*Econometric Theory*, **29**, 2013, 1289–1313.
doi:10.1017/S0266466613000066

# THE IMPACT OF PERSISTENT CYCLES ON ZERO FREQUENCY UNIT ROOT TESTS

TOMÁS DEL BARRIO CASTRO
*University of the Balearic Islands*

PAULO M.M. RODRIGUES
*Banco de Portugal*
*Universidade Nova de Lisboa*
*and*
*CEFAGE*

A.M. ROBERT TAYLOR
*University of Essex*

In this paper we investigate the impact of persistent (nonstationary or near nonstationary) cycles on the asymptotic and finite-sample properties of standard unit root tests. Results are presented for the augmented Dickey–Fuller (ADF) normalized bias and $t$-ratio-based tests (Dickey and Fuller, 1979, *Journal of the American Statistical Association* 745, 427–431; Said and Dickey, 1984; *Biometrika* 71, 599–607). the variance ratio unit root test of Breitung (2002, *Journal of Econometrics* 108, 343–363), and the $M$ class of unit-root tests introduced by Stock (1999, in Engle and White (eds.), *A Festschrift in Honour of Clive W.J. Granger*) and Perron and Ng (1996, *Review of Economic Studies* 63, 435–463). We show that although the ADF statistics remain asymptotically pivotal (provided the test regression is properly augmented) in the presence of persistent cycles, this is not the case for the other statistics considered and show numerically that the size properties of the tests based on these statistics are too unreliable to be used in practice. We also show that the $t$-ratios associated with lags of the dependent variable of order greater than two in the ADF regression are asymptotically normally distributed. This is an important result as it implies that extant sequential methods (see Hall, 1994, *Journal of Business & Economic Statistics* 17, 461–470; Ng and Perron, 1995, *Journal of the American Statistical Association* 90, 268–281) used to determine the order of augmentation in the ADF regression remain valid in the presence of persistent cycles.

We thank the editor, Peter Phillips, the co-editor, Pentti Saikkonen, two anonymous referees, and Bent Nielsen for their helpful and constructive comments on earlier versions of this paper. Tomás del Barrio Castro and Paulo Rodrigues gratefully acknowledge financial support from the Ministerio de Educación y Ciencia ECO2011-23934 and the Fundação para a Ciência e a Tecnologia (POCTI program), respectively. Address correspondence to Robert Taylor, Essex Business School, University of Essex, Colchester CO4 3SQ, United Kingdom, e-mail: rtaylor@essex.ac.uk.

© Cambridge University Press 2013 

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

## 1. INTRODUCTION

Peaks at low nonzero frequencies imply the existence of cycles in time series, a feature present in many macroeconomic, financial, and other time series; see, among others, Conway and Frame (2000), Birgean and Kilian (2002), and Priestley (1981). The importance of and interest in trend growth, cyclicality, and seasonal fluctuations in economic and financial time series date back several decades; see Burns and Mitchell (1946). For instance, Canova (1996) discusses the literature on Bayesian learning (Nyarko, 1992) and on noisy traders in financial markets (Campbell and Kyle, 1993) where models that generate irregularly spaced but significant cycles in economic activity and asset prices are proposed. The presence of cycles is also documented in the political economy literature (electoral cycles in government variables; Alesina and Roubini, 1992) and naturally arises in the business cycle literature (Pagan, 1997).

The existence of both complex and real unit (or near-unit) roots can induce growth cycles similar to those observed in economic data; see Allen (1997). Autoregressive (AR) processes with roots on or in a neighborhood of the complex unit circle display persistent cyclical behavior similar to that of persistent business cycles; see Bierens (2001), who reports evidence that business cycles may indeed be due to complex unit roots. Shibayama (2008) studies inventories and monetary policy by estimating vector AR (VAR) models and detects complex roots generating cycles of around 55–70 months, which are close to business cycle lengths. Hence, complex eigenvalues would appear to be a feature frequently found in aggregate time series. Azariadis, Bullard, and Ohanian (2004) provide an overview of evidence on the eigenvalues from the AR and VAR representations of macroeconomic variables. They stress that economic time series typically indicate an autocorrelation structure that may originate complex or negative eigenvalues in empirical reduced forms. In a different context, Estrella (2004) uses an AR process with complex roots to examine the cyclical behavior of optimal bank capital, where a cycle is defined as a particular predictable pattern that unfolds over time.

Given that cycles would appear to be a common feature in economic and financial variables it is important to investigate their implications on the performance of pretesting procedures, in particular on the limiting null distributions and finite-sample properties of zero frequency unit root test statistics. Consequently, following the body of empirical evidence summarized earlier, in this paper we will focus on the case where the cyclical component is characterized by a second-order AR component with complex roots in the neighborhood of unity, although generalizations to higher order nonstationary factors will also be discussed.

The asymptotic distributions of the least squares estimates of an AR(2) model with complex unit roots have been addressed in Ahtola and Tiao (1987), Gregoir (2004), and Tanaka (2008); the latter generalizes the results in Ahtola and Tiao (1987) by allowing the error term in the AR(2) model to follow a stationary process. For asymptotic results for general AR processes see, for example,

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

Chan and Wei (1988). Chan and Wei (1988) considered the limiting distributions of the least squares estimate of a general nonstationary AR model of order $p$ (AR($p$)) with characteristic roots at different frequencies on or outside the unit circle, each of which may have different multiplicities (see also Nielsen, 2001). This was the first comprehensive treatment of least squares estimates for a general nonstationary AR($p$) model. Chan and Wei (1988) showed that the location of the roots of the time series played an important role in characterizing the limiting distributions. Jeganathan (1991) generalized this idea to the near-integrated context where the limiting distributions of the least squares estimators are expressed in terms of Ornstein–Uhlenbeck (OU) processes. Extensions to VAR processes are provided in Tsay and Tiao (1990) and to processes with deterministic trends in Chan (1989).

We will focus attention on the conventional augmented Dickey–Fuller (ADF) tests (Dickey and Fuller, 1979; Said and Dickey, 1984) and the variance ratio unit root test of Breitung (2002) and will also consider the trinity of so-called $M$ unit root tests introduced by Stock (1999) and popularized by Perron and Ng (1996). The $M$ tests have been increasingly popular in the unit root literature; indeed, in discussing the ADF tests Haldrup and Jansson (2006, p. 267) argue that "practitioners ought to abandon the use of these tests" in favor of the $M$ tests because of "the excellent size properties and 'nearly' optimal power properties" of the latter. The results in this paper, however, suggest that quite the opposite conclusion holds in cases where the time series process admits (near-) unit roots at cyclical frequencies. In particular we show that although the limiting distributions of the two ADF statistics are unaltered by the presence of persistent cycles (provided the test regression is properly augmented), this is not the case for the variance ratio or the three $M$ statistics, all of which become too unreliable to be used in practice. We also show that the standard sequential methods (see Hall, 1994; Ng and Perron, 1995) used to determine the order of augmentation in the ADF regression remain asymptotically valid when cyclical (near-) unit roots are present in the data.

The remainder of the paper is organized as follows. In Section 2 we outline our reference time series model that allows for cyclical near-integrated roots and briefly outline the ADF, variance ratio, and $M$ unit root tests. In Section 3, in the case where persistent cycles are present in the data, we discuss the large-sample behavior of these tests together with those of the conventional $t$-ratios for testing the significance of lagged dependent variables in the ADF test regression. Finite-sample simulations are reported in Section 4. Section 5 concludes. All proofs are collected in an Appendix.

In what follows, $\lfloor \cdot \rfloor$ denotes the integer part of its argument, $\Rightarrow$ denotes weak convergence, $\xrightarrow{p}$ denotes convergence in probability, and $x := y$ ($x =: y$) indicates that $x$ is defined by $y$ ($y$ is defined by $x$).

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

## 2. THE MODEL AND UNIT ROOT TESTS

### 2.1. The Time Series Model

We consider a univariate time series $\{x_t\}$ generated according to the following data generation process (DGP):

$$\Psi(L)(1-\varphi_{0n}L)x_t = \varepsilon_t, \qquad \varepsilon_t \sim IID\left(0,\sigma^2\right), \qquad t = 1,2,\ldots,n, \tag{1}$$

where $L$ denotes the usual lag operator and *IID* denotes independent and identical distribution. It is assumed throughout that the process is initialized at $x_{-2} = x_{-1} = x_0 = 0$, although weakening this to allow these starting values to be of $o_p(n^{1/2})$ would not change any of the asymptotic results that follow. In (1) the AR polynomial $\Psi(L) = (1 - 2\cos(\phi)\varphi_{\phi n}L + \varphi_{\phi n}^2L^2)$, with $\phi \in (0,\pi)$ and $\varphi_{\phi n} := \exp\left(c_\phi/n\right) \simeq \left(1+c_\phi/n\right)$ with $c_\phi \leq 0$ and fixed. Consequently, when $c_\phi = 0$ ($c_\phi < 0$), $\Psi(L)$ admits the complex conjugate pair of unit (near-unit) roots, $\exp(\pm i\phi) \equiv \cos(\phi) \pm i\sin(\phi)$, at the spectral frequency $\phi$. Notice that when $c_\phi < 0$ a complex conjugate pair of stable roots at frequency $\phi$ is obtained for any finite $n$. The polynomial $\Psi(L)$ generates a persistent cycle of $2\pi/\phi$ periods. Indeed, and as noted by Díaz-Emparanza (2004), it also generates an aliased cycle of $2\pi/(2\pi-\phi)$ periods, and as such our analysis in fact covers $\phi \in (0,2\pi) - \{\pi\}$.

In (1) we also allow for a (near-) unit root at the zero frequency through the parameter $\varphi_{0n} := \exp(c_0/n) \simeq (1 + c_0/n)$ with $c_0 \leq 0$ and fixed; a zero frequency unit (near-unit) root is obtained when $c_0 = 0$ ($c_0 < 0$). In the case where $c_0 = c_\phi = 0$, $\{x_t\}$ is therefore integrated of order one at both the zero and $\phi$ spectral frequencies, denoted $I_0(1)$ and $I_\phi(1)$, respectively. In this case, it follows that $z_t := \Delta x_t$, where $\Delta := (1-L)$, will be $I_\phi(1)$ but $I_0(0)$, whereas $u_t := \Delta_\phi x_t$, where $\Delta_\phi := \left(1 - 2\cos(\phi)L + L^2\right)$, will be $I_0(1)$ but $I_\phi(0)$. Our focus in this paper is on testing the standard zero frequency unit root null hypothesis that $x_t \sim I_0(1)$, $H_0 : \varphi_{0n} = 1$, against the alternative that $x_t \sim I_0(0)$, $H_1 : |\varphi_{0n}| < 1$, in the case where $\Psi(L)$ admits a pair of near-integrated complex roots at frequency $\phi$.

**Remark 1.** The model in (1) is arguably quite simple. However, it can be extended to allow for deterministic components, weak dependence in $\{\varepsilon_t\}$, and unit roots at other cyclical frequencies lying in $(0,\pi)$ and/or at the Nyquist ($\pi$) frequency, without altering the qualitative conclusions that can be drawn from the analysis of (1). For expositional purposes we will therefore focus on (1), but we will discuss how our results generalize to these cases.

### 2.2. Zero Frequency Unit Root Tests

A large number of procedures have been proposed to test for zero frequency unit roots; see, for example, Stock (1994), Maddala and Kim (1998), and Phillips and Xiao (1998) for excellent overviews. In this section we briefly review three popular classes of such tests.

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

The first tests we consider are the ADF normalized bias and $t$-ratio tests. These are computed from the fitted (ordinary least squares (OLS)) regression,

$$\Delta x_t = \widehat{\rho} x_{t-1} + \sum_{j=1}^{k} \widehat{\alpha}_j \Delta x_{t-j} + \widehat{\varepsilon}_{k,t}. \tag{2}$$

In (2), $k$ denotes the lag truncation order chosen to account (parametrically) for $\Psi(L)$ and any weak dependence in $\{\varepsilon_t\}$; in the simplest form of the DGP given in (1), where $\varepsilon_t$ is independently and identically distributed (i.i.d.) and $k = 2$. More generally, where $\varepsilon_t$ is a linear process satisfying standard summability and moment conditions (see Chang and Park, 2002), $k$ needs to be such that $1/k + k^3/n \to 0$ as $n \to \infty$; see Said and Dickey (1984) and Chang and Park (2002). The ADF $t$-ratio for testing $H_0$ against $H_1$ will be denoted $t_{\widehat{\rho}} := \widehat{\rho}/se\left(\widehat{\rho}\right)$ and the associated normalized bias statistic as $Z_{\widehat{\rho}} := n\widehat{\rho}/(1 - \sum_{i=1}^{k} \widehat{\alpha}_i)$. The ADF tests remain the most popularly applied unit root tests in part because of their ease of construction. At this point we also outline the sequential lag selection method of Hall (1994) and Ng and Perron (1995). Here one starts from a maximum lag length, $k_{\max}$, say, in (2), satisfying the preceding rate condition. One then runs the standard $t$-test for the significance of the terminal lag using critical values from the standard normal distribution. If the null is rejected the ADF statistics are computed from (2) with lag length $k_{\max}$; otherwise, the lag length is reduced to $k_{\max} - 1$ and the procedure is repeated until the lag length cannot be reduced further, or if a prespecified value for the minimum lag length is attained, at which point the ADF statistics are computed for that lag length.

Second, we will consider the trinity of so-called $M$ unit root tests of Stock (1999) and Perron and Ng (1996), among others:

$$MSB := \left( \hat{\lambda}^{-2} n^{-2} \sum_{t=1}^{n} x_{t-1}^2 \right)^{1/2}, \qquad MZ_\alpha := \frac{n^{-1} x_n^2 - \hat{\lambda}^2}{2n^{-2} \sum_{t=1}^{n} x_{t-1}^2},$$

$$MZ_t := MSB \times MZ_\alpha, \tag{3}$$

where $\hat{\lambda}^2$ is an estimator of the long-run variance of $\{\varepsilon_t\}$. Following Perron and Ng (1996) we can consider two alternative estimators for the long-run variance: First, a nonparametric kernel estimator based on the sample autocovariances, $\hat{\lambda}^2 = s_{WA}^2$, with $s_{WA}^2 := \sum_{h=-n+1}^{n-1} \omega(h/m) \hat{\gamma}_h$, $\hat{\gamma}_h := n^{-1} \sum_{t=1}^{n-|h|} \hat{\varepsilon}_{0,t} \hat{\varepsilon}_{0,t+|h|}$, where $\hat{\varepsilon}_{0,t}$ are the OLS residuals from regressing $x_t$ on $x_{t-1}$, with kernel function $\omega(\cdot)$ satisfying, for example, the general conditions reported in Jansson (2002, Assump. A3) and the bandwidth parameter $m \in (0, \infty)$ satisfying $1/m + m^2/n \to 0$ as $n \to \infty$ (which corresponds to Assumption A4 of Jansson, 2002). Second, a parametric AR spectral density estimator, $\hat{\lambda}^2 := s_{AR}^2$, of the form suggested by Berk (1974), where $s_{AR}^2 := \hat{\sigma}_k^2 / \left(1 - \sum_{i=1}^{k} \widehat{\alpha}_i\right)^2$ and $\hat{\sigma}_k^2 := n^{-1} \sum \widehat{\varepsilon}_{k,t}^2$, where $k$ in (2) satisfies the same conditions as given before in the context of the ADF tests. As noted in Section 1, it has been suggested by some authors (e.g., Haldrup and Jansson, 2006) that the

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

$M$ tests, when coupled with the modified Akaike information criterion (AIC) lag selection method of Ng and Perron (2001), are preferable to the standard ADF tests outlined previously because of their superior size properties, relative to the latter, in the presence of weak dependence in $\{\varepsilon_t\}$.

Finally, we will also consider the variance ratio test proposed by Breitung (2002):

$$VRT := n^{-2}\left(\sum_{t=1}^{n} x_t^2\right)^{-1} \sum_{t=1}^{n}\left(\sum_{j=1}^{t} x_j\right)^2. \tag{4}$$

The variance ratio test has some appealing properties. First of all it requires no correction, parametric or nonparametric, for serial correlation from $\{\varepsilon_t\}$ and/or $\Psi(L)$. Second, by virtue of its lack of such a correction factor, it has been advocated by some authors (see, e.g., Müller, 2008) as a unit root test that avoids the criticisms of Faust (1996) regarding the (theoretical) uncontrollability of the size of unit root tests based around (parametric or nonparametric) corrections for general weak dependence in $\{\varepsilon_t\}$.

## 3. ASYMPTOTIC (LARGE-SAMPLE) BEHAVIOR UNDER PERSISTENT CYCLES

In this section we discuss the large-sample behavior of the various unit root tests presented in the previous section. In Theorem 1 we detail the null distributions of the unit root tests from Section 2.2 when the DGP contains persistent cycles. Subsequently, in Theorem 2, we will establish corresponding results for the $t$-ratios on the lagged dependent variables appearing in the ADF regression (2).

THEOREM 1. *Let $\{x_t\}$ be generated by (1) with $\varphi_{0n} = 1 + c_0/n$ and $\varphi_{\phi n} = 1 + c_\phi/n$, and $c_0 \le 0$, $c_\phi \le 0$ and fixed. Then for any $\phi \in (0, \pi)$ and for $k \ge 2$ in (2), the following results hold:*

$$t_{\widehat{\rho}} \Rightarrow \frac{\int_0^1 W_{0,c_0}(r)dW_{0,c_0}(r)}{\left\{\int_0^1 \left[W_{0,c_0}(r)\right]^2 dr\right\}^{1/2}} := \mathcal{DF}_1, \quad Z_{\widehat{\rho}} \Rightarrow \frac{\int_0^1 W_{0,c_0}(r)dW_{0,c_0}(r)}{\int_0^1 \left[W_{0,c_0}(r)\right]^2 dr} := \mathcal{DF}_2, \tag{5}$$

$$VRT \Rightarrow \frac{\int_0^1 \left[\int_0^r W_{0,c_0}(s)ds\right]^2 dr}{\int_0^1 W_{0,c_0}(r)^2 dr + \frac{(1-\cos\phi)}{2\sin^2\phi}\int_0^1 \left(\left[W^{\alpha}_{\phi,c_\phi}(r)\right]^2 + \left[W^{\beta}_{\phi,c_\phi}(r)\right]^2\right) dr}. \tag{6}$$

*In the case where $\hat{\lambda}^2 = s^2_{WA}$,*

$$MSB = O_p\left([mn]^{-1/2}\right), \qquad MZ_\alpha = O_p(mn), \qquad MZ_t = O_p\left([mn]^{1/2}\right). \tag{7}$$

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

*In the case where* $\hat{\lambda}^2 = s^2_{AR}$, $MZ_\alpha$ *and* $MZ_t$ *are both of* $O_p(1)$ *with limiting distributions that depend on* $\phi$, *whereas*

$$MSB \Rightarrow \left\{ \int_0^1 W_{0,c_0}(r)^2 dr + \frac{(1-\cos\phi)}{2\sin^2\phi} \left\{ \int_0^1 \left[ W^{\alpha}_{\phi,c_\phi}(r) \right]^2 dr + \int_0^1 \left[ W^{\beta}_{\phi,c_\phi}(r) \right]^2 dr \right\} \right\}^{1/2}, \quad \textbf{(8)}$$

*where* $m$ *is the bandwidth used to compute* $s^2_{WA}$ *and* $W_{0,c_0}(r)$, $W^{\alpha}_{\phi,c_\phi}(r)$, *and* $W^{\beta}_{\phi,c_\phi}(r)$ *are independent standard OU processes.*

**Remark 2.** In the case where no (near) nonstationary cycles are present in $\Psi(L)$ in (1), such that $\Psi(L)$ is a stable $AR(2)$ polynomial for all $n$, it is well known (see, e.g., Hamilton, 1994, and Haldrup and Jansson, 2006) that both $t_{\widehat{\rho}}$ and $MZ_t$ weakly converge to $\mathcal{DF}_1$, whereas $Z_{\widehat{\rho}}$ and $MZ_\alpha$ weakly converge to $\mathcal{DF}_2$, with $\mathcal{DF}_1$ and $\mathcal{DF}_2$ as defined in (5) of Theorem 1. Moreover, in this case $MSB \Rightarrow \left( \int_0^1 W^2_{0,c_0}(r) \right)^{1/2} =: \mathcal{MSB}$ and $VRT \Rightarrow \left( \int_0^1 \left[ \int_0^r W_{0,c_0}(s) ds \right]^2 dr \right) / \left( \int_0^1 W_{0,c_0}(r)^2 dr \right) =: \mathcal{VRT}$. A comparison with Theorem 1 shows that only the two ADF statistics, $t_{\widehat{\rho}}$ and $Z_{\widehat{\rho}}$, retain their usual limiting distributions in the presence of (near-) integrated cycles. Moreover, the convergence in probability of $\widehat{\rho}$ to zero at rate $n$ is also unaffected by the presence of persistent cycles. Of the remaining statistics considered, only the $VRT$ statistic and the $MSB$ statistic with an AR spectral density estimator of the long-run variance have well-defined limiting distributions. However, these are both nonpivotal, their functional forms comprising additional terms arising from the persistent cycle at frequency $\phi$.

**Remark 3.** Where a kernel-based estimator of the long-run variance is used, (7) shows that the trinity of $M$ statistics have degenerate limiting null distributions in the presence of near-integrated cycles. This result obtains because $\Delta x_t$ is nonstationary, which causes the kernel-based long-run variance estimator to diverge to $+\infty$, which in turn causes both $MZ_\alpha$ and $MZ_t$ statistics to diverge to $-\infty$ and causes $MSB$ to converge in probability to zero. Consequently, the three $M$ tests based on these statistics will all have asymptotic size of unity in the presence of persistent cycles.

**Remark 4.** The three $M$ statistics no longer diverge when $s^2_{AR}$ is used. However, these statistics do not have the pivotal limiting distributions given in Remark 2 that obtain in the conventional $I(1)$ case. Although the limiting distribution for $MSB$ in (8) is easily obtained, deriving those for the $MZ_\alpha$ and $MZ_t$ statistics is considerably complicated by the presence of the term $n^{-1}x_n^2$ in the numerator of these statistics. In particular, to demonstrate convergence of this object one needs to use a standard central limit theorem (CLT)–based approach, whereby it can be shown that $n^{-1/2}x_n$ converges to a mean zero normal distribution whose

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

variance is a function of $\phi$. The limiting distributions for the denominators of these two statistics both follow immediately from the first result in Lemma A.3 in the Appendix. Although these results are not sufficient for us to provide explicit representations for the limiting distributions of the $MZ_\alpha$ and $MZ_t$ statistics, they do inform us on the key point that these limiting distributions, like that of the $MSB$ statistic, also do not have the pivotal form given in Remark 2 when a pair of (near) unit roots is present at frequency $\phi$. The effects of this will be explored further using Monte Carlo simulation methods in Section 4.

**Remark 5.** In this paper we have not included the $Z_\alpha$ and $Z_t$ unit root tests of Phillips (1987) and Phillips and Perron (1988) in the set of tests under discussion. However, from expressions (2.7) and (2.10) of Perron and Ng (1996, p. 437) we have that $Z_\alpha = MZ_\alpha - (n/2)(\widehat{\alpha}-1)^2$ and $Z_t = MZ_t - 0.5(\sum_{t=1}^{n} x_{t-1}^2/s_{WA}^2)^{1/2}(\widehat{\alpha}-1)^2$, where $\widehat{\alpha}$ is the OLS estimate from regressing $x_t$ on $x_{t-1}$, in each case for the version of the $M$ test formed using $s_{WA}^2$. It then follows that both $Z_\alpha$ and $Z_t$ will also diverge to $-\infty$, at the same rates as are given for $MZ_\alpha$ and $MZ_t$, respectively, in (7), noting that the link term, $(n/2)(\widehat{\alpha}-1)^2$, is of $O_p(n)$ under the conditions of Theorem 1.

**Remark 6.** A corresponding result to that given for $t_{\widehat{\rho}}$ in (5) is derived in the context of the likelihood ratio test for a unit root in Nielsen (2001). A similar result has also been established for the case of seasonally integrated data in Ghysels, Lee, and Noh (1994) and Rodrigues (1999).

**Remark 7.** It is straightforward, using Lemma 2 of Bierens (2000), to show that the results given in (5) and (7) of Theorem 1 will not alter if we allow for weaker linear process conditions on $\{\varepsilon_t\}$ provided that Assumptions 1.1 and 1.2 of Gregoir (2004), adapted slightly to our situation, are satisfied. Precisely, these conditions entail that $\varepsilon_t = d(L)e_t$, where $(\varepsilon_t, \mathcal{F}_t)$ is a martingale difference sequence, with filtration $(\mathcal{F}_t)$, such that $\mathbb{E}\left(\varepsilon_t^2|\mathcal{F}_{t-1}\right) = \sigma^2$ and $\sup_t \mathbb{E}\left(|\varepsilon_t|^{2+\kappa}|\mathcal{F}_{t-1}\right) < \infty$ almost surely for some $\kappa > 0$, and where $d(L) := 1 + \sum_{j=1}^{\infty} d_j z^j$ is such that $d(z) \neq 0$ for $z = 0$ and $z = \phi$, and $\sum_{j=1}^{\infty} j\left|d_j\right| < \infty$. In contrast, the limiting null distributions for $VRT$ and the trinity of $M$ statistics with AR spectral density estimators of the long-run variance would now depend on additional nuisance parameters arising from the moving average parameters, $\{d_j\}_{j=1}^{\infty}$. Moreover, the results in Theorem 1 are qualitatively unchanged if we allow $\Psi(L)$ to be a $p$th-order polynomial containing additional unit roots with frequencies in the range $(0, \pi]$, provided $k \geq p$ in (2). Specifically, in this case the results in (5) will continue to hold, as will the order results in (7), whereas the limiting null distributions for $VRT$ and the trinity of $M$ statistics with AR spectral density estimators of the long-run variance will now depend on stochastic and nonstochastic functions relating to these frequencies but will remain nondegenerate.

In Theorem 2 we now establish the large-sample behavior of the $t$-ratios on the lagged dependent variables in (2).

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

THEOREM 2. *Let the conditions of Theorem 1 hold and define the vector of parameters from (2) as* $\beta' := [\rho, \alpha_1, \alpha_2, \alpha_3 \ldots \alpha_k] =: [\rho, \alpha_1, \alpha_2, \Phi']$. *For* $k \geq 2$ *in (2), then,*

$$n(\hat{\alpha}_1 - \alpha_1) \Rightarrow \frac{2(\Psi_1 \sin\phi + \Psi_2 \cos\phi)}{\mathcal{A}^2 + \mathcal{B}^2} + \frac{\Psi_0 (2\cos\phi - 1)}{\mathcal{V}_{c_0}^2}, \tag{9}$$

$$n(\hat{\alpha}_2 - \alpha_2) \Rightarrow -\left(\frac{\Psi_0}{\mathcal{V}_{c_0}^2} + \frac{2\Psi_2}{\mathcal{A}^2 + \mathcal{B}^2}\right), \tag{10}$$

$$n^{1/2}\left(\hat{\Phi} - \Phi\right) \Rightarrow N\left(0, \sigma^2 \mathbf{H}_2 \Gamma^{-1} \mathbf{H}_2\right), \tag{11}$$

*where* $\Psi_0 := \int_0^1 W_{0,c_0}(r) dW_{0,c_0}(r)$, $\Psi_1 := \int_0^1 \left[W^{\alpha}_{\phi,c_\phi}(r) dW^{\beta}_{\phi,c_\phi}(r) - W^{\beta}_{\phi,c_\phi}(r) dW^{\alpha}_{\phi,c_\phi}(r)\right]$, $\Psi_2 := \int_0^1 \left[W^{\alpha}_{\phi,c_\phi}(r) dW^{\alpha}_{\phi,c_\phi}(r) + W^{\beta}_{\phi,c_\phi}(r) dW^{\beta}_{\phi,c_\phi}(r)\right]$, $\mathcal{V}_{c_0}^2 := \int_0^1 W_{0,c_0}^2(r) dr$, $\mathcal{A}^2 := \int_0^1 \left[W^{\alpha}_{\phi,c_\phi}(r)\right]^2 dr$, $\mathcal{B}^2 := \int_0^1 \left[W^{\beta}_{\phi,c_\phi}(r)\right]^2 dr$, *with* $W_{0,c_0}(r)$, $W^{\alpha}_{\phi,c_\phi}(r)$, *and* $W^{\beta}_{\phi,c_\phi}(r)$ *as in Theorem 1;* $\Gamma$ *is defined in (A.8) in the Appendix; and*

$$\mathbf{H}_2 := \begin{bmatrix} 1 & -2\cos\phi & \cdots & 0 \\ 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & -2\cos\phi & 1 \\ 0 & 0 & 1 & -2\cos\phi \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

*Moreover, for* $2 < i \leq k$,

$$t_{\widehat{\alpha}_i} := (\widehat{\alpha}_i - \alpha_i)/s.e.(\widehat{\alpha}_i) \Rightarrow N(0,1). \tag{12}$$

**Remark 8.** It follows from the results in (9) and (10) that $\widehat{\alpha}_1$ and $\widehat{\alpha}_2$ converge, at rate $n$, to $2\cos\phi$ and $-1$, respectively. From (11), it is seen that the parameters on the lagged difference terms from lag three onward are root-$n$ consistent asymptotically normal (CAN). Under the conditions of Theorem 2, $\Phi = \mathbf{0}$, because $\varepsilon_t \sim IID(0, \sigma^2)$; however, the stated results also hold when $\varepsilon_t$ is a stationary $AR(k^*)$ process provided $k \geq k^*$. Moreover, where $\varepsilon_t$ displays the general weak dependence of the form given in Remark 7 the foregoing results still remain valid, although here, for a given lag truncation $k$, the parameters $\alpha_j$, $j \geq 3$, take the role of pseudo-parameters in the same sense as in, for example, Chang and Park (2002) and Ng and Perron (1995).

**Remark 9.** The key result in Theorem 2 is that given in (12), which establishes that standard $t$-tests for the significance of the lagged dependent variables of order three and above can be conducted using standard normal critical values (see also Nielsen, 2006). Moreover, the result in (10) implies that the $t$-statistic for $\alpha_2 = 0$

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

in (2) will diverge to $-\infty$ at rate $n$. Taken together these results imply that the sequential lag specification methods of Hall (1994) and Ng and Perron (1995), as outlined in Section 2.2, remain asymptotically valid in the presence of nonstationary cycles. In the scenario considered in Remark 7 where $\Psi(L)$ is a $p$th-order polynomial of near-integrated factors then it is straightforward but tedious to show that the OLS estimators associated with the first $p$ lagged dependent variables in (2) will have nonstandard limiting distributions (now being functionals of $p$ independent standard OU processes) whereas those for $p+1$ onward will again be root-$n$ CAN with their associated $t$-ratios having standard normal limiting null distributions.

**Remark 10.** The previous results extend straightforwardly to cases where a deterministic trend is present in the DGP. Here (2) needs to be constructed from detrended data, denoted $\hat{x}_t$, obtained as the OLS residuals from the regression of $x_t$ onto the deterministic trend; principally either a constant (demeaned data) or a constant and linear trend (linear detrended data). Similarly the variance ratio and $M$ unit root tests are constructed using $\hat{x}_t$, although in the definition of the $MZ_\alpha$ statistic the term $-n^{-1}\hat{x}_0^2$ needs to be added to the numerator of $MZ_\alpha$ in (3); see, for example, Müller and Elliott (2003). In both of these cases the results in Theorems 1 and 2 remain valid provided $W_{0,c_0}(r)$ in (5) is replaced by its OLS detrended counterpart. Finally, for the corresponding unit root tests based on local generalized least squares (GLS) detrending (see, among others, Elliott, Rothenberg, and Stock, 1996; Ng and Perron, 2001) the previous results again hold but now replacing $W_{0,c_0}(r)$ by its local GLS demeaned or linear detrended counterparts; see Elliott et al. (1996, pp. 824–825) for precise details.

## 4. MONTE CARLO EXPERIMENTS

In this section, we use simulation methods to investigate the finite-sample properties of the unit root tests discussed in Section 2.2 when (near-) nonstationary cycles are present in the data. All reported results are based on 20,000 Monte Carlo replications using the RNDN function of Gauss 9.0. Unless otherwise stated, results are presented for both demeaned and linear detrended data.

Our first set of experiments, reported in Tables 1 (OLS demeaned data) and 2 (linear OLS detrended data), relate to data generated according to the DGP

$$\left(1-\left(1+\frac{c_0}{n}\right)L\right)\left(1-2\cos(\phi)\left(1+\frac{c_\phi}{n}\right)L+\left(1+\frac{c_\phi}{n}\right)^2L^2\right)x_t = \epsilon_t \sim NID(0,1) \tag{13}$$

with $c_0 = 0$, where *NID* denotes normally and independently distributed, so that the unit root null hypothesis holds, and with $\phi \in \{\pi/16, \pi/8, \pi/4, \pi/2\}$ and $c_\phi \in \{0, -2.5, -5, -10, -20\}$, in each case for a sample size of $n = 200$, initialized at $x_{-2} = x_{-1} = x_0 = 0$. These values of $\phi$ were considered because they

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press



**TABLE 1.** Empirical rejection frequencies, OLS demeaned data, DGP (13)

| | $t_{\widehat{\rho}}$ | $t_{\widehat{a}_1}$ | $t_{\widehat{a}_2}$ | $Z_{\widehat{\rho}}$ | $VRT$ | $MSB^b$ | $MSB^q$ | $MSB^{AR}$ | $MZ_\alpha^b$ | $MZ_\alpha^q$ | $MZ_\alpha^{AR}$ | $MZ_t^b$ | $MZ_t^q$ | $MZ_t^{AR}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | $c_\phi = 0$ | | | | | | | |
| 0 | 0.050 | 0.049 | 0.046 | 0.058 | 0.048 | 0.043 | 0.043 | 0.039 | 0.045 | 0.046 | 0.042 | 0.046 | 0.047 | 0.043 |
| $\phi$ | | | | | | | | | | | | | | |
| $\pi/16$ | 0.052 | 1.000 | 1.000 | 0.066 | 0.313 | 0.418 | 0.059 | 0.000 | 0.343 | 0.014 | 0.000 | 0.119 | 0.001 | 0.001 |
| $\pi/8$ | 0.048 | 1.000 | 1.000 | 0.060 | 0.337 | 0.926 | 0.918 | 0.000 | 0.917 | 0.907 | 0.000 | 0.874 | 0.863 | 0.001 |
| $\pi/4$ | 0.048 | 1.000 | 1.000 | 0.061 | 0.376 | 0.972 | 0.995 | 0.000 | 0.967 | 0.994 | 0.000 | 0.947 | 0.989 | 0.000 |
| $\pi/2$ | 0.047 | 0.228 | 1.000 | 0.058 | 0.483 | 1.000 | 1.000 | 0.000 | 1.000 | 1.000 | 0.000 | 1.000 | 1.000 | 0.000 |
| | | | | | | | $c_\phi = -2.5$ | | | | | | | |
| $\phi$ | | | | | | | | | | | | | | |
| $\pi/16$ | 0.049 | 1.000 | 1.000 | 0.063 | 0.162 | 0.106 | 0.003 | 0.000 | 0.084 | 0.002 | 0.000 | 0.025 | 0.002 | 0.002 |
| $\pi/8$ | 0.049 | 1.000 | 1.000 | 0.062 | 0.185 | 0.751 | 0.729 | 0.000 | 0.734 | 0.714 | 0.000 | 0.648 | 0.620 | 0.002 |
| $\pi/4$ | 0.048 | 1.000 | 1.000 | 0.061 | 0.210 | 0.880 | 0.968 | 0.000 | 0.869 | 0.966 | 0.000 | 0.812 | 0.945 | 0.002 |
| $\pi/2$ | 0.049 | 0.174 | 1.000 | 0.062 | 0.283 | 0.998 | 0.999 | 0.000 | 0.998 | 0.998 | 0.000 | 0.997 | 0.997 | 0.001 |
| | | | | | | | $c_\phi = -5$ | | | | | | | |
| $\phi$ | | | | | | | | | | | | | | |
| $\pi/16$ | 0.049 | 1.000 | 1.000 | 0.062 | 0.111 | 0.031 | 0.001 | 0.000 | 0.026 | 0.001 | 0.000 | 0.014 | 0.005 | 0.004 |
| $\pi/8$ | 0.049 | 1.000 | 1.000 | 0.062 | 0.132 | 0.572 | 0.541 | 0.000 | 0.556 | 0.525 | 0.000 | 0.459 | 0.428 | 0.004 |
| $\pi/4$ | 0.049 | 1.000 | 1.000 | 0.064 | 0.155 | 0.767 | 0.919 | 0.000 | 0.755 | 0.911 | 0.000 | 0.675 | 0.873 | 0.004 |
| $\pi/2$ | 0.046 | 0.148 | 1.000 | 0.059 | 0.202 | 0.995 | 0.996 | 0.000 | 0.994 | 0.995 | 0.000 | 0.989 | 0.989 | 0.002 |

*Table continues on overleaf*

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

**TABLE 1.** *continued*

| | $t_{\widehat{\rho}}$ | $t_{\widehat{\alpha}_1}$ | $t_{\widehat{\alpha}_2}$ | $Z_{\widehat{\rho}}$ | $VRT$ | $MSB^b$ | $MSB^q$ | $MSB^{AR}$ | $MZ_\alpha^b$ | $MZ_\alpha^q$ | $MZ_\alpha^{AR}$ | $MZ_t^b$ | $MZ_t^q$ | $MZ_t^{AR}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | $c_\phi = -10$ | | | | | | | |
| $\phi$ | | | | | | | | | | | | | | |
| $\pi/16$ | 0.049 | 1.000 | 1.000 | 0.059 | 0.068 | 0.005 | 0.000 | 0.003 | 0.006 | 0.000 | 0.003 | 0.011 | 0.009 | 0.009 |
| $\pi/8$ | 0.050 | 1.000 | 1.000 | 0.062 | 0.095 | 0.315 | 0.281 | 0.002 | 0.301 | 0.267 | 0.002 | 0.225 | 0.195 | 0.010 |
| $\pi/4$ | 0.050 | 1.000 | 1.000 | 0.063 | 0.110 | 0.565 | 0.781 | 0.001 | 0.547 | 0.771 | 0.001 | 0.458 | 0.703 | 0.009 |
| $\pi/2$ | 0.049 | 0.119 | 1.000 | 0.062 | 0.143 | 0.972 | 0.972 | 0.000 | 0.969 | 0.970 | 0.000 | 0.948 | 0.950 | 0.004 |
| | | | | | | | $c_\phi = -20$ | | | | | | | |
| $\phi$ | | | | | | | | | | | | | | |
| $\pi/16$ | 0.053 | 1.000 | 1.000 | 0.064 | 0.044 | 0.001 | 0.000 | 0.018 | 0.001 | 0.000 | 0.018 | 0.013 | 0.017 | 0.022 |
| $\pi/8$ | 0.049 | 1.000 | 1.000 | 0.061 | 0.064 | 0.098 | 0.078 | 0.010 | 0.097 | 0.077 | 0.012 | 0.070 | 0.058 | 0.020 |
| $\pi/4$ | 0.045 | 1.000 | 1.000 | 0.057 | 0.076 | 0.324 | 0.521 | 0.006 | 0.313 | 0.509 | 0.007 | 0.240 | 0.431 | 0.015 |
| $\pi/2$ | 0.047 | 0.087 | 1.000 | 0.060 | 0.098 | 0.869 | 0.878 | 0.001 | 0.861 | 0.870 | 0.001 | 0.811 | 0.820 | 0.010 |

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press



**TABLE 2.** Empirical rejection frequencies, OLS linear Detrended data, DGP (13)

| | $t_{\widehat{\rho}}$ | $t_{\widehat{\alpha}_1}$ | $t_{\widehat{\alpha}_2}$ | $Z_{\widehat{\rho}}$ | $VRT$ | $MSB^b$ | $MSB^q$ | $MSB^{AR}$ | $MZ_\alpha^b$ | $MZ_\alpha^q$ | $MZ_\alpha^{AR}$ | $MZ_t^b$ | $MZ_t^q$ | $MZ_t^{AR}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | $c_\phi = 0$ | | | | | | | |
| 0 | 0.046 | 0.052 | 0.051 | 0.076 | 0.052 | 0.072 | 0.073 | 0.061 | 0.042 | 0.043 | 0.034 | 0.034 | 0.034 | 0.029 |
| $\phi$ | | | | | | | | | | | | | | |
| $\pi$ /16 | 0.050 | 1.000 | 1.000 | 0.090 | 0.547 | 0.189 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| $\pi/8$ | 0.050 | 1.000 | 1.000 | 0.080 | 0.639 | 0.977 | 0.973 | 0.000 | 0.961 | 0.953 | 0.000 | 0.936 | 0.925 | 0.000 |
| $\pi/4$ | 0.049 | 1.000 | 1.000 | 0.080 | 0.693 | 0.996 | 1.000 | 0.000 | 0.992 | 0.999 | 0.000 | 0.984 | 0.999 | 0.000 |
| $\pi/2$ | 0.049 | 0.474 | 1.000 | 0.078 | 0.812 | 1.000 | 1.000 | 0.000 | 1.000 | 1.000 | 0.000 | 1.000 | 1.000 | 0.000 |
| | | | | | | | $c_\phi = -2.5$ | | | | | | | |
| $\phi$ | | | | | | | | | | | | | | |
| $\pi/16$ | 0.048 | 1.000 | 1.000 | 0.085 | 0.267 | 0.022 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| $\pi/8$ | 0.049 | 1.000 | 1.000 | 0.083 | 0.378 | 0.903 | 0.885 | 0.000 | 0.847 | 0.820 | 0.000 | 0.773 | 0.738 | 0.000 |
| $\pi/4$ | 0.048 | 1.000 | 1.000 | 0.077 | 0.427 | 0.977 | 0.998 | 0.000 | 0.959 | 0.996 | 0.000 | 0.932 | 0.992 | 0.000 |
| $\pi/2$ | 0.049 | 0.380 | 1.000 | 0.080 | 0.579 | 1.000 | 1.000 | 0.000 | 1.000 | 1.000 | 0.000 | 1.000 | 1.000 | 0.000 |
| | | | | | | | $c_\phi = -5$ | | | | | | | |
| $\phi$ | | | | | | | | | | | | | | |
| $\pi/16$ | 0.051 | 1.000 | 1.000 | 0.082 | 0.153 | 0.005 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 |
| $\pi/8$ | 0.050 | 1.000 | 1.000 | 0.081 | 0.250 | 0.768 | 0.729 | 0.000 | 0.666 | 0.622 | 0.000 | 0.560 | 0.508 | 0.001 |
| $\pi/4$ | 0.048 | 1.000 | 1.000 | 0.078 | 0.302 | 0.937 | 0.993 | 0.000 | 0.895 | 0.986 | 0.000 | 0.840 | 0.974 | 0.000 |
| $\pi/2$ | 0.046 | 0.315 | 1.000 | 0.078 | 0.438 | 1.000 | 1.000 | 0.000 | 1.000 | 1.000 | 0.000 | 1.000 | 1.000 | 0.000 |

*Table continues on overleaf*

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

**TABLE 2.** *continued*

| | $t_{\widehat{\rho}}$ | $t_{\widehat{\alpha}_1}$ | $t_{\widehat{\alpha}_2}$ | $Z_{\widehat{\rho}}$ | $VRT$ | $MSB^b$ | $MSB^q$ | $MSB^{AR}$ | $MZ_\alpha^b$ | $MZ_\alpha^q$ | $MZ_\alpha^{AR}$ | $MZ_t^b$ | $MZ_t^q$ | $MZ_t^{AR}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | $c_\phi = -10$ | | | | | | | |
| $\phi$ | | | | | | | | | | | | | | |
| $\pi/16$ | 0.050 | 1.000 | 1.000 | 0.080 | 0.069 | 0.001 | 0.000 | 0.002 | 0.000 | 0.000 | 0.000 | 0.002 | 0.001 | 0.001 |
| $\pi/8$ | 0.048 | 1.000 | 1.000 | 0.077 | 0.148 | 0.453 | 0.397 | 0.001 | 0.331 | 0.281 | 0.000 | 0.242 | 0.200 | 0.001 |
| $\pi/4$ | 0.045 | 1.000 | 1.000 | 0.074 | 0.186 | 0.797 | 0.955 | 0.000 | 0.708 | 0.924 | 0.000 | 0.612 | 0.886 | 0.001 |
| $\pi/2$ | 0.045 | 0.233 | 1.000 | 0.075 | 0.282 | 0.999 | 1.000 | 0.000 | 0.999 | 0.999 | 0.000 | 0.997 | 0.997 | 0.000 |
| | | | | | | | $c_\phi = -20$ | | | | | | | |
| $\phi$ | | | | | | | | | | | | | | |
| $\pi/16$ | 0.053 | 1.000 | 1.000 | 0.080 | 0.030 | 0.000 | 0.000 | 0.016 | 0.000 | 0.000 | 0.006 | 0.003 | 0.003 | 0.008 |
| $\pi/8$ | 0.050 | 1.000 | 1.000 | 0.078 | 0.083 | 0.139 | 0.110 | 0.006 | 0.080 | 0.060 | 0.003 | 0.055 | 0.042 | 0.006 |
| $\pi/4$ | 0.049 | 1.000 | 1.000 | 0.079 | 0.118 | 0.518 | 0.775 | 0.003 | 0.407 | 0.693 | 0.001 | 0.316 | 0.610 | 0.003 |
| $\pi/2$ | 0.048 | 0.161 | 1.000 | 0.078 | 0.179 | 0.987 | 0.989 | 0.000 | 0.977 | 0.980 | 0.000 | 0.961 | 0.966 | 0.001 |

generate cycles of duration 32 and 16 quarters for quarterly data ($\pi/16$ and $\pi/8$, respectively) and the equivalent lengths of 8 and 4 years for annual data ($\pi/4$ and $\pi/2$, respectively). Burns and Mitchell (1946) argue that business cycle durations typically vary between in excess of 1 year up to 10 or 12 years. Results for other values of $\phi$ can additionally be found in the accompanying working paper, del Barrio Castro, Rodrigues, and Taylor (2011) (CRT). As regards the selected values for $c_\phi$, we have found that the values $-2.5$ and $-5$ are in line with the periodogram of quarterly U.S. real gross domestic product computed for a sample starting in the first quarter of 1947 and ending in the fourth quarter of 2011, whereas the other values are included to illustrate the effects of lower levels of persistence and of pure cyclical unit root behavior. For comparative purposes, results for the conventional random walk, $(1-L)x_t = \epsilon_t \sim NID(0,1)$, initialized at $x_0 = 0$, are also provided in the rows labeled 0.

The results in Tables 1 and 2 report the empirical (null) rejection frequencies of the unit root tests from Section 2.2 in each case for a nominal 5% significance level using the asymptotic critical values appropriate to the case where (near-) nonstationary cycles are not present in the data; that is, from $\mathcal{DF}_1$, $\mathcal{DF}_2$, $\mathcal{MSB}$, or $\mathcal{VRT}$, as appropriate. In Table 1, the ADF $t_{\hat{\rho}}$ and $Z_{\hat{\rho}}$ tests were computed from the ADF regression (2) for the true lag length, $k = 2$. In the context of the trinity of $M$ tests, the additional superscript used in the nomenclature of Tables 1 and 2 denotes the long-run variance estimator used; the superscripts $b$ and $q$ indicate that $\hat{\lambda}^2 := s^2_{WA}$ with the Bartlett and quadratic spectral kernels, respectively (using the data-dependent bandwidth formulations for these kernels suggested in Newey and West, 1994, eqns. (3.8)–(3.15) and Table 1), whereas the superscript $AR$ indicates that $\hat{\lambda}^2 := s^2_{AR}$ (calculated for $k = 2$). Also reported in Tables 1 and 2 are the empirical rejection frequencies for the conventional $t$-ratio tests on the lagged dependent variables $\Delta x_{t-1}$ and $\Delta x_{t-2}$ from (2), denoted $t_{\hat{a}_1}$ and $t_{\hat{a}_2}$, in each case compared to the 0.05 level critical value from the standard normal distribution (a 5% rule). As predicted by the results in Theorem 1, it is only the ADF $t_{\hat{\rho}}$ and $Z_{\hat{\rho}}$ tests that display finite-sample size properties that are robust to the presence of (near-) integrated cycles in the data. The size properties of the $t_{\hat{\rho}}$ test are somewhat better than those of $Z_{\hat{\rho}}$, which is a little oversized, most notably so in the case of linear detrended data, but both show no significant variations in size from the random walk base case under nonstationary or near nonstationary cycles for all of the frequencies considered.

Also in line with the predictions from the asymptotic theory, we see that this is not the case for the other tests considered. As predicted by the results in Theorem 1 and Remark 2, the empirical sizes of the $VRT$ test vary considerably across both $\phi$ and $c_\phi$. As might be expected, for a given frequency, $\phi$, the size distortions in $VRT$ decrease as $c_\phi$ increases; this is because the cyclical component at frequency $\phi$ moves further away from the nonstationarity boundary as $c_\phi$ increases. For each of the three $M$ tests based on an AR spectral density estimator empirical size is always very close to zero for all of the values of $\phi$ and $c_\phi$ considered. Overall though, even though the $VRT$ and $M$ tests with $\hat{\lambda}^2 := s^2_{AR}$ are

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

not asymptotically degenerate under (near-) nonstationary cycles, the results in Tables 1 and 2 suggest that none of these tests could be reliably used in practice. The degeneracy of the kernel-based $M$ tests is clearly reflected in Tables 1 and 2, for $\phi \in \{\pi/8, \pi/4, \pi/2\}$ with very high empirical rejection frequencies close to unity in a number of cases. Notice that, as expected, these oversizing effects tend to decrease as $c_\phi$ and $\phi$ increase and decrease, respectively. Interestingly, for $\phi = \pi/16$ these patterns differ with undersizing seen for most cases with the Bartlett kernel and throughout for the tests based on the quadratic spectral kernel. This appears attributable to the fact that for $\pi/16$ only six complete cycles are completed for this sample size, making it hard to distinguish this component from the zero frequency. This phenomenon is, however, quickly reversed as the sample size is increased; for example, for $c_\phi = 0$ and $n = 400$ in the linear detrended data case, the empirical sizes of the $MSB$, $MZ_\alpha$, and $MZ_t$ tests with the Bartlett kernel are 0.950, 0.914 and 0.858, respectively, and 0.854, 0.749, and 0.587, respectively, for the quadratic spectral kernel.

Finally, we observe that the empirical rejection frequency of the $t_{\hat{\alpha}_2}$ test with a 5% rule is unity throughout (except in the random walk case where it is essentially correctly sized), implying that in the presence of (near-) nonstationary cycles the sequential method of Hall (1994) and Ng and Perron (1995) will always retain both lagged dependent variables in (2), as would be hoped. Indeed, we obtained the same outcome when a tighter 1% rule was used. More generally, in unreported simulations we found that in the case outlined in Remark 7, where $\Psi(L)$ is a $p$th-order polynomial of nonstationary factors, the same finding is valid for the $t$-test on the $p$th lag, so that $p$ lagged dependent variables would always be included in (2).

In the accompanying working paper, CRT, the impact of a number of conventional lag selection methods on the empirical size and power properties of the $t_{\widehat{\rho}}$ test is investigated both for OLS and local GLS detrended data. In particular, results are reported for the sequential method (SQ) outlined in Section 2.2 using a 10% rule, the standard AIC rule, and also the modified AIC (MAIC) rule of Ng and Perron (2001). The results reported there suggest that the empirical size of the ADF test with OLS detrending is reasonably close to the nominal level throughout for both SQ and AIC. In the case of the MAIC rule, a degree of undersizing is seen throughout, perhaps not unexpectedly, given that the penalty function on which the MAIC rule is based is misspecified when nonstationary cycles are present. Similar comments apply when local GLS detrending is used, although here tests based on MAIC can be very badly undersized in the case of linear detrending; indeed all tests display a tendency to undersize here. The results in CRT also show that the empirical power of $t_{\widehat{\rho}}$ varies little with the frequency at which the nonstationary cyclical roots occur. For the OLS detrended tests power is also fairly insensitive as to whether a nonstationary cycle occurs or not. The same is not true for the local GLS detrended tests (which are based on the assumption that $\Delta x_t$ is stochastically stationary), which show considerably lower power for $c_\phi = 0$ than for $c_\phi < 0$, other things equal.

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

## 5. CONCLUSIONS

In this paper we have shown that among popularly applied unit root test statistics, only the ADF $t$-ratio and normalized bias statistics have pivotal limiting null distributions in the presence of (near-) nonstationary cycles in the data. Other commonly employed unit root statistics, such as the variance ratio statistic of Breitung (2002) and the trinity of $M$ statistics of Stock (1999) and Perron and Ng (1996), were shown either to admit nonpivotal limiting null distributions or to have nondegenerate limiting null distributions, in the latter case yielding tests with an asymptotic size of one, when (near-) nonstationary cycles are present. Additionally, we have shown that the $t$-ratios on the lagged dependent variables within the ADF test regression also retain standard normal limiting null distributions such that sequential lag specification remains valid under (near-) nonstationary cycles. Consequently, we strongly recommend the use of ADF unit root tests coupled with sequential lag selection in cases where it is suspected that (near-) nonstationary cycles may be present in the data.

## *REFERENCES*

Ahtola, J.A. & G.C. Tiao (1987) Distribution of the least squares estimators of autoregressive parameters for a process with complex roots on the unit circle. *Journal of Time Series Analysis* 8, 1–14.

Alesina, A. & N. Roubini (1992) Politics cycles in OECD countries. *Review of Economic Studies* 59, 663–688.

Allen, D.S. (1997) Filtering Permanent Cycles with Complex Unit Roots. Working paper 1997-001A, Federal Reserve Bank of St. Louis (http://research.stlouisfed.org/wp/1997/97-001.pdf).

Azariadis, C., J. Bullard, & L. Ohanian (2004) Trend-reverting fluctuations in the life-cycle model. *Journal of Economic Theory* 119, 334–356.

Berk, K.N. (1974) Consistent autoregressive spectral estimates. *Annals of Statistics* 2, 489–502.

Bierens, H.J. (2000) Complex Unit Roots and Business Cycles: Are They Real? Separate appendix. http://econ.la.psu.edu/~hbierens/PAPERS.HTM.

Bierens, H.J. (2001) Complex unit roots and business cycles: Are they real? *Econometric Theory* 17, 962–983.

Birgean, L. & L. Kilian (2002) Data-driven nonparametric spectral density estimators for economic time series: A Monte Carlo study. *Econometric Reviews* 21, 449–476.

Breitung, J. (2002) Nonparametric tests for unit roots and cointegration. *Journal of Econometrics* 108, 343–363.

Burns, A.F. & W.C. Mitchell (1946) *Measuring Business Cycles*. National Bureau of Economic Research.

Campbell, J. & A. Kyle (1993) Smart money, noise trading and stock price behaviour. *Review of Economics Studies* 60, 1–34.

Canova, F. (1996) Three tests for the existence of cycles in time series. *Ricerche Economiche* 50, 135–162.

Chan, N.H. (1989) Asymptotic inference for unstable autoregressive time series with drifts. *Journal of Statistical Planning and Inference* 23, 301–312.

Chan, N.H. & C.Z. Wei (1988) Limiting distribution of least squares estimates of unstable autoregressive processes. *Annals of Statistics* 16, 367–401.

Chang, Y. & J.Y. Park (2002) On the asymptotics of ADF tests for unit roots. *Econometric Reviews* 21, 431–448.

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

Choi, I. (1993) Asymptotic normality of the least squares estimates for higher order autoregressive integrated processes with some applications. *Econometric Theory* 9, 263–282.
Conway, P. & D. Frame (2000) A Spectral Analysis of New Zealand Output Gaps Using Fourier and Wavelet Techniques. Discussion paper DP2000/06, Reserve Bank of New Zealand.
del Barrio Castro, T., P.M.M. Rodrigues, & A.M.R. Taylor (2011) The impact of persistent cycles on zero frequency unit root tests, Working paper 24/2011, Banco de Portugal.
Díaz-Emparanza, I. (2004) A note on the paper by H.J. Bierens: "Complex unit roots and business cycles: are they real?" *Econometric Theory* 20, 636–637.
Dickey, D.A. & W.A. Fuller (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427–431.
Elliott, G., T.J. Rothenberg, & J.H. Stock (1996) Efficient tests for an autoregressive unit root. *Econometrica* 64, 813–836.
Estrella, A. (2004) The cyclical behavior of optimal bank capital. *Journal of Banking and Finance* 28, 1469–1498.
Faust, J. (1996) Near observational equivalence and theoretical size problems with unit root tests. *Econometric Theory* 12, 724–731.
Ghysels, E., H.S. Lee, & J. Noh (1994) Testing for unit roots in seasonal time series: Some theoretical extensions and a Monte Carlo investigation. *Journal of Econometrics* 62, 415–442.
Gregoir, S. (2004) Efficient tests for the presence of a pair of complex conjugate unit roots in real time series. *Journal of Econometrics* 130, 45–100.
Haldrup, N. & M. Jansson (2006) Improving power and size in unit root testing. In T.C. Mills & K. Patterson (eds.), *Palgrave Handbooks of Econometrics*, vol. 1, chap. 7. Palgrave MacMillan, pp. 252–277.
Hall, A.R. (1994) Testing for a unit root in time series with pretest data-based model selection. *Journal of Business & Economic Statistics* 12, 461–470.
Hamilton, J.D. (1994) *Time Series Analysis*. Princeton University Press.
Jansson, M. (2002) Consistent covariance matrix estimation for linear processes. *Econometric Theory* 18, 1449–1459.
Jeganathan, P. (1991) On the asymptotic behavior of least-squares estimators in AR time series with roots near the unit circle. *Econometric Theory* 7, 269–306.
Maddala, G.S. & I. Kim (1998) *Unit Roots, Cointegration, and Structural Change.* Cambridge University Press.
Müller, U. (2008) The impossibility of consistent discrimination between I(0) and I(1) processes. *Econometric Theory* 24, 616–630.
Müller, U.K. & G. Elliott (2003) Tests for unit roots and the initial condition. *Econometrica* 71, 1269–1286.
Newey, W.K. & K.D. West (1994) Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61, 631–653.
Ng, S. & P. Perron (1995) Unit root tests in ARMA models with data dependent methods for selection of the truncation lag. *Journal of the American Statistical Association* 90, 268–281.
Ng, S. & P. Perron (2001) Lag length selection and the construction of unit root tests with good size and power. *Econometrica* 69, 1519–1554.
Nielsen, B. (2001) The asymptotic distribution of unit root tests of unstable autoregressive processes. *Econometrica* 69, 211–219.
Nielsen, B. (2006) Order determination in general vector autoregressions. In H.-C. Ho, C.-K. Ing, & T.L. Lai (eds.), *Time Series and Related Topics: In Memory of Ching-Zong Wei*, IMS Lecture Notes and Monograph Series 52, 93–112.
Nyarko, Y. (1992) Learning in misspecified models and the possibility of cycles. *Journal of Economic Theory* 55, 416–427.
Pagan, A. (1997) Towards understanding some business cycle characteristics. *Australian Economic Review* 30, 1–15.
Perron, P. & Ng, S. (1996) Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. *Review of Economic Studies* 63, 435–463.

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

Phillips, P.C.B. (1987) Time series regression with a unit root, *Econometrica* 55, 277–301.
Phillips, P.C.B. & P. Perron (1988) Testing for a unit root in time series regression. *Biometrika* 75, 335–346.
Phillips, P.C.B. & Z. Xiao (1998) A primer on unit root testing. *Journal of Economic Surveys* 12, 423–470.
Priestley, M.B. (1981) *Spectral Analysis and Time Series*. Academic Press.
Rodrigues, P.M.M. (1999) A note on the application of the DF test to seasonal data. *Statistics and Probability Letters* 47, 171–175.
Said, S.E. & D.A. Dickey (1984) Testing for unit roots in autoregressive–moving average models of unknown order. *Biometrika* 71, 599–607.
Shibayama, K. (2008) On the periodicity of inventories. Department of Economics, University of Kent at Canterbury.
Stock, J. (1994) Unit roots, structural breaks and trends. In *Handbook of Econometrics,* vol. IV, chap. 46, pp. 2739–2841.
Stock, J.H. (1999) A class of tests for integration and cointegration. In R.F. Engle & H. White (eds.), *Cointegration, Causality and Forecasting. A Festschrift in Honour of Clive W.J. Granger*, pp. 137–167. Oxford University Press.
Tanaka, K. (2008) Analysis of models with complex roots on the unit circle. *Journal of the Japanese Statistical Society* 38, 145–155.
Taylor, A.M.R. (2003) Robust stationarity tests in seasonal time series processes. *Journal of Business & Economic Statistics* 21, 156–163.
Tsay, R.S. & G.C. Tiao (1990) Asymptotic properties of multivariate nonstationary processes with applications to autoregressions. *Annals of Statistics* 12, 321–364.

# APPENDIX

It will prove useful in what follows to consider first frequency specific orthogonal decompositions of $x_t$ and $\Delta x_t$. These are collected in Lemma A.1, the proof of which, along with those of Lemmas A.2 and A.3, can be found in CRT. Lemma A.1 builds on results in Bierens (2001, pp. 963–964) appropriate for the case where $\varphi_{0n} = c_\phi = 0$ in (1).

LEMMA A.1. *Let the conditions of Theorem 1 hold. Then for any* $\phi \in (0, \pi)$ *we may state the following decompositions:*

$$n^{-1/2}x_t = \frac{\delta_0}{n^{1/2}} S_{0,c_0}(t) + \frac{\delta_\phi^\alpha}{n^{1/2}\sin\phi}\left\{\sin[\phi(t+1)]\, S_{\phi,c_\phi}^\alpha(t) - \cos[\phi(t+1)]\, S_{\phi,c_\phi}^\beta(t)\right\}$$
$$+ \frac{\delta_\phi^\beta}{n^{1/2}\sin\phi}\left\{\cos[\phi(t+1)] S_{\phi,c_\phi}^\alpha(t) + \sin[\phi(t+1)]\, S_{\phi,c_\phi}^\beta(t)\right\} + O_p\left(n^{-1/2}\right), \quad \textbf{(A.1)}$$

$$n^{-1/2}\Delta x_t = \frac{1}{n^{1/2}\sin\phi}\left\{\sin[\phi(t+1)]\, S_{\phi,c_\phi}^\alpha(t) - \cos[\phi(t+1)]\, S_{\phi,c_\phi}^\beta(t)\right\} + O_p\left(n^{-1/2}\right), \quad \textbf{(A.2)}$$

*where* $S_{0,c_0}(t) := \sum_{j=1}^t \varphi_{0n}^{t-j}\varepsilon_j$, $S_{\phi,c_\phi}^\alpha(t) := \sum_{j=1}^t \varphi_{\phi n}^{t-j}\varepsilon_j \cos(j\phi)$, $S_{\phi,c_\phi}^\beta(t) := \sum_{j=1}^t \varphi_{\phi n}^{t-j}\varepsilon_j \sin(j\phi)$, $\delta_0 := \frac{1}{2}(1-\cos\phi)$, $\delta_\phi^\alpha := (1 - 2\cos\phi + (\varphi_{0n}\cos\phi)/\varphi_{\phi n})/2(1-\cos\phi)$, $\delta_\phi^\beta := -\sin(\phi)\,\varphi_{0n}/2(1-\cos\phi)\,\varphi_{\phi n}$, *and the* $O_p\left(n^{-1/2}\right)$ *terms in (A.1) and (A.2) are both uniform in* $t = 1, \ldots, n$.

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

**Remark A.1.** It is straightforward to show that the results stated in Lemma A.1 continue to hold under the weaker linear process conditions on $\{\varepsilon_t\}$ given in Remark 7.

Next in Lemma A.2 we provide some additional results that will be needed for establishing the limiting distributions of the unit root statistics from Section 2.2.

LEMMA A.2. *Let the conditions of Theorem 1 hold. Then, for any $\phi \in (0, \pi)$,*

$$(i)\quad n^{-2}\sum_{t=1}^{n} x_{t-1}^2 = \frac{1}{4n^2(1-\cos\phi)^2}\sum_{t=1}^{n}\left(S_{0,c_0}(t-1)\right)^2 + \frac{1}{2n^2(1-\cos\phi)} \times \frac{1}{2\sin^2\phi}\sum_{t=1}^{n}\left[\left(S^{\alpha}_{\phi,c_\phi}(t-1)\right)^2 + \left(S^{\beta}_{\phi,c_\phi}(t-1)\right)^2\right] + o_p(1),$$

$$(ii)\quad n^{-2}\sum_{t=1}^{n}\Delta x_{t-j}^2 = n^{-2}\sum_{t=1}^{n}\frac{\left(S^{\alpha}_{\phi,c_\phi}(t-j)\right)^2 + \left(S^{\beta}_{\phi,c_\phi}(t-j)\right)^2}{2\sin^2\phi} + o_p(1), \quad j=0,1,2,$$

$$(iii)\quad n^{-2}\sum_{t=1}^{n} x_{t-1}\Delta x_{t-1} = \frac{1}{4n^2\sin^2\phi}\sum_{t=1}^{n}\left[\left(S^{\alpha}_{\phi,c_\phi}(t-1)\right)^2 + \left(S^{\beta}_{\phi,c_\phi}(t-1)\right)^2\right] + o_p(1),$$

$$(iv)\quad n^{-2}\sum_{t=1}^{n} x_{t-1}\Delta x_{t-2} = \frac{\cos\phi+\frac{1}{2}}{2n^2\sin^2\phi}\sum_{t=1}^{n}\left[\left(S^{\alpha}_{\phi,c_\phi}(t-1)\right)^2 + \left(S^{\beta}_{\phi,c_\phi}(t-1)\right)^2\right] + o_p(1).$$

Using the results provided in Lemma A.2, we are now in a position to state the large-sample behavior of the elements that form the unit root tests under analysis in the near-integrated context considered. These results are collected in Lemma A.3.

LEMMA A.3. *Let the conditions of Theorem 1 hold. Then, for any $\phi \in (0, \pi)$,*

$$n^{-2}\sum_{t=1}^{n} x_{t-1}^2 \Rightarrow \frac{\sigma^2\int_0^1 W_{0,c_0}^2(r)dr}{4(1-\cos\phi)^2} + \frac{\sigma^2\int_0^1\left(\left[W^{\alpha}_{\phi,c_\phi}(r)\right]^2 + \left[W^{\beta}_{\phi,c_\phi}(r)\right]^2\right)dr}{8(1-\cos\phi)\sin^2\phi},$$

$$n^{-2}\sum_{t=1}^{n}\Delta x_t^2 = \frac{1}{n^2}\sum_{t=1}^{n}\Delta x_{t-1}^2 + o_p(1) = \frac{1}{n^2}\sum_{t=1}^{n}\Delta x_{t-2}^2 + o_p(1) \Rightarrow \frac{\sigma^2\int_0^1\left(\left[W^{\alpha}_{\phi,c_\phi}(r)\right]^2 + \left[W^{\beta}_{\phi,c_\phi}(r)\right]^2\right)dr}{4\sin^2\phi},$$

$$n^{-2}\sum_{t=1}^{n} x_{t-1}\Delta x_{t-1} \Rightarrow \frac{\sigma^2\int_0^1\left(\left[W^{\alpha}_{\phi,c_\phi}(r)\right]^2 + \left[W^{\beta}_{\phi,c_\phi}(r)\right]^2\right)dr}{8\sin^2\phi},$$

$$n^{-2}\sum_{t=1}^{n} x_{t-1}\Delta x_{t-2} \Rightarrow \frac{\sigma^2\left(\cos\phi+\frac{1}{2}\right)\int_0^1\left(\left[W^{\alpha}_{\phi,c_\phi}(r)\right]^2 + \left[W^{\beta}_{\phi,c_\phi}(r)\right]^2\right)dr}{4\sin^2\phi},$$

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

$$n^{-1}\sum_{t=1}^{n} x_{t-1}\varepsilon_t \Rightarrow \frac{\sigma^2 \int_0^1 W_{0,c_0}(r)dW_{0,c_0}(r)}{2(1-\cos\phi)} + \frac{\sigma^2 \int_0^1 \left(W^{\alpha}_{\phi,c_\phi}(r)dW^{\beta}_{\phi,c_\phi}(r) - W^{\beta}_{\phi,c_\phi}(r)dW^{\alpha}_{\phi,c_\phi}(r)\right)}{4\sin\phi} - \frac{\sigma^2 \int_0^1 \left(W^{\alpha}_{\phi,c_\phi}(r)\,dW^{\alpha}_{\phi,c_\phi}(r) + W^{\beta}_{\phi,c_\phi}(r)dW^{\beta}_{\phi,c_\phi}(r)\right)}{4(1-\cos(\phi))},$$

$$n^{-1}\sum_{t=1}^{n} \Delta x_{t-1}\varepsilon_t \Rightarrow \frac{\sigma^2 \int_0^1 \left(W^{\alpha}_{\phi,c_\phi}(r)\,dW^{\beta}_{\phi,c_\phi}(r) - W^{\beta}_{\phi,c_\phi}(r)dW^{\alpha}_{\phi,c_\phi}(r)\right)}{2\sin\phi},$$

$$n^{-1}\sum_{t=1}^{n} \Delta x_{t-2}\varepsilon_t \Rightarrow \frac{\sigma^2\cos\phi}{2\sin\phi}\int_0^1 \left(W^{\alpha}_{\phi,c_\phi}(r)dW^{\beta}_{\phi,c_\phi}(r) - W^{\beta}_{\phi,c_\phi}(r)dW^{\alpha}_{\phi,c_\phi}(r)\right) - \frac{\sigma^2}{2}\int_0^1 \left(W^{\alpha}_{\phi,c_\phi}(r)dW^{\alpha}_{\phi,c_\phi}(r) + W^{\beta}_{\phi,c_\phi}(r)dW^{\beta}_{\phi,c_\phi}(r)\right),$$

*where* $W_{0,c_0}(r)$, $W^{\alpha}_{\phi,c_\phi}(r)$, *and* $W^{\beta}_{\phi,c_\phi}(r)$ *are independent standard OU processes.*

We are now in a position to prove our main results.

**Proof of Theorems 1 and 2.** First we define $\mathbf{z}_{t-1} := [x_{t-1}\ \Delta x_{t-1}\ \Delta x_{t-2}\ \Delta x_{t-3}\ \ldots\ \Delta x_{t-k}]'$ and $\beta := (\rho, \alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_k)'$. Hence, to analyze the convergence of the OLS parameter estimates from (2) we will make use of the following expression for the OLS estimation error:

$$\widehat{\beta}-\beta = \left[\sum_{t=1}^{n} \mathbf{z}_{t-1}\mathbf{z}'_{t-1}\right]^{-1}\left[\sum_{t=1}^{n} \mathbf{z}_{t-1}\varepsilon_t\right]. \tag{A.3}$$

Notice that under the conditions of Theorems 1 and 2, the true values of the parameters satisfy $\beta \overset{a}{=} (0, 2\cos\phi, -1, 0, \ldots, 0)'$, where $\overset{a}{=}$ denotes asymptotic equivalence. Moreover, using a similar approach to Choi (1993), we define the $(k+1)\times(k+1)$ matrix

$$\mathbf{A} := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & -2\cos\phi & 1 & \cdots & 0 \\ 0 & 0 & 0 & 1 & -2\cos\phi & \cdots & 0 \\ 0 & 0 & 0 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & -2\cos\phi & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2\cos\phi \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} =: [\mathbf{A}_1\ \mathbf{A}_2],$$

where $\mathbf{A}_1$ is a $(k+1)\times 3$ matrix and $\mathbf{A}_2$ a $(k+1)\times(k-2)$ matrix. Notice that $\mathbf{A}_2$ is a filtration matrix because $\mathbf{A}'_2\mathbf{z}_{t-1} = \left(\Delta_\phi\Delta x_{t-1}, \Delta_\phi\Delta x_{t-2}, \ldots, \Delta_\phi\Delta x_{t-k}\right)'$. Using the matrix $\mathbf{A}$ and introducing the scaling matrix $\Upsilon_n := \mathrm{diag}\{\Upsilon_{1n}, \Upsilon_{2n}\}$ where $\Upsilon_{1n} := \mathrm{diag}\{n, n, n\}$ and $\Upsilon_{2n} := \mathrm{diag}\{\sqrt{n}, \ldots, \sqrt{n}\}$, the latter a $(k-2)\times(k-2)$ matrix, we can

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

rewrite the scaled estimator from (A.3) as

$$\begin{aligned}\Upsilon_n\left(\widehat{\beta}-\beta\right) &= \mathbf{A}\left\{\Upsilon_n^{-1}\begin{bmatrix}\sum_{t=1}^{n}\mathbf{A}_1'\mathbf{z}_{t-1}\mathbf{z}_{t-1}'\mathbf{A}_1 & \sum_{t=1}^{n}\mathbf{A}_1'\mathbf{z}_{t-1}\mathbf{z}_{t-1}'\mathbf{A}_2\\ \sum_{t=1}^{n}\mathbf{A}_2'\mathbf{z}_{t-1}\mathbf{z}_{t-1}'\mathbf{A}_1 & \sum_{t=1}^{n}\mathbf{A}_2'\mathbf{z}_{t-1}\mathbf{z}_{t-1}'\mathbf{A}_2\end{bmatrix}^{-1}\Upsilon_n^{-1}\right\}\\ &\quad\times\Upsilon_n^{-1}\begin{bmatrix}\sum_{t=1}^{n}\mathbf{A}_1'\mathbf{z}_{t-1}\varepsilon_t\\ \sum_{t=1}^{n}\mathbf{A}_2'\mathbf{z}_{t-1}\varepsilon_t\end{bmatrix}\\ &= \mathbf{A}\begin{bmatrix}\Upsilon_{1n}^{-1}\sum_{t=1}^{n}\mathbf{A}_1'\mathbf{z}_{t-1}\mathbf{z}_{t-1}'\mathbf{A}_1\Upsilon_{1n}^{-1} & o_p(1)\\ o_p(1) & \Upsilon_{2n}^{-1}\sum_{t=1}^{n}\mathbf{A}_2'\mathbf{z}_{t-1}\mathbf{z}_{t-1}'\mathbf{A}_2\Upsilon_{2n}^{-1}\end{bmatrix}^{-1}\\ &\quad\times\begin{bmatrix}\Upsilon_{1n}^{-1}\sum_{t=1}^{n}\mathbf{A}_1'\mathbf{z}_{t-1}\varepsilon_t\\ \Upsilon_{2n}^{-1}\sum_{t=1}^{n}\mathbf{A}_2'\mathbf{z}_{t-1}\varepsilon_t\end{bmatrix}. \end{aligned} \tag{A.4}$$

We now establish convergence results for the elements in (A.4). First we observe, using results in Lemma A.3, that

$$\Upsilon_{1n}^{-1}\sum_{t=1}^{n}\mathbf{A}_1'\mathbf{z}_{t-1}\mathbf{z}_{t-1}'\mathbf{A}_1\Upsilon_{1n}^{-1} \Rightarrow \sigma^2\int_0^1 \Xi(r)dr, \tag{A.5}$$

where

$$\int_0^1 \Xi(r)dr := \begin{bmatrix}\frac{\mathcal{V}_{c_0}^2}{4(1-\cos\phi)^2}+\frac{1}{2(1-\cos\phi)}\frac{(\mathcal{A}^2+\mathcal{B}^2)}{4\sin^2\phi} & \frac{(\mathcal{A}^2+\mathcal{B}^2)}{8\sin^2\phi} & \frac{\left(\cos\phi+\frac{1}{2}\right)(\mathcal{A}^2+\mathcal{B}^2)}{4\sin^2\phi}\\ \frac{(\mathcal{A}^2+\mathcal{B}^2)}{8\sin^2\phi} & \frac{(\mathcal{A}^2+\mathcal{B}^2)}{4\sin^2\phi} & \frac{\cos\phi(\mathcal{A}^2+\mathcal{B}^2)}{4\sin^2\phi}\\ \frac{\left(\cos\phi+\frac{1}{2}\right)(\mathcal{A}^2+\mathcal{B}^2)}{4\sin^2\phi} & \frac{\cos\phi(\mathcal{A}^2+\mathcal{B}^2)}{4\sin^2\phi} & \frac{(\mathcal{A}^2+\mathcal{B}^2)}{4\sin^2\phi}\end{bmatrix}$$

with $\mathcal{V}_{c_0}^2$, $\mathcal{A}^2$, and $\mathcal{B}^2$ as defined in Theorem 2. Consequently, for the inverse of the right member of (A.5) we have that $\left[\sigma^2\int_0^1\Xi(r)dr\right]^{-1} =$

$$\frac{1}{\sigma^2\mathcal{V}_{c_0}^2}\begin{bmatrix}4(1-\cos\phi)^2 & 2(2\cos\phi-1)(1-\cos\phi) & -2(1-\cos\phi)\\ 2(2\cos\phi-1)(1-\cos\phi) & \frac{4\mathcal{V}_{c_0}^2}{(\mathcal{A}^2+\mathcal{B}^2)}+(2\cos\phi-1)^2 & (1-2\cos\phi)-\frac{4\mathcal{V}_{c_0}^2\cos(\phi)}{(\mathcal{A}^2+\mathcal{B}^2)}\\ -2(1-\cos\phi) & (1-2\cos\phi)-\frac{4\mathcal{V}_{c_0}^2\cos(\phi)}{(\mathcal{A}^2+\mathcal{B}^2)} & \frac{4\mathcal{V}_{c_0}^2}{(\mathcal{A}^2+\mathcal{B}^2)}+1\end{bmatrix}. \tag{A.6}$$

Moreover, using results in Lemma A.3, we have that

$$\Upsilon_{1n}^{-1}\sum_{t=1}^{n}\mathbf{A}_1'\mathbf{z}_{t-1}\varepsilon_t \Rightarrow \sigma^2 Q(\phi)\Psi, \tag{A.7}$$

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

where

$$Q(\phi) := \begin{bmatrix} \frac{1}{2(1-\cos\phi)} & \frac{1}{4\sin\phi} & -\frac{1}{4(1-\cos(\phi))} \\ 0 & \frac{1}{2\sin\phi} & 0 \\ 0 & \frac{\cos\phi}{2\sin\phi} & -\frac{1}{2} \end{bmatrix}$$

and

$$\Psi := \begin{bmatrix} \int_0^1 W_{0,c_0}(r)dW_{0,c_0}(r) \\ \int_0^1 \left[ W^{\alpha}_{\phi,c_\phi}(r)dW^{\beta}_{\phi,c_\phi}(r) - W^{\beta}_{\phi,c_\phi}(r)\,dW^{\alpha}_{\phi,c_\phi}(r) \right] \\ \int_0^1 \left[ W^{\alpha}_{\phi,c_\phi}(r)dW^{\alpha}_{\phi,c_\phi}(r) + W^{\beta}_{\phi,c_\phi}(r)\,dW^{\beta}_{\phi,c_\phi}(r) \right] \end{bmatrix}.$$

Next we have $\Upsilon_{2n}^{-1} \sum \mathbf{A}_2' \mathbf{z}_{t-1} \mathbf{z}_{t-1}' \mathbf{A}_2 \Upsilon_{2n}^{-1}$ for which by a standard law of large numbers result we have that

$$\Upsilon_{2n}^{-1} \sum_{t=1}^{n} \mathbf{A}_2' \mathbf{z}_{t-1} \mathbf{z}_{t-1}' \mathbf{A}_2 \Upsilon_{2n}^{-1} \xrightarrow{p} \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{k-3} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{k-3} & \gamma_{k-2} & \cdots & \gamma_0 \end{bmatrix} =: \mathbf{\Gamma}, \tag{A.8}$$

where $\gamma_j$ denotes the $jth$-order autocovariance of $\left\{ \Delta \Delta_\phi x_t \right\}$. Moreover, from the CLT it follows that

$$\Upsilon_{2n}^{-1} \sum_{t=1}^{n} \mathbf{A}_2' \mathbf{z}_{t-1} \varepsilon_t \Rightarrow N(0, \sigma^2 \mathbf{\Gamma}). \tag{A.9}$$

Consequently, putting the results in (A.6)–(A.9) together we obtain that

$$\Upsilon_n \left( \widehat{\beta} - \beta \right) \Rightarrow \mathbf{A}\ \mathrm{diag} \left\{ \left[ \int_0^1 \Xi(r)dr \right]^{-1}, \Gamma^{-1} \right\} \begin{bmatrix} Q(\phi)\Psi \\ N(0, \sigma^2 \mathbf{\Gamma}) \end{bmatrix}. \tag{A.10}$$

Before progressing we need to introduce some further notation. Define the $(k+1) \times 1$ selection vector $\mathbf{e}_i$ to have a 1 at the $i$th position and zeros elsewhere. Moreover, partition this selection vector as $\mathbf{e}_i = (\mathbf{e}_{1j}', \mathbf{e}_{2v}')'$ where $\mathbf{e}_{1j}$ is a $3 \times 1$ vector and $\mathbf{e}_{2v}$ is a $(k-2) \times 1$ vector, where the indexes $j$ and $v$ serve to indicate the position at which the one occurs. Note that if one of the first three parameters, namely, $\rho$, $\alpha_1$, or $\alpha_2$, is to be selected then a one will occur in the $\mathbf{e}_{1j}$ vector and $\mathbf{e}_{2v}$ will be a vector of zeros and vice versa.

For each element of $\widehat{\beta}$ we therefore obtain from (A.10) the results in (9)–(11) in Theorem 2 and that

$$n\widehat{\rho} \Rightarrow 2(1-\cos\phi) \frac{\int_0^1 W_{0,c_0}(r)dW_{0,c_0}(r)}{\int_0^1 W_{0,c_0}^2(r)dr}. \tag{A.11}$$

Then from the result in (A.11) coupled with the consistency of the $\alpha_j$, $j = 1, \ldots, k$, estimators, we obtain that $Z_{\hat{\rho}} \Rightarrow \int_0^1 W_{0,c_0}(r)dW_{0,c_0}(r) / \int_0^1 W_{0,c_0}^2(r)dr$. This completes the

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

proof for $Z_{\widehat{\rho}}$. Turning to $t_{\widehat{\rho}}$, the appropriate selection vector is given by $\mathbf{e}_{11} := (1,0,0)'$, and hence

$$t_{\widehat{\rho}} = \frac{n\widehat{\rho}}{\left\{\widehat{\sigma}_k^2 \mathbf{e}_{11}' \left[\Upsilon_{1n}^{-1} \sum_{t=1}^{n} \mathbf{A}_1' \mathbf{z}_{t-1} \mathbf{z}_{t-1}' \mathbf{A}_1 \Upsilon_{1T}^{-1}\right]^{-1} \mathbf{e}_{11}\right\}^{1/2}}$$

$$\Rightarrow \frac{2(1-\cos\phi)\int_0^1 W_{0,c_0}(r)dW_{0,c_0}(r)}{\int_0^1 W_{0,c_0}^2(r)dr} \left[\frac{\int_0^1 W_{0,c_0}^2(r)dr}{4(1-\cos\phi)^2}\right]^{1/2} = \frac{\int_0^1 W_{0,c_0}(r)dW_{0,c_0}(r)}{\left[\int_0^1 W_{0,c_0}^2(r)dr\right]^{1/2}}$$

as required, using the fact that $\widehat{\sigma}_k^2 \xrightarrow{p} \sigma^2$.

We now turn to establishing the result for the $VRT$ statistic in (6). To do so we will require a representation for the limit of $n^{-3/2}\sum_{k=1}^{t} x_k$. Using the frequency specific orthogonal decomposition of $x_t$ from Lemma A.1 it follows that

$$\begin{aligned} n^{-3/2}\sum_{k=1}^{t} x_k = {} & \delta_0 n^{-3/2}\sum_{k=1}^{t} S_{0,c_0}(k) + \frac{\delta_\phi^\alpha}{\sin\phi} n^{-3/2}\sum_{k=1}^{t} \sin[\phi(k+1)]\, S_{\phi,c_\phi}^{\alpha}(k) \\ & - \frac{\delta_\phi^\alpha}{\sin\phi} n^{-3/2}\sum_{k=1}^{t} \cos[\phi(k+1)]\, S_{\phi,c_\phi}^{\beta}(k) \\ & - \frac{\delta_\phi^\beta}{\sin\phi} n^{-3/2}\sum_{k=1}^{t} \cos[\phi(k+1)]\, S_{\phi,c_\phi}^{\alpha}(k) \\ & + \frac{\delta_\phi^\beta}{\sin\phi} n^{-3/2}\sum_{k=1}^{t} \sin[\phi(k+1)]\, S_{\phi,c_\phi}^{\beta}(k) + o_p(1). \end{aligned} \tag{A.12}$$

Next noting, as demonstrated in Chan and Wei (1988, expression 3.3.10 of Lem. 3.3.6, p. 388, and Rmk, 1 p. 389), that $\sum_{k=1}^{t} \sin[\phi(k+1)]\, S_{\phi,c_\phi}^{j}(k)$ and $\sum_{k=1}^{t} \cos[\phi(k+1)]\, S_{\phi,c_\phi}^{j}(k)$, $j = \alpha, \beta$, are each of $o_p\left(n^{3/2}\right)$, it follows immediately that (A.12) simplifies to

$$n^{-3/2}\sum_{k=1}^{t} x_k = \delta_0 n^{-3/2}\sum_{k=1}^{t} S_{0,c_0}(k) + o_p(1). \tag{A.13}$$

The result in (6) then follows from (A.13), (A.3), and applications of the continuous mapping theorem (CMT), after simplification.

For the trinity of $M$ statistics based on the AR-based estimator, $\hat{\lambda}^2 = s_{AR}^2$, the consistency of the OLS parameter estimates from (2), established previously, yields that $s_{AR}^2 \xrightarrow{p} \sigma^2/(1-2\cos\phi+1)^2 = \sigma^2/(4(1-\cos\phi)^2)$. For $MSB$ this coupled with the result in (A.3) and an application of the CMT yields the result in (8). For $MZ_\alpha$ and $MZ_t$ we also require the large-sample behavior of $n^{-1}x_n^2$; see the discussion in Remark 4. For the $M$ tests that employ the kernel-based estimator, $\hat{\lambda}^2 = s_{WA}^2$, it follows, as shown in the proof of Theorem 2 in Taylor (2003), that $s_{WA}^2 = O_p(mn)$, because $z_t := (1-L)x_t$ is nonstationary, from which the stated results follow immediately.

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press

We conclude by establishing the result in (12) in Theorem 2. Here for $1 \leq \upsilon \leq k-2$ we have that

$$
\begin{aligned}
t_{\widehat{\alpha}_{\upsilon+2}} &= \frac{\sqrt{n}\left(\widehat{\alpha}_{\upsilon+2}-\alpha_{\upsilon+2}\right)}{\widehat{\sigma}_k \mathbf{e}_{2\upsilon}' \left\{ \mathbf{H}_2 \left[ \Upsilon_{2n}^{-1} \sum_{t=1}^{n} \mathbf{A}_2' \mathbf{z}_{t-1} \mathbf{z}_{t-1}' \mathbf{A}_2 \Upsilon_{2n}^{-1} \right]^{-1/2} \right\} \mathbf{e}_{2\upsilon}} \\
&= \left\{ \widehat{\sigma}_k \mathbf{e}_{2\upsilon}' \left[ \mathbf{H}_2 \left( \Upsilon_{2n}^{-1} \sum_{t=1}^{n} \mathbf{A}_2' \mathbf{z}_{t-1} \mathbf{z}_{t-1}' \mathbf{A}_2 \Upsilon_{2n}^{-1} \right)^{-1/2} \right] \mathbf{e}_{2\upsilon} \right\}^{-1} \\
&\quad \times N\left(0, \sigma^2 \mathbf{e}_{2\upsilon}' \mathbf{H}_2 \Gamma^{-1} \mathbf{H}_2' \mathbf{e}_{2\upsilon}\right) + o_p(1) \\
&\Rightarrow N\left(0, \frac{\mathbf{e}_{2\upsilon}' \mathbf{H}_2 \Gamma^{-1} \mathbf{H}_2' \mathbf{e}_{2\upsilon}}{\mathbf{e}_{2\upsilon}' \mathbf{H}_2 \Gamma^{-1} \mathbf{H}_2' \mathbf{e}_{2\upsilon}}\right) = N(0,1). \qquad \blacksquare
\end{aligned}
$$

https://doi.org/10.1017/S0266466613000066 Published online by Cambridge University Press