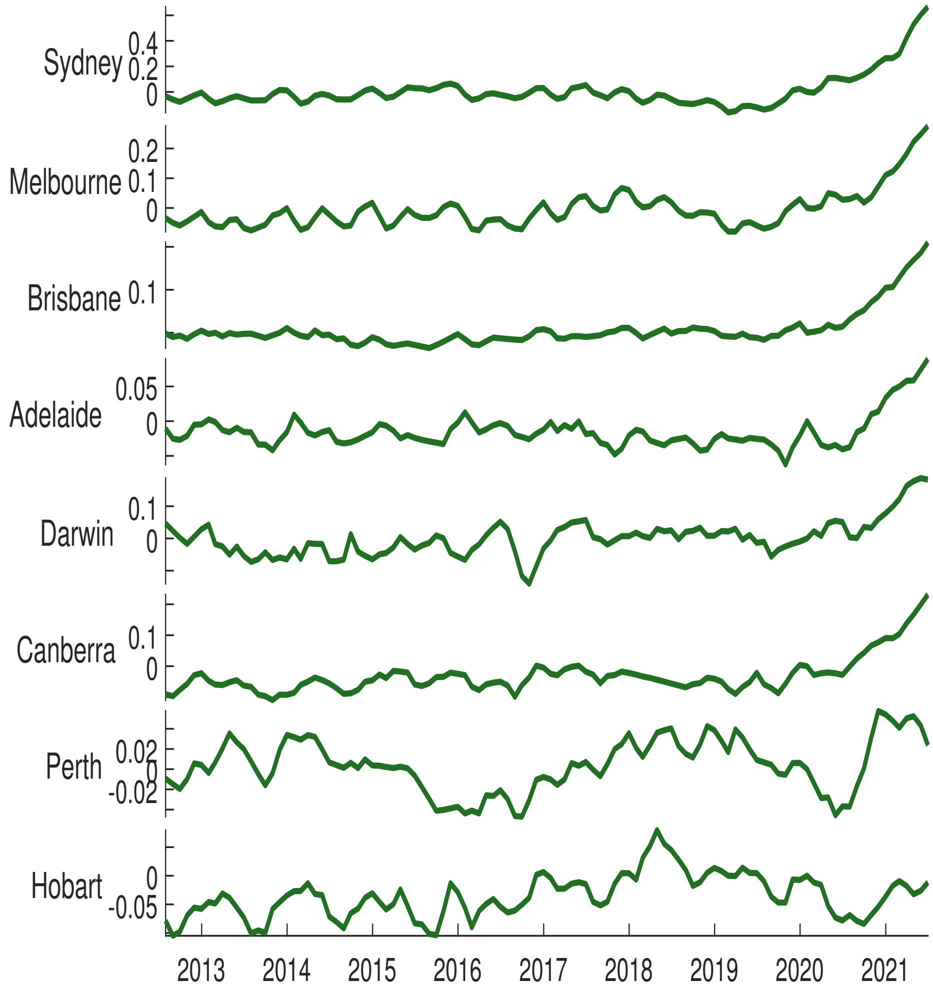


Figure 11: Monthly Australian metropolitan city house price/rent ratios (controlled by IVX regression estimation to remove the effects of economic fundamentals, including real disposable income and real mortgage interest rates) over the period 2013–2021. Data source: https://www.housing-fever.com/

Table 1: Autoregressive coefficient estimates $\hat{\theta}_n$, localizing rate estimates $\{\hat{\gamma}_n, \tilde{\gamma}_n\}$ and nominal 95% confidence intervals of γ_n for the non-fundamental components of house price/rental ratios in Australian State capitals

Capital city	Years	$\hat{\theta}_n$	$\hat{\gamma}_n$	$CI_{\hat{\gamma}}$	$\tilde{\gamma}_n$	$CI_{\tilde{\gamma}}$
Sydney	2012–2021	1.101	0.4897	$(0.4896, 0.4899)$		
	2018–2021	1.134	0.5564	$(0.4895, 0.6233)$		
Melbourne	2012–2021	1.021	0.8258	$(0.2457, 1.4061)$		
	2018–2021	1.112	0.6070	$(0.4671, 0.7468)$		
Brisbane	2012–2021	1.107	0.4768	$(0.4767, 0.4769)$		
	2018–2021	1.132	0.5587	$(0.4893, 0.6281)$		
Adelaide	2012–2021	0.974	0.7815	$(0.4265, 1.1365)$	0.7679	$(0.4240, 1.1117)$
	2018–2021	1.107	1.1146	$(0.000^*, 4.7690)$		
Darwin	2012–2021	0.922	0.5456	$(0.3413, 0.7500)$	0.5302	$(0.3331, 0.7273)$
	2018–2021	1.051	0.8196	$(0.000^*, 1.9047)$		
Canberra	2012–2021	1.012	0.9419	$(0.000^*, 2.4149)$		
	2018–2021	1.090	0.6661	$(0.3783, 0.9538)$		
Perth	2012–2021	0.919	0.5373	$(0.3369, 0.7377)$	0.4694	$(0.2984, 0.6404)$
	2018–2021	0.907	0.6582	$(0.2440, 1.0723)$	0.5169	$(0.1960, 0.8378)$
Hobart	2012–2021	0.924	0.5516	$(0.3444, 0.7588)$	0.5645	$(0.3509, 0.7781)$
	2018–2021	0.896	0.6273	$(0.2356, 1.0190)$	0.5997	$(0.2270.0.9723)$

Notes:
(i) $CI_{\hat{\gamma}_n}$ and $CI_{\tilde{\gamma}_n}$ indicate confidence intervals constructed using the estimate $\hat{\gamma}_n$ and bias corrected estimate $\tilde{\gamma}_n$, respectively (see Section 2.4 for $\tilde{\gamma}_n$ and $CI_{\tilde{\gamma}_n}$). Entries in the columns for $\tilde{\gamma}_n$ and $CI_{\tilde{\gamma}_n}$ are shown only for cases where $\hat{\theta}_n < 1$.
(ii) 0.000^* signifies that the lower limit of the constructed confidence interval for γ_n is negative and therefore lies below the natural zero boundary for the left limit of the rate coefficient.
(iii) Confidence intervals are constructed using (3.15) when $\hat{\theta}_n > 1$ and (2.24) when $\hat{\theta}_n < 1$. Confidence intervals robust to weak dependence when $\hat{\theta}_n < 1$ are obtained using the bias corrected estimator $\tilde{\gamma}_n < 1$ as shown in (2.36).
(iv) Data: 108 monthly observations from July 31, 2012 to June 30, 2021; 36 monthly observations from July 31, 2018 to June 30, 2021.

6 Concluding Remarks

Early research in the 1980s on time series regression with unit roots revealed the advantages of function space limit theory in delivering general properties for estimation and inference in regressions that involve time series variables with nonstationarities that can be captured by unit autoregressive roots or roots that may be local to unity. The methods in that research relied on non-standard limit theory involving nonlinear functionals of stochastic processes such as Brownian motions and diffusions. The present paper shows that valid inference about the character of nonstationary time series in a wider class than the unit root and local unit root class can be conducted using pivotal Gaussian and heavy-tailed Cauchy limit theory. The methods facilitate the study of time series that may have more divergent behavior or milder wandering behavior than random walks. These characteristics can be identified, estimated and used for inference about the particular form of nonstationarity in the data without conducting tests such as unit root or KPSS tests. As the empirical application to the housing market illustrates, the techniques may be particularly useful in studying episodes of financial and asset market exuberance where it is useful to distinguish different forms of nonstationarity. Rather than confining attention to unit root and local unit root processes in designing inference, the methods focus on the implied localizing rate parameter that measures the extent of divergence from unit root behavior, thereby adding to the econometric toolkit for detecting multiple different forms of nonstationarity in economic data. In future work these tools of inference with near unit roots can be applied in recursive analyses and empirical dating algorithms to characterize changes that may occur in the character of nonstationary data.

Appendix

Proof of Lemma 2.2. Since the rate parameter sequence is defined by $\gamma_n = \alpha - \frac{\log|c|}{\log n}$, we have

$$n^{\frac{1+\gamma_n}{2}} = n^{\frac{1+\alpha}{2}} n^{-\frac{1}{2}\frac{\log|c|}{\log n}} = n^{\frac{1+\alpha}{2}} |c|^{-1/2} \tag{A-1}$$

because

$$\log\left(n^{\frac{1}{2}\frac{\log|c|}{\log n}}\right) = \frac{1}{2}\frac{\log|c|}{\log n} \times \log n = \log|c|^{1/2}$$

and so $n^{-\frac{1}{2}\frac{\log|c|}{\log n}} = |c|^{-1/2}$. The results of Lemma 2.2 follow directly by use of the equivalence (A-1) in Lemma 2.1. □

Proof of Theorem 2.1. **Part (i):** Working in the expanded probability space where (2.9) holds and using the definition $\hat{A}_n = \hat{\theta}_n - 1$ and the fact that $c < 0$ and $\alpha \in (0,1)$ are fixed we have, as $n \to \infty$,

$$\begin{aligned}
\log\left|\hat{A}_n\right| &= \log\left|\frac{c}{n^\alpha} + \frac{\xi_c}{n^{\frac{1+\alpha}{2}}}\{1 + o_p(1)\}\right| \\
&= \log\left(\frac{1}{n^\alpha}\left|c + \frac{\xi_c}{n^{\frac{1-\alpha}{2}}}\{1 + o_p(1)\}\right|\right) \\
&= -\alpha\log n + \log\left|c + \frac{\xi_c}{n^{\frac{1-\alpha}{2}}}\{1 + o_p(1)\}\right| \\
&= -\alpha\log n + \log|c| + \log\left|1 + \frac{\xi_c}{cn^{\frac{1-\alpha}{2}}}\{1 + o_p(1)\}\right| \\
&= -\alpha\log n + \log|c| + \frac{\xi_c}{cn^{\frac{1-\alpha}{2}}}\{1 + o_p(1)\}.
\end{aligned} \tag{A-2}$$

Then

$$\hat{\alpha} = -\frac{\log\left|\hat{A}_n\right|}{\log n} = \alpha - \frac{\log|c|}{\log n} - \frac{\xi_c}{cn^{\frac{1-\alpha}{2}}\log n}\{1 + o_p(1)\}, \tag{A-3}$$

so that $\hat{\alpha} \to_p \alpha$ in both the expanded and the original space. Moreover, in the original space we have the weak convergence

$$n^{\frac{1-\alpha}{2}}\log n\left\{\hat{\alpha} - \alpha + \frac{\log|c|}{\log n}\right\} \Rightarrow \frac{\xi_c}{-c} =_d \mathcal{N}\left(0, \frac{2}{|c|}\right),$$

giving the required result (i).

Part (ii): When $\theta_n = 1 + \frac{c_n}{n^\alpha} = -\frac{1}{n^{1-\delta}}$ with $c_n = -\frac{1}{n^{1-\alpha-\delta}}$ with fixed $\delta > 0$, then $A_n = \theta_n - 1 = -\frac{1}{n^{1-\delta}}$, and in view of (2.9) we have

$$\hat{A}_n = \hat{\theta}_n - 1 = -\frac{1}{n^{1-\delta}} + \frac{\xi}{n^{1-\delta/2}}\{1 + o_p(1)\},$$

with $\xi =_d \mathcal{N}(0,2)$. In place of (A-2), we now find that

$$\log\left|\hat{A}_n\right| = \log\left|-\frac{1}{n^{1-\delta}} + \frac{\xi}{n^{1-\delta/2}}\left\{1+o_p(1)\right\}\right|$$
$$= \log\left|\frac{1}{n^{1-\delta}}\right| + \log\left|1 - \frac{\xi}{n^{\delta/2}}\left\{1+o_p(1)\right\}\right|$$
$$= -(1-\delta)\log n - \frac{\xi}{n^{\delta/2}}\left\{1+o_p(1)\right\}. \tag{A-4}$$

Then

$$\hat{\alpha} = -\frac{\log\left|\hat{A}_n\right|}{\log n} = 1-\delta + \frac{\xi}{n^{\delta/2}\log n}\left\{1+o_p(1)\right\}, \tag{A-5}$$

so that $\hat{\alpha} \to_p 1-\delta$ in both the expanded and the original space, with the following limit theory in the original space

$$n^{\delta/2}\log n\left\{\hat{\alpha} - (1-\delta)\right\} \rightsquigarrow \xi =_d \mathcal{N}(0,2),$$

giving (ii). When $\delta = 1-\alpha$ we have $n^{(1-\alpha)/2}\log n\left\{\hat{\alpha}-\alpha\right\} \rightsquigarrow \mathcal{N}(0,2)$ as in (i) with $|c| = 1$.

Part (ii) SV Extension: In a similar manner, when $c_n = -\frac{L_n}{n^{1-\alpha}}$ with an SV function $L_n \to \infty$ as $n \to \infty$ we have $\theta_n = 1 - \frac{L_n}{n}$. In the notation of Phillips and Magdalinos (2007a) this formulation is $\theta_n = 1 + \frac{c}{k_n}$ with $k_n = |c|n/L_n$, or more simply $k_n = n/L_n =$ where the SV function absorbs the constant $|c|$. From Phillips and Magdalinos (2007a) we have the limit theory

$$\sqrt{nk_n}(\hat{\theta}_n - \theta_n) \rightsquigarrow \mathcal{N}(0,-2c) = \xi_c. \tag{A-6}$$

In our present notation $\frac{n}{\sqrt{L_n}}(\hat{\theta}_n - 1 + \frac{L_n}{n}) \rightsquigarrow \mathcal{N}(0,2) = \xi$, which in the expanded probability space we write as

$$\hat{\theta}_n - 1 = -\frac{L_n}{n} + \frac{\sqrt{L_n}}{n}\xi\{1+o_p(1)\}. \tag{A-7}$$

It follows that

$$\log\left|\hat{A}_n\right| = \log|\hat{\theta}_n - 1| = \log\left|-\frac{L_n}{n} + \frac{\sqrt{L_n}}{n}\xi\left\{1+o_p(1)\right\}\right|$$
$$= \log\left|\frac{L_n}{n}\right| + \log\left|1 - \frac{\xi}{\sqrt{L_n}}\left\{1+o_p(1)\right\}\right|$$
$$= -\log n + \log L_n - \frac{\xi}{\sqrt{L_n}}\left\{1+o_p(1)\right\}. \tag{A-8}$$

Hence

$$\hat{\alpha} = -\frac{\log\left|\hat{A}_n\right|}{\log n} = 1 - \frac{\log L_n}{\log n} + \frac{\xi}{L_n^{1/2}\log n}\left\{1+o_p(1)\right\}, \tag{A-9}$$

which leads to the limit theory

$$L_n^{1/2} \log n \left(\hat{\alpha} - \left\{ 1 - \frac{\log L_n}{\log n} \right\} \right) \rightsquigarrow \xi, \tag{A-10}$$

giving the required extension of (ii) stated in Remark 2.1. The downward bias term in this near unit root case is $-\frac{\log L_n}{\log n}$ which is nonnegligible given the convergence rate $L_n \log n$.

Part (iii): When $\theta_n = 1 + \frac{c}{n}$ for some fixed $c \in (-\infty, \infty)$ and $\alpha = 1$ in (2.1) the limit theory for $\hat{\theta}_n$ is (see Phillips, 1987b)

$$n \left(\hat{\theta}_n - \theta_n \right) \rightsquigarrow \frac{\int_0^1 J_c dW}{\int_0^1 J_c^2} =: \xi_{J_c}, \quad as\ n \to \infty, \tag{A-11}$$

where $J_c(r) = \int_0^r e^{c(r-s)} dW(s)$ is a standard linear diffusion and W is standard Brownian motion. Then, proceeding as above on a suitably expanded probability space we have $\hat{\theta}_n = \theta_n + \frac{1}{n}\xi_{J_c}\{1 + o_p(1)\}$ and $\hat{A}_n := \hat{\theta}_n - 1 = \frac{c}{n} + \frac{1}{n}\xi_{J_c}\{1 + o_p(1)\} = \frac{1}{n}[c + \xi_{J_c}\{1 + o_p(1)\}]$ so that

$$\log \left| \hat{A}_n \right| = -\log n + \log |c + \xi_{J_c}\{1 + o_p(1)\}| = -\log n + \log |c + \xi_{J_c}| + o_p(1).$$

Hence,

$$\hat{\alpha} = -\frac{\log \left| \hat{A}_n \right|}{\log n} = 1 - \frac{\log |c + \xi_{J_c}|}{\log n} + o_p \left(\frac{1}{\log n} \right)$$

and then $\hat{\alpha} \to_p 1$ in both the expanded and original spaces. The convergence rate of $\hat{\alpha}$ is $\log n$ and in the original space we have weak convergence and the following limit theory for the normalized and centered estimator

$$(\log n)\{\hat{\alpha} - 1\} \rightsquigarrow -\log |c + \xi_{J_c}|,$$

giving the stated result (iii). □

Proof of Theorem 2.2. From Theorem 2.1 $\hat{\alpha} - \alpha = -\frac{\log |c|}{\log n}\{1 + o_p(1)\}$ and so $n^{\hat{\alpha}-\alpha} = n^{\frac{\log(1/|c|)}{\log n}\{1+o_p(1)\}}$ from which we have

$$\log \left(n^{\hat{\alpha}-\alpha} \right) = (\hat{\alpha} - \alpha) \log n = \frac{\log(1/|c|)}{\log n}\{1 + o_p(1)\} \log n = \log(1/|c|)\{1 + o_p(1)\}. \tag{A-12}$$

Then

$$\hat{c} = -\frac{n^{\hat{\alpha}-\alpha}\hat{\sigma}^2}{\frac{2}{n^{1+\alpha}}\sum_{t=1}^n X_{t-1}^2} = -\frac{\frac{1}{|c|}\hat{\sigma}^2}{\frac{2}{n^{1+\alpha}}\sum_{t=1}^n X_{t-1}^2}\{1 + o_p(1)\} \to_p -\frac{\frac{1}{|c|}\sigma^2}{\frac{\sigma^2}{-c}} = -1, \tag{A-13}$$

so that $\hat{c} \to c_* = -1$ as required. To find the convergence rate, note first that by standard calculations $\hat{\sigma}^2 = \sigma^2 + O_p(1/\sqrt{n})$. Next $\log n^{\hat{\alpha}-\alpha} = (\hat{\alpha}-\alpha)\log n = O_p\left(n^{-\frac{1}{(1-\alpha)/2}}\right)$ from Theorem 2.1, so that $n^{\hat{\alpha}-\alpha} = 1 + O_p\left(n^{-\frac{1}{(1-\alpha)/2}}\right)$. Further, as in the proof of PM(2007a, Lemma 3.1), we find that

$$\frac{1-\theta^2}{n}\sum_{t=1}^{n} X_{t-1}^2 = \sigma^2 + O_p(n^{-1/2}) + O_p\left(n^{-(1-\alpha)/2}\right) + O_p\left(n^{-\alpha/2}\right),$$

from which it follows that

$$\hat{c} = -\frac{n^{\hat{\alpha}-\alpha}\hat{\sigma}^2}{\frac{2}{n^{1+\alpha}}\sum_{t=1}^{n} X_{t-1}^2} = -1 + O_p\left(n^{-(1-\alpha)/2} + n^{-\alpha/2}\right) = -1 + O_p(\min\{n^{-(1-\alpha)/2}, n^{-\alpha/2}\}). \tag{A-14}$$

□

Proof of Corollary 2.1. In (i) we have $\theta = 1 + \frac{c}{n^\alpha}$. So, from (A-2) and using the fact that $n^{\gamma_n} = n^\alpha/|c|$ we have

$$\begin{aligned}\frac{\log|\hat{A}_n|}{\log n} &= -\alpha + \frac{\log|c|}{\log n} + \frac{\xi_c}{cn^{\frac{1-\alpha}{2}}\log n}\{1+o_p(1)\} = -\gamma_n + \frac{\xi_c}{cn^{\frac{1-\alpha}{2}}\log n}\{1+o_p(1)\} \\ &= -\gamma_n + \frac{\xi_c|c|^{\frac{1}{2}}n^{\frac{\gamma_n}{2}}}{cn^{\frac{1}{2}}\log n}\{1+o_p(1)\} = -\gamma_n + \frac{\xi}{n^{\frac{1-\gamma_n}{2}}\log n}\{1+o_p(1)\}\end{aligned} \tag{A-15}$$

where $\xi_c = \mathcal{N}(0,-2c)$ from (2.5) and $\xi = \mathcal{N}(0,2)$ from Lemma 2.2. Then

$$\hat{\gamma}_n - \gamma_n = -\frac{\log\left|\hat{A}_n\right|}{\log n} - \gamma_n = -\frac{\xi}{n^{\frac{1-\gamma_n}{2}}\log n}\{1+o_p(1)\}, \tag{A-16}$$

from which it follows that

$$n^{\frac{1-\gamma_n}{2}}\log n\,(\hat{\gamma}_n - \gamma_n) \rightsquigarrow \xi =_d \mathcal{N}(0,2), \tag{A-17}$$

giving the stated result. The same result holds directly in case (ii) because $c_n = -\frac{1}{n^{1-\alpha-\delta}}$ with $\delta \in (0, 1-\alpha]$ and then $\gamma_n = 1-\delta$.

In the extension of (ii) discussed in Remark 2.1 where $\theta_n = 1 - \frac{L_n}{n}$ involving the SV function L_n we have $c_n = -\frac{L_n}{n^{1-\alpha}} = -\frac{n^{\frac{\log L_n}{\log n}}}{n^{1-\alpha}}$ so that

$$\theta_n = 1 + \frac{c_n}{n^\alpha} = 1 - \frac{1}{n^{1-\frac{\log L_n}{\log n}}} = 1 - \frac{1}{n^{\gamma_n}},$$

with $\gamma_n = 1 - \frac{\log L_n}{\log n}$. The following limit theory then holds when $L_n \to \infty$ as $n \to \infty$

$$L_n^{1/2}\log n\,(\hat{\gamma}_n - \gamma_n) \rightsquigarrow \xi, \tag{A-18}$$

where the centering in (A-10) is adjusted by the term $-\frac{\log L_n}{\log n}$ which becomes absorbed in the uni-parameter $\gamma_n = 1 - \frac{\log L_n}{\log n}$. Result (A-18) continues to hold for choices of the SV function L_n, such as $L_n = \log n$, for which $\frac{\log L_n}{\log n} \to 0$ and $\gamma_n = 1 - \frac{\log L_n}{\log n} \to 1$ as $n \to \infty$.

Finally, in the LUR case where $\theta_n = 1 + \frac{c}{n}$ with fixed c, the result follows directly from Theorem 2.1(iii) and the fact that $\hat{\gamma}_n = \hat{\alpha}$. □

Proof of Corollary 2.2. Assume $0 < \hat{\theta}_n < 1$ and write

$$\begin{aligned}\hat{\gamma}_n - \gamma_n &= -\frac{\log\left(1-\hat{\theta}_n\right)}{\log n} + \frac{\log\left(1-\theta_n\right)}{\log n} = \frac{1}{\log n}\left\{\log\frac{1}{1-\hat{\theta}_n} - \log\frac{1}{1-\theta_n}\right\} \\ &= \frac{1}{\log n}\left\{\log\frac{1+\hat{\theta}_n}{1-\hat{\theta}_n} - \log\frac{1+\theta_n}{1-\theta_n}\right\} - \frac{1}{\log n}\log\frac{1+\hat{\theta}_n}{1+\theta_n}.\end{aligned}$$

Now $\theta_n = 1 - \frac{1}{n^{\gamma_n}}$ and, from Lemma 2.2, $\hat{\theta}_n = \theta_n + O_p\left(\frac{1}{n^{(1+\gamma_n)/2}}\right)$ so that

$$\log\frac{1+\hat{\theta}_n}{1+\theta_n} = \log\frac{1+\theta_n+O_p\left(\frac{1}{n^{(1+\gamma_n)/2}}\right)}{1+\theta_n} = \log\left\{1+O_p\left(\frac{1}{n^{(1+\gamma_n)/2}}\right)\right\} = O_p\left(\frac{1}{n^{(1+\gamma_n)/2}}\right).$$

It follows that

$$\hat{\gamma}_n - \gamma_n = \frac{1}{\log n}\left\{\log\frac{1+\hat{\theta}_n}{1-\hat{\theta}_n} - \log\frac{1+\theta_n}{1-\theta_n}\right\} + O_p\left(\frac{1}{n^{(1+\gamma_n)/2}\log n}\right)$$

and thus

$$n^{\frac{1-\gamma_n}{2}}\log n\left(\hat{\gamma}_n - \gamma_n\right) = n^{\frac{1-\gamma_n}{2}}\left\{\log\frac{1+\hat{\theta}_n}{1-\hat{\theta}_n} - \log\frac{1+\theta_n}{1-\theta_n}\right\} + O_p\left(\frac{1}{n^{\gamma_n}}\right) \rightsquigarrow \mathcal{N}(0,2),$$

as stated. □

Proof of Theorem 2.3. **Part (i):** For model (2.1) and (2.2) with fixed $c < 0$, fixed $\alpha \in \left(\frac{1}{3}, 1\right)$ and u_t satisfying Assumption **LP** as $n \to \infty$ we have the following limit theory for $\hat{\theta}_n$ from PM(2007b[Theorem 4.2, equation (24)])

$$n^{\frac{1+\alpha}{2}}\left(\hat{\theta}_n - \theta_n - \frac{1}{n^\alpha}\frac{-2c\lambda}{\omega^2}\right) \rightsquigarrow \mathcal{N}\left(0, -2c\varphi^2\right) =: \xi_{c\varphi^2} \quad \textit{as } n \to \infty, \tag{A-19}$$

where $\omega^2 = \sigma^2 C(1)^2$ is the long run variance of u_t, $\lambda = (\omega^2 - \sigma^2)/2$ is the one-sided long run covariance of u_t, and $\varphi = \frac{\sigma^2}{\omega^2}$. Expanding the probability space as before, the convergence (A-19) holds in probability and we have

$$\hat{\theta}_n = \theta_n + \frac{1}{n^\alpha}\frac{2c\lambda}{\omega^2} + \frac{\xi_{c\varphi^2}}{n^{\frac{1+\alpha}{2}}}\left\{1 + o_p(1)\right\}.$$

Setting $\hat{A}_n = \hat{\theta}_n - 1$ it follows that for $c < 0$

$$\begin{aligned}
\log\left|\hat{A}_n\right| &= \log\left|\frac{c}{n^\alpha} - \frac{1}{n^\alpha}\frac{2c\lambda}{\omega^2} + \frac{\xi_{c\varphi^2}}{n^{\frac{1+\alpha}{2}}}\{1+o_p(1)\}\right| \qquad \text{(A-20)}\\
&= \log\left\{\left|\frac{1}{n^\alpha}\right|\left|c - \frac{2c\lambda}{\omega^2} + \frac{\xi_{c\varphi^2}}{n^{\frac{1-\alpha}{2}}}\{1+o_p(1)\}\right|\right\}\\
&= -\alpha\log n + \log\left|c - \frac{2c\lambda}{\omega^2} + \frac{\xi_{c\varphi^2}}{n^{\frac{1-\alpha}{2}}}\{1+o_p(1)\}\right|\\
&= -\alpha\log n + \log\left\{|c|\left|1 - \frac{2\lambda}{\omega^2} + \frac{\xi_{c\varphi^2}}{cn^{\frac{1-\alpha}{2}}}\{1+o_p(1)\}\right|\right\}\\
&= -\alpha\log n + \log|c| + \log\left|\frac{\sigma^2}{\omega^2} + \frac{\xi_{c\varphi^2}}{cn^{\frac{1-\alpha}{2}}}\{1+o_p(1)\}\right|\\
&= -\alpha\log n + \log|c| + \log\varphi + \log\left|1 + \frac{\xi_{c\varphi^2}}{c\varphi n^{\frac{1-\alpha}{2}}}\{1+o_p(1)\}\right|\\
&= -\alpha\log n + \log|c| + \log\varphi + \frac{1}{c}\frac{\xi_c}{n^{\frac{1-\alpha}{2}}}\{1+o_p(1)\}. \qquad \text{(A-21)}
\end{aligned}$$

Then, using the fact that $n^\alpha = |c|n^{\gamma_n}$, we have

$$\begin{aligned}
\hat{\gamma}_n &= -\frac{\log\left|\hat{A}_n\right|}{\log n} = \alpha - \frac{\log|c| + \log\varphi}{\log n} - \frac{\xi_c}{cn^{\frac{1-\alpha}{2}}\log n}\{1+o_p(1)\}\\
&= \gamma_n - \frac{\log\varphi}{\log n} - \frac{\xi_c}{c^{1/2}n^{\frac{1-\gamma_n}{2}}\log n}\{1+o_p(1)\} = \gamma_n - \frac{\log\varphi}{\log n} - \frac{\xi}{n^{\frac{1-\gamma_n}{2}}\log n}\{1+o_p(1)\}, \qquad \text{(A-22)}
\end{aligned}$$

where $\xi =_d \mathcal{N}(0,2)$, from which it follows that $\hat{\gamma}_n - \gamma_n \to_p 0$, in both the expanded and the original space at a $\log n$ rate. In the original space we have the weak convergence

$$n^{\frac{1-\gamma_n}{2}}\log n\left\{\hat{\gamma}_n - \gamma_n + \frac{\log\varphi}{\log n}\right\} \rightsquigarrow \xi =_d \mathcal{N}(0,2), \qquad \text{(A-23)}$$

giving the stated result.

Parts (ii) & (iii): Here we have $\theta_n = 1 + \frac{c_n}{n^\alpha} = 1 - \frac{1}{n^{\gamma_n}}$ with $c_n = -\frac{1}{n^{1-\alpha-\delta}}$, $\gamma_n = 1 - \delta$ and $0 < \delta \leq 1 - \alpha$. Using the same derivations as above but with these values of (c_n, γ_n) we obtain

$$n^{\delta/2}\log n\left\{\hat{\gamma}_n - \gamma_n + \frac{\log\varphi}{\log n}\right\} \rightsquigarrow \xi =_d \mathcal{N}(0,2),$$

giving (2.32). Similarly, in the SV case (iii) where $\theta_n = 1 + \frac{c_n}{n^\alpha} = 1 - \frac{L_n}{n}$ with $c_n = -\frac{L_n}{n^{1-\alpha}}$ and $L_n \to \infty$ is slowly varying at infinity, we have as $n \to \infty$

$$L_n^{1/2}\log n\left(\hat{\gamma}_n - \gamma_n + \frac{\log\varphi}{\log n}\right) \rightsquigarrow \xi =_d \mathcal{N}(0,2). \qquad \text{(A-24)}$$

Part (iv): When $\theta_n = 1 + \frac{c}{n}$ with fixed c we have on a suitably expanded space $\hat{\theta}_n =$

$\theta_n + \frac{1}{n}\xi_{J_c}\{1+o_p(1)\}$ so that

$$\log\left|\hat{A}_n\right| = -\log n + \log\left|c + \xi_{J_c}\{1+o_p(1)\}\right| = -\log n + \log\left|c+\xi_{J_c}\right| + o_p(1).$$

Hence,

$$\hat{\gamma}_n = -\frac{\log\left|\hat{A}_n\right|}{\log n} = 1 - \frac{\log\left|c+\xi_{J_c}\right|}{\log n} + o_p\left(\frac{1}{\log n}\right)$$

from which $\hat{\gamma}_n \to_p 1$ and the limit theory $(\log n)\{\hat{\gamma}_n - 1\} \rightsquigarrow -\log|c+\xi_{J_c}|$, holds as stated. □

Proof of equation (2.35). In the weak dependent error case, the bias corrected estimator is $\tilde{\gamma}_n = \hat{\gamma}_n + \frac{\log\hat{\varphi}}{\log n}$ where $\hat{\varphi} = \hat{\sigma}^2/\hat{\omega}^2$ in which $\hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^n \hat{u}_t^2$ with $\hat{u}_t = X_t - \hat{\theta}_n X_{t-1}$ and $\hat{\omega}^2$ is a standard kernel estimator of the long run variance, ω^2, of u_t . Note that $\hat{\sigma}^2 - \sigma^2 = O_p\left(n^{-1/2}\right)$ and by standard HAC estimator asymptotics we have $\hat{\omega}^2_{tri} - \sigma^2 = O_p\left(n^{-1/3}\right)$ for the triangular lag kernel estimator and $\hat{\omega}^2_{quad} - \sigma^2 = O_p\left(n^{-2/5}\right)$ for the quadratic lag kernel estimator with optimal bandwidth choices. It follows directly that $\hat{\varphi} - \varphi = O_p(n^{-\kappa})$ with $\kappa \in \{1/3, 2/5\}$ for these two types of kernels. We then have

$$\begin{aligned}
&n^{\frac{1-\gamma_n}{2}}\log n\{\tilde{\gamma}_n - \gamma_n\} = n^{\frac{1-\gamma_n}{2}}\log n\left\{\hat{\gamma}_n + \frac{\log\hat{\varphi}}{\log n} - \gamma_n\right\} \\
&= n^{\frac{1-\gamma_n}{2}}\log n\left\{\hat{\gamma}_n - \gamma_n + \frac{\log\varphi}{\log n} + O_p(n^{-\kappa})\right\} \\
&= n^{\frac{1-\gamma_n}{2}}\log n\left\{\hat{\gamma}_n - \gamma_n + \frac{\log\varphi}{\log n}\right\} + O_p\left(n^{\frac{1-\gamma_n-2\kappa}{2}}\log n\right) \rightsquigarrow \xi =_d \mathcal{N}(0,2).
\end{aligned}$$

Since $n^{\gamma_n} = n^{\alpha - \log|c|/\log n} = n^{\alpha}/|c|$, we have $n^{\frac{1-\gamma_n-2\kappa}{2}} = n^{\frac{1-\alpha-2\kappa}{2}}|c|^{1/2}$, so that $n^{\frac{1-\gamma_n-2\kappa}{2}}\log n = o_p(1)$ whenever $\alpha > 1-2\kappa$ or $\alpha > \frac{1}{3}$ for $\kappa \geq 1/3$, which includes both triangular and quadratic lag kernel choices. It follows that the bias corrected estimator $\tilde{\gamma}_n = \hat{\gamma}_n + \frac{\log\hat{\varphi}}{\log n}$ satisfies

$$n^{\frac{1-\gamma_n}{2}}\log n\{\tilde{\gamma}_n - \gamma_n\} = n^{\frac{1-\gamma_n}{2}}\log n\left\{\hat{\gamma}_n - \gamma_n + \frac{\log\varphi}{\log n}\right\} + o_p(1) \rightsquigarrow \xi =_d \mathcal{N}(0,2), \tag{A-25}$$

for all $\alpha > \frac{1}{3}$, and the result holds for all standard consistent HAC estimators of ω^2 under the conditions of Theorem 2.3. □

Proof of Theorem 3.1. **Part (i):** Defining $\hat{A}_n = \hat{\theta}_n - 1$ as before, it follows that for $c > 0$

$$\hat{A}_n = \frac{c}{n^{\alpha}} + \frac{1}{n^{\alpha}}\frac{2c}{\theta_n^n}\mathbb{C}\{1+o_p(1)\} = \frac{c}{n^{\alpha}}\left[1 + \frac{2}{\theta_n^n}\mathbb{C}\{1+o_p(1)\}\right],$$

so that

$$\log\left|\hat{A}_n\right| \quad = \quad -\alpha\log n + \log|c| + \log\left[1 + \frac{2}{\theta_n^n}\mathbb{C}\{1+o_p(1)\}\right]$$

$$= \quad -\alpha \log n + \log |c| + \frac{2}{\theta_n^n} \mathbb{C} \{1 + o_p(1)\}.$$

Hence

$$\hat{\alpha} = -\frac{\log \left|\hat{A}_n\right|}{\log n} = \alpha - \frac{\log |c|}{\log n} - \frac{2}{\theta_n^n \log n} \mathbb{C} \{1 + o_p(1)\},$$

and then $\hat{\alpha} \to_p \alpha$ in both the expanded and original spaces. The convergence rate to α is $\log n$ but upon recentering and normalizing we have

$$\left(1 + \frac{c}{n^\alpha}\right)^n \log n \left\{\hat{\alpha} - \alpha + \frac{\log |c|}{\log n}\right\} \rightsquigarrow 2c\mathbb{C}, \tag{A-26}$$

as stated. Then, noting $\left(1 + \frac{c}{n^\alpha}\right)^n = e^{cn^{1-\alpha}} \left(1 - O(n^{1-2\alpha})\right) \sim_a e^{cn^{1-\alpha}}$ when $\alpha > \frac{1}{2}$, we have

$$e^{cn^{1-\alpha}} \log n \left\{\hat{\alpha} - \alpha + \frac{\log |c|}{\log n}\right\} \rightsquigarrow 2c\mathbb{C}, \tag{A-27}$$

giving (3.6).

Part (ii): When $c_n = \frac{1}{n^{1-\alpha-\delta}}$ with $0 < \delta \leq 1 - \alpha$, we have $\theta_n = 1 + \frac{1}{n^\delta}$. It then follows as in Part (i) that $\hat{\alpha} \to_p 1 - \delta$ as $n \to \infty$. Further, since $\log |c_n| = \log \frac{1}{n^{1-\alpha-\delta}} = (-1 + \alpha + \delta) \log n$, we have $\alpha - \frac{\log |c_n|}{\log n} = 1 - \delta$ and then

$$\left(1 + \frac{1}{n^{1-\delta}}\right)^n \log n \{\hat{\alpha} - (1 - \delta)\} \rightsquigarrow 2\mathbb{C}. \tag{A-28}$$

Noting that $\left(1 + \frac{1}{n^{1-\delta}}\right)^n = e^{n^\delta} \left(1 - O(\frac{1}{n^{1-2\delta}})\right) \sim_a e^{n^\delta}$ when $\delta < \frac{1}{2}$, we have

$$e^{n^{1-\delta}} \log n \{\hat{\alpha} - (1 - \delta)\} \rightsquigarrow 2\mathbb{C},$$

as in (3.7).

Part (ii) SV Extension: When the autoregressive coefficient $\theta_n = 1 + \frac{L_n}{n}$ differs from unity by a slowly varying function at infinity, L_n, we have $c_n = \frac{L_n}{n^{1-\alpha}} = \frac{n^{\frac{\log L_n}{\log n}}}{n^{1-\alpha}}$ so that

$$\theta_n = 1 + \frac{L_n}{n} = 1 + \frac{c_n}{n^\alpha} = 1 + \frac{1}{n^{1-\frac{\log L_n}{\log n}}} =: 1 + \frac{1}{n^{\gamma_n}},$$

with uni-parameter $\gamma_n = 1 - \frac{\log L_n}{\log n}$, as in (3.8). In the same way as in (A-28), the following limit theory then holds when $L_n \to \infty$ as $n \to \infty$

$$\left(1 + \frac{1}{n^{1-\frac{\log L_n}{\log n}}}\right)^n \log n \left(\hat{\alpha} - \left\{1 - \frac{\log L_n}{\log n}\right\}\right) \rightsquigarrow 2\mathbb{C}, \tag{A-29}$$

which can be rewritten entirely in terms of the uni-parameter γ_n as

$$\left(1+\frac{1}{n^{\gamma_n}}\right)^n \log n\,(\hat{\gamma}_n-\gamma_n)=\left(1+\frac{L_n}{n}\right)^n \log n\,(\hat{\gamma}_n-\gamma_n)\rightsquigarrow 2\mathbb{C}, \tag{A-30}$$

where $\hat{\gamma}_n=-\frac{\log\left|\hat{A}_n\right|}{\log n}$. In (A-30) the centering in (A-29) is absorbed in the uni-parameter $\gamma_n=1-\frac{\log L_n}{\log n}$ because when $c_n=\frac{L_n}{n^{1-\alpha}}$

$$\gamma_n=\alpha-\frac{\log c_n}{\log n}=\alpha-\frac{\log L_n-(1-\alpha)\log n}{\log n}=1-\frac{\log L_n}{\log n},$$

and $\frac{1}{n^{\gamma_n}}=\frac{L_n}{n}$. □

References

Chan, N. H. and C.-Z. Wei (1987): "Asymptotic inference for nearly nonstationary AR(1) processes," *The Annals of Statistics*, 1050–1063.

Duffy, J. and I. Kasparis (2021): "Regressions with fractional d= 1/2 and weakly nonstationary processes," Tech. rep., Mimeo, arXiv.

Elliott, G. (1998): "On the robustness of cointegration methods when regressors almost have unit roots," *Econometrica*, 66, 149–158.

Fisher, R. A. (1921): "On the 'probable error' of a coefficient of correlation deduced from a small sample," *Metron*, 1, 1–32.

Giraitis, L. and P. C. B. Phillips (2006): "Uniform limit theory for stationary autoregression," *Journal of Time Series Analysis*, 27, 51–60.

Han, C., P. C. B. Phillips, and D. Sul (2011): "Uniform asymptotic normality in stationary and unit root autoregression," *Econometric Theory*, 27, 1117–1151.

Hansen, B. E. (1999): "The grid bootstrap and the autoregressive model," *Review of Economics and Statistics*, 81, 594–607.

Hotelling, H. (1953): "New light on the correlation coefficient and its transforms," *Journal of the Royal Statistical Society. Series B (Methodological)*, 15, 193–232.

Jenkins, G. (1954): "An angular transformation for the serial correlation coefficient," *Biometrika*, 41, 261–265.

Konishi, S. (1981): "Normalizing transformations of some statistics in multivariate analysis," *Biometrika*, 68, 647–651.

Kostakis, A., T. Magdalinos, and M. P. Stamatogiannis (2015): "Robust econometric inference for stock return predictability," *The Review of Financial Studies*, 28, 1506–1553.

Lin, Y. and Y. Tu (2020): "Robust inference for spurious regressions and cointegrations involving processes moderately deviated from a unit root," *Journal of Econometrics*, 219, 52–65.

Mikusheva, A. (2007): "Uniform inference in autoregressive models," *Econometrica*, 75, 1411–1452.

——— (2012): "One-dimensional Inference in Autoregressive Models with the Potential Presence of a Unit Root," *Econometrica*, 80, 173–212.

Phillips, P. C. B. (1977): "Approximations to some finite sample distributions associated with a first-order stochastic difference equation," *Econometrica*, 45, 463–485.

——— (1979): "Normalizing transformation for the serial correlation coefficient $\hat{\alpha}$," *Research Note I and Note II - University of Birmingham.*

——— (1987a): "Time series regression with a unit root," *Econometrica*, 55, 277–301.

——— (1987b): "Towards a unified asymptotic theory for autoregression," *Biometrika*, 74, 535–547.

——— (1988): "Regression theory for near-integrated time series," *Econometrica*, 56, 1021–1043.

——— (2011): "Estimation of the Localizing Rate for Mildly Integrated and Mildly Explosive Processes," *Working Paper, Yale University.*

——— (2014): "On confidence intervals for autoregressive roots and predictive regression," *Econometrica*, 82, 1177–1195.

——— (2015): "Halbert White Jr. Memorial JFEC Lecture: Pitfalls and possibilities in predictive regression," *Journal of Financial Econometrics*, 13, 521–555.

Phillips, P. C. B. and T. Magdalinos (2007a): "Limit theory for moderate deviations from a unit root," *Journal of Econometrics*, 136, 115–130.

——— (2007b): "Limit Theory for Moderate Deviations from a Unit Root Under Weak Dependence," in *The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis*, ed. by G. D. A. Phillips and E. Tzavalis, Cambridge: Cambridge University Press, 123–162.

——— (2009): "Econometric inference in the vicinity of unity," *Singapore Management University, CoFie Working Paper*, 7.

Phillips, P. C. B., T. Magdalinos, and L. Giraitis (2010): "Smoothing local-to-moderate unit root theory," *Journal of Econometrics*, 158, 274–279.

Phillips, P. C. B., S. Shi, and J. Yu (2015a): "Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500," *International Economic Review*, 56, 1043–1078.

——— (2015b): "Testing for multiple bubbles: Limit theory of real-time detectors," *International Economic Review*, 56, 1079–1134.

Phillips, P. C. B. and V. Solo (1992): "Asymptotics for linear processes," *The Annals of Statistics*, 971–1001.

Phillips, P. C. B., Y. Wu, and J. Yu (2011): "Explosive Behavior in the 1990s NASDAQ: When did Exuberance Escalate Asset Values?" *International Economic Review*, 52, 201–226.

Shi, S. and P. C. B. Phillips (2021): "Diagnosing housing fever with an econometric thermometer," *Journal of Economic Surveys, forthcoming.*

Stock, J. H. (1991): "Confidence intervals for the largest autoregressive root in US macroeconomic time series," *Journal of Monetary Economics*, 28, 435–459.

Winterbottom, A. (1979): "A note on the derivation of Fisher's transformation of the correlation coefficient," *The American Statistician*, 33, 142–143.