Testing for unit roots in bounded time series*

Giuseppe Cavaliere
University of Bologna

Fang Xu
European University Institute
Christian-Albrechts-University of Kiel

May 2011

Abstract

Many key economic and financial series are bounded either by construction or through to policy controls. Conventional unit root tests are potentially unreliable in the presence of bounds, since they tend to over-reject the null hypothesis of a unit root, even asymptotically. So far, very little work has been undertaken to develop unit root tests which can be applied to bounded time series. In this paper we address this gap in the literature by proposing unit root tests which are valid in the presence of bounds. We present new augmented Dickey-Fuller type tests as well as new versions of the modified 'M' tests developed by Ng and Perron (2001, Econometrica 69, pp. 1519-1554) and demonstrate how these tests, combined with a simulation-based method to retrieve the relevant critical values, make it possible to control size asymptotically. A Monte Carlo study suggests that the proposed tests perform well in finite samples. Moreover, the tests outperform the Phillips-Perron type tests originally proposed in Cavaliere (2005, Econometric Theory 21, 907-945). An illustrative application to U.S. interest rate data is provided.

1 Introduction

In his latest contribution Clive Granger (2010) suggests that the analysis of time series which, despite being non-stationary, are *bounded*, is a topic which deserves further attention. Specifically, he argues that the unsolved issue is how a concept such as I(1) can be extended to bounded processes.

According to Granger (2010), '*a limited process is one that has bounds either below (at zero, say) or above (full capacity) or both*'. Indeed, many important economic and financial series are bounded in this sense. Examples are expenditure and budget shares, unemployment rates, nominal interest rates, target zone exchange rates. Although limited time series cannot be integrated in the usual sense, see the discussion in Granger (2010), in many theoretical and applied studies they are modelled as pure I(1) processes.

Cavaliere (2005) is the only attempt to explain how the concept of I(1) can coexist with the constraints of a bounded process. He shows that in the presence of (one or two) bounds, the well known Phillips-Perron (1988) [PP], unit root test statistics are characterized by a quite different asymptotic behavior. In general, the limiting null distributions depend upon nuisance parameters related to the position of the bounds: the tighter the bounds, the more shifted to the left the distributions of the unit root statistics. As a consequence, unit root tests based on standard asymptotic critical values become over-sized. Only when the bounds are sufficiently far away, conventional unit root methods behave according to the standard asymptotic theory. Cavaliere (2005) also proposes a two-stage procedure where the nuisance parameters related to the position

*We thank seminar participants at the European University Institute, the London School of Economics, CREATES (Aarhus), and participants at the Conference in Honour of Clive Granger held in Nottingham, 21-22 May 2010, for useful comments. Cavaliere acknowledges financial supports from Italian PRIN 2007 grants. Xu acknowledges the financial support of Fritz Thyssen Stiftung (Az.10.08.1.088).

of the bounds are first estimated. These estimates are then employed to retrieve bound-robust (asymptotic) critical values which can be applied to the standard PP tests.

Although it allows to obtain asymptotically valid tests, the approach proposed in Cavaliere (2005) suffers of the well-known finite sample size problems affecting PP unit root tests as well as most of the tests based on sum-of-covariances estimators of the long-run variance. More robust approaches, such as tests based on spectral estimators of the long run variance (Ng and Perron, 1995, 2001) or the well-known augmented Said-Dickey-Fuller [ADF] tests, could be applied. Unfortunately, no theory for such tests is available for bounded time series.

In this paper we aim at filling this gap in the existing literature by proposing a new approach to unit root testing in bounded time series which leads to tests which are asymptotically valid and possess good finite sample properties. By focusing on the ADF tests as well as on the autocorrelation-robust 'M' unit root tests of Perron and Ng (1996), Stock (1999) and Ng and Perron (2001) – although the approach we outline can equally well be applied to any of the commonly used unit root statistics – we propose a numerical solution to the inference problem. Specifically, direct simulation methods – based on new consistent estimators of the nuisance parameters related to the bounds – are employed to obtain approximate p-values from the asymptotic null distributions of the standard unit root statistics. A variety of algorithms are also suggested to account for potential autocorrelation and heteroskedasticity in the error terms. In addition, we demonstrate that the simulation-based ADF and M tests possess good finite sample properties, outperforming the PP tests considered in Cavaliere (2005).

As for the test discussed in Cavaliere (2005), our tests can be applied to series which have either one bound (above, or below, such as for the much discussed case of nominal interest rates) or two bounds. Moreover, we allow the errors to be general linear processes driven by martingale difference innovations, hence allowing for conditional heteroskedasticity e.g. of the ARCH type.

The paper is organized as follows. The next section introduces bounded integrated processes and discusses the main assumptions. In Section 3 the asymptotic distributions of the ADF and M test statistics are derived and their dependence on nuisance parameters related to the position of the bounds is established. The simulation-based unit root tests that account for the presence of bounds are presented in Section 4. The finite sample properties are investigated in Section 5. A brief illustrative application to US interest rates is reported in Section 6. Section 7 concludes. All proofs are collected in the Appendix.

The following notation is used through out the paper. '$\lfloor\cdot\rfloor$' denotes the integer part of its argument; '$\xrightarrow{w}$' denotes weak convergence and '$\xrightarrow{p}$' convergence in probability, in each case as the sample size diverges to positive infinity; '$x := y$' ('$x =: y$') indicates that x is defined by y (y is defined by x); $\mathcal{D} := D[0,1]$ is the space of right continuous with left limit (càdlàg) processes on $[0,1]$, equipped with the Skorohod metric; $\|x\|$ denotes the standard Euclidean norm of a column vector x, and the norm of a matrix B is defined as $\|B\| = \sup_{\|x\|<1} \|Bx\|$.

2 Bounded unit root processes

This section introduces the reference class of bounded non-stationary processes. We consider processes that behave similarly to random walks but, at the same time, they are bounded either above or below, or both. Processes belonging to this class will be referred to as 'bounded I(1)' or 'bounded unit root' processes, BI(1) hereafter. Bounded I(1) processes are discussed in Cavaliere (2005) and Granger (2010).

In general, a bounded time series X_t, with (fixed) bounds at $\underline{b}, \overline{b}$ ($\underline{b} < \overline{b}$), is a stochastic process satisfying $X_t \in [\underline{b}, \overline{b}]$ almost surely for all t. This requires that, at each t, the increment ΔX_t lies within the interval $[\underline{b} - X_{t-1}, \overline{b} - X_{t-1}]$. Focusing on the case of a constant deterministic component, a simple and relatively general way to extend the notion of bounded processes to the unit root case is to assume that (see Cavaliere, 2005)

$$X_t = \theta + Y_t \tag{2.1}$$
$$Y_t = \alpha Y_{t-1} + u_t,\ \alpha = 1 \tag{2.2}$$

initialized at $Y_0 = O_p(1)$. The term u_t is further decomposed as follows:

$$u_t = \varepsilon_t + \underline{\xi}_t - \overline{\xi}_t\ , \tag{2.3}$$

where ε_t is a (weakly dependent) zero-mean unbounded process and $\underline{\xi}_t, \overline{\xi}_t$ are non-negative processes such that $\underline{\xi}_t > 0$ if and only if $Y_{t-1}+\varepsilon_t < \underline{b}-\theta$ and, similarly, $\overline{\xi}_t > 0$ if and only if $Y_{t-1}+\varepsilon_t > \overline{b}-\theta$. Since any truncated, censored or reflected random variable can be represented as in (2.3) for some unbounded ε_t (see the discussion in Cavaliere, 2005), this assumption is quite general.

A BI(1) process reverts because of the bounds only. It behaves as a unit root process when it is far away from the bounds. Conversely, in the neighborhood of the bounds it differs from standard I(1) processes because of the presence of the terms $\underline{\xi}_t$ and $\overline{\xi}_t$, which force X_t to lie between $\underline{b}$ and $\overline{b}$. In the stochastic control literature, see Harrison (1985), $\underline{\xi}_t$ and $\overline{\xi}_t$ are referred to as 'regulators', as they control the path of X_t by keeping it between $\underline{b}$ and $\overline{b}$.

Throughout the paper we assume that ε_t is a general linear process [LP] of the form

$$\varepsilon_t = C(L)\, v_t \tag{2.4}$$

where v_t is a martingale difference sequence [MDS] and $C(z) := \sum_{j=0}^{\infty} c_j z^j$. We make use of the following standard assumption on ε_t, see e.g. Chang and Park (2002, pp.433–4).

Assumption $\mathcal{A}$: $\mathcal{A}_1$. (a) $\{v_t, \mathcal{F}_t\}$ *is a MDS with respect to some filtration* $\mathcal{F}_t$*, such that* $E(v_t^2) = \sigma^2 < \infty$, (b) $T^{-1}\sum_{t=1}^{T} v_t^2 \xrightarrow{p} \sigma^2$*, and* (c) $E|v_t|^r < \infty$ *for some* $r > 4$; $\mathcal{A}_2$. *The lag polynomial satisfies* $C(z) \neq 0$ *for all* $|z| \leq 1$ *and* $\sum_{j=0}^{\infty} j^s |c_j| < \infty$ *for some* $s \geq 1$.

In addition, we consider two further conditions related to the bounds. The first is a technical condition needed to prevent $\{X_t\}$ from 'jumping' too much at the bounds. The second formalizes a relation between the positions of the bounds and the sample size.

Assumption $\mathcal{B}$: $\mathcal{B}_1$. $\sup_{t=1,\dots,T} E|\underline{\xi}_t|^r < \infty$ *and* $\sup_{t=1,\dots,T} E|\overline{\xi}_t|^r < \infty$*, with* r *given in* $\mathcal{A}_1$; $\mathcal{B}_2$. $(\underline{b}-\theta)/\left(\lambda T^{1/2}\right) = \underline{c} + o(1)$ *and* $(\overline{b}-\theta)/\left(\lambda T^{1/2}\right) = \overline{c} + o(1)$*, where* $\underline{c} \leq 0 \leq \overline{c}$*,* $\underline{c} \neq \overline{c}$*, and* $\lambda^2 := \sigma^2 C(1)^2$ *denotes the long-run variance of* ε_t.

Some remarks are due.

Remark 2.1. Under Assumption $\mathcal{A}$, $C(z)^{-1} =: \alpha(z) = 1 - \sum_{j=1}^{\infty} \alpha_j z^j$ is well defined. By letting $\underline{\xi}_t^* := \alpha(L)\underline{\xi}_t$ and $\overline{\xi}_t^* := \alpha(L)\overline{\xi}_t$ we can write

$$u_t = C(L)v_t + \underline{\xi}_t - \overline{\xi}_t = C(L)v_t^*,\ v_t^* := v_t + \underline{\xi}_t^* - \overline{\xi}_t^*\ . \tag{2.5}$$

The differenced process ΔX_t therefore admits the LP representation $\Delta X_t = C(L)v_t^*$. Different from the standard I(1) case, v_t^* depends both on the innovations ε_t and the (current and past) regulators, $\underline{\xi}_t$ and $\overline{\xi}_t$.

Remark 2.2. As is standard, via the Beveridge-Nelson [BN] representation (cf. Phillips and Solo, 1992) ε_t can be written as $\varepsilon_t = C(1)v_t + \tilde{\varepsilon}_{t-1} - \tilde{\varepsilon}_t$, with $\tilde{\varepsilon}_t = \sum_{j=0}^{\infty} \tilde{c}_j v_{t-j}$ $(\tilde{c}_j := \sum_{i=j+1}^{\infty} c_i)$ being well defined in the L^r sense. Consequently, X_t can be decomposed as

$$X_t = \theta + C(1)\sum_{i=1}^{t} v_i + \sum_{i=1}^{t}(\underline{\xi}_i - \overline{\xi}_i) + \tilde{\varepsilon}_0 - \tilde{\varepsilon}_t\ . \tag{2.6}$$

Eq. (2.6) implies that the non-stationary component of the BI(1) process can be decomposed into a standard random walk, $\sum_{i=1}^{t} v_i$, and the cumulated regulators, $\sum_{i=1}^{t}(\underline{\xi}_i - \overline{\xi}_i)$. Under Assumption $\mathcal{B}_2$, these two terms are of the same order ($\sim T^{1/2}$). Therefore, the behavior of X_t is influenced by the regulators not only at short horizons, but in the long run as well. Notice also that, due to the presence of the cumulated regulators, standard convergence tools for I(1) processes (such as FCLTs) are not expected to hold in general.

Remark 2.3. For the unit root tests which will be discussed later it is useful to notice that an autoregressive [AR] approximation can be given to u_t. Since (2.5) implies the AR(∞) representation $\alpha(L)u_t = v_t^*$, we may write

$$u_t = \sum\nolimits_{i=1}^{k} \alpha_i u_{t-i} + v_{t,k} \ , \ v_{t,k} := v_t^* + \sum\nolimits_{i=k+1}^{\infty} \alpha_i u_{t-i} \ .$$

As in Chang and Park (2002,p.434), the moment restrictions on $v_t, \overline{\xi}_t, \underline{\xi}_t$ and the decaying rate for the coefficients of $C(z)$ (cf. Assumptions $\mathcal{A}$ and $\mathcal{B}$) imply that the approximation error, i.e. $v_{t,k} - v_t$, satisfies $E|v_{t,k} - v_t|^r = o(k^{-rs})$, even in the presence of the regulators $\underline{\xi}_t, \overline{\xi}_t$.

Remark 2.4. Assumption $\mathcal{B}_2$ relates the position of the bounds $\underline{b}$ and $\overline{b}$ (relatively to the location parameter θ) to the sample size T. As noted in Granger (2010, p.4), '*the qualifying feature of an I(1) process is the strong relationship between now and the distant past, so that corr*$(X_t, X_{t-k}) = 1$ *for any k*'. Under Assumption $\mathcal{B}_2$, this fundamental property is not altered by the presence of the bounds. Additionally, it is a key condition: (i) for establishing the asymptotic behavior of any unit root test statistics in the presence of bounds, see Cavaliere (2005), and (ii) for constructing proper unit root tests that take account of the bounds without making any parametric assumption on the behavior of X_t near the bounds.

Remark 2.5. One-sided bounds can be treated as a special case by setting $\overline{c} = \infty$ (lower bound only) or $\underline{c} = -\infty$ (upper bound only). By construction, $\overline{c} = \infty$ ($\underline{c} = -\infty$) implies that the upper regulator $\overline{\xi}_t$ (the lower regulator $\underline{\xi}_t$) equals zero, almost surely, for all t.

Remark 2.6. Since under $\mathcal{B}_2$ the bound parameters $\underline{b}$ and $\overline{b}$ depends on T, X_t formally constitutes a triangular array of the type $\{X_{Tt} : t = 0, 1, ..., T;\ T = 0, 1, ...\}$. The double index notation is not essential and bounded integrated processes will be simply denoted as $\{X_t\}$. □

3 Unit root asymptotics for bounded processes

In this section we discuss how the presence of bounds affects the asymptotic null distribution of the well-known augmented (Said-)Dickey-Fuller tests and of the M unit root tests.

For a given sample $\{X_t\}_0^T$, the ADF statistics are based on the OLS regression

$$\hat{X}_t = \alpha \hat{X}_{t-1} + \sum_{i=1}^{k} \alpha_i \Delta \hat{X}_{t-i} + \varepsilon_{t,k} \tag{3.7}$$

and are defined as

$$\mathcal{ADF}_\alpha := \frac{T(\hat{\alpha} - 1)}{\hat{\alpha}(1)}, \ \mathcal{ADF}_t := \frac{\hat{\alpha} - 1}{s(\hat{\alpha})}$$

where $\hat{\alpha}(1) := 1 - \sum_{i=1}^{k} \hat{\alpha}_i$, with $\hat{\alpha}_i$ denoting the OLS estimator of α_i in (3.7) and $s(\hat{\alpha})$ the usual OLS standard error of $\hat{\alpha}$. Here $\hat{X}_t$ denotes the OLS residuals from the regression of X_t on a constant term. Alternatively $\hat{X}_t$ can be taken as the pseudo GLS de-meaned series, see Elliott, Rothenberg and Stock (1996).

The M statistics are defined as

$$\mathcal{MZ}_\alpha := \frac{T^{-1}\hat{X}_T^2 - T^{-1}\hat{X}_0^2 - s_{AR}^2(k)}{2T^{-2}\sum_{t=1}^{T}\hat{X}_{t-1}^2}, \ \ \mathcal{MSB} := \left(T^{-2}\sum_{t=1}^{T}\hat{X}_{t-1}^2 / s_{AR}^2(k)\right)^{1/2}$$

and $\mathcal{MZ}_t := \mathcal{MZ}_\alpha \times \mathcal{MSB}$,[1] where $s^2_{AR}(k)$ is an autoregressive estimator of the (non-normalized) spectral density at frequency zero of $\{u_t\}$. Specifically,

$$s^2_{AR}(k) := \hat{\sigma}^2 / \hat{\alpha}(1)^2 \tag{3.8}$$

where $\hat{\alpha}(1)$ is as defined above and $\hat{\sigma}^2$ is the OLS variance estimator from the ADF regression (3.7). For all tests the lag truncation parameter is required to satisfy the following assumption (Lewis and Reinsel, 1985).

Assumption $\mathcal{K}$. *As $T \to \infty$, $1/k + k^2/T \to 0$.*

It is well known that if Assumptions $\mathcal{A}$ and $\mathcal{B}$ hold on (2.1)-(2.2) and if X_t is unbounded (i.e., $\overline{b} = -\underline{b} = \infty$), the asymptotic (null) distributions of the $\mathcal{ADF}$ and M statistics are as follows (see Ng and Perron, 2001; Chang and Park, 2002):

$$\begin{aligned} \mathcal{ADF}_\alpha, \mathcal{MZ}_\alpha &\xrightarrow{w} \tfrac{1}{2}\left(F_B(1)^2 - F_B(0)^2 - 1\right)\left(\int_0^1 F_B(s)^2\, ds\right)^{-1} =: \zeta_1 \\ \mathcal{MSB} &\xrightarrow{w} \left(\int_0^1 F_B(s)^2\, ds\right)^{1/2} =: \zeta_2 \\ \mathcal{ADF}_t, \mathcal{MZ}_t &\xrightarrow{w} \tfrac{1}{2}\left(F_B(1)^2 - F_B(0)^2 - 1\right)\left(\int_0^1 F_B(s)^2\, ds\right)^{-1/2} =: \zeta_3 \end{aligned} \tag{3.9}$$

with $F_B := B - \int_0^1 B(r)\, dr$ and B a demeaned Brownian motion and a Brownian motion, respectively. In the case of the $\mathcal{ADF}$ and $\mathcal{MZ}$ tests, the unit root null is rejected for large negative values of the statistics, while a test based on $\mathcal{MSB}$ rejects for small values of the statistic.

In Theorem 1 below we now provide representations for the asymptotic null distributions of the test statistics considered in the presence of bounds. A key role in the asymptotic distributions of the statistics is played by the process $B_{\underline{c}}^{\overline{c}}$, a Brownian motion, *regulated* at $\underline{c}, \overline{c}$. The regulated Brownian motion [RBM] behaves like a standard BM except in the neighborhood of the bounds, where it is forced to revert; see Harrison (1985) or Cavaliere (2005) for a technical definition.

Theorem 1 *Let $\{X_t\}_0^T$ be generated as in (2.1) with $\alpha = 1$, under Assumptions $\mathcal{A}$ and $\mathcal{B}$. Then: (i) $T^{-1/2} Y_{\lfloor T\cdot \rfloor} := T^{-1/2} \sum_{t=1}^{\lfloor T\cdot \rfloor} u_t \xrightarrow{w} \lambda B_{\underline{c}}^{\overline{c}}(\cdot)$ in $\mathcal{D}$; (ii) if Assumption $\mathcal{K}$ also holds, $s^2_{AR}(k) \xrightarrow{p} \lambda^2 := \sigma^2 C(1)^2$, $\mathcal{ADF}_\alpha, \mathcal{MZ}_\alpha \xrightarrow{w} 0.5(F_{B_{\underline{c}}^{\overline{c}}}(1)^2 - F_{B_{\underline{c}}^{\overline{c}}}(0)^2 - 1)(\int_0^1 F_{B_{\underline{c}}^{\overline{c}}}(s)^2\, ds)^{-1} =: \zeta_1^{\underline{c},\overline{c}}$, $\mathcal{MSB} \xrightarrow{w} (\int_0^1 F_{B_{\underline{c}}^{\overline{c}}}(s)^2\, ds)^{1/2} =: \zeta_2^{\underline{c},\overline{c}}$, and $\mathcal{ADF}_t, \mathcal{MZ}_t \xrightarrow{w} 0.5(F_{B_{\underline{c}}^{\overline{c}}}(1)^2 - F_{B_{\underline{c}}^{\overline{c}}}(0)^2 - 1)(\int_0^1 F_{B_{\underline{c}}^{\overline{c}}}(s)^2\, ds)^{-1/2} =: \zeta_3^{\underline{c},\overline{c}}$, where $F_{B_{\underline{c}}^{\overline{c}}} := B_{\underline{c}}^{\overline{c}} - \int_0^1 B_{\underline{c}}^{\overline{c}}(s) ds$.*

The following remarks collect some of the implications of Theorem 1.

Remark 3.1. The results in Theorem 1 differ from standard I(1) asymptotics mainly because the limiting process is not a standard Brownian motion, but a *regulated* Brownian motion. The sample paths of the limiting process are therefore bounded between $\underline{c}$ and $\overline{c}$, with the well known case of no bounds following as a special case by setting $-\underline{c}$ and $\overline{c}$ equal to infinity.

Remark 3.2. Under Assumption $\mathcal{B}$, the usual spectral estimator of the long run variance is still consistent for λ^2, the long-run variance of ε_t. However, this result alone does not suffice for the unit root statistics to have the usual Dickey-Fuller type distributions. Specifically, the asymptotic distributions $\zeta_i^{\underline{c},\overline{c}}$ $i = 1, 2, 3$ are non-standard and depend on the nuisance parameters $\underline{c}, \overline{c}$. Therefore, inference based on the usual quantiles is generally invalid. Only for bounds sufficiently far away the quantiles of the distributions in Theorem 1 are well approximated by the quantiles of ζ_i, $i = 1, 2, 3$. Conversely, unit root tests based on standard critical values are

[1] As in Müller and Elliott (2003), we include the term $-T^{-1}\hat{X}_0^2$ in the numerator of $\mathcal{MZ}_\alpha$ and $\mathcal{MZ}_t$, so that the $\mathcal{ADF}_\alpha$ ($\mathcal{ADF}_t$) and the $\mathcal{MZ}_\alpha$ ($\mathcal{MZ}_t$) statistics have the same limiting distributions.

over-sized, with the degree of oversizing depending on the two parameters $\underline{c}, \overline{c}$ (the narrower the limits, the higher the degree of oversizing).

Remark 3.4. When pseudo-GLS de-meaning is used and X_t is unbounded, the results in (3.9) hold with F_B replaced by B. Similarly, in the bounded case Theorem 1 can be generalized to tests based on pseudo-GLS de-trending. In the case of GLS de-meaned data, results for the ADF and M tests are as those given in Theorem 1 but with $B_{\underline{c}}^{\overline{c}}$ replacing $F_{B_{\underline{c}}^{\overline{c}}}$.

Remark 3.5. The results given in Theorem 1 can be readily extended to the near-integrated case, $\alpha := 1 - \kappa/T$, $0 < \kappa < \infty$ in (2.1)-(2.2). It is straightforward to demonstrate that Theorem 1 continues to hold but with $B_{\underline{c}}^{\overline{c}}$ replaced by the Ornstein-Uhlenbeck [OU] process, $J^{\kappa}(s) := \int_0^s \exp(-\kappa(s-r))\, dB(r)$, regulated at $\underline{c}, \overline{c}$ (see also Cavaliere, 2005, Theorem 4). Consequently, the asymptotic local power function of the unit root tests will also be affected by the bounds. □

4 Testing for unit roots in the presence of bounds

As discussed in the previous section, standard unit root inference is affected by the presence of bounds, as the null asymptotic distributions of the commonly employed test statistics are non-standard, leading to over-sized tests. Hence, in the presence of bounds, when the null hypothesis is rejected on the basis of standard critical values it is not possible to assess whether the rejection depends on the absence of a unit root or, conversely, on the presence of the bounds only.

Despite the fact that the asymptotic distributions of the unit root test statistics depend on the two nuisance parameters $\underline{c}$ and $\overline{c}$, see Theorem 1, in this section we are able to propose a simulation-based approach which allows to obtain proper asymptotic p-values for unit root tests when the time series of interest is bounded. Both one-sided and two-sided bounds are covered.

The derivation of unit root tests for bounded time series takes two steps. First (Section 4.1), we construct two simple, consistent estimators of the nuisance parameters $\underline{c}$ and $\overline{c}$. Second (Sections 4.2 and 4.3), we define a simulation-based approach which draws on such estimators and which can be used to retrieve the relevant p-values. Extensions to cases of (unconditionally) heteroskedastic shocks are discussed in Section 4.4.

4.1 Consistent estimation of the bound parameters

Since the bounds $\underline{b}, \overline{b}$ are assumed to be known, consistent estimation of the nuisance parameters $\underline{c}, \overline{c}$ is actually feasible. To this aim, it suffices to define the estimators $\widehat{\underline{c}}$ and $\widehat{\overline{c}}$ as follows

$$\widehat{\underline{c}} := \frac{\underline{b} - X_0}{s_{AR}(k)\, T^{1/2}}, \widehat{\overline{c}} := \frac{\overline{b} - X_0}{s_{AR}(k)\, T^{1/2}} \tag{4.10}$$

where $s_{AR}^2(k)$ is the spectral AR estimator of the long run variance as defined in Section 2. The main result on the consistency of $\widehat{\underline{c}}$ and $\widehat{\overline{c}}$ is given in the next lemma, which generalizes Corollary 5 of Cavaliere (2005) to the present framework.

Lemma 1 *Let the assumptions of Theorem 1 hold. Then,* $\widehat{\underline{c}} \xrightarrow{p} \underline{c}$, $\widehat{\overline{c}} \xrightarrow{p} \overline{c}$.

Hence, given that the bounds $(\underline{b}, \overline{b})$ are known, the two nuisance parameters of the limiting distributions in Theorem 1, $\underline{c}$ and $\overline{c}$, can be consistently estimated via $\widehat{\underline{c}}$ and $\widehat{\overline{c}}$, respectively. These estimators are the two key ingredients for our simulation-based tests.

Remark 4.1. Notice that in (4.10) the deterministic term θ is implicitly estimated under the null, as advocated in e.g. Schmidt and Phillips (1992). If θ is estimated by standard OLS, i.e. by replacing X_0 of (4.10) by $T^{-1}\sum_{t=1}^{T} X_t$, the resulting estimators of $\underline{c}, \overline{c}$ becomes inconsistent. □

4.2 Simulation-based tests

In this section we show how direct simulation methods can be used to retrieve p-values from the limiting null distributions of the standard ADF and M statistics given in Theorem 1.

As noted in Section 3, the limiting distributions depend on the regulated Brownian motion, $B_{\underline{c}}^{\overline{c}}$. Our method is based on the construction of a càdlàg process B_n^* that satisfies $B_n^* \xrightarrow{w} B_{\underline{c}}^{\overline{c}}$ with probability tending to one. We can then approximate quantiles from the non-pivotal limiting null distributions in Theorem 1 by simple numerical simulation methods based on approximating the limiting process $B_{\underline{c}}^{\overline{c}}$ through the càdlàg process B_n^*. The simulation-based versions of the ADF and M tests, which we denote generically as ADF^* and M^* in what follows, only require the computation of the standard ADF and M statistics of Section 3 and of the associated Monte Carlo [MC] p-values. Taking the ADF_α test to illustrate, the simulation-based test is constructed according to the following algorithm.

Algorithm 1 *Step* (i). *Let ε_t^* be an i.i.d.$(0,1)$ sequence (independent of $(X_0, ..., X_T)$);*
Step (ii). *For some $n \geq T$, let X_t^*, $t = 1, ..., n$ be recursively defined as*

$$X_t^* := \begin{cases} \widehat{\overline{c}} & \text{if } X_{t-1}^* + n^{-1/2}\varepsilon_t^* > \widehat{\overline{c}} \\ \widehat{\underline{c}} & \text{if } X_{t-1}^* + n^{-1/2}\varepsilon_t^* < \widehat{\underline{c}} \\ X_{t-1}^* + n^{-1/2}\varepsilon_t^* & \text{otherwise} \end{cases} \tag{4.11}$$

with initial condition $X_0 = 0$. The corresponding càdlàg process is $X_n^*(s) := X_{\lfloor ns \rfloor}^*$, $s \in [0,1]$.
Step (iii). *Compute the Monte Carlo statistics*

$$\mathcal{ADF}_\alpha^* := \frac{\tilde{X}_n^*(1)^2 - \tilde{X}_n^*(0)^2 - 1}{2\int_0^1 \tilde{X}_n^*(s)^2 ds}, \; \tilde{X}_n^*(s) := X_n^*(s) - \int_0^1 X_n^*(u)\, du$$

Step (iv). *Define the Monte Carlo p-value as $p_n^* := G_n^*(\mathcal{ADF}_\alpha)$, where G_n^* denotes the cumulative distribution function of $\mathcal{ADF}_\alpha^*$, conditional on $\widehat{\underline{c}}, \widehat{\overline{c}}$. Similarly, for any significance level η, cv_η that solves $G_n^*(cv_\eta) = \eta$ is the Monte Carlo critical value.*

Then, the following theorem holds as T diverges.

Theorem 2 *Let $\{X_t\}_0^T$ be generated as in* (2.1) *with $\alpha = 1$, under Assumptions $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{K}$. Then, as $T \to \infty$:* (i) *$X_n^* \xrightarrow{w} B_{\underline{c}}^{\overline{c}}$ in probability, and* (ii) *$\mathcal{ADF}_\alpha^* \xrightarrow{w} \zeta_1^{\underline{c},\overline{c}}$ in probability. Finally,* (iii) *$p_n^* \xrightarrow{w} U[0,1]$.*

Theorem 2(i)-(ii) shows that for T diverging to infinity, the simulated process X_n^* is distributed as the limiting process $B_{\underline{c}}^{\overline{c}}$ of Theorem 1(i) and that the MC statistic $\mathcal{ADF}_\alpha^*$ is asymptotically distributed as $\mathcal{ADF}_\alpha$ under the unit root null hypothesis. Consequently, see (iii), even if the unit root statistics are not pivotal in the presence of bounds, the $\mathcal{ADF}_\alpha^*$ test has correct (asymptotic) size. That is, for any chosen significance level η, as T diverges it holds that $P(p_n^* \leq \eta) \to \eta$ for any value of $\underline{c}, \overline{c}$, and a test which rejects the null hypothesis when $p_n^* \leq \eta$ has asymptotic size equal to η. Some remarks are due.

Remark 4.2. Algorithm 1 works under fairly general conditions, as it only requires two consistent estimators of $\underline{c}$ and $\overline{c}$. The term n, which can be interpreted as the discretization step used for approximating the limiting regulated Brownian motion, only needs to be bounded below by T. In principle, Algorithm 1 can be applied to any unit root test with null limiting distribution depending on the regulated Brownian motion $B_{\underline{c}}^{\overline{c}}$.

Remark 4.3. As is standard with simulation-based tests, the MC p-value p_n^* can be computed with any desired degree of accuracy by generating B (conditionally) independent samples $\{X_{T:b}^*\}$, $b = 1, ..., B$, and by computing $\mathcal{ADF}_{\alpha:b}^*$ as above on each sample. The simulated p-value is then computed as $\tilde{p}_n^* := B^{-1}\sum_{n=1}^{B}\mathbb{I}\left(\mathcal{ADF}_{\alpha:b}^* < \mathcal{ADF}_\alpha\right)$, and is such that $\tilde{p}_n^* \overset{a.s.}{\to} p_n^*$ as $B\to\infty$. An asymptotic standard error is given by $(\tilde{p}_n^*(1-\tilde{p}_n^*)/B)^{1/2}$; cf. Hansen (1996, p.419).

REMARK 4.4. In Theorem 2, any number of steps n used to construct the process X_n^* such that $n \geq T$ is admissible. Given that the simulation-based approach is used to retrieve p-values from the asymptotic distribution of the test statistics, it appears natural to consider a large number of steps. However, we found that setting $n = T$ generally provides better approximations to the finite sample distribution of the test statistic, see Section 5 below.

REMARK 4.5. The procedure outlined above can be applied in the one bound case as well. In the case of a single lower (upper) bound, it suffices to set $\widehat{\overline{c}} = +\infty$ $(\widehat{\underline{c}} = -\infty)$. □

4.3 Re-coloured simulation-based tests

Because the limiting distribution of $\mathcal{ADF}_\alpha$ does not depend on serial correlation nuisance parameters, the MC errors ε_t^* (see Step (i) of Algorithm 1) are uncorrelated and the MC statistic $\mathcal{ADF}_\alpha^*$ in Algorithm 1 does not require a correction for serial correlation. However, an improved finite sample approximation in the presence of serially correlated errors might be anticipated from replacing $\mathcal{ADF}_\alpha^*$ in step (iii) with the analogue, say $\mathcal{ADF}_\alpha^{**}$, of the original $\mathcal{ADF}_\alpha$ statistic, computed from the OLS regression

$$\hat{X}_t^* = \alpha\hat{X}_{t-1}^* + \sum_{i=1}^{k}\alpha_i\Delta\hat{X}_{t-i}^* + e_t^*$$

with $\hat{X}_t^*$ the de-meaned counterpart of X_t^*. The p-value in Step (iv) is then computed using the cdf of $\mathcal{ADF}_\alpha^{**}$, say $G_n^*()$. The results in Theorem 2 would be unaltered.

In a further attempt to improve finite sample performance in the case of correlated shocks, the basic MC approach outlined above can be extended to include a re-colouring (or sieve) component, without altering the large sample theory given in Theorem 2. As in Ferretti and Romo (1996), Chang and Park (2003) and Cavaliere and Taylor (2009), *inter alia*, this involves re-building stationary serial correlation into the MC innovations. This can be done by using the estimated stationary lag dynamics obtained from fitting the ADF regression

$$\hat{X}_t = \alpha\hat{X}_{t-1} + \sum_{i=1}^{k_{rc}}\alpha_i\Delta\hat{X}_{t-i} + \varepsilon_{t,k_{rc}}\ , \tag{4.12}$$

where k_{rc} is the lag truncation used for the purposes of re-colouring. Accordingly, with $\hat{\alpha}_{k_{rc}}(z) := 1-\sum_{i=1}^{k_{rc}}\hat{\alpha}_i z^i$, the recursion in Step (ii) of Algorithm 1 can be replaced by the re-coloured recursion

$$X_t^* := \begin{cases} \widehat{\overline{c}} & \text{if } X_{t-1}^* + n^{-1/2}u_{t,k_{rc}}^* > \widehat{\overline{c}} \\ \widehat{\underline{c}} & \text{if } X_{t-1}^* + n^{-1/2}u_{t,k_{rc}}^* < \widehat{\underline{c}} \\ X_{t-1}^* + n^{-1/2}u_{t,k_{rc}}^* & \text{otherwise} \end{cases} \tag{4.13}$$

where $u_{t,k_{rc}}^*$ is the re-coloured innovation process defined through the difference equation

$$\frac{\hat{\alpha}_{k_{rc}}(L)}{\hat{\alpha}_{k_{rc}}(1)}u_{t,k_{rc}}^* = \varepsilon_t^*,\ t = 1,...,T, \tag{4.14}$$

initialized at 0.[2] The scheme in (4.13)-(4.14) differs from that in Algorithm 1 in that the estimated AR lag polynomial, $\hat{\alpha}_{k_{rc}}(L)$, is incorporated into the algorithm to re-colour the MC innovations ε_t^*. Obviously, since the simulated errors are autocorrelated, the $\mathcal{ADF}_\alpha^{**}$ statistic should be considered in step (iii) of the algorithm.

[2] Notice that the $\hat{\alpha}_{k_{rc}}(1)^{-1}$ factor appearing on the left hand side of (4.14) ensures that $u_{t,k_{rc}}^*$ has unit long run variance. This normalization guarantees that $X_n^*(\cdot) := X_{\lfloor n\cdot\rfloor}^*$ converges weakly to $B_{\underline{c}}^{\overline{c}}(\cdot)$ in probability, as required.

REMARK 4.6. Notice that k_{rc} need not diverge to infinity with the sample size, nor it has to be equal to the truncation lag k used in the original ADF regression, since the re-colouring device is motivated from purely finite sample concerns. The results established in Theorem 2 also apply when re-coloured MC errors are used, provided $k_{rc} = o\left(T^{1/2}\right)$.

REMARK 4.7. In small samples there is the possibility that the estimated lag polynomial could have one or more explosive roots. We found that the performance of the algorithm was improved if any such root was shrunk to have modulus less than unity. In our experiments reported in Section 5 below we scaled such estimated roots to have modulus equal to .99. □

4.4 Extension to heteroskedastic shocks

Assumption $\mathcal{A}_1$ allows for cases where the innovation process $\{v_t\}$ in (2.4) is a (second order) stationary martingale difference sequence. This assumption therefore allows for certain forms of conditional heteroskedasticity. Unconditional heteroskedasticity, as considered by Cavaliere and Taylor (2007, 2008), can alter the large sample results given in this paper. Precisely, consider the decomposition $v_t := \sigma_t z_t$ with z_t an i.i.d.$(0, 1)$ process (with bounded fourth order moment) and σ_t satisfying $\sigma_t := \omega(t/T) > 0$ for all $t = 1, ..., T$, where $\omega(\cdot) \in \mathcal{D}$ is deterministic, Cavaliere and Taylor (2007) show that $T^{-1/2}Y_{\lfloor T\cdot\rfloor} := T^{-1/2}\sum_{t=1}^{\lfloor T\cdot\rfloor} u_t = T^{-1/2}C(1)\sum_{t=1}^{\lfloor T\cdot\rfloor} \varepsilon_t + o_p(1) \xrightarrow{w} \lambda_\omega M(\cdot)$ where, for $\bar{\omega} := (\int_0^1 \omega^2)^{1/2}$, $\lambda_\omega^2 := \bar{\omega}^2 C(1)^2$ and M is the continuous time Martingale $M(\cdot) := \bar{\omega}^{-1}\int_0^{\cdot} \omega dB$ (B being a standard Brownian motion). It can be shown that, in this case, Theorem 1 is no longer appropriate. Rather, the limiting distribution of $T^{-1/2}Y_{\lfloor T\cdot\rfloor}$ and of the unit root statistics are as given in Theorem 1 but with the regulated Brownian motion $B_{\underline{c}}^{\overline{c}}$ replaced by a Martingale process, regulated at $\underline{c}$ and $\overline{c}$, say $M_{\underline{c}}^{\overline{c}}$.[3] Consequently, the simulation-based tests earlier proposed are no longer valid, as they do not allow to replicate the time-varying behaviour of the unconditional variance of the shocks.

Nevertheless, a simple way of accounting for (possible) unconditional heteroskedasticity can be achieved by using a wild-bootstrap type construction of the simulated innovations, ε_t^*. Specifically, instead of generating ε_t^* as an i.i.d. process, we can set $\varepsilon_t^* := \hat{\varepsilon}_{t,k_{rc}} z_t$, where z_t is an i.i.d. $N(0, 1)$ sequence (independent of the original sample) and $\hat{\varepsilon}_{t,k_{rc}}$ are the residuals from the ADF regression (4.12); see Cavaliere and Taylor (2008, 2009). Given the preceding results, it can reasonably be conjectured that the large sample results of Theorem 2 remain valid even under unconditional heteroskedasticity of the type described here. Although a full asymptotic analysis of this case is beyond the scope of this paper, extensive simulation results support this view.

It is worth noting, however, that the type of heteroskedasticity which can be allowed using the wild bootstrap approach does not cover cases where the volatility of the innovations is related to the levels of the process. For instance, EMS target zone exchange rates tend to be more volatile as the exchange rate approaches the bound; conversely, for nominal interest rates volatility is positively related to the levels. Unfortunately, most studies in unit root and co-integration seem to neglect this possible relation between levels and volatility and – to our knowledge – no asymptotic theory is available for these processes in the non-stationary case (for a small class of level-dependent heteroskedastic, but stationary, processes, see e.g. Ling, 2002). Nevertheless, in a Monte Carlo study, Rodrigues and Rubia (2005) show that level-dependent heteroskedasticity does not seem to affect the size of the unit root tests. In the bounded case, a number of Monte Carlo experiments (available from the authors upon request) show that the size of our tests are also only marginally affected by level-dependent heteroskedasticity.

[3] This process can be constructed as a regulated Brownian motion, see Harrison (1985) or Cavaliere (2005), but with the standard Brownian motion replaced by the M.

5 Finite sample simulations

In this section we use Monte Carlo simulation methods to analyze the finite sample size of the bound-corrected unit root tests of Section 4 for a variety of bounded integrated processes.

Data are generated as in (2.1)-(2.2) for $T = 100, 500$ under the unit root hypothesis $\alpha = 1$, where we set $Y_0 = \theta = 0$ without loss of generality. Both the case of two (symmetric) bounds ($\bar{c} = -\underline{c} =: c > 0$), and a single lower bound ($\bar{c} = \infty$, $-\underline{c} =: c > 0$) are considered. All experiments are conducted using $2,000$ replications and using the `rndKMn` function of Gauss 9.0.[4]

Following Cavaliere (2005) and Ng and Perron (2001), the errors v_t in (2.4) are generated as i.i.d. $N(0,1)$. The (conditional) distribution of $u_t = \Delta X_t$ is then obtained by reflecting[5] the distribution of $\varepsilon_t := C(L)\, v_t$ at $\underline{b} - X_{t-1}$ and $\bar{b} - X_{t-1}$. Results for different truncation mechanisms do not alter the result reported in this section.

In Section 5.1, to analyze the effects of the presence of the bounds uncontaminated by serial dependence, we set $C(L) = 1$ in (2.4) and, correspondingly, $k = 0$ in (3.8). The analysis is then extended in Section 5.2 to allow for weak dependence in ε_t. In this case, the number of lags in the spectral AR estimator of the long run variance (3.8) is chosen according to the MAIC lag length selection criterion of Ng and Perron (2001) with $k \leq \lfloor 12(T/100)^{0.25} \rfloor$.

Three different version of the simulation-based unit root tests are employed. In the first version we consider the ADF^* and M^* tests constructed according to Algorithm 1, where we set the discretization step n to $20,000$.[6] The second version differs from the first since we set $n = T$, so that the simulated test statistic reflects exactly the corresponding length of the original sample. Although the two variants are asymptotically equivalent, we aim at assessing whether using a lower discretization step improves the finite sample size of the test. Finally, in the autocorrelated case we also consider the effect of adding the re-coloured device described in Section 4 to the algorithm. For all variants, the MC errors ε_t^* are $N(0,1)$. Moreover, each test rejects the null hypothesis when the corresponding simulated p-value is below the nominal asymptotic 5% level. p-values are computed as in Remark 4.3 with $B = 499$.

For space constraints, only results for the OLS de-trended statistics are reported. Results for the pseudo-GLS de-meaned statistics do not substantially differ.

5.1 Uncorrelated errors

Table 1 reports the (empirical) size results, for $\alpha = 1$ in (2.1)-(2.2), of the simulation-based ADF^* and M^* tests of Section 4.2. The tests based on $n = 20,000$ are denoted by 'a' while tests based on $n = T$ are denoted by 'b'. In order to evaluate the impact of (neglected) bounds on the size of standard unit root tests, we also report the (empirical) size for the standard ADF and M tests.

[*Table 1 about here*]

The upper panel of Table 1 reports the size of the various tests in the two-bound case (the case of no bounds corresponds to the '∞' entry), while single bound case is reported in the lower panel.

Consider first the standard ADF and M tests, where the bounds are neglected. In the presence of bounds, they are generally over-sized, relative to the benchmark case of no bounds ($c = \infty$). For instance, the empirical sizes of the $\mathcal{ADF}_\alpha$ and $\mathcal{MZ}_\alpha$ tests, which are quite accurate for $c = \infty$ and $T = 500$, increase to 30% and 29% when there are two bounds with $c = 0.4$ and $T = 500$. The $\mathcal{ADF}_t$ and $\mathcal{MZ}_t$ tests are also over-sized in the presence of two bounds, with empirical size

[4] The Gauss procedure for computing the simulation-based p-values is available from the authors upon request.

[5] Specifically, we set $\underline{\xi}_t := 2\left(\underline{b} - (X_{t-1} + u_t)\right) \mathbb{I}\left(X_{t-1} + u_t < \underline{b}\right)$ and $\bar{\xi}_t := 2\left((X_{t-1} + u_t) - \bar{b}\right) \mathbb{I}\left(X_{t-1} + u_t > \bar{b}\right)$.

[6] Since the case $n = 20,000$ is computationally burdensome, we implemented this algorithm as follows. We applied Algorithm 1 with $n = 20,000$, $B = 50,000$ and setting $\widehat{\bar{c}} = -\widehat{\underline{c}} = c$ ($-\widehat{\underline{c}} = c$ in the single bound case), with c taking values on the grid $0.01, 0.02, ..., 10$. For each c, the corresponding critical values were stored. Simulation-based critical values were then retrieved through a linear interpolation of the stored critical values.

around 20% in the two-bound case with $c = 0.4$ and $T = 500$. The $\mathcal{MSB}$ test appears to be the worst affected (its size is up to 34%). Significant oversizing can also be observed in the one-bound case, where in the case of a process starting at the lower bound ($c = 0$), most tests have size between 18% and 21%. These results show that standard unit root tests can be unreliable in the presence of bounds.

Turning to the simulation-based tests, the results in Table 1 show that the size accuracy of the tests is extremely good. In the two-bound case, the size of the ADF^* and M^* tests based on $n = 20,000$ (columns 'a' in the table) is as accurate as the size of the standard ADF and M test in the unbounded case, for all the values of c considered. For instance, in the unbounded case, the size of the $\mathcal{ADF}_t^*$ and the $\mathcal{ADF}_t$ tests is only slightly above 5% for both $T = 100$ and $T = 500$. Conversely, for $c = 0.4$ the size of $\mathcal{ADF}_t^*$ is still about 5% (specifically, 5.4% for $T = 100$ and 4.7% for $T = 500$), whereas the standard $\mathcal{ADF}_t$ test has size above 20% for both $T = 100$ and $T = 500$. Some of the simulation-based tests, in particular the M^* tests, appear to be marginally undersized, especially in the one-bound case; however, the overall performance is largely satisfactory, even when the bounds are tight.

Improved size properties are obtained by considering the simulation-based tests with discretization step $n = T$ (columns 'b' in the table). Almost all the ADF^* and M^* tests have size very close to 5%. Even the M^* tests, which appear to be slightly undersized for $n = 20,000$, display size close to 5% when $n = T$. The superiority of the ADF^* and M^* tests with $n = T$ can be observed both in the two-bound and in the one-bound cases. Finally, it is worth noting that the tests based on $n = T$ outperform the asymptotic tests discussed in Cavaliere (2005).

5.2 Autocorrelated errors

The size properties of the ADF^* and M^* tests are now examined for ε_t following a linear process. Two cases are considered. First, ε_t is a stationary AR(1) process, i.e. $\varepsilon_t = \phi\varepsilon_{t-1} + \nu_t$, so that $C(L) = \left(1 + \phi L + \phi^2 L^2 + ...\right)$ in (2.4). The term ν_t is i.i.d. $N\left(0, (1-\phi)^2\right)$ and the AR parameter $\phi \in \{-0.5, 0.5\}$. Second, ε_t is $MA(1)$, i.e. $\varepsilon_t = \theta\nu_{t-1} + \nu_t$, so that $C(L) = 1 + \theta L$ in (2.4). In this case ν_t is i.i.d. $N(0, 1/(1+\theta^2))$ with $\theta \in \{-0.5; 0.5\}$. In all cases the long-run variance of $\{\varepsilon_t\}$ is unity.

Together with the ADF^* and M^* tests of Section 4.2 with $n = 20,000$ and $n = T$, we also report the simulation-based tests employing the re-colouring device of Section 4.3. For the latter tests, the re-colouring lag truncation parameter is $k_{rc} = k$.[7]

[*Tables 2 & 3 about here*]

Results for the two-bound case are reported in Table 2. With AR(1) errors, the ADF^* and M^* tests based on $n = 20,000$ (columns 'a' in the table) are slightly undersized both for positive autocorrelation and for negative autocorrelation, particularly when the bounds are tight. The $\mathcal{ADF}_\alpha^*$ and $\mathcal{ADF}_t^*$ tests seem to be the most reliable, in particular for $T = 100$. It is worth recalling, however, that even when there are no bounds, the standard ADF and M tests with automatic data-dependent lag selection rules also tend to be slightly conservative in the presence of autocorrelated errors.

Setting $n = T$ in Algorithm 1 (columns 'b' in the table) improves the finite sample size for almost all tests, in particular for $T = 100$. For instance, when the bounds are very close ($c = 0.4$), $\mathcal{MSB}_t^*$ has size 0.3% for $T = 100$ and $\phi = -0.5$, while the size increases to 2.5% when the discretization step is $n = T$. This effect characterizes almost all tests considered.

Massive size improvements are obtained when the re-colouring device is added to Algorithm 1 (columns 'c' in the table). Now, even for moderate sample sizes, the size properties of the tests

[7] As in Cavaliere and Taylor (2009), in unreported simulations we also set $k_{rc} = 4$. The corresponding sizes were quite close to those obtained for $k_{rc} = k$; however, tests based on $k_{rc} = k$ are generally preferable.

are largely satisfactory, with almost all tests having size close to 5%, even for small values of c. Taking again the case of $\mathcal{MSB}_t^*$ to illustrate, for $T = 100$, $\phi = -0.5$ and $c = 0.4$, the size grows from 0.3% to 4% when the proposed re-colouring device is used.

The results for the case of MA errors are comparable to those obtained for AR errors. The proposed simulation-based tests perform particularly well, in particular when the re-colouring device is used. There is evidence of oversizing for $\phi = -0.5$, but only for $c = 0.4$ and $T = 100$. Also in this case the re-colouring device improves the size of all tests.

The results for one single (lower) bound, reported in Table 3, confirm those obtained for the case of two bounds. The proposed simulation-based tests seem to control size properly even in the presence of autocorrelated disturbances. For some parameter configurations tests without re-colouring device tend to be undersized. However, tests based on re-colouring appear to be largely satisfactory for all the error processes considered and for all values of the bound parameter c.

6 Empirical illustration

In this section we illustrate the methods discussed in this paper with a short application to nominal interest rate dynamics. Well-known examples are e.g. econometric models for the term structure (see Campbell and Shiller, 1987, for an early reference), tests of the so-called Fisher hypothesis (Rose, 1988), joint tests of PPP and UIP (Johansen and Juselius, 1992). In this framework, unit root (and cointegration) tests on nominal interest rates have been extensively applied. Despite the existence of a vast literature on the time series properties of nominal interest rates, it is somewhat surprising that most of the papers do not emphasize that nominal interest rates are non-negative and hence cannot be I(1) in the usual sense[8]. Focusing on tests for a unit root, in this section we shed some light on this issue by explicitly taking the lower bound at zero into account.

We consider monthly data of 3-month U.S. Treasury Bill (T-bill) rate from January 1957 to September 2008 ($T = 621$), see Figure 1. Data are obtained from the International Financial Statistics CD-Rom (2008) of the International Monetary Fund.

[*Figure 1 and Table 4 about here*]

In the left panel of Table 4, standard ADF and M tests are reported along with the corresponding (standard) asymptotic p-values. For all tests, the lag truncation parameter k, selected according to the MAIC criterion of Ng and Perron (2001) with $k \leq \lfloor 12(T/100)^{0.25} \rfloor$, equals 16; the corresponding estimate of the long run variance is $s_{AR}^2(k) = 0.63$. All statistics were computed on OLS de-meaned data. Results for pseudo GLS de-meaned data do not differ and are omitted for brevity.

Without exceptions, all standard tests strongly reject the unit root hypothesis. All p-values are below 1%, with the strongest rejection obtained from $\mathcal{MZ}_\alpha$ and $\mathcal{MSB}$ tests (the corresponding p-values are about 0.3%). Although this result seems to point against the presence of a unit root in nominal interest rates data, it may actually be affected by the fact that standard unit root tests are not reliable when applied to time series bounded below. This is confirmed by looking at the estimate of the bound parameter $\underline{c}$. Using $\widehat{\underline{c}}$ of Section 4.1, we obtain $\widehat{\underline{c}} = -0.16$. According to the simulation results in Table 1,[9] the size of a nominal 5% asymptotic test based on $\mathcal{ADF}_\alpha$ or $\mathcal{MZ}_\alpha$ is not less than 18% when the bound parameter $\underline{c}$ is -0.20 (or above) and $T = 500$. Similarly, the $\mathcal{ADF}_t$, $\mathcal{MZ}_t$ and $\mathcal{MSB}$ tests are also over-sized, with size exceeding 14%, 13% and 19%, respectively.[10] Hence, because of the lower bound at zero, on the basis of standard unit

[8] An exception is Nicolau (2002).

[9] Although the results in Table 1 refer to the case of white noise errors, they still provide a clear indication of the tendency of standard unit root tests to be seriously oversized when there is a lower bound at -0.16.

[10] This is in agreement with the findings in Cavaliere (2005, figure 5), where for $\underline{c} \in [-0.20, 0]$ the asymptotic size of most unit root tests at the nominal 5% (asymptotic) level is about 20%.

root tests it is not possible to assess whether the rejection of the unit root hypothesis is due to the presence of the bound (i.e., the DGP is a bounded unit root process) or whether the rejection should be taken as evidence of no unit roots (i.e., the DGP is a bounded process with no unit roots).

Test of the unit root hypothesis where the effects of the bounds are properly taken into account can be performed using the simulation-based approach of Section 4. In the right panel of Table 4 we report the simulation-based p-values for the ADF^* and M^* tests of Section 4.2 with $n = 20,000$ (column 'a') and $n = T$ (column 'b'). Moreover, we also report the re-coloured simulation-based p-values of Section 4.3 (columns 'c'), where we set $k_{rc} = k$.

The proposed simulation-based tests reverse the conclusion of standard unit root tests. It can immediately be noticed that all p-values become much higher when the lower bound at zero is taken into account. This is clearly in favour of the maintenance of the unit root hypothesis. The tests based on the re-colouring device (the most reliable in finite samples, according to the Monte Carlo experiment in Section 5), have p-values in the range 0.08-0.12.

Several points can be made out of this analysis. First of all, the presence of bounds affects the outcome of standard unit root tests, as predicted by the asymptotic theory. This can be immediately seen from the inspection of the p-values obtained with and without taking account of the bound: when the bound is considered, all p-values increase remarkably. Second, standard unit root tests are not useful for understanding whether the rejection of the unit root null hypothesis should be attributed to the presence of the bound or to the absence of a unit root. In the special case considered in this section, standard unit root tests lead to the conclusion that the interest rate considered is not a unit root process. Conversely, when the bound is accounted for, this conclusion is reversed. Interestingly, this result is not at odd with the conclusions in Aït-Sahalia (1996), who suggests (using high frequency data) that the US interest rate is likely to behave as a unit root process most of the time, but it reverts toward its mean when it reaches low values.

7 Conclusions

When applied to bounded time series, conventional unit root tests have to be treated with care. This paper shows that the popular ADF unit root tests as well as the so-called M tests can be unreliable when applied to bounded time series. Specifically, the asymptotic distributions of the corresponding test statistics depend on nuisance parameters related to the position of the bounds; the null distributions are shifted to the left. As a consequence, the rejection of the unit root hypothesis based on standard p-values might be due to the fact that the time series of interest is actually bounded.

To rectify this problem, in this paper we discuss a new approach for computing p-values (and critical values) for unit root tests in time series which are bounded above, or below, or both. Our approach combines the standard ADF and M statistics with a simulation-based approach to constructing the relevant p-values. It allows to test statistically whether a bounded time series reverts because of the presence of the bounds alone or because it does not have a unit root. Numerical evidence suggests that our proposed simulation-based procedure works extremely well in finite samples, in particular when it is used in conjunction with a re-colouring device. Moreover, the new tests outperforms the Phillips-Perron type tests analyzed in Cavaliere (2005).

Although the class of processes considered here is rather general, some important features of bounded time series are still left aside for future research.

Throughout it has been assumed that the bounds are fixed. Hence, cases such as target zone exchange rates under realignments of the central parity are not covered. Nevertheless, our analysis can be generalized to cases of time-varying (known) bounds; see Cavaliere (2000) for linearly trending bounds and Carrion-i-Silvestre and Gadea (2010) for the case of exogenous changes in the bound location. Specifically, most of the results given in this paper continue to hold when the bounds are time varying and satisfy $(\underline{b}_t - \theta) / \left(\lambda T^{1/2}\right) = \underline{f}\,(t/T) + o\,(1)$ and $(\bar{b}_t - \theta) / \left(\lambda T^{1/2}\right) = \bar{f}\,(t/T) + o\,(1)$, with $\underline{f} \leq 0$ and $\bar{f} \geq 0$ general càdlàg functions on [0,1].

This paper deals with the (most likely) case of that the bounds are *known*. However, even for unknown bounds the framework developed in this paper can provide useful insights. Where it is known that the time series of interest is regulated, but levels at which regulation occurs are unknown, a reasonable range for the bounds can often be inferred from historical observations and/or from the relevant economic theory. Moreover, by using our approach one can determine a minimum range under which the ADF and M tests do not suffer from oversizing, see Herwartz and Xu (2008) for the analysis of current account imbalances. The minimum range is defined by 'break even' bounds which approximately equalize the p-value of the unit root test considering bounds and the p-value obtained ignoring bounds. Unreasonably large break-even bounds signifies that neglecting the bounds when testing for a unit root might be misleading. In addition, it is possible to construct a (conservative) test for the unit root hypothesis by taking the maximum of the simulation-based p-values over a grid of admissible bound locations.

It is worth emphasising that only a constant deterministic term is allowed in the paper and extension to more general deterministic components is not straightforward. For instance, the presence of a linear trend has strong implications for bounded variables, as the trend might imply that, as T increases, the series is absorbed at one of the bounds or, in the one-bound case, that it drifts away from the bound (with the latter becoming irrelevant). Given our main assumption that the location of the bounds is related to $T^{1/2}$, local linear trends of the form $\theta_t = \theta + \tau t$ with $\tau := \kappa T^{-1/2}$ (κ being a fixed constant) represent a reasonable solution to introduce linear trends in bounded time series. Alternatively, piecewise-constant deterministic terms can also be considered, see Carrion-i-Silvestre and Gadea (2010). Both extensions are not trivial and beyond the scope of the present paper.

Finally, given that all the results discussed here hold for univariate time series only, an important and necessary extension is to generalize the proposed simulation-based tests to the case of multiple time series and co-integration tests. Suggestions for this step – currently under investigation by the authors – are given in Granger (2010).

References

Aït-Sahalia, Y., 1996. Testing Continuous-Time Models of the Spot Interest Rate. *Review of Financial Studies* 9, 385-426.

Berk, K.N., 1974. Consistent Autoregressive Spectral Estimates. *Annals of Statistics* 2, 489-502.

Campbell, J.Y. and Shiller, R.J., 1987. Cointegration and Tests of Present Value Models. *Journal of Political Economy* 95, 1062-1088.

Carrion-i-Silvestre and Gadea, M.D., 2010. Bounds, Breaks and Unit Root Tests. Working paper, University of Barcelona.

Cavaliere, G., 2000. A Rescaled Range Statistics Approach to Unit Root Tests, Econometric Society World Congress 2000 Contributed Papers 0318.

Cavaliere, G., 2005. Limited Time Series With A Unit Root. *Econometric Theory* 21. 907-945.

Cavaliere, G. and Taylor, A.M.R, 2007. Testing for Unit Roots in Time Series Models with Non-Stationary Volatility. *Journal of Econometrics* 140, 919–947.

Cavaliere, G. and Taylor, A.M.R, 2008. Bootstrap Unit Root Tests for Time Series Models with Non-Stationary Volatility. *Econometric Theory* 24, 43–71.

Cavaliere, G. and Taylor, A. M. R., 2009. Heteroskedastic Time Series with a Unit Root. *Econometric Theory* 25, 1228-1276.

Chang, Y. and Park, J.Y., 2002. On the Asymptotics of ADF Tests for Unit Roots. *Econometric Reviews* 21, 431-447.

Chang, Y. and Park, J.Y., 2003. A Sieve Bootstrap for the Test of a Unit Root. *Journal of Time Series Analysis* 24, 379-400.

Dickey, D.A. and Fuller, W.A., 1979. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association* 74, 427-431.

Elliott, G., Rothenberg, T.J. and Stock, J.H., 1996. Efficient Tests for an Autoregressive Unit Root. *Econometrica* 64, 813-836.

Ferretti, N. and Romo, J., 1996. Bootstrap Tests for Unit Root AR(1) Models. *Biometrika* 84, 849-860.

Granger, C.W.J., 2010. Some Thoughts on the Development of Cointegration. *Journal of Econometrics* 158, 3-6.

Hansen, B. E., 1996. Inference When a Nuisance Parameter Is Not Identified under the Null Hypothesis. *Econometrica* 64, 413-430.

Hansen, B.E. (2000). Testing for Structural Change in Conditional Models. *Journal of Econometrics* 97, 93–115.

Harrison, J.M., 1985. *Brownian Motion and Stochastic Flow Systems.* John Wiley & Sons.

Herwatz H. and Xu, F. 2008. Reviewing the Sustainability/Stationarity of Current Account Imbalances with Tests for Bounded Integration, *Manchester School* 76, 267–278.

Johansen, S. and Juselius, K., 1992. Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and the UIP for UK. *Journal of Econometrics* 53, 211-244.

Lewis, R. and Reinsel, G.C., 1985. Prediction of Multivariate Time Series by Autoregressive Model Fitting. *Journal of Multivariate Analysis* 16, 393-411.

Müller, U.K. and Elliott, G., 2003. Tests for Unit Roots and the Initial Condition. *Econometrica* 71, 1269-1286.

Ng, S. and Perron, P., 1995. Unit Root Tests in ARMA Models with Data Dependent Methods for the Selection of the Truncation lag. *Journal of the America Statistical Association* 90, 268-281.

Ng, S. and Perron, P., 2001. LAG Length Selection and the Construction of Unit Root Tests with Good Size and Power. *Econometrica* 69, 1519-1554.

Nicolau, J., 2002. Stationary Processes that Look Like Random Walks – The Bounded Random Walk Process in Discrete and Continuous Time. *Econometric Theory* 18, 99–118.

Perron, P. and Ng, S., 1996. Useful Modifications to some Unit Root Tests with Dependent Errors and their Local Asymptotic Properties. *Review of Economic Studies* 63, 435-463.

Phillips, P.C.B, and Perron, P., 1988. Testing for a Unit Root in Time Series Regressions. *Biometrika* 75, 335-346.

Phillips, P.C.B, and Solo, V., 1992. Asymptotics for Linear Processes. *Annals of Statistics* 20, 971-1001.

Rodrigues, P.M.M. and Rubia, A., 2005. The Performance of Unit Root Tests under Level-Dependent Heteroskedasticity. *Economics Letters* 89, 262–268.

Rose, A. K., 1988. Is the Real Interest Rate Stable? *Journal of Finance* 43, 1095-1112.

Said, E.S. and Dickey, D.A., 1984. Testing for Unit Roots in ARMA(p,q) Models with Unknown p and q. *Biometrika* 71, 599–607.

Schmidt, P. and Phillips, P.C.B. 1992. LM Test for a Unit Root in the Presence of Deterministic Trends, *Oxford Bulletin of Economics and Statistics* 54, 257–287.

Stock, J. H., 1999. A Class of Tests for Integration and Cointegration, in R. Engle and H. White (eds.), *Cointegration, Causality and Forecasting: A Festschrift for Clive W.J. Granger*, ch. 6, Oxford University Press.

A Appendix

This appendix is organized as follows. Section A.1 introduces some preliminary asymptotic results. Section A.2 contains the proofs of Theorem 1 and of the related lemmas. Section A.3 reports the proofs of the simulation-based results of Theorem 2.

A.1 Preliminary Lemmata

Let $v_t^* := v_t + \underline{\xi}_t^* - \overline{\xi}_t^*$ and $w_t^* := \sum_{i=1}^t v_i^*$, with $\underline{\xi}_t^*$, $\overline{\xi}_t^*$ as defined in Section 2. Furthermore, let $r_t^* := \underline{\xi}_t^* - \overline{\xi}_t^*$. The following results hold as T diverges.

Lemma A.1 *Under the assumptions of Theorem 1,* $(\sigma^2 T)^{-1/2} w^*_{\lfloor T\cdot\rfloor} \xrightarrow{w} B_{\underline{c}}^{\overline{c}}(\cdot)$.

Proof. By the BN decomposition of Remark 2.2,

$$T^{-1/2}\sum_{t=1}^{\lfloor T\cdot\rfloor} u_t = C(1)T^{-1/2}\sum_{t=1}^{\lfloor T\cdot\rfloor} v_t^* + T^{-1/2}\tilde{u}_0 - T^{-1/2}\tilde{u}_{\lfloor T\cdot\rfloor}, \tag{A.1}$$

with $\tilde{u}_t = \sum_{j=0}^{\infty}\tilde{c}_j v^*_{t-j}$ ($\tilde{c}_j := \sum_{i=j+1}^{\infty} c_i$). Since, under Assumptions $\mathcal{A}$ and $\mathcal{B}$, $\sup_t|\tilde{u}_t| = o_p(T^{1/2})$, we have that $T^{-1/2}\sum_{t=1}^{\lfloor T\cdot\rfloor} u_t - C(1)T^{-1/2}\sum_{t=1}^{\lfloor T\cdot\rfloor} v_t^* \xrightarrow{p} 0$. From Theorem 1 in Cavaliere (2005) it holds that $(\lambda^2 T)^{-1/2}\sum_{t=1}^{\lfloor T\cdot\rfloor} u_t \xrightarrow{w} B_{\underline{c}}^{\overline{c}}(\cdot)$. Hence, $\lambda^{-1}C(1)T^{-1/2}\sum_{t=1}^{\lfloor T\cdot\rfloor} v_t^* \xrightarrow{w} B_{\underline{c}}^{\overline{c}}(\cdot)$. The proof is completed by noticing that $\lambda^{-1}C(1) = \sigma^{-1}$. □

Lemma A.2 *Under the assumptions of Theorem 1,* $T^{-1}\sum_{t=1}^{T} v_t^{*2} \xrightarrow{p} \sigma^2$.

Proof. First, notice that

$$\frac{1}{T}\sum_{t=1}^{T} v_t^{*2} = \frac{1}{T}\sum_{t=1}^{T}(v_t + r_t^*)^2 = \frac{1}{T}\sum_{t=1}^{T} v_t^2 + \frac{1}{T}\sum_{t=1}^{T}(r_t^{*2} + 2v_t r_t^*).$$

where $T^{-1}\sum_{t=1}^{T} v_t^2 \xrightarrow{p} \sigma^2$ by Assumption $\mathcal{A}_1(b)$ and $T^{-1}\sum_{t=1}^{T}(r_t^{*2} + 2v_t r_t^*) = o_p(1)$. To show the latter result it suffices to notice that

$$\left|\frac{1}{T}\sum_{t=1}^{T}(r_t^{*2} + 2v_t r_t^*)\right| \le \frac{1}{T}\sum_{t=1}^{T}\left|r_t^{*2} + 2v_t r_t^*\right| \le \frac{1}{T}\left\{\max|2v_t| + \max|r_t^*|\right\}\sum_{t=1}^{T} 2\left|r_t^*\right| = o_p(1),$$

as $\max|v_t|$, $\max|r_t^*|$ are of $o_p(T^{1/2})$, and $\sum_{t=1}^{T}|r_t^*| = O_p(T^{1/2})$ under Assumption $\mathcal{B}$. □

Lemma A.3 *Under the assumptions of Theorem 1,* $T^{-1/2}\sum_{t=1}^{T} r_t^* = O_p(1)$.

Proof. It follows as $T^{-1/2}\sum_{t=1}^{T} r_t^* = T^{-1/2}\sum_{t=1}^{T} v_t^* - T^{-1/2}\sum_{t=1}^{T} v_t = O_p(1)$ according to Lemma A.1 and a standard FCLT (Phillips and Solo, 1992). □

A.2 Proof of Theorem 1 and related Lemmas

PROOF OF THEOREM 1. Throughout, to simplify the proof it is assumed (without loss of generality) that $v_t = \underline{\xi}_t^* = \overline{\xi}_t^* = 0$ for all $t \leq 0$ and that no deterministics are included in the model and in the estimation. Furthermore, we sketch the proof for the $\mathcal{ADF}_t$ and $\mathcal{ADF}_\alpha$ statistics; results for the other statistics follow similarly.

Let $Z_{t,k} := (\Delta X_{t-1}, ..., \Delta X_{t-k})'$, $\beta := (\alpha_1, ..., \alpha_k)'$ and recall that the ADF regression is

$$X_t = \alpha X_{t-1} + \beta' Z_{t,k} + v_{t,k}$$

with $v_{t,k} = v_t^* + \sum_{i=k+1}^{\infty} \alpha_i \Delta X_{t-i} = v_t^* + \sum_{i=k+1}^{\infty} \alpha_i u_{t-i}$. The proof for $\mathcal{ADF}_t$ is as follows. First, under the null hypothesis $\alpha = 1$, using Lemma A.2 and Lemmas A.4, A.5 below, we can proceed as in Chang and Park (2002) to prove that

$$\mathcal{ADF}_t = \sigma^{-1} \left(T^{-2} \sum_{t=1}^{T} w_{t-1}^{*2} \right)^{-1/2} \left(T^{-1} \sum_{t=1}^{T} w_{t-1}^* v_t^* \right) + o_p(1)$$

with w_t^* and v_t^* as previously defined. Then, Lemma A.1, the continuous mapping theorem [CMT] and Lemma A.2 imply

$$T^{-1} \sum_{t=1}^{T} w_{t-1}^* v_t^* = \frac{1}{2T} w_T^{*2} - \frac{1}{2T} \sum_{t=1}^{T} v_t^{*2} \xrightarrow{w} \frac{\sigma^2}{2} \left(B_{\underline{c}}^{\overline{c}}(1)^2 - 1 \right)$$
$$T^{-2} \sum_{t=1}^{T} w_{t-1}^{*2} \xrightarrow{w} \sigma^2 \int_0^1 B_{\underline{c}}^{\overline{c}}(s)^2 ds$$

which completes the proof for $\mathcal{ADF}_t$. Similarly,

$$\mathcal{ADF}_\alpha = \frac{T(\hat{\alpha} - 1)}{\hat{\alpha}(1)} = \left(T^{-2} \sum_{t=1}^{T} w_{t-1}^{*2} \right)^{-1} \left(T^{-1} \sum_{t=1}^{T} w_{t-1}^* v_t^* \right) + o_p(1) \; .$$

Lemma A.5 then implies the consistency of $\hat{\alpha}(1)$, which completes the proof of the above equation.□

Lemma A.4 *Under the assumptions of Theorem 1, (a)* $T^{-1} \sum_{t=1}^{T} X_{t-1} v_{t,k} = C(1) T^{-1} \sum_{t=1}^{T} w_{t-1}^* v_t^* + o_p(1)$*;(b)*$T^{-2} \sum_{t=1}^{T} X_{t-1}^2 = C(1)^2 T^{-2} \sum_{t=1}^{T} {w_{t-1}^*}^2 + o_p(1)$*;(c)*$T^{-1} \sum_{t=1}^{T} v_{t,k}^2 = T^{-1} \sum_{t=1}^{T} {v_t^*}^2 + o_p(1)$*.*

PROOF. Part (a). We have that

$$\begin{aligned}
\sum_{t=1}^{T} X_{t-1} v_{t,k} &= \sum_{t=1}^{T} X_{t-1} v_t^* + \sum_{t=1}^{T} X_{t-1} (v_{t,k} - v_t^*) \\
&= \sum_{t=1}^{T} (C(1) w_{t-1}^* + \tilde{u}_0 - \tilde{u}_{t-1}) v_t^* + \sum_{t=1}^{T} (C(1) w_{t-1}^* + \tilde{u}_0 - \tilde{u}_{t-1})(v_{t,k} - v_t^*) \\
&= C(1) \sum_{t=1}^{T} w_{t-1}^* v_t^* + \sum_{t=1}^{T} \tilde{u}_0 v_t^* - \sum_{t=1}^{T} \tilde{u}_{t-1} v_t^* + \\
&\quad C(1) \sum_{t=1}^{T} w_{t-1}^* (v_{t,k} - v_t^*) + \sum_{t=1}^{T} \tilde{u}_0 (v_{t,k} - v_t^*) - \sum_{t=1}^{T} \tilde{u}_{t-1} (v_{t,k} - v_t^*) \\
&= C(1) \sum_{t=1}^{T} w_{t-1}^* v_t^* + R_1 + R_2 + R_3 + R_4 + R_5.
\end{aligned}$$

By showing $R_1 = O_p(T^{1/2})$ and $R_2 + R_3 + R_4 + R_5 = o_p(T)$, the statement (a) follows.

First, $R_1 = \sum_{t=1}^{T} \tilde{u}_0 v_t^* = \tilde{u}_0 (\sum_{t=1}^{T} v_t + \sum_{t=1}^{T} r_t^*) = O_p(T^{1/2})$ due to Lemma A.3 and a standard FCLT. Second,

$$\begin{aligned} R_2 &= \sum_{t=1}^{T} \tilde{u}_{t-1} v_t^* = \sum_{t=1}^{T} \sum_{j=0}^{\infty} \tilde{c}_j v_{t-1-j}^* v_t^* = \sum_{t=1}^{T} \sum_{j=0}^{\infty} \tilde{c}_j (v_{t-1-j} + r_{t-1-j}^*)(v_t + r_t^*) \\ &= \sum_{t=1}^{T} \sum_{j=0}^{\infty} \tilde{c}_j v_{t-1-j} v_t + \sum_{t=1}^{T} \sum_{j=0}^{\infty} \tilde{c}_j (v_{t-1-j} r_t^* + r_{t-1-j}^* v_t + r_{t-1-j}^* r_t^*) = o_p(T). \end{aligned}$$

Since $\sum_{j=0}^{\infty} \tilde{c}_j v_{t-1-j} v_t$ is a MDS with finite variance, it can be shown that $\sum_{t=1}^{T} \sum_{j=0}^{\infty} \tilde{c}_j v_{t-1-j} v_t = o_p(T)$. Also,

$$\left| \sum_{t=1}^{T} \sum_{j=0}^{\infty} \tilde{c}_j (v_{t-1-j} r_t^* + r_{t-1-j}^* v_t + r_{t-1-j}^* r_t^*) \right| \leq \sum_{t=1}^{T} \sum_{j=0}^{\infty} |\tilde{c}_j| \left| v_{t-1-j} r_t^* + r_{t-1-j}^* v_t + r_{t-1-j}^* r_t^* \right|$$

$$\leq \sum_{t=1}^{T} \sum_{j=0}^{\infty} |\tilde{c}_j| \, 3 \, |r_t^*| \left\{ \max |v_t| + \max |r_t^*| \right\} = \sum_{j=0}^{\infty} |\tilde{c}_j| \left\{ \max |v_t| + \max |r_t^*| \right\} \sum_{t=1}^{T} 3 \, |r_t^*| = o_p(T)$$

since $\max |v_t|$ and $\max |r_t^*|$ are of $o_p(T^{1/2})$, $\sum_{j=0}^{\infty} |\tilde{c}_j| < \infty$ and $\sum_{t=1}^{T} |r_t^*| = O_p(T^{1/2})$.

Third, defining $c_{k,j} = \alpha_j C(L)$, we have that

$$\begin{aligned} R_3 &= \sum_{t=1}^{T} w_{t-1}^* (v_{t,k} - v_t^*) = \sum_{t=1}^{T} \sum_{i=1}^{t-1} v_i^* \sum_{j=k+1}^{\infty} c_{k,j} v_{t-j}^* = \sum_{j=k+1}^{\infty} c_{k,j} \sum_{t=1}^{T} v_{t-j}^* \sum_{i=1}^{t-1} v_i^* \\ &= o_p(k^{-s}) O_p(T^{1/2}) O_p(T^{1/2}) = o_p(T k^{-s}), \end{aligned}$$

given Assumption $\mathcal{A}_2$ and Lemma A.1. Next,

$$\begin{aligned} R_4 &= \sum_{t=1}^{T} \tilde{u}_0 (v_{t,k} - v_t^*) = \sum_{t=1}^{T} \tilde{u}_0 \sum_{j=k+1}^{\infty} c_{k,j} v_{t-j}^* = \tilde{u}_0 \sum_{j=k+1}^{\infty} c_{k,j} \sum_{t=1}^{T} v_{t-j}^* \\ &= o_p(k^{-s}) O_p(T^{1/2}) = o_p(T^{1/2} k^{-s}) \\ R_5 &= \sum_{t=1}^{T} \tilde{u}_{t-1} (v_{t,k} - v_t^*) = \sum_{t=1}^{T} \left(\sum_{j=0}^{\infty} \tilde{c}_j v_{t-1-j}^* \right) \left(\sum_{j=K+1}^{\infty} c_{k,j} v_{t-j}^* \right) \\ &= \sum_{j=0}^{\infty} \sum_{j=k+1}^{\infty} \tilde{c}_j c_{k,j} \sum_{t=1}^{T} v_{t-1-j}^* v_{t-j}^*. \end{aligned}$$

where

$$\begin{aligned} \sum_{t=1}^{T} v_{t-1-j}^* v_{t-j}^* &= \sum_{t=1}^{T} (v_{t-1-j} + r_{t-1-j}^*)(v_{t-j} + r_{t-j}^*) \\ &= \sum_{t=1}^{T} (v_{t-1-j} v_{t-j}) + \sum_{t=1}^{T} (v_{t-1-j} r_{t-j}^* + r_{t-1-j}^* v_{t-j} + r_{t-1-j}^* r_{t-j}^*) = o_p(T) \end{aligned}$$

since $\sum_{t=1}^{T} v_{t-1-j} v_{t-j} = o_p(T)$ as $v_{t-1-j} v_{t-j}$ is a MDS with finite variance. Besides, $\sum_{t=1}^{T} (v_{t-1-j} r_{t-j}^* + r_{t-1-j}^* v_{t-j} + r_{t-1-j}^* r_{t-j}^*) = o_p(T)$ according to similar arguments as those for R_2. Furthermore,

$$\left| \sum_{j=0}^{\infty} \sum_{j=k+1}^{\infty} \tilde{c}_j c_{k,j} \right| \leq \sum_{j=0}^{\infty} \sum_{j=k+1}^{\infty} |\tilde{c}_j c_{k,j}| = \sum_{j=0}^{\infty} |\tilde{c}_j| \sum_{j=k+1}^{\infty} |c_{k,j}| = o_p(k^{-s}).$$

Therefore, $R_5 = o_p(Tk^{-s})$. The proof of statement (a) is then complete.
Part (b). According to the BN representation,

$$\begin{aligned}\frac{1}{T^2}\sum_{t=1}^{T} X_{t-1}^2 &= \frac{1}{T^2}\sum_{t=1}^{T}\left(C(1)w_{t-1}^* + \tilde{u}_0 - \tilde{u}_{t-1}\right)^2 \\ &= C(1)^2\frac{1}{T^2}\sum_{t=1}^{T} w_{t-1}^{*2} + \frac{1}{T^2}\sum_{t=1}^{T}\left(\tilde{u}_0^2 - \tilde{u}_{t-1}^2 - 2\tilde{u}_0\tilde{u}_{t-1} + 2C(1)w_{t-1}^*\tilde{u}_0 - 2C(1)w_{t-1}^*\tilde{u}_{t-1}\right) \\ &= C(1)^2\frac{1}{T^2}\sum_{t=1}^{T} w_{t-1}^{*2} + o_p(1),\end{aligned}$$

since $\tilde{u}_{t-1} = o_p(T^{1/2})$ and $w_{t-1}^* = O_p(T^{1/2})$ as shown in Lemma A.1.
Part (c). The result can be obtained by applying

$$\left|\left(\frac{1}{T}\sum_{t=1}^{T} v_{t,k}^2\right)^{1/2} - \left(\frac{1}{T}\sum_{t=1}^{T} v_t^{*2}\right)^{1/2}\right| \leq \left[\frac{1}{T}\sum_{t=1}^{T}\left(v_{t,k} - v_t^*\right)^2\right]^{1/2}.$$

The right hand side of the previous inequality satisfies

$$\begin{aligned}E\left[\frac{1}{T}\sum_{t=1}^{T}\left(v_{t,k} - v_t^*\right)^2\right] &= E\left[\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{j=k+1}^{\infty} c_{k,j}v_{t-j}^*\right)^2\right] = E\left[\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{j=k+1}^{\infty} c_{K,j}(v_{t-j} + r_{t-j}^*)\right)^2\right] \\ &= E\left[\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=k+1}^{\infty}\sum_{j=k+1}^{\infty} c_{k,j}c_{k,j}(v_{t-i}v_{t-j} + r_{t-i}^*r_{t-j}^* + 2v_{t-i}r_{t-j}^*)\right)\right] \\ &= \frac{1}{T}\sum_{t=1}^{T}\left(E(v_{t-i}v_{t-j}) + E(r_{t-i}^*r_{t-j}^*) + 2E(v_{t-i}r_{t-j}^*)\right)\sum_{i=k+1}^{\infty}\sum_{j=k+1}^{\infty} c_{k,i}c_{k,j}\end{aligned}$$

Note that $\sum_{i=k+1}^{\infty}\sum_{j=k+1}^{\infty} c_{k,i}c_{k,j} = \left(\sum_{j=k+1}^{\infty} c_{k,j}\right)^2 = o(k^{-2s})$. Furthermore, $E(v_{t-i}v_{t-j}) + E(r_{t-i}^*r_{t-j}^*) + 2E(v_{t-i}r_{t-j}^*) < \infty$ due to the stated moment conditions on ε_t^* and r_t^*. Therefore, $E\left[T^{-1}\sum_{t=1}^{T}\left(v_{t,k} - v_t^*\right)^2\right] = o(k^{-2s})$, and $T^{-1}\sum_{t=1}^{T}\left(v_{t,k} - v_t^*\right)^2 = o_p(k^{-2s})$. The statement (c) is then proved. □

Lemma A.5 *Under the assumptions of Theorem 1, as $T \to \infty$, (a) $||(T^{-1}\sum_{t=1}^{T} Z_{t,k}Z_{t,k}')^{-1}|| = O_p(1)$; (b) $||\sum_{t=1}^{T} Z_{t,k}X_{t-1}|| = O_p(Tk^{1/2})$; (c)$||T^{-1}\sum_{t=1}^{T} Z_{t,k}v_{t,k}|| = o_p(k^{-1/2})$.*

Proof. Part (a). Let $\gamma_i := E(\varepsilon_t\varepsilon_{t-i})$ be the autocovariance function of ε_t and $\Gamma_k = (\gamma_{i-j})_{i,j=1}^{k}$. Denote $T^{-1}\sum_{t=1}^{T} Z_{t,k}Z_{t,k}'$ as $\hat{\Gamma}_k$, we have

$$\left\|\hat{\Gamma}_k^{-1}\right\| \leq \left\|\hat{\Gamma}_k^{-1} - \Gamma_k^{-1}\right\| + \left\|\Gamma_k^{-1}\right\|.$$

Note $\left\|\Gamma_k^{-1}\right\|$ is uniformly bounded above by a positive constant F for all k (see e.g. equation (2.14) in Berk (1974)). As the next step, we show that $\left\|\hat{\Gamma}_k^{-1} - \Gamma_k^{-1}\right\| = o_p(1)$ under Assumption $\mathcal{K}$, which completes the statement. First, defining $\epsilon_{t,k} := (\varepsilon_{t-1}, ..., \varepsilon_{t-k})'$ and $R_{t,k} := (r_{t-1}, ..., r_{t-k})'$ with $r_t = \underline{\xi}_t - \overline{\xi}_t$, we have

$$\left\|\hat{\Gamma}_k - \Gamma_k\right\| \leq \left\|T^{-1}\sum_{t=1}^{T}\epsilon_{t,k}\epsilon_{t,k}' - \Gamma_k\right\| + \left\|T^{-1}\sum_{t=1}^{T}(\epsilon_{t,k}R_{t,k}' + R_{t,k}\epsilon_{t,k}' + R_{t,k}R_{t,k}')\right\| = o_p(1).$$

Since $E\left(\left\|T^{-1}\sum_{t=1}^{T}\epsilon_{t,k}\epsilon_{t,k}'-\Gamma_k\right\|\right)^2 \leq \text{constant}k^2/(T-k)$ as can be seen e.g. from equation (2.10) and (2.11) in Berk (1974), $\left\|T^{-1}\sum_{t=1}^{T}\epsilon_{t,k}\epsilon_{t,k}'-\Gamma_k\right\| = O_p(k/T^{1/2}) = o_p(1)$ under Assumption $\mathcal{K}$. Because $T^{-1}\sum_{t=1}^{T}\varepsilon_{t-i}r_{t-j}$ and $T^{-1}\sum_{t=1}^{T}r_{t-i}r_{t-j}$ are $O_p(T^{-1/2})$, $\left\|T^{-1}\sum_{t=1}^{T}(\epsilon_{t,k}R_{t,k}'+R_{t,k}\epsilon_{t,k}'+R_{t,k}R_{t,k}')\right\| = \sqrt{k^2(O_p(T^{-1/2})^2} = O_p(k/T^{1/2}) = o_p(1)$ under Assumption $\mathcal{K}$. Then, from

$$\left\|\hat{\Gamma}_k^{-1}-\Gamma_k^{-1}\right\| = \left\|\hat{\Gamma}_k^{-1}(\Gamma_k-\hat{\Gamma}_k)\Gamma_k^{-1}\right\| \leq \left\|\hat{\Gamma}_k^{-1}\right\|\left\|\Gamma_k-\hat{\Gamma}_k\right\|\left\|\Gamma_k^{-1}\right\| \leq \left(\left\|\hat{\Gamma}_k^{-1}-\Gamma_k^{-1}\right\|+F\right)\left\|\Gamma_k-\hat{\Gamma}_k\right\|F,$$

we have $M_{k,T} \leq \left\|\Gamma_k-\hat{\Gamma}_k\right\| = o_p(1)$ with $M_{k,T} := \left\|\hat{\Gamma}_k^{-1}-\Gamma_k^{-1}\right\|/\left(\left\|\hat{\Gamma}_k^{-1}-\Gamma_k^{-1}\right\|+F\right)F$. Thus, $\left\|\hat{\Gamma}_k^{-1}-\Gamma_k^{-1}\right\| = F^2M_{k,T}/(1-FM_{k,T}) = o_p(1)$, which completes the statement (a).

Part (b). To show this statement we can use exactly the same arguments for Lemma 3.2 (b) in Chang and Park (2002) and the fact that

$$\sum_{t=1}^{T}(u_{t-i}u_{t-j}-\gamma_{i-j}) = \sum_{t=1}^{T}\left(\varepsilon_{t-i}\varepsilon_{t-j}-\gamma_{i-j}\right)+\sum_{t=1}^{T}\left(\varepsilon_{t-i}r_{t-j}+r_{t-i}\varepsilon_{t-j}+r_{t-i}r_{t-j}\right) = O_p(T^{1/2}).$$

Part (c). Note that $\left\|\sum_{t=1}^{T}Z_{t,k}v_{t,k}\right\| \leq \left\|\sum_{t=1}^{T}Z_{t,k}(v_{t,k}-v_t^*)\right\|+\left\|\sum_{t=1}^{T}Z_{t,k}v_t^*\right\|$. For $q=1,...,k$,

$$\begin{aligned}
E\left|\sum_{t=1}^{T}u_{t-q}(v_{t,k}-v_t^*)\right|^2 &= \sum_{t=1}^{T}\sum_{r=1}^{T}E\left|u_{t-q}(v_{t,k}-v_t^*)u_{r-q}(u_{k,r}^*-v_r^*)\right| \\
&= \sum_{t=1}^{T}\sum_{r=1}^{T}E\left|\left(\sum_{i=0}^{\infty}c_iv_{t-q-i}^*\right)\left(\sum_{j=k+1}^{\infty}c_{k,j}v_{t-j}^*\right)\left(\sum_{m=0}^{\infty}c_mv_{r-q-m}^*\right)\left(\sum_{n=k+1}^{\infty}c_{k,n}v_{r-n}^*\right)\right| \\
&\leq \sum_{t=1}^{T}\sum_{r=1}^{T}\left[\sum_{i=0}^{\infty}\sum_{j=k+1}^{\infty}\sum_{m=0}^{\infty}\sum_{n=k+1}^{\infty}\left|c_ic_{k,j}c_mc_{k,n}\right|E\left|v_{t-q-i}^*v_{t-j}^*v_{r-q-m}^*v_{r-n}^*\right|\right] \\
&\leq \sup_t E\left|v_t^*\right|^4\sum_{t=1}^{T}\sum_{r=1}^{T}\left[\sum_{i=0}^{\infty}|c_i|\sum_{j=k+1}^{\infty}|c_{k,j}|\sum_{m=0}^{\infty}|c_m|\sum_{n=k+1}^{\infty}|c_{k,n}|\right] \\
&= T^2\left(\sum_{i=0}^{\infty}|c_i|\right)^2\left(\sum_{j=k+1}^{\infty}|c_{k,j}|\right)^2\sup E\left|v_t^*\right|^4 = O(T^2k^{-2s}),
\end{aligned}$$

as v_t and r_t^* have bounded fourth moments and $\sum_{j=k+1}^{\infty}|c_{k,j}| = o_p(k^{-s})$. Therefore,

$$\begin{aligned}
E\left\|\sum_{t=1}^{T}Z_{t,k}(v_{t,k}-v_t^*)\right\|^2 &= E\left(\sum_{t=1}^{T}u_{t-1}(v_{t,k}-v_t^*)\right)^2+\cdots+E\left(\sum_{t=1}^{T}u_{t-k}(v_{t,k}-v_t^*)\right)^2 \\
&= O(T^2k^{1-2s}),
\end{aligned}$$

and $\left\|\sum_{t=1}^{T}Z_{t,k}(v_{t,k}-v_t^*)\right\| = O_p(Tk^{1/2-s})$. Because $Tk^{1/2-s} = o\left(Tk^{-1/2}\right)$ due to Assumption $\mathcal{K}$, $\left\|\sum_{t=1}^{T}Z_{t,k}(v_{t,k}-v_t^*)\right\| = O_p(Tk^{-1/2})$.

Furthermore, we have $\left\|\sum_{t=1}^{T}Z_{t,k}v_t^*\right\| = \sqrt{\left(\sum_{t=1}^{T}u_{t-1}v_t^*\right)^2+...+\left(\sum_{t=1}^{T}u_{t-k}v_t^*\right)^2}$. It can be shown that

$$\sum_{t=1}^{T}u_{t-q}v_t^* = \sum_{t=1}^{T}\varepsilon_{t-q}v_t+\sum_{t=1}^{T}\varepsilon_{t-q}r_t^*+\sum_{t=1}^{T}r_{t-q}r_t^*+\sum_{t=1}^{T}r_{t-q}v_t = O_p(T^{1/2}),\ \text{for } q=1,...,k.$$

Since $E\left(\sum_{t=1}^{T}\varepsilon_{t-q}v_t\right)^2 = \sum_{t=1}^{T}E(\varepsilon_{t-q}^2v_t^2) \le \max\{E(\varepsilon_t^2)\}\sum_{t=1}^{T}E(v_t^2) = T\max\{E(\varepsilon_t^2)\}\sigma^2 = O(T)$, we have $\sum_{t=1}^{T}\varepsilon_{t-q}v_t = O_p(T^{1/2})$. Similarly, $\sum_{t=1}^{T}r_{t-q}v_t = O_p(T^{1/2})$ because $E\left(\sum_{t=1}^{T}r_{t-q}v_t\right)^2 = \sum_{t=1}^{T}E(r_{t-q}^2v_t^2) \le \max\{E(r_t^2)\}\sum_{t=1}^{T}E(v_t^2) = T\max\{E(r_t^2)\}\sigma^2 = O(T)$. Since $\sum_{t=1}^{T}\varepsilon_{t-q}r_t^*$ and $\sum_{t=1}^{T}r_{t-q}r_t^*$ are also $O_p(T^{1/2})$, we have that $\sum_{t=1}^{T}u_{t-q}v_t^* = O_p(T^{1/2})$ and $\left\|\sum_{t=1}^{T}Z_{t,k}v_t^*\right\| = O_p(k^{1/2}T^{1/2})$. Because $k^{1/2}T^{1/2}$ is $O(Tk^{-1/2})$ under Assumption $\mathcal{K}$, $\left\|\sum_{t=1}^{T}Z_{t,k}v_t^*\right\| = O_p(Tk^{-1/2})$ and the proof is complete.

A.3 Proof of Theorem 2 and related results

Proof of Lemma 1. It follows from the consistency property of $s_{AR}^2(k)$, which can be established using the consistency of $\hat{\sigma}$ and $\hat{\alpha}(1)$, see the proof of Theorem 1. □

Proof of Theorem 2. Part (i). The proof of Theorem 2(i) consists of two steps. First, we show how to construct a càdlàg process $\tilde{X}_n^*$, independent of the original data $(X_0, ..., X_T)$, such that $\tilde{X}_n^* \xrightarrow{w} B_{\underline{c}}^{\overline{c}}$. Second, we show that $\tilde{X}_n^*$ and X_n^* are 'close', in the sense that $\sup_{s\in[0,1]}|\tilde{X}_n^*(s) - X_n^*(s)| \xrightarrow{p} 0$. Taken together, these two results imply that $X_n^* \xrightarrow{w} B_{\underline{c}}^{\overline{c}}$ in probability, as required.

For the first part, in order to define $\tilde{X}_n^*$ it suffices to consider the following construction, for $t = 1, ..., n$:

$$\tilde{X}_t^* := \begin{cases} \overline{c} & \text{if } \tilde{X}_{t-1}^* + n^{-1/2}\varepsilon_t^* > \overline{c} \\ \underline{c} & \text{if } \tilde{X}_{t-1}^* + n^{-1/2}\varepsilon_t^* < \underline{c} \\ \tilde{X}_{t-1}^* + n^{-1/2}\varepsilon_t^* & \text{otherwise} \end{cases}$$

with initial condition $X_0 = 0$ and ε_t^* as in (4.11). By setting $\tilde{X}_n^*(s) := \tilde{X}_{\lfloor ns\rfloor}^*$, since $n \to \infty$ we can proceed as in the proof of Theorem 6 of Cavaliere (2005) to obtain that $\tilde{X}_n^* \xrightarrow{w} B_{\underline{c}}^{\overline{c}}$. This completes the first part.

To show that $\sup_{s\in[0,1]}|\tilde{X}_n^*(s) - X_n^*(s)| = \max_{t=0,1,..,n}|\tilde{X}_t^* - X_t^*| = o_p(1)$ we can make use of an inductive argument to prove that for all $t = 0, ..., n$, $|\tilde{X}_t^* - X_t^*| \le |\widehat{\underline{c}} - \underline{c}| + |\widehat{\overline{c}} - \overline{c}|$. By Lemma 1 and the normality assumption on ε_t^*, this implies that $\sup_{s\in[0,1]}|\tilde{X}_n^*(s) - X_n^*(s)| \le |\widehat{\underline{c}} - \underline{c}| + |\widehat{\overline{c}} - \overline{c}|$, as required. We consider the one-bound case only, i.e. we set $\overline{c}$, $\widehat{\overline{c}} = \infty$; the proof for the two-bound case is substantially identical. Furthermore, we let (without loss of generality) $\widehat{\underline{c}} > \underline{c}$. For $t = 0$,the relation is trivially satisfied. Now, suppose that the relation holds at time $t-1$, i.e. $|\tilde{X}_{t-1}^* - X_{t-1}^*| \le |\widehat{\underline{c}} - \underline{c}|$. To prove that the relation holds at time t as well it is useful to distinguish the following cases.

(a) $X_{t-1}^* + n^{-1/2}\varepsilon_t^* \ge \widehat{\underline{c}}$ and $\tilde{X}_{t-1}^* + n^{-1/2}\varepsilon_t^* \ge \underline{c}$. In this case we have that

$$|X_t^* - \tilde{X}_t^*| = |X_{t-1}^* + n^{-1/2}\varepsilon_t^* - (\tilde{X}_{t-1}^* + n^{-1/2}\varepsilon_t^*)| = |X_{t-1}^* - \tilde{X}_{t-1}^*| \le |\widehat{\underline{c}} - \underline{c}|,$$

as required.

(b) $X_{t-1}^* + n^{-1/2}\varepsilon_t^* \ge \widehat{\underline{c}}$ and $\tilde{X}_{t-1}^* + n^{-1/2}\varepsilon_t^* < \underline{c}$. This implies $\varepsilon_t^* < 0$, $X_{t-1}^* \ge \tilde{X}_{t-1}^*$, and $X_t^* \ge \tilde{X}_t^*$. Therefore,

$$\begin{aligned} X_t^* - \tilde{X}_t^* &= X_{t-1}^* + n^{-1/2}\varepsilon_t^* - \underline{c} = X_{t-1}^* - \tilde{X}_{t-1}^* + \left(\tilde{X}_{t-1}^* + n^{-1/2}\varepsilon_t^* - \underline{c}\right) \\ &\le X_{t-1}^* - \tilde{X}_{t-1}^* \le \widehat{\underline{c}} - \underline{c}. \end{aligned}$$

(c) $X_{t-1}^* + n^{-1/2}\varepsilon_t^* < \widehat{\underline{c}}$ and $\tilde{X}_{t-1}^* + n^{-1/2}\varepsilon_t^* \ge \underline{c}$. Both $X_t^* \ge \tilde{X}_t^*$ and $X_t^* < \tilde{X}_t^*$ are possible. In the former case, $X_t^* - \tilde{X}_t^* = \widehat{\underline{c}} - \tilde{X}_t^* \le \widehat{\underline{c}} - \underline{c}$. The latter case implies $X_{t-1}^* < \tilde{X}_{t-1}^*$, and thus,

$$\begin{aligned} \tilde{X}_t^* - X_t^* &= (\tilde{X}_{t-1}^* + n^{-1/2}\varepsilon_t^*) - \widehat{\underline{c}} = \tilde{X}_{t-1}^* - X_{t-1}^* + X_{t-1}^* + n^{-1/2}\varepsilon_t^* - \widehat{\underline{c}} \\ &\le \widehat{\underline{c}} - \underline{c} + X_{t-1}^* + n^{-1/2}\varepsilon_t^* - \widehat{\underline{c}} = X_{t-1}^* + n^{-1/2}\varepsilon_t^* - \underline{c} \le \widehat{\underline{c}} - \underline{c}. \end{aligned}$$

(d) $X^*_{t-1} + n^{-1/2}\varepsilon^*_t < \widehat{\underline{c}}$ and $\tilde{X}^*_{t-1} + n^{-1/2}\varepsilon^*_t < \underline{c}$. Then $X^*_t = \widehat{\underline{c}}$ and $\tilde{X}^*_t = \underline{c}$, which obviously implies that $|X^*_t - \tilde{X}^*_t| \leq |\widehat{\underline{c}} - \underline{c}|$.

Taken together (a)–(d) implies that $|X^*_t - \tilde{X}^*_t| \leq |\widehat{\underline{c}} - \underline{c}|$ at time t, hence completing the proof.

Part (ii). It follows from Part (i) using standard continuous mapping arguments.

Part (iii). It suffices to follow the proof of Theorem 5 in Hansen (2000). □

Table 1: Finite-sample null rejection probabilities white noise model, one-bound at $-c$ and two bounds at $-c$ and c cases

c	T	$\mathcal{ADF}_\alpha$	$\mathcal{ADF}^*_\alpha$		$\mathcal{ADF}_t$	$\mathcal{ADF}^*_t$		$\mathcal{MZ}_\alpha$	$\mathcal{MZ}^*_\alpha$		$\mathcal{MZ}_t$	$\mathcal{MZ}^*_t$		$\mathcal{MSB}$	$\mathcal{MSB}^*$	
			(a)	(b)		(a)	(b)		(a)	(b)		(a)	(b)		(a)	(b)
Two symmetric bounds, no lags																	
∞	100	4.2	4.2	5.3	5.3	5.3	5.0	3.0	3.0	5.3	3.0	3.0	4.9	3.0	3.0	5.2
	500	5.0	5.0	5.2	5.6	5.6	5.3	4.8	4.8	5.2	5.0	5.0	5.4	4.5	4.5	5.2
0.8	100	9.3	4.4	6.2	11.2	5.8	5.6	7.0	2.8	6.2	7.1	3.2	5.6	4.5	2.4	5.8
	500	11.6	4.9	5.5	11.3	5.2	5.4	11.1	4.5	5.5	10.4	4.7	5.4	7.8	4.1	5.4
0.6	100	15.7	4.3	6.5	14.5	5.9	6.1	12.5	2.6	6.5	9.2	2.7	5.9	11.4	2.4	7.0
	500	18.4	4.4	5.3	14.6	4.8	5.3	17.8	4.0	5.3	13.4	4.0	5.3	17.2	4.1	5.5
0.4	100	26.1	3.4	5.9	20.5	5.4	5.6	21.4	1.4	5.9	12.7	1.6	5.5	25.0	1.5	6.5
	500	30.2	4.3	5.4	20.4	4.7	5.3	29.1	3.8	5.4	18.9	3.8	5.3	33.8	3.6	5.5
One single bound, no lags																	
0.8	100	7.0	4.3	5.8	8.3	5.5	5.5	5.1	3.0	5.8	5.0	3.2	5.3	3.7	2.6	5.5
	500	8.2	4.9	5.4	8.3	5.3	5.4	7.8	4.5	5.4	7.6	4.8	5.4	6.1	4.2	5.4
0.6	100	10.1	4.3	6.1	9.9	5.7	5.6	7.9	2.8	6.1	6.3	2.9	5.5	7.4	2.5	6.4
	500	11.3	4.6	5.5	9.8	5.1	5.3	10.9	4.2	5.5	8.9	4.3	5.3	10.4	4.1	5.4
0.4	100	13.2	3.9	5.7	11.7	5.3	5.4	10.9	2.4	5.7	7.6	2.4	5.2	12.3	2.4	6.2
	500	15.0	4.1	5.1	12.0	4.5	5.1	14.6	3.8	5.1	11.1	3.9	5.1	16.1	3.8	5.1
0.2	100	15.8	2.9	5.3	14.1	4.5	5.4	13.5	1.5	5.3	9.3	1.6	5.3	14.9	1.7	5.5
	500	17.7	3.8	4.8	14.1	4.2	4.9	17.2	3.5	4.8	13.2	3.5	4.9	18.8	3.6	4.8
0	100	17.4	2.5	5.1	18.8	4.5	5.0	14.9	1.0	5.1	13.4	1.1	5.1	14.2	1.1	5.1
	500	20.0	3.7	5.1	19.1	4.2	5.1	19.4	3.4	5.1	18.0	3.3	5.1	18.5	3.3	5.0

Notes: (i) Nominal 5% asymptotic level. (ii) Tests based on OLS de-meaned data. (iii) Columns a and b denote simulation-based p-values with $n = 20,000$ (a), $n = T$ (b).

Table 2: Finite-sample null rejection probabilities AR(1) and MA(1) models; two-bound (at $-c$ and c) case.

c	T	$\mathcal{ADF}_\alpha^*$			$\mathcal{ADF}_t^*$			$\mathcal{MZ}_\alpha^*$			$\mathcal{MZ}_t^*$			$\mathcal{MSB}^*$		
		(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
AR(1) with $\phi = 0.5$																
∞	100	4.2	5.2	5.0	4.1	3.9	5.4	4.0	6.0	4.8	3.7	5.9	5.2	4.3	6.5	4.8
	500	4.8	4.8	4.9	4.9	4.7	5.2	4.7	5.0	4.9	4.8	4.9	4.9	4.5	4.9	4.8
0.8	100	2.7	4.0	4.9	4.9	4.9	6.7	2.0	4.5	4.9	4.0	6.7	6.7	2.3	5.0	4.2
	500	3.7	4.2	5.0	4.4	4.6	5.3	3.5	4.2	5.0	4.4	4.8	5.2	3.2	3.8	4.2
0.6	100	2.2	3.6	5.2	3.9	4.0	6.1	1.5	4.2	5.3	3.2	5.9	6.3	1.3	4.0	3.9
	500	3.2	3.7	5.0	3.5	3.7	5.0	2.9	3.8	4.9	3.5	4.1	5.1	2.6	3.6	4.8
0.4	100	1.5	2.6	4.6	2.1	2.4	4.7	1.0	3.3	4.7	1.5	3.7	4.9	0.6	2.6	3.9
	500	2.4	3.1	4.4	2.4	2.8	4.5	2.2	3.3	4.6	2.3	3.4	4.5	2.1	3.2	4.6
AR(1) with $\phi = -0.5$																
∞	100	3.9	5.0	5.0	3.7	3.7	5.2	1.5	3.2	4.5	1.8	3.4	4.2	1.6	3.2	4.3
	500	5.0	4.8	4.9	4.8	4.6	5.3	4.2	4.2	4.9	4.3	4.5	5.1	3.9	4.2	4.7
0.8	100	4.0	5.9	5.2	3.4	3.3	4.6	1.0	3.3	4.3	1.3	3.0	4.0	0.9	3.4	4.2
	500	4.8	5.6	5.6	4.7	4.7	5.6	3.8	4.6	5.6	3.9	4.5	5.4	3.5	4.5	5.6
0.6	100	3.7	5.8	5.3	2.8	2.8	4.7	0.7	2.7	4.1	0.6	2.2	3.7	0.6	3.2	4.5
	500	4.3	5.2	5.7	3.8	4.1	5.4	3.1	4.2	5.6	3.1	3.8	5.2	3.2	4.5	5.8
0.4	100	2.8	6.1	6.2	2.2	2.5	5.8	0.2	2.2	3.8	0.2	2.0	3.7	0.2	2.5	3.9
	500	3.6	4.4	5.6	3.1	3.5	5.5	2.4	3.5	5.5	2.3	3.4	5.3	2.3	3.6	5.7
MA(1) with $\theta = 0.5$																
∞	100	4.5	4.9	4.6	3.6	3.5	4.8	3.7	5.4	4.3	3.4	5.5	4.7	4.0	5.7	4.2
	500	5.1	5.1	5.0	4.7	4.4	5.1	4.9	5.2	4.9	5.0	5.2	5.0	4.5	5.1	4.9
0.8	100	3.0	4.4	4.9	4.0	4.1	5.6	1.9	4.5	4.6	3.3	5.9	5.8	1.9	4.3	3.5
	500	4.0	4.7	4.7	4.1	4.3	5.0	3.6	4.4	4.6	4.4	4.6	4.9	3.4	4.1	4.1
0.6	100	2.3	4.1	5.5	3.1	3.2	5.4	1.4	4.2	5.2	2.3	5.0	5.5	1.0	3.8	4.2
	500	3.4	4.0	4.8	3.3	3.4	4.8	3.1	3.9	4.8	3.3	4.2	4.9	2.9	3.9	4.7
0.4	100	1.3	2.6	4.8	1.6	1.8	4.2	0.8	3.0	4.8	0.9	3.0	4.4	0.6	3.2	4.9
	500	2.7	3.3	4.3	2.3	2.5	4.4	2.4	3.3	4.4	2.5	3.3	4.4	2.2	3.3	4.4
MA(1) with $\theta = -0.5$																
∞	100	8.1	9.5	7.6	5.5	5.3	6.3	3.7	5.7	5.9	3.1	5.0	5.4	3.8	5.8	6.1
	500	6.4	6.7	6.1	5.7	5.3	5.8	5.0	5.5	5.9	4.8	5.0	5.6	4.8	5.3	5.8
0.8	100	6.5	9.7	7.1	3.3	3.7	5.4	1.3	4.6	5.2	1.2	3.8	4.7	1.5	5.8	5.4
	500	6.3	7.3	6.4	4.7	4.9	5.6	4.0	5.1	5.9	3.8	4.5	5.6	4.4	5.6	5.9
0.6	100	5.5	9.0	7.6	3.1	3.5	5.6	1.0	4.5	5.4	1.0	4.0	5.1	1.0	5.1	5.5
	500	5.6	6.8	6.1	3.9	4.3	5.8	3.2	4.2	5.7	3.0	3.7	5.4	3.6	4.8	5.8
0.4	100	16.3	29.6	19.8	9.4	10.4	11.9	0.8	14.7	12.6	0.9	14.3	12.5	0.9	15.0	12.8
	500	6.0	7.8	7.6	3.5	4.2	7.2	2.6	4.4	6.9	2.4	4.2	6.7	2.6	4.6	7.0

Notes: (i) Nominal 5% asymptotic level. (ii) Tests based on OLS de-meaned data. (iii) The lag truncation parameter k is selected according to the MAIC criterion of Ng and Perron (2001) with $k \leq \lfloor 12(T/100)^{0.25} \rfloor$. Columns a, b and c denote simulation-based p-values with $n = 20,000$ (a), $n = T$ (b) and with re-colouring device (c).

Table 3: Finite-sample null rejection probabilities AR(1) and MA(1) models; one-bound (at $-c$) case.

c	T	$\mathcal{ADF}_\alpha^*$			$\mathcal{ADF}_t^*$			$\mathcal{MZ}_\alpha^*$			$\mathcal{MZ}_t^*$			$\mathcal{MSB}^*$		
		(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
								AR(1) with $\phi = 0.5$								
0.8	100	2.9	4.3	4.7	4.4	4.3	6.2	2.5	5.0	4.8	3.7	6.3	6.0	2.8	5.3	4.4
	500	4.0	4.4	4.8	4.5	4.5	5.4	3.8	4.5	4.8	4.4	4.9	5.2	3.5	4.1	4.4
0.6	100	2.6	3.7	5.0	3.5	3.8	5.9	2.1	4.5	5.0	3.1	5.6	5.7	1.8	4.5	4.2
	500	3.3	3.9	4.8	3.7	3.9	5.0	3.2	3.9	4.8	3.6	4.2	4.9	2.8	3.8	4.5
0.4	100	1.8	3.0	5.0	2.4	2.4	5.2	1.4	3.9	5.0	1.9	4.4	5.0	1.2	3.5	4.5
	500	2.6	3.4	4.6	2.7	3.1	4.9	2.6	3.5	4.6	2.8	3.8	4.8	2.4	3.3	4.6
0.2	100	0.8	1.8	4.0	1.0	1.3	4.4	0.6	2.5	4.2	0.7	2.7	4.3	0.7	2.5	4.4
	500	2.0	2.7	4.3	1.9	2.4	4.6	1.9	2.9	4.4	2.1	2.9	4.5	1.9	2.7	4.2
0	100	0.4	1.3	3.9	0.5	0.7	3.5	0.4	1.8	4.1	0.4	1.8	3.6	0.4	2.1	4.4
	500	1.7	2.3	3.9	1.5	1.7	4.0	1.7	2.5	4.0	1.6	2.5	3.7	1.8	2.6	4.1
								AR(1) with $\phi = -0.5$								
0.8	100	4.1	5.4	5.2	3.6	3.4	5.0	1.1	3.2	4.5	1.6	3.4	4.2	1.2	3.2	4.4
	500	4.7	5.2	5.3	4.5	4.7	5.5	3.8	4.6	5.4	4.1	4.7	5.4	3.7	4.7	5.4
0.6	100	3.8	5.4	5.3	3.1	3.2	5.0	1.1	2.9	4.3	1.3	3.0	4.1	1.1	3.3	4.6
	500	4.2	4.9	5.5	4.1	4.3	5.3	3.4	4.2	5.4	3.5	4.2	5.2	3.4	4.4	5.6
0.4	100	2.8	4.6	5.1	2.3	2.6	5.1	0.5	2.3	4.1	0.6	2.4	4.0	0.6	2.4	4.2
	500	3.6	4.4	5.3	3.3	3.7	5.3	2.6	3.7	5.4	2.6	3.7	5.2	2.8	3.8	5.6
0.2	100	2.0	3.6	4.6	2.0	2.4	5.1	0.4	1.8	3.8	0.5	2.1	4.1	0.4	1.7	3.8
	500	3.0	3.9	5.0	2.7	3.4	5.4	2.0	3.2	5.1	2.2	3.2	5.2	1.9	3.2	4.9
0	100	1.5	3.3	5.0	1.6	2.1	5.7	0.4	1.9	3.9	0.7	2.8	5.0	0.3	1.6	3.7
	500	2.6	3.6	5.1	2.4	3.1	5.9	1.9	3.1	5.1	2.1	3.3	5.7	1.7	2.9	5.1
								MA(1) with $\theta = 0.5$								
0.8	100	3.2	4.5	4.5	3.8	3.6	5.2	2.3	4.6	4.3	3.1	5.5	5.2	2.4	4.6	3.7
	500	4.2	4.5	4.6	4.1	4.1	5.0	3.9	4.5	4.6	4.4	4.9	5.1	3.7	4.4	4.4
0.6	100	2.7	4.2	5.1	3.1	3.2	5.2	2.0	4.3	4.9	2.6	5.1	5.2	1.7	4.0	4.1
	500	3.5	4.2	4.6	3.4	3.7	4.7	3.2	4.0	4.6	3.7	4.3	4.7	3.0	3.9	4.4
0.4	100	1.7	3.2	5.1	1.8	2.1	5.0	1.1	3.6	5.0	1.4	3.8	4.8	1.1	3.5	5.0
	500	2.9	3.6	4.6	2.5	3.0	4.7	2.7	3.7	4.7	2.9	3.8	4.8	2.6	3.5	4.6
0.2	100	0.9	2.4	4.6	1.1	1.4	4.7	0.5	2.7	4.8	0.6	2.9	4.9	0.5	2.6	4.8
	500	2.3	3.1	4.3	1.9	2.4	4.7	2.0	3.1	4.4	2.2	3.2	4.5	2.0	3.0	4.3
0	100	0.5	1.8	4.5	0.8	1.1	4.1	0.3	2.3	4.7	0.4	2.3	4.3	0.4	2.4	4.8
	500	2.0	2.8	4.3	1.5	1.8	4.6	1.9	3.0	4.4	1.9	2.9	4.2	2.0	3.0	4.3
								MA(1) with $\theta = -0.5$								
0.8	100	7.0	9.4	7.4	4.2	4.6	6.1	2.1	5.0	5.7	1.9	4.3	5.3	2.3	5.6	5.8
	500	6.3	7.1	6.2	4.9	5.2	5.8	4.1	5.2	5.9	4.2	5.0	5.8	4.3	5.6	6.0
0.6	100	5.9	8.4	7.2	3.6	4.1	6.1	1.6	4.5	5.6	1.6	4.2	5.3	1.7	4.8	5.8
	500	5.5	6.4	6.0	4.2	4.6	5.7	3.5	4.4	5.7	3.4	4.3	5.6	3.7	4.7	5.6
0.4	100	4.5	7.4	7.0	2.9	3.4	5.9	1.2	3.8	5.4	1.3	3.9	5.3	1.2	3.8	5.4
	500	4.7	5.8	5.9	3.4	4.0	6.0	2.8	4.0	5.7	2.8	3.8	5.7	2.8	4.0	5.8
0.2	100	3.8	6.6	6.9	2.8	3.5	6.5	1.1	3.7	5.4	1.2	4.1	5.6	1.0	3.4	5.0
	500	4.0	4.9	5.9	2.9	3.6	6.2	2.1	3.2	5.6	2.3	3.4	5.8	1.8	3.0	5.5
0	100	3.5	6.0	6.8	2.7	3.0	6.6	1.1	3.8	5.2	1.7	5.1	7.0	1.0	3.2	4.8
	500	3.4	4.4	5.8	2.2	2.9	6.1	2.1	3.2	5.7	2.4	4.1	6.6	1.7	2.9	5.4

Notes: see Table 2.

Table 4. Standard and simulation-based unit root tests, U.S. 3-month Treasury Bills rate, monthly data 1957–2008

Standard Unit Root Tests			Simulation-based Unit Root Tests			
	statistic	p-values		p-values		
				(a)	(b)	(c)
$\mathcal{ADF}_\alpha$	−22.580	0.006	$\mathcal{ADF}^*_\alpha$	0.073	0.068	0.086
$\mathcal{ADF}_t$	−2.945	0.004	$\mathcal{ADF}^*_t$	0.142	0.128	0.118
$\mathcal{MZ}_\alpha$	−25.195	0.003	$\mathcal{MZ}^*_\alpha$	0.053	0.061	0.088
$\mathcal{MZ}_t$	−3.502	0.009	$\mathcal{MZ}^*_t$	0.057	0.061	0.092
$\mathcal{MSB}$	0.139	0.003	$\mathcal{MSB}^*$	0.053	0.062	0.084

Notes: (i) The number of lags determined by MAIC is $k = 16$. Columns a, b and c denote simulation-based p-values with $n = 20,000$ (a), $n = T$ (b) and with recolouring device (c). (ii) The estimated long run variance is $s^2_{AR}(k) = 0.63$. (iii) The estimated bound parameter is $\widehat{\underline{c}} = -0.16$.

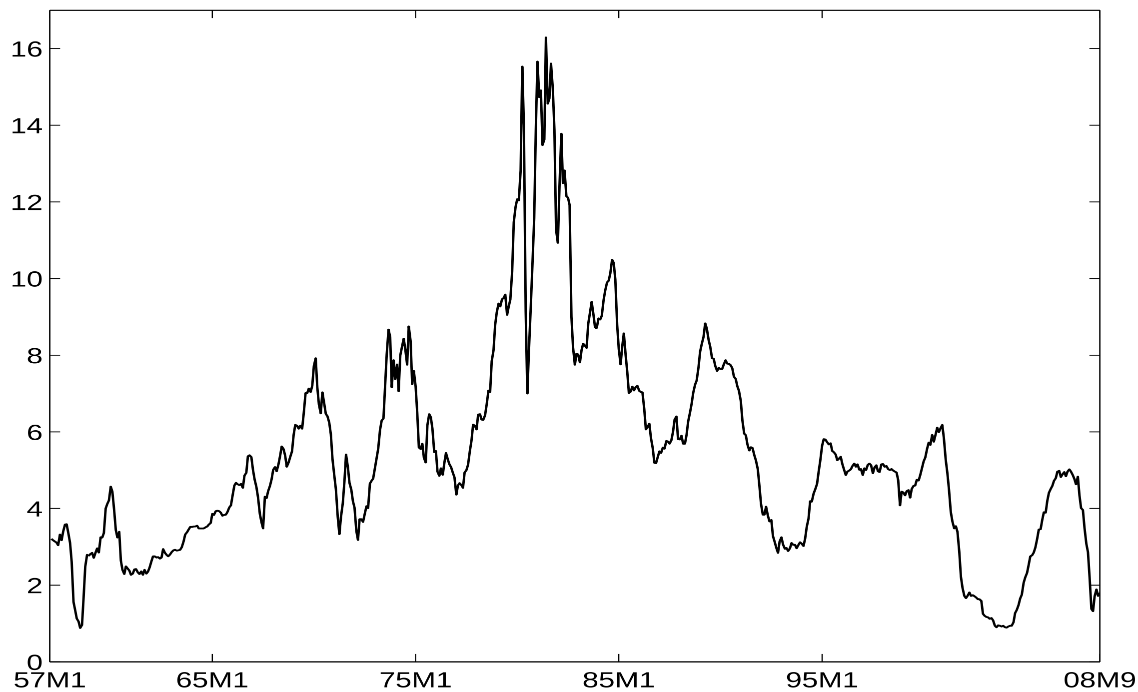

Figure 1: U.S. 3-month Treasury Bill rate, monthly data 1957:01-2008:09