Testing for seasonal unit roots in monthly panels of time series

Robert M. Kunst *
and
Philip Hans Franses †

Econometric Institute Report EI 2009-05

Abstract

We consider the problem of testing for seasonal unit roots in monthly panel data. To this aim, we generalize the quarterly CHEGY test to the monthly case. This parametric test is contrasted with a new nonparametric test, which is the panel counterpart to the univariate RURS test that relies on counting extrema in time series. All methods are applied to an empirical data set on tourism in Austrian provinces. The power properties of the tests are evaluated in simulation experiments that are tuned to the tourism data.

JEL Codes: C12, C14, C23.
Keywords: seasonality, nonparametric test, unit roots, panel, tourism.

*Institute for Advanced Studies Vienna, Stumpergasse 56, 1060 Wien (Vienna), Austria. Telephone +43-1-59991-255, fax +43-1-59991-163, e-mail kunst@ihs.ac.at
†Erasmus University Rotterdam

1 Introduction

While unit roots at the zero frequency are related to concepts such as trend or mean reversion, testing for seasonal unit roots essentially means investigating the degree of reversion of seasonal cycles to time-constant equilibrium patterns. In all unit-root tests on economic time series of limited length, discriminatory power is notoriously low. Repeated observations on series with comparable properties, as they are given in panel data, may serve convenient in increasing that power. Compared to the sizeable literature on panel unit roots, tests on seasonal roots in panels have hitherto drawn much less attention. We just mention the contributions of Otero *et al.* (2005,2007) as well as of Ucar and Guler (2007) and of Dreger and Reimers (2004). Most of this research focuses on the case of quarterly data. In this paper, we consider the monthly case.

The panel literature tends to describe unit-root tests under the assumption of homogeneity and independence in the cross-section dimension as the 'first-generation' tests and tests that admit heterogeneity and static cross-section correlation as 'second-generation' tests (see Hlouskova and Wagner, 2006). In this sense, the above authors already focus on second-generation tests, as seasonal unit roots have drawn little attention in the age of the first generation. In the following, we will adopt the CHEGY test by Otero *et al.* (2007) and we will contrast it with a nonparametric test that follows the univariate RURS test introduced by Kunst (2009) as a seasonal generalization of the RUR ('record unit-root') test by Aparicio *et al.* (2006). Due to its construction, the RURS panel test is unlikely to be much affected by heterogeneity and cross-section correlation.

There is no general agreement on the correct definition of nonparametric tests. In fact, most unit-root tests contain semiparametric elements, such as the Phillips-Perron test and its generalizations to multivariate and panel problems. Truly nonparametric tests tend to radically digress from the concept of approximating likelihood-ratio tests based on a backdrop parametric model. They achieve high robustness to some deviations from usual assumptions, such as strong breaks and outliers, at the cost of low power in standard situations. Variance-bounds tests, as considered by Breitung (2002), constitute an intermediate approach that may deserve attention.

Panels vary a lot with regard to the proportion of their two dimensions. Microeconometric panels may have small time dimension and an enormous cross-section dimension. By contrast, panels of collected time series, as they are typical for macroeconometrics, have a fixed number of cross-section units, usually not more than 10 or 20, and a time-series dimension that increases gradually as new information arrives. This distinction affects the focus of asymptotic analysis for panel tests. While much of the literature considers extending the two dimensions jointly to infinity, in varying proportions, sometimes achieving asymptotic normal laws, this approach hardly appears justifiable for panels of time series. For this reason, we view the cross-section dimension as fixed and study the small-sample behavior in this respect by means of simulation.

As a role-model case for seasonal time-series panels, we study a monthly panel of overnight stays in nine Austrian provinces. Due to long-run changes in tourists' preferences, cheaper airfares that permit spending summer vacations in more distant destinations, a stronger focus on sports and on sightseeing in cities, the increased female labor force participation that rules out long-term family stays with children in rural Austrian regions, and various other reasons, this is a data set where the distribution over the annual cycle is not constant and is unlikely to revert to any original pattern. For this reason, one may presume that it is well described by a model with seasonal unit roots. We focus on questions such as which unit roots are found at which frequencies and how these unit-root events are modified by the application of different tests.

From an estimated structure of autoregressions fitted to the nine individual series, connected by a covariance matrix of their residuals, we simulate pseudo-data to study the local power of the panel unit-root tests. We generally find that power improves at higher frequencies and that the nonparamatric tests face considerable problems. We also find that size bias is a persistent feature of panel unit-root tests, even after using the correction suggested by Pesaran (2007) that is taken up in the CHEGY test by Otero *et al.* (2007).

The plan of this paper is as follows. Section 2 expounds the univariate and panel tests that we wish to consider. Section 3 considers an empirical application to Austrian tourism data. Section 4 reports some parametric bootstrap simulation to study the size and power properties of the tests under a design that closely corresponds to the empirical data set. Section 5 draws some tentative conclusions and summarizes our results.

2 The testing procedures

2.1 The testing problem

Consider a panel of N real-valued time-series variables that are available at a monthly frequency for $t = 1, \ldots, T$. Denote the typical element as X_{jt} for $j = 1, \ldots, N$ and $t = 1, \ldots, T$. Note that we use j rather than the conventional i for the individual index, to keep it apart from $i = \sqrt{-1}$. The testing problem is to determine whether the autoregressive operators Φ_j in

$$\Phi_j(B)X_{jt} = \varepsilon_{jt},$$

with notation $\Phi_j(z) = \sum_{k=0}^{p_j} \phi_{k,j} z^k$ and B denoting the lag operator, contain roots at the locations $\exp(ik\pi/6)$ for $k = 0, \ldots 6$. Such autoregressive representations of order p_j are assumed to exist in the sense that error processes ε_{jt} are white noise for all j. Lag orders and shapes of the polynomials may vary across the cross-section dimension. We also allow for a non-diagonal covariance matrix $\Sigma = \mathrm{E}\,(\varepsilon_t \varepsilon_t')$, where $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$. For the applications of the panel versions of the tests, we however assume that unit-root events are homogeneous across the cross-section dimension in the sense that existence of a unit root at a frequency $\exp(ik\pi/6)$ in any Φ_j implies the existence for all j. Additionally, the general construction of our test statistics imposes the constraint $p_j \geq 12$.

Thus, for the N variables $X_{1t}, \ldots, X_{Nt}$ some common structure holds across the cross-section index j. A customary assumption is that unit-root events are identical across j, while other time-series characteristics, including the lag order p_j, may vary. Whether the incidence of a specific unit root in some members of the panel and absence in others should be classified into the alternative of panel unit-root tests, is a debated issue in the literature. The answer to this question decides whether test power in the case of an absent unit root for some j only is viewed as beneficial. We generally assume that unit-root events are identical across j, with the argument that otherwise the problem can be handled more appropriately by univariate tests.

2.2 The monthly HEGY test

Consider a real-valued time-series variable X that is available at a monthly frequency for $t = 1, \ldots, T$. The testing problem is to determine whether the autoregressive operator Φ in

$$\Phi(B)X_t = \varepsilon_t \tag{1}$$

contains roots at the locations $\exp(ik\pi/6)$ for $k = 0, \ldots 6$. Hylleberg *et al.* (1990, HEGY) drew attention to the fact that all of these roots appear in the polynomial Φ if, under some conditions, in the equivalent representation

$$\Delta_{12} X_t = \alpha' \left(X_{t-1}, \ldots, X_{t-12}\right)' + \gamma' \left(\Delta_{12} X_{t-1}, \ldots, \Delta_{12} X_{t-p}\right)' + \varepsilon_t, \tag{2}$$

the 12–vector α is a zero vector. We use the symbol Δ_{12} to denote the seasonal differencing operator $1 - B^{12}$. The condition $\alpha = 0$ can be checked by a corresponding least-squares regression and by considering the F–statistic on α, which, under the null hypothesis $\alpha = 0$, has an asymptotic non-standard distribution that differs from the usual F–distribution with 12 numerator degrees of freedom.

The representation can be made more informative by applying an additional transformation to the vector of level lags $X_t^- = (X_{t-1}, \ldots, X_{t-12})'$. In order to define that transformation matrix, consider the 12–vectors c_k, $k = 0, \ldots, 6$, filled by $\cos(lk\pi/6)$ for $l = 1, \ldots, 12$. Then, consider the 12–vectors d_k, $k = 1, \ldots, 5$, filled by $\sin(lk\pi/6)$ for $l = 1, \ldots, 12$. These are used to define a transformation matrix $\mathbf{M}$ by

$$\mathbf{M} = (c_0, c_1, d_1, c_2, d_2, \ldots, d_5, c_6)\,,$$

in this order. The matrix is nonsingular, and $(Y_t^-)' = (X_t^-)'\mathbf{M}$ can serve as an alternative regressor to define the specification

$$\Delta_{12} X_t = \beta' Y_t^- + \gamma' \left(\Delta_{12} X_{t-1}, \ldots, \Delta_{12} X_{t-p}\right)' + \varepsilon_t. \tag{3}$$

Note that γ is identical to (2) and ε_t is identical to (1) and to (2). In the form (3), the entries β_k in $\beta = (\beta_1, \ldots, \beta_{12})'$ correspond to unit-root events as follows. If $\beta_1 = 0$, there is a unit root at $+1$, and if $\beta_{12} = 0$, there is a unit root at -1. If $\beta_{2k} = \beta_{2k+1} = 0$ for any $k = 1, \ldots, 5$, there is a unit root at $\exp(ik\pi/6)$. The particular unit-root events can then be checked empirically by t– and F–statistics, again with non-standard asymptotic laws. The t– and F–statistics for the tests on the unit roots at $\exp(ik\pi/6)$ will be denoted by $t_0, F_1, \ldots, F_5, t_6$ in the following.

In general, we focus on null hypotheses of unit roots and on stationary alternatives, in most cases represented by $\beta_k < 0$ for $k = 1$ and k even, and thus we view the HEGY test as one-sided. For a digression from that principle to allow for unstable alternatives, see Section 4. For a summary on all properties of the HEGY statistics, see Ghysels and Osborn (2001).

2.3 The monthly CHEGY test

The CHEGY or cross-sectionally augmented HEGY test was introduced by Otero *et al.* (2007) who take up an idea developed by Pesaran (2007) and apply it to the problem of testing for seasonal unit roots in quarterly data. They call the test CHEGY-IPS, as it is historically related to the IPS test by Im *et al.* (2003). It avoids the calculation of correction factors for means and variances of the original IPS test, however, and hence we will refer to it simply as CHEGY test.

Panel unit-root tests are known to be sensitive to cross-section heterogeneity. In order to tackle this feature, various methods and corrections to existing methods have been suggested in the literature. The simple idea of the CHEGY test is to add cross-section averages of the Y^- variables as well as of the seasonal differences as additional regressors to the basic HEGY regression. Denote these variables as $\bar{Y}^-$ and as $\Delta_{12}\bar{X}$, respectively. Then, the HEGY regression for the monthly case reads

$$\begin{aligned} \Delta_{12}X_{jt} &= \beta' Y_{j,t}^- + \gamma' \left(\Delta_{12}X_{j,t-1}, \ldots, \Delta_{12}X_{j,t-p_j}\right)' \\ &\quad + \bar{\beta}' \bar{Y}_t^- + \bar{\gamma}' \left(\Delta_{12}\bar{X}_t, \ldots, \Delta_{12}\bar{X}_{t-p_j}\right)' \\ &\quad + \varepsilon_{jt}, \end{aligned} \tag{4}$$

with the noteworthy restriction that the lag order with regard to the averages is identical to the one for the individual regressors $\Delta_{12}X$. Note the simultaneous regressor $\Delta_{12}\bar{X}_t$.

The parts $\bar{\beta}$ and $\bar{\gamma}$ are designed to capture the effect of common factor structures. For small N, the correlation of a single X_{jt} and $\bar{X}_t$ may be sizeable. However, Otero *et al.* (2007) show that test performance is satisfactory in general.

The CHEGY statistics are then defined as arithmetic averages of individual t– and F–statistics for the β_k elements over the N individual values. Otero *et al.* (2007) provide simulated significance points for the quarterly CHEGY test at various values of N and T. We report some further quantiles tailored to our empirical application with $N = 9$ and $T = 406$ as Table 1, where we also add the cases $N = 5$ and $N = 18$ for a comparison. These distributional characteristics are based on the simulation design of a seasonal random walk $\Delta_{12}X_{jt} = \varepsilon_{jt}$ with $\varepsilon_{jt} \sim N(0, 1)$ independent across j and on 10,000 replications. The statistics are calculated for deterministic regressors containing monthly dummy constants and a linear time trend, and lag orders p_j are determined via BIC. Due to these specifications, our quantiles differ slightly from the ones given by Otero *et al.* (2007).

2.4 The monthly RURS test

Various nonparametric tests for unit roots have been considered in the literature (see Granger and Hallman, 1991, So and Shin, 2001, Aparicio *et al.*, 2006, Choi and Moh, 2007). By the observation that seasonal unit root tests are essentially unit-root tests on transforms of the original series, most of these tests can be generalized to the seasonal unit root problem. The RURS test by Kunst (2009) follows the 'record unit root' RUR test by Aparicio *et al.*

Table 1: Empirical distribution of the CHEGY statistic for $T = 406$ and $N = 5, 9, 18$.

	mean	0.05	0.10	0.50	0.90	0.95
$N = 5$						
t_0	-2.388	-3.063	-2.908	-2.394	-1.853	-1.699
F_1	4.117	2.219	2.569	4.022	5.780	6.327
F_2	4.117	2.208	2.577	4.039	5.745	6.265
F_3	4.103	2.199	2.549	4.040	5.725	6.274
F_4	4.096	2.231	2.570	4.015	5.707	6.244
F_5	4.085	2.190	2.551	3.995	5.760	6.276
t_6	-1.820	-2.567	-2.409	-1.827	-1.216	-1.039
$N = 9$						
t_0	-2.386	-2.895	-2.786	-2.385	-1.998	-1.892
F_1	4.122	2.666	2.934	4.076	5.376	5.789
F_2	4.127	2.647	2.923	4.080	5.387	5.783
F_3	4.110	2.605	2.908	4.060	5.362	5.751
F_4	4.118	2.623	2.924	4.071	5.384	5.783
F_5	4.112	2.618	2.899	4.064	5.369	5.762
t_6	-1.816	-2.370	-2.254	-1.820	-1.366	-1.242
$N = 18$						
t_0	-2.389	-2.750	-2.671	-2.390	-2.103	-2.024
F_1	4.098	2.927	3.168	4.078	5.055	5.356
F_1	4.118	2.953	3.187	4.087	5.094	5.380
F_2	4.114	2.932	3.170	4.082	5.088	5.391
F_4	4.115	2.972	3.196	4.084	5.082	5.366
F_5	4.112	2.951	3.177	4.080	5.090	5.412
t_6	-1.815	-2.247	-2.156	-1.818	-1.471	-1.373

Note: Columns correspond to quantiles of the empirical distribution generated using 10,000 replications of seasonal random walks with $N(0, 1)$ errors.

(2006, AES). The main idea of the AES test is to count the occasions of new records, maxima and minima, in a series, starting from the beginning of the series and also from the end. Ultimately, the forward and backward counts are averaged. If the series under investigation is a random walk, the statistic $T^{-1/2}R$, where R is the averaged number of extremum counts and T is the sample size, converges to a nonstandard distribution as $T \to \infty$. If the series is stationary, the statistic converges to 0.

In the construction of the RURS test, Kunst (2009) suggests to overcome the problem that the asymptotic distribution is known only for the case of a pure random walk by conditioning out lagged differences, in the tradition of the parametric Dickey-Fuller test. In order to preserve reasonable power properties for the test, the number of lags for this conditioning step is determined via the BIC criterion within a restrictive maximum order proportional to $T^{1/4}$.

In detail, the transforms Y^- are pure unit-root processes if the original process X is a seasonal random walk, in the sense that a single operator of the form $1 + aB + bB^2$ suffices to transform them to white noise. For the frequencies 0 and π, these operators are the simple first-order operators $1 - B$ and $1 + B$, respectively. To cope with the potential autocorrelation in a seasonally integrated process rather than a seasonal random walk, each of the transforms is regressed on its lags, which yields residuals $\hat{u}_t^{(k)}$. Then, new series are defined by accumulating the residuals, in symbols $Z_t^{(k)} = \sum_{s=1}^{t} \hat{u}_s^{(k)}$. If the original process is seasonally integrated, the processes $Z^{(k)}$, $k = 1, \ldots, 12$, will approximate random walks. If it is not integrated at the respective frequency, they will be stationary.

In a second step, the statistic $J_k = T^{-1/2}R^{(k)}$ is calculated, with T denoting the sample size and $R^{(k)}$ denoting the number of new maxima or minima in the series $Z^{(k)}$, where k may vary from 1 to 12 for the monthly case. This step is repeated for the series backward, with the ensuing statistics denoted by J'_k. For all frequencies $k \neq 1, 12$, this yields two different statistics with generally very similar values. In the following, we use the simple average over these two statistics.

In a third step, the forward and backward statistics are averaged according to

$$J_{*k} = 2^{-1/2}(J_k + J'_k), \tag{5}$$

with the weight $2^{-1/2}$ following the tradition of AES.

The test rejects whenever the count of new records is less than the 5% quantile of the typical distribution for a random walk, that is, the test is one-sided and rejects the unit-root null if too few new records are found. In this regard, the RURS test by Kunst (2009) digresses from the RUR test by AES, which is a two-sided test that rejects whenever there are too few record counts—indicative of a stationary process—or too many—indicative of a process with a deterministic time trend. Of course, too many records may even be found in the RURS version, which may be taken as indicating a superlinear trend in the original data.

In standard situations, the power of such tests is necessarily lower than that of comparable parametric tests, such as the HEGY test. Their virtue is that they are resilient to many deviations from the standard design, such as local outliers, structural breaks, and particularly some nonlinear models whose properties can be analyzed within the linear I(0)/I(1) framework (see Granger and Hallman, 1991, for this latter point).

Some simulated quantiles for the RURS test are provided in the upper panel of Table 2.

2.5 The panel RURS test

The RURS test is robust against seasonal deterministic cycles, and it is also invariant to heterogeneity across j or to non-diagonal Σ. As long as dependence in the cross-section dimension does not invalidate laws of large numbers, an average of the N RURS statistics will, under the unit-root null at the considered frequency, converge to the first moment of the RURS null distribution as $N \to \infty$. Of course, for small N it makes sense to study the null distribution of

$$\bar{J}_{*k} = N^{-1} \sum_{j=1}^{N} J_{*k}^{(j)},$$

where $J_{*k}^{(j)}$ denotes the RURS statistic at frequency k for individual series j. In simulations and bootstrap experiments tuned to actual data, we found that the distribution quickly condenses around its mode of around 2.3 as N gets larger. In other words, a value of less than 2 becomes conspicuous for $N = 10$ in the average, while it would not be significant for a single series.

The left-sided test based on the average RURS statistic $\bar{J}_{*k}$ will be called the RURS-p test in the following. Some simulated quantiles for the RURS-p test are provided in the lower panels of Table 2. Note that empirical means are identical but that the distribution is much more concentrated than for the univariate RURS statistics. This concentration becomes sharper as N increases.

Table 2: Quantiles for the RURS statistics based on $T = 406$ and for the RURS-p statistics based on $T = 406$ and $N = 5, 9, 18$. Monte Carlo results from 10,000 replications of seasonal random walks with $N(0, 1)$ errors.

	mean	0.05	0.1	0.5	0.9	0.95
RURS						
0	2.369	1.613	1.757	2.330	3.047	3.263
$\pi/6$	2.329	1.577	1.721	2.295	3.012	3.227
$\pi/3$	2.384	1.649	1.757	2.330	3.047	3.263
$\pi/2$	2.313	1.784	1.887	2.301	2.766	2.922
$2\pi/3$	2.384	1.649	1.757	2.330	3.047	3.298
$5\pi/6$	2.330	1.577	1.721	2.295	3.012	3.227
π	2.371	1.613	1.757	2.330	3.047	3.263
RURS-p	$N = 5$					
0	2.375	2.022	2.094	2.366	2.667	2.753
$\pi/6$	2.335	1.979	2.051	2.330	2.632	2.718
$\pi/3$	2.389	2.036	2.108	2.381	2.689	2.775
$\pi/2$	2.314	2.058	2.115	2.311	2.513	2.580
$2\pi/3$	2.390	2.029	2.108	2.381	2.682	2.768
$5\pi/6$	2.335	1.972	2.051	2.323	2.632	2.725
π	2.366	2.008	2.079	2.359	2.660	2.753
RURS-p	$N = 9$					
0	2.369	2.095	2.155	2.366	2.589	2.653
$\pi/6$	2.329	2.052	2.111	2.326	2.549	2.613
$\pi/3$	2.384	2.111	2.167	2.382	2.601	2.665
$\pi/2$	2.313	2.129	2.166	2.313	2.462	2.508
$2\pi/3$	2.384	2.107	2.167	2.378	2.605	2.669
$5\pi/6$	2.330	2.052	2.115	2.326	2.549	2.617
π	2.371	2.103	2.159	2.370	2.585	2.653
RURS-p	$N = 18$					
0	2.368	2.179	2.219	2.366	2.520	2.567
$\pi/6$	2.328	2.137	2.177	2.324	2.482	2.532
$\pi/3$	2.382	2.195	2.233	2.380	2.536	2.581
$\pi/2$	2.313	2.179	2.206	2.311	2.419	2.448
$2\pi/3$	2.382	2.193	2.233	2.380	2.534	2.579
$5\pi/6$	2.328	2.137	2.177	2.326	2.484	2.530
π	2.369	2.173	2.217	2.366	2.524	2.569

2.6 Deterministic terms

Deterministic terms can affect seasonal time-series models in two respects. Firstly, intercepts and trends in the autoregressive form generate linear and quadratic trends in the time series if there is a unit root at +1, like in non-seasonal unit-root models. Second, however, seasonal dummy constants generate ever-increasing seasonal cycles if seasonal unit roots are present. In particular, 'spectral' transforms of seasonal dummy constants using $\mathbf{M}$ imply expanding seasonal cycles at specific frequencies. This feature has led to the suggestion of restricting these constants per frequency in multivariate seasonal models (see FRANSES AND KUNST, 1999).

Within this paper, we add deterministic terms to the basic autoregressions as follows. In all HEGY-type tests, a linear time trend and 12 seasonal dummy constants are used as additional regressors. In the nonparametric tests, linear time trends and seasonal dummy constants are inessential, as they drop out in the test construction due to the augmenting step, where a constant is used in the regression. In all artificially generated data, coefficients on dummy constants are restricted to zero if seasonal unit roots are present and trend coefficients are restricted to zero if a unit root at +1 is present. This avoids the generation of systematically unstable or quadratically trending trajectories that we do not view as realistic.

3 An empirical example

3.1 The data

Data are from the Austrian Wifo data base. They are monthly and cover the time range January 1970 to October 2008. Variables are the registered overnight stays in the nine Austrian 'länder' (provinces).

Quantities are quite heterogeneous, as are the sizes and populations of the nine länder. Burgenland and Vorarlberg are the smallest Austrian provinces, and Lower Austria and Vienna are the largest ones, according to their populations. A similar heterogeneity applies to the nature of tourism across the regions. The capital of Vienna and the city of Salzburg, which is the capital of the province bearing the same name, attract visitors all around the year, because of cultural activities. The mountainous regions of Tyrol, Vorarlberg, and to a lesser degree also Carinthia, Salzburg, and Styria, attract winter tourism due to their skiing facilities. On the other hand, regions like Burgenland and much of the lowlands of Lower Austria offer no skiing facilities. Traditionally, Upper and Lower Austria served as summer resorts for longer-term stays of German and Viennese inhabitants, the so-called 'Sommerfrische'. This form of vacation is in a long-run decline. Carinthia and Burgenland tend to concentrate on their lakes and on summer sports, such as swimming and water skiing. Generally, the picture is quite colorful and heterogeneous, but it is dominated by a long-term tendency toward more intensive winter tourism and all-year city tourism and, sadly enough, away from traditional summer vacation.

Figures 1 and 2 demonstrate these developments. Particularly in Salzburg, Tyrol, and Vorarlberg has winter tourism overtaken summer tourism. With the exception of the main tourist cities Vienna and Salzburg, it appears that Austria on the whole is on its way to a winter destination.

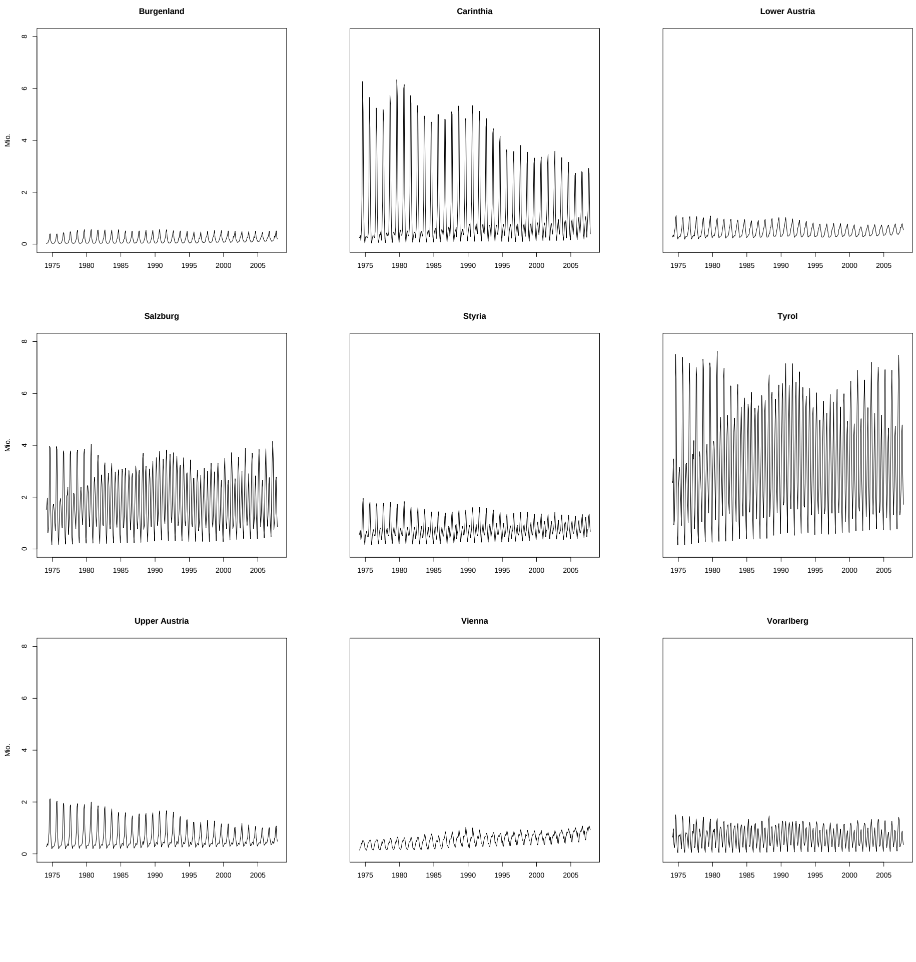

Figure 1: Overnight stays in the nine Austrian provinces, on a common scale.

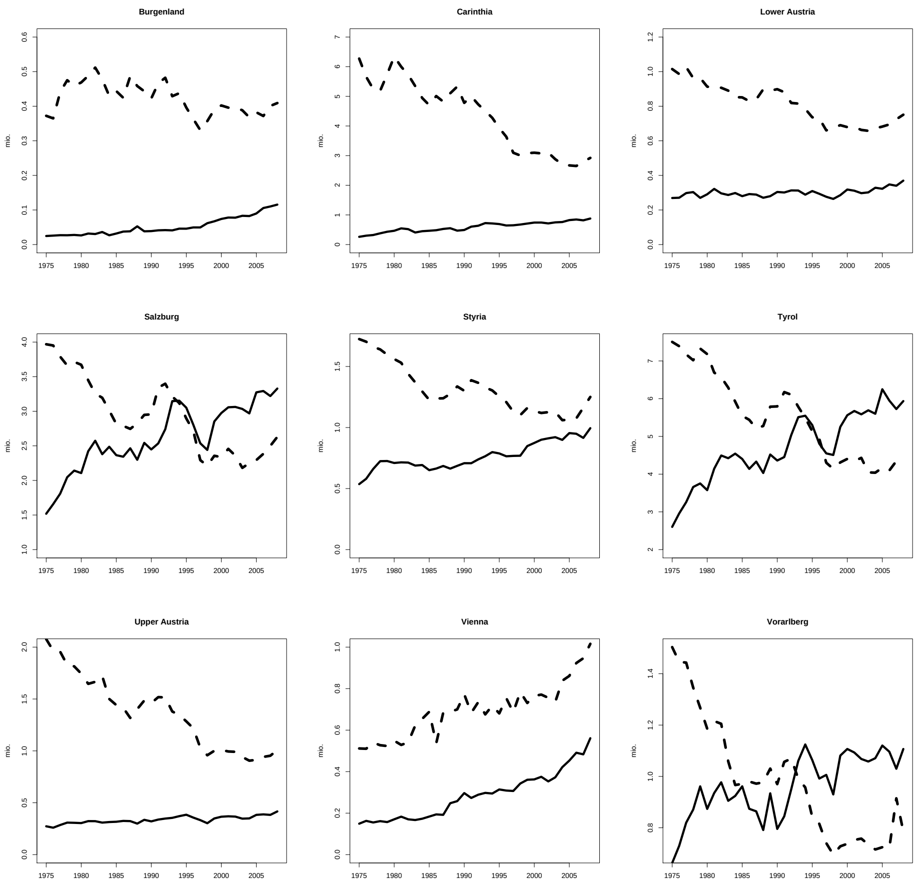

Figure 2: Overnight stays in January (solid) and in July (dashed) in the nine Austrian provinces, individual scales.

Table 3: HEGY statistics for individual series in the panel and CHEGY statistics for the panel.

Region	p	t_0	F_1	F_2	F_3	F_4	F_5	t_6
Burgenland	2.00	-1.64	0.54	6.37	15.50*	23.08*	31.62*	-6.85*
Carinthia	2.00	-1.97	1.15	5.08	7.00*	13.26*	37.78*	-7.01*
L. Austria	3.00	-1.51	0.16	7.55*	10.91*	29.74*	41.24*	-5.33*
Salzburg	10.00	-3.04	3.27	3.90	5.52	12.16*	33.15*	-5.63*
Styria	2.00	-1.13	1.13	2.86	7.69*	11.74*	31.24*	-5.96*
Tyrol	1.00	-2.53	2.37	6.18	7.02*	18.42*	38.55*	-6.66*
U. Austria	2.00	-1.52	0.90	1.52	12.05*	13.10*	28.90*	-6.88*
Vienna	2.00	-2.31	0.76	3.29	6.93*	7.12*	17.63*	-5.14*
Vorarlberg	1.00	-2.21	2.06	6.82	6.86*	21.37*	34.59*	-6.63*
CHEGY		-2.40	5.78	5.95*	7.87*	9.93*	8.21*	-2.09

Note: Asterisks denote significance at 5%.

Whether the peak is in summer or in winter, seasonality is strong for all provinces. Modelling the seasonal structure appears to be crucial for the task of modelling the time series.

3.2 Univariate tests

Table 3 shows the HEGY statistics for the series on the nine provinces, after applying a logarithmic transformation. For most series, a BIC search found a small lag order p, the exception is Salzburg. Simulations comparable to the ones reported elsewhere in the paper for $T = 406$ give 5% points of -3.43 and -2.92 for the two t–tests and of around 6.8 for the F–tests, in good correspondence to values reported in the literature (see, for example, Beaulieu and Miron, 1993).

For none or almost none of the series does the HEGY test reject the unit-root null at the frequencies $0,\pi/6$, and $\pi/3$, i.e. at the long-run frequency and at the annual and semi-annual seasonal cycles. In short, seasonality appears to be unit-root stochastic at the longer frequencies. By contrast, the HEGY tests do reject almost unanimously at the short frequencies π, $5\pi/6$, $2\pi/3$, and $\pi/2$. This means that seasonal cycles at these frequencies are deterministic or that the contribution from these frequency components is small.

The latter conjecture is confirmed by spectral estimates that are presented in Figure 3. Typically, all seasonal frequencies appear as peaks but the frequencies $\pi/6$ and $\pi/3$ tend to dominate. Thus, the traditional parametric HEGY test finds those unit roots that correspond to the most dominant components of the seasonal cycle. A different view of the higher seasonal frequencies is that these represent month-to-month patterns, such as November to December, which are relatively stable over time, while the importance of January relative to August, for example, experiences sizeable long-run movements.

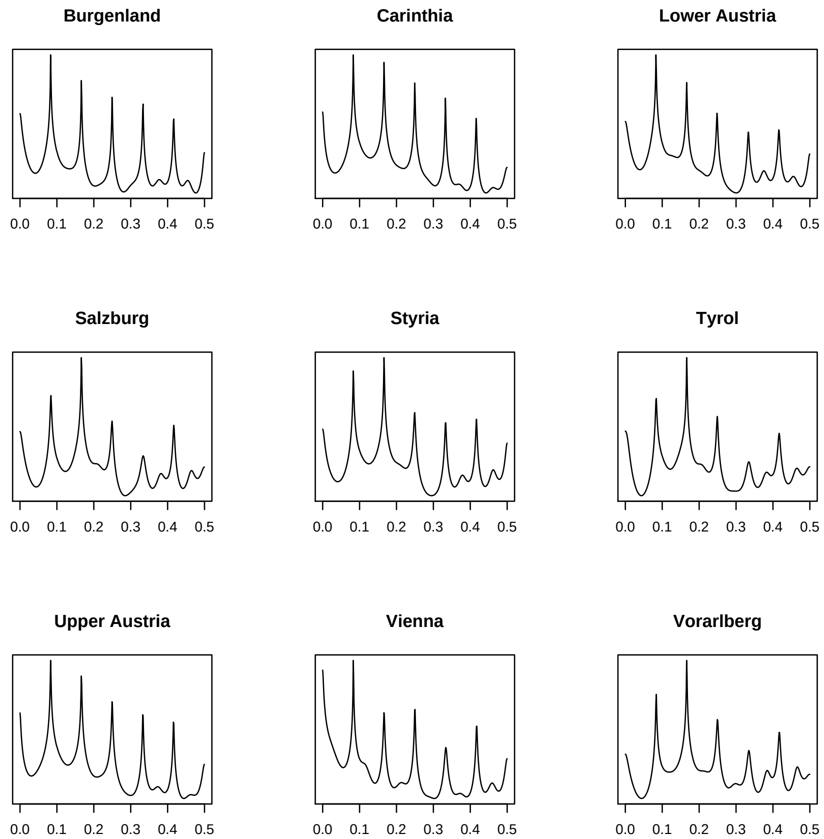

Figure 3: Estimated spectral densities for the series from all Austrian provinces, individual scales.

3.3 The CHEGY panel test

We implemented the CHEGY panel test that was outlined in Section 2.2. The lag order p is again determined by BIC. Because of the averages as additional regressors, the penalty for larger p increases and hence lag orders are generally smaller than in the univariate HEGY tests. This leads to increased F statistics for some provinces and, consequently, to CHEGY statistics that are generally above the naive averages of the individual HEGY statistics. For this reason, the CHEGY panel test rejects at 5% for the frequencies $\pi/3$ to $5\pi/6$, while it still supports the unit-root null for the 0 and π frequencies as well as for $\pi/6$. In detail, the F–statistics are provided in the bottom row of Table 3, to be checked against the 5% significance point provided in Table 1, which renders all CHEGY statistics significant at 5% except at 0, $\pi/6$, and π.

In summary, the panel test indicates unit-root seasonality at the Nyqvist and annual frequencies and at the zero frequency—indeed, this is the non-seasonal traditional unit root—while it rejects unit roots at all intermediate seasonal frequencies. While these results appear to be in conflict with univariate tests, note that the rejection at $\pi/3$ is fragile and that the p–values tend to decrease as the frequency shortens. On the other hand, they re-increase for $5\pi/6$ and particularly at π, where the univariate test rejects but the CHEGY test does not. The reason are small, and even positively signed, contributions by Carinthia and Styria. In these two cases, the additional covariates constructed from cross-section averages eliminate the correlation of increments and alternating cumulated sums of the levels that is used as evidence against a unit root at π in the HEGY test.

3.4 The nonparametric tests

Table 4 gives the RURS statistics at all frequencies for the nine Austrian provinces. Quantiles given in the upper panel of Table 2 are around 1.65, which implies that the null of a seasonal unit root is rejected at the frequencies $\pi/2$ and π for most cases, otherwise the data apparently support the null.

This conclusion is confirmed by the average RURS-p statistics. The corresponding panel of Table 2 for $N = 9$ shows a simulated 5% point of 2.1 at all frequencies, thus seasonal unit roots at $\pi/2$ and π are rejected for the panel. At the remaining frequencies, *more* extrema are found than would be typical for unit-root processes, which indicates an expansion of seasonal cycles beyond the random-walk rate. Particularly at the zero frequency, a growth rate of the series beyond the constant expansion rate covered by the linear trend—which is eliminated in the construction of the test statistic—is reflected in the high value of 3.25.

Table 4: RURS statistics for all countries.

	0	$\pi/6$	$\pi/3$	$\pi/2$	$2\pi/3$	$5\pi/6$	π
Burgenland	3.585	2.976	2.868	1.474	3.621	2.976	1.291
Carinthia	3.119	2.151	2.868	1.887	3.083	2.868	2.079
L. Austria	3.729	3.012	2.581	1.577	3.513	2.581	1.183
Salzburg	3.513	3.119	3.191	1.474	3.693	3.191	1.291
Styria	3.191	2.940	2.223	1.603	3.298	2.223	1.434
Tyrol	3.693	3.513	3.693	1.525	3.693	3.693	1.542
U. Austria	3.047	2.402	2.581	1.784	3.083	2.581	1.398
Vienna	2.653	2.008	2.474	1.810	2.474	2.330	1.613
Vorarlberg	2.761	2.617	2.617	1.577	2.617	2.617	1.398
RURS-p	3.255	2.749	2.788	1.635	3.231	2.784	1.470

4 Post-sample evidence on size and power

4.1 The simulation design

The local power close to the null of seasonal unit roots will be studied on the basis of a simulation design that is adapted to the tourism data set that we investigated in the last section. Unrestricted estimates serve as the alternative design, and a restriction of this structure serves as the null design. Weighted averages of these two benchmarks, in other words on an arc between the two models, will also be evaluated.

In detail, autoregressive models of the form

$$\begin{aligned}\Delta_{12}X_t &= \alpha'\left(X_{t-1},\ldots,X_{t-12}\right)' + \gamma'\left(\Delta_{12}X_{t-1},\ldots,\Delta_{12}X_{t-p}\right)' \\ &\quad + \sum_{j=1}^{12}\delta_j D_{jt} + \kappa t + \varepsilon_t \qquad (6)\end{aligned}$$

are fitted to the nine observed series with the lag order p selected by BIC, and the covariance matrix of the individual error terms is estimated by maximum likelihood. These parameters $\hat{\Sigma}$, $\hat{\alpha}$, $\hat{\gamma}$, $\hat{\delta}$, and $\hat{\kappa}$ are then used to simulate the model, with errors drawn from a nine-variate $N(0,\hat{\Sigma})$ distribution. 10,000 replications of this parametric bootstrap alternative design are generated and all statistics are recorded.

Starting from this 'realistic' alternative design, we then shrink the crucial parameters in α to 0 by multiplying them by a constant $\tau \in$

$[0,1]$, in symbols we simulate the following model:

$$\Delta_{12}X_t = \tau\alpha' \left(X_{t-1}, \ldots, X_{t-12}\right)' + \gamma' \left(\Delta_{12}X_{t-1}, \ldots, \Delta_{12}X_{t-p}\right)' \\ + \tau \sum_{j=1}^{12} \delta_j D_{jt} + \tau\kappa t + \varepsilon_t. \quad (7)$$

Thus, the value $\tau = 0$ defines a null model with unit roots, and rejection frequencies of tests for $\tau = 0$ are size values rather than power. Note that the deterministic terms are also shrunk toward zero, in order to avoid deterministically unstable models under the null with quadratic trends and seasonally divergent trends. It is conceivable to admit a non-zero constant under the null or to shrink coefficients at different rates but we prefer to keep the design simple.

Note that the implied null design for $\tau = 0$ does not correspond to the maximum-likelihood estimate under the null. Nevertheless, this simulation design appears to be more informative on the power of the tests than a constructed artificial design.

In fact, for the given data the fitted structure is unstable for the case of Burgenland with regard to its stochastic part, in the sense that the α part generates unstable modes. The modulus of the unstable roots only slightly exceeds unity, though, such that trajectories do not digress visually from the observed data. It is known that stochastic instability leads to a shift of the distribution of HEGY–type statistics to their 'other' tail and thus to an increased tendency to accept the null in a one-sided test.

4.2 The parametric tests

It is under these caveats that Table 5 is to be interpreted. The test is considerably oversized except at the Nyqvist frequency π. The best power occurs at the highest frequencies, as is to be expected, as these admit the best information in the sample.

Table 5 demonstrates the local power of the CHEGY procedure according to the experimental design expounded above. The simulation design corresponds to a point in the—generalized, because partially unstable— alternative of the test, and the column headed $\tau = 1$ shows that the rejection rate is 100% at all frequencies higher than $\pi/3$, while the test does not reject at all at the annual frequency $\pi/6$ and achieves around 50% rejection at the zero frequency.

Test performance may be unduly affected by the fact that the generating model is unstable for pseudo-Burgenland at a frequency close to the annual cycle. For this reason, we repeat the experiment with replacing the unstable mode at $|\zeta| > 1$ by a stable mode at ζ^{-1}

Table 5: Rejection frequency of CHEGY test around the null.

τ	0.00	0.20	0.40	0.60	0.80	1.00
unstable unrestricted estimate						
t_0	0.166	0.063	0.023	0.069	0.221	0.486
F_1	0.164	0.020	0.005	0.001	0.000	0.000
F_2	0.153	0.050	0.196	0.450	0.716	0.908
F_3	0.178	0.315	0.774	0.987	1.000	1.000
F_4	0.169	0.645	0.999	1.000	1.000	1.000
F_5	0.164	0.994	1.000	1.000	1.000	1.000
t_6	0.069	0.979	1.000	1.000	1.000	1.000
stabilized design						
t_0	0.166	0.062	0.025	0.074	0.231	0.500
F_1	0.164	0.036	0.048	0.074	0.108	0.156
F_2	0.153	0.050	0.201	0.466	0.735	0.917
F_3	0.178	0.316	0.776	0.986	1.000	1.000
F_4	0.169	0.646	0.999	1.000	1.000	1.000
F_5	0.164	0.994	1.000	1.000	1.000	1.000
t_6	0.069	0.979	1.000	1.000	1.000	1.000

and report the results in the lower panel of Table 5. The differences are small, excepting the frequency $\pi/6$. The original design adapted to the data is indeed unstable but the mode as well as its inverse are so close to one that power remains small even for the artificially stabilized design.

In detail, the size bias at $\tau = 0$ turns into a strongly dichotomous behavior even at $\tau = 0.2$: good power at the high frequencies, a marked fall in rejection rates at lower frequencies. Strong support for the null prevails for all τ at and around $\pi/6$. The main explanation is that the design at $\tau = 1$ is not really in the expected alternative. The stochastic instability occurs at approximately $\pi/6$, in other words an unstable annual cycle. That instability apparently aliases into the neighboring frequencies at 0 and at $\pi/3$.

The power of the panel CHEGY test is to be compared to the results of a standard HEGY test procedure applied to the individual states. Table 6 gives such a comparison, where rejection frequencies are averaged across states. Thus, the distinction among the individuals is lost. In fact, we found considerable heterogeneity, in the sense that some cases have a rejection frequency of 1, while others have low frequency, at least for $\tau = 1$. Note that the HEGY test has approximately correct size at $\tau = 0$. It is immune to the size distortions of

Table 6: Average rejection frequency of HEGY test around the null.

τ	0.00	0.20	0.40	0.60	0.80	1.00
unstable unrestricted estimate						
t_0	0.049	0.236	0.562	0.549	0.362	0.228
F_1	0.056	0.040	0.043	0.057	0.155	0.194
F_2	0.060	0.069	0.120	0.181	0.268	0.377
F_3	0.061	0.131	0.282	0.477	0.675	0.838
F_4	0.065	0.273	0.633	0.858	0.947	0.980
F_5	0.058	0.558	0.964	0.998	1.000	1.000
t_6	0.043	0.403	0.883	0.984	0.999	1.000
stabilized design						
t_0	0.049	0.236	0.562	0.549	0.363	0.229
F_1	0.056	0.042	0.044	0.054	0.070	0.088
F_2	0.060	0.069	0.121	0.181	0.269	0.378
F_3	0.061	0.131	0.282	0.477	0.675	0.838
F_4	0.065	0.273	0.633	0.858	0.947	0.980
F_5	0.058	0.558	0.964	0.998	1.000	1.000
t_6	0.043	0.403	0.883	0.984	0.999	1.000

the panel CHEGY test. Its power increases more slowly at the high frequencies, while support for the unit root at the lower frequencies becomes more fragile.

Again, the general test performance is hardly affected at all by the fact that one of the countries, Burgenland, is unstable at the annual frequency, as can be seen from the lower panel of Table 6, which is roughly identical to the upper panel.

4.3 The nonparametric tests

In this subsection, we report the results of simulation experiments that investigate the power of the RURS and RURS-p tests. These experiments are analogous to those for the HEGY and CHEGY tests reported in the last subsection.

Tables 7 and 8 show that the performance of the nonparametric tests is generally disappointing. While the generating model is in the alternative for all $\tau \neq 0$, the procedure rejects for the frequencies π and $\pi/2$ only. The univariate test even fails to reject for these two frequencies for a substantial portion of the replications, whereas the HEGY-type tests find the missing roots with a probability close to unity. Even at π, where the power of the panel RURS-p test appears acceptable, is the CHEGY test faster in picking up the information than the RURS-p test. CHEGY rejection is close to unity at $\tau = 0.2$, whereas the RURS-p test needs $\tau = 0.4$ to attain a comparable discriminatory power.

Conversely, the nonparametric tests have a considerable size bias at $\tau = 0$, far beyond the problems encountered by the CHEGY test. This bias can be reduced, however, by liberalizing the lag-order search, substituting AIC for BIC, or increasing the upper bound for p.

Thus, the local-power simulation corroborates the findings for the sample at $\tau = 1$ but it helps to make it more precise. At a point in the parameter space, where a traditional parametric test tends to support the alternative for the high seasonal frequencies, will the nonparametric record-counting test be unable to reject the unit-root null. A first explanation is that the nonparametric tests, by construction, process less information than the parametric rivals and thus have less power. This explanation, however, only suffices for the behavior at the frequency π, where the difference in power is merely quantitative. At the intermediate frequencies, however, the difference is qualitative. The record-counting tests interpret the typical shape change in a role-model seasonal time series as being composed of a pattern-reverting deformation at backbone frequencies at π and $\pi/2$ and persistent unit-root cycles at the annual and various intermediate frequencies, where the HEGY-type tests are more prone to see a substantial amount of pattern reversion.

Both the univariate and the panel variant were also applied to the artificially stabilized simulation design that was introduced in the last subsection. These experiments are reported in the lower panels of Tables 7 and 8. There are only minor differences between the upper and lower panels. The local instability in the Burgenland series does not affect the overall performance of the nonparametric tests.

5 Discussion

We present a generalization of seasonal unit-root tests to monthly panels, and we illustrate the properties on an empirical data set on Austrian tourism data. The data set permits us to inspect the size and power properties of a parametric and a nonparametric test procedure in a realistic simulation design.

Whereas generally the discriminatory power of the CHEGY test appears acceptable, it is not immune to size bias due to heterogeneous dynamic effects in the component time series. The optimum determination of lag orders continues to be a problem, and it may well be that alternatives to our BIC search deserve attention.

Table 7: Average rejection frequency of RURS test around the null.

τ	0.00	0.20	0.40	0.60	0.80	1.00
unstable unrestricted estimate						
0	0.031	0.000	0.000	0.006	0.018	0.066
$\pi/6$	0.198	0.005	0.006	0.024	0.050	0.141
$\pi/3$	0.143	0.010	0.024	0.017	0.064	0.209
$\pi/2$	0.074	0.082	0.137	0.205	0.309	0.416
$2\pi/3$	0.215	0.006	0.001	0.021	0.042	0.128
$5\pi/6$	0.145	0.000	0.000	0.005	0.018	0.077
π	0.040	0.202	0.439	0.632	0.754	0.836
stabilized design						
0	0.031	0.000	0.000	0.006	0.018	0.044
$\pi/6$	0.198	0.005	0.006	0.024	0.050	0.122
$\pi/3$	0.143	0.010	0.024	0.017	0.064	0.177
$\pi/2$	0.074	0.081	0.128	0.200	0.295	0.405
$2\pi/3$	0.215	0.006	0.001	0.021	0.041	0.106
$5\pi/6$	0.145	0.000	0.000	0.005	0.018	0.060
π	0.040	0.204	0.445	0.635	0.762	0.843

Table 8: Rejection frequency of RURS-p test around the null.

τ	0.00	0.20	0.40	0.60	0.80	1.00
unstable unrestricted estimate						
0	0.049	0.000	0.000	0.000	0.000	0.001
$\pi/6$	0.478	0.000	0.000	0.000	0.000	0.034
$\pi/3$	0.296	0.000	0.000	0.000	0.000	0.096
$\pi/2$	0.222	0.282	0.486	0.705	0.890	0.974
$2\pi/3$	0.479	0.000	0.000	0.000	0.000	0.004
$5\pi/6$	0.347	0.000	0.000	0.000	0.000	0.002
π	0.134	0.893	0.997	1.000	1.000	1.000
stabilized design						
0	0.049	0.000	0.000	0.000	0.000	0.000
$\pi/6$	0.478	0.000	0.000	0.000	0.000	0.026
$\pi/3$	0.296	0.000	0.000	0.000	0.000	0.059
$\pi/2$	0.222	0.278	0.471	0.697	0.882	0.969
$2\pi/3$	0.479	0.000	0.000	0.000	0.000	0.003
$5\pi/6$	0.347	0.000	0.000	0.000	0.000	0.001
π	0.134	0.894	0.997	1.000	1.000	1.000

The discriminatory power of the nonparametric test is lower than that of the CHEGY test. However, it appears to offer an interesting alternative in the presence of instabilities. The HEGY-type tests treat such instabilities in a symmetric fashion to stationary deviations from the null. By contrast, the record-based test views them as evidence on trend expansions and finds considerably more records than would be typical under the null. Thus, the tests may arrive at different conclusions in realistic data sets, and it may be worth while to apply both concepts in order to get a more reliable impression of the nature of seasonal cycles in the data.

References

[1] Aparicio, F., Escribano, A., and A.E. Sipols, 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 J.A. Miron, 1993, Seasonal unit roots in aggregate U.S. data. Journal of Econometrics 55, 305–328.

[3] Breitung, J., 2002, Nonparametric tests for unit roots and cointegration. Journal of Econometrics 108, 343–363.

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

[5] Dreger, C., and H.E. Reimers, 2004, Panel Seasonal Unit Root Test with an Application for Unemployment Data. Discussion paper No. 191, IWH, Halle.

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

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

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

[9] Hlouskova, J., and M. Wagner, 2006, The Performance of Panel Unit Root and Stationarity Tests: Results from a Large Scale Simulation Study. Econometric Reviews 25 (1), 85–116.

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

[11] Im, K., Pesaran, M.H., and Shin, Y., 2003, Testing for unit roots in heterogeneous panels. Journal of Econometrics 115, 53–74.

[12] Kunst, R.M., 2009, A nonparametric test for seasonal unit roots. Economics Series No. 233, Institute for Advanced Studies.

[13] Otero, J., Smith, J., and M. Giulietti, 2005, Testing for Seasonal Unit Roots in Heterogeneous Panels. Economics Letters 86, 229–235.

[14] Otero, J., Smith, J., and M. Giulietti, 2007, Testing for Seasonal Unit Roots in Heterogeneous Panels in the Presence of Cross Section Dependence. Economics Letters 97, 179–84.

[15] Pesaran, M.H., 2007, A simple panel unit root test in the presence of cross-section dependence. Journal of Applied Econometrics 22, 265–312.

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

[17] Ucar, N., and H. Guler, 2007, Testing Stochastic Convergence Hypothesis in Seasonal Heterogeneous Panels. Working paper, PGlobal Advisory Services, Ankara.