Econometric Theory, **21**, 2005, 907–945. Printed in the United States of America.
DOI: 10.1017/S0266466605050462

LIMITED TIME SERIES WITH A UNIT ROOT

GIUSEPPE CAVALIERE
University of Bologna

This paper develops an asymptotic theory for integrated and near-integrated time series whose range is constrained in some ways. Such a framework arises when integration and cointegration analyses are applied to time series that are bounded either by construction or because they are subject to control. The asymptotic properties of some commonly used integration tests are discussed; the bounded unit root distribution is introduced to describe the limiting distribution of the sample first-order autoregressive coefficient of a random walk under range constraints. The theoretical results show that the presence of such constraints can lead to drastically different asymptotics. Because deviations from the standard unit root theory are measured through two noncentrality parameters that can be consistently estimated, simple measures of the impact of range constraints on the asymptotic distributions are obtained. Generalizations of standard unit root tests that are robust to the presence of range constraints are also provided. Finally, it is shown that the proposed asymptotic framework provides an adequate approximation to the finite-sample properties of the unit root statistics under range constraints.

1. INTRODUCTION

Despite the extensive literature on modeling nonstationary economic time series and on limited-dependent variables, a controversial and rarely discussed topic is how to interpret and analyze time series whose behaviors can be well approximated by means of integrated processes, I(1), but are "limited" in the sense that their range is constrained by fixed bounds. Common cases arise in the context of composition ratios, such as expenditure shares or unemployment rates, or in the presence of nonnegativity restrictions, e.g., for nominal interest rates. Moreover, range constraints represent the standard framework in the context of target zone exchange rates and, more generally, of time series that are subject to control. In the following discussion, time series satisfying (one-sided or two-sided) range restrictions are called "limited" or "bounded."

Partial financial support from Italian MIUR grants is gratefully acknowledged. I thank, without implicating, Pentti Saikkonen (the co-editor), an anonymous referee, Attilio Gardini, Martin Jacobsen, Robert de Jong, Paolo Paruolo, Anders Rahbek, and participants at the 58th European Meeting of the Econometric Society, Stockholm, August 21–24, 2003, for helpful comments. I also thank the Bank of International Settlements for providing the European monetary system exchange rate data. Address correspondence to Giuseppe Cavaliere, Department of Statistical Sciences, University of Bologna, Via Belle Arti 41, 40126 Bologna, Italy; e-mail: cavaliere@stat.unibo.it.

 0266-4666/05 $12.00

Although limited time series cannot be integrated in the usual sense, in many theoretical and econometric studies they are modeled in the I(1) framework. For example, several empirical models of European Monetary System exchange rates have been specified by using (co-)integrated vector autoregressive (VAR) models without taking account of the presence of the target zones (see, among many others, Anthony and MacDonald, 1998; Svensson, 1993; see also Phillips, 2001, and references therein). Similarly, in their influential paper Nelson and Plosser (1982) reject the unit root hypothesis in the U.S. unemployment rate (see the discussions in Caner and Hansen, 2001; Abadir and Taylor, 1999), whereas several authors have looked for possible cointegrating relations linking unemployment rates to other variables. Some theoretical models consistent with a unit root in the unemployment rate have also been proposed (Blanchard and Summers, 1986; Lindbeck and Snower, 1989). In some cases, preliminary data transformations are applied to deal with time series with infinite support. However, such transformations (i) may worsen the overall model fit (Caner and Hansen, 2001, pp. 1581–1582, for the U.S. male unemployment rate) and (ii) cannot be applied when the time series of interest reaches the upper or the lower bound.

A partial attempt to define I(1) processes with range constraints is made by Barr and Cuthbertson (1991, n. 21), regarding investment shares. They observe that

> although shares are I(1) in the data there is a theoretical problem in that shares are bounded. Shares cannot be a random walk since such a series is unbounded. However, a random walk is a very special case of an I(1) series, namely linear with an additive Gaussian error. Near the boundary shares must have a non-Gaussian error. Similar considerations apply to variables such as the percentage unemployment and bilateral exchange rates (which are bounded below) which have been examined using cointegration techniques.

A more formal explanation is given by Nicolau (2002), who shows that (quasi-) integrated dynamics can arise when there are nominal bounds also. What is not generally discussed in the literature is (i) why in some cases the constraints can reasonably be neglected and therefore standard I(1) modeling is still appropriate, (ii) to what extent the misspecification depending on the omission of the range constraints affects standard (co-)integration tests, and (iii) how to test for unit roots (or cointegration) in the presence of range constraints.

The existence of range constraints makes the interpretation of the outcome of unit root tests controversial. Suppose, e.g., that the unit root hypothesis is rejected. Should such a rejection be attributed to the absence of a unit root, or does it depend on the presence of range constraints? Standard unit root tests cannot provide an answer to this question.

In this paper all of these issues will be addressed within a unified framework. This aim will be achieved by developing a new asymptotic theory that accounts for both nonstationarity and presence of bounds. This approach allows us to generalize the large-sample theory for integrated and near-integrated processes to the case of range constraints. The unit root distribution can be extended to the case of bounds; the standard theory (see, e.g., Phillips, 1987a, 1987b) is obtained as a special case. The limiting "bounded unit root" distribution depends on two noncentrality parameters that are expressed in terms of the position of the bounds. Such parameters can be estimated and allow a rapid evaluation of the impact of the range constraints on the properties of unit root tests. Moreover, they allow generalization of unit root inference to limited time series.

The paper is organized as follows. The next section defines integrated and near-integrated processes under range constraints, and some basic asymptotic results are obtained. In Section 3 the bounded unit root distribution is derived and analyzed in the context of the unit root statistics. Implications for unit root tests are discussed in Section 4. In Section 5 the problem of testing for unit roots in the presence of bounds is tackled. Two illustrative applications are reported in Section 6. Section 7 concludes with a discussion on some possible extensions of the results obtained.

2. LIMITED AUTOREGRESSIVE PROCESSES

We start by considering a standard, integrated or near-integrated (NI(1)), real-valued AR(1) process $\{S_t\}$

$$S_t = \rho_T S_{t-1} + \varepsilon_t,$$

where ρ_T is unity or near unity, and a stochastic process $\{X_t\}$ that is obtained by properly mapping the sample paths of $\{S_t\}$ onto an interval $[\underline{b}, \bar{b}]$. This can be done by requiring that $\{\Delta X_t\}$ depends on $\{\Delta S_t\}$ in such a way that the constraint

$$X_t \in [\underline{b}, \bar{b}]$$

holds almost surely for all t (i.e., $\Delta X_t \in [\underline{b} - X_{t-1}, \bar{b} - X_{t-1}]$ a.s., all t). The simplest process satisfying these requirements is obtained by assuming that, conditionally on X_{t-1}, ΔX_t is determined by truncating ΔS_t at $\underline{b} - X_{t-1}$ and $\bar{b} - X_{t-1}$; alternatively, by censoring ΔS_t at $\underline{b} - X_{t-1}$ and $\bar{b} - X_{t-1}$. Both (censored and truncated) types of behavior near the bounds can be nested within a more general class of limited processes. The conditions that define such a class are given as follows.[1]

DEFINITION 1. *A stochastic process* $\{X_t\}_0^T$ *is called "limited near-integrated of order 1," or "bounded near-integrated of order 1," briefly* BNI(1), *if the range constraint*

$$X_t \in [\underline{b}, \bar{b}], \quad a.s., \quad all\ t \tag{1}$$

and the following assumptions hold:

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

(A1) the differenced process $\Delta X_t =: u_t$ *can be decomposed as* $u_t = \Delta S_t + \underline{\xi}_t - \bar{\xi}_t$, *where*

$$S_t = \rho_T S_{t-1} + \varepsilon_t, \qquad \rho_T := \exp(-\alpha/T), \qquad \alpha \geq 0, \qquad S_0 \overset{a.s.}{=} 0 \tag{2}$$

and $\{\underline{\xi}_t\},\{\bar{\xi}_t\}$ *are nonnegative processes such that* $\underline{\xi}_t > 0$ *iff* $X_{t-1} + \Delta S_t < \underline{b}$ *and* $\bar{\xi}_t > 0$ *iff* $X_{t-1} + \Delta S_t > \bar{b}$; *specifically, if* $X_{t-1} + \Delta S_t < \underline{b}$, *then* $\underline{\xi}_t$ *satisfies the constraint*

$$\underline{b} - (X_{t-1} + \Delta S_t) \leq \underline{\xi}_t \leq \bar{b} - (X_{t-1} + \Delta S_t) \quad a.s., \tag{3}$$

and, similarly, if $X_{t-1} + \Delta S_t > \bar{b}$, *then* $\bar{\xi}_t$ *satisfies the constraint*

$$(X_{t-1} + \Delta S_t) - \bar{b} \leq \bar{\xi}_t \leq (X_{t-1} + \Delta S_t) - \underline{b} \quad a.s.; \tag{4}$$

(A2) $\{\varepsilon_t\}$ *(see eqn. (2)), is a zero-mean process satisfying restrictions sufficient to ensure that, for some* $\lambda \in (0,\infty)$, $S_T(\cdot) := (\lambda^2 T)^{-1/2} \sum_{t=1}^{[\cdot T]} \varepsilon_t \overset{w}{\longrightarrow} B(\cdot)$, *B being a standard Brownian motion;*

(A3) $\{\underline{\xi}_t\}$ *and* $\{\bar{\xi}_t\}$ *satisfy restrictions sufficient to ensure that* $\max_{t=1,\ldots,T}|\underline{\xi}_t|$ *and* $\max_{t=1,\ldots,T}|\bar{\xi}_t|$ *are of* $o_p(T^{1/2})$;

(A4) $\underline{b} = \underline{c}\lambda T^{1/2}$, $\bar{b} = \bar{c}\lambda T^{1/2}$, $X_0 = c_0 \lambda T^{1/2}$, $\underline{c} \leq c_0 \leq \bar{c}$, $|c_0| < \infty$, $\underline{c} \neq \bar{c}$.

When $\alpha = 0$, $\{X_t\}$ *is called "limited integrated of order 1" or "bounded integrated of order 1," briefly* BI(1).

Remarks.

2.1. Assumption $\mathcal{A} := \{\mathrm{A1, A2, A3, A4}\}$ defines the time series behavior of $\{X_t\}$. The basic idea is to separate the effect of the bounds from the dynamics that characterize $\{\Delta X_t\}$ in the absence of range restrictions. This is achieved through the decomposition of the differenced process $\{u_t\}$ given in Assumption (A1). Specifically, (A1) implies that $\{X_t\}$ has the component representation $\Delta X_t = \Delta S_t + \underline{\xi}_t - \bar{\xi}_t$, and hence $X_t = X_0 + S_t + M_t$, $M_t := \sum_{i=1}^{t}(\underline{\xi}_i - \bar{\xi}_i)$. Because in the absence of bounds $X_t = X_0 + S_t$, the difference $M_t = (X_t - X_0) - S_t$ represents the (cumulated) amount that controls the trajectory of $\{X_t\}$ to satisfy the range constraint (1). Note that $\underline{\xi}_t$ and $\bar{\xi}_t$ are different from zero if and only if $X_{t-1} + \Delta S_t$ does not respect the range constraint. When $X_{t-1} + \Delta S_t > \bar{b}$ ($X_{t-1} + \Delta S_t < \underline{b}$), $\bar{\xi}_t$ ($\underline{\xi}_t$) is large enough to ensure that X_t satisfies (1). The rationale behind Assumption $\mathcal{A}$ is that any truncated/censored/reflected random variable (r.v.) can be written as a random transformation of a r.v. with infinite support.[2]

2.2. The process $\{\varepsilon_t\}$ (see (A2)) satisfies an invariance principle and is therefore Rényi-mixing (Phillips and Ouliaris, 1990), or I(0) in a broad sense; see also Davidson (2002). Hence, the local-to-unity asymptotics of Phillips (1987b) imply that $(\lambda^2 T)^{-1/2} S_{[sT]} \overset{w}{\longrightarrow} J_\alpha(s) := \int_0^s \exp\{-\alpha(s-r)\}\,dB(r)$, i.e., an Ornstein–Uhlenbeck process. For $\alpha > 0$, $\{S_t\}$ is therefore near-integrated. In the special case $\alpha = 0$, $\Delta S_t = \varepsilon_t$ and $\{S_t\}$ is I(1). Note that (A2) also

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

implies that the running maximum of $|\varepsilon_t|$ does not diverge too much, i.e., that $\max_{t=1,\dots,T}|\varepsilon_t|$ is of $o_p(T^{1/2})$.[3]

2.3. Assumption (A3) is a technical condition that is needed to prevent $\{X_t\}$ from "jumping" at the bounds. Specifically, (A2) and (A3) together imply that $\max_{t=1,\dots,T}|\Delta X_t|$ is of $o_p(T^{1/2})$. It is also worth noting that for some bounded processes Assumption (A3) automatically follows from (A1), (A2), and (A4). Consider, e.g., the simple bounded random walk $\{X_t\}$, recursively obtained by censoring $X_{t-1}+\varepsilon_t$ at $\underline{b}$ and $\bar{b}$ (see Cox and Miller, 1965), which is a bounded I(1) process with $\underline{\xi}_t := [\underline{b}-(X_{t-1}+\varepsilon_t)]\mathbb{I}\{X_{t-1}+\varepsilon_t<\underline{b}\}$ and $\bar{\xi}_t := [(X_{t-1}+\varepsilon_t)-\bar{b}]\mathbb{I}\{X_{t-1}+\varepsilon_t>\bar{b}\}$. In this case $\underline{\xi}_t,\bar{\xi}_t \le |\varepsilon_t|$ and by Assumption (A2) $\max_{t=1,\dots,T}\underline{\xi}_t, \max_{t=1,\dots,T}\bar{\xi}_t \le \max_{t=1,\dots,T}|\varepsilon_t| = o_p(T^{1/2})$ (see Remark 2.2); hence, (A3) holds. In general, however, (A3) does not follow from (A1), (A2), and (A4) as these assumptions do not rule out jumps of magnitude $O_p(T^{1/2})$ when the process breaches the bounds. To see this fact, suppose, e.g., that $\underline{\xi}_t$ is defined in a slightly different way:

$$\underline{\xi}_t := [\underline{b}-(X_{t-1}+\varepsilon_t)+\mathcal{P}_t]\mathbb{I}\{X_{t-1}+\varepsilon_t<\underline{b}\},$$

where $\mathcal{P}_t := (\bar{b}-\underline{b})/2 = T^{1/2}\lambda(\bar{c}-\underline{c})/2$. In this case, once the lower bound is breached, the process jumps to the middle of the interval $[\underline{b},\bar{b}]$. It therefore follows that

$$\begin{aligned}\max_{t=1,\dots,T}\underline{\xi}_t &= \max_{t=1,\dots,T}((\underline{b}-(X_{t-1}+\varepsilon_t))\mathbb{I}\{X_{t-1}+\varepsilon_t<\underline{b}\}+\mathcal{P}_t\mathbb{I}\{X_{t-1}+\varepsilon_t<\underline{b}\})\\ &\ge \max_{t=1,\dots,T}\mathcal{P}_t\mathbb{I}\{X_{t-1}+\varepsilon_t<\underline{b}\} = T^{1/2}\frac{\lambda(\bar{c}-\underline{c})}{2}A_T,\end{aligned}$$

with $A_T := 1-\prod_{t=1}^{T}\mathbb{I}\{X_{t-1}+\varepsilon_t\ge\underline{b}\}$, so that if $\Pr\{A_T=1\}\ne o_p(1)$, Assumptions (A1), (A2), and (A4) do not imply (A3). In this respect, Assumption (A3) allows us to rule out jumps that are not asymptotically negligible (i.e., that are not of $o_p(T^{1/2})$).

2.4. Assumption (A4) may appear unusual. It states a relation between the position of the bounds $(\underline{b},\bar{b})$ and the sample size T. (A4) is a key condition to assess both empirically and theoretically to what extent the bounds impact on the behavior of the process. As will be stressed later on, $\underline{c}$ and $\bar{c}$ in (A4) provide a way to measure the influence of such bounds in finite samples. Moreover, they allow derivation of an asymptotic theory in the presence of range restrictions without modeling the behavior of the process near the bounds in a parametric fashion. Finally, (A4) enables us to obtain a convenient unification of the (near-)unit root asymptotic theory with the limited-dependent variable framework, and also to modify standard unit root inference to take account of the range constraints properly.

2.5. Because the bound parameters $\underline{b}$ and $\bar{b}$ (and also the initial value X_0) depend on T, a time series generated according to Definition 1 formally constitutes a triangular array of the type $\{X_{Tt}: t=0,1,\dots,T;\ T=0,1,\dots\}$ (see, e.g., Phillips, 1987b). This notation is not essential to the discussion that follows; hence, limited (near-)integrated processes will be simply denoted as $\{X_t\}$. Note that this richer notation would also allow us to justify Assumption (A4) by referring to the so-called triangular array asymptotics (Andrews and McDermott, 1995), where the sample size is fixed at T_0 and the model is imbedded in the triangular array $\{X_{Tt}\}$ (see also Saikkonen and Choi, 2004).

A bounded near-integrated process reverts (i) because of the bounds $[\underline{b}, \bar{b}]$ and, if $\alpha > 0$ in equation (2), (ii) because its driving process $\{S_t\}$ has no unit roots. In the special case of bounded integrated dynamics, i.e., $\alpha = 0$, the process is mean reverting in the close neighborhood of $\underline{b}$ and $\bar{b}$ only; hence BI(1) processes have a unit root but differ from standard I(1) processes because of the range constraints. In Section 5 it will be shown how the constraint $\alpha = 0$ implied by the BI(1) model can be tested against the alternative hypothesis of limited autoregressive dynamics without a unit root.

A basic result of this paper is that BI(1) processes satisfy an invariance principle as in the standard I(1) framework; in this case, however, the limiting process is not a Brownian motion in general but depends on the parameters c_0, $\underline{c}$, and $\bar{c}$ of Assumption (A4). To derive this property, we need to introduce the following definition.

DEFINITION 2. *Let Z be a stochastic process in $\mathcal{C}$ with $Z(0) \in [a,b]$, $a < b$. The bivariate process (L,U) is said to be a "two-sided regulator" of Z, with bounds a,b, if (i) $Z_a^b(s) := Z(s) + L(s) - U(s) \in [a,b]$, (ii) L and U are increasing and continuous with $L(0) = U(0) = 0$ a.s., (iii) L and U increase only when $Z_a^b = a$ and $Z_a^b = b$, respectively. If $Z = B$, i.e., a standard Brownian motion, then Z_a^b is called "regulated Brownian motion."*

The two-sided regulator controls the trajectory of a process in $\mathcal{C}$ by keeping its sample paths between the given bounds a,b; the regulated process lies in $\mathcal{C}$ also. Note that for $s = 0$ the (nondecreasing) regulators L and U equal 0 and hence $Z_a^b = Z$ until the first time at which Z is about to cross one of the two bounds a,b. Then, although Z could escape out of the interval $[a,b]$, the regulated process is forced to lie within the interval because of the two-sided regulator (L,U). The reader can refer to Harrison (1985) and Dixit (1993) for further insights.

Consider the following continuous-time approximant of $\{X_t\}$:

$$X_T(s) := \frac{1}{\lambda T^{1/2}} X_{[sT]}, \qquad s \in [0,1]. \tag{5}$$

Then, in the unit root case ($\alpha = 0$) the next theorem follows.

THEOREM 1. *Let $\{X_t\}$ be a BNI(1) process (see Definition 1). Moreover, let $X_T(\cdot)$ be defined as in (5). Then, if $\alpha = 0$, as $T \uparrow \infty$ $X_T(\cdot) - X_T(0) \xrightarrow{w}$*

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

$B_{\underline{c}-c_0}^{\bar{c}-c_0}(\cdot)$, *where* $B_{\underline{c}-c_0}^{\bar{c}-c_0}$ *is a regulated Brownian motion with bounds at* $\underline{c} - c_0$, $\bar{c} - c_0$.

This result differs from standard I(1) asymptotics mainly because the limiting process is not a standard Brownian motion, as, e.g., in Phillips (1987a), but is a *regulated* Brownian motion. The sample paths of the limiting process are therefore bounded between $\underline{c} - c_0$ and $\bar{c} - c_0$. Theorem 1 obviously nests usual asymptotics because, for $c_0 - \underline{c}$ and $\bar{c} - c_0$ equal to infinity, the standard invariance principle follows. Finally, note also that the weak convergence given in Theorem 1 allows us to consider a BI(1) process as a discrete analog of the regulated Brownian motion.

Remark.

2.7. When $b = \infty$ and a is finite, the regulator in Definition 2 is said to be a "one-sided regulator" and the regulated process can equivalently be defined as

$$Z_a^{+\infty}(s) := Z(s) + L(s), \qquad L(s) := -\left\{0 \wedge \inf_{0 \le s' \le s} (Z(s') - a)\right\} \tag{6}$$

(see Harrison, 1985, Prop. 2.3). Hence, if $\bar{c} = +\infty$ the (one-sided) regulated Brownian motion in Theorem 1 can be expressed as $B_{\underline{c}-c_0}^{+\infty}(s) := B(s) - \{0 \wedge \inf_{0 \le s' \le s}(B(s') - (\underline{c} - c_0))\}$. The case of an upper bound follows similarly.

3. THE "BOUNDED UNIT ROOT" DISTRIBUTION

In this section we will show how the presence of range constraints modifies the asymptotic framework of unit root tests. Specifically, by relying on the weak convergence result given in the previous section, the so-called unit root distribution will be generalized to the case of bounds.

Given a sample $\{X_t\}_0^T$ drawn from a BI(1) process, with $X_0 = 0$, let $\hat{\rho}_T$ be the sample first-order autoregressive coefficient, which solves $\hat{\rho}_T \sum_{t=1}^T X_{t-1}^2 = \sum_{t=1}^T X_t X_{t-1}$. It is well known that if no range constraints are imposed, namely, if $\underline{c} = -\infty$ and $\bar{c} = +\infty$, under well-known conditions[4] $\hat{\rho}_T$ has the following asymptotic distribution (see Phillips, 1987a):

$$T(\hat{\rho}_T - 1) \xrightarrow{w} \frac{B(1)^2 - \sigma^2/\lambda^2}{2\int_0^1 B(s)^2\, ds}. \tag{7}$$

In the special case $\lambda^2 = \sigma^2$, the asymptotic distribution (7) is known as the unit root (or Dickey–Fuller) distribution, $\mathcal{Z}$ in the following discussion. By referring to Theorem 1, it is straightforward to extend the asymptotics summarized in (7) to the case of bounds. To this purpose, we strengthen Definition 1 by requiring that the following assumption holds.

Assumption $\mathcal{B}$. (B1) $\{\varepsilon_t\}$ (see (A2) of Definition 1) is the linear process $\varepsilon_t := \sum_{j=0}^{\infty} c_j v_{t-j}$, where $C(z) := \sum_{j=0}^{\infty} c_j z^j \neq 0$ for all $|z| \leq 1$ and $\sum_{j=0}^{\infty} j|c_j| < \infty$; $\{v_t, \mathcal{F}_t\}$ is a martingale difference sequence with respect to some filtration $\mathcal{F}_t$ satisfying $E(v_t^2|\mathcal{F}_{t-1}) = \sigma_v^2 < \infty$ and $\sup_{t\in\mathbb{Z}} E|v_t|^{2+\eta} < \infty$ for some $\eta > 0$; (B2) for some $\eta, \bar{\eta} > 0$, $\{\underline{\xi}_t\}$ and $\{\bar{\xi}_t\}$ (see (A3) of Definition 1) satisfy $\sup_{t=1,\dots,T} E|\underline{\xi}_t|^{2+\underline{\eta}} < \infty$ and $\sup_{t=1,\dots,T} E|\bar{\xi}_t|^{2+\bar{\eta}} < \infty$.

The unit root distribution in the presence of range constraints is presented in the next theorem.

THEOREM 2. *Let $\{X_t\}$ be a BNI(1) process (see Definition 1); moreover, suppose that Assumption $\mathcal{B}$ also holds. Then, if $\alpha = 0$ and $X_0 = 0$, as $T \uparrow \infty$,*

$$Z_\rho := T(\hat{\rho}_T - 1) \xrightarrow{w} \frac{B_{\underline{c}}^{\bar{c}}(1)^2 - \sigma^2/\lambda^2}{2\int_0^1 B_{\underline{c}}^{\bar{c}}(s)^2\, ds}, \tag{8}$$

where $B_{\underline{c}}^{\bar{c}}$ is a regulated Brownian motion with bounds at $\underline{c}, \bar{c}$, $\sigma^2 := \sigma_v^2 \sum_{j=0}^{\infty} c_j^2$, and $\lambda^2 := \sigma_v^2 C(1)^2$.

For $\lambda^2 = \sigma^2$ the sample autoregressive coefficient is asymptotically distributed as $(2\int B_{\underline{c}}^{\bar{c}}(s)^2\, ds)^{-1}(B_{\underline{c}}^{\bar{c}}(1)^2 - 1)$, which differs from the usual unit root distribution only because it is expressed in terms of functionals of a regulated Brownian motion and not of a standard Brownian motion. This distribution is called a "bounded unit root distribution," with parameters $\underline{c}$ and $\bar{c}$, and denoted as $\mathcal{Z}(\underline{c}, \bar{c})$.

Kernel estimates of the probability density function (p.d.f.) associated with the bounded unit root distribution for various values of $\bar{c} = -\underline{c} =: c > 0$, i.e., under symmetric bounds around the origin, are reported in Figure 1. These are based on 50,000 Monte Carlo (MC) replications where the limiting regulated Brownian motion is obtained as $B_{\underline{c}}^{\bar{c}}(s) = \varphi_{\underline{c}}^{\bar{c}}(B(s))$, where $\varphi_{\underline{c}}^{\bar{c}}(\cdot)$ is the reflec-

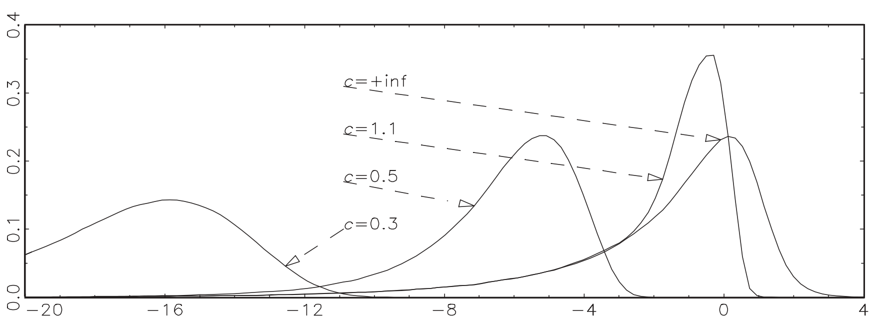


FIGURE 1. Probability density function of the bounded unit root distribution $\mathcal{Z}(-c, c)$ for $c = 0.3, 0.5, 1.1, +\infty$ (standard unit root distribution). Kernel estimates with Epanechnikov weights.

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

tion function (Karatzas and Shreve, 1988, p. 97, for $\underline{c} = 0$ and $\bar{c} = a$) and the Brownian motion is approximated by its discrete realization from a sample of size $N = 20{,}000$; the rationale behind this algorithm is that $B_{\underline{c}}^{\bar{c}} \overset{d}{=} \varphi_{\underline{c}}^{\bar{c}}(B)$ (see Harrison, 1985).[5] As expected, for c sufficiently large the bounded unit root distribution approaches the standard unit root distribution. The presence of bounds translates the asymptotic distribution of the autoregressive coefficient toward negative values: the smaller c is, the more the bounded unit root distribution is skewed toward negative values.

The 5% quantile of the bounded unit root distribution $\mathcal{Z}(-c, c)$ is reported in Figure 2 for various values of c.[6] A selection of 5% quantiles of the unit root distribution under symmetric bounds is also reported in Table 1 (third column, with $c_0/\underline{c} = 0$). Again, for c sufficiently large the quantiles correspond to those of the standard unit root distribution. The 5% quantile diverges to $-\infty$ as c approaches 0.

Several implications can be derived from the preceding results. First, in the presence of range constraints the large-sample distribution of the first-order autoregressive coefficient is nonstandard. With respect to the usual unit root distribution, the limiting "bounded unit root" distribution has two more noncentrality parameters, $\underline{c}$ and $\bar{c}$. Second, the quantiles of the bounded unit root distribution can be extremely different from those of the standard unit root distribution. To which extent the quantiles differ depends (i) on the distance of the bounds from the initial value of the process (through the parameters $\underline{c}, \bar{c}$) and (ii) on the variability of the innovations to $\{X_t\}$ (through the long-run variance λ^2). Third, only for bounds sufficiently far from the starting value of the process the quantiles of the bounded unit root distribution are well approximated by the quantiles of the standard unit root distribution.

Remarks.

3.1 (Initial conditions). The previous derivation of the bounded unit root distribution is based on the condition $X_0 = 0$. However, if the process starts in $X_0 := c_0 \lambda T^{1/2}$, where $c_0 \in [\underline{c}, \bar{c}]$, the weak convergence (8) still holds with $(\underline{c}, \bar{c})$ replaced by $(\underline{c} - c_0, \bar{c} - c_0)$, provided that $\hat{\rho}_T$ is based on the deviations from the initial value, i.e., on $\{X_t - X_0\}$. Note that the (5%) quantiles of the bounded unit root distribution are highly sensitive to the presence of asymmetric bounds (i.e., $-(\underline{c} - c_0) \neq \bar{c} - c_0$); see Table 1, third column (Z_ρ).

3.2. (One-sided bounds). One-sided bounds can be treated as a special case by setting $\bar{c} = +\infty$ (lower bound only) or $\underline{c} = -\infty$ (upper bound only). By letting $X_0 = 0$ and $\bar{c} = +\infty$ the 5% quantile of the bounded unit root distribution $\mathcal{Z}(\underline{c}, +\infty)$ is reported in Figure 3 for various values of $c := -\underline{c}$, and a selection of 5% quantiles is reported in Table 2, second column (Z_ρ). It is interesting to observe that, as $\underline{c} \uparrow 0$, i.e., the process starts at the lower bound, the quantiles converge to those of the standard unit root distribution. This result—which has already been pointed out in Cavaliere (2003) under more restrictive assumptions—follows from the distribution equality $B_0^{+\infty} \overset{d}{=} |B|$ (see Harrison, 1985, p. xii), which implies that $\mathcal{Z}(0, +\infty) \overset{d}{=} \mathcal{Z}$.

TABLE 1. Critical values of the Phillips–Perron unit root tests in the two-bound (at $-\underline{c}$ and at $\bar{c}$) case

		Tests based on $\{X_t - X_0\}$		Tests based on $\{X_t - \bar{X}\}$	
$\bar{c} = -\underline{c}$	$c_0/\underline{c}$	Z_ρ	Z_t	$Z_{\rho.c}$	$Z_{\rho.t}$
0.3	0	−22.715	−3.340	−31.834	−3.972
	$\frac{1}{4}$	−25.992	−3.569	−31.607	−3.967
	$\frac{1}{2}$	−24.700	−3.472	−31.438	−3.977
	$\frac{3}{4}$	−15.945	−2.773	−31.504	−4.025
	1	−9.093	−2.082	−32.339	−4.136
0.5	0	−10.889	−2.290	−25.078	−3.616
	$\frac{1}{4}$	−13.639	−2.564	−25.246	−3.624
	$\frac{1}{2}$	−21.750	−3.262	−25.989	−3.655
	$\frac{3}{4}$	−18.905	−3.019	−26.778	−3.709
	1	−8.106	−1.946	−27.971	−3.857
0.7	0	−8.309	−1.979	−20.159	−3.372
	$\frac{1}{4}$	−9.550	−2.132	−21.205	−3.411
	$\frac{1}{2}$	−14.753	−2.675	−23.689	−3.497
	$\frac{3}{4}$	−23.416	−3.372	−26.030	−3.613
	1	−8.106	−1.946	−27.822	−3.808
0.9	0	−8.106	−1.946	−16.851	−3.172
	$\frac{1}{4}$	−8.277	−1.975	−18.209	−3.228
	$\frac{1}{2}$	−11.021	−2.299	−21.846	−3.383
	$\frac{3}{4}$	−23.547	−3.388	−25.455	−3.559
	1	−8.106	−1.946	−27.822	−3.793
1.1	0	−8.106	−1.946	−15.132	−3.054
	$\frac{1}{4}$	−8.112	−1.947	−16.353	−3.109
	$\frac{1}{2}$	−9.164	−2.090	−20.042	−3.274
	$\frac{3}{4}$	−19.648	−3.100	−24.646	−3.507
	1	−8.106	−1.946	−27.822	−3.787
1.5	0	−8.106	−1.946	−14.205	−2.930
	$\frac{1}{4}$	−8.106	−1.946	−14.606	−2.969
	$\frac{1}{2}$	−8.140	−1.952	−17.040	−3.117
	$\frac{3}{4}$	−13.616	−2.562	−23.166	−3.426
	1	−8.106	−1.946	−27.822	−3.786

Notes: Nominal level 5%. When $c_0/\underline{c} = 0$ the bounds are symmetric around the starting value. When $c_0/\underline{c} = 1$ the process starts on the lower bound. Critical values have been obtained through MC simulation by discretizing the limiting regulated Brownian motion over $T = 20{,}000$ segments and using 50,000 replications.

3.3 (Deterministic term corrections). If the computation of the sample first-order autoregressive coefficient $\hat{\rho}_T$ is based on the demeaned series $\{X_t - \bar{X}\}$, the bounded unit root distribution has the form

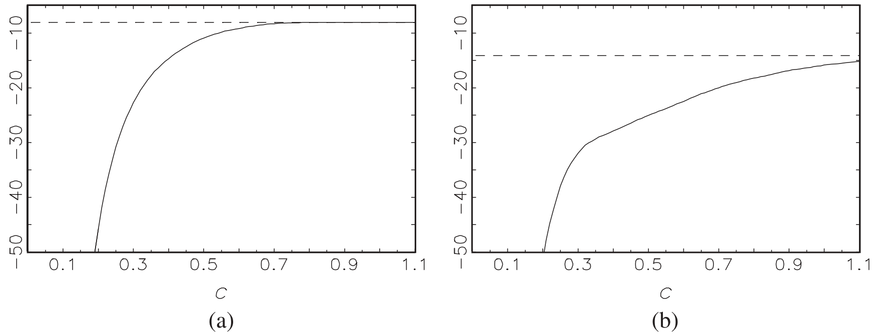


FIGURE 2. The 5% quantile of the bounded unit root distribution $\mathcal{Z}(-c, c)$ for various values of c: (a) raw data; (b) demeaned data.

$$Z_{\rho.c} \xrightarrow{w} \frac{\tilde{B}_{\underline{c}}^{\bar{c}}(1)^2 - \tilde{B}_{\underline{c}}^{\bar{c}}(0)^2 - \sigma^2/\lambda^2}{2\int_0^1 \tilde{B}_{\underline{c}}^{\bar{c}}(s)^2\, ds}, \tag{9}$$

where $\tilde{B}_{\underline{c}}^{\bar{c}} := B_{\underline{c}}^{\bar{c}} - \int_0^1 B_{\underline{c}}^{\bar{c}}(s)\, ds$ is a demeaned regulated Brownian motion. On the other hand, the Z_ρ statistic calculated after local generalized least squares (GLS) demeaning at $\bar{\rho} := 1 - \bar{\alpha}T^{-1}$ (see Elliott, Rothenberg, and Stock, 1996) has the limiting representation given in (8). In a similar way, if the computation of $\hat{\rho}_T$ involves fitting a linear time trend by ordinary least squares (OLS) the limiting distribution is given by (9) with $\tilde{B}_{\underline{c}}^{\bar{c}}$ replaced by a demeaned and detrended regulated Brownian motion, namely, $B_{\underline{c}}^{\bar{c}}(s) - \hat{\alpha} - \hat{\beta}s$, $(\hat{\alpha}, \hat{\beta}) := \arg\min_{\alpha,\beta} \int_0^1 (B_{\underline{c}}^{\bar{c}}(r) - \alpha - \beta r)^2\, dr$. Finally, Z_ρ computed on data obtained from local GLS linear detrending at $\bar{\rho} := 1 - \bar{\alpha}T^{-1}$ has the limiting representation given in (8) with $B_{\underline{c}}^{\bar{c}}(s)$ replaced by $V_{\underline{c},\bar{\alpha}}^{\bar{c}}(s) := B_{\underline{c}}^{\bar{c}}(s) - s(\bar{\lambda}_{\bar{\alpha}} B_{\underline{c}}^{\bar{c}}(1) + 3(1 - \bar{\lambda}_{\bar{\alpha}}) \int_0^1 r B_{\underline{c}}^{\bar{c}}(r)\, dr)$, where $\bar{\lambda}_{\bar{\alpha}} := (1 + \bar{\alpha})/(1 + \bar{\alpha} + \bar{\alpha}^2/3)$.

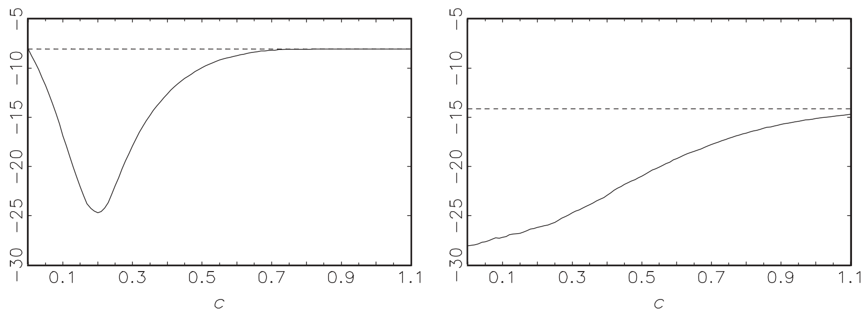
3.4. It is straightforward to notice that Assumption $\mathcal{B}$ is stronger than (A2) and (A3) of Definition 1. Specifically, under (B1) $\{\varepsilon_t\}$ satisfies the weak convergence given in (A2) with λ^2 as in Theorem 2 (see Phillips and Solo, 1992). Similarly, by standard arguments

$$\Pr\left\{\max_{t=1,\ldots,T} \underline{\xi}_t > \varepsilon T^{1/2}\right\} \le \sum_{t=1}^{T} \Pr\{\underline{\xi}_t > \varepsilon T^{1/2}\} \le \frac{\sup_{t=1,\ldots,T} E|\underline{\xi}_t|^{2+\underline{\eta}}}{\varepsilon^{2+\underline{\eta}} T^{\underline{\eta}/2}} \to 0$$

and

$$\Pr\left\{\max_{t=1,\ldots,T} \bar{\xi}_t > \varepsilon T^{1/2}\right\} \le \frac{\sup_{t=1,\ldots,T} E|\underline{\xi}_t|^{2+\bar{\eta}}}{\varepsilon^{2+\bar{\eta}} T^{\bar{\eta}/2}} \to 0;$$

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press



(a) (b)

FIGURE 3. The 5% quantile of the bounded unit root distribution $\mathcal{Z}(-c,+\infty)$ for various values of c: (a) raw data; (b) demeaned data.

TABLE 2. Critical values of the Phillips–Perron unit root tests in the one-bound (at $-\underline{c}$) case

	Tests based on $\{X_t - X_0\}$		Tests based on $\{X_t - \bar{X}\}$	
$-\underline{c}$	Z_ρ	Z_t	$Z_{\rho.c}$	$Z_{t.c}$
0.0	−8.106	−1.946	−27.822	−3.786
0.1	−16.275	−2.800	−26.615	−3.646
0.2	−24.177	−3.431	−25.813	−3.575
0.3	−17.780	−2.946	−24.350	−3.486
0.4	−12.596	−2.462	−22.761	−3.400
0.5	−9.965	−2.182	−20.974	−3.314
0.6	−8.663	−2.029	−19.169	−3.223
0.7	−8.207	−1.964	−17.715	−3.144
0.8	−8.112	−1.948	−16.525	−3.087
0.9	−8.106	−1.946	−15.668	−3.034
1.0	−8.106	−1.946	−15.067	−2.995
1.2	−8.106	−1.946	−14.448	−2.945
1.4	−8.106	−1.946	−14.216	−2.912
1.6	−8.106	−1.946	−14.144	−2.891
1.8	−8.106	−1.946	−14.139	−2.878
2.0	−8.106	−1.946	−14.139	−2.872
2.5	−8.106	−1.946	−14.139	−2.866
3.0	−8.106	−1.946	−14.139	−2.864

Notes: Nominal level 5%. When $\underline{c} = 0$ the process starts on the lower bound. Critical values have been obtained through MC simulation by discretizing the limiting regulated Brownian motion over $T = 20{,}000$ segments and using 50,000 replications.

hence, (B2) implies (A3). Under Assumption $\mathcal{B}$, the process $\{\varepsilon_t\}$ is a quite general linear process in terms of a martingale difference. It is worth noting that the conditions given in Assumption (B1), although allowing a neat derivation of the results presented in Theorem 2, are not strictly necessary. For instance, the strong mixing conditions given in Phillips (1987a, Thm. 3.1), could have also been considered. Similarly, (B2) is not strictly necessary for the results given in Theorem 2 and could have been replaced by any set of conditions ensuring that both $\max_{t=1,\dots,T} \underline{\xi}_t$ and $\max_{t=1,\dots,T} \bar{\xi}_t$ are of $o_p(T^{1/2})$.

There are important implications of the convergence results outlined previously for unit root testing. Such implications are examined in the next section.

4. IMPLICATIONS FOR UNIT ROOT TESTS

The most common approach to testing for a unit root against stable alternatives is to refer to statistic (8) as a left-sided test, i.e., to reject the null of a unit root for large negative values of $Z_\rho := T(\hat{\rho}_T - 1)$. By using the distribution results of the previous section it can be reasonably argued that in some cases the rejection rate of the test can be substantially affected by the range constraints.

To stress this result, assume that the data generating process (DGP) is BI(1) with uncorrelated homoskedastic innovations ($\sigma^2 = \lambda^2$). The rejection probabilities of the Z_ρ test when standard critical values are employed (estimated through MC simulation; see the previous section) are reported in Figure 4; the significance level is set to 5%. As expected, the rejection frequency is strongly related to the position of the bounds. There are at least two important consequences deriving from this result.

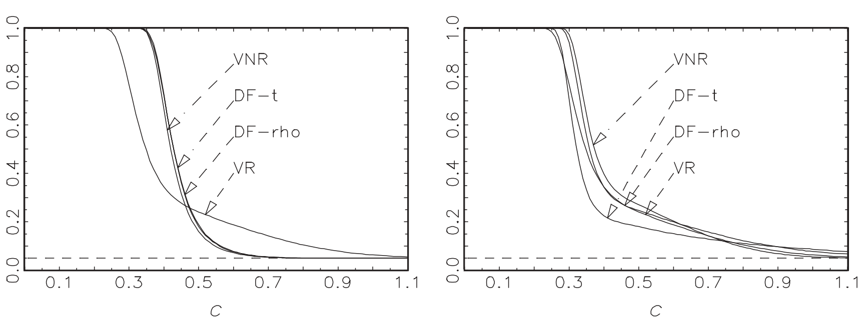


FIGURE 4. Size of the unit root tests in the two-bound case for various values of $\bar{c} = -\underline{c} =: c$ and $c_0 = 0$. DF-rho: Z_ρ unit root test; DF-t: Z_t unit root test; VNR: von Neumann ratio test; VR: variance-ratio test, $\tau = 0.5$. (a) Tests based on raw data; (b) tests based on demeaned data.

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

On the one hand, in the framework of limited time series tests based on conventional critical values could point to the rejection of the unit root hypothesis and the researcher might erroneously conclude that the process has no unit roots, whereas it *has* a unit root but it is also subject to (one-sided or two-sided) range constraints. When testing for unit roots in the presence of limited time series, an analysis of the "negligibility" of the bounds is therefore a necessary step before interpreting the outcome of standard unit root tests in the usual way. Such a step is usually missing when unit root techniques are applied to the analysis of bounded time series.

On the other hand, Figure 4 also explains why the practitioner might fail to reject the (false) standard I(1) model in the presence of bounds. When the bounds are sufficiently far away, the rejection frequencies of the unit root test when the DGP is I(1) or BI(1) are in fact identical. That is, unit root tests are not always able to detect the presence of the bounds.

A further important implication is that, despite the fact that the bounded unit root distribution (and also the asymptotic distributions of other unit root test statistics; see the discussion that follows) depends on three nuisance parameters, namely, $(\underline{c} - c_0, \bar{c} - c_0, \sigma/\lambda)$, the unit root statistic (8) can be rearranged to eliminate σ/λ from its asymptotic distribution; this can be achieved by following the approach of Phillips (1987a). Let $\hat{\lambda}^2, \hat{\sigma}^2$ be two consistent estimators for λ^2, σ^2 in the absence of bounds (i.e., when $\bar{c} = +\infty$ and $\underline{c} = -\infty$); for simplicity we assume that $\hat{\sigma}^2 := (1/T)\sum_{t=1}^{T} \hat{u}_t^2$, where $\{\hat{u}_t\}$ denotes the OLS residuals obtained by regressing X_t on X_{t-1} (and, if necessary, on a constant term), and that $\hat{\lambda}^2$ is a conventional sum-of-covariances (SC) estimator of the form

$$\hat{\lambda}^2 := \frac{1}{T} \sum_{j=-T+1}^{T-1} \omega\left(\frac{j}{q_T}\right) \sum_{t=|j|+1}^{T} \hat{u}_t \hat{u}_{t-|j|} \tag{10}$$

(see, e.g., de Jong and Davidson, 2000, and references therein). We further require the following assumption to hold (see Jansson, 2002).

Assumption $\mathcal{K}$. (K1) For all $x \in \mathbb{R}$, $\omega(x) = \omega(-x)$; $\omega(0) = 1$; $\omega(\cdot)$ is continuous at zero; $\bar{\omega}(0) < \infty$, and $\int_{[0,\infty)} \bar{\omega}(x)\,dx < \infty$, where $\bar{\omega}(x) := \sup_{x' \geq x} |\omega(x')|$; (K2) $q_T \in (0,\infty)$ and $\lim_{T\to\infty}(q_T^{-1} + T^{-\delta} q_T) = 0$ for some $\delta \in (0, \frac{1}{2}]$.

Assumption $\mathcal{K}$ places some restrictions on the kernel function $\omega(\cdot)$ and on the bandwidth/lag truncation parameter q_T, which needs to grow just more slowly than $T^{1/2}$; such restrictions, however, are rather general and almost standard in the literature. When Assumption $\mathcal{K}$ holds and $\{\varepsilon_t\}$ satisfies (B1) of Assumption $\mathcal{B}$, then the idealized estimator

$$\tilde{\lambda}^2 := \frac{1}{T} \sum_{j=-T+1}^{T-1} \omega\left(\frac{j}{q_T}\right) \sum_{t=|j|+1}^{T} \varepsilon_t \varepsilon_{t-|j|} \tag{11}$$

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

satisfies $\tilde{\lambda}^2 \xrightarrow{p} \lambda^2$ (Jansson, 2002, Thm. 2), and, if there are no bounds (i.e., $\underline{c} = -\infty$ and $\bar{c} = +\infty$) and $\{X_t\}$ has a unit or near unit root, also $\hat{\lambda}^2$ is a consistent estimator of λ^2.[7]

When there are bounds, consistency can still be preserved given that the bandwidth/lag truncation parameter q_T satisfies a regularity condition that is slightly stronger than the one given in (K2) of Assumption $\mathcal{K}$. The following lemma summarizes this result.

LEMMA 3. *Under the assumptions of Theorem 2, as* $T \uparrow \infty$, $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$. *Moreover, if Assumption* $\mathcal{K}$ *holds with (K2) replaced by the following assumption:*

(K2′). $q_T \in (0,\infty)$ *and, for some* $\delta \in (0,\delta^*]$, $\lim_{T\to\infty}(q_T^{-1} + T^{-\delta}q_T) = 0$, *where* $\delta^* := \frac{1}{2}[\eta^*/(2+\eta^*)]$, $\eta^* := \min\{\eta, \underline{\eta}, \bar{\eta}\}$;

then, as $T \uparrow \infty$, $\hat{\lambda}^2 \xrightarrow{p} \lambda^2$.

Note that (K2′) simply slows down the growth rate of q_T by establishing a trade-off between the rate at which q_T is allowed to diverge and the moments of ε_t, $\underline{\xi}_t$, and $\bar{\xi}_t$ (recall from Assumption $\mathcal{B}$ that η, $\underline{\eta}$, and $\bar{\eta}$ control the existing moments of ε_t, $\underline{\xi}_t$, and $\bar{\xi}_t$, respectively). Now, q_T is allowed to grow just more slowly than $T^{1/2}$ only in the very special case that $\eta^* = +\infty$, i.e., all moments exist. If, e.g., $\eta^* = 2$ (i.e., fourth-order moments exist) then $\delta^* = \frac{1}{4}$, and the well-known $o(T^{1/4})$ rate suggested in Phillips (1987a) is obtained.

The main consequence of Lemma 3 is that the unit root statistic of Phillips (1987a)

$$\hat{Z}_\rho := T(\hat{\rho}_T - 1) - \frac{\hat{\lambda}^2 - \hat{\sigma}^2}{2T^{-2}\sum_{t=1}^{T} X_{t-1}^2}$$

and also the asymptotically equivalent "modified" statistic (Ng and Perron, 2001)

$$MZ_\rho := \frac{T^{-1}X_T^2 - \hat{\lambda}^2}{2T^{-2}\sum_{t=1}^{T} X_{t-1}^2}$$

satisfy $\hat{Z}_\rho, MZ_\rho \xrightarrow{w} \mathcal{Z}(\underline{c}, \bar{c})$, i.e., weak convergence to a bounded unit root distribution with parameters $(\underline{c}, \bar{c})$. Hence, the presence of range constraints does not affect the consistency of the estimators of the nuisance parameters (σ^2, λ^2), which do not enter the asymptotic distribution of the test statistics. However, $\underline{c}$ and $\bar{c}$ are still two noncentrality parameters affecting the asymptotic distributions and, consequently, the outcome of the tests.

Remarks.

4.1. (One-sided bounds). Consider the one-sided bound case (see Remark 3.2). Such a case arises, e.g., in the (co-)integration analysis of nominal inter-

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

est rates (see, among others, Bec and Rahbek, 2004). For $X_0 = 0$ and $\bar{c} = +\infty$ the asymptotic rejection frequencies of the Z_ρ ($\hat{Z}_\rho$ for $\sigma^2 \neq \lambda^2$) unit root test are reported in Figure 5 for various values of $-\underline{c} := c$; the nominal level of the test is 5%. As in the two-bound case, the rejection frequency essentially depends on the distance between the starting value of the process and the position of the bound. However, if such a distance is negligible, i.e., $\underline{c} \approx 0$, the quantiles are identical to those of the standard unit root distribution (see Remark 3.2); this result does not apply when the test involves OLS demeaning (or detrending) of the original time series. In general, in the one-bound case the rejection rate is not as high as in the two-bound case; nevertheless, it can considerably exceed the significance level.

4.2 (Other unit root tests). Comparable evidence affects most of the procedures usually employed to test for unit roots. The t-ratio unit root test (Z_t) based on the t-statistic associated to $\hat{\pi}$ in the regression equation $\Delta X_t = \pi X_{t-1} +$ error has asymptotic distribution $(\lambda/\sigma)(4\int B_{\underline{c}}^{\bar{c}}(s)^2\,ds)^{-1/2}(B_{\underline{c}}^{\bar{c}}(1)^2 - (\sigma/\lambda)^2)$. By Lemma 3, Phillips' t-test $\hat{Z}_t := (\hat{\sigma}/\hat{\lambda})Z_t - (\hat{\sigma}^2 - \hat{\lambda}^2)(4\hat{\lambda}^2 T^{-2}\sum_{t=1}^{T} X_{t-1}^2)^{-1/2}$ and the corresponding "modified" statistic $MZ_t := (T^{-1}X_T^2 - \hat{\lambda}^2) \times (4T^{-2}\sum_{t=1}^{T} X_{t-1}^2)^{-1/2}$ converge weakly to $(4\int_0^1 B_{\underline{c}}^{\bar{c}}(s)^2\,ds)^{-1/2}(B_{\underline{c}}^{\bar{c}}(1)^2 - 1)$, which differs from the usual asymptotic distribution because $B_{\underline{c}}^{\bar{c}}$ replaces the Brownian motion B. The von Neumann ratio test of Sargan and Bhargava (1983), based on the statistic $VN := \hat{\lambda}^2 T^2/\sum_{t=1}^{T} X_t^2$, satisfies the convergence $VN \xrightarrow{w} (\int_0^1 B_{\underline{c}}^{\bar{c}}(s)^2\,ds)^{-1}$, which assumes larger values with respect to the standard limiting distribution $(\int_0^1 B(s)^2\,ds)^{-1}$. The variance-ratio test, based on $VR(\tau) := (\hat{\lambda}T)^{-2}\sum_{t-[\tau T]}^{T}(X_t - X_{t-[\tau T]})^2$, satisfies $VR(\tau) \xrightarrow{w} \int_\tau^1 (B_{\underline{c}}^{\bar{c}}(s) - B_{\underline{c}}^{\bar{c}}(s-\tau))^2\,ds$, which is closer to 0 than the standard limiting distribution $\int_\tau^1 (B(s) - B(s-\tau))^2\,ds$. Asymptotic sizes at the 5% nominal level are plotted in Fig-

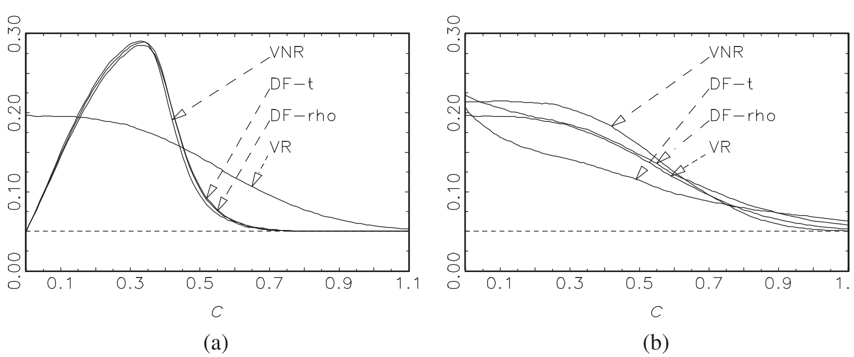


FIGURE 5. Size of the unit root tests in the one-bound case for various values of $-\underline{c} =: c$ and $c_0 = 0$. DF-rho: Z_ρ unit root test; DF-t: Z_t unit root test; VNR: von Neumann ratio test; VR: variance-ratio test, $\tau = 0.5$. (a) Tests based on raw data; (b) tests based on demeaned data.

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

ures 4 (two-bound case) and 5 (one-bound case). See also Tables 1 and 2 for the 5% quantiles of the $\hat{Z}_t$ test.

4.3 (Deterministic term corrections). As noticed for the bounded unit root distribution (see Remark 3.1), if the initial condition is $X_0 = c_0 \lambda T^{1/2}$, the results do not change provided that the tests are based on $\{X_t - X_0\}$ (or on GLS-demeaned data) and $\underline{c}, \bar{c}$ are replaced by $\underline{c} - c_0, \bar{c} - c_0$. If the tests are based on OLS-demeaned (demeaned and detrended) variables the limiting distributions depend on a demeaned (demeaned and detrended) regulated Brownian motion $\tilde{B}_{\underline{c}}^{\bar{c}}$ as follows: $\hat{Z}_t \xrightarrow{w} (4\int_0^1 \tilde{B}_{\underline{c}}^{\bar{c}}(s)^2\,ds)^{-1/2}(\tilde{B}_{\underline{c}}^{\bar{c}}(1)^2 - \tilde{B}_{\underline{c}}^{\bar{c}}(0)^2 - 1)$; $MZ_t \xrightarrow{w} (4\int_0^1 \tilde{B}_{\underline{c}}^{\bar{c}}(s)^2\,ds)^{-1/2}(\tilde{B}_{\underline{c}}^{\bar{c}}(s)^2 - 1)$, $VN^{-1} \xrightarrow{w} \int_0^1 \tilde{B}_{\underline{c}}^{\bar{c}}(s)^2\,ds$; $VR(\tau) \xrightarrow{w} \int_\tau^1 (\tilde{B}_{\underline{c}}^{\bar{c}}(s) - \tilde{B}_{\underline{c}}^{\bar{c}}(s-\tau))^2\,ds$. When the tests are based on GLS linear detrending, the limiting distributions are those given in Remark 4.2 with the regulated Brownian motion replaced by the functional $V_{\underline{c},\bar{\alpha}}^{\bar{c}}$ given in Remark 3.3.

4.4 (Empirical assessment of the impact of the bounds). Figures 2 and 4 provide simple tools for understanding to what extent in the basic BI(1) model with symmetric bounds the rejection of the unit root hypothesis could depend on the presence of the bounds. Figures 3 and 5 can be referred to in the case of a lower (upper) bound; similar pictures can be easily obtained by simulation for any value of $(\underline{c} - c_0, \bar{c} - c_0)$. Note, however, that because $\underline{c} - c_0$ and $\bar{c} - c_0$ are not generally known, they should at least be consistently estimated. Section 5 tackles this issue.

We end this section by briefly investigating the power function of the Z_ρ unit root test in the presence of near-integrated dynamics and range constraints. The DGP is therefore BNI(1) with $\alpha > 0$. Note that the I(1) hypothesis is violated (i) because the DGP has no unit roots and (ii) because of the range constraints. We might therefore expect the rejection frequency to be higher than in the usual near-integrated, NI(1), case.

To explore this issue we need to derive the asymptotic distribution of the unit root test statistics when the DGP is BNI(1) with $\alpha > 0$. The next result provides the result for the $Z_\rho, \hat{Z}_\rho$ statistics.

THEOREM 4. *Under the assumptions of Theorem 2 with* $\alpha = 0$ *replaced by* $\alpha > 0$ *then, as* $T \uparrow \infty$,

$$Z_\rho := T(\hat{\rho}_T - 1) \xrightarrow{w} \frac{J_{\underline{c}.\alpha}^{\bar{c}}(1)^2 - \sigma^2/\lambda^2}{2\displaystyle\int_0^1 J_{\underline{c}.\alpha}^{\bar{c}}(s)^2\,ds}, \tag{12}$$

where $J_{\underline{c}.\alpha}^{\bar{c}}(s) := J_\alpha(s) + L(s) - U(s)$, J_α *being the diffusion* $J_\alpha(s) := \int_0^s \exp\{-\alpha(s-r)\}\,dB(r)$ *and* $\{L, U\}$ *being the two-sided regulator of* J_α *with bounds at* $\underline{c}, \bar{c}$. *The heteroskedasticity and autocorrelation robust statistic* $\hat{Z}_\rho$ *satisfies (12) with* σ^2/λ^2 *replaced by unity, provided that Assumption* $\mathcal{K}$ *holds with (K2) replaced by (K2′).*

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

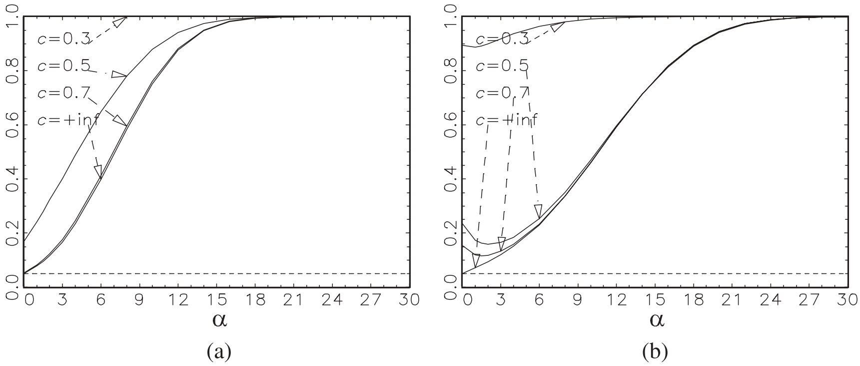


FIGURE 6. Asymptotic power function of the unit root Z_ρ test against bounded near-integrated alternatives for various values of $\bar{c} = -\underline{c} =: c$, $c_0 = 0$, and α ranging from 0 to 30: (a) tests based on raw data; (b) tests based on demeaned data.

Hence, with respect to the near-integrated framework, the asymptotic distribution of Z_ρ depends on a *regulated* Ornstein–Uhlenbeck process. For $\sigma^2 = \lambda^2$ the distribution (12) can be denoted as "bounded near-unit root distribution," $\mathcal{Z}_\alpha(\underline{c}, \bar{c})$. The theorem can be easily extended to the various unit root tests and to the case of (OLS or GLS) deterministic corrections.

The asymptotic rejection frequency of the Z_ρ ($\hat{Z}_\rho$ for $\sigma^2 \neq \lambda^2$) test is plotted for various values of α and $c := \bar{c} = -\underline{c}$ under $c_0 = 0$ (symmetric case) in Figure 6. The figures are based on an MC experiment with 50,000 replications where the limiting regulated Ornstein–Uhlenbeck process is obtained by applying the two-sided regulator (see Definition 2) to a discrete realization of the Ornstein–Uhlenbeck process over a grid of 20,000 points. In the left panel of the figure, the test is based on raw data, whereas in the right panel it is based on OLS-demeaned data. For $c = +\infty$ the usual rejection rate of the Z_ρ test when the DGP is near-integrated is obtained (see, e.g., Elliott et al., 1996, Figs. 1,2). In addition, the figure shows that in the presence of symmetric bounds, the unit root Z_ρ test tends to reject more often than in the absence of constraints.

5. TESTING THE "BOUNDED UNIT ROOT" HYPOTHESIS

When there are bounds, instead of testing the (trivially false) I(1) hypothesis the researcher should be more interested in testing the "bounded I(1)" hypothesis against a bounded alternative with no unit roots, e.g., the bounded, near-I(1) model. Our framework allows us to tackle this testing problem. Specifically, despite the fact that the bounded unit root distribution (and also the asymptotic distributions of the other unit root test statistics) depends on three nuisance parameters, $(\underline{c} - c_0, \bar{c} - c_0, \sigma/\lambda)$, it is possible to define a proper rejection region to test the BI(1) hypothesis at a given significance level, hence avoiding spurious rejections caused by the presence of the bounds.

The main result needed to develop a BI(1) test is given by the following corollary of Lemma 3, which shows that the two unknown parameters $(\underline{c} - c_0, \bar{c} - c_0)$ can be consistently estimated.

COROLLARY 5. *Let the assumptions of Theorem 2 hold and suppose that Assumption $\mathcal{K}$ holds with (K2) replaced by (K2′). Then, as $T \uparrow \infty$, $\hat{\underline{c}} \xrightarrow{p} \underline{c}$, $\hat{\bar{c}} \xrightarrow{p} \bar{c}$, and $\hat{c}_0 \xrightarrow{p} c_0$, where $\hat{\underline{c}} := \underline{b}(\hat{\lambda}^2 T)^{-1/2}$, $\hat{\bar{c}} := \bar{b}(\hat{\lambda}^2 T)^{-1/2}$, and $\hat{c}_0 := X_0(\hat{\lambda}^2 T)^{-1/2}$.*

Hence, given that the bounds $(\underline{b}, \bar{b})$ are known, from the consistency of $\hat{\lambda}^2$ it follows that the two nuisance parameters of the bounded unit root distribution, $\underline{c} - c_0$ and $\bar{c} - c_0$, can be consistently estimated by $\hat{\underline{c}} - \hat{c}_0$ and $\hat{\bar{c}} - \hat{c}_0$, respectively.

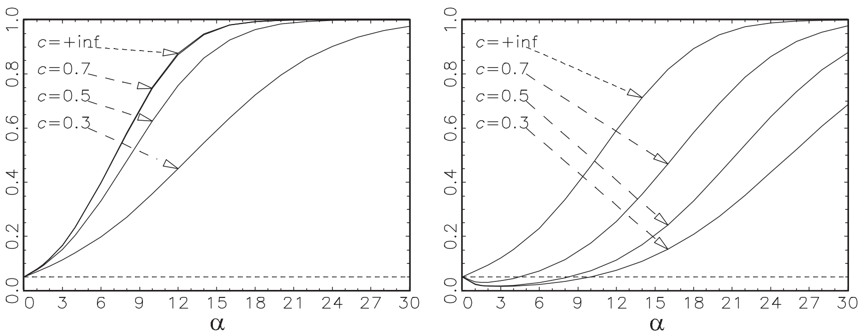
Therefore, if the DGP is a BI(1) process, the rejection frequency of the $\hat{Z}_\rho$ unit root test equalizes the selected significance level in large samples as far as the quantiles of the $\mathcal{Z}(\hat{\underline{c}} - \hat{c}_0, \hat{\bar{c}} - \hat{c}_0)$ distribution are used; such quantiles can easily be computed by MC simulation.[8] Moreover, the same procedure can be applied to all the unit root tests previously discussed. In the following discussion it will be shown that this asymptotic framework also provides excellent results in small samples. The asymptotic local power properties of the BI(1) test will be investigated beforehand.

5.1. Asymptotic Local Power

It is interesting to compare the asymptotic power function of BI(1) tests against BNI(1) with the asymptotic power function of standard unit root tests in the absence of range constraints. To analyze this issue we do not need any further theoretical result because (i) the distribution of the unit root statistics under BNI(1) dynamics has already been obtained in Section 4, Theorem 4, and (ii) Corollary 5 remains valid.

To assess to what extent the triple $\{\alpha, \underline{c}, \bar{c}\}$ affects the asymptotic power function of the “bounded unit root” tests, in Figure 7 the asymptotic power function of the BI(1) Z_ρ test ($\hat{Z}_\rho$ for $\sigma^2 \neq \lambda^2$) against BNI(1) is plotted for various values of α and $c := \bar{c} = -\underline{c}$ under $c_0 = 0$ (symmetric case). In the left panel of the figure the test is applied to raw data, whereas in the right panel the test refers to OLS-demeaned data.

As expected, tests of the BI(1) hypothesis are less powerful than standard I(1) tests; i.e., in the presence of range constraints it is more difficult to assess whether a given series has a unit root than in the usual case of no constraints. Specifically, the smaller c is, the lower the power of the test is. It is also worth noting that, in the case of symmetric bounds, OLS demeaning reduces the power of the test (for α close to zero the test has no power against BNI(1) alterna-



(a) (b)

FIGURE 7. Asymptotic power function of the "bounded unit root" Z_ρ test against bounded near-integrated alternatives for various values of $\bar{c} = -\underline{c} =: c$, $c_0 = 0$, and α ranging from 0 to 30: (a) tests based on raw data; (b) tests based on demeaned data.

tives). This evidence, however, does not necessarily hold when the bounds are either asymmetric or one-sided; in particular, when the bounds are asymmetric, the tests based on OLS demeaning can outperform the tests based either on the deviations from the initial value or on GLS demeaning. This result suggests that the practitioner should compute both OLS and GLS demeaning-based tests when the null of interest is the BI(1) hypothesis.

5.2. Finite-Sample Properties

In this section we briefly report the outcome of a set of MC simulations on the small-sample size of the bound-corrected unit root tests outlined earlier. Because a key assumption ((A4) of Definition 1) of BI(1) asymptotics is that the position of the bounds depends on T, one could reasonably argue that the small-sample accuracy of the test (at least in terms of size) might be inadequate. In the following discussion it will be shown that this is not the case.[9]

Initially, a BI(1) process $\{X_t\}$ is chosen with $\{\varepsilon_t\}$ being a Gaussian independent and identically distributed (i.i.d.) process with zero mean and unit variance. The (conditional) distribution of $\{\varepsilon_t\}$ is reflected at the bounds (see note 2); the results for different truncation mechanisms do not substantially differ. Attention is paid to the case of two symmetric bounds ($\bar{c} = -\underline{c} =: c > 0$ and $X_0 = 0$) and also to the case of a single bound ($\bar{c} = \infty$, $-\underline{c} =: c > 0$ and $X_0 = 0$). We consider the Z_ρ and Z_t unit root tests, both based on the deviations of the observed series from the initial value and from the sample average; the results for tests based on GLS demeaned variables, and also for the modified MZ tests, do not differ from those discussed here. Because $\lambda^2 = \sigma^2$, c is esti-

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

mated by $\hat{c} = \underline{b}(\hat{\sigma}^2 T)^{-1/2} = \bar{b}(\hat{\sigma}^2 T)^{-1/2}$. The critical values of the asymptotic distribution under the null hypothesis are then retrieved through a linear interpolation of the critical values obtained by simulation in Sections 3 and 4. The selected sample sizes are $T = 50, 100, 250, 500$, and the number of MC replications is 20,000.

The results are summarized in Table 3. The small-sample performance of the tests is excellent. The empirical rejection frequency exceeds the significance level only in the two-bound case, for $c = 0.3$ and $T = 50$, when the tests are based on raw data. Nevertheless, when $c = 0.3$ and $T = 100$ the rejection rate is already very close to the nominal size for most tests. For the other values of (c, T) the asymptotic approximation of the distribution of the tests considered is extremely well behaving.

We now turn to the two-bound case in the presence of autocorrelated errors.[10] Table 4 concerns the case of AR(1) errors; specifically, $\varepsilon_t = \phi\varepsilon_{t-1} + \nu_t$ where

TABLE 3. Finite-sample null rejection probabilities white noise (WN) model in the one-bound (at $-c$) and in the two-bound (at $-c$ and c) cases

		Two-sided constraints $(-c, c)$				One-sided constraints $(-c, +\infty)$			
c	T	Z_ρ	Z_t	$Z_{\rho.c}$	$Z_{t.c}$	Z_ρ	Z_t	$Z_{\rho.c}$	$Z_{t.c}$
∞	50	4.630	5.355	3.635	6.465	4.630	5.355	3.635	6.465
	100	4.545	4.760	4.185	5.345	4.545	4.760	4.185	5.345
	250	4.805	4.910	4.765	5.045	4.805	4.910	4.765	5.045
	500	4.840	4.865	4.975	5.380	4.840	4.865	4.975	5.380
0.9	50	4.630	5.360	3.840	7.400	4.630	5.355	3.780	7.015
	100	4.550	4.765	4.155	5.880	4.545	4.765	4.095	5.485
	250	4.810	4.915	4.545	5.330	4.805	4.910	4.435	5.115
	500	4.840	4.865	4.950	5.440	4.840	4.865	4.845	5.310
0.7	50	4.730	5.695	3.870	7.480	4.650	5.555	3.765	6.975
	100	4.640	4.985	4.280	6.080	4.575	4.825	4.095	5.655
	250	4.825	4.995	4.765	5.490	4.865	4.995	4.375	5.050
	500	4.905	4.850	4.895	5.300	4.870	4.815	4.775	5.230
0.5	50	5.270	7.735	3.380	6.530	4.765	6.420	3.460	6.410
	100	4.970	6.070	3.970	5.690	4.525	5.470	4.005	5.565
	250	4.800	5.330	4.335	5.070	4.950	5.220	4.320	4.835
	500	5.010	5.265	4.455	4.865	4.730	5.025	4.595	4.850
0.3	50	12.370	28.320	2.255	8.130	5.150	8.390	2.430	5.285
	100	8.015	15.315	3.350	6.165	4.685	6.630	3.345	4.755
	250	5.870	8.490	3.875	5.000	4.525	5.305	3.940	4.445
	500	5.380	6.425	4.010	4.570	4.600	5.090	4.315	4.570

Notes: Z_ρ and Z_t denote the standard coefficient and t-tests, respectively (without heteroskedasticity and autocorrelation correction); $Z_{\rho.c}$ and $Z_{t.c}$ are based on OLS-demeaned data.

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

TABLE 4. Finite-sample null rejection probabilities AR(1) model; two-bound (at $-c$ and c) case

ϕ	c	T	$\hat{Z}_\rho$	$\hat{Z}_t$	MZ_ρ^{AR}	$\hat{Z}_{\rho.c}$	$\hat{Z}_{t.c}$	$MZ_{\rho.c}^{AR}$	$\overline{MZ}_\rho^{AR}$
0.3	∞	50	2.425	2.530	2.370	1.520	3.370	1.345	4.485
		100	2.845	2.950	3.410	2.235	3.335	2.320	5.175
		250	3.745	3.755	4.410	3.280	3.680	4.050	5.260
		500	4.010	3.985	4.550	3.955	4.315	4.530	5.020
	0.7	50	2.260	2.390	1.475	1.795	4.685	0.505	3.260
		100	2.775	2.895	2.865	2.440	3.900	1.165	4.650
		250	3.610	3.615	4.095	2.990	3.895	2.835	4.870
		500	3.945	3.905	4.420	3.405	3.910	3.415	4.940
	0.5	50	1.845	2.160	0.680	2.355	4.170	0.600	1.805
		100	2.390	2.530	1.905	2.330	3.420	0.925	3.425
		250	2.895	2.975	3.160	2.555	3.010	2.135	4.050
		500	3.370	3.400	3.675	2.805	3.120	2.875	4.235
	0.3	50	1.720	2.825	0.375	1.680	2.840	0.155	1.720
		100	1.145	1.625	0.355	1.505	2.065	0.405	0.900
		250	1.480	1.670	1.240	1.830	2.040	1.010	1.765
		500	1.840	1.915	1.695	2.095	2.255	2.050	2.000
−0.3	∞	50	9.095	10.630	3.045	12.265	15.135	2.035	5.885
		100	7.730	8.410	3.150	10.915	12.090	2.435	4.785
		250	6.950	7.245	4.030	9.080	9.160	3.350	4.750
		500	6.200	6.275	4.420	7.990	7.910	4.260	4.935
	0.7	50	10.865	14.295	2.755	9.415	14.895	0.600	7.060
		100	8.690	10.240	3.190	9.615	12.790	1.515	5.445
		250	7.175	7.920	4.140	8.415	9.535	2.940	5.065
		500	6.475	6.795	4.480	7.580	8.155	3.715	5.055
	0.5	50	18.190	29.490	2.315	7.725	15.470	0.165	6.960
		100	13.785	19.520	3.555	9.760	13.870	1.020	6.675
		250	9.960	11.935	4.305	8.685	10.405	2.180	5.990
		500	8.115	9.005	5.020	7.755	8.690	2.965	5.840
	0.3	50	51.555	82.650	1.115	16.235	49.355	0.025	2.550
		100	43.055	71.245	2.815	19.175	41.825	0.525	5.110
		250	27.830	44.475	5.740	15.085	24.080	1.700	7.585
		500	19.160	27.245	7.110	11.390	14.875	2.450	8.335

Notes: $\hat{Z}_\rho$ and $\hat{Z}_t$, respectively, denote the Phillips–Perron coefficient and t-tests with long-run variance λ^2 estimated according to the Andrews (1991) SC estimator with quadratic spectral kernel and AR(1)-automatic bandwidth selection; MZ_ρ^{AR} is the modified coefficient test based on the AR estimator of the long-run variance with the number of lags k, $0 \le k \le [12\cdot(T/100)^{0.25}]$ chosen according to the MAIC criterion of Ng and Perron (2001); $.c$ denotes tests based on OLS-demeaned data; $\overline{MZ}_\rho^{AR}$ denotes the modified coefficient test based on GLS-demeaned data.

ν_t is i.i.d. $N(0,(1-\phi)^2)$ and $\phi = \pm 0.3$. Table 5 concerns the case of MA(1) errors; specifically, $\varepsilon_t = \theta\nu_{t-1} + \nu_t$ where ν_t is i.i.d. $N(0,1/(1+\theta^2))$ and $\theta = \pm 0.3$. Note that in both cases the long-run variance of $\{\varepsilon_t\}$ is equal to unity. In

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

TABLE 5. Finite sample null rejection probabilities MA(1) model; two-bound (at $-c$ and c) case

θ	c	T	$\hat{Z}_\rho$	$\hat{Z}_t$	MZ_ρ^{AR}	$\hat{Z}_{\rho.c}$	$\hat{Z}_{t.c}$	$MZ_{\rho.c}^{AR}$	$\overline{MZ}_\rho^{AR}$
0.3	∞	50	3.195	3.340	2.350	2.180	3.970	1.210	4.515
		100	3.610	3.675	3.540	2.975	3.835	2.360	5.280
		250	4.225	4.290	4.675	4.000	4.250	4.185	5.495
		500	4.475	4.430	4.615	4.500	4.755	4.595	5.115
	0.7	50	3.040	3.210	1.515	2.375	4.630	0.475	3.110
		100	3.455	3.570	2.970	2.930	4.140	1.300	4.910
		250	4.045	4.095	4.270	3.560	4.270	2.990	5.135
		500	4.465	4.355	4.460	3.850	4.330	3.580	5.025
	0.5	50	2.590	3.025	0.785	2.285	4.160	0.465	2.250
		100	3.065	3.270	2.055	2.705	3.660	1.000	3.545
		250	3.455	3.585	3.295	3.080	3.485	2.275	4.380
		500	3.960	4.015	3.845	3.300	3.545	3.025	4.475
	0.3	50	2.715	5.285	0.385	1.400	2.720	0.110	1.925
		100	2.070	3.025	0.490	1.735	2.475	0.345	1.210
		250	2.125	2.435	1.545	2.315	2.615	1.190	2.180
		500	2.455	2.770	2.115	2.625	2.800	2.280	2.580
−0.3	∞	50	12.335	13.665	3.830	17.470	19.795	3.120	7.550
		100	11.115	11.850	3.895	15.920	16.570	3.335	5.900
		250	9.290	9.720	4.785	13.140	12.585	4.290	5.730
		500	8.125	8.240	4.880	10.910	10.560	4.940	5.315
	0.7	50	15.860	21.015	3.540	10.645	17.270	0.550	9.525
		100	13.070	15.760	3.955	12.920	16.780	1.690	7.165
		250	10.405	11.345	4.840	11.340	12.855	3.235	6.075
		500	8.570	8.955	5.100	10.095	10.740	4.000	5.775
	0.5	50	26.275	42.425	3.080	8.800	19.275	0.155	7.815
		100	21.810	30.910	4.505	12.980	18.690	1.165	8.385
		250	15.225	18.880	5.550	12.345	14.685	2.325	7.675
		500	11.780	13.380	6.110	10.850	12.035	3.205	7.125
	0.3	50	62.460	90.860	0.935	23.790	65.030	0.005	2.425
		100	58.210	85.785	3.540	31.740	62.390	0.625	5.590
		250	45.995	67.405	8.160	27.915	44.605	2.190	10.485
		500	33.040	45.810	10.180	20.360	28.250	3.065	11.745

Note: See the previous table.

the autocorrelated case the issue of precise estimation of the long-run variance (and hence of c, which is now estimated by $\hat{c} = \underline{b}(\hat{\lambda}^2 T)^{-1/2} = \bar{b}(\hat{\lambda}^2 T)^{-1/2}$; see Corollary 5) becomes crucial. Hence, together with the Phillips–Perron tests based on the SC estimator of Andrews (1991)[11] of λ^2 $(\hat{Z}_\rho, \hat{Z}_t)$, we also report the results obtained by the modified coefficient test (MZ_ρ^{AR}) based on an autoregressive estimator of λ^2, $\hat{\lambda}_{AR}^2$ in the following discussion, with the number of

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

lags chosen according to the modified Akaike information criterion (MAIC) defined in Ng and Perron (2001). Finally, together with the tests based on raw data, tests based on OLS and GLS demeaning ($MZ_{\rho.c}^{AR}$ and $\overline{MZ}_{\rho.c}^{AR}$, respectively) are included in the tables.

In the AR(1) case (Table 4) with positively autocorrelated errors, the simulation evidence is comparable to that obtained in the white noise case, although the tests tend to be slightly conservative when the bounds are very close to each other ($c = 0.3$). The various tests have almost the same size. When the errors are negatively autocorrelated, the tests based on $\hat{\lambda}_{AR}^2$ perform better than the tests based on the SC estimator (as in the standard I(1) framework, $c = \infty$ in the table). The MA(1) case (Table 5) does not significantly differ from the AR(1) case when the errors are positively correlated, but size distortions are more pronounced when the errors are negatively correlated and the SC estimator of λ^2 is employed. On the contrary, the tests based on $\hat{\lambda}_{AR}^2$ behave very well also in the presence of negatively correlated MA errors. In general, when the errors are autocorrelated, tests employing an autoregressive estimator of the long-run variance deliver the best results in terms of size accuracy.

Overall, the small-sample performance of the asymptotic approximation combined with the use of appropriate empirical estimates of the bound parameters seems to be adequate, and the small-sample size of the BI(1) test appears to be largely satisfactory.

6. EMPIRICAL ILLUSTRATIONS

In this section, we briefly discuss two common applications of unit root tests to limited time series, namely, testing for exchange rate mean reversion within a target zone and testing for a unit root in the unemployment rate.

6.1. European Monetary System Target Zone Exchange Rates

We start by examining an empirical problem that has often been tackled in the literature, i.e., testing for exchange rate mean reversion in the presence of a target zone. The reader can refer to Svensson (1993) and Anthony and MacDonald (1998). Economic theories of target zone exchange rates usually associate the rejection of the unit root hypothesis with the presence of either intramarginal Central Bank interventions or mean reverting fundamentals (Delgado and Dumas, 1992). However, as noticed by Svensson (1993), the presence of the target zone can be the source of mean reversion of the exchange rate. In this framework, we will briefly show how the outcome of unit root tests in the presence of (target zone–) range constraints can lead to wrong economic conclusions and how the researcher can properly modify the conventional test procedures to take the target zone into account.

The exchange rates of four currencies are considered, namely, the Danish krone (DK), the French franc (FF), the Irish pound (IP), and the Dutch guilder (NG), all against the Deutsche mark (DM);[12] see Figure 8. The exchange rates have been transformed by taking logs and multiplying by 100; the observation frequency is daily. The selected sample starts on 87:01:12 and ends on 93:01:29. During such a period, all the (bilateral) exchange rates considered were bounded within $\pm 2.25\%$ target zones, which had not been realigned.

In Table 6 the (constant-corrected) Phillips–Perron coefficient and *t*-tests ($\hat{Z}_{\rho.c}$ and $\hat{Z}_{t.c}$) are reported; the long-run variance λ^2 is estimated according to the SC estimator of Andrews (1991) with quadratic spectral kernel and AR(1)-automatic bandwidth selection; see the previous section. Together with the Phillips–Perron tests, we also show the results obtained by using the modified coefficient test, $MZ_{\rho.c}^{AR}$, which employs an autoregressive estimator of the long-run variance with the number of lags k chosen by both the MAIC criterion of Ng and Perron (2001) and the Bayesian information criterion (BIC) ($MZ_{\rho.c}^{AR}$ and $MZ_{\rho.c}^{bic}$, respectively) under the constraint $k \leq [12(T/100)^{1/4}]$.

Evidence against the I(1) hypothesis is generally found for the French franc, the Irish pound, and the Dutch guilder at the 5% significance level and for the Danish krone at the 10% significance level. Therefore, the researcher might erroneously conclude that the four European monetary system exchange rates considered do not have a unit root and that they are mean reverting *within* the band.

However, because there are nominal bounds, the relevant question is whether the rejection of the I(1) hypothesis depends on the presence of the target zone (i.e., the DGP is a bounded process with a unit root, BI(1)) or whether it should be interpreted as evidence of mean reversion *within* the band (i.e., the DGP is a bounded process with no unit roots, BNI(1)). Standard unit root analysis does not provide an answer to this question. However, an answer can be given by employing the BI(1) test introduced in Section 5. In the second half of Table 6 the estimates of the bound parameters ($\bar{c} - c_0, \underline{c} - c_0$) are reported, in addition to the asymptotic 5% and 10% quantiles associated to the corresponding test statistics under the "bounded unit root" null hypothesis. For the DM/IP and the DM/NG exchange rates, almost all tests reject the null of a unit root even when bound-corrected critical values are employed. Therefore, for these two exchange rates there is evidence of mean reversion that cannot be attributed to the presence of the target zone alone (i.e., the DM/IP and the DM/NG exchange rates are mean reverting *within* the band).

The results for the DM/DK exchange rate and for the DM/FF exchange rates are opposite. Contrary to the results obtained when standard critical values are employed, the "bounded unit root" tests do not substantially lead to the rejection of the hypothesis of a unit root: the observed mean reversion can be explained by the presence of the target zone alone (i.e., the DM/DK and the DM/FF exchange rates are *not* mean reverting within the band).

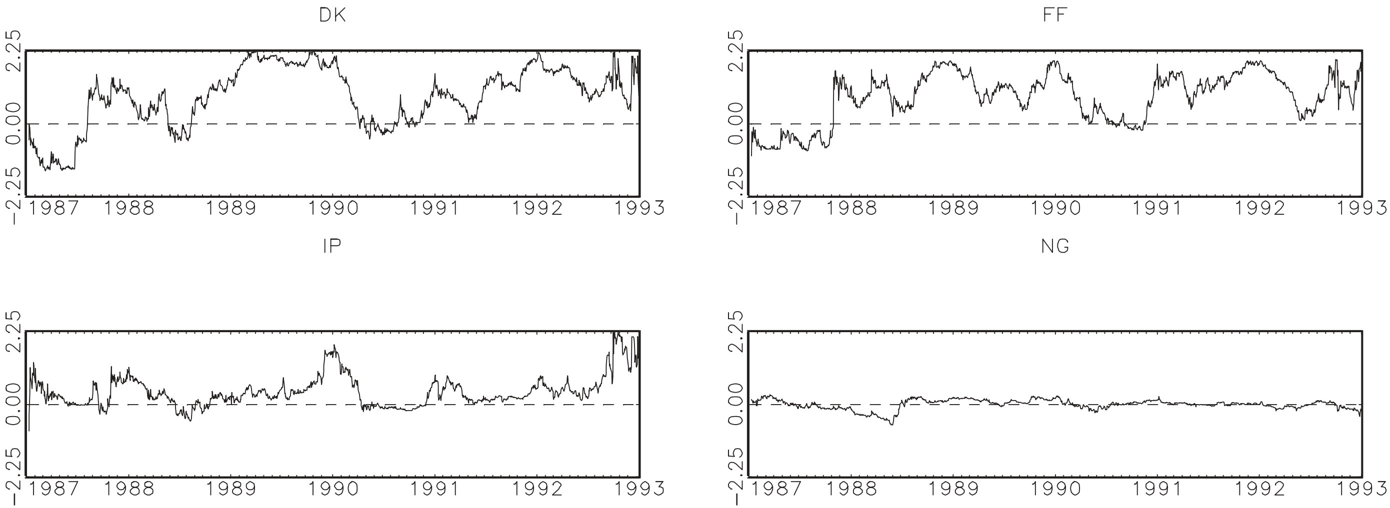


FIGURE 8. European monetary system exchange rates of the Danish krone (DK), the French franc (FF), the Irish pound (IP), and the Dutch guilder (NG), all against the Deutsche mark, 1987:01–1993:01. Percentage deviations from the central parity.

TABLE 6. European monetary system exchange rates: Unit root tests, estimates of the bound parameters, and bound-corrected critical values

	DK				FF			
	$\hat{Z}_{\rho.c}$	$\hat{Z}_{t.c}$	$MZ^{AR}_{\rho.c}$	$MZ^{bic}_{\rho.c}$	$\hat{Z}_{\rho.c}$	$\hat{Z}_{t.c}$	MZ^{AR}_{ρ}	$MZ^{bic}_{\rho.c}$
	-12.918^{b}	-2.655^{b}	-9.887	-12.960^{b}	-17.153^{a}	-3.230^{a}	-17.623^{a}	-17.623^{a}
$\hat{\lambda}$	0.113	0.113	0.098	0.103	0.101	0.101	0.103	0.103
$\underline{\hat{c}} - \hat{c}_0$	-0.401	-0.401	-0.463	-0.399	-0.322	-0.322	-0.315	-0.315
$\hat{\bar{c}} - \hat{c}_0$	0.615	0.615	0.711	0.612	0.809	0.809	0.793	0.793
$c.v._{0.05}$	-25.353	-3.635	-23.442	-25.401	-25.016	-3.593	-25.224	-3.608
$c.v._{0.10}$	-20.324	-3.229	-18.834	-20.388	-19.718	-3.196	-19.902	-19.902
	IP				NG			
	$\hat{Z}_{\rho.c}$	$\hat{Z}_{t.c}$	$MZ^{AR}_{\rho.c}$	$MZ^{bic}_{\rho.c}$	$\hat{Z}_{\rho.c}$	$\hat{Z}_{t.c}$	MZ^{AR}_{ρ}	$MZ^{bic}_{\rho.c}$
	-34.424^{a}	-4.023^{a}	-12.958^{b}	-34.385^{a}	-20.546^{a}	-3.238^{a}	-16.177^{a}	-21.872^{a}
$\hat{\lambda}$	0.109	0.109	0.071	0.110	0.025	0.025	0.022	0.026
$\underline{\hat{c}} - \hat{c}_0$	-0.341	-0.341	-0.524	-0.339	-2.470	-2.470	-2.778	-2.383
$\hat{\bar{c}} - \hat{c}_0$	0.708	0.708	1.088	0.705	2.105	2.105	2.368	2.031
$c.v._{0.05}$	-25.340	-3.630	-20.661	-25.372	-14.193	-2.879	-14.193	-14.193
$c.v._{0.10}$	-20.210	-3.217	-16.479	-20.229	-11.354	-2.583	-11.351	-11.354

Notes: All exchange rates are against the Deutsche mark. Standard 5% (10%) critical values are -14.10 (-11.35) for the $\hat{Z}_{\rho.c}$ and $MZ_{\rho.c}$ tests, -2.86 (-2.57) for the $\hat{Z}_{t.c}$ test; ^{a} (^{b}) denotes a statistic significant at the 5% (10%) significance level when standard I(1) critical values are used. Critical values for the BI(1) hypothesis have been obtained through MC simulation by discretizing the limiting regulated Brownian motion over $T = 20{,}000$ segments and using 50,000 replications.

6.2. U.S. Unemployment Rate

In this section we analyze the monthly U.S. unemployment rate among adult males from January 1948 through August 1999; see Figure 9. These data have recently been analyzed by Caner and Hansen (2001) by means of threshold autoregression methods; the reader can refer to their paper for further details on data definition.

Our first question is whether the I(1) hypothesis is rejected over the considered sample. By construction, the unemployment rate is bounded, and therefore the I(1) specification should not provide an adequate representation of the data. We explore this issue by referring to the tests previously discussed; we also consider the rescaled range unit root test based on the statistic $\hat{r}_\mu := (\hat{\lambda}^2 T)^{-1/2} \times (\max_{t=1,\dots,T} X_t - \min_{t=1,\dots,T} X_t)$, which has been applied to U.S. unemployment by Cavaliere (2001). For the $\hat{Z}_{\rho.c}$, $\hat{Z}_{t.c}$, and $\hat{r}_\mu$ tests, the long-run variance λ^2 is estimated through the Andrews (1991) quadratic spectral kernel SC estimator based on first-order autoregression residuals with 12 lags to take account of the peak at the 12-month frequency in the autocorrelation function of the unemployment rate changes. For the modified coefficient test, $MZ^{AR}_{\rho.c}$, λ^2 is

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

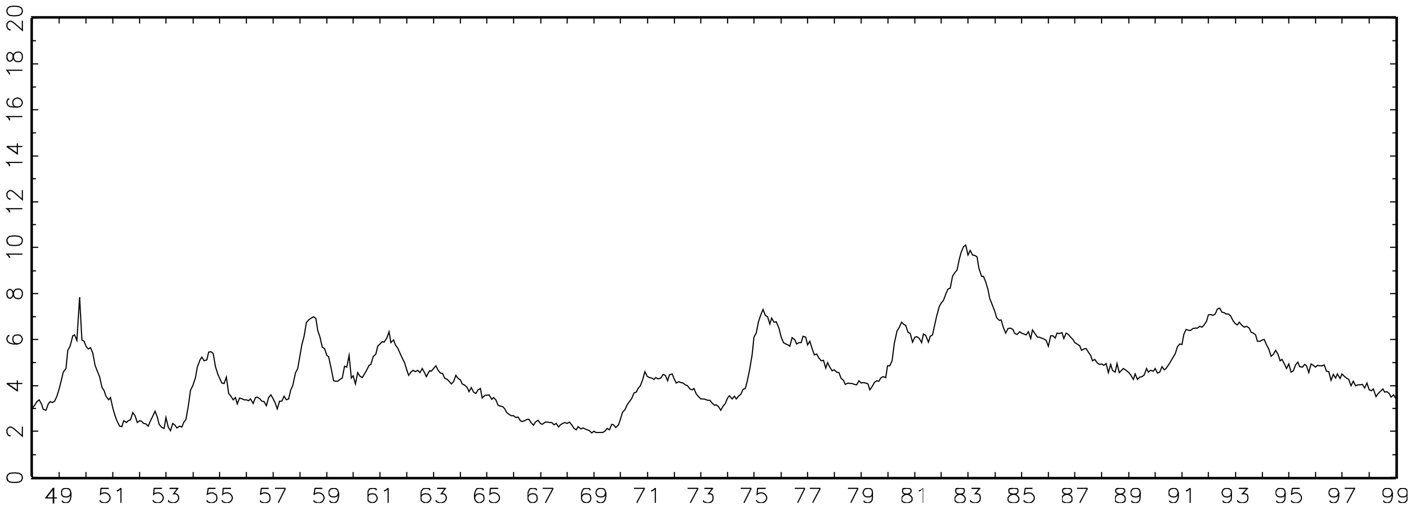


FIGURE 9. Monthly U.S. percentage unemployment rate among adult males from January 1948 through August 1999.

estimated by the autoregressive estimator $\hat{\lambda}^2_{AR}$ with number of lags k equal to 12, which corresponds to the lag length selected by means of the BIC and the MAIC criteria; $k = 12$ is also the lag length selected by Caner and Hansen (2001). The range statistic based on $\hat{\lambda}^2_{AR}$, $\hat{r}^{AR}_{\mu}$, is also reported.

The results are summarized in Table 7, first row. All tests suggest rejecting the I(1) model at the 5% significance level, with the $\hat{r}_\mu$ test rejecting at the 1% level (the standard 1% critical value of the $\hat{r}_\mu$ test is 0.833). The researcher should therefore investigate whether such a rejection depends on the existence of an upper and a lower bound or if it can be attributed to the absence of a unit root. In the third, fourth, and fifth rows of Table 7 the estimates of λ^2, $\underline{c} - c_0$, and $\bar{c} - c_0$ are reported. Because $\underline{c} - c_0$ is between -0.32 and -0.33 whereas $\bar{c} - c_0$ is between 9.86 and 10.28, the presence of the upper barrier at $\bar{b} = 100$ seems to be negligible. Conversely, according to Figure 5, a lower barrier at -0.32 inflates the asymptotic rejection frequency of the (5% nominal level) OLS-demeaned tests to 18% for $\hat{Z}_{\rho.c}/MZ^{AR}_{\rho.c}$ and to 13% for $\hat{Z}_{t.c}$; similarly, the asymptotic rejection frequency of $\hat{r}_\mu$ is found to be about 30%. Therefore, the outcome of unit root tests is likely to be affected by the presence of the lower bound.

Bound-corrected critical values can easily be obtained by referring to the criteria outlined in Section 5. The estimates of the 5% and 10% critical values for testing the BI(1) hypothesis are presented in the last two rows of Table 7. Now, two tests reject the BI(1) null hypothesis at the 10% level, with the remaining tests leading to the maintenance of the null hypothesis even at the 10% level. Hence, once one has properly taken account of the bounds, the evidence against the unit root hypothesis in the postwar U.S. male unemployment rate is much weaker than the evidence obtained without accounting for the bounds.

TABLE 7. U.S. male unemployment rate: Unit root tests, estimates of the bound parameters, and bound-corrected critical values

	$\hat{Z}_{\rho.c}$	$\hat{Z}_{t.c}$	$\hat{r}_\mu$	$MZ^{AR}_{\rho.c}$	$\hat{r}^{AR}_\mu$
	-19.172^a	-3.115^a	0.827^a	-17.342^a	0.869^a
λ^2	0.158	0.158	0.158	0.143	0.143
$\underline{c} - c_0$	−0.317	−0.317	−0.317	−0.334	−0.334
$\bar{c} - c_0$	9.864	9.864	9.864	10.282	10.282
$c.v._{0.05}$	−24.567	−3.494	0.648	−24.200	0.651
$c.v._{0.10}$	−18.927	−3.081	0.728	−18.710	0.732

Notes: Standard 5% (10%) critical values are −14.10 (−11.35) for the $\hat{Z}_{\rho.c}$ and $MZ_{\rho.c}$ tests, −2.86 (−2.57) for the $\hat{Z}_{t.c}$ test, and 0.975 (1.063) for the $\hat{r}_\mu$ test; [a] ([b]) denotes a statistic significant at the 5% (10%) significance level when standard I(1) critical values are used. Critical values for the BI(1) hypothesis have been obtained through MC simulation by discretizing the limiting regulated Brownian motion over $T = 20{,}000$ segments and using 50,000 replications.

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

7. EXTENSIONS AND CONCLUDING REMARKS

This paper shows how the presence of range constraints affects the asymptotic distribution of unit root tests. Testing for unit roots in bounded time series should always be carried out with caution because, when the unit root hypothesis is rejected, the range constraints can be the leading cause of the rejection. The approach suggested in the paper provides a way to assess the role of range constraints, and it can be implemented easily. It allows a quick evaluation of the relevance of the bounds and also allows the researcher to test statistically whether a given limited time series reverts because of the presence of the bounds alone (the "bounded unit root" hypothesis) or because it does not have a unit root (the "bounded, near-unit root" hypothesis). Moreover, the proposed asymptotic framework provides an adequate approximation of the finite-sample properties of unit root tests under range constraints.

The asymptotics and the finite-sample methods discussed in this paper can be extended to a multivariate framework. Specifically, it can be shown that the stronger the range constraints on the data, the higher the probability that cointegration tests will point toward (spurious) cointegration.

Finally, it is worth noting that the asymptotics obtained also provide a basis for (asymptotic) power comparisons when the researcher wants to test whether a given time series with integrated behavior is bounded by unobservable bounds. For example, in the context of floating nominal exchange rates one might be interested in testing whether a given bilateral exchange rate is regulated within an undeclared target zone (see Nicolau, 2002, and references therein). The asymptotics of Section 4 allow us to understand which tests are preferable in terms of power when the standard I(1) hypothesis is tested against the bounded I(1) alternative.

NOTES

1. In the following discussion, $\mathcal{C} := C[0,1]$ denotes the space of all continuous, real valued functions on $[0,1]$, whereas $\mathcal{D} := D[0,1]$ denotes the space of all real valued "cadlag" functions on $[0,1]$, i.e., real valued functions that are right continuous at each point of $[0,1)$ with left limit existing at each point of $(0,1]$. Finally, $\xrightarrow{w}$ and $\xrightarrow{p}$ denote weak convergence and convergence in probability, respectively.

2. Consider for ease of notation the lower bound case only, and suppose that a r.v. u is censored at $\underline{b}$ and that its cumulative distribution function (c.d.f.) has the form $F_u(x) = G(x)\mathbb{I}\{x \geq \underline{b}\}$, where $G(\cdot)$ is an uncensored c.d.f. The same distribution can be obtained by taking $u := \varepsilon + \underline{\xi}$, where ε is a r.v. with c.d.f. $G(\cdot)$ on all the real set and $\underline{\xi} := (\underline{b} - \varepsilon)\mathbb{I}\{\varepsilon < \underline{b}\}$. Truncation at $\underline{b}$, i.e., $F_u(x) = [1 - G(\underline{b})]^{-1}[G(x) - G(\underline{b})]\mathbb{I}\{x \geq \underline{b}\}$ can be obtained by simply taking ε as a r.v. with c.d.f. $G(\cdot)$ and by taking $\underline{\xi}$ as a r.v. such that $F_{\underline{\xi}|\varepsilon<\underline{b}}(x) = F_u(x - \varepsilon)$ and $F_{\underline{\xi}|\varepsilon\geq\underline{b}}(x) = \mathbb{I}\{x \geq 0\}$. Reflection at $\underline{b}$, i.e., $F_u(x) = [G(x) - G(2\underline{b} - x)]\mathbb{I}\{x \geq \underline{b}\}$ is given by taking ε with c.d.f. $G(\cdot)$ and defining $\underline{\xi} := (2\underline{b} - x)\mathbb{I}\{\varepsilon < \underline{b}\}$.

3. A formal proof is provided in the Appendix, Proof of Theorem 1.

4. See, e.g., Assumption (B1) in this section.

5. It is worth noting that, although this algorithm allows exact simulation of the regulated Brownian motion over a discrete grid, simulation of functionals such as $\int_0^1 g(B_{\underline{c}}^{\bar{c}}(s))\,ds$ requires the

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

discretized time increment $1/N$ to be extremely fine to obtain an accurate assessment of the limiting distribution (see Asmussen, Glynn, and Pitman, 1995).

6. Quantiles of the bounded unit root distribution (and also the quantiles of the various statistics discussed in the paper) have been estimated over a grid of values for $c = 0.20, 0.21, \ldots, 1.00, 1.02, \ldots, 2.00$.

7. Consistency follows, e.g., from Theorem 3 of Hansen (1992).

8. A selection of 5% quantiles and a GAUSS program for simulating the asymptotic quantiles of unit root tests for any choice of $(\underline{c}, \bar{c})$ are available from the web page http://www2.stat.unibo.it/cavaliere/rconstr/.

9. We do not report the results related to the power properties of the tests in finite samples because the asymptotic local power analysis provides a good description of the finite-sample performance, with all tests approaching the corresponding asymptotic local power function relatively fast as T increases.

10. When the errors are autocorrelated and there is one bound only, the size performance of the tests (not reported) is very close to the size performance in the no-bound case ($c = \infty$).

11. Specifically, the SC estimator is based on the OLS residuals obtained by regressing X_t on X_{t-1} (on $(X_{t-1}{:}1)'$ for the constant-corrected tests) and employs a quadratic spectral kernel with bandwidth parameter chosen according to the Andrews (1991) automatic data-dependent procedure using the plug-in method based on an AR(1) model fit to the data. Note that the quadratic spectral kernel satisfies Assumption (K1); cf. Andrews (1991, p. 837) and Jansson (2002, p. 1450).

12. The data were obtained from Ecu rates extracted from the Bank of International Settlements (BIS) database. All exchange rates are spot Ecu rates recorded at a daily Central Bank telephone conference at 2.30 p.m. Swiss time. The bilateral exchange rates have been calculated from these Ecu rates.

REFERENCES

Abadir, K.M. & A.M.R. Taylor (1999) On the definition of (co-)integration. *Journal of Time Series Analysis* 20, 129–137.

Andrews, D.W.K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–858.

Andrews, D.W.K. & C.J. McDermott (1995) Nonlinear econometric models with deterministically trending variables. *Review of Economic Studies* 62, 343–360.

Anthony, M. & R. MacDonald (1998) On the mean-reverting properties of target zone exchange rates: Some evidence from the ERM. *European Economic Review* 42, 1493–1523.

Asmussen, S., P. Glynn, & J. Pitman (1995) Discretization error in simulation of one-dimensional reflecting Brownian motion. *Annals of Applied Probability* 5, 875–896.

Barr, D.G. & K. Cuthbertson (1991) Neo-classical consumer demand theory and the demand for money. *Economic Journal* 101, 855–876.

Bec, F. & A. Rahbek (2004) Vector equilibrium correction models with non-linear discontinuous adjustments. *Econometrics Journal* 7, 628–651.

Billingsley, P. (1968) *Convergence of Probability Measures*. Wiley.

Blanchard, O.J. & L.H. Summers (1986) Hysteresis and European unemployment. In S. Fischer (ed.), *NBER Macroeconomics Annual*, pp. 15–77. MIT Press.

Caner, M. & B.E. Hansen (2001) Threshold autoregression with a unit root. *Econometrica* 69, 1555–1596.

Cavaliere, G. (2001) Testing the unit root hypothesis using generalized rescaled range statistics. *Econometrics Journal* 4, 70–88.

Cavaliere, G. (2003) The asymptotic distribution of the Dickey-Fuller statistic under a non-negativity constraint. *Econometric Theory*, Problem, 19, 691–692; Solution, 20, 808–810.

Cox, D.R. & H.D. Miller (1965) *The Theory of Stochastic Processes*. Chapman and Hall.

Davidson, J. (2002) Establishing conditions for the functional central limit theorem in nonlinear and semiparametric time series processes. *Journal of Econometrics* 106, 243–269.

De Jong, R.J. & J. Davidson (2000) Consistency of kernel estimators of heteroskedastic and autocorrelated covariance matrices. *Econometrica* 68, 407–424.

Delgado, F. & B. Dumas (1992) Target zones, broad and narrow. In P. Krugman & M. Miller (eds.), *Exchange Rate Targets and Currency Bands*, pp. 35–56. Cambridge University Press.

Dixit, A. (1993) *The Art of Smooth Pasting*. Harwood Academic Publishers.

Elliott, G., T.J. Rothenberg, & J.H. Stock (1996) Efficient tests for an autoregressive unit root. *Econometrica* 64, 813–836.

Hansen, B.E. (1992) Consistent covariance matrix estimation for dependent heterogeneous processes. *Econometrica* 60, 967–972.

Harrison, M.J. (1985) *Brownian Motion and Stochastic Flow Systems*. Wiley.

Jansson, M. (2002) Consistent covariance matrix estimation for linear processes. *Econometric Theory* 18, 1449–1459.

Karatzas, I. & S.E. Shreve (1988) *Brownian Motion and Stochastic Calculus*. Springer-Verlag.

Lindbeck, A. & D.J. Snower (1989) Macroeconomic policy and insider power. *American Economic Review* 79, 370–376.

Nelson, C.R. & C.I. Plosser (1982) Trends and random walks in macroeconomic time series. *Journal of Monetary Economics* 10, 139–162.

Ng, S. & P. Perron (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.

Phillips, P.C.B. (1987a) Time series regression with a unit root. *Econometrica* 55, 277–301.

Phillips, P.C.B. (1987b) Toward a unified asymptotic theory for autoregression. *Biometrika* 74, 535–547.

Phillips, P.C.B. (2001) Descriptive econometrics for non-stationary time series with empirical illustrations. *Journal of Applied Econometrics* 16, 389–413.

Phillips, P.C.B. & S. Ouliaris (1990) Asymptotic properties of residual based tests for cointegration. *Econometrica* 58, 165–193.

Phillips, P.C.B. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.

Pollard, D. (1990) *Empirical Processes: Theory and Applications*. NSF-CBMS Regional Conference Series in Probability and Statistics, vol. 2. Hayward: Institute of Mathematical Statistics; Alexandria: American Statistical Association.

Saikkonen, P. & I. Choi (2004) Cointegrating smooth transition regressions. *Econometric Theory* 20, 301–340.

Sargan, J.D. & A. Bhargava (1983) Testing residuals from least squares regression for being generated by the Gaussian random walk. *Econometrica* 51, 153–174.

Svensson, L.E.O. (1993) Assessing target zone credibility: Mean reversion and devaluation expectations in the ERM, 1979–1992. *European Economic Review* 37, 763–802.

APPENDIX

Proof of Theorem 1. Consider the process $\{\widetilde{X}_t\}$ defined recursively as

$$\widetilde{X}_t = \begin{cases} \widetilde{X}_{t-1} + \varepsilon_t & \text{if } \widetilde{X}_{t-1} + \varepsilon_t \in [\underline{b}, \bar{b}] \\ \bar{b} & \text{if } \widetilde{X}_{t-1} + \varepsilon_t > \bar{b} \\ \underline{b} & \text{if } \widetilde{X}_{t-1} + \varepsilon_t < \underline{b} \end{cases} \quad \textbf{(A.1)}$$

with initial condition $\widetilde{X}_0 = X_0$. By setting $\widetilde{X}_T(\cdot) := (\lambda^2 T)^{-1/2}\widetilde{X}_{[\cdot T]}$ the following theorem holds.

THEOREM 6. *Under the conditions of Theorem 1, as* $T \uparrow \infty$ $\widetilde{X}_T(\cdot) - \widetilde{X}_T(0) \xrightarrow{w} B_{\underline{c}-c_0}^{\bar{c}-c_0}(\cdot)$, *where* $B_{\underline{c}-c_0}^{\bar{c}-c_0}$ *is a regulated Brownian motion with bounds at* $\underline{c} - c_0, \bar{c} - c_0$.

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

Proof. The proof consists of two steps. First, we define a continuous approximant of $\widetilde{X}_T(\cdot)$ that satisfies Harrison's construction of the regulated Brownian motion, and weak convergence is proved. Then, it is shown that weak convergence holds for $\widetilde{X}_T(\cdot)$ too.

The process $\{\widetilde{X}_t\}$ can be recursively defined as $\widetilde{X}_t = \widetilde{X}_0 + S_t + L_t - V_t$, where $S_t := \sum_{i=1}^t \varepsilon_i$, $L_t := \sum_{i=1}^t l_i$, $V_t := \sum_{i=1}^t v_i$, $l_t := -(\widetilde{X}_{t-1} + \varepsilon_t)\mathbb{I}\{\widetilde{X}_{t-1} + \varepsilon_t < \underline{c}\lambda T^{1/2}\}$, and $v_t := (\widetilde{X}_{t-1} + \varepsilon_t - \bar{c}\lambda T^{1/2})\mathbb{I}\{\widetilde{X}_{t-1} + \varepsilon_t > \bar{c}\lambda T^{1/2}\}$. Obviously, $\widetilde{X}_t \in [\underline{c}\lambda T^{1/2}, \bar{c}\lambda T^{1/2}]$, all t. To define a continuous approximant of $\widetilde{X}_t$, say, $\widetilde{X}_T^\star(\cdot)$, let us define continuous approximants for all its normalized components, i.e., $(\lambda^2 T)^{-1/2} S_t$, $(\lambda^2 T)^{-1/2} L_t$, and $(\lambda^2 T)^{-1/2} V_t$. For the partial sum S_t we can set

$$S_T^\star(s) := \frac{1}{\lambda T^{1/2}} \sum_{t=1}^{[sT]} \varepsilon_t + \varepsilon_{[sT]+1} \frac{(sT - [sT])}{\lambda T^{1/2}} = S_T(s) + \varepsilon_{[sT]+1} \frac{(sT - [sT])}{\lambda T^{1/2}}, \qquad \textbf{(A.2)}$$

$$S_T^\star(1) := \frac{1}{\lambda T^{1/2}} \sum_{t=1}^{T} \varepsilon_t = S_T(1),$$

which represents the process obtained by joining the points $(t/T, (\lambda^2 T)^{-1/2} S_t)$ by means of straight lines. For V_t and L_t we define this approximation in a slightly different way:

$$V_T^\star(s) := \begin{cases} \dfrac{1}{\lambda T^{1/2}} \displaystyle\sum_{t=1}^{[sT]} v_t & \text{if } v_{[sT]+1} = 0 \\ \dfrac{1}{\lambda T^{1/2}} \displaystyle\sum_{t=1}^{[sT]} v_t + \dfrac{1}{\lambda T^{1/2}} \dfrac{sT - [sT] - \dfrac{\varepsilon_{[sT]+1} - v_{[sT]+1}}{\varepsilon_{[sT]+1}}}{1 - \dfrac{\varepsilon_{[sT]+1} - v_{[sT]+1}}{\varepsilon_{[sT]+1}}} v_{[sT]+1} & \\ \qquad \times \mathbb{I}\left\{ sT \geq [sT] + \dfrac{\varepsilon_{[sT]+1} - v_{[sT]+1}}{\varepsilon_{[sT]+1}} \right\} & \text{if } v_{[sT]+1} > 0 \end{cases},$$

$$V_T^\star(1) := \frac{1}{\lambda T^{1/2}} \sum_{t=1}^{T} v_t,$$

and

$$L_T^\star(s) := \begin{cases} \dfrac{1}{\lambda T^{1/2}} \displaystyle\sum_{t=1}^{[sT]} l_t & \text{if } l_{[sT]+1} = 0 \\ \dfrac{1}{\lambda T^{1/2}} \displaystyle\sum_{t=1}^{[sT]} l_t + \dfrac{1}{\lambda T^{1/2}} \dfrac{sT - [sT] - \dfrac{\varepsilon_{[sT]+1} + l_{[sT]+1}}{\varepsilon_{[sT]+1}}}{1 - \dfrac{\varepsilon_{[sT]+1} + l_{[sT]+1}}{\varepsilon_{[sT]+1}}} l_{[sT]+1} & \\ \qquad \times \mathbb{I}\left\{ sT \geq [sT] + \dfrac{\varepsilon_{[sT]+1} + l_{[sT]+1}}{\varepsilon_{[sT]+1}} \right\} & \text{if } l_{[sT]+1} > 0 \end{cases},$$

$$L_T^\star(s) := \frac{1}{\lambda T^{1/2}} \sum_{t=1}^{T} l_t.$$

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

With respect to linear interpolations like (A.2), this construction still interpolates $(\lambda^2 T)^{-1/2} X_t$ but also satisfies the following properties:

(1) $L_T^\star(\cdot)$ and $V_T^\star(\cdot)$ are increasing and continuous with $L_T^\star(0) = V_T^\star(0) = 0$;
(2) $\widetilde{X}_T^\star(s) = c_0 + S_T^\star(s) + L_T^\star(s) - V_T^\star(s) \in [\underline{c}, \bar{c}]$, all $s \in [0,1]$;
(3) $L_T^\star(\cdot)$ and $V_T^\star(\cdot)$ increase only when $\widetilde{X}_T^\star(\cdot) = \underline{c}$ and $\widetilde{X}_T^\star(\cdot) = \bar{c}$, respectively.

From Harrison (1985, Prop. 2.6), the continuous mapping $\widetilde{X}_T^\star(\cdot) = g_{\underline{c}}^{\bar{c}}(c_0 + S_T^\star(\cdot)) = c_0 + S_T^\star(\cdot) + L_T^\star(\cdot) - V_T^\star(\cdot)$ is the *unique* functional that regulates $c_0 + S_T^\star(\cdot)$ to lie within the interval $[\underline{c}, \bar{c}]$ and that satisfies properties 1–3. This allows us to obtain the limiting distribution of $\widetilde{X}_T^\star(\cdot)$ by applying the continuous mapping theorem (CMT) (see Billingsley, 1968) to the limit of $c_0 + S_T^\star(\cdot)$. Specifically, if we prove that $S_T^\star(\cdot) \xrightarrow{w} B(\cdot)$ under Assumption (A2), then the CMT would imply that $\widetilde{X}_T^\star(\cdot) = g_{\underline{c}}^{\bar{c}}(c_0 + S_T^\star(\cdot)) \xrightarrow{w} g_{\underline{c}}^{\bar{c}}(c_0 + B(\cdot))$, which is a Brownian motion with initial value c_0, regulated at $\underline{c}$ and at $\bar{c}$. To show that $S_T^\star(\cdot) \xrightarrow{w} B(\cdot)$, it suffices to consider the following result:

$$\sup_{s\in[0,1]} |S_T^\star(s) - S_T(s)| \le \frac{1}{\lambda T^{1/2}} \max_{t=1,\dots,T} |\varepsilon_t| = o_p(1) \tag{A.3}$$

under Assumption (A2), so that from Billingsley (1968, Theorem 4.1), $S_T(\cdot) \xrightarrow{w} B(\cdot)$ implies $S_T^\star(\cdot) \xrightarrow{w} B(\cdot)$. To see that the last equality in (A.3) holds, note that because weak convergence on a compact space implies stochastic equicontinuity (see, e.g., Pollard, 1990), we have for all $\varepsilon > 0$,

$$\limsup_{\delta\to 0} \limsup_{T\to\infty} \Pr\left\{ \sup_{s\in[0,1]} \sup_{s'\in[0,1]:|s'-s|<\delta} |S_T(s) - S_T(s')| > \varepsilon \right\} = 0. \tag{A.4}$$

Now, $\max_{t=1,\dots,T}|\varepsilon_t| = \lambda T^{1/2} \max_{t=1,\dots,T}|S_T(t/T) - S_T((t-1)/T)|$, and hence

$$\Pr\left\{ \max_{t=1,\dots,T} |\varepsilon_t| > \varepsilon\lambda T^{1/2} \right\} = \Pr\left\{ \max_{t=1,\dots,T} |S_T(t/T) - S_T((t-1)/T)| > \varepsilon \right\}$$
$$\le \Pr\left\{ \sup_{s\in[0,1]} \sup_{s'\in[0,1]:|s'-s|<T^{-1}} |S_T(s) - S_T(s')| > \varepsilon \right\},$$

which converges to 0 as T diverges; see (A.4).

To prove weak convergence of $\widetilde{X}_T(\cdot)$ it is sufficient to prove that the process

$$\widetilde{X}_T^\star(s) - \widetilde{X}_T(s) = \left(L_T^\star(s) - \frac{L_{[sT]}}{\lambda T^{1/2}} \right) - \left(V_T^\star(s) - \frac{V_{[sT]}}{\lambda T^{1/2}} \right) + \frac{1}{\lambda T^{1/2}} \varepsilon_{[sT]+1}(sT - [sT]) \tag{A.5}$$

converges to 0 uniformly in probability. Because both $|L_T^\star(s) - (\lambda^2 T)^{-1/2} L_{[sT]}|$ and $|V_T^\star(s) - (\lambda^2 T)^{-1/2} V_{[sT]}|$ are smaller than $(\lambda^2 T)^{-1/2}|\varepsilon_{[sT]+1}|$, and because the set of increasing points of $L_T^\star(\cdot)$ and the set of increasing points of $V_T^\star(\cdot)$ are disjoint, it follows that

$$|\widetilde{X}_T^\star(s) - \widetilde{X}_T(s)| \le \frac{2}{\lambda T^{1/2}} |\varepsilon_{[sT]+1}|(sT - [sT]) \le \frac{2}{\lambda T^{1/2}} |\varepsilon_{[sT]+1}|; \tag{A.6}$$

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

therefore, $\sup_{s\in[0,1]}|\widetilde{X}_T^\star(s) - \widetilde{X}_T(s)| \le 2(\lambda^2 T)^{-1/2}\max_{t=1,\dots,T}|\varepsilon_t|$, which is of $o_p(1)$ (see equation (A.3)), and hence the weak convergence of $\widetilde{X}_T(\cdot)$ follows. The proof is completed by noting that because $\widetilde{X}_T(\cdot)$ converges weakly to a Brownian motion starting at c_0 and regulated at $\underline{c},\bar{c}$, the convergence of $\widetilde{X}_T(\cdot) - \widetilde{X}_T(0) = \widetilde{X}_T(\cdot) - c_0$ to a standard Brownian motion, regulated at $\underline{c} - c_0, \bar{c} - c_0$, also follows. ■

To complete the proof of Theorem 1 it is sufficient to refer to the following lemma.

LEMMA 7. $\{X_t\}$ *and* $\{\widetilde{X}_t\}$ *satisfy the relation*

$$\max_{t=0,\dots,T}|X_t - \widetilde{X}_t| \le \max_{t=0,\dots,T}\bar{\xi}_t + \max_{t=0,\dots,T}\underline{\xi}_t \tag{A.7}$$

for all $T \ge 0$.

Proof. Let $d_t := \max_{t'\le t}|X_t - \widetilde{X}_t|$. As $\max_{t=0,\dots,T}|X_t - \widetilde{X}_t| = \max_{t=0,\dots,T} d_t$ we prove the lemma by showing that the relation

$$d_t \le \max_{t'=0,\dots,t}\bar{\xi}_{t'} + \max_{t'=0,\dots,t}\underline{\xi}_{t'} \tag{A.8}$$

holds for all $t = 0,1,\dots,T$. The relation (A.8) is proved by induction. When $t = 0$, $X_0 = \widetilde{X}_0$, $d_0 = \bar{\xi}_0 = \underline{\xi}_0 = 0$, and (A.8) holds; therefore it suffices to show that $d_{t+1} \le \max_{t'=0,\dots,t+1}\bar{\xi}_{t'} + \max_{t'=0,\dots,t+1}\underline{\xi}_{t'}$ given that relation (A.8) holds at time t. Suppose that $X_t \ge \widetilde{X}_t$ and $\varepsilon_{t+1} < 0$. If $\widetilde{X}_t + \varepsilon_{t+1} \ge \underline{b}$, also $X_t + \varepsilon_{t+1} \ge \underline{b}$, so that $\underline{\xi}_{t+1} = \bar{\xi}_{t+1} = 0$, $d_{t+1} = d_t$, and relation (A.8) holds because $\max_{t'=0,\dots,t}\bar{\xi}_{t'} + \max_{t'=0,\dots,t}\underline{\xi}_{t'} = \max_{t'=0,\dots,t+1}\bar{\xi}_{t'} + \max_{t'=0,\dots,t+1}\underline{\xi}_{t'}$. If $\widetilde{X}_t + \varepsilon_{t+1} < \underline{b}$ and $X_t + \varepsilon_{t+1} \ge \underline{b}$, then $\underline{\xi}_{t+1} = \bar{\xi}_{t+1} = 0$; hence, as $X_{t+1} - \widetilde{X}_{t+1} = X_t - \widetilde{X}_t + (\widetilde{X}_t + \varepsilon_{t+1} - \underline{b})$ with $X_{t+1} - \widetilde{X}_{t+1} \ge 0$, $X_t - \widetilde{X}_t > 0$, and $(\widetilde{X}_t + \varepsilon_{t+1} - \underline{b}) < 0$, relation (A.8) holds as $d_{t+1} < d_t = \max_{t'=0,\dots,t}\bar{\xi}_{t'} + \max_{t'=0,\dots,t}\underline{\xi}_{t'} = \max_{t'=0,\dots,t+1}\bar{\xi}_{t'} + \max_{t'=0,\dots,t+1}\underline{\xi}_{t'}$. If both $\widetilde{X}_t + \varepsilon_{t+1}$ and $X_t + \varepsilon_{t+1}$ are smaller than $\underline{b}$, then $\widetilde{X}_{t+1} = \underline{b}$, $X_{t+1} = X_t + \varepsilon_{t+1} + \underline{\xi}_{t+1}$, and $\bar{\xi}_{t+1} = 0$; because $0 \le X_{t+1} - \widetilde{X}_{t+1} = (X_t + \varepsilon_{t+1} - \underline{b}) + \underline{\xi}_{t+1} < \underline{\xi}_{t+1}$ (as $X_t + \varepsilon_{t+1} < \underline{b}$), we have $d_{t+1} \le \max_{t'=0,\dots,t+1}\bar{\xi}_{t'} + \max_{t'=0,\dots,t+1}\underline{\xi}_{t'}$, so that relation (A.8) holds. By similar arguments, the induction proof also holds when $X_t < \widetilde{X}_t$ and, symmetrically, when $\varepsilon_{t+1} \ge 0$. ■

Under (A3), $\max_{t=0,\dots,T}\bar{\xi}_t + \max_{t=0,\dots,T}\underline{\xi}_t$ is of $o_p(T^{1/2})$, and hence Lemma 7 implies that $\sup_{s\in[0,1]}|X_T(s) - X_T(0) - (\widetilde{X}_T(s) - \widetilde{X}_T(0))| = \sup_{s\in[0,1]}|X_T(s) - \widetilde{X}_T(s)| \xrightarrow{p} 0$. Consequently, we can apply Billingsley (1968, Theorem 4.1), to conclude that Theorem 6 holds for $X_T(\cdot) - X_T(0)$ also. ■

Proof of Theorem 2. As in Phillips (1987a), we begin by writing the unit root statistics as the ratio between $(\frac{1}{2})(T^{-1}X_T^2 - \hat{\sigma}_0^2)$, $\hat{\sigma}_0^2 := T^{-1}\sum_{t=1}^{T}(\Delta X_t)^2$, and $T^{-2}\sum_{t=1}^{T}X_{t-1}^2$. Under Assumption $\mathcal{B}$, the weak convergence in (A2) of Definition 1 holds with $\lambda^2 := \sigma_v^2 C(1)^2$ (Phillips and Solo, 1992, Thm. 3.15); consequently, Theorem 1 and the CMT imply the joint weak convergence of $(T^{-1}X_T^2, T^{-2}\sum_{t=1}^{T}X_{t-1}^2)'$ to $\lambda^2(B_{\underline{c}}^{\bar{c}}(1)^2, \int_0^1 B_{\underline{c}}^{\bar{c}}(s)^2\,ds)'$ with λ^2 as previously defined. Finally, the convergence $\hat{\sigma}_0^2 \xrightarrow{p} \sigma^2 := \sigma_v^2\sum_{j=1}^{\infty}c_j^2$ (see the proof of Lemma 3, which follows), the CMT, and Theorem 4.1 of Billingsley (1968) give the desired result. ■

Proof of Lemma 3. For ease of notation and without loss of generality, we prove the lemma for the case of one bound at $\underline{c} < c_0$. Initially, we consider the estimator of λ^2 based on first-differenced data:

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

$$\hat{\lambda}_0^2 := \frac{1}{T}\sum_{j=-T+1}^{T-1}\omega\left(\frac{j}{q_T}\right)\sum_{t=|j|+1}^{T}\Delta X_t \Delta X_{t-|j|} = \frac{1}{T}\sum_{j=-T+1}^{T-1}\omega\left(\frac{j}{q_T}\right)\sum_{t=|j|+1}^{T} u_t u_{t-|j|};$$

later, we generalize the proof to the residual-based estimator. First, decompose $\hat{\lambda}_0^2$ as

$$\begin{aligned}
\hat{\lambda}_0^2 &= \frac{1}{T}\sum_{j=-T+1}^{T-1}\omega\left(\frac{j}{q_T}\right)\sum_{t=|j|+1}^{T} u_t u_{t-|j|} \\
&= \frac{1}{T}\sum_{j=-T+1}^{T-1}\omega\left(\frac{j}{q_T}\right)\sum_{t=|j|+1}^{T} (\varepsilon_t + \underline{\xi}_t)(\varepsilon_{t-|j|} + \underline{\xi}_{t-|j|}) \\
&= \frac{1}{T}\sum_{j=-T+1}^{T-1}\omega\left(\frac{j}{q_T}\right)\sum_{t=|j|+1}^{T} \varepsilon_t \varepsilon_{t-|j|} \\
&\quad + \frac{1}{T}\sum_{j=-T+1}^{T-1}\omega\left(\frac{j}{q_T}\right)\sum_{t=|j|+1}^{T} (\underline{\xi}_t \varepsilon_{t-|j|} + \varepsilon_t \underline{\xi}_{t-|j|} + \underline{\xi}_t \underline{\xi}_{t-|j|}) \\
&= \tilde{\lambda}^2 + \frac{1}{T}\sum_{j=-T+1}^{T-1}\omega\left(\frac{j}{q_T}\right)\sum_{t=|j|+1}^{T} (\underline{\xi}_t \varepsilon_{t-|j|} + \varepsilon_t \underline{\xi}_{t-|j|} + \underline{\xi}_t \underline{\xi}_{t-|j|}), \qquad \textbf{(A.9)}
\end{aligned}$$

where $\tilde{\lambda}^2$ is the idealized estimator given in equation (11). Under Assumptions $\mathcal{K}$ and (A2), Theorem 2 in Jansson (2002) shows that $\tilde{\lambda}^2 \xrightarrow{p} \lambda^2$, and hence to complete the proof we only need to prove that the last term on the right-hand side of equation (A.9) goes to 0 in probability. As $\underline{\xi}_t \geq 0$, all t, we can focus on the following inequalities:

$$\begin{aligned}
&\left|\frac{1}{T}\sum_{j=-T+1}^{T-1}\omega\left(\frac{j}{q_T}\right)(\underline{\xi}_t \varepsilon_{t-|j|} + \varepsilon_t \underline{\xi}_{t-|j|} + \underline{\xi}_t \underline{\xi}_{t-|j|})\right| \\
&\quad \leq \frac{1}{T}\sum_{j=-T+1}^{T-1}\left|\omega\left(\frac{j}{q_T}\right)\right|\sum_{t=|j|+1}^{T} |\underline{\xi}_t \varepsilon_{t-|j|} + \varepsilon_t \underline{\xi}_{t-|j|} + \underline{\xi}_t \underline{\xi}_{t-|j|}| \\
&\quad \leq \frac{1}{T}\sum_{j=-T+1}^{T-1}\left|\omega\left(\frac{j}{q_T}\right)\right| \max\left\{\max_{t=1,\dots,T}|\varepsilon_t|, \max_{t=1,\dots,T}\underline{\xi}_t\right\}\sum_{t=1}^{T} 3\underline{\xi}_t \\
&\quad \leq \frac{3L_T}{T^{1/2}}\left(\max_{t=1,\dots,T}|\varepsilon_t| + \max_{t=1,\dots,T}\underline{\xi}_t\right)\sum_{j=-T+1}^{T-1}\left|\omega\left(\frac{j}{q_T}\right)\right|\frac{1}{T^{1/2}} \qquad \textbf{(A.10)}
\end{aligned}$$

with $L_T := \sum_{t=1}^{T}\underline{\xi}_t$. Recursive substitutions allow us to express X_T as $X_T = c_0\lambda T^{1/2} + \sum_{t=1}^{T}\varepsilon_t + \sum_{t=1}^{T}\underline{\xi}_t$, implying that $T^{-1/2}L_T = T^{-1/2}\sum_{t=1}^{T}\underline{\xi}_t = T^{-1/2}X_T - c_0\lambda - T^{-1/2}\sum_{t=1}^{T}\varepsilon_t$ converges weakly to the well-defined random variable $\lambda(B_{\underline{c}-c_0}^{+\infty}(1) - B(1))$. Moreover, consider the following equality:

$$\begin{aligned}
&\frac{1}{T^{1/2}}\left(\max_{t=1,\dots,T}|\varepsilon_t| + \max_{t=1,\dots,T}\underline{\xi}_t\right)\sum_{j=-T+1}^{T-1}\omega\left(\frac{j}{q_T}\right) \\
&\quad = \frac{q_T}{T^{\delta^*}}\left(\max_{t=1,\dots,T}\frac{|\varepsilon_t|}{T^{1/2-\delta^*}} + \max_{t=1,\dots,T}\frac{\underline{\xi}_t}{T^{1/2-\delta^*}}\right)\left(\frac{1}{q_T}\sum_{j=-T+1}^{T-1}\omega\left(\frac{j}{q_T}\right)\right). \qquad \textbf{(A.11)}
\end{aligned}$$

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

By (B1), (B2), and (K2′),

$$\frac{\max_{t=0,\dots,T}|\varepsilon_t|}{T^{1/2-\delta^*}}, \frac{\max_{t=0,\dots,T}|\underline{\xi}_t|}{T^{1/2-\delta^*}} = O_p(1). \tag{A.12}$$

To check this result note that because (K2′) implies that δ^*, η^* solve $(\frac{1}{2}-\delta^*)\times(2+\eta^*)-1=0$, it follows that

$$\Pr\left\{\frac{\max_{t=0,\dots,T}|\varepsilon_t|}{T^{1/2-\delta^*}} > \varepsilon\right\} = \Pr\left\{\max_{t=0,\dots,T}|\varepsilon_t| > \varepsilon T^{1/2-\delta^*}\right\} \le \frac{\sup_{t=1,\dots,T} E|\varepsilon_t|^{2+\eta}}{\varepsilon^{2+\eta}T^{(1/2-\delta^*)(2+\eta)-1}}$$

$$\le \frac{\sup_{t=1,\dots,T} E|\varepsilon_t|^{2+\eta}}{\varepsilon^{2+\eta}T^{(1/2-\delta^*)(2+\eta^*)-1}} = \frac{\sup_{t=1,\dots,T} E|\varepsilon_t|^{2+\eta}}{\varepsilon^{2+\eta}} \tag{A.13}$$

and, similarly,

$$\Pr\left\{\frac{\max_{t=0,\dots,T}\underline{\xi}_t}{T^{1/2-\delta^*}} > \varepsilon\right\} \le \frac{\sup_{t=1,\dots,T} E|\underline{\xi}_t|^{2+\underline{\eta}}}{\varepsilon^{2+\underline{\eta}}T^{(1/2-\delta^*)(2+\underline{\eta})-1}} \le \frac{\sup_{t=1,\dots,T} E|\underline{\xi}_t|^{2+\underline{\eta}}}{\varepsilon^{2+\underline{\eta}}}. \tag{A.14}$$

As (A.13) and (A.14) can be made arbitrarily small by choosing ε appropriately, the result (A.12) is proved. Consequently, because (i) $q_T T^{-\delta^*} \to 0$ (see (K2′)); (ii) (K1) implies $\sup_{T\ge 1}(1/q_T)\sum_{j=-T+1}^{T-1}|\omega(j/q_T)| < \infty$ (see Jansson, 2002); (iii) $T^{-(1/2-\delta^*)}\times(\max_{t=1,\dots,T}|\varepsilon_t| + \max_{t=1,\dots,T}|\underline{\xi}_t|) = O_p(1)$; (iv) $T^{-1/2}L_T = O_p(1)$, it holds that (A.10), and hence $\hat{\lambda}_0^2 - \lambda^2$, are of $o_p(1)$.

By setting $q_T = 1$, $\omega(x) := \mathbb{I}\{x=0\}$ and using the weak law of large numbers $\tilde{\sigma}^2 := (1/T)\sum_{t=1}^T \varepsilon_t^2 \xrightarrow{p} \sigma^2 := \sigma_v^2 \sum_{j=1}^\infty c_j^2$ (which holds under (B1), see, e.g., Jansson, 2002, p. 1450), it also follows that $\hat{\sigma}_0^2 := (1/T)\sum_{t=1}^T(\Delta X_t)^2 = (1/T)\sum_{t=1}^T u_t^2 \xrightarrow{p} \sigma^2$; note that this result holds irrespective of Assumption $\mathcal{K}$.

Now, consider the difference $\hat{\lambda}^2 - \hat{\lambda}_0^2$, where $\hat{\lambda}^2$ is given in (10). Because $\hat{u}_t = u_t - (\hat{\rho}_T - 1)X_{t-1}$, simple algebra allows us to show that

$$\begin{aligned}
|\hat{\lambda}^2 - \hat{\lambda}_0^2| &= \left|\sum_{j=-T+1}^{T-1}\omega\left(\frac{j}{q_T}\right)\frac{1}{T}\sum_{t=|j|+1}^{T}(\hat{u}_t\hat{u}_{t-|j|} - u_t u_{t-|j|})\right| \\
&= \Bigg|(\hat{\rho}_T - 1)\sum_{j=-T+1}^{T-1}\omega\left(\frac{j}{q_T}\right) \\
&\quad \times \frac{1}{T}\sum_{t=|j|+1}^{T}((\hat{\rho}_T-1)X_{t-1}X_{t-1-|j|} - u_t X_{t-1-|j|} - X_{t-1}u_{t-|j|})\Bigg| \\
&\le |\hat{\rho}_T - 1|\sum_{j=-T+1}^{T-1}\left|\omega\left(\frac{j}{q_T}\right)\right| \\
&\quad \times \left(|\hat{\rho}_T - 1|\max_{t=1,\dots,T}X_t^2 + 2\max_{t=1,\dots,T}|u_t|\max_{t=1,\dots,T}|X_t|\right) \\
&= T|\hat{\rho}_T - 1|\left(\frac{1}{q_T}\sum_{j=-T+1}^{T-1}\left|\omega\left(\frac{j}{q_T}\right)\right|\right) \\
&\quad \times \frac{q_T}{T^{\delta^*}}\left(T^{\delta^*}|\hat{\rho}_T - 1|\frac{\max_{t=1,\dots,T}X_t^2}{T} + 2\frac{\max_{t=1,\dots,T}|u_t|}{T^{1/2-\delta^*}}\frac{\max_{t=1,\dots,T}|X_t|}{T^{1/2}}\right),
\end{aligned}$$

which is $o_p(1)$ because (i) $T(\hat{\rho}_T - 1) = O_p(1)$ (Theorem 2); (ii) $T^{-\delta^*} q_T \to 0$ and $q_T^{-1} \sum_{j=-T+1}^{T-1} |\omega(jq_T^{-1})|$ is asymptotically bounded (Assumption $\mathcal{K}$); (iii) $\max_{t=1,\dots,T} T^{-1/2}|X_t|$ is of $o_p(1)$ (Theorem 1 and the CMT); and (iv) $\max_{t=1,\dots,T} T^{-(1/2-\delta^*)}|u_t| \le \max_{t=1,\dots,T} T^{-(1/2-\delta^*)}|\underline{\xi}_t| + \max_{t=1,\dots,T} T^{-(1/2-\delta^*)}|\varepsilon_t| = O_p(1)$ (see equation (A.12)). Therefore, $\hat{\lambda}^2 \xrightarrow{p} \lambda^2$ also. As before, by setting $q_T = 1$ and $\omega(x) := \mathbb{I}\{x = 0\}$ it also follows that $\hat{\sigma}^2 - \hat{\sigma}_0^2 = o_p(1)$, which does not require Assumption $\mathcal{K}$ to hold.

Extension to the two-bound case and to the case of demeaned data follows similarly. ∎

Proof of Theorem 4. We initially prove that when $\alpha > 0$, Theorem 1 holds with $X_T(\cdot) \xrightarrow{w} B_{\underline{c}}^{\bar{c}}(\cdot)$ replaced by $X_T(\cdot) \xrightarrow{w} J_{\underline{c}.\alpha}^{\bar{c}}(\cdot)$. First, under (A2), Lemma 1(a) of Phillips (1987b) holds and implies that $(\lambda^2 T)^{-1/2} S_{[\cdot T]} \xrightarrow{w} J_\alpha(\cdot)$. The same weak convergence holds for the approximant

$$S_T^\alpha(s) := \frac{1}{\lambda T^{1/2}} S_{[sT]} + (S_{[sT]+1} - S_{[sT]}) \frac{(sT - [sT])}{\lambda T^{1/2}},$$

$$S_T^\alpha(1) := \frac{1}{\lambda T^{1/2}} S_T,$$

because

$$\sup_{s\in[0,1]} \left| S_T^\alpha(s) - \frac{1}{\lambda T^{1/2}} S_{[sT]} \right|$$

$$\le \frac{1}{\lambda T^{1/2}} \sup_{s\in[0,1]} |S_{[sT]+1} - S_{[sT]}|$$

$$\le (e^{-\alpha/T} - 1) \sup_{s\in[0,1]} \left| \frac{S_{[sT]}}{\lambda T^{1/2}} \right| + \frac{1}{\lambda T^{1/2}} \max_{t=1,\dots T} |\varepsilon_t| = o_p(1).$$

Second, convergence to the regulated diffusion $J_{\underline{c}.\alpha}^{\bar{c}}$ is obtained by following the proof of Theorem 1, where (A.2) is now replaced by $S_T^\alpha(s)$; see the preceding discussion.

As in the proof of Theorem 2, the weak convergence in (12) holds if $\hat{\sigma}_0^2 := T^{-1} \sum_{t=1}^T (\Delta X_t)^2$ converges to σ^2 in probability. But this result follows from the decomposition

$$\hat{\sigma}_0^2 = \frac{1}{T} \sum_{t=1}^T (\Delta S_t + \underline{\xi}_t - \bar{\xi}_t)^2 = \frac{1}{T} \sum_{t=1}^T ((\rho_T - 1)S_{t-1} + \varepsilon_t + \underline{\xi}_t - \bar{\xi}_t)^2$$

$$= \frac{1}{T} \sum_{t=1}^T (\varepsilon_t + \underline{\xi}_t - \bar{\xi}_t)^2 + \frac{(\rho_T - 1)^2}{T} \sum_{t=1}^T S_{t-1}^2 + \frac{2(\rho_T - 1)}{T} \sum_{t=1}^T (\varepsilon_t + \underline{\xi}_t - \bar{\xi}_t) S_{t-1}$$

https://doi.org/10.1017/S0266466605050462 Published online by Cambridge University Press

because (i) $T^{-1}\sum_{t=1}^{T}(\varepsilon_t + \underline{\xi}_t - \bar{\xi}_t)^2 \xrightarrow{p} \sigma^2$ (see Lemma 3); (ii) $T^{-1}(\rho_T - 1)^2 \times \sum_{t=1}^{T} S_{t-1}^2 = O_p(T^{-1})$ (as $\rho_T - 1 = -\alpha/T + O(T^{-2})$); and (iii)

$$\begin{aligned}
&\left| \frac{\rho_T - 1}{T} \sum_{t=1}^{T} (\varepsilon_t + \underline{\xi}_t - \bar{\xi}_t) S_{t-1} \right| \\
&\quad \leq \left| \frac{\rho_T - 1}{T} \right| \sum_{t=1}^{T} |\varepsilon_t + \underline{\xi}_t - \bar{\xi}_t| |S_{t-1}| \\
&\quad \leq \left| \frac{\rho_T - 1}{T} \right| \max_{t=1,\ldots,T} |\varepsilon_t + \underline{\xi}_t - \bar{\xi}_t| \sum_{t=1}^{T} |S_{t-1}| \\
&\quad \leq T|\rho_T - 1| \frac{\max_{t=1,\ldots,T}(|\varepsilon_t| + \underline{\xi}_t + \bar{\xi}_t)}{T^{1/2}} \sup_{s\in[0,1]} \left| \frac{S_{[sT]}}{T^{1/2}} \right| = o_p(1).
\end{aligned}$$

To prove the second part of the theorem, we only need to show that if the DGP is BNI(1), Lemma 3 continues to hold, and hence $\hat{\lambda}^2 \xrightarrow{p} \lambda^2$. This can be done by following mechanically the proof of Lemma 3, and it is therefore omitted for brevity. ■

Proof of Corollary 5. Because by Assumption (A4), $\hat{\underline{c}} := \underline{b}(\hat{\lambda}^2 T)^{-1/2} = \underline{c}\lambda T^{1/2} \times (\hat{\lambda}^2 T)^{-1/2} = \underline{c}\lambda/\hat{\lambda}$, consistency follows from $\hat{\lambda}^2 \xrightarrow{p} \lambda^2 > 0$; see Lemma 3. The same proof applies to $\hat{\bar{c}}$ and to $\hat{c}_0$. ■