233

Reihe Ökonomie

Economics Series

A Nonparametric Test for Seasonal Unit Roots

Robert M. Kunst

233
Reihe Ökonomie
Economics Series

A Nonparametric Test for Seasonal Unit Roots

Robert M. Kunst

January 2009

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna

Contact:

Robert M. Kunst
Department of Economics and Finance
Institute for Advanced Studies
Stumpergasse 56
1060 Vienna, Austria
and
University of Vienna
Department of Economics
Brünner Straße 72
1210 Vienna, Austria
email: robert.kunst@univie.ac.at

Founded in 1963 by two prominent Austrians living in exile – the sociologist Paul F. Lazarsfeld and the economist Oskar Morgenstern – with the financial support from the Ford Foundation, the Austrian Federal Ministry of Education and the City of Vienna, the Institute for Advanced Studies (IHS) is the first institution for postgraduate education and research in economics and the social sciences in Austria. The **Economics Series** presents research done at the Department of Economics and Finance and aims to share "work in progress" in a timely way before formal publication. As usual, authors bear full responsibility for the content of their contributions.

Das Institut für Höhere Studien (IHS) wurde im Jahr 1963 von zwei prominenten Exilösterreichern – dem Soziologen Paul F. Lazarsfeld und dem Ökonomen Oskar Morgenstern – mit Hilfe der Ford-Stiftung, des Österreichischen Bundesministeriums für Unterricht und der Stadt Wien gegründet und ist somit die erste nachuniversitäre Lehr- und Forschungsstätte für die Sozial- und Wirtschaftswissenschaften in Österreich. Die **Reihe Ökonomie** bietet Einblick in die Forschungsarbeit der Abteilung für Ökonomie und Finanzwirtschaft und verfolgt das Ziel, abteilungsinterne Diskussionsbeiträge einer breiteren fachinternen Öffentlichkeit zugänglich zu machen. Die inhaltliche Verantwortung für die veröffentlichten Beiträge liegt bei den Autoren und Autorinnen.

Abstract

We consider a nonparametric test for the null of seasonal unit roots in quarterly time series that builds on the RUR (records unit root) test by Aparicio, Escribano, and Sipols. We find that the test concept is more promising than a formalization of visual aids such as plots by quarter. In order to cope with the sensitivity of the original RUR test to autocorrelation under its null of a unit root, we suggest an augmentation step by autoregression. We present some evidence on the size and power of our procedure and we illustrate it by applications to a commodity price and to an unemployment rate.

Keywords
Seasonality, nonparametric test, unit roots

JEL Classification
C12, C14, C22

Comments
The author wishes to thank Joerg Breitung, Philip Hans Franses, Helmut Luetkepohl, Bent Nielsen, Werner Mueller, and Martin Wagner for helpful comments. The usual proviso applies

Contents

1 Introduction

Since the possibly first contribution to the literature by Dickey *et al.* (1984), various authors have developed statistical test procedures for the discrimination of seasonal unit roots and purely deterministic seasonality in time series (for a survey on this literature and for the properties of the two model classes, see Ghysels and Osborn, 2001). Like the popular HEGY test by Hylleberg *et al.* (1990), most tests were created in a parametric autoregressive framework, where unit roots are the null and their non-existence is the alternative. The reverse tests by Caner (1998), Canova and Hansen (1995), and Tam and Reinsel (1997) can be viewed as semi-parametric, however, with their null of a deterministic cycle and the background of an unobserved-components structure, as it is more typical of the literature on seasonal adjustment rather than seasonal models. Hitherto, fully non-parametric tests on seasonal unit roots have not been in usage. It is the aim of this paper to consider this possibility.

For non-seasonal unit roots, So and Shin (2001) and Aparicio *et al.* (2006, AES) developed genuine non-parametric testing procedures. Such tests are most commonly applied in the presence of non-linear features (for example, see Choi and Moh, 2007). The tests exploit properties of time series that are not linked to moments but are nevertheless characteristic of integrated versus stationary processes, such as sign changes, zero crossings, or the occurrence of new extrema. By construction, sign and extrema features are robust or even entirely invariant to nonlinear transformations of variables and to occasional but rare outliers and level shifts. For this reason, nonparametric tests can be helpful if nonlinear transformations of linear time-series models are plausible data-generating mechanisms and also if the error process indicates strong deviations from the normal distribution. The obvious drawback is that nonparametric tests lack power as compared with parametric tests if the investigated time series indeed conform to standard assumptions.

A less obvious drawback for nonparametric testing for unit roots is that the ideas typically rely on properties of pure random walks. The generalization of random walks to integrated processes imposes problems, as the serial correlation of increments invalidates the underlying statistical results. Thus, while the frequency of new extrema in the logarithm of a random walk with positive values is the same as in an untransformed random walk, it is severely affected if the first difference of the observed variable follows a stable autoregression. A suggestion is then to filter the original data in order to remove autocorrelation. To that aim, one may consider frequency-domain filters or conditioning on lags, as in the 'augmentation' of the Dickey-Fuller test. In this paper, we follow the latter route and condition on some lags determined by an auxiliary regression.

Intuitively, it may appear simple to construct a non-parametric test for seasonal unit roots on the basis of by-season time-series plots, as they have been suggested by Franses (1994, 1996) and are being implemented into major software programs. Usually, occurring intersections among seasons—in most cases, quarters—are interpreted as evidence on changes in the seasonal pattern. However, this informal test is unlikely to be very rigorous, as variables with a weak seasonal pattern will intersect even more frequently. For this reason, we consider a test that parallels the known test procedure by Aparicio *et al.* (2006) in a context of possible seasonal unit roots.

This paper is organized as follows. Section 2 sets up the testing problem of interest and defines the maintained model. Section 3 considers the simple testing idea of monitoring seasonal crossings that was outlined above. Section 4 introduces our RURS (records unit root seasonal) test, a seasonal variant of the AES procedure. Section 5 presents some Monte Carlo simulations for assessing the size and power properties of the test. Section 6 considers the issue of robustness of the test to extended maintained hypotheses. Section 7 addresses the RURS test at the monthly frequency. Section 8 reports the application of the test to economics examples. Section 9 concludes.

2 The testing problem

We wish to consider time-series variables that are generated by autoregressions—possibly with unit roots—that are superseded with deterministic cycles and possibly a linear trend function. That is, x_t for $t = 1, 2, \ldots, n$ is assumed to follow

$$x_t = x_t^* + \sum_{j=1}^{4} \gamma_j^* D_{jt} + c^* t, \qquad (1)$$

with the purely stochastic autoregressive process (x_t^*), where D_{jt} denotes the usual seasonal dummy variables. For $t = 0, \ldots, -3$, x_t has some given non-stochastic starting values. At first, we constrain the autoregressive lag order to be less equal four, such that the AR(4) model

$$x_t^* = \sum_{j=1}^{4} a_j^* x_{t-j}^* + z_t$$

with white-noise z_t and starting values $x_t^* = 0$ for $t \leq 0$ is correctly specified. The process (z_t) is Gaussian white noise in the most restricted specification. More general conditions will be introduced below.

In the spirit of testing for seasonal unit roots, the focus is not on the properties of the deterministic cycle. In other words, a model with $\gamma_j^* = \gamma^*$ is seen as a special case of a model with 'deterministic seasonality' if the purely stochastic part (x_t^*) does not have seasonal unit roots.

The presentation here is tuned to the case of quarterly time series with the periodicity of four. The generalization to the monthly case or to any other periodicity is straight forward. However, these cases demand for some complexity in notation and introduce some subtle problems that will be addressed in Section 7.

In the representation (1), the interpretation of the constant $4^{-1}\sum_{j=1}^{4}\gamma_j^*$ is independent of the unit-root event $H_0 : \sum_{j=1}^{4} a_j^* = 1$. It is an average mean of the process (x_t) following linear de-trending by c^*t. Potential drift under H_0 is handled by the trend coefficient c^*.

The process (x_t) has an interesting alternative parameterization that is sometimes called the 'spectral' or 'cycles' representation:

$$\begin{aligned}\Delta_4 x_t &= \mu + a_1\left(x_{t-1}^{(1)} + ct\right) + a_2\left(x_{t-1}^{(2)} - \gamma_2(-1)^t\right) \\ &\quad + a_3\Delta_2 x_{t-2} + a_4\Delta_2 x_{t-1} \\ &\quad - (a_3+a_4)^2\left(\gamma_3 s_{t-2} + \gamma_4 s_{t-1}\right) + z_t, \end{aligned} \tag{2}$$

with the notation

$$\begin{aligned} x_t^{(1)} &= x_t + x_{t-1} + x_{t-2} + x_{t-3}, \\ x_t^{(2)} &= x_t - x_{t-1} + x_{t-2} - x_{t-3}, \\ \Delta_m x_t &= x_t - x_{t-m}, \quad m = 2,4, \\ s_t &= \sin \pi t/2. \end{aligned}$$

It is not difficult to see that equations (1) and (2) describe the same model class. The equivalence of the customary parameterization $(a_1^*, \ldots, a_4^*)$ of a fourth-order autoregression and $(a_1, \ldots, a_4)$ is the basis for the traditional HEGY test. The correspondence between the drift term c^* and the intercept μ is rather trivial, and the transition between $(\gamma_j^*, j = 1, \ldots, 4)$ and $(\gamma_j, j = 1, \ldots, 4)$ follows from the details in Smith and Taylor (1999). Note that the parameter dimension of representation (2) exactly corresponds to model (1), with $4^{-1}\mu$ denoting the drift under H_0 and the added linear trend only activated under H_0^C.

The handling of the deterministic terms serves as a 'bridle' that contains implausible expansion at the seasonal frequencies. Similar bridle devices are common in multivariate models (see Johansen,

1995) to contain trending behavior at the zero frequency in the case of unit roots. For a similar purpose, Dickey and Fuller (1979) suggested F–tests for the joint hypothesis of a unit root and zero constant. Note, however, that model (1) does not restrict the trend at the zero frequency. The effects of constants at the trend frequency and of deterministic cycles at the seasonal frequencies are different. A constant together with a unit root at +1 generates a linear trend in x_t, which is not an implausible specification for economic data. A trigonometric cycle together with a unit root at -1 or at $\pm i$ generates linear expansion of seasonal cycles, which may be regarded as implausible. See Franses and Kunst (1999) for a similar argument in multivariate seasonal models.

The spectral representation (2) is important for our purposes, as it directly motivates testing for the a_j coefficients in order to check events of seasonal unit roots. It is obvious that (x_t) 'has a unit root at +1' if and only if $a_1 = 0$, and it 'has a unit root at -1' iff $a_2 = 0$. The unit roots at $\pm i$ can only occur jointly, and they do so if and only if $(a_3, a_4) = (0, 0)$. The spectral representation also serves as a device for generating alternatives in the reported simulations, where γ_j rather than γ_j^* values will be set.

The range of admissible (z_t) terms is constrained by the so-called Berman condition, which suffices for the derivation of extremal properties. Our first assumption concerns the parametric space that we wish to investigate as the maintained hypothesis, the second assumption concerns the properties of the errors process. Throughout, we use B to denote the lag operator.

Assumption 1 *There is a representation*

$$(1-B)^{m_1}(1+B)^{m_2}(1+B^2)^{m_3}x_t = \tilde{x}_t,$$

such that $(\tilde{x}_t)$ *is stationary, where* $m_j \in \{0,1\}$ *for* $j = 1, \ldots, 3$. *The representation is unique if* m_j *is defined as the minimum value that achieves stationarity in* $\tilde{x}_t$.

The word 'stationary' is meant to include the possibility of transitory deviations from strict stationarity due to starting values and to allow for an added four-periodic deterministic cycle and a linear time trend. Assumption 1 is roughly equivalent to the following one:

Assumption 1 (′) *The polynomial* $\Phi(z)$ *defined by*

$$\begin{aligned}\Phi(z) &= 1 - z^4 - a_1(z+z^2+z^3+z^4) - a_2(z-z^2+z^3-z^4) \\ &\quad - a_3(z^2-z^4) - a_4(z-z^3)\end{aligned}$$

has no roots outside the set

$$\{z \in \mathbb{C} : |z| > 1\} \cup \{\pm 1, \pm i\},$$

and all roots within the latter set have multiplicity one.

Note that $\Phi(z)$ can be even uniquely defined if the generating law is an AR process with $p > 4$, for example by splitting off the roots with largest modulus. This would take care of the case of more general stationary but serially correlated z_t, for example stationary ARMA.

These assumptions exclude explosive roots, multiple unit roots, and unit roots at bizarre frequencies. We use the term 'bizarre' to describe unit roots at values different from the main spectral frequencies. For example, the second-order autoregressive model $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$ has the well known triangular stability area for its coefficients. One of the boundaries of the triangle, the open horizontal line segment $\{(\phi_1, -1), -2 < \phi_1 < 2\}$, describes models with bizarre frequencies, excepting the point $(0, -1)$. Such unit roots are rare in empirical applications.

Assumption 1 is satisfied for small negative a_1 and small positive a_2 and a_3. The coefficient a_4 interacts with a_3 in a complicated way. For $a_3 > 0$, the model will remain stable for a relatively wide range of a_4 values.

Assumption 2 *If at least one of the values m_j is 0, the stationary process $(\tilde{x}_t)$ considered in Assumption 1 fulfills a Berman condition. A stationary process with autocorrelations c_k is said to fulfill a Berman condition if $c_k \log k \to 0$ as $k \to \infty$.*

This condition is adopted by AES to define the alternative model for their unit-root test. For seasonal tests, it may make sense to demand this property not only from processes under the 'total' alternative of stationarity but also from transformed processes that are in the alternative of some hypothesis but in the null of others. For example, $(a_1, a_2, a_3, a_4) = (0, 1/2, 1/2, 1/2)$ defines a process with a unit root at $+1$ but no other unit root, in the symbols of Assumption 1 $(m_1, m_2, m_3) = (1, 0, 0)$. Its first differences $\tilde{x}_t = \Delta x_t$ are stationary and should have the Berman property.

Obviously, the usage of the liberal Berman condition to define the properties of the classes of interest entails that the tests are not designed to discriminate unit roots in the traditional I(1) sense from fractional alternatives. The Berman condition holds for stationary ARMA as well as for $I(d)$ with $d < 0.5$.

Technically, the Berman condition ensures the asymptotic validity of extremal distributions under the I(0) hypothesis if additionally normality is assumed. For non-normal distributions, general mixing conditions are required. A sufficient mixing condition is the $D(u_n)$ condition by Leadbetter and Rootzén (1988).

Depending on appropriate assumptions on the error process (z_t) that should cover stable autoregressions as well as invertible moving-average processes, the model class defined in (1) is pretty general. It comprises the seasonal random walk (SRW) for $(a_1, a_2, a_3, a_4) = (0, 0, 0, 0)$ with or without drift, and it also covers I(1) processes with added deterministic seasonal cycles. It also permits the occurrence of, for example, a seasonal unit root at -1 but not at $\pm i$ together with a quarterly deterministic seasonal cycle that can be expressed as the sum of s_t terms. The bridle excludes processes with expanding deterministic seasonal variation, and we also do not wish to consider superlinear trends.

It would be convenient if processes with seasonal unit roots, in generalization of the usual I(1) hypothesis, could be used as null hypotheses. Unfortunately, even in the non-seasonal model little is known about the properties of basic test statistics and their asymptotic distributions for general I(1) processes. The tests considered by AES and by Burridge and Guerre (1996) rely on the assumption of pure random walks, i.e. cumulative sums of i.i.d. random variables, under their null. This special case is much too restrictive for typical econometric applications, where increments show considerable positive serial correlation. We will introduce a correction that serves to at least approximately retain the distributional properties for the random-walk null in the correlated case. Robustness in this direction will be explored with the help of simulations.

Our test mainly builds on the 'range unit-root' (RUR) test that was suggested by AES. The test bears little relation to the range of the variables, so we prefer to read its acronym as 'records unit-root'. Its idea is to count the 'records', i.e. new extrema in the time-series sequence. Consider the maximum $x_{j,j}$ and the minimum $x_{1,j}$ of j successive time-series observations $x_k, k = 1, \ldots, j$. If a new minimum or maximum is encountered, as j is increased, this is called a *record*. In the following, the number of such records that occur until time point j will be denoted by $R(j)$, or, to mark the time-series label, as $R^{(x)}(j)$.

As j increases, the record count $R(j)$ will also increase. For random walks, it can be shown that $R(n) = O(n^{1/2})$ as $n \to \infty$. For i.i.d. sequences, the slower rate $R(n) = O(\log n)$ can be established. This proposition can be generalized to many stationary processes. AES give as the weakest possible condition the $D(u_n)$ mixing property that was mentioned above. At the other end of the scale, for drifting I(1) processes, the faster rate $R(n) = O(n)$ holds on quite general grounds.

For $a_1 = 0$ and $a_2 = a_3 = a_4 = 1/2$, $\Phi(z) = 1 - z$ and (x_t) becomes a random walk, assuming white noise (z_t). For the random walk with i.i.d. increments, we have the following theorem by AES:

Theorem 1 *If (x_t) is a random walk with i.i.d. increments, and if the probability law of these increments fulfills some regularity conditions (p.d.f. is bounded and continuous, second moments are finite, expectation is zero), the statistic $J_0^{(n)} = n^{-1/2} R^{(x)}(n)$ converges to a well-defined probability law as $n \to \infty$.*

AES provide a characterization of the asymptotic distribution and they show that it is well approximated in samples of moderate size. Roughly, the unimodal p.d.f. peaks at the value 2 and its left-tail critical points are around the value of 1. We will refer to this distribution as the 'AES distribution' in the following.

When there are no unit roots, i.e. if all roots of $\Phi(z)$ are confined to the set $\{z \in \mathbb{C} : |z| > 1\}$, AES prove another result:

Theorem 2 *If (x_t) is a stationary Berman process, the statistic $J_0^{(n)}$ converges to 0 in probability as $n \to \infty$.*

This theorem establishes the consistency of the RUR test that relies on the statistic J_0 with regard to its left tail against a stationary alternative. In fact, the test is also consistent in its right tail against a drifting alternative. However, this property will not be in focus in this paper. We further note that it is obvious that the procedure may have little power in the presence of sub-linear trends

Rather than in testing $H_0 : (a_1, \ldots, a_4) = (0, 1/2, 1/2, 1/2)$, we will be concerned with testing the more general hypotheses

$$\begin{aligned} H_{0+} &: a_1 = 0, \\ H_{0-} &: a_2 = 0, \\ H_{0i} &: a_3 = a_4 = 0, \end{aligned} \tag{3}$$

which correspond to unit roots at $+1$ ($m_1 = 1$), at -1 ($m_2 = 1$) and at $\pm i$ ($m_3 = 1$), respectively. In the following, we will address the hypotheses by these roots of $\Phi(z)$ and also alternatively by the corresponding angular frequencies at $\omega = 0$, $\omega = \pi$, and $\omega = \pi/2$.

Before we will fully introduce an adequate test for the hypotheses H_{0+}, H_{0-}, H_{0i} that parallels the AES idea, we consider a suggestion from the literature in the following section.

3 Non-parametric tests based on quarter graphs

Several monographs on seasonality recommend plots of time series by quarters (see, e.g., FRANSES, 1994, 1996, or GHYSELS AND OSBORN, 2001). Apparently, the frequency of rank changes among quarters is viewed as an indicator that eases the classification of the time series with respect to the main seasonal models. One may presume that seasonal unit-root processes entail infrequent rank changes, while deterministic seasonality might imply few changes when it is pronounced, and very frequent changes when it is weak.

Table 1 displays the rank changes counted in a small Monte Carlo experiment. First, seasonal random walks (SRW) were generated from zero starting values, and changes in ranking were counted. Then, random walks were generated with an added seasonal pattern that was fixed within one trajectory and drawn from four further normal random numbers. In 'RW+', the same variance was used for the increments and the seasonal pattern. In 'RW++', the variance of the seasonal pattern was 100 times the variance of the increments, which is not unusual in some empirical examples. 10,000 replications were used to determine quantiles of the empirical distribution and the mean.

The results conform to expectations and indicate that counting rank changes cannot be a very reliable test. For the weak seasonal pattern in RW+, the average frequency of rank changes is close to $n/2$, while for the strong seasonal pattern in RW++, the distribution is very asymmetric and its mean is of a magnitude that is comparable to the case of a seasonal random walk.

4 The RURS test

4.1 Records in ranges of seasonals

The results by AES on the RUR test suggest using the record count as a tool for discriminating stationary and integrated variables. The fact that the limit distribution for the random walk has an established although not simple form suggests using the random walk as a null hypothesis and rejecting for too small values—which are representative of the stationary alternative—as well as for too large values—which may represent processes with a deterministic drift. However, economists are rarely interested in testing for the existence of a drift and prefer to test for unit roots that reflect permanent effects of shocks. Therefore, we will use the tests in their one-sided versions and ex-

Table 1: Changes of ranking in simulated trajectories.

model	5%	10%	median	90%	95%	mean
SRW $n = 100$	6.0	7.0	14.0	24.0	27.0	15.06
SRW $n = 500$	17.0	21.0	37.0	58.0	65.0	38.41
SRW $n = 1000$	25.0	31.0	54.0	84.0	94.0	56.15
RW+ $n = 100$	22	26	43	64	69	44.4
RW+ $n = 500$	127	144	224	319	340	227.8
RW+ $n = 1000$	256	292	449	637	680	456.4
RW++ $n = 100$	0	0	3	14	19	5.80
RW++ $n = 500$	0	0	17	69	93	29.4
RW++ $n = 1000$	0	0	34	139	193	59.35

Note: SRW denotes that the generating model is a seasonal random walk. RW+ denotes a random walk with an added deterministic cycle that is drawn with the same variance as the increments. RW++ denotes that the seasonal standard error is ten times the standard deviation of the increments. Columns headed by percentages collect empirical quantiles.

tract potential deterministic time trends, in analogy to So and Shin (2001).

If x_t is an SRW $x_t = x_{t-4} + \varepsilon_t$ rather than a random walk, the autoregressive operator contains four unit roots at $\pm 1, \pm i$. Within the model class defined by (2) and Assumptions 1 and 2, one may be interested in considering the three null hypotheses H_{0+}, H_{0-}, H_{0i}. These hypotheses can be addressed after transforming the original SRW in order to eliminate all other unit roots that are not under immediate consideration.

First, the four-quarter moving average

$$x_t^{(1)} = x_t + x_{t-1} + x_{t-2} + x_{t-3}$$

is a pure random walk if x_t is a SRW. Parametric unit-root tests or the RUR test can then be applied to $x_t^{(1)}$. This statistic is denoted by J_1 here and is designed to test for H_{0+}. The properties of the RUR statistic under H_{0+} depend on the validity of the other unit-root hypotheses H_{0-} and H_{0i}. For example, if x_t is a random walk, it should be classified under H_{0+} but the increments of $x_t^{(1)}$ are autocorrelated and follow a third-order non-invertible moving-average process. The limiting distribution of the RUR statistic is not known for this case. For technical reasons, we convene the notation $x_t^{[1]} = x_t^{(1)}$.

Second, the four-quarter alternating moving average

$$x_t^{(2)} = x_t - x_{t-1} + x_{t-2} - x_{t-3}$$

is a ‘mirror’ random walk $x_t^{(2)} = -x_{t-1}^{(2)} + \varepsilon_t$ if (x_t) is a SRW. The adjusted process $x_t^{[2]} = (-1)^t x_t^{(2)}$ is a regular random walk and can be subjected to the RUR test. The corresponding RUR statistic is denoted by J_2 here and is designed to test for H_{0-}. The process $(x_t^{[2]})$ contains a unit root at $+1$ even if (x_t) is not a SRW, as long as (x_t) conforms to H_{0-}. Then, however, again the distribution of the RUR statistic will change.

Third, the second-order difference

$$\Delta_2 x_t = x_t - x_{t-2}$$

follows the simple law $\Delta_2 x_t = -\Delta_2 x_{t-2} + \varepsilon_t$ if x_t really is a SRW. Two separate random walks $x_t^{[3]}$ and $x_t^{[4]}$ can be constructed by sampling only every other observation and multiplying every other observation of either process by -1. Again, these two random walks can be subjected to RUR tests, and the RUR statistics will follow the AES distribution for either process. These statistics are denoted by J_3 and J_4 and they are designed to test for H_{0i}. However, they will deviate from the AES distribution if the original x_t contains a unit root at $\pm i$ but is not a SRW. Note that the sample size is halved for this test. In the following, the collection of the four tests based on $J_j, j = 1, \ldots, 4$ for the unit-root hypotheses H_{0+}, H_{0-}, H_{0i} will be called the RURS test, for ‘records unit-root seasonal’.

Table 2 gives simulated quantiles for all four statistics if the generating process is a SRW. The quantiles correspond well to those given by AES. Statistics J_1 and J_2 as well as J_3 and J_4 have almost identical empirical distributions even in fairly small samples, and their simulation results are given only once.

4.2 Lag augmentation

We first summarize the fundamental asymptotic properties in a theorem.

Theorem 3 *Under assumptions 1 and 2, the following two properties hold:*

(a) If (x_t) is a seasonal random walk with regular i.i.d. increments and thus is an element of $H_{0+} \cap H_{0-} \cap H_{0i}$, the distribution of all statistics $J_j, j = 1, \ldots, 4$ converges to the law indicated by AES;

(b) If (x_t) is in the alternative of any of the three hypotheses H_{0+}, H_{0-}, H_{0i}, the corresponding test statistic J_j will converge to 0 as $n \to \infty$.

Table 2: RURS test. Simulated quantiles from 10,000 replications.

model		1%	5%	10%	median	90%	95%	99%
SRW $n = 1000$	± 1	0.95	1.14	1.23	1.67	2.24	2.40	2.75
	$\pm i$	1.07	1.25	1.38	2.06	2.99	3.35	4.11
SRW $n = 500$	± 1	0.94	1.11	1.20	1.65	2.18	2.41	2.72
	$\pm i$	1.01	1.26	1.39	2.02	2.97	3.28	3.91
SRW $n = 100$	± 1	0.78	0.98	1.08	1.57	2.16	2.26	2.65
	$\pm i$	0.84	1.12	1.26	1.82	2.66	3.08	3.50

Note: SRW denotes that the generating model is a seasonal random walk. Rows ± 1 correspond to the J_1 and J_2 statistic with identical performance, while rows $\pm i$ correspond to the J_3 and J_4 statistics.

The proof of this theorem is obvious from AES. For an SRW, all transforms $(x_t^{[j]})$, $j = 1, \ldots, 4$ are random walks, and under the alternatives the Berman conditions will hold even for dynamic transforms of the processes that obey Berman's conditions because of assumption 2. Note that the theorem is silent, for example, on the properties of $J_j, j \neq 2$ on $H_{0+} \cap H_{0-}^c$, and even on processes in $H_{0+} \cap H_{0-} \cap H_{0i}$ that are not pure SRW.

Unfortunately, even for the standard RUR test by AES very little can be said about general I(1) processes that serve as the null hypothesis. While the limiting distribution is valid for nonlinear transformations of random walks, it is invalidated by deviations from the i.i.d. assumption on increments. A simple simulation shows that the AES distribution becomes indeed completely distorted if increments are serially correlated.

Thus, the test is *not similar* unless the null hypothesis is restricted to the very special case of a SRW with i.i.d. increments. In order to achieve approximate similarity for the more general null hypotheses of interest, we suggest a parametric autoregressive correction in the spirit of the ADF test. To this aim, we first consider the four variables $x_t^{[j]}, j = 1, \ldots, 4$

$$\begin{aligned} x_t^{[1]} &= x_t^{(1)}, \quad t = 1, \ldots, n, \\ x_t^{[2]} &= (-1)^t x_t^{(2)}, \quad t = 1, \ldots, n, \\ x_t^{[3]} &= (-1)^t \Delta_2 x_{2t}, \quad t = 1, \ldots, n/2, \\ x_t^{[4]} &= (-1)^t \Delta_2 x_{2t-1}, \quad t = 1, \ldots, n/2. \end{aligned} \tag{4}$$

Ideally, all of these variables follow random walks and their first differences are i.i.d. white noise. If there is low-order autoregressive autocorrelation, then the residuals

$$u_{j,t} = \Delta x_t^{[j]} - \mu - \sum_{k=1}^{p} \phi_k \Delta x_{t-k}^{[j]}$$

will be white noise. In practice, these true residuals will be replaced by least-squares regression residuals $\hat{u}$, although alternative estimation techniques may deserve consideration. In a final step, the residuals $\hat{u}_{j,t}$ can be accumulated again, and the cumulative sums,

$$\tilde{x}_t^{[j]} = x_p^{[j]} + \sum_{r=p+1}^{t} \hat{u}_{j,r}, \tag{5}$$

say, are then subjected to the original RUR test.

There is a danger that the 'augmenting' correction does more harm than good. If the analyzed process is really an ARIMA(p,1,0) in the familiar Box-Jenkins notation, consistent information criteria will find the true p, and at least for large n the procedure will yield a pure random walk. Unfortunately, if the original series does not have unit roots at all different frequencies, the process at hand will have a non-invertible ARIMA(p,1,q) structure and fitting autoregressions will result in large p and will nonetheless be unable to correct completely. This again will hamper the power of the test in case the alternative model at the investigated frequency holds.

Therefore, we tend to take care that p remains reasonably low. Parsimonious criteria such as the consistent BIC rather than liberal ones like AIC are an obvious choice. Also, we restrict the maximum order in the BIC search by a slow function of the sample size. Some experimentation with different upper bounds shows that, for moderate n, a bound of $n^{1/4}$ appears to be a good choice. In particular, $n^{1/4}$ yields better power properties than customary bounds proportional to $n^{1/3}$ at values distant from the null, where the augmentation tends to over-correct, at little expense for the size properties.

Essentially, we find indications for good performance but we also find critical issues in extensive simulations that we can only partially report. For example, Figure 1 demonstrates that the procedure works well with respect to the null distribution around the mode. The design for this simulation is a SRW with first-order autoregressive errors, and the evaluated test statistic is the one for the root at -1. The picture shows that the 10%, 50%, and 90% quantiles of the null distribution are unaffected by the autoregressive coefficient, even for non-stationary cases such as $+1$ and -1.

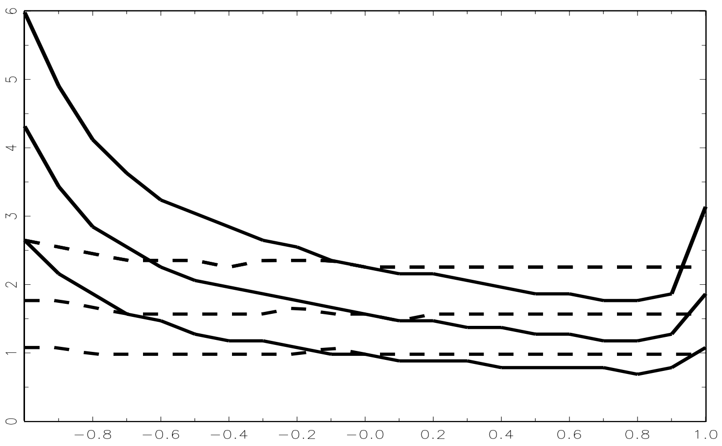

Figure 1: 10%, 50%, and 90% quantiles for the uncorrected (solid) and for the augmentation-corrected (dashed) RURS statistic $J_2^{(n)}$ if it is calculated from trajectories of length $n = 100$ from the data-generation process $\Delta_4 x_t = \phi \Delta_4 x_{t-1} + \varepsilon_t$ and ϕ is varied over the interval $[-1, 1]$. ϕ values on the abscissa.

The situation is comparable in the presence of moving-average errors, as Figure 2 demonstrates. The generation process is defined as $x_t = x_{t-4} + \varepsilon_t + \theta \varepsilon_{t-1}$, and θ varies over the interval $[-1, 1]$. Because the seasonal differences are a first-order moving average process, statistics J_3 and J_4 are unaffected. For J_1, $\theta = -1$ actually defines a process without a unit root at $+1$, as the unit roots in the autoregressive and the moving-average operator cancel. Similarly, for J_2, $\theta = 1$ defines an element of the alternative.

Both figures show that the uncorrected statistic $J_2^{(100)}$ is severely distorted. For the augmentation-corrected statistic, shown quantiles remain flat over all negative θ values and react for larger positive values. Whereas the resilience in the negative area is to be appreciated, the reduced sensitivity in the positive area points to a loss in power relative to the uncorrected original statistic. For example, the 1% quantile falls from the value $\theta = 0$ to $\theta = 1$ by a sizeable amount of 0.5 in the uncorrected case, while this difference reduces to 0.4 in the corrected case.

Similar simulations were conducted for the statistics J_1, J_3, J_4, and the performance was comparable to the H_{0-} case reported in Figures 1 and 2.

We note that the inclusion of a constant in the set of conditioning regressors, even if BIC suggests no augmentation, results in eliminating a possible drift as well as all seasonal deterministic terms. The latter property is grounded in the transformations that are used in

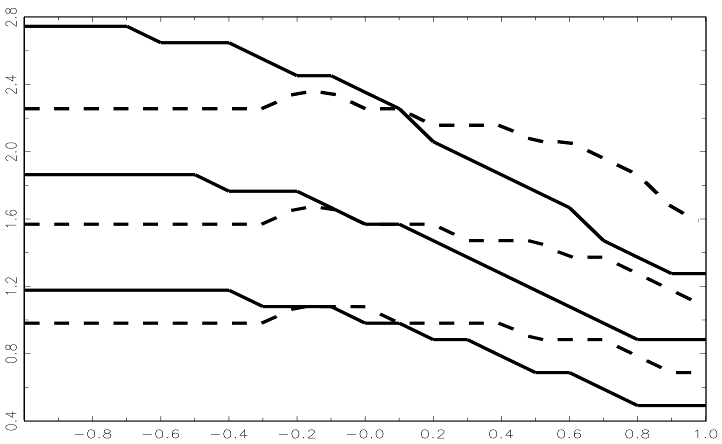

Figure 2: 10%, 50%, and 90% quantiles for the uncorrected (solid) and for the augmentation-corrected (dashed) RURS statistic $J_2^{(n)}$ if it is calculated from trajectories of length $n = 100$ from the data-generation process $\Delta_4 x_t = \varepsilon_t + \theta\varepsilon_{t-1}$ and θ is varied over the interval $[-1, 1]$. θ values on the abscissa.

constructing $x_t^{[j]}, j = 1, \ldots, 4$, which together with the sign-change and sample-split operations eliminate all cycles of length 2 and 4 observations. We also note that the constructed processes $x_t^{[j]}$ do not necessarily have zero mean, as the cumulation is started from actual starting values. It is obvious, however, that the counting of records is unaffected by universal level shifts. It is exactly these robustness properties that have motivated the usage of non-parametric tests like RUR in the presence of potential features that are so typical for many economic time series, such as local aberrations, outliers, and local level shifts.

A critical issue that may deserve further study, however, is the occurrence of outliers and the reaction of the augmenting step. It is known from the time-series literature that least-squares estimation is robust to *innovations outliers* but is critically affected by *additive outliers* (see also KLEINER *et al.*, 1976). Hence, whereas the correction works well for non-normal distributions of innovations, it is easy to construct examples where it fails for added outliers. If added outliers are a known feature of a data set, it may be advisable to replace the least-squares estimator by a robust procedure.

Table 3 reports a casual power simulation that can be compared to the results on the graphical test of Section 3. The RW+ design of Table 1 is used at the sample sizes $n = 100, 500, 1000$. For this process, J_1 should not reject, while $J_j, j = 2, \ldots, 4$ should reject the non-existing seasonal unit roots. This case is particularly interesting, as processes

$(x_t^{[j]})$ by construction have non-invertible moving-average parts that are a handicap for autoregressive augmentation. Augmentations are sizeable, double-digit lag orders are the rule for $n = 1000$. Table 3 demonstrates that power properties at the seasonal roots are acceptable but that the BIC augmentation is slightly too weak to preserve the size for the test based on J_1. Experiments with stronger augmentation, however, indicate a severe loss of power for the seasonal tests for H_{0-} and H_{0i} and tend to discourage more liberal augmentation.

Table 3: Power of RURS test. Simulated quantiles from 10,000 replications and rejection frequencies.

model		10%	median	90%	$r(0.01)$	$r(0.05)$	$r(0.1)$
RW+ $n = 100$	J_1	0.98	1.47	2.06	0.02	0.07	0.18
	J_2	0.78	1.08	1.57	0.10	0.26	0.52
	J_3, J_4	0.56	0.98	1.40	0.27	0.62	0.77
RW+ $n = 500$	J_1	0.95	1.38	1.96	0.13	0.23	0.30
	J_2	0.62	0.85	1.11	0.70	0.88	0.94
	J_3, J_4	0.44	0.63	0.88	0.98	1.00	1.00
RW+ $n = 1000$	J_1	0.88	1.36	1.96	0.14	0.30	0.36
	J_2	0.57	0.76	0.95	0.92	0.99	1.00
	J_3, J_4	0.36	0.49	0.67	1.00	1.00	1.00

Note: RW+ denotes that the generating model is a random walk with added seasonal constants drawn from a $N(0, 1)$ distribution. Row J_3, J_4 corresponds to both the J_3 and J_4 statistic with almost identical performance. $r(p)$ denotes rejection frequency when the RURS test is used at a significance level of p.

A more systematic study of size and power properties will be the topic of Section 5.

4.3 Forward and backward

In order to increase the power of their RUR statistic, AES suggest to run the test in both directions, that is to count records from $t = 1$ to $t = n$ and also from $t = n$ to $t = 1$, and then to average the two thus obtained test statistics. Indeed, the laws of motion of random walks as well as of stationary processes have known time reversibility properties, and the additional information can serve in improving test performance.

In line with the AES notation, we denote the thus obtained RURS-fb ('fb' for 'forward-backward') statistics by $J_{*j}, j = 1, \ldots, 4$. We also keep the convention by AES to define these statistics via

$$J_{*j} = 2^{-1/2}(J_j + J_j'),$$

rather than by the arithmetic mean, where J_j' denotes the reverse version of J_j. Note that, if J_j and J_j' were independent, this definition would increase scales by a factor $\sqrt{2}$. Table 4 provides a collection of corresponding significance points.

Table 4: Empirical null distribution of the RURS-fb test statistic.

statistic	n	1%	2.5%	5%	10%
J_{*1}, J_{*2}	100	0.81	1.03	1.18	1.77
J_{*3}, J_{*4}	100	1.07	1.28	1.49	1.92
J_{*1}, J_{*2}	200	0.82	1.08	1.23	1.85
J_{*3}, J_{*4}	200	1.11	1.40	1.47	1.99
J_{*1}, J_{*2}	300	0.88	1.08	1.25	1.88
J_{*3}, J_{*4}	300	1.19	1.43	1.55	2.08
J_{*1}, J_{*2}	400	0.90	1.11	1.26	1.87
J_{*3}, J_{*4}	400	1.18	1.44	1.54	2.10
J_{*1}, J_{*2}	500	0.90	1.12	1.28	1.89
J_{*3}, J_{*4}	500	1.23	1.42	1.56	2.15
J_{*1}, J_{*2}	600	0.91	1.11	1.26	1.90
J_{*3}, J_{*4}	600	1.25	1.46	1.58	2.13

Note: Simulated quantiles from 20,000 replications of seasonal random walks. n is sample size, columns denoted 'p%' denote quantiles of the empirical distribution.

In the remainder of the paper, the focus will be on this RURS-fb version.

5 Simulation evidence

5.1 Some remarks on distributions

In our simulations, we generally found that the empirical distribution does not coincide with the one reported by AES, neither at the frequency zero nor at $\pm i$. In this subsection, we attempt to explain the discrepancies.

Firstly, the limit distribution derived by AES builds on properties of Brownian motion. In a rough interpretation, the limit law describes the probability of range expansions over a Brownian motion process. This is not an accurate statement, as the concept of range expansion

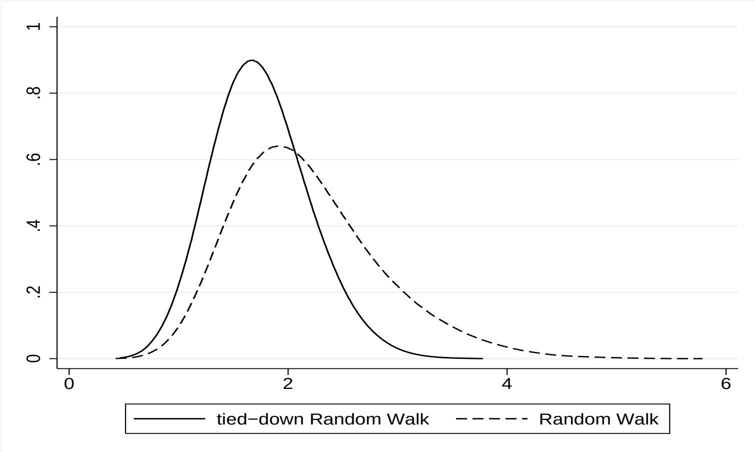

Figure 3: Empirical densities of the RUR statistic for 500 observations. Random walk and drift-adjusted random walk, 20000 replications, Epanechnikov kernel estimation.

makes little sense for a continuous-time process but it comes close to the asymptotic behavior.

By contrast, the intermediate augmenting step eliminates all trends and drifts in any random-walk trajectory. Therefore, the limit law does not build on a Brownian motion but rather on a Brownian bridge. Figure 3 compares the empirical distributions of the RUR statistic for random walks and for adjusted random walks. The adjustment will reduce findings of new extrema for many cases. Roughly, random walks sometimes expand near-monotonically and then extrema are becoming much sparser after correction. Conversely, sometimes random walks tend to generate pseudo-cycles, in which case extrema may become even more frequent after adjustment. The first effect appears to be stronger.

Second, the distribution of the RURS statistics does not even correspond to the Brownian bridge version due to the effects of the augmenting step.

A technical detail concerns the handling of the first few observations $x_t^{[j]}$ for $t < p$ (see equation (5)). We experimented with keeping them within the extrema search and also with excluding them. On the whole, we recommend running the search for records after excluding the starting observations, as these may mask other extrema in the presence of outliers.

Figure 4 conveys an impression of the empirical distributions for the actual statistics that are used in the RURS test.

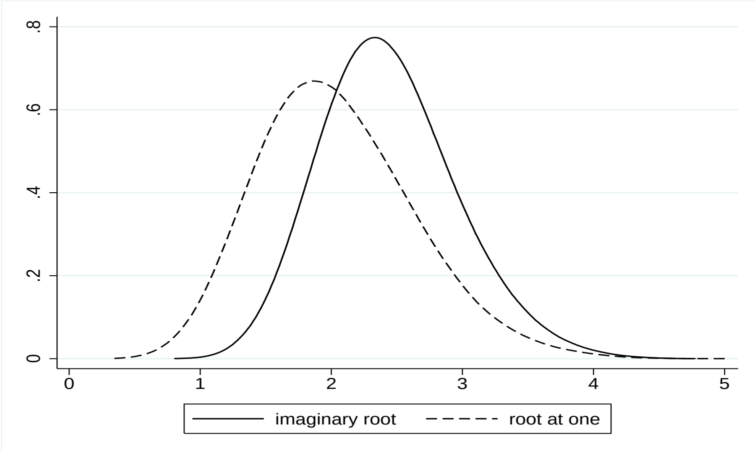

Figure 4: Empirical densities of the RURS statistic for 500 observations. Seasonal random walk, 20000 replications, Epanechnikov kernel estimation.

5.2 Size and power

In our size simulations, we compare the properties of the HEGY test and of the RURS test in its forward-backward (fb) variant. Quarterly seasonal random walks are generated in 10,000 replications for the (extended) sample sizes $n = 100$ to $n = 600$. These numbers correspond to 25 to 150 years and they are designed to cover the typical applications of economic relevance.

In detail, HEGY statistics are calculated from regressions

$$\begin{aligned}\Delta_4 x_t &= \mu + \sum_{j=1}^{4} \gamma_j D^*_{tj} + a_1 x^{(1)}_{t-1} + a_2 x^{(2)}_{t-1} + a_3 \Delta_2 x_{t-2} + a_4 \Delta_2 x_{t-1} \\ &\quad + \sum_{j=1}^{p} \zeta_j \Delta_4 x_{t-j} + u_t, \qquad (6)\end{aligned}$$

where the t–statistics on a_1 and a_2 and the F–statistics on (a_3, a_4) constitute the HEGY statistics on the three frequencies of concern. The lag order p is found by minimizing BIC in the Schwarz variant, just as described before for the RURS test.

Characteristics of the empirical distributions are obtained from these null simulations. Because of the BIC–guided lag orders, they differ slightly from the small-sample quantiles reported in the literature.

Next, instead of seasonal random walks, we simulate seasonally integrated processes of the type

$$x_t = x_{t-4} + \phi \Delta_4 x_{t-1} + \varepsilon_t,$$

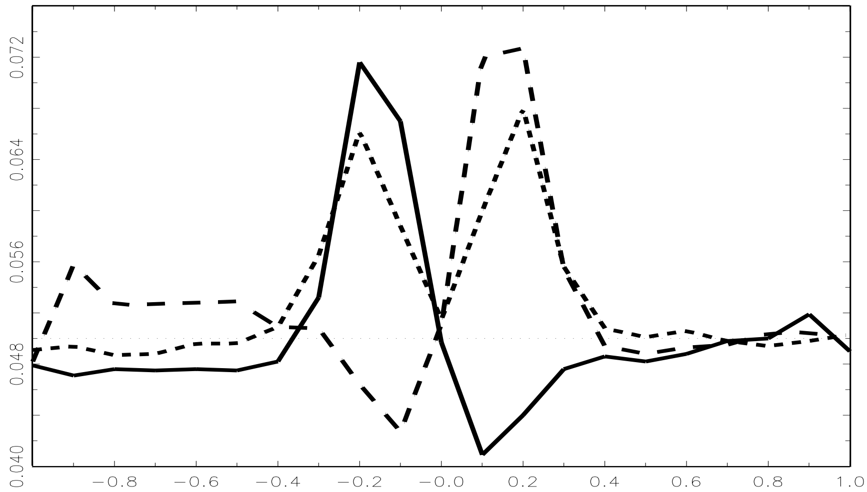

Figure 5: Empirical rejection frequency of unit-root null hypotheses using HEGY t– and F–tests for 200 observations at nominal 5% significance level. Solid curve for t_1, long dashes for t_2, and short dashes for F_{34}. Generating model is $\Delta_4 y_t = \phi \Delta_4 y_{t-1} + \varepsilon_t$ with N(0,1) errors ε_t, ϕ on the abscissa axis, 10000 replications.

with ϕ varying over the range $[-1, 1]$. Note that these processes would demand for a lag augmentation of $p = 1$ and that they have additional unit roots for the cases $\phi = \pm 1$.

Generally, the lag order search yields the expected results, with estimated p only slightly above 0 for $n = 200$ and $\phi = 0$ and increasing smoothly to mean estimates of $\hat{p} = 1$ as ϕ deviates from zero. Nevertheless, size distortions can be considerable. The exemplary Figure 5 shows rejection frequencies for the nominal 5% points for $n = 200$. The frequency-zero t_1 test is oversized for small negative ϕ values, the t_2 test for small positive ϕ values, and the F_{34} test for all small values, irrespective of their sign. By contrast, the performance at the extreme values $\phi = \pm 1$ is surprisingly good. The main qualitative features in Figure 5 persist at larger n.

The corresponding performance for the RURS test in its fb variant is shown in Figure 6 and is less convincing. The test under-rejects in some areas but this is of minor importance. The empirical distribution is discrete, and many values of the test statistic exactly match the significance points. For the same reason, the test for the 'annual' unit root at $\pm i$ shows a slight positive size bias that can be ignored. By contrast, the positive size bias for positive autocorrelation and the unit root at $+1$ and for negative autocorrelation at -1 is considerable and persists for larger samples. The augmenting step effectively improves

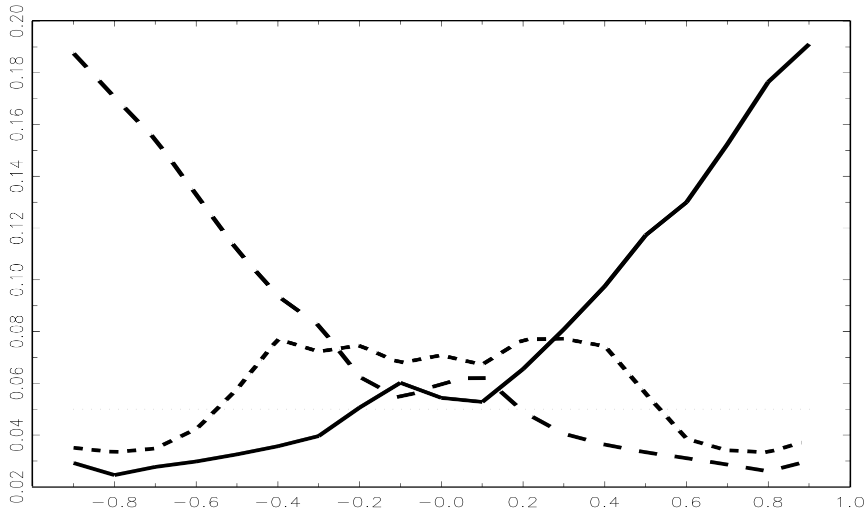

Figure 6: Empirical rejection frequency of unit-root null hypotheses using RURS-fb tests for 200 observations at nominal 5% significance level. Solid curve for $\omega = 0$, long dashes for $\omega = \pi$, and short dashes for $\omega = \pi/2$. Generating model is $\Delta_4 y_t = \phi \Delta_4 y_{t-1} + \varepsilon_t$ with N(0,1) errors ε_t, ϕ on the abscissa axis, 20,000 replications.

upon the original size bias but comes far from eliminating it, due to the small-sample bias of time-series coefficient estimates.

In the power simulations, we used the alternatives

$$\Delta_4 x_t = a_1 x_{t-1}^{(1)} + \varepsilon_t$$

for $a_1 \in (-1, 0]$ and

$$\Delta_4 x_t = a_2 x_{t-1}^{(2)} + \varepsilon_t$$

for $a_2 \in [0, 1)$. Excepting the boundary cases $a_1 = 0$ and $a_2 = 0$, the former alternative model is not in H_{0+} but in $H_{0-} \cap H_{0i}$, and the latter model is in $H_{0+} \cap H_{0i}$ but not in H_{0-}. Again, we used 20,000 replications and evaluated rejection frequencies for the RURS-fb and the HEGY tests against the significance points that were obtained from the size simulations.

For $n = 200$, the RURS-fb test yields the power curves depicted in Figures 7 and 8. The rejection frequency surpasses 20% as the value of a_1 or a_2 reaches 0.3 but it does not increase beyond that as the coefficients increase further. These graphs confirm the observation by AES that records tests have acceptable power close to the null but they are dominated by parametric tests at a larger distance from the null.

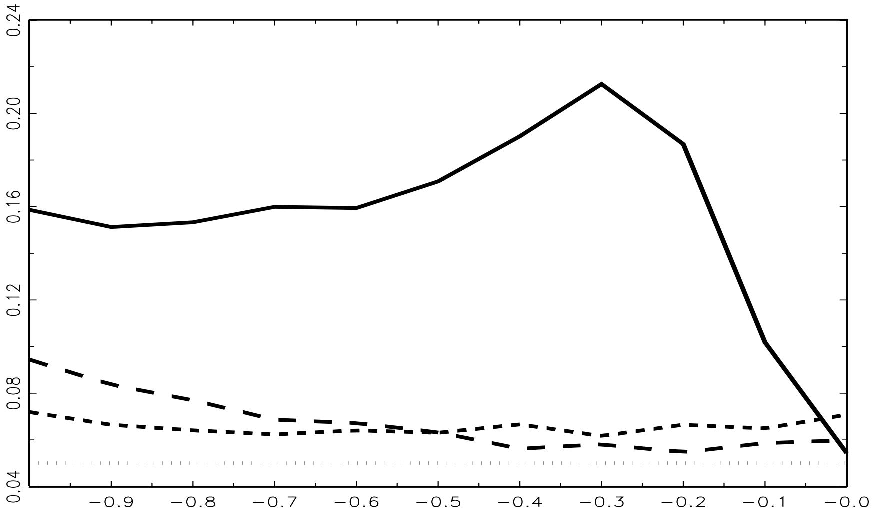

Figure 7: Empirical rejection frequency of unit-root null hypotheses using RURS-fb tests for 200 observations at nominal 5% significance level. Solid curve for $\omega = 0$, long dashes for $\omega = \pi$, and short dashes for $\omega = \pi/2$. Generating model is $\Delta_4 x_t = a_1 x_{t-1}^{(1)} + \varepsilon_t$ with N(0,1) errors ε_t, a_1 on the abscissa axis, 20,000 replications.

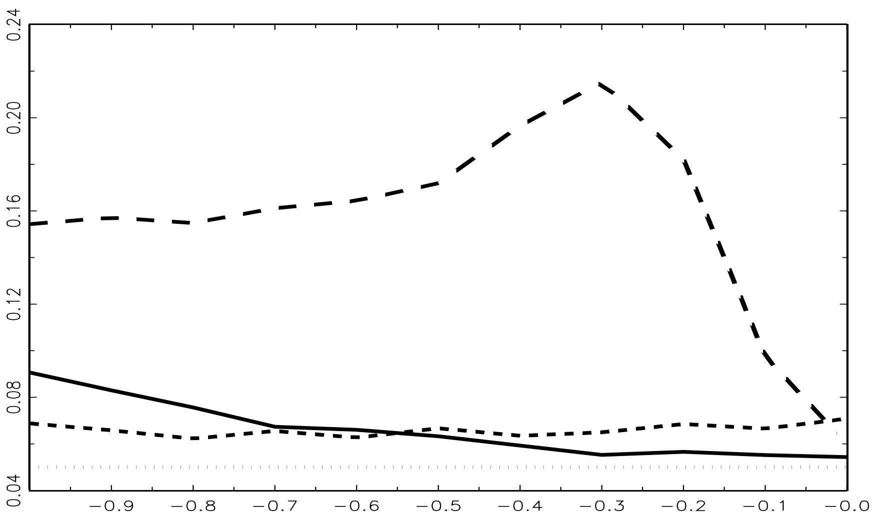

Figure 8: Empirical rejection frequency of unit-root null hypotheses using RURS-fb tests for 200 observations at nominal 5% significance level. Solid curve for $\omega = 0$, long dashes for $\omega = \pi$, and short dashes for $\omega = \pi/2$. Generating model is $\Delta_4 x_t = a_2 x_{t-1}^{(2)} + \varepsilon_t$ with N(0,1) errors ε_t, a_2 on the abscissa axis, 20000 replications.

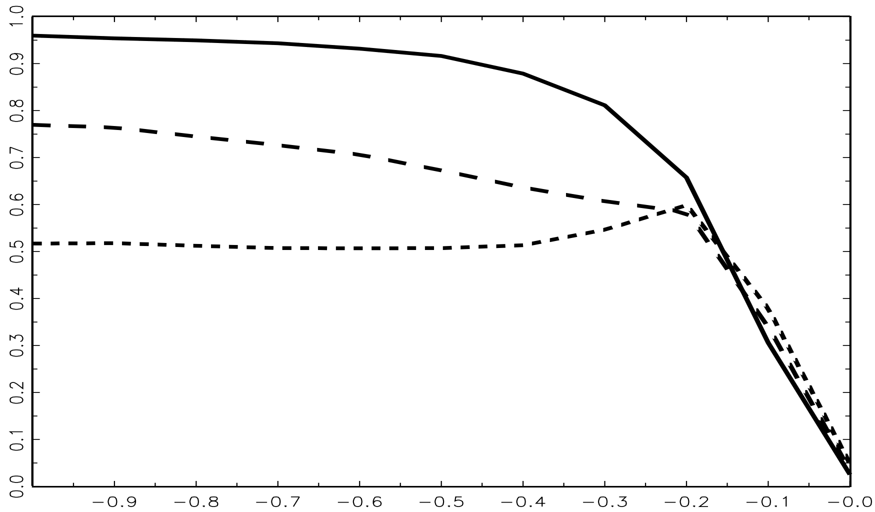

Figure 9: Empirical rejection frequency of unit-root null hypotheses at $\omega = \pi$ using RURS-fb tests for 500 observations at nominal 5% significance level. Upper bound for augmentation varies. Solid curve for the case with no augmentation, dashes for $N = n^{1/4}$, and short dashes for $N = 2n^{1/3}$. Generating model is $\Delta_4 x_t = a_2 x_{t-1}^{(2)} + \varepsilon_t$ with N(0,1) errors ε_t, a_2 on the abscissa axis, 20000 replications.

Note that these graphs were generated with the upper bound for conditioning lags of $2n^{1/3}$. For $n^{1/4}$, power becomes monotonic, and this feature is reproduced for larger n, such as $n = 500$ or $n = 1000$. For these larger samples, the RURS-fb test power indeed increases, thus corroborating our conjecture that the augmented version of the test is indeed consistent.

For moderate sample sizes, our simulations point to a severe lack of power for the nonparametric test even more if the test is compared to the parametric HEGY test. Even for $n = 200$, the HEGY test has an 86% rejection rate for values as small as $a_1 = 0.1$. This comparative impression is mitigated if deterministic seasonal patterns are introduced. For example, the design

$$\Delta_4 x_t = a_2 (x_{t-1}^{(2)} - \gamma_2 (-1)^t) + \varepsilon_t$$

tends to decrease HEGY power, while RURS-fb power, although still low, is not affected.

6 Extending the null and alternative

The application of nonparametric tests is typically motivated by their 'robustness', which formally means that the null and alternative of the test can be extended without affecting the properties of the test 'too much'. Particularly in the econometric literature, such extensions are rarely discussed and they are viewed as 'obvious'.

For example, if a random walk is subjected to 'structural breaks', a 'robust' test is casually defined as a test whose null distribution is not too much affected, such that it does not reject the random walk null. Similarly, a robust test would reject the random-walk or unit-root hypothesis for stationary processes to which deterministic elements such as breaking trends or outliers are added. We emphasize that it is not immediately obvious that such properties on extended null and alternative hypotheses are to be seen as beneficial.

A systematic extension of the investigated hypotheses was suggested by Burridge and Guerre (1996) who, following earlier work by Granger and Hallman (1991), consider extending the null of I(1) processes to all monotonic and continuous functions of I(1) processes. Under the alternative, a transformation of a strictly stationary ergodic process is stationary anyway.

Adopting this casual definition of natural extensions for our purposes, we would see it as beneficial if the null distribution of the RURS test statistic were immune to monotonic transformations of seasonally integrated variables. For example, a logit transform of a process in $H_{0+} \cap H_{0i}$ should generate the typical null distribution for the statistics J_{*1} and J_{*3}, J_{*4}, while J_{*2} should be sensitive to violations of H_{0-}. We evaluate the test performance in this direction by a small Monte Carlo experiment.

The variable x is defined as the logit of y, i.e.

$$x_t = \frac{\exp y_t}{1 + \exp y_t}, \tag{7}$$

$$\Delta_4 y_t = a_2 y_{t-1}^{(2)} + \varepsilon_t. \tag{8}$$

The coefficient a_2 is varied over (-1,0].

Figure 10 shows that the performance of the seasonal test statistics $J_{*j}, j = 3, 4$ is hardly affected by the logit transformation. The test based on J_{*2} has notoriously low power. The test based on J_{*1} tends to reject H_{0+}, while the data generating process is, in a sense, a member of H_{0+}. The test construction involves a dynamic filter that operates on a nonlinear transformation, which is not the same as the transformation of the filtered data. For large $|a_2|$, the test rejects the unit root at +1, although it exists in the extended definition, and

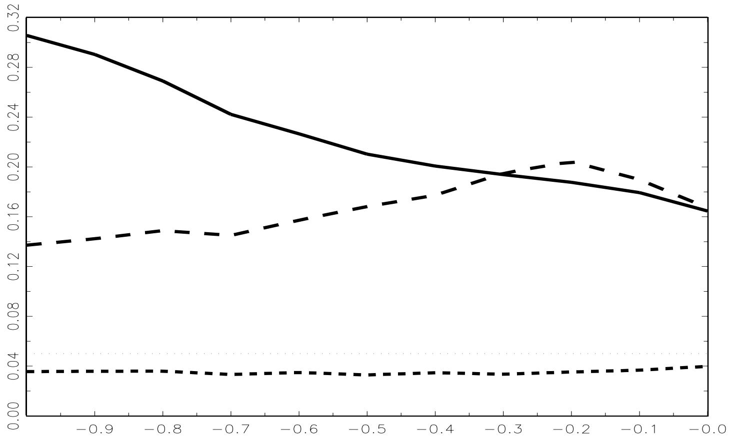

Figure 10: Empirical rejection frequency of unit-root null hypotheses using RURS-fb tests for 200 observations at nominal 5% significance level. Solid curve for $\omega = 0$, long dashes for $\omega = \pi$, and short dashes for $\omega = \pi/2$. Generating model is the logit transformation of x for $\Delta_4 x_t = a_2 x_{t-1}^{(2)} + \varepsilon_t$ with N(0,1) errors ε_t. a_2 on the abscissa axis, 20000 replications.

tends to accept the unit root at -1, even though it does not exist in the data generating process.

The same features—low power for deviations from H_{0-} and excessive size for the test on H_{0+}—concern the parametric HEGY test. At least with regard to the logit transformation, neither of the two test concepts has the kind of robustness that may be in the spirit of the extended hypothesis definition of Granger and Hallman (1991).

Another, albeit more promising, experiment concerns the robustness of tests to outliers. Figure 11 reports an arguably extreme design, in which a value of 10.0 was added to the generated series at the location $n/2$. Otherwise, the design is identical to the hitherto used process $\Delta_4 x_t = a_2 x_{t-1}^{(2)} + \varepsilon_t$ with $a_2 \in [-1, 0]$. The performance of the parametric HEGY test is affected severely. The test rejects in nearly all replications even for a comparatively large sample of $n = 500$. Note that it does not only reject H_{0-}, which it is supposed to reject for $a_2 \neq 0$, it also rejects at the null case $a_2 = 0$, and it also rejects H_{0+}, which is supposed to be valid for the extended interpretation of the model. By contrast, the nonparametric RURS-fb test displays its—concededly low—power curve at π and hardly any size bias at 0.

We also ran comparable experiments with smaller added outliers and with structural breaks, and these generally correspond to this pattern. The parametric tests over-reject the hypotheses H_{0+} and H_{0i}, whereas the RURS-fb test is immune to the 'aliasing' feature

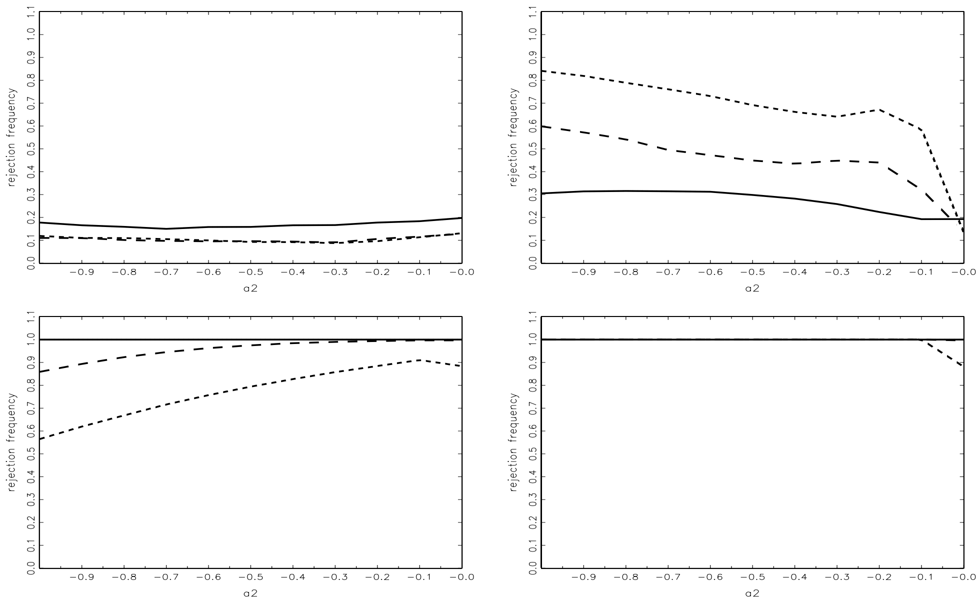

Figure 11: Empirical rejection frequency of hypothesis H_{0+} (left panels) and H_{0-} (right panels) at nominal 10% significance level. Data generation process is $\Delta_4 x_t = a_2 x_{t-1}^{(2)} + \varepsilon_t$ with N(0,1) errors ε_t. Top panels for RURS-fb test and bottom panels for HEGY test. Solid curve for 100 observations, dashes for 300, and short dashes for 500 observations. 20000 replications.

and attains reasonable power properties, as the sample size increases. Note that both testing procedures involve a least-squares regression step that can be replaced by a robust regression procedure. This modification tends to improve the statistical properties of both tests.

7 The monthly version of the RURS test

It is straight forward to generalize the basic model of equation (1) to

$$x_t = x_t^* + \sum_{j=1}^{12} \gamma_j^* D_{jt} + c^* t, \qquad (9)$$

for the process x_t^* now generated by a twelfth-order autoregression. The representation by cycles (or 'spectral' representation) in equation (2) can then be modified accordingly. For the monthly case, the model contains 12 'spectral' coefficients, an unrestricted intercept, and 11 'spectral' constants. The contribution of the deterministic cycles is contained by quadratic expressions that de-activate the contribution in the presence of corresponding seasonal unit roots.

In detail, we consider the representation:

$$\begin{aligned}\Delta_{12} x_t &= \mu + a_0 \left(x_{t-1}^{(0)} + ct \right) + a_6 \left(x_{t-1}^{(6)} - \gamma_6 (-1)^t \right) \\ &+ \sum_{j=1}^{5} \left\{ a_j x_{t-1}^{(j,\alpha)} + b_j x_{t-1}^{(j,\beta)} \right. \\ &\left. + (a_j + b_j)^2 \left(\gamma_{j,1} \cos \frac{j\pi}{6} + \gamma_{j,2} \sin \frac{j\pi}{6} \right) \right\} + z_t, \qquad (10)\end{aligned}$$

with the notation

$$\begin{aligned} x_t^{(0)} &= \sum_{j=0}^{11} x_{t-j}, \\ x_t^{(6)} &= \sum_{j=0}^{11} (-1)^j x_{t-j}, \\ \Delta_m x_t &= x_t - x_{t-m}, \\ x_t^{(j,\alpha)} &= -\nabla_j \left(\cos \frac{j\pi}{6} - B \right) x_t, \\ x_t^{(j,\beta)} &= -\nabla_j \left(\sin \frac{j\pi}{6} \right) x_t, \\ \nabla_j &= (1 - B^2) \prod_{k=1, k \neq j}^{5} \left(1 - 2 \cos \frac{k\pi}{6} + B^2 \right). \end{aligned}$$

Details of this representation were analyzed by SMITH AND TAYLOR (1999), and the model was used for developing HEGY-type tests by BEAULIEU AND MIRON (1993). Again, $(a_0, a_1, \ldots, a_6, b_1, \ldots, b_5) =$

$(0, \ldots, 0)$ defines a process with unit roots at all frequencies, a seasonal random walk if additionally $\mu = 0$ and white-noise (z_t) is assumed.

The objective of testing is obvious. Apart from analyzing the inherited hypotheses $H_{0+} = H_{00} : a_0 = 0$, $H_{0-} = H_{06} : a_6 = 0$, and $H_{0i} = H_{03} : a_3 = b_3 = 0$, we also wish to consider the hypotheses $H_{0j} : a_j = b_j = 0$ at $j = 1, 2, 4, 5$ that correspond to the angular frequencies $\pi/6, \pi/3, 2\pi/3, 5\pi/6$.

In analogy to the quarterly case, parametric testing based on regressions yields HEGY-type F–statistics on $a_j = b_j = 0$ for $j = 1. \ldots, 5$, and t–statistics on $a_0 = 0$ and $a_6 = 0$. These hypotheses are then equivalent to seasonal unit-root events at frequencies $\omega = j\pi/6$ for $j = 0, \ldots, 6$.

Similarly, time-series operators that eliminate unit roots at all frequencies excepting a specific one can be used to construct RURS and RURS-fb statistics. For the backdrop example of a seasonal random walk, however, only the cases $x_t^{(j)}$ for $j = 0, 3, 6$ are random walks or are simple transforms thereof. In all other cases, these series will not be random walks. The non-trivial fact that random elements of the type $n^{-1/2} x_t^{(j)}$ obey functional limit theorems with standard Brownian motion limits has been proofed by Chan and Wei (1988). As a consequence, the asymptotic properties of extremal statistics such as J_j and J_{*j} will quantitatively match those of the RUR statistics. Small-sample quantiles, however, vary somewhat across frequencies, and a detailed analysis of their limit distribution has not yet been conducted.

Table 5: Empirical significance points for the monthly RURS test statistics for $n = 100$.

statistic	frequency	n	1%	5%	10%	median
J_0	0	100	0.995	1.194	1.293	2.059
J_1	$\pi/6$	100	0.697	0.995	1.095	1.806
J_2	$\pi/3$	100	0.498	0.796	0.995	1.678
J_3	$\pi/2$	100	0.885	1.180	1.327	1.959
J_4	$2\pi/3$	100	0.597	0.896	0.995	1.816
J_5	$5\pi/6$	100	0.498	0.697	0.896	1.601
J_6	π	100	0.995	1.194	1.393	2.058

Note: Empirical quantiles from 10,000 replications. Generating model is the SRW $x_t = x_{t-12} + \varepsilon_t$ with i.i.d. N(0,1) errors.

Note that there is no direct counterpart to the processes $x_t^{[j]}$ at j different from $0, 3, 6$ in the monthly test. The processes $x_t^{(j)}$ will not even be random walks under $\bigcap_{j=0}^{6} H_{0j}$. There are two obvious suggestions to conduct the accumulation step after identification of a lag order and collecting residuals. The first is to accumulate according to the law of motion under the null, which leads to processes with a pure seasonal unit root at given periodicity. This suggestion was followed for Table 5.

Table 5 gives the empirical significance points for the case $n = 100$. As for the quarterly case, power performance is satisfactory against stationary alternatives of the type $x_t = \phi x_{t-12} + \varepsilon_t$ with $|\phi| < 1$ but much less so if individual unit-root hypotheses are considered.

Table 6: Empirical significance points for the monthly RURS-fb test statistics for $n = 200$.

statistic	frequency	n	1%	5%	10%	median
J_0	0	200	1.439	1.652	1.812	2.345
J_1	$\pi/6$	200	1.439	1.652	1.759	2.345
J_2	$\pi/3$	200	1.439	1.652	1.759	2.345
J_3	$\pi/2$	200	1.410	1.632	1.773	2.337
J_4	$2\pi/3$	200	1.439	1.652	1.759	2.345
J_5	$5\pi/6$	200	1.439	1.652	1.759	2.345
J_6	π	200	1.439	1.652	1.812	2.345

Note: Empirical quantiles from 10,000 replications. Generating model is the SRW $x_t = x_{t-12} + \varepsilon_t$ with i.i.d. N(0,1) errors.

An alternative suggestion is to accumulate the residuals to random walks under their nulls H_{0j}, i.e. in exact analogy to equation (5) in the quarterly model. This variant leads to identical significance points across frequencies, excepting $\omega = \pi/2$, where less effective observations are utilized. Otherwise, unreported power simulations did not find any consistent evidence in favor of either version.

Table 6 gives empirical quantiles for the RURS-fb version of the test at $n = 200$. These simulations are based on the random-walk accumulation method in the augmenting step as described above, which yields almost perfect homogeneity across frequencies. These quantiles experience little changes if the sample size is modified, thus they can serve as benchmarks for tests at various empirically relevant n.

8 Empirical applications

Figure 12 shows the price of Belgian barley, in quarterly observations from 1971 to the first quarter of 2003. The series has a clearly recog-

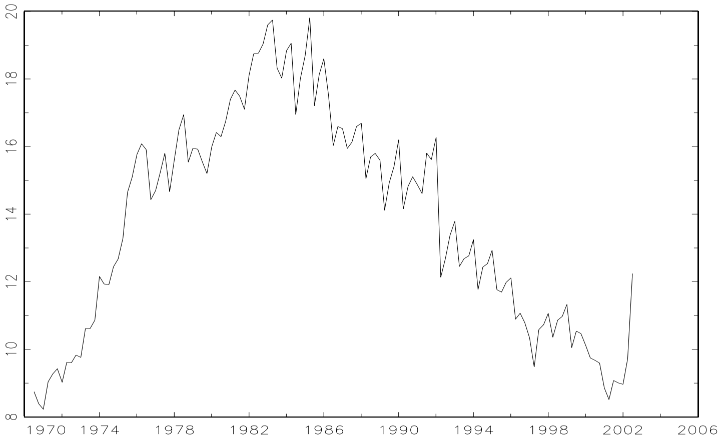

Figure 12: Belgian barley prices, 1971:1–2003:1, quarterly observations.

nizable seasonal structure but it is subject to many irregularities.

There are 129 observations, so we utilize the quantiles from Table 2 for $n = 100$. The BIC search procedure finds lags of $p = 6$ for the J_1 and the J_2 test but only $p = 3$ and $p = 4$ for the J_3 and J_4 tests. The calculated statistics are $J_1 = 2.03$, $J_2 = 1.85$, $J_3 = J_4 = 1.00$. The unit roots at $+1$ and at -1 are supported while the unit-root pair at $\pm i$ is rejected at the 5% although not at the 1% level. Therefore, the RURS test indicates that the observed seasonal variation is composed of a deterministic component at the annual frequency and a persistently changing semi-annual component. The finding of a unit root at $+1$ conforms to intuition.

By contrast, the application of traditional parametric tests leads to unclear results. The HEGY test rejects unit roots at -1 and also at $\pm i$ at the 5% level, while the CH tests by Canova and Hansen (1995) tend to reject deterministic seasonality. Also, these tests are sensitive to transformations of the original data, for example by logarithms, while the RURS test is robust in this direction.

For the monthly version of the test, we analyze the Austrian unemployment rate shown in Figure 13 from January 1950 to December 2005. The series displays strong seasonal variation, and the cycles do not appear to be constant over time.

For the unemployment rate, the corresponding RURS statistics are, ordered from the lowest frequency to the Nyqvist frequency: 0.89, (0.70, 0.43), (0.31, 0.35), (0.94, 0.78), (0.86, 1.01), (0.58, 0.47), 0.70. Here, statistics for the same frequency have been collected in parentheses. A comparison with Table 5 shows that most values are significant

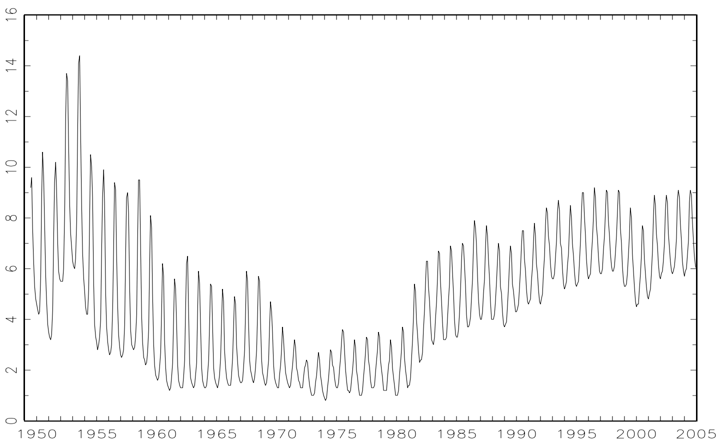

Figure 13: Austrian unemployment rate, 1950:1–2005:12, monthly observations.

at 5% , excepting the frequency $2\pi/3$. Three hypotheses are even rejected at the 1% level, including the long-run root at +1. This result indicates that the seasonality of the unemployment rate is primarily deterministic and that most apparent pattern changes can be well described by an unspecified nonlinear transformation of a stationary variable with an added deterministic cycle.

The fb version of the RURS test insinuates a slightly different conclusion. Test statistics are 2.85, 2.40, 2.93, 1.493, 2.958, 2.958, 1.340. Seasonal unit roots are accepted at most frequencies, including the long-run frequency, but not at -1 and at $\pm i$. In other words, there is a random-walk type trend in the data, and the main backbone of the annual seasonal cycle is purely deterministic, while the subtle details of variation from month to month experience persistent changes over the observation period.

The conclusion of the parametric HEGY test differs from both RURS versions. The HEGY test does not reject unit roots at the low frequencies but rejects them at $\pi/2$, i.e. at four-month cycles, and at all higher frequencies. The parametric test sees the unemployment rate as driven by longer stochastic waves and by short-run cycles of fixed structure and considerable importance. Of course, the true data generating process for this variable is unknown. We presume that the RURS and RURS-fb results are more credible, however, as these tests allow for nonlinear transformations. Additionally, the HEGY-test classification of the bounded rate as a traditional I(1) process does not make much sense.

9 Concluding remarks

Nonparametric unit-root tests, such as the RUR test by AES, are a welcome addendum to the toolbox of time-series analysis. While these tests cannot be as reliable as their parametric counterparts for regular well-behaved Gaussian linear processes, many economic variables indicate that they do not conform to these standard assumptions. Then, however, nonparametric procedures that circumvent moment estimation show their advantages.

We conjecture that tests for seasonality suffer from such irregularities even more commonly than tests for trend behavior. Seasonal cycles are subject to sudden and also gradual changes, then shift back to their original form or disappear for a few years before turning up in a different shape. Therefore, if the concept of seasonal unit roots is of interest to the researcher, there is sufficient motivation for tools beyond traditional tests.

While the traditional plot of time series decomposed into quarters aids a lot in getting a feeling for the data, we find that the visual impression is difficult to quantify in a statistical test procedure. Rather, we consider a generalization of the recently developed RUR (records unit roots) test by AES to the RURS test for seasonal unit roots. We suggest to utilize a parametric correction in its calculation, as the test is sensitive to autocorrelation under its null hypothesis. In simulations, we find some support for our augmentation correction. The optimal tradeoff between size properties that would demand for stronger augmentation and power properties that would demand for less augmentation is unknown as yet.

In particular for mixtures of the null hypothesis at one seasonal frequency and the alternative at other frequencies did we find disappointingly low power for the RURS test. For this reason, we would not recommend to replace currently used parametric test by nonparametric tests, not even if features such as breaks, outliers, and nonlinear transformations are suspected. As an additional information, these tests may prove useful to researchers, however.

The development of nonparametric tests is much more informal and guided by creative intuition than that of parametric tests that live in a world dictated by likelihood theory. The occurrence of records in the range of a time series is only one out of many conceivable criteria that can be used for discriminating I(1) from I(0) processes. The study of records is attractive, as the corresponding statistics have been shown to be invariant to distributional properties and robust to outliers. These important results, however, do not rule out the continuing search for new procedures.

References

[1] Aparicio, F., Escribano, A., and Sipols, A.E. (2006) 'Range unit-root (RUR) tests: robust against nonlinearities, error distributions, structural breaks and outliers'. *Journal of Time Series Analysis* **27**, 545–576.

[2] Beaulieu, J.J., and Miron, J.A. (1993) 'Seasonal unit roots in aggregate U.S. data,' *Journal of Econometrics* **55**, 305–328.

[3] Burridge, P., and Guerre, E. (1996) 'The limit distribution of level crossings of a random walk,' *Econometric Theory* **12**, 705–723.

[4] Caner, M. (1998) 'A Locally Optimal Seasonal Unit-Root Test'. *Journal of Business and Economic Statistics* **16**, 349–356.

[5] Canova, F., and Hansen, B.E. (1995) 'Are Seasonal Patterns Constant over Time? A Test for Seasonal Stability', *Journal of Business and Economic Statistics* **13**, 237–252.

[6] Chan, N.H., and Wei, C.Z. (1988) 'Limiting Distributions of Least Squares Estimates of Unstable Autoregressive Processes,' *The Annals of Statistics* **16**, 367–401.

[7] Choi, C.Y., and Moh, Y.K. (2007) 'How useful are tests for unit-root in distinguishing unit-root processes from stationary but non-linear processes?' *Econometrics Journal* **10**, 82–112.

[8] Dickey, D.A., and Fuller, W. (1979) 'Distribution of the Estimators for Autoregressive Time Series with a Unit Root,' *Journal of the American Statistical Association* **74**(366), 427–431.

[9] Dickey, D.A., Hasza, D.P., and Fuller, W.A. (1984) 'Testing for unit roots in seasonal time series,' *Journal of the American Statistical Association* **79**, 355–367.

[10] Franses, P.H.F. 'A Multivariate Approach to Modeling Univariate Seasonal Time Series,' *Journal of Econometrics* **63**, 133–151.

[11] Franses, P.H.F. (1996) *Periodicity and Stochastic Trends in Economic Time Series*. Oxford University Press.

[12] Franses, P.H.F., and Kunst, R.M. (1999) 'On the Role of Seasonal Intercepts in Seasonal Cointegration,' *Oxford Bulletin of Economics and Statistics* **61**(3), 409–433.

[13] Ghysels, E., and Osborn, D.R. (2001). *The Econometric Analysis of Seasonal Time Series*, Cambridge University Press.

[14] Granger, C.W.J., and Hallman, J. (1991) 'Nonlinear transformations of integrated time series,' *Journal of Time Series Analysis* **12**, 207–224.

[15] Hylleberg, S., Engle, R.F., Granger, C.W.J. and Yoo, B.S. (1990). 'Seasonal integration and cointegration', *Journal of Econometrics* **44**, 215–238.

[16] Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press.

[17] Kleiner, B., Martin, R.D., and Thomson, D.J. (1976) 'Three approaches towards making power spectra less vulnerable to outliers,' *ASA Proceedings of the Business and Economics Statistics Section* 386–391.

[18] Leadbetter, M.R., and Rootzén, H. (1988) 'Extremal theory for stochastic processes,' *Annals of Probability* **16**, 431–478.

[19] Smith, R.J., and Taylor, A.M.R. (1999) 'Likelihood ratio tests for seasonal unit roots,' *Journal of Time Series Analysis* **20**, 453–476.

[20] So, B.S., and Shin, D.W. (2001) 'An invariant sign test for random walks based on recursive median adjustment,' *Journal of Econometrics* **102**, 197–229.

[21] Tam, W.-K., and Reinsel, G.C. (1997) 'Tests for Seasonal Unit Roots in ARIMA Models,' *Journal of the American Statistical Association* **92**, 725–738.

Author: Robert M. Kunst

Title: A Nonparametric Test for Seasonal Unit Roots

Reihe Ökonomie / Economics Series 233

Editor: Robert M. Kunst (Econometrics)
Associate Editors: Walter Fisher (Macroeconomics), Klaus Ritzberger (Microeconomics)

ISSN: 1605-7996

Stumpergasse 56, A-1060 Vienna • ☎ +43 1 59991-0 • Fax +43 1 59991-555 • http://www.ihs.ac.at

ISSN: 1605-7996