A test for strict stationarity in a random coefficient autoregressive model of order 1

Lorenzo Trapani*

* *University of Nottingham*

e-mail: lorenzo.trapani@nottingham.ac.uk

Abstract: We propose a test for the null of strict stationarity in a Random Coefficient AutoRegression (RCAR) of order 1. The test can also be used in the case of a standard AR(1) model, and it can be applied under minimal requirements on the existence of moments - in both cases without requiring any modifications or prior knowledge.

Keywords: Random Coefficient AutoRegression, Stationarity, Heavy Tails.

AMS 2000 subject classification: Primary 62F05; secondary 62M10.

1. Introduction

In this paper, we propose a test for the null of strict stationarity of a series generated by a Random Coefficient AutoRegressive (RCAR) model of order 1:

$$X_t = (\varphi + b_t)\, X_{t-1} + e_t, \quad t \in \mathbb{Z} < \infty. \tag{1.1}$$

Our procedure requires minimal moment conditions, and it can be applied as long as $\sup_t E\,|X_t|^{\delta} < \infty$ for some $\delta > 0$. Further, the test can also be used when $b_t = 0$ a.s., thus being valid also in a standard autoregressive set-up. No prior knowledge is needed as to whether the variance of X_t is finite or not, or as to whether the autoregressive coefficient is random or not.

The RCAR model was firstly proposed by Anděl (1976), and subsequently extensively studied in the monograph by Nicholls and Quinn (2012); see also the recent review by Regis et al. (2021). Since its initial appearance, equation (1.1) has attracted considerable attention by the literature, mainly due to its flexibility. Indeed, (1.1) nests the AR(1) model as a special case, with the advantage that it can be viewed as a more flexible competitor for a model with an abrupt break in the autoregressive root (see, especially, a related paper by Giraitis et al., 2014). (1.1) also belongs in the class of conditional heteroskedasticity models: Tsay (1987) showed that the popular ARCH model can be viewed as a second-order equivalent to an RCAR model; further, (1.1) is also closely related to the Double

AutoRegressive (DAR) model $X_t = \varphi X_{t-1} + v_t$ with $v_t = \sqrt{a + bX_{t-1}^2}\epsilon_t$. Finally, a restricted version of (1.1) with $\varphi = 1$ has become a popular alternative to deterministic unit root processes - this is known in the literature as the Stochastic Unit Root (STUR) process, and we refer to the contributions by Granger and Swanson (1997), McCabe and Tremayne (1995) and Leybourne et al. (1996), among others, for an overview. On account of such flexibility, (1.1) has been used ubiquitously in applied sciences, with several examples in such diverse fields as biology (Stenseth et al., 1998), medicine (Fryz, 2017), and physics (Ślezak et al., 2019). The RCAR model has also been applied successfully in the analysis of economic and financial data, where the STUR variant of (1.1) is viewed as a convenient way of modelling possibly nonstationary time series: see, *inter alia*, Tsay (2005) and Banerjee et al. (2013).

Some aspects of the inference on (1.1) are generally well-developed. As far as estimation of φ is concerned, Quasi-ML estimation has been developed by Aue et al. (2006) and Berkes et al. (2009); Koul and Schick (1996) and Horváth and Trapani (2016) study Weighted Least Squares (WLS) estimation, whereas Hill and Peng (2014) and Hill et al. (2016) consider the Empirical Likelihood (EL) estimator. In addition, Akharif and Hallin (2003), Nagakura (2009) and Horváth and Trapani (2019) propose tests to check whether the autoregressive coefficient is genuinely random.

Other aspects of the inference on (1.1) are not fully established. In particular, only few results are available as far as testing for the stationarity/ergodicity of X_t is concerned. This is in stark contrast with the very well-developed literature on unit roots in the AR(1) case, and also with the rest of the literature on nonlinear models (see e.g. Tsay, 1997, and Kapetanios et al., 2003, *inter alia*). Indeed, most contributions which refer to the RCAR model focus on restricted versions of (1.1), e.g. testing whether X_t is a genuine unit root process versus the alternative of a STUR process - see McCabe and Tremayne (1995), Leybourne et al. (1996), Distaso (2008) and Nagakura (2009). Moreover, even in the AR(1) case, testing procedures usually require the existence of at least the first two moments (see however Phillips, 1990).

To the best of our knowledge, there are very few contributions which propose a test for the stationarity of X_t as generated by (1.1). All such contributions have one (restrictive) assumption in common: the finiteness of the second moment of the data. Indeed, available tests for the RCAR model are typically based on verifying the necessary and sufficient condition for stationarity (see Nicholls and Quinn, 2012) $\varphi^2 + E\left(b_0^2\right) < 1$. For example, Aue and Horváth (2011) propose a test based on QMLE, and Zhao and Wang (2012) study the Empirical Likelihood estimator for (1.1), developing, as a by-product, a test for the null of stationarity. In addition to requiring finite variance, both tests also require that $E(b_0^2) > 0$, thus not being valid in the case of a pure AR model. Trapani (2021) proposes a procedure to decide between the null of (strict) stationarity and the alternative of nonstationarity, which does not require having finite second moments and can be used irrespective of whether $b_t = 0$ a.s. or not. In particular, Trapani (2021) constructs a family of statistics which converge to a nonzero constant under the null, and drift to zero under the alternative. Exploiting such rates, a randomised test is proposed, thus obtaining a test statistic which converges to a well-defined limiting law under the null, and diverges under the alternative. Despite its generality, the procedure proposed by Trapani (2021) suffers from one major drawback: the randomness added by each researcher in constructing the test statistic does not vanish asymptotically. This has two potentially undesirable consequences. Firstly, different researchers using the same data will have different outcomes, dashing hopes of reproducibility - in fact, it could be shown that if an infinite number of tests were carried out, the p-values would follow a $U[0,1]$ distribution. Secondly, the randomised test is constructed in a non-traditional way, in that the randomisation amounts to reject randomly H_0, with probability 5%, whenever the test statistic takes a small value, whatever the evidence in favour of the null. Conversely, the test proposed in this contribution is constructed so as to reject with asymptotic probability 5% at the boundary of the null hypothesis, but with asymptotic probability 0 at a point belonging to the interior of the null hypothesis.

This note complements the paper by Trapani (2021). Under the same assumptions, we develop a test which does not require a randomisation, thus yielding the same outcome for all researchers. The test is based on the same statistic as in Trapani (2021), say D_T. By proving a CLT for the suitably normed $D_T - E(D_T)$, a test is proposed for the null that $E(D_T) > 0$, which corresponds to strict stationarity, versus the alternative that $E(D_T) = 0$, which corresponds to nonstationarity. The relevant limiting law is shown to hold for all cases considered (finite or infinite variance, random or nonrandom autoregressive root), with no need for estimation of nuisance parameters. The corresponding testing procedure is shown to ensure pointwise size control, and a small scale Monte Carlo experiment shows that the test has good size and power.

NOTATION Henceforth, c_0, c_1,... denote positive and finite constants that do not depend on the sample size (unless otherwise stated), and whose value may change from line to line. We use "$\rightarrow$" to denote the ordinary limit; "$\stackrel{\mathcal{D}}{\rightarrow}$" to denote convergence in distribution; "$\stackrel{\mathcal{D}}{=}$" to indicate equality in distribution; "a.s." stands for almost surely.

2. Testing for strict stationarity

As mentioned in the introduction, we propose a test for

$$\begin{cases} H_0: & X_t \text{ is strictly stationary} \\ H_A: & X_t \text{ is nonstationary} \end{cases} \tag{2.1}$$

Under the assumption that logarithmic moments exist for e_0 and $\varphi + b_0$, the strict stationarity of X_t depends on whether $\theta = E \ln |\varphi + b_0|$ is negative or not. Indeed, when $-\infty \leq \theta < 0$, then it is well known (Aue et al., 2006) that X_t is strictly stationary, in the sense that it converges (exponentially fast and for all initial values X_0) to a strictly stationary solution. Conversely, if $\theta \geq 0$, then X_t is nonstationary. In particular, when $\theta > 0$, X_t exhibits an explosive behaviour, i.e. $\exp(-C_0 t) |X_t| \to \infty$ a.s. for all $0 < C_0 < \theta$ (see e.g. Corollary 1 in Berkes et al., 2009). On the other hand, when $\theta = 0$, X_t is also nonstationary, although in this case $|X_t|$ diverges at a slower rate than exponential (Horváth and Trapani, 2016; Horváth and Trapani, 2019).

We require the following assumptions, which are the same as in Trapani (2021).

Assumption 1. *It holds that: (i) $\{b_t, -\infty < t < \infty\}$ and $\{e_t, -\infty < t < \infty\}$ are independent sequences; (ii) $\{b_t, -\infty < t < \infty\}$ are independent and identically distributed random variables; (iii) $\{e_t, -\infty < t < \infty\}$ are independent and identically distributed random variables; (iv) b_0 and e_0 are symmetric random variables; (v) $E |b_0|^{\nu} < \infty$ and $E |e_0|^{\nu} < \infty$ for some $\nu > 0$; (vi) X_0 is independent of $\{e_t, b_t, t \geq 1\}$ with $E |X_0|^{\nu} < \infty$.*

Assumption 2. *If $\theta < 0$, it holds that (i) $P\left(|\bar{X}_0| = 0\right) < 1$; (ii) $P\left(|e_0| = 0\right) < 1$.*

Assumption 3. *If $\theta \geq 0$, it holds that: (i) e_0 has bounded density; (ii) when $P\left(b_0 = 0\right) < 1$, $E \left|\ln |\varphi + b_0|\right|^k < \infty$ for some $k > 2$; (iii) $EX_0^2 < \infty$.*

Assumption 4. *When $\theta = 0$ with $b_0 = 0$ a.s., it holds that either (i) $E |X_0|^2 < \infty$ and $E |e_0|^{\nu'} < \infty$ for some $\nu' > 2$; or (ii) (a) $\{e_t, -\infty < t < \infty\}$ are symmetric random variables with common distribution $F(x)$ such that*

$$1 - F(x) = C_0 x^{-\gamma} + \varsigma(x) x^{-\gamma}, \quad x \geq x_0,$$

with $C_0 > 0$, $\gamma \in (0, 2]$ and $\varsigma(x) \to 0$ as $x \to \infty$, with $\varsigma(x) x^{-\gamma}$ decreasing for all $x \geq x_0$; and (b) $E |X_0|^{\gamma'} < \infty$ for all $\gamma' < \gamma$.

Assumptions 1-4 are the same as in Trapani (2021). Comments and examples are in the Supplement.

In order to propose a "universal" test for H_0 in (2.1), recall that - under all possible circumstances, i.e. whether $b_t = 0$ a.s. or not and/or whether $E|X_t|^2 < \infty$ or not - it holds that $|X_t| \to \infty$ a.s. or not according as X_t is non-stationary or stationary. Given a sample $1 \leq t \leq T$, and along the same lines as in Trapani (2021), therefore, we propose using the transformation

$$Y_t(a) = \frac{a}{a + X_t^2}, \tag{2.2}$$

with $0 < a < \infty$. In particular, Trapani (2021) shows that, as $T \to \infty$

$$\frac{1}{T}\sum_{t=1}^{T} Y_t(a) = \left\{ \begin{array}{ll} c_0(a) + o_{a.s.}(1) \\ O(T^{-\epsilon}) \end{array} \right. \text{ according as } \begin{array}{c} \theta < 0 \\ \theta \geq 0 \end{array}, \tag{2.3}$$

with $0 < c_0(a) < \infty$ and for all $\epsilon > 0$ and all $a > 0$.

In order to ensure scale invariance, we propose

$$Y_t = \int_0^{\infty} Y_t(a)\, dF(a), \tag{2.4}$$

where $F(a)$ is a distribution.

Theorem 1. *We assume that Assumptions 1 and 4 are satisfied. Then, as $T \to \infty$, under Assumption 2 it holds that*

$$\frac{\sum_{t=1}^{T}(Y_t - E(Y_t))}{\left(E\left(\sum_{t=1}^{T}(Y_t - EY_t)\right)^2\right)^{1/2}} \xrightarrow{\mathcal{D}} Z, \tag{2.5}$$

with $E(Y_t) > 0$, where $Z \sim N(0,1)$. Under Assumption 3, it holds that, for all $\beta \geq 1$

$$\frac{1}{T}\sum_{t=1}^{T} Y_t^{\beta} = o_{a.s.}(1). \tag{2.6}$$

Based on Theorem 1, a test for the null of strict stationarity could be based on the one-sided α-level confidence interval

$$CI_{\alpha} = \left(T^{-1}\sum_{t=1}^{T} Y_t - c_{\alpha}T^{-1/2}V^{1/2}, \infty\right),$$

where

$$V = \lim_{T\to\infty} E\left(T^{-1/2}\sum_{t=1}^{T}(Y_t - EY_t)\right)^2.$$

If CI_{α} contains 0, this indicates that X_t is nonstationary, whereas if it does not, then it may be concluded that X_t is stationary. Equation (2.5) ensures that, as $T \to \infty$, under H_0 it holds that $P(0 \in CI_{\alpha}) = \alpha$ for all $\theta < 0$; similarly, by (2.6), as $T \to \infty$, under H_A it holds that $P(0 \in CI_{\alpha}) = 1$.

In order to make the test feasible, we propose a (standard) weighted-sum-of-covariances estimator for V

$$\widehat{V}=\widehat{r}_0+2\sum_{j=1}^{H}\left(1-\frac{j}{H+1}\right)\widehat{r}_j,$$

having defined

$$\widehat{r}_j=\frac{1}{T}\sum_{t=j+1}^{T}\left(Y_t-\frac{1}{T}\sum_{t=1}^{T}Y_t\right)\left(Y_{t-j}-\frac{1}{T}\sum_{t=1}^{T}Y_t\right),$$

for $j=0,1,\ldots$.

Lemma 1. *We assume that Assumptions 1 and 2 are satisfied. Then, as* $\min(H,T)\to\infty$ *with* $\frac{H^3}{T}\to 0$, *it holds that* $\widehat{V}=V+o_P(1)$. *Under Assumptions 1 and 3-4, it holds that* $T^{-1/2}\widehat{V}=o_P(1)$.

3. Simulations

We assess the performance of the proposed method through a small Monte Carlo exercise, using (1.1) as a DGP. We generate b_t and e_t as independent of each other and *i.i.d.* with a power law distribution with tail index γ - that is, letting $F(x)$ denote the common distribution of b_0 and e_0, $F(x)=Cx^{-\gamma}$; we simulate the power law distribution as indicated in Clauset et al. (2009), by setting

$$\epsilon_t = \sigma_\epsilon\left(1-v_{e,t}\right)^{-1/\gamma}, \tag{3.1}$$

$$b_t = \sigma_b\left(1-v_{b,t}\right)^{-1/\gamma}, \tag{3.2}$$

where $v_{e,t}$ and $v_{b,t}$ are generated independently and as *i.i.d.* from a uniform distribution on the interval $[0,1]$, and then we center ϵ_t and b_t around zero; σ_ϵ and σ_b are scale parameters. Across all experiments, we have used $\gamma=1.2$, and 1.5, and we have also considered, as benchmark, the case where b_t and e_t are generated as $N\left(0,\sigma_b^2\right)$ and $N(0,1)$, respectively. As far as the implementation of the test is concerned, we have used $a=\pm 0.5$, drawn with equal probability. Preliminary experiments show that using $c_\alpha=1.64$ (corresponding to a 95% confidence interval) works well in general, but it yields very low power in the unit root case, where $\varphi=1$ and $b_t=0$ a.s.; in order to enhance power, we propose a less conservative approach based on using the interval $\left(T^{-1}\sum_{t=1}^{T}Y_t-\sqrt{\frac{\ln T}{T}}V^{1/2},\infty\right)$. In addition to reporting results for this test, we also consider, by way of comparison, the test developed by Trapani (2021), the test proposed in Zhao and Wang (2012), and also a test, based on the QML estimator, proposed in Aue and Horváth (2011). We point out that the hypothesis testing in Zhao and Wang (2012) (similarly to Aue and Horváth, 2011) is spelt out as

$$\begin{aligned} H_0 &: \varphi^2+\sigma_b^2\leq 1-h \\ H_A &: \varphi^2+\sigma_b^2\geq 1+h, \end{aligned} \tag{3.3}$$

where $h > 0$ is a user defined truncation (in our experiments, we have used $h = 1/8$, which is one of the values proposed in Zhao and Wang, 2012). Based on this, it can be anticipated that the test may have size distortion when $\varphi^2 + \sigma_b^2$ is smaller than 1, but close to the boundary. Finally, the number or replications is set equal to $1,000$ - this entails that, when considering empirical rejection frequencies under the null, these have a confidence interval of $[0.036, 0.064]$.

Table 1 reports results for "ordinary" cases, where the null and the alternative are not on the boundary. In these case, the test has the correct behaviour under both the null and the alternative. Indeed, the test has very low probability of Type 1 error for all cases (with the partial exception of the Gaussian case with $(\varphi, \sigma_b) = (1.05, 0.1)$, although as T increases this distortion vanishes) under the null. The power of the testing procedure is satisfactorily high whenever $\theta > 0$, and it seems to be only marginally affected by γ; note that the test by Zhao and Wang (2012), especially for smaller values of T, has higher power, although in some cases the power of that test, puzzlingly, dips as T increases. Note also that the test by Trapani (2021) does not have vanishing empirical rejection frequencies for θ very faraway from 1. This is due to the fact that, as mentioned in the introduction, the test is constructed so as to reject the null with 5% probability irrespective of whether θ is very close to the boundary (i.e., $\theta = -\epsilon$) or well inside the parameter space defined by the null (i.e. $\theta << 0$).

Boundary cases are explored in Table 2. The results show that the test has the best performance among all tests considered as far as probability of Type 1 error is concerned, at least when $T \geq 1,000$. In general, it can be noted that the empirical rejection frequencies decline as T increases, but less quickly as γ decreases. The power of the testing procedure is satisfactorily high for all cases considered, and it seems to be only marginally affected by γ. Interestingly, as could be predicted on the grounds of (3.3), the test by Zhao and Wang (2012) has severe size distortion when the null is satisfied but very close to the boundary; conversely, their test is very powerful versus the alternative, but this is bound to be affected by the size distortion. In addition to studying several boundary cases, in Table 3, we have run a small scale experiment to assess the behaviour of the test in the case of the popular STUR specification, where $\varphi = 1$, but having $\sigma_b > 0$ may entail that the series is stationary. Again, the test has low rejection frequency under the null and satisfactory power - indeed, the test by Zhao and Wang (2012) has higher power, but occasionally, especially when θ is close to zero but negative, this is at the expense of size. Interestingly, in the STUR set-up, the procedure based on Aue and Horváth (2011) does not seem to work even in the Gaussian case.

Finally, in the Supplement we report further evidence concerning the case of a time-varying, deterministic coefficient[1].

Appendix A: Proofs and technical lemmas

We report some preliminary lemmas and the proof of Theorem 1. The proofs of the lemmas, and the proof of Lemma 1, which is quite repetitive, are in the Supplement.

We know from the proof of Lemma 2 in Aue et al. (2006) that there exists a $k \in (0, \nu)$ such that $0 \leq \delta = E\left|\varphi + b_0\right|^k < 1$. Also, when $-\infty \leq \theta < 0$

$$\overline{X}_t = \sum_{s=-\infty}^{t} e_s \prod_{z=s+1}^{t} \left(\varphi + b_z\right), \tag{A.1}$$

converges a.s.; the process $\left\{\overline{X}_t, t \in \mathbb{Z}\right\}$ is the unique, stationary, causal solution of (1.1) - see Aue et al. (2006). In the proofs, we will also use the construction

$$\widehat{X}_t = \sum_{s=t-\lfloor \beta \ln T \rfloor}^{t} e_s \prod_{z=s+1}^{t} \left(\varphi + b_z\right) + \sum_{s=-\infty}^{t-\lfloor \beta \ln T \rfloor - 1} \widehat{e}_s \prod_{z=s+1}^{t} \left(\varphi + \widehat{b}_z\right), \tag{A.2}$$

where β is chosen as

$$\beta = \frac{1 + k + \epsilon}{1 - \delta}, \tag{A.3}$$

with $\epsilon > 0$, and $\left\{\widehat{e}_t, \widehat{b}_t\right\}_{t=-\infty}^{\infty}$ are completely independent and independent of $\{e_t, b_t\}_{t=-\infty}^{\infty}$, with $\widehat{e}_t \overset{\mathcal{D}}{=} e_t$ and $\widehat{b}_t \overset{\mathcal{D}}{=} b_t$. Clearly, $\widehat{X}_t \overset{\mathcal{D}}{=} \overline{X}_t$. Finally, we define the short-hand notation

$$z_t' = \int \frac{a}{a + \overline{X}_t^2} dF(a) - E \int \frac{a}{a + \overline{X}_0^2} dF(a). \tag{A.4}$$

$$z_t = \int \frac{a}{a + \widehat{X}_t^2} dF(a) - E \int \frac{a}{a + \overline{X}_0^2} dF(a). \tag{A.5}$$

We now report some preliminary lemmas.

Lemma A.1. *We assume that Assumptions 1 and 4 are satisfied. Then it holds that*

$$\sum_{t=1}^{T} \left(\int \frac{a}{a + \widehat{X}_t^2} dF(a) - \int \frac{a}{a + \overline{X}_t^2} dF(a) \right) = O_P\left(T \delta^{\lfloor \beta \ln T \rfloor / (1+k)}\right) = O\left(T^{-\epsilon}\right). \tag{A.6}$$

Lemma A.2. *We assume that Assumptions 1 and 4 are satisfied. Then it holds that*

$$E\left(\sum_{t=1}^{T} z_t'\right)^2 = O(T). \tag{A.7}$$

Let $B_T^2 = Var\left(\sum_{t=1}^{T} z_t\right)$.

[1] I wish to thank a Referee for asking the question which led to this idea.

Lemma A.3. *We assume that Assumptions 1 and 4 are satisfied. Then it holds that*

$$E\left(\sum_{t=1}^{m} z_t\right)^2 = O\left(m^2\right), \tag{A.8}$$

$$B_T^2 = O\left(T \ln T\right), \tag{A.9}$$

$$B_T^2 \geq c_0 T. \tag{A.10}$$

Lemma A.4. *We assume that Assumptions 1 and 4 are satisfied. Then it holds that*

$$\frac{\sum_{t=1}^{T} z_t'}{\left[E\left(\sum_{t=1}^{T} z_t'\right)^2\right]^{1/2}} = \frac{\sum_{t=1}^{T} z_t}{\left[E\left(\sum_{t=1}^{T} z_t\right)^2\right]^{1/2}} + o_P\left(1\right).$$

Lemma A.5. *We assume that Assumptions 1 and 4 are satisfied. Then it holds that*

$$\frac{\sum_{t=1}^{T} z_t}{\left[E\left(\sum_{t=1}^{T} z_t\right)^2\right]^{1/2}} \stackrel{\mathcal{D}}{\to} N\left(0,1\right).$$

Let $\widetilde{z}_t = Y_t - E\left(Y_t\right)$; then equation (7.7) in Horváth and Trapani (2016) yields $E\left|\widetilde{z}_t - z_t'\right| = O\left(\delta^t\right)$, for some $0 < \delta < 1$, so that it follows immediately that

$$\sum_{t=1}^{T} \widetilde{z}_t = \sum_{t=1}^{T} z_t' + O_P\left(1\right). \tag{A.11}$$

Also note that

$$E\left(\sum_{t=1}^{T} \widetilde{z}_t\right)^2 - E\left(\sum_{t=1}^{T} z_t'\right)^2 = E\left[\left(\sum_{t=1}^{T} \left(\widetilde{z}_t + z_t'\right)\right)\left(\sum_{t=1}^{T} \left(\widetilde{z}_t - z_t'\right)\right)\right]$$
$$\leq \left[E\left(\sum_{t=1}^{T} \left(\widetilde{z}_t + z_t'\right)\right)^2\right]^{1/2} \left[E\left(\sum_{t=1}^{T} \left(\widetilde{z}_t - z_t'\right)\right)^2\right]^{1/2}.$$

We now have that

$$E\left(\sum_{t=1}^{T} \left(\widetilde{z}_t - z_t'\right)\right)^2 = E\left(\sum_{t=1}^{T}\sum_{s=1}^{T} \left(\widetilde{z}_t - z_t'\right)\left(\widetilde{z}_s - z_s'\right)\right) \leq \sum_{t=1}^{T}\sum_{s=1}^{T} \left[E\left(\widetilde{z}_t - z_t'\right)^2\right]^{1/2} \left[E\left(\widetilde{z}_s - z_s'\right)^2\right]^{1/2} \tag{A.12}$$
$$\leq \left(\sum_{t=1}^{T} \left[E\left(\widetilde{z}_t - z_t'\right)^2\right]^{1/2}\right)^2 \leq \left(2\sum_{t=1}^{T} \left(E\left|\widetilde{z}_t - z_t'\right|\right)^{1/2}\right)^2 = O\left(1\right),$$

by the same logic as above. Also, noting that $E\left(\sum_{t=1}^{T} \left(\widetilde{z}_t + z_t'\right)\right)^2 \leq 2E\left(2\sum_{t=1}^{T} z_t'\right)^2 + 2E\left(\sum_{t=1}^{T} \left(\widetilde{z}_t - z_t'\right)\right)^2$ $= O\left(T\right)$, by virtue of Lemma A.2 and (A.12), it follows that

$$E\left(\sum_{t=1}^{T} \widetilde{z}_t\right)^2 = E\left(\sum_{t=1}^{T} z_t'\right)^2 + O\left(T^{1/2}\right). \tag{A.13}$$

Thus, (A.11) and (A.13) yield

$$\frac{\sum_{t=1}^{T}\widetilde{z}_t}{\left[E\left(\sum_{t=1}^{T}\widetilde{z}_t\right)^2\right]^{1/2}}=\frac{\sum_{t=1}^{T}z_t'}{\left[E\left(\sum_{t=1}^{T}z_t'\right)^2\right]^{1/2}}+o_P(1). \tag{A.14}$$

Lemmas A.4 and A.5 now yield the desired result.

We now show that

$$E\left(\frac{a}{a+\widehat{X}_0^2}\right)>0. \tag{A.15}$$

Indeed, for all $0<c_0<\infty$

$$\begin{aligned}E\left(\frac{a}{a+\widehat{X}_0^2}\right) &= E\left(\frac{a}{a+\widehat{X}_0^2}|\widehat{X}_0^2\geq c_0\right)P\left(\widehat{X}_0^2\geq c_0\right)+E\left(\frac{a}{a+\widehat{X}_0^2}|\widehat{X}_0^2<c_0\right)P\left(\widehat{X}_0^2<c_0\right)\\ &\geq E\left(\frac{a}{a+\widehat{X}_0^2}|\widehat{X}_0^2<c_0\right)P\left(\widehat{X}_0^2<c_0\right)\geq c_1P\left(\widehat{X}_0^2<c_0\right).\end{aligned}$$

Assumption 2*(i)* immediately entails that there exists a $d>0$ such that $E\left|\overline{X}_0\right|^d>0$; thus, $P\left(\widehat{X}_0^2\geq c_0\right)\leq c_0^{-d/2}E\left|\overline{X}_0\right|^d$. Upon choosing $c_0^{d/2}=b+E\left|\overline{X}_0\right|^d$, for some $b>0$, it follows that $P\left(\widehat{X}_0^2\geq c_0\right)<1$, which yields (A.15). This readily entails that $E\left(\int\frac{a}{a+\widehat{X}_0^2}dF(a)\right)>0$; in turn, this entails $E(Y_t)>0$.

Finally, (2.6) has been shown in Lemma 10 in Trapani (2021).

TABLE 1
Empirical rejection frequencies

		Gaussian case															
			$T=500$				$T=1000$				$T=2000$				$T=4000$		LT
		This paper	ZW	AH	LT	*This paper*	ZW	AH	LT	*This paper*	ZW	AH	LT	*This paper*	ZW	AH	LT
$(\varphi,\sigma_b)=(0.5,0.1)$	$\theta=-0.922$	0.000	0.000	0.000	0.055	0.000	0.000	0.000	0.058	0.000	0.000	0.000	0.051	0.000	0.000	0.000	0.051
$(\varphi,\sigma_b)=(0.5,0.0)$	$\theta=-0.693$	0.000	0.000	0.000	0.045	0.000	0.000	0.000	0.049	0.000	0.000	0.000	0.054	0.000	0.000	0.000	0.051
$(\varphi,\sigma_b)=(0.75,0.1)$	$\theta=-0.412$	0.000	0.000	0.000	0.048	0.000	0.000	0.000	0.052	0.000	0.000	0.000	0.050	0.000	0.000	0.000	0.051
$(\varphi,\sigma_b)=(1.05,0.1)$	$\theta=-0.007$	0.080	0.130	0.996	0.335	0.076	0.090	1.000	0.480	0.050	0.050	1.000	0.592	0.022	0.005	1.000	0.972
$(\varphi,\sigma_b)=(1.05,0.0)$	$\theta=0.049$	0.742	1.000	1.000	0.980	0.887	1.000	1.000	0.998	0.951	0.566	1.000	1.000	0.958	0.766	1.000	1.000
		$\gamma=1.5,\sigma_e=0.1$															
			$T=500$				$T=1000$				$T=2000$				$T=4000$		
		This paper	ZW	AH	LT	*This paper*	ZW	AH	LT	*This paper*	ZW	AH	LT	*This paper*	ZW	AH	LT
$(\varphi,\sigma_b)=(0.5,0.1)$	$\theta=-1.801$	0.006	0.001	0.029	0.053	0.000	0.000	0.048	0.047	0.006	0.000	0.081	0.051	0.001	0.000	0.131	0.046
$(\varphi,\sigma_b)=(0.5,0.0)$	$\theta=-0.693$	0.000	0.000	0.001	0.045	0.000	0.000	0.002	0.053	0.000	0.000	0.005	0.059	0.000	0.000	0.007	0.046
$(\varphi,\sigma_b)=(0.75,0.1)$	$\theta=-0.921$	0.006	0.001	0.096	0.048	0.000	0.000	0.125	0.052	0.004	0.000	0.195	0.058	0.000	0.000	0.240	0.046
$(\varphi,\sigma_b)=(1.05,0.1)$	$\theta=-0.299$	0.008	0.001	0.464	0.062	0.001	0.002	0.472	0.044	0.003	0.000	0.495	0.060	0.000	0.000	0.495	0.046
$(\varphi,\sigma_b)=(1.05,0.0)$	$\theta=0.049$	0.969	0.974	0.950	0.999	0.992	1.000	0.984	0.998	0.998	0.788	0.983	1.000	1.000	1.000	1.000	1.000
		$\gamma=1.2,\sigma_e=0.1$															
			$T=500$				$T=1000$				$T=2000$				$T=4000$		
		This paper	ZW	AH	LT	*This paper*	ZW	AH	LT	*This paper*	ZW	AH	LT	*This paper*	ZW	AH	LT
$(\varphi,\sigma_b)=(0.5,0.1)$	$\theta=-0.514$	0.002	0.009	0.172	0.044	0.001	0.000	0.187	0.042	0.001	0.000	0.250	0.059	0.009	0.001	0.270	0.048
$(\varphi,\sigma_b)=(0.5,0.0)$	$\theta=-0.693$	0.012	0.000	0.025	0.046	0.014	0.000	0.027	0.047	0.002	0.000	0.063	0.058	0.001	0.000	0.084	0.048
$(\varphi,\sigma_b)=(0.75,0.1)$	$\theta=-1.069$	0.001	0.003	0.210	0.039	0.001	0.000	0.215	0.041	0.009	0.000	0.253	0.057	0.008	0.000	0.250	0.048
$(\varphi,\sigma_b)=(1.05,0.1)$	$\theta=-0.959$	0.000	0.000	0.309	0.042	0.000	0.000	0.299	0.040	0.000	0.000	0.301	0.055	0.000	0.000	0.275	0.048
$(\varphi,\sigma_b)=(1.05,0.0)$	$\theta=0.049$	0.991	1.000	0.918	0.994	1.000	1.000	0.954	1.000	0.999	0.704	0.955	1.000	1.000	1.000	1.000	1.000

The tests considered in the table are, in addition to the one developed in this paper: Zhao and Wang (2012) (denoted as ZW); the one based on the QML estimator studied in Aue and Horváth (2011) (denoted as AH); and the one developed in Trapani (2021) (denoted as LT).
As in the remainder of the paper, we use the notation $\vartheta = E\ln|\phi+b_0|$. Routines are written using Gauss 21. In all experiments, $T+1,000$ datapoints have been generated, discarding the first $1,000$ observations to remove dependence on initial conditions.

TABLE 2
Empirical rejection frequencies

			$T = 500$				$T = 1000$				$T = 2000$				$T = 4000$		
		This paper	ZW	AH	LT	*This paper*	ZW	AH	LT	*This paper*	ZW	AH	LT	*This paper*	ZW	AH	LT
Gaussian case																	
$(\varphi, \sigma_b) = (1, 0.0)$	$\theta = 0.000$	0.659	0.707	0.192	0.526	0.680	0.823	0.221	0.688	0.692	0.812	0.200	0.755	0.698	0.876	0.205	0.904
$(\varphi, \sigma_b) = (1, 0.05)$	$\theta = -0.001$	0.087	0.104	0.638	0.066	0.029	0.122	0.833	0.055	0.000	0.093	0.964	0.058	0.000	0.082	1.000	0.048
$(\varphi, \sigma_b) = (1, 0.1)$	$\theta = -0.062$	0.000	0.000	0.819	0.081	0.000	0.002	0.957	0.068	0.000	0.000	0.999	0.054	0.000	0.001	1.000	0.048
$(\varphi, \sigma_b) = (1, 0.2)$	$\theta = -0.141$	0.000	0.001	0.934	0.058	0.000	0.002	0.999	0.062	0.000	0.001	1.000	0.059	0.000	0.000	1.000	0.048
$\gamma = 1.5, \sigma_e = 0.1$																	
$(\varphi, \sigma_b) = (1, 0.0)$	$\theta = 0.000$	0.409	0.740	0.064	0.566	0.505	0.795	0.076	0.703	0.543	0.811	0.106	0.786	0.639	0.838	0.149	0.876
$(\varphi, \sigma_b) = (1, 0.05)$	$\theta = -0.030$	0.004	0.000	0.361	0.073	0.001	0.000	0.401	0.064	0.001	0.000	0.433	0.058	0.000	0.000	0.492	0.049
$(\varphi, \sigma_b) = (1, 0.1)$	$\theta = -0.381$	0.004	0.001	0.340	0.071	0.002	0.000	0.422	0.071	0.002	0.000	0.457	0.055	0.000	0.000	0.458	0.049
$(\varphi, \sigma_b) = (1, 0.2)$	$\theta = -0.707$	0.014	0.001	0.354	0.053	0.009	0.000	0.406	0.055	0.003	0.001	0.445	0.061	0.000	0.000	0.461	0.049
$\gamma = 1.2, \sigma_e = 0.1$																	
$(\varphi, \sigma_b) = (1, 0.0)$	$\theta = 0.000$	0.428	0.684	0.098	0.554	0.521	0.691	0.116	0.683	0.655	0.712	0.137	0.779	0.737	0.991	0.168	1.000
$(\varphi, \sigma_b) = (1, 0.05)$	$\theta = -0.089$	0.037	0.000	0.290	0.059	0.026	0.001	0.322	0.064	0.020	0.000	0.304	0.057	0.028	0.000	0.329	0.046
$(\varphi, \sigma_b) = (1, 0.1)$	$\theta = -1.019$	0.034	0.003	0.271	0.079	0.039	0.001	0.292	0.072	0.004	0.000	0.328	0.063	0.004	0.000	0.315	0.046
$(\varphi, \sigma_b) = (1, 0.2)$	$\theta = -0.677$	0.009	0.004	0.301	0.070	0.007	0.000	0.321	0.066	0.001	0.001	0.369	0.059	0.001	0.000	0.343	0.046

See Table 1 for details.

TABLE 3
Empirical rejection frequencies

			$T = 500$				$T = 1000$				$T = 2000$				$T = 4000$		
		This paper	ZW	AH	LT	*This paper*	ZW	AH	LT	*This paper*	ZW	AH	LT	*This paper*	ZW	AH	LT
Gaussian case																	
$(\varphi, \sigma_b) = (0.95, 0.0)$	$\theta = -0.051$	0.000	0.031	0.000	0.056	0.000	0.095	0.000	0.059	0.000	0.239	0.000	0.051	0.000	0.461	0.000	0.051
$(\varphi, \sigma_b) = (0.98, 0.0)$	$\theta = -0.020$	0.012	0.997	0.000	0.058	0.000	1.000	0.000	0.054	0.000	1.000	0.000	0.054	0.000	1.000	0.000	0.051
$(\varphi, \sigma_b) = (0.99, 0.0)$	$\theta = -0.011$	0.126	1.000	0.399	0.038	0.027	1.000	0.540	0.038	0.000	1.000	0.726	0.058	0.000	1.000	0.910	0.051
$(\varphi, \sigma_b) = (0.99, 0.05)$	$\theta = -0.010$	0.028	0.018	0.401	0.043	0.001	0.021	0.713	0.055	0.000	0.027	0.777	0.056	0.000	0.015	0.923	0.051
$(\varphi, \sigma_b) = (1.01, 0.01)$	$\theta = 0.005$	0.828	1.000	0.812	0.966	0.928	1.000	0.959	0.978	0.971	1.000	0.997	1.000	0.994	1.000	1.000	1.000
$(\varphi, \sigma_b) = (1.01, 0.0)$	$\theta = 0.010$	0.960	1.000	0.902	0.998	0.997	1.000	0.992	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\gamma = 1.5, \sigma_e = 0.1$																	
$(\varphi, \sigma_b) = (0.95, 0.0)$	$\theta = -0.051$	0.008	0.199	0.003	0.037	0.000	0.349	0.002	0.040	0.003	0.504	0.013	0.058	0.000	0.762	0.035	0.046
$(\varphi, \sigma_b) = (0.98, 0.0)$	$\theta = -0.020$	0.043	0.990	0.166	0.044	0.027	1.000	0.186	0.041	0.016	1.000	0.173	0.059	0.001	1.000	0.200	0.046
$(\varphi, \sigma_b) = (0.99, 0.0)$	$\theta = -0.011$	0.126	1.000	0.344	0.052	0.080	1.000	0.405	0.037	0.040	1.000	0.416	0.058	0.022	1.000	0.424	0.046
$(\varphi, \sigma_b) = (0.99, 0.05)$	$\theta = -0.032$	0.004	0.001	0.385	0.044	0.001	0.002	0.412	0.042	0.001	0.000	0.442	0.058	0.000	0.000	0.445	0.046
$(\varphi, \sigma_b) = (1.01, 0.01)$	$\theta = -0.055$	0.116	0.128	0.812	0.062	0.031	0.142	0.959	0.062	0.007	0.118	0.997	0.059	0.000	0.152	1.000	0.046
$(\varphi, \sigma_b) = (1.01, 0.0)$	$\theta = 0.010$	0.977	0.993	0.598	0.994	0.997	1.000	0.726	1.000	0.999	1.000	0.824	1.000	1.000	1.000	0.906	1.000
$\gamma = 1.2, \sigma_e = 0.1$																	
$(\varphi, \sigma_b) = (0.95, 0.0)$	$\theta = -0.051$	0.018	0.315	0.015	0.041	0.007	0.483	0.041	0.053	0.002	0.648	0.096	0.060	0.006	0.825	0.152	0.049
$(\varphi, \sigma_b) = (0.98, 0.0)$	$\theta = -0.020$	0.080	0.992	0.219	0.044	0.034	1.000	0.197	0.054	0.020	1.000	0.210	0.061	0.012	1.000	0.267	0.049
$(\varphi, \sigma_b) = (0.99, 0.0)$	$\theta = -0.011$	0.169	0.998	0.353	0.048	0.138	1.000	0.397	0.051	0.090	1.000	0.388	0.061	0.047	1.000	0.421	0.049
$(\varphi, \sigma_b) = (0.99, 0.05)$	$\theta = -0.090$	0.037	0.001	0.315	0.046	0.026	0.000	0.283	0.038	0.020	0.000	0.295	0.061	0.028	0.000	0.282	0.049
$(\varphi, \sigma_b) = (1.01, 0.01)$	$\theta = -0.293$	0.020	0.000	0.000	0.055	0.011	0.000	0.000	0.057	0.011	0.000	0.000	0.061	0.007	0.000	0.000	0.049
$(\varphi, \sigma_b) = (1.01, 0.0)$	$\theta = 0.010$	0.957	0.994	0.584	0.991	0.997	1.000	0.704	1.000	1.000	1.000	0.761	1.000	1.000	1.000	0.853	1.000

See Table 1 for details.

References

Akharif, A. and M. Hallin (2003). Efficient detection of random coefficients in autoregressive models. *The Annals of Statistics 31*(2), 675–704.

Anděl, J. (1976). Autoregressive series with random parameters. *Mathematische Operationsforschung und Statistik 7*(5), 735–741.

Aue, A. and L. Horváth (2011). Quasi-likelihood estimation in stationary and nonstationary autoregressive models with random coefficients. *Statistica Sinica*, 973–999.

Aue, A., L. Horváth, and J. Steinebach (2006). Estimation in random coefficient autoregressive models. *Journal of Time Series Analysis 27*(1), 61–76.

Banerjee, A. N., G. Chevillon, and M. Kratz (2013). Detecting and forecasting large deviations and bubbles in a near-explosive random coefficient model.

Berkes, I., L. Horváth, and S. Ling (2009). Estimation in nonstationary random coefficient autoregressive models. *Journal of Time Series Analysis 30*(4), 395–416.

Clauset, A., C. R. Shalizi, and M. E. Newman (2009). Power-law distributions in empirical data. *SIAM review 51*(4), 661–703.

Distaso, W. (2008). Testing for unit root processes in random coefficient autoregressive models. *Journal of Econometrics 142*(1), 581–609.

Fryz, M. (2017). Conditional linear random process and random coefficient autoregressive model for eeg analysis. In *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, pp. 305–309. IEEE.

Giraitis, L., G. Kapetanios, and T. Yates (2014). Inference on stochastic time-varying coefficient models. *Journal of Econometrics 179*(1), 46–65.

Granger, C. W. and N. R. Swanson (1997). An introduction to stochastic unit-root processes. *Journal of Econometrics 80*(1), 35–62.

Hill, J., D. Li, and L. Peng (2016). Uniform interval estimation for an AR(1) process with AR errors. *Statistica Sinica 26*(1), 119–136.

Hill, J. and L. Peng (2014). Unified interval estimation for random coefficient autoregressive models. *Journal of Time Series Analysis 35*(3), 282–297.

Horváth, L. and L. Trapani (2016). Statistical inference in a random coefficient panel model. *Journal of Econometrics 193*(1), 54–75.

Horváth, L. and L. Trapani (2019). Testing for randomness in a random coefficient autoregression. *Journal of Econometrics 209*, 338–352.

Kapetanios, G., Y. Shin, and A. Snell (2003). Testing for a unit root in the nonlinear STAR framework. *Journal of Econometrics 112*(2), 359–379.

Koul, H. L. and A. Schick (1996). Adaptive estimation in a random coefficient autoregressive model. *The Annals of Statistics 24*(3), 1025–1052.

Leybourne, S. J., B. P. McCabe, and A. R. Tremayne (1996). Can economic time series be differenced to stationarity? *Journal of Business & Economic Statistics 14*(4), 435–446.

McCabe, B. P. and A. R. Tremayne (1995). Testing a time series for difference stationarity. *The Annals of Statistics*, 1015–1028.

Nagakura, D. (2009). Asymptotic theory for explosive random coefficient autoregressive models and inconsistency of a unit root test against a stochastic unit root process. *Statistics & Probability Letters 79*(24), 2476–2483.

Nicholls, D. F. and B. G. Quinn (2012). *Random Coefficient Autoregressive Models: An Introduction: An Introduction*, Volume 11. Springer Science & Business Media.

Phillips, P. C. B. (1990). Time series regression with a unit root and infinite-variance errors. *Econometric Theory 6*(01), 44–62.

Regis, M., P. Serra, and E. R. Heuvel (2021). Random autoregressive models: A structured overview. *Econometric Reviews*. forthcoming.

Ślezak, J., K. Burnecki, and R. Metzler (2019). Random coefficient autoregressive processes describe brownian yet non-gaussian diffusion in heterogeneous systems. *New Journal of Physics 21*(7), 073056.

Stenseth, N. C., W. Falck, K.-S. Chan, O. N. Bjørnstad, M. O?Donoghue, H. Tong, R. Boonstra, S. Boutin, C. J. Krebs, and N. G. Yoccoz (1998). From patterns to processes: phase and density dependencies in the canadian lynx cycle. *Proceedings of the National Academy of Sciences 95*(26), 15430–15435.

Trapani, L. (2021). Testing for strict stationarity in a random coefficient autoregressive model. *Econometric Reviews*. forthcoming.

Tsay, R. S. (1987). Conditional heteroscedastic time series models. *Journal of the American Statistical Association 82*(398), 590–604.

Tsay, R. S. (1997). Unit root tests with threshold innovations. *University of Chicago*.

Tsay, R. S. (2005). *Analysis of financial time series*, Volume 543. John wiley & sons.

Zhao, Z.-W. and D.-H. Wang (2012). Statistical inference for generalized random coefficient autoregressive model. *Mathematical and Computer Modelling 56*(7), 152–166.