TESTING FOR CYCLICAL NON-STATIONARITY IN AUTOREGRESSIVE PROCESSES

By Robert M. Kunst
Institute for Advanced Studies, Vienna and Johannes Kepler University, Linz

*Journal of Time Series Analysis **12**, 137–156 (1997)*

Abstract

This paper deals with the distributions evolving from the likelihood-ratio test for the factor $1-B^n$ in the lag polynomial $\Phi(B)$ under the basic assumption that the data series is generated by the autoregressive model $\Phi(B)X_t = \varepsilon_t$ where $\{\varepsilon_t\}$ denotes Gaussian white noise. A characterization of the statistic and its asymptotic properties is given. Asymptotic and finite sample significance points are tabulated. The test procedure is illustrated by an economics example.

Keywords. Time series; unit roots; seasonality; autoregressive processes.

1 INTRODUCTION

Recently, the problem of testing for unit roots in the autoregressive polynomial operator of an autoregressive (AR) or, this case requiring some extensions, an autoregressive-moving average (ARMA) process has gained widespread attention. The typical tests have as null hypothesis a certain kind of non-stationarity, such as exactly one root at 1—the so-called integrated process of order 1. Rejection then usually indicates stationary behavior (possibly with added deterministic components). The single unit root at 1 has been treated most extensively, mainly because of its popularity in modeling real economic growth (see Nelson and Plosser (1982) for the economics and Dickey and Fuller (1979) for the statistics). Less attention has been given to the multiple root at 1 although it is sometimes reported as being consistent with price series (see for example Hendry, 1995).

Another interesting non-stationary model is given by n equidistant complex roots on the unit circle, one of which is exactly 1. This model has been considered already by Box and Jenkins (1976), there only tested by visual evidence on the sample autocorrelation function. Dickey *et al.* (1984) have suggested explicit statistical tests on this model for the cases $n = 2, 4, 12$, corresponding to semi-annual, quarterly and monthly data.

This paper is concerned with analogous tests for arbitrary n. Although certain periodicities are suggested by the seasonal structure of the series in typical economic applications (such as the example given in Section 3) and in temperature or precipitation data for example, this need not necessarily be so in all data series. Moreover, multiples of the assumed frequency (e.g. 6 months instead of 12) may still give rise to poles in the spectrum even though the complete factor (in this case year-to-year differences) has been rejected, a possibility implicitly treated for quarterly data by Hylleberg *et al.* (1990). Their technique of decomposing the polynomial factors can be generalized to similar cases, such as monthly or bimonthly data (see for example Franses, 1991).

If the investigated polynomial can be fully decomposed, such as $1 - B^4$ into $1 - B, 1 + B, 1 + B^2$, the asymptotics of tests on the single factors are known in many cases, e.g. on $1 - B$ from the work of Dickey and Fuller (1979) and on $1 + B^2$ from Ahtola and Tiao (1987). In practice, an additional test on the joint hypothesis, i.e. on the complete factor $1 - B^n$, is still advisable, by similar arguments as for the joint use of F and t–tests in regression analysis. In the case of prime n, e.g. if one wants to test for weekly 5–days cycles in daily data on asset prices, only the first differences operator $1 - B$ can be split off $1 - B^n$ and testing on the full factor becomes indispensable.

The remainder of this paper is organized as follows. The relevant class of distributions is defined and characterized in Section 2. The testing problem leading to these distributions is considered in Section 3 and it is applied to an empirical example in Section 4. Section 5 concludes.

2 CIRCULAR χ^2 DISTRIBUTIONS

In order to define the circular χ^2 family, let us first consider a finite sequence of n mutually independent random walks $\{X_{1,t}\}, \{X_{2,t}\}, ..., \{X_{n,t}\}$ with independent and identically distributed (i.i.d.) Gaussian increments $\{\varepsilon_{i,t}\}$, $i = 1, ..., n$:

$$X_{i,t} = X_{i,t-1} + \varepsilon_{i,t}, \quad t = 1, 2, \dots. \tag{2.1}$$

For convenience, let us assume $X_{i,0} = 0$ and denote $\mathrm{E}\varepsilon_{t-1}^2$ by σ^2 which is assumed to be constant across i. Next, we focus attention on the process $\{Y_t\}$ which is derived from the $\{X_{i,t}\}$ by circular merging. This means that

$$(Y_1, Y_2, ..., Y_n, Y_{n+1}, ...) = (X_{1,1}, X_{2,1}, ..., X_{n,1}, X_{1,2}, ...)$$

or, in short (using $[x]$ for integer x),

$$Y_t = X_{t-[(t-1)/n],[(t-1)/n]+1}. \tag{2.2}$$

Let $\mathbf{H}_n(T)$ be the matrix of T^2–weighted sums of squares and cross-products of the process $\{Y_t\}$ with its first $n - 1$ lags, constructed from the observations $Y_1, ..., Y_T$, i.e.

$$\mathbf{H}_n(T) = (H_{ij}(T))_{i,j=1,\dots,n} = \left(\left(\sum_{t=\max(i,j)+1}^{T} Y_{t-i} Y_{t-j} \right) \Big/ T^2 \right). \tag{2.3}$$

It can be shown that $\mathbf{H}_n(T)$ approaches a non-trivial limit distribution if $T \to \infty$. To demonstrate this, we use the standard result that

$$T^{-2} \sum_{t=1}^{T} X_{i,t} X_{j,t} \Rightarrow \sigma^2 \int_0^1 B_i(\omega) B_j(\omega) d\omega$$

where $\Rightarrow$ denotes convergence in distribution and $B(\omega)$ denotes standard Brownian motion. For a derivation of this result under general conditions on the error process, see for example Davidson (1994, p. 498). For the circularly merged random process Y_t we obtain, assuming $j > i$ without loss of generality,

$$
\begin{aligned}
&T^{-2} \sum_{t=j+1}^{T} Y_{t-i} Y_{t-j} \\
&= T^{-2} \sum_{t=j+1}^{T} X_{t-i-[(t-i-1)/n],[(t-i-1)/n]+1} X_{t-j-[(t-j-1)/n],[(t-j-1)/n]+1} \\
&= \mathrm{o_p}(T^{-1}) + T^{-2} \sum_{l=1}^{n} \sum_{t=1}^{[T/n]} X_{l,t} X_{l+j-i-[(l+j-i-1)/n],t} \\
&\qquad \Rightarrow \sigma^2 n^{-1} \sum_{l=1}^{n} \int_0^1 B_l(\omega) B_{l+j-i-[(l+j-i-1)/n]}(\omega)\, d\omega.
\end{aligned}
$$

For the diagonal elements we have simply

$$
\begin{aligned}
T^{-2} \sum_{t=1}^{T} Y_t^2 &= T^{-2} \sum_{t=1}^{T} X^2_{t-[(t-1)/n],t-[(t-1)/n]+1} \\
&= \mathrm{O_p}(T^{-1}) + T^{-2} \sum_{l=1}^{n} \sum_{t=1}^{[T/n]} X^2_{l,t} \Rightarrow \sigma^2 n^{-1} \sum_{l=1}^{n} \int_0^1 B_l^2(\omega) d\omega.
\end{aligned}
$$

Individually, all off-diagonal elements converge to the same limit law. Regarding the joint asymptotic distribution, the random matrices $\mathbf{H}_n(T)$ converge to a stochastic limit matrix $\mathbf{H}_n$ with a special symmetric circular Toeplitz structure, a property which will be treated in more detail below.

The n–vector $\mathbf{h}_n(T)$ is defined to contain the T–weighted cross-products of the n–step increments of Y_t, i.e. $\Delta_n Y_t = Y_t - Y_{t-n}$ with the first n lags of Y_t, again constructed from the observations $Y_1, \ldots, Y_T$, i.e.

$$
\mathbf{h}_n(T) = \left(\left(\sum_{t=2}^{T} Y_{t-1} \Delta_n Y_t \right) \Big/ T, \ldots, \left(\sum_{t=n+1}^{T} Y_{t-n} \Delta_n Y_t \right) \Big/ T \right)'. \tag{2.4}
$$

Note that the n–step increments are a serially uncorrelated Gaussian process which is also independent with respect to all past realizations of $\{Y_t\}$. Hence the result that

$$
T^{-1} \sum_{t=1}^{T} X_t \varepsilon_t \Rightarrow \sigma^2 \int_0^1 B(\omega) dB_0(\omega)
$$

for random walk $\{X_t\}$ and independent noise $\{\varepsilon_t\}$ can be used to derive the limit law of $\mathbf{h}_n(T)$. In detail,

$$\begin{aligned}
T^{-1}\sum_{t=1}^{T} Y_{t-i}\Delta_n Y_t &= T^{-1}\sum_{t=1}^{T} X_{t-i-[(t-i-1)/n][(t-i-1)/n]+1}\varepsilon_{t-[(t-1)/n],[(t-1)/n]+1} \\
&= \mathrm{O_p}(T^{-1/2}) + T^{-1}\sum_{l=1}^{n}\sum_{t=1}^{[T/n]} X_{l,t}\varepsilon_{l+i-[(l+i-1)/n],t} \\
&\Rightarrow \sigma^2 n^{-1}\sum_{l=1}^{n}\int_0^1 B_l(\omega)dB_{l+i-[(l+i-1)/n]}(\omega).
\end{aligned}$$

For the individual distributions, the first index of the noise process plays no role. Hence, all elements of the limit vector $\mathbf{h}_n$ are identically distributed. In contrast to the $\mathbf{H}_n$ elements, this limit law has a standard form and is a transform of χ^2 with n degrees of freedom. We note, however, that the $\mathbf{h}_n$ elements are not independent.

From the joint stochastic limit laws for the $n \times n$–matrix $\mathbf{H}_n$ and for the n–vector $\mathbf{h}_n$, we define the circular χ^2 distribution with periodicity n by the distribution of the scale-free quadratic form

$$\beta_n = \sigma^{-2}\mathbf{h}_n'\mathbf{H}_n^{-1}\mathbf{h}_n. \tag{2.5}$$

A special case is $n = 1$. In this case, β_1 gives the distribution of

$$\lim_{T\to\infty} T\left(\sum_{t=2}^{T} X_{t-1}\varepsilon_t\right)^2 \Bigg/ \sum_{t=2}^{T} X_{t-1}^2 \tag{2.6}$$

with random walk X_t and, for example, $\sigma^2 = 1$. β_1 is closely related to the distribution of the unit root test statistic suggested by Dickey and Fuller (1979) and some of its fractiles have been tabulated by Johansen (1988). It is well known that the limit distribution in (2.6) can be written in a closed expression

$$\left\{\int_0^1 B(\omega)dB(\omega)\right\}^2 \Bigg/ \left\{\int_0^1 B(\omega)^2 d\omega\right\}. \tag{2.7}$$

If $n = 2$ the matrix $\mathbf{H}_2$ consists of four elements. We note that $h_{12} = h_{21}$ and that the individual distributions of h_{11} and h_{22} are identical. The inverse of this simple Toeplitz matrix is given by

$$\mathbf{H}_2^{-1} = (h_{11}^2 - h_{12}^2)^{-1}\left[\begin{pmatrix} h_{11} & -h_{12} \\ -h_{12} & h_{11}\end{pmatrix}\right].$$

The general asymptotic form of $\mathbf{H}_n$ is

$$\begin{bmatrix}
a_1 & a_2 & a_3 & \cdots & a_3 & a_2 \\
a_2 & a_1 & a_2 & \cdots & a_4 & a_3 \\
\vdots & & & & \vdots & \\
a_3 & a_4 & a_5 & \cdots & a_1 & a_2 \\
a_2 & a_3 & a_4 & \cdots & a_2 & a_1
\end{bmatrix}. \tag{2.8}$$

Such matrices are known in the literature as symmetric circular Toeplitz matrices. For their properties, see Basilevsky (1983, p. 223) for example, but note that the formula for their determinant given there is incorrect and appears to hold for $n = 1$ and $n = 3$ only. We therefore summarize the determinants $d_i = \det(\mathbf{H}_i)$ for some exemplary and important cases:

$$\begin{array}{ll} n=1 & d_1 = a_1 \\ n=2 & d_2 = (a_1+a_2)(a_1-a_2) \\ n=3 & d_3 = (a_1+2a_2)(a_1-a_2)^2 \\ n=4 & d_4 = (a_1-a_3)^2(a_1+a_3+2a_2)(a_1+a_3-2a_2) \\ n=6 & d_6 = (a_1+2a_2+2a_3+a_4)(a_1-2a_2+2a_3-a_4). \end{array}$$

The circular Toeplitz matrices are more easily accessed through their eigenvalues. These eigenvalues r_k and corresponding eigenvectors v_k can be expressed via sine and cosine functions:

$$r_k = \sum_{j=0}^{n-1} a_{j+1} \cos\left(\frac{2\pi kj}{n}\right) \tag{2.9}$$

$$\begin{aligned} v_k = \; & n^{1/2}\left\{\cos\left(\frac{2\pi k}{n}\right)+\sin\left(\frac{2\pi k}{n}\right),\cos\left(\frac{4\pi k}{n}\right)\right. \\ & \left.+\sin\left(\frac{4\pi k}{n}\right),\ldots,\cos\left(\frac{2n\pi k}{n}\right)+\sin\left(\frac{2n\pi k}{n}\right)\right\}' \end{aligned} \tag{2.10}$$

The matrix $\mathbf{V}_n = (v_1, ..., v_n)$ is symmetric and $\mathbf{V}_n^2$ is the identity matrix. This provides us with the essential elements for the quadratic form (2.5). $\mathbf{R}_n$ will denote the diagonal matrix with the elements $r_1, ..., r_n$ in the following.

In general, properties of the circular χ^2 distribution can only be obtained by Monte Carlo simulation. To this end, first n random walks have to be generated to calculate the a_n elements and to approximate the $\mathbf{h}_n$ vector. Then, the property

$$\mathbf{h}_n'\mathbf{H}_n^{-1}\mathbf{h}_n = \mathbf{h}_n'\mathbf{V}_n\mathbf{R}_n^{-1}\mathbf{V}_n\mathbf{h}_n \tag{2.11}$$

can be used to approximate the required distribution without having to invert $\mathbf{H}_n$. The approximation can be improved by increasing the length of the circularly merged random walk or by increasing the number of replications. For the characteristics of the circular χ^2 distribution reported in Table I, the length of the series was set at 1000 and the number of replications at 10 000. Table I shows that the mean and variance of $\mathbf{H}_n$ can be approximated roughly by n and $2n$, respectively, which would correspond to χ^2 distributions with n degrees of freedom. We also note that the deviation from the Gaussian shape measured by third and fourth moments decreases with the periodicity n.

The fractiles reported in Table I do not coincide with those from the finite-sample analog ($\hat{\sigma}^2$ denotes the least-squares or conditional maximum likelihood estimate of σ^2)

$$\beta_n(T) = \hat{\sigma}^{-2}\mathbf{h}_n(T)'\,\mathbf{H}_n(T)^{-1}\,\mathbf{h}_n(T) \tag{2.12}$$

which could be called a circular Fisher Z distribution. $\mathbf{H}_n(T)$ does not have the circular Toeplitz property. Hence, the distribution of $\beta_n(T)$ must be simulated directly from the cross-sums of lagged Y_t. In Table II, simulated fractiles are displayed for selected values of n and T. For $T = 1000$, there is little difference from Table I, while the fractiles seem to be smaller for smaller values of T. We note that, in practical applications, the small-sample distribution of $\beta_n(T)$ is influenced by a variety of possible deviations from the independent normal assumption for $\{\varepsilon_t\}$ and also by plausible modifications of (2.12).

Table I: Characteristics of the Circular χ^2 Distribution.

	Fractiles								Moments			
n	0.1	0.25	0.5	0.75	0.9	0.95	0.975	0.99	Mean	Variance	Skewness	Kurtosis
1	0.0	0.1	0.6	1.6	2.9	4.1	5.1	6.8	1.1	2.1	2.5	9.6
2	0.3	0.8	1.7	3.2	5.0	6.5	7.9	9.5	2.3	4.4	1.8	5.0
3	0.7	1.4	2.7	4.4	6.8	8.5	10.0	12.0	3.3	6.7	1.5	3.2
4	1.3	2.2	3.8	5.9	8.4	10.1	11.9	14.2	4.4	9.0	1.4	3.2
5	1.9	3.0	4.8	7.2	9.9	11.7	13.5	16.0	5.4	10.9	1.2	1.9
6	2.6	3.9	5.9	8.7	11.6	13.8	15.7	17.9	6.7	13.8	1.1	1.7
7	3.2	4.7	6.9	9.8	12.8	14.9	17.2	19.8	7.6	15.5	1.0	1.5
8	4.0	5.6	8.1	11.1	14.5	16.8	19.0	21.9	8.8	18.6	1.1	2.0
9	4.7	6.5	9.1	12.3	15.8	18.3	20.4	23.0	9.8	20.3	0.9	1.2
10	5.5	7.4	10.3	13.6	17.3	19.6	21.8	24.6	10.9	22.3	0.9	1.1
11	6.1	8.4	11.3	14.8	18.4	21.1	23.4	26.8	11.9	24.8	0.9	1.3
12	7.1	9.4	12.6	16.3	20.4	23.0	25.2	28.5	13.3	28.2	0.8	0.9

Notes: n is the periodicity parameter. Numbers are based on a Monte Carlo experiment with Gaussian random walks of length 1000 and 10 000 replications. Kurtosis has been standardized so that it is 0 for the Gaussian case.

Table II: Characteristics of the Finite-Sample Distribution.

		Fractiles								Moments			
	n	0.1	0.25	0.5	0.75	0.9	0.95	0.975	0.99	Mean	Variance	Skewness	Kurtosis
$T = 1000$	1	0.0	0.1	0.6	1.6	2.9	4.1	5.1	6.9	1.1	2.1	2.5	9.5
	2	0.3	0.8	1.7	3.2	5.0	6.4	7.8	9.5	2.3	4.4	1.8	4.9
	3	0.7	1.4	2.7	4.5	6.7	8.4	10.0	11.8	3.3	6.6	1.5	3.2
	4	1.3	2.2	3.8	5.9	8.3	10.0	11.7	14.0	4.4	8.7	1.4	3.1
	6	2.6	3.9	5.9	8.6	11.4	13.5	15.4	17.4	6.6	13.1	1.1	1.5
	12	7.0	9.2	12.2	15.8	19.7	22.1	24.3	27.2	12.9	25.6	0.7	0.7
$T = 100$	1	0.0	0.2	0.6	1.5	2.9	3.9	5.2	6.8	1.1	2.1	2.5	8.8
	2	0.3	0.8	1.7	3.1	4.8	6.3	7.6	9.5	2.2	4.2	1.8	5.0
	3	0.7	1.4	2.6	4.4	6.5	8.0	9.4	11.5	3.2	6.1	1.5	3.1
	4	1.3	2.2	3.6	5.6	7.9	9.6	11.2	13.2	4.2	7.9	1.4	3.2
	6	2.4	3.7	5.7	8.1	10.9	12.8	14.6	16.8	6.3	11.9	1.1	1.7
	12	6.5	8.6	11.4	14.7	18.5	20.8	23.2	25.9	12.0	22.6	0.8	1.0

Notes: n is the periodicity parameter. Numbers are based on a Monte Carlo experiment with Gaussian random walks of length T and 10 000 replications. Kurtosis has been standardized so that it is 0 for the Gaussian case.

3 THE LIKELIHOOD-RATIO TEST FOR NON-STATIONARY CYCLES

Assume that a time series follows an autoregressive process of order p

$$Y_t = \Phi_1 Y_{t-1} + ... + \Phi_p Y_{t-p} + \varepsilon_t \tag{3.1}$$

which can be written by a polynomial in the lag operator B:

$$\Phi(B) Y_t = \varepsilon_t. \tag{3.2}$$

It is a common assumption that the roots of $\Phi(z) = 0$ are restricted to the area outside of the unit disk which warrants stationary behavior of $\{Y_t\}$. There are also some popular models which allow for certain violations of this assumption to model non-stationary behavior. If n roots lie at $z = 1$, the process is called integrated of order n. The first- and second-order integrated models are used frequently to represent trending or random-walk-like behavior. If n roots at equidistant points on the unit disk are allowed and $z = 1$ is one of them, the model reflects non-stationary cyclical fluctuations with a length of n observations. In this case, $\Delta_n = 1 - B^n$ is a factor of $\Phi(B)$ and the application of Δ_n to Y_t yields a stationary time series. It is important to know whether this kind of non-stationarity is actually present as unjustified application of Δ_n severely complicates the time series properties of $\{Y_t\}$. If the true model is finite-order autoregressive, i.e. AR(p) as in (3.1), the filtered process is ARMA(p, n) with unit roots in the MA lag polynomial. A test for this kind of non-stationarity is suggested by a transformation of (3.1):

$$\Delta_n Y_t = b_1 \Delta_n Y_{t-1} + \ldots + b_{p-n} \Delta_n Y_{t-p+n} + c_1 Y_{t-p+n-1} + \ldots + c_n Y_{t-p} + \varepsilon_t. \tag{3.3}$$

If (and only if) the null hypothesis holds that Δ_n is a factor of $\Phi(B)$ then the coefficients on the lagged level terms must be zero, i.e. $c_1 = ... = c_n = 0$. Provided that $\{\varepsilon_t\}$ is a temporally uncorrelated Gaussian process with variance σ^2, the distribution of the likelihood-ratio test statistic LR_n on H_0: $c_1 = ... = c_n = 0$ ($L(.)$ denotes the logarithmic likelihood function conditional on fixed starting values $Y_{-p+1}, ..., Y_0$)

$$\begin{aligned} \mathrm{LR}_n = -2\{ & L(b_1, ..., b_{p-n}, 0, ..., 0, \sigma^2 | Y_1, ..., Y_T) \\ & - L(b_1, ..., b_{p-n}, c_1, ..., c_n, \sigma^2 | Y_1, ..., Y_T)\} \end{aligned}$$

and of the handier F–type regression statistic ($\mathbf{u}_0$ and $\mathbf{u}$ are the respective residual vectors estimated under the null and the general hypothesis)

$$F_n = (T - p)(\mathbf{u}_0' \mathbf{u}_0 - \mathbf{u}' \mathbf{u}) / (\mathbf{u}' \mathbf{u}) \tag{3.4}$$

would both tend to χ^2 with n degrees of freedom asymptotically if all the variables were stationary. Note that F_n has not been standardized to approximate unit mean by division through n, as is common for the ordinary regression F statistic. The non-stationarity of $\{Y_t\}$ complicates the limiting behavior of LR_n and F_n considerably. For $n = 1$, Dickey and Fuller (1979) have analyzed the limiting distribution and tabulated simulated fractiles. For $n = 4$, Dickey *et al.* (1984) have published some results and the problem has been taken up again by Hylleberg *et al.* (1990). For general n and $p = n$, F_n clearly corresponds to the quadratic form β_n in (2.12) and thus converges in distribution to the circular χ^2 law defined via (2.5). The analogy to the asymptotic χ^2 law in the standard stationary case justifies the name 'circular χ^2 distribution' adopted in Section 2.

The tables published by Dickey *et al.* (1984) differ from our tables as they use a different representation of the AR process, regressing the level variate Y_t on one lagged level Y_{t-n}, possibly augmenting this regression with differenced terms. Testing the coefficient of Y_{t-n} against one is essentially equivalent to the F_n test suggested here but it entails the use of Studentized coefficients instead of the F–like restriction test on (3.3). The distribution of the Studentized coefficients is two-sided but non-symmetric which makes a direct algebraic transition from Dickey's tables to Tables I and II impossible.

Dickey *et al.* also presented a test version for AR processes with intercepts, entailing subtracting the means in the construction of F_n. Even if the true constant term is zero, an extraction of means affects the limit distribution where Brownian motion changes to de-meaned Brownian motion. Under the null hypothesis, the series is non-stationary and does not have true first moments, and therefore the sample mean does not converge but influences the distribution even in the limit. In Table III, simulated fractiles are displayed for the case that the true process has zero drift but an intercept has been included in the regression model. These fractiles are the same as those for the case where sample means have been extracted preliminarily. The differences between Table II and Table III are obvious. They decrease with larger n and this coincides well with the observation that they also decrease with larger dimension in the multivariate versions of this test (cf. Johansen, 1994).

Table III: CHARACTERISTICS OF THE FINITE-SAMPLE DISTRIBUTION BASED ON CENTERED VARIATES.

		Fractiles								Moments			
	n	0.1	0.25	0.5	0.75	0.9	0.95	0.975	0.99	Mean	Var.	Skew.	Kurt.
$T = 1000$	1	0.2	1.0	2.4	4.2	6.6	8.2	9.8	11.6	3.0	6.9	1.4	2.7
	2	0.9	1.9	3.5	5.7	8.1	9.9	11.6	13.7	4.1	9.0	1.2	2.2
	3	1.4	2.7	4.5	6.9	9.8	11.6	13.5	15.6	5.1	11.4	1.1	1.6
	4	2.1	3.5	5.6	8.3	11.2	13.1	15.0	17.5	6.2	13.3	1.1	1.7
	6	3.6	5.3	7.7	10.8	14.2	16.3	18.3	20.9	8.4	17.8	0.9	1.2
	12	8.1	10.8	14.1	17.9	22.2	24.7	27.1	30.1	14.7	30.3	0.7	0.7
$T = 100$	1	0.2	1.0	2.4	4.2	6.4	8.0	9.5	11.5	2.9	6.6	1.4	2.9
	2	0.8	1.8	3.4	5.5	8.0	9.7	11.5	13.3	4.0	8.7	1.3	2.4
	3	1.4	2.6	4.4	6.7	9.3	11.0	12.6	14.9	5.0	10.2	1.1	1.6
	4	2.0	3.4	5.4	7.8	10.6	12.5	14.4	16.5	6.0	12.1	1.1	1.9
	6	3.4	5.0	7.4	10.2	13.2	15.4	17.1	19.0	7.9	15.5	0.8	0.9
	12	7.6	9.9	13.0	16.5	20.3	22.9	25.3	27.9	13.5	25.9	0.7	0.8

Notes: n is the periodicity parameter. Numbers are based on a Monte Carlo experiment with Gaussian random walks of length T and 10 000 replications. Kurtosis has been standardized so that it is 0 for the Gaussian case.

Up to this point, only the simple model (3.3) with $p = n$ has been accounted for, for which (2.5) is the straightforward limit expression for the F_n statistic. If some non-zero short-term b_i coefficients appear in (3.3), i.e. $p > n$, restricted and unrestricted estimation is done conditional on the corresponding lagged differences. The arguments of Dickey *et al.* (1984, p. 366) that this conditioning on stationary regressors cannot affect the limit distribution carry over to our problem. In small samples, however, the effect can be troublesome, particularly when it is difficult to determine the order of the AR process. If the true process is ARMA rather than finite-order AR, all AR models will be misspecified and the low power reported by Schwert (1989) for the Dickey-Fuller (F_1) test will be also felt for the cycles test. In general, whenever the true errors process in (3.3) is not white noise, the limit distribution is affected by nuisance parameters and the fractiles given are invalid. The reason for this phenomenon is in the limiting behavior of the sums $\sum Y_{t-i}\varepsilon_t$ where absence of any correlation between Y_{t-i} and ε_t is essential (cf. Park and Phillips (1988)).

4 AN EXAMPLE: THE BRITISH UNEMPLOYMENT SERIES

To provide an example of the way the test procedure detects cyclical non-stationarity, it was applied to the OECD series of unemployment in the United Kingdom (monthly data from 1960 to September of 1989, measured in 1000 persons). To stabilize the variance, the popular transformation of taking logarithms was applied to the series. A referee has suggested comparing the series obtained with an alternative transformation via the quartic root. The two resulting series LUR and QUR are shown in Figure 1. Although no seasonal adjustment has been applied, it is difficult for the eye to see repetitive seasonal patterns in the series. This would be easier in countries whose economy is heavily influenced by harsh winters, such as Sweden or Austria.

Table IV shows the test statistics under the assumption of an AR model of order 20. This means that $20 - n$ differences Δ_n have been used as conditioning terms for the test for cycles of length n. The AR order of 20 for the levels is clearly suggested by several criteria and is the same for both specifications (LUR and QUR).

The factor $1 - B^n$ is rejected at the 1% level of significance for all n which are not factors of 12. At the same time, the factor cannot be rejected even at the 10% level of significance for $n = 1, 2, 3, 4, 6, 12$. This result confirms the assumption that seasonal differencing is consistent with the cyclical structure of the data series.

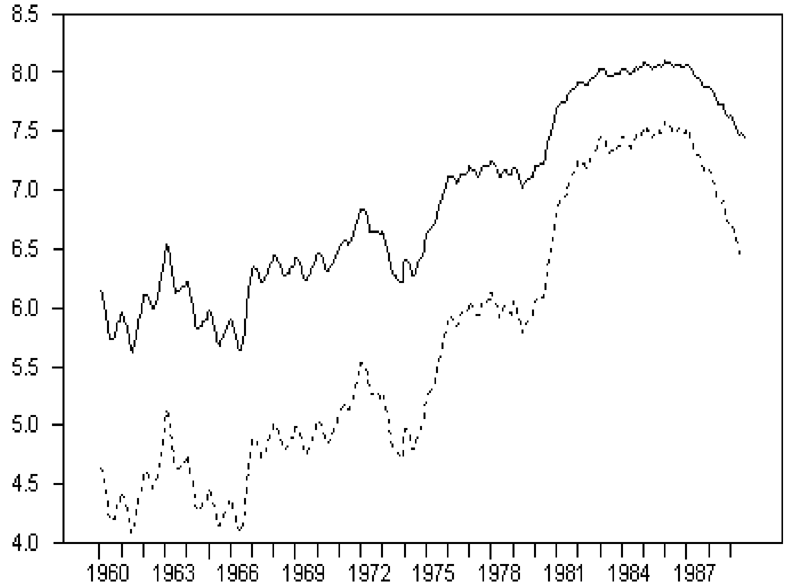

Figure 1: Time series plots of UK unemployment rate. The solid line corresponds to the logarithmic transformation of the original variable (LUR), the dashed line to the transformation by the fourth root (QUR).

Table IV: Values of Test Statistics for UK Unemployment Series.

n	T	F_n(LUR)	F_n(QUR)	Significance
1	337	1.139	1.108	–
2	337	1.738	1.210	–
3	337	2.534	2.195	–
4	337	4.615	3.526	–
5	337	30.680	31.399	1%
6	337	2.929	2.113	–
7	337	30.110	32.098	1%
8	337	27.151	27.518	1%
9	337	38.274	40.184	1%
10	337	94.463	100.950	1%
11	337	79.824	85.742	1%
12	337	4.864	3.777	–

Notes: n is the periodicity length, T is the number of observations, F_n is the statistic defined in (3.3) and (3.4), and LUR and QUR denote the logarithmic and quartic-root transformation of the unemployment rate. –, not significant at the 10% level.

5 OUTLOOK

There remain some unsolved problems. Firstly, extensive Monte Carlo simulations will be necessary to assess the power of the test in the presence of MA terms. In that case, any finite-order AR model would be misspecified. More general types of misspecification could also be interesting. Certain kinds of misspecification could give rise to extensions of the model class under consideration and to the evolution of new tests in the more general class. This observation also applies to distributional robustness.

Secondly, even within the bounds of the AR models, the true order is not known in applications. Many procedures are available to estimate the AR order but, as the testing power is not independent of the order selection, the validity of these procedures will have to be assessed jointly with the cycles test. Asymptotic results can certainly be obtained analytically on this topic but it is doubtful whether these are of direct practical use.

Thirdly, a field of applications for the tests on unusual cycles, like $n = 7$, is still unknown. Probably, such cycles could appear in data series stemming from technical and engineering sciences.

References

[1] Ahtola, J. and Tiao, G.C. (1987) Distributions of least squares estimators of autoregressive parameters for a process with complex roots on the unit circle. *J. Time Ser. Anal.* 8, 1–14.

[2] Basilevsky, A. (1983) *Applied Matrix Algebra in the Statistical Sciences.* New York, Amsterdam: North-Holland.

[3] Box, G.E.P. and Jenkins, G.M. (1976) *Time Series Analysis, Forecasting and Control.* San Francisco: Holden Day.

[4] Davidson, J. (1994) *Stochastic Limit Theory.* Oxford: Oxford University Press.

[5] Dickey, D.A. and Fuller, W.A. (1979) Distribution of the estimators for autoregressive time series with unit roots. *J. Am. Stat. Assoc.* 74, 427–431.

[6] ______ Hasza, D.P. and Fuller, W.A. (1984) Testing for unit roots in seasonal time series. *J. Am. Stat. Assoc.* 79, 355–67.

[7] Franses, P.H. (1991) *Model Selection and Seasonality in Time Series.* Amsterdam: Thesis Publishers.

[8] Hendry, D.F. (1995) *Dynamic Econometrics.* Oxford: Oxford University Press.

[9] Hylleberg, S., Engle, R.F., Granger , C.W.J. and Yoo, B.S. (1990) Seasonal integration and cointegration. *J. Economet.* 44, 215–38.

[10] Johansen, S. (1988) Statistical analysis of cointegration vectors. *J. Econ. Dynam. Control* 12, 199–201.

[11] ______ (1994) The role of the constant and linear terms in cointegration analysis of nonstationary variables. *Economet. Rev.* 13, 205–29.

[12] Nelson, C.R. and Plosser, C. (1982) Trends and random walks in macroeconomic time series: some evidence and implications. *J. Monet. Econ.* 10, 139–62.

[13] Park, J.Y. and Phillips, P.C.B. (1988) Statistical inference in regressions with integrated processes: Part I. *Economet. Theory* 4,3, 468–97.

[14] Schwert, G.W. (1989) Tests for unit roots: a Monte Carlo investigation. *J. Bus. Econ. Stat.* 7, 147–59.