CESifo *Working Paper Series*

DECISIONS ON SEASONAL UNIT ROOTS

Robert M. Kunst
Michael Reutter

Working Paper No. 286

April 2000

CESifo
Poschingerstr. 5
81679 Munich
Germany
Phone: +49 (89) 9224-1410/1425
Fax: +49 (89) 9224-1409
http://www.CESifo.de

CESifo *Working Paper No. 286*
April 2000

DECISIONS ON SEASONAL UNIT ROOTS

Abstract

Decisions on the presence of seasonal unit roots in economic time series are commonly taken on the basis of statistical hypothesis tests. Some of these tests have absence of unit roots as the null hypothesis, while others use unit roots as their null. Following a suggestion by Hylleberg (1995) to combine such tests in order to reach a clearer conclusion, we evaluate the merits of such test combinations on the basis of a Bayesian decision setup. We find that the potential gains over a pure application of the most common test due to Hylleberg et al. (1990) are small.

JEL Classification: C12, C15

Robert M. Kunst
Institute for Advanced Studies
Stumpergasse 56
1060 Vienna
Austria
e-mail: kunst@ihs.ac.at

Michael Reutter
CES
University of Munich
Schackstr. 4
Munich
Germany

1 Introduction

Many time series that are sampled at a quarterly or monthly frequency, in economics and in other disciplines, show remarkable seasonal variation. Recent statistical research has focused on the problem of discriminating models with seasonal unit roots from models with deterministic seasonal cycles. Such discrimination is relevant for prediction but it may also allow a more substantial interpretation. If seasonal unit roots are present, shocks to the seasonal cycle have a permanent effect. If deterministic cycles explain seasonal variation, all shocks to the seasonal cycle are transitory. This distinction may permit a detailed interpretation of the sources of seasonal variation. For example, tourism demand may shift permanently from summer to winter tourism if skiing facilities are developed in a mountainous region. Construction may increase permanently during the cold season if a new technology can be applied that copes with icy conditions. Conversely, all seasonality may be classified as being deterministic if the direct effect of the climatic cycle dominates.

For this discrimination problem, a variety of statistical testing procedures have been suggested in the literature. The test by Hylleberg *et al.* (1990, HEGY) was the first one to separately treat seasonality at different frequencies. We will exclusively focus on the quarterly case and we will abstain from testing for a non-seasonal unit root at frequency $\omega = 0$, hence the HEGY test may classify series by having a unit root pair $\pm i$ at the annual cycle $\omega = \pi/2$ or not, and simultaneously by having a unit root of -1 at $\omega = \pi$ or not. The HEGY test uses an autoregressive model frame and has the presence of unit roots as null hypotheses. The tests by Tam and Reinsel (1997), Canova and Hansen (1995, CH), and Caner (1998) use different model frames that allow an apparent interchange of null and alternative hypotheses. For example, the CH test sees the data generation process as the weighted sum of deterministic, stationary stochastic, and pure unit-root components. Only if the weight of the latter component is zero, the unit root will be rejected. In this paper, we focus on the CH test and on the Caner test, as the procedure by Tam and Reinsel does not treat the two seasonal frequencies separately.

The question whether the HEGY test or the tests by CH or Caner are to be preferred has no trivial answer. Hylleberg (1995) compares the performance of the HEGY and the CH procedures by Monte Carlo simulation and finds that either test dominates with respect to discriminatory power if the true data generation mechanism corresponds to the hypothesized model frame. This result gives no clear guidance for the case of conflicting results from the two procedures, as the true generation mechanism is unknown.

We focus on interpreting all testing problems in a parametric framework, although we are aware that particularly the CH test was presented in a semi-parametric one. Seen as parametric tests, we note that the null hypotheses of the HEGY test and the alternatives of the CH and Caner tests are not simply equivalent. Starting from different parametric structures, each test considers specific restrictions on their individual parameter spaces, which finally imply the persistence of shocks to seasonal cycles. We take this feature, i.e., whether or not seasonal shocks are persistent, to represent the main hypotheses of interest. The application of any of these tests is then interpreted as an indication that the researcher is also willing to accept the union of the suggested model classes as his/her working frame. We note that similar remarks apply to the more common problem of testing a unit root at 1, and that a fully semi-parametric interpretation of unit-root testing encounters subtle difficulties (see Abadir and Taylor, 1999).

In this paper we develop a decision-theoretic setup that aims at constructing a new test from considering a pair of available seasonal unit-root statistics jointly. The researcher is assumed to observe a data series that was generated from either the HEGY or the CH generation model, either with or without unit roots, with identical probability. The researcher wishes to classify the generation mechanisms as to whether seasonal shocks are persistent or not. We use Monte Carlo simulation in order to maximize the frequency of correct classification. In the field of testing for unit roots at frequency 0, joint usage of two tests with seemingly complementary hypotheses was also analyzed by Hatanaka (1996) and Hatanaka and Koto (1995) who recommend a comparison of p–values from both tests. We build on their suggestion but we extend their idea in order to explicitly quantify the test decision.

Although, strictly speaking, the results of our analysis are only valid conditional on our weighting priors within the assumed classes, we feel that at least our qualitative conclusions are of interest to the practitioner. Firstly, although every second trajectory is generated from the unobserved-components frame that was suggested by Canova and Hansen (1995), it appears that the test statistic suggested by HEGY allows a much more precise classification than both the CH and the Caner statistics if any of these is used on its own. Although the alternative tests may dominate on their home grounds, they fail to do so on a mixture of their generation models with the simple autoregressive model. Secondly, the information contained in the CH and in the CANER statistics is roughly equivalent, hence there is not much to choose between these two approaches. Thirdly, the situation at the two seasonal frequencies is very different. For $\omega = \pi/2$, our procedure suggests to rely solely on the HEGY statistic. For $\omega = \pi$, a combination of the HEGY statistic with one of the alternative statistics is likely to improve the task. This latter issue is in line with HYLLEBERG (1995).

The remainder of this paper is organized as follows. Section 2 reviews the test procedures and our decision setup. Section 3 reports and analyzes the results of our Monte Carlo simulation. Section 4 concludes.

2 Testing for seasonal unit roots

2.1 Three testing procedures

The most common test for seasonal unit roots has been suggested by HYLLEBERG *et al.* (1990, HEGY). It starts by assuming a low-order autoregression as the data generation process, possibly with added deterministic terms such as a constant, a linear time trend, and a strictly periodic cycle

$$\Phi(B)X_t = D_t + \varepsilon_t \quad .$$

We denote the lag operator by B and we use D_t for the deterministic part. $\Phi(.)$ is a polynomial of order p. The error term ε_t is specified as Gaussian white noise. The latter assumption can be relaxed considerably without affecting the asymptotic performance of the test but we will not focus on this issue here. By an algebraic one-one transformation, the autoregressive model can be expanded around the values of ± 1 and $\pm i$ to yield the representation

$$\begin{aligned} \Delta_4 X_t &= a_1 S(B) X_{t-1} + a_2 A(B) X_{t-1} + a_3 \Delta_2 X_{t-1} + a_4 \Delta_2 X_{t-2} \\ &\quad + \sum_{j=1}^{p-4} \phi_j \Delta_4 X_{t-j} + D_t + \varepsilon_t \quad . \end{aligned} \tag{1}$$

We use $S(B) = 1 + B + B^2 + B^3$ for the seasonal moving average operator, $A(B) = 1 - B + B^2 - B^3$ for the alternating summation, and $\Delta_j = 1 - B^j$ for the seasonal differencing operator. HEGY (1990) show that the polynomial $\Phi(.)$ contains roots at +1, -1, and $\pm i$ iff $a_1 = 0$, $a_2 = 0$, and $(a_3, a_4) = (0, 0)$, respectively. Therefore, they suggest running a least-squares regression based on (1) and using t–values on the coefficients $a_j, j = 1, \ldots, 4$ in order to test for the unit-root hypotheses. For the hypothesis $\Phi(\pm i) = 0$, i.e., $(a_3, a_4) = (0, 0)$, they also consider the corresponding F–statistic. The asymptotic distributions of these statistics are mixtures of normal or chi-square random variables that can be expressed as rational functions of Brownian motion integrals. Tabulated fractiles of these distributions can be used for testing the hypotheses at specified conventional significance levels, such as 5%. In summary, the persistence of seasonal shocks corresponds to sharp restrictions on the parameter space and therefore to the null hypotheses of the HEGY tests.

Although an autoregression is often a convenient first step in time-series modeling and the HEGY procedure remains valid asymptotically for ARMA structures provided that $p \to \infty$, some authors consider alternative approaches. A general non-parametric frame was analyzed by JOYEUX (1992), an unobserved-components approach was adopted by CANOVA AND HANSEN (1995, CH), and a test based on checking for seasonal over-differencing was presented by TAM AND REINSEL (1997). CANER (1998) presented another test that draws on similar ideas and shows that it has higher local power than the CH test.

Inspired by a traditional unobserved-components model that was suggested by HANNAN (1970), CH view the observed variable as the sum of a typically time-changing seasonal component and a purely stochastic short-memory stationary remainder

$$X_t = \mu + f_t' \alpha_t + e_t \quad . \tag{2}$$

The variable f_t consists of deterministic seasonal cycles at the frequencies $\omega = \pi, \pi/2$, i.e., $f_t' = (\cos(\pi t), \cos(\pi t/2), \sin(\pi t/2))$. The error term e_t is short-memory stationary but generally serially correlated. The coefficients α_t are specified to follow a random walk

$$\mathbf{A}' \alpha_t = \mathbf{A}' \alpha_{t-1} + u_t \quad ,$$

with a white-noise error term with its variance described by the scalar parameter τ^2. The matrix $\mathbf{A}$ is a selection matrix that can be modified in order to test for seasonal variation at a specific frequency and thus defines the dimensionality of u_t. In our notation, $\tau^2 = 0$ is equivalent to the *absence* of a seasonal unit root at $\omega = \pi$ for $\mathbf{A}_\pi' = (1, 0, 0)$ and to the *absence* of a seasonal unit root at $\omega = \pi/2$ for $\mathbf{A}_{\pi/2}' = (\mathbf{0}, \mathbf{I}_2)$ such that $\dim(\mathbf{A}_{\pi/2}) = 3 \times 2$. A Lagrange-multiplier (LM) statistic for this hypothesis can be constructed on the basis of a non-parametric generalized least-squares estimate of the regression equation (2). CH show that the asymptotic distribution of this LM statistic under the null hypothesis of no seasonal unit roots can be characterized by an integral over Brownian bridge variables. They call this distribution a 'generalized Von Mises distribution' and tabulate critical points.

Caner (1998) considers a very similar idea, starting with an autoregressive variant of the unobserved-components (UC) model (2),

$$\Phi(B)X_t = \mu + f_t'\alpha_t + e_t \quad , \quad \alpha_t = \alpha_{t-1} + u_t \quad . \tag{3}$$

The error terms e_t and u_t are white noise. This fully parametric version of Hannan's UC model is equivalent to a seasonal ARMA model

$$\Phi(B)S(B)X_t = \tilde{\mu} + \Theta(B)\zeta_t \quad . \tag{4}$$

The error series ζ_t is a function of the UC errors e_t and u_t. The null hypothesis of no seasonal unit roots can be expressed by findings of unit roots in the MA polynomial. Hence, the test incorporates the idea of over-differencing, which is also considered by Tam and Reinsel (1997, TR). The comparative advantage of Caner's test relative to the TR test is its ability to separate the cases of unit roots at $\omega = \pi$ and $\omega = \pi/2$. Each one of these tests has a counterpart in the literature for the case of testing a unit root at $\omega = 0$. A main difference to the CH test is that the model (4) is estimated fully parametrically, without taking resort to a spectral estimate. Again, the asymptotic distribution under the null can be expressed as a generalized Von Mises distribution, i.e., as an integral over Brownian bridge elements. For all details, see Caner (1998).

2.2 The general discrimination problem

All seasonal unit-root tests consider a discrimination problem among the basic types of seasonality: firstly, seasonal cycles generated by trigonometric functions of time; secondly, seasonal cycles generated by seasonal unit roots in linear time-series models. For convenience we refer to these two modeling ideas as 'deterministic' and 'stochastic' seasonality. With quarterly observations, the type of seasonality at the two seasonal frequencies is not necessarily the same. For example, one may have a seasonal unit root at $\omega = \pi$ and a deterministic cycle proportional to $\cos(\pi t/2)$. However, the joint occurrence of stochastic and deterministic seasonality at the *same* frequency is unlikely as it causes seasonal cycles of increasing amplitude that appear implausible (see FRANSES AND KUNST, 1999, for details).

Whereas the HEGY test uses the stochastic model as its null hypothesis and the deterministic one as its alternative, the CH and CANER tests use the opposite approach. As it is common in binary discrimination tests, the choice of null and alternative hypotheses apparently depends on technical requirements and not on theoretical considerations. We see no *a priori* reason to prefer any of the two testing ideas. It seems attractive to combine the two tests in order to increase the efficiency of the statistical discrimination procedure. Such a combination of the HEGY and the CH test was recommended by HYLLEBERG (1995) who, however, did not specify his recommendation in detail.

In the parallel problem of discriminating unit-root processes and trend-stationary processes, HATANAKA AND KOTO (1995) and HATANAKA (1996) indicate a formal solution. They consider the tests by DICKEY AND FULLER (1979) and by SAIKKONEN AND LUUKKONEN (1993), the former of which has the unit root as a null hypothesis and trend stationarity as the alternative, whereas the roles are reversed in the latter test. Transforming the sample evidence from the primary statistics into nominal p–values enables viewing the decision problem in the square $[0,1] \times [0,1]$. HATANAKA (1995) distinguishes three main empirical situations: (a) clear cases where one p–value is very small (say, less than 0.05 or 0.1) and the other is inconspicuous; (b) unclear cases where both hypotheses are seemingly rejected; (c) two inconspicuous p–values. When one is *a priori* unlikely to accept the joint occurrence of both features—as it would lead to implausible long-run behavior—or to accept the joint absence of them—as it conflicts with plain visual evidence or with the primary stimulus of the analysis—decisions have to be defined for the unclear cases (b) and (c). HATANAKA considers preferring hypotheses on the basis of a simple comparison of p–values. It is known from the statistical literature that such decision rule is, under certain conditions, equivalent to a Bayes test after assigning a prior probability of 0.5 to each of the hypotheses. Although HATANAKA does not use a formal Bayesian framework, some of his remarks—particularly the one on 'fair' tests—reveal that he actually uses Bayesian ideas.

The seasonal problem is very similar to HATANAKA's problem and deserves a similar treatment. We assign prior probabilities of 0.5 to each of the specific hypotheses of seasonal unit roots at $\omega = \pi$ and $\omega = \pi/2$ and the 'alternatives' of deterministic cycles at these frequencies. The nominal

p–values of both the HEGY and the CH test are calculated. Additional to the specification of prior probabilities, Bayesian decision procedures require the specification of a loss criterion. In line with HATANAKA, we specify loss by the simple 0–1 loss function in the spirit of a model selection problem. In other words, we assume that we are interested in maximizing the probability of finding the true model.

2.3 The frame

The frame is a *stochastic process* that defines the design for the decision problem. If a decision is searched among n hypotheses and each hypothesis is represented by a collection of stochastic processes $(X_t(\omega, \theta), \theta \in \Theta_i)$ for $i = 1, \ldots, n$, the basic rule is that the probability that a trajectory is taken from any of the n collections is n^{-1}. The index i can be viewed as a discrete parameter and this rule can be viewed as a discrete uniform prior over this discrete parameter space.

Within each of the n collections of processes or model classes, a weighting prior has to be defined for the parameter space Θ_i. There is no unanimous agreement about how to represent prior lack of information over arbitrary continuous parameter spaces. We have decided for simple technical priors. Bounded parameters will generally be drawn from uniform distributions, which we denote by $U(A)$ for the bounded set A, while unbounded nuisance parameters will be drawn from standard normal $N(0, 1)$ distributions. We are aware of the Bayesian argument that uniform priors are unlikely to reflect prior beliefs in testing sharp hypotheses—particularly for unit-root testing, see MARRIOTT AND NEWBOLD (1998)—but we note that the aim of our analysis differs from a usual Bayes test. Rather, our method crucially relies on well-defined simple priors that are comparable across experiments.

Note that neither the prior distributions nor the parameter spaces Θ_i are necessarily related across the classes and that θ may have different dimension in different model classes. In accordance with the Bayesian literature it is always assumed that ω and θ are independent.

With these definitions, the frame is a stochastic process that is generated by a sequence of hierarchical random steps. First, the discrete class parameter is drawn from a uniform distribution. Second, a continuous parameter θ is drawn from a technical prior. Third, a trajectory symbolically denoted by ω is drawn from the probability law of the fixed-parameter process $(X_t(\omega, \theta)|\theta)$. The first-step discrete decision problem is to infer the first random draw from the finite observed part of the trajectory. This problem is simplified by condensing the relevant information in statistics.

The effect of the discrete decisions is commonly evaluated by a loss function. Technical loss functions consider the correctness of the classification only. In other words, the purpose of the discrimination is maximizing the probability of correctly classifying the observed trajectory. Different loss functions are conceivable if the purpose of modeling is known, as for example in forecasting, but we note that technical loss directly matches the classical hypothesis testing approach.

In detail, in accordance with the discrete prior probability of 1/2 for each of the two hypotheses of deterministic and unit-root seasonality at both seasonal frequencies $\omega = \pi$ and $\omega = \pi/2$, we allot a prior probability of 0.25 to each of the following events: processes with seasonal unit roots at both frequencies, processes without seasonal unit roots, processes with a unit root at $\omega = \pi$ only, processes with a unit root at $\omega = \pi/2$ only. Additionally, within each of these four cases, we draw with a probability of 0.5 from those models where the HEGY test feels most at home and with a probability of 0.5 from those models that were suggested by Canova and Hansen (1995). For both partial frames, we extend the most primitive data generation design by two autoregressive and two moving-average lags. We feel that this setup gives equal chances to both procedures and also to each of the four partial hypotheses.

As a 'HEGY DGP' we choose the general model

$$\begin{aligned} & (1-L)(1+\rho_1 L)(1+\rho_2 L^2)(1-\phi_1 L)(1-\phi_2 L)X_t \\ = \; & \mu + f_t'\alpha + (1+\theta_1 L)(1+\theta_2 L)\varepsilon_t \end{aligned} \qquad (5)$$

with the 3–vector of seasonal constants $f_t = (\cos(\pi t/2), \sin(\pi t/2), \cos(\pi t))'$ and the corresponding coefficients vector $\alpha = (\alpha_1, \alpha_2, \alpha_3)'$. The four hypotheses correspond to the following cases:

1. unit roots at $\omega = \pi, \pi/2$: $\rho_1 = \rho_2 = 1$, $\alpha = 0$,
2. unit root at $\omega = \pi$: $\rho_1 = 1$, $\rho_2 \sim U(-1, 1)$, $\alpha_3 = 0$,
3. unit root at $\omega = \pi/2$: $\rho_1 \sim U(-1, 1)$, $\rho_2 = 1$, $\alpha_1 = \alpha_2 = 0$,
4. no seasonal unit roots: ρ_1, $\rho_2 \sim U(-1, 1)$.

Unless it is stated otherwise explicitly, all random drawings of parameters are mutually independent. The remaining parameters—for α_i, unless already specified by the above assumptions—are taken from simple non-informative distributions:

$$\begin{aligned} \mu, \alpha_i, i = 1, \ldots, 3 &\sim N(0,1), \\ \phi_1, \phi_2, \theta_1, \theta_2 &\sim U(-1,1), \\ \mathbb{E}(\varepsilon_t^2) = \sigma^2 &\sim \chi_1^2. \end{aligned}$$

As a 'CH DGP' we consider an unobserved-components structure as a general model

$$(1-L)(1-\phi_1 L)(1-\phi_2 L)y_t = \mu + f_t'\alpha_t + (1+\theta_1 L)(1+\theta_2 L)\varepsilon_t \qquad (6)$$

where α_t follows a vector AR(1) process with diagonal coefficient matrix $\mathbf{D}$ and Gaussian errors υ_t, i.e. $\alpha_t = \mathbf{D}\alpha_{t-1} + \upsilon_t$ with $\mathbb{E}(\upsilon_t \upsilon_t') = \sigma_\upsilon^2 \mathbf{I}_3$. This model generates seasonal unit roots whenever the diagonal matrix $\mathbf{D} = \text{diag}(d_1, d_2, d_3)$ with $|d_1|, |d_2|, |d_3| \leq 1$ contains a one as an autoregressive coefficient for a scalar component of α_t. In particular, one has

1. unit roots at $\omega = \pi, \pi/2$ for $d_1 = d_2 = d_3 = 1$,
2. unit root at $\omega = \pi/2$ for $d_1 = d_2 = 1$, $d_3 \sim U(-1,1)$,
3. unit root at $\omega = \pi$ for $d_1, d_2 \sim U(-1,1)$, $d_3 = 1$,
4. no seasonal unit roots for $d_1, d_2, d_3 \sim U(-1,1)$.

In all cases, we draw starting values for α_0 from a standard Gaussian $N(0, \mathbf{I}_3)$ distribution. The distributional specifications for the remaining parameters are:

$$\begin{aligned} \mu &\sim N(0,1), \\ \phi_1, \phi_2, \theta_1, \theta_2 &\sim U(-1,1), \\ \mathbb{E}(\varepsilon_t^2) = \sigma^2 &\sim \chi_1^2, \\ \sigma_\upsilon^2 &\sim \chi_1^2. \end{aligned}$$

3 The simulation results

For each DGP we simulate 100,000 realizations of the frame, such that 25,000 realizations correspond to each of the basic cases: unit roots at $\omega = \pi$ and $\omega = \pi/2$; unit root at $\omega = \pi$ only; unit root at $\omega = \pi/2$ only; no unit root. Each generated trajectory has a length ('sample size') of $T = 100$. For each trajectory, HEGY, CH, and CANER statistics for both seasonal frequencies are calculated and are transformed into p–values, which have been numerically calculated from a base-line simulation. The bivariate distributions of p–values for pairs of statistics at the same seasonal frequency are discretized in bins of length 0.05. This value has to be chosen large enough, such that the bins are not empty.

Each bin contains a certain finite amount of pairs of statistics. A certain proportion of these pairs stem from processes that have been generated with a unit root at the critical frequency, whereas the remaining pairs stem from processes without that unit root. When a bin is dominated by unit-root processes, this can be interpreted as evidence that a unit root is more likely or more probable, given that a pair in that bin area is observed and that trajectories have been generated from the known frame. In other words, the posterior probability for the unit-root hypothesis is larger than 0.5, hence one would be inclined to decide for a unit root if, as we assume, one is interested in maximizing the probability of a correct decision.

Ideally, the areas where a decision for a unit root is suggested and those where the opposite decision is obtained are separated by smooth curves. Due to sampling variation, the empirical counterparts to these *decision contours* often have a ragged and split appearance. Rather than confusing the evidence by arbitrary smoothing, we prefer to present the raw results, assuming that their asymptotic counterparts can be guessed from the figures.

These empirical decision contours are displayed for four cases that correspond to tests for unit roots at $\omega = \pi$ and $\omega = \pi/2$, based on pairs of the HEGY and CANER and of the HEGY and CH statistic. For the first case, i.e., the HEGY and CANER pair and $\omega = \pi$, we also present the empirical density for each of the eight model subclasses that constitute the frame, which permits some more detailed analysis of the evidence.

3.1 Caner and HEGY: Testing for a unit root at π

In this case, we give all results in some detail, including a full set of graphs. In order to save space, for the other cases we just outline the main features. It turns out that the main qualitative results are quite similar for all four experiments.

Figures 1,3,5,7 correspond to cases *with* a unit root at π, whereas Figures 2,4,6,8 display cases *without* that unit root. Correspondingly, the figures with odd numbers correspond to the null hypothesis of the CANER test and to the alternative of the HEGY test, while these roles are reversed for the even-numbered figures. Ideally, one would expect that the p–values are uniformly distributed under the null, while they are concentrated in the area close to 0 under the alternative.

Figure 1 approximately corresponds to this expectation. There is some concentration of the CANER p–values for *large* values, which just indicates that the distribution under the alternative has additional mass in the 'wrong' end relative to the null distribution and would be of little concern to the decision problem. The visual impression of the HEGY test performance does not deviate significantly from uniformity.

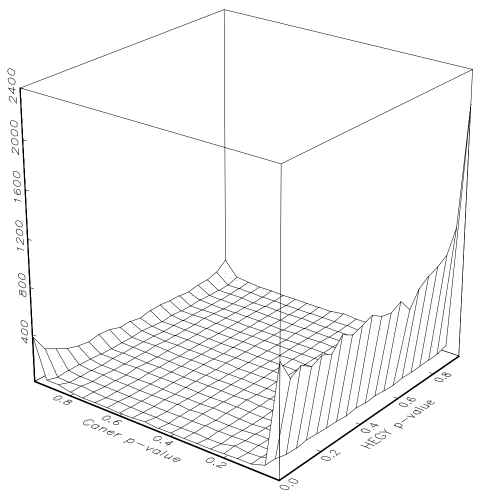

Figure 1: p–values for HEGY and Caner test procedure for a unit root at $\omega = \pi$ when the generating process is the HEGY model with unit roots at $\omega = \pi, \pi/2$.

In contrast, Figure 2 represents a less satisfactory situation. The HEGY test performs according to its construction and tends to reject its null. However, the CANER test *also rejects*, although it is not supposed to do so. A possible reason for this behavior is the generation process, which is based on the HEGY idea and a parametric autoregression and may bias the results in favor of the HEGY test. This feature had also been observed by HYLLEBERG (1995) for the CH and HEGY tests and hence represents no surprise.

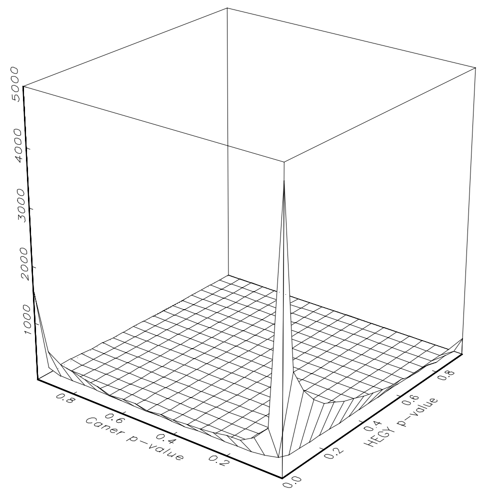

Figure 2: p–values for HEGY and Caner test procedure for a unit root at $\omega = \pi$ when the generating process is the HEGY model without unit roots at $\omega = \pi, \pi/2$.

Figure 3 is almost perfect. The fat tails of the CANER alternative have disappeared. Figure 4 confirms that they are caused by the unit root at the other seasonal frequency. The root at $\pi/2$ tends to tip the distribution away from the critical area for $\omega = \pi$. Unfortunately, the CANER test distribution is U–shaped indicating an incorrect 'size'. The apparently low 'power' of the HEGY test is a visual artifact due to the strong cross-frequency shift effect in the CANER test.

Figures 5 and 6 are the counterparts to Figures 1 and 2 with a modified generating mechanism that is supposedly more favorable to the CANER test. These figures indicate that the conclusion from HYLLEBERG's (1995) simulations that the CH and HEGY tests are most powerful in their individual 'habitats' may not be corroborated. The problems that we analyzed for Fig-

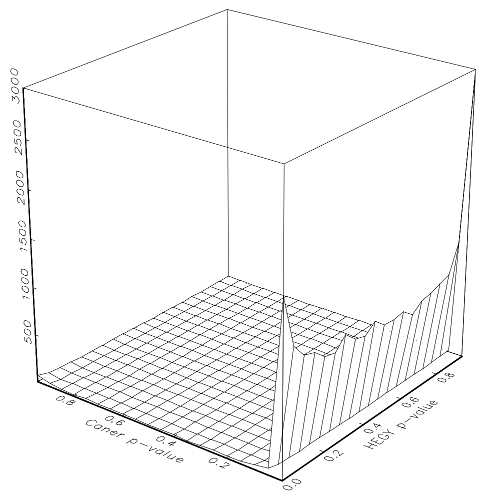

Figure 3: p–values for HEGY and Caner test procedure for a unit root at $\omega = \pi$ when the generating process is the HEGY model with a unit root at $\omega = \pi$ and no unit root at $\omega = \pi/2$.

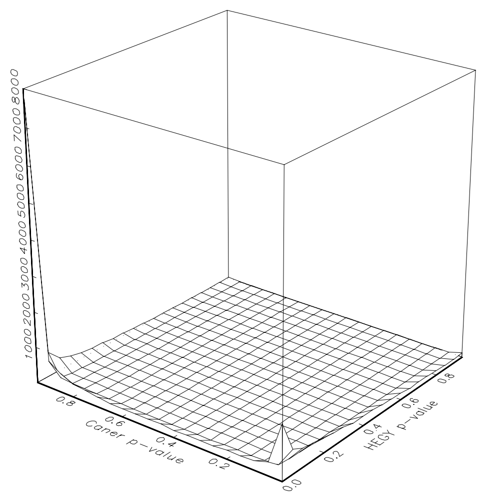

Figure 4: p–values for HEGY and Caner test procedure for a unit root at $\omega = \pi$ when the generating process is the HEGY model with a unit root at $\omega = \pi/2$ and no unit root at $\omega = \pi$.

ure 2 persist in Figure 6, although they have weakened. Similarly, Figures 7 and 8 come close to Figures 3 and 4, respectively.

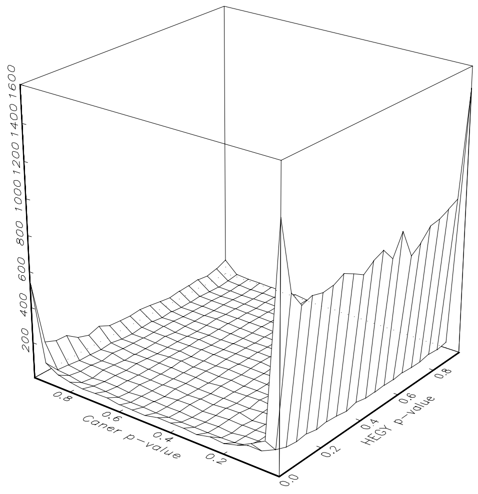

Figure 5: p–values for HEGY and Caner test procedure for a unit root at $\omega = \pi$ when the generating process is the CH model with unit roots at $\omega = \pi, \pi/2$.

Figure 9 summarizes all eight experiments by averaging aggregation. A wiggly curve delimits the area where the joint decision supports a unit root at π. The rough shape is likely to be due to sampling variation in our experiment. The curve could be smoothed by increasing the number of experimental replications, by smoothing the primary bins, or by smoothing the curve itself. Some tentative application of the latter two ideas resulted in a boundary that runs parallel to the ordinate (i.e., the p–value of the HEGY test is 0) axis, at a p–value of around 0.2, excepting the area close to the origin where the boundary moves in to smaller HEGY p–values. This would mean that the 'optimal' decision on whether a unit root at π is present or not mainly relies on the HEGY p–value. Only if the Caner p–value is very small (significant), the decision maker should decide for a unit root, even when the HEGY test rejects. For less significant Caner p–values, the decision maker should reject the unit root in the presence of even weak evidence against it on the basis of the HEGY statistic.

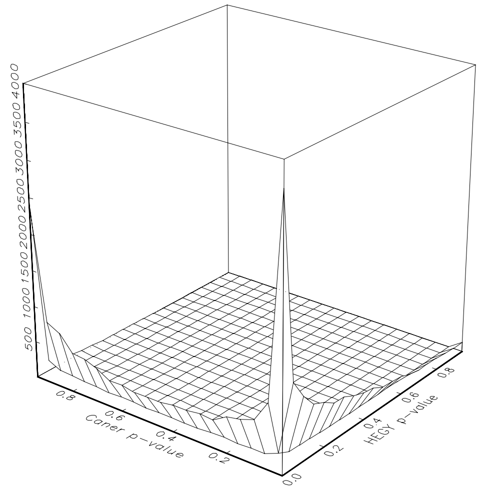

Figure 6: p–values for HEGY and Caner test procedure for a unit root at $\omega = \pi$ when the generating process is the CH model without unit roots at $\omega = \pi, \pi/2$.

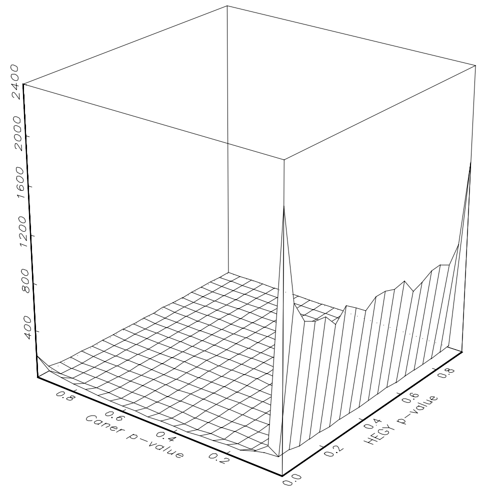

Figure 7: p–values for HEGY and Caner test procedure for a unit root at $\omega = \pi$ when the generating process is the CH model with a unit root at $\omega = \pi$ and no unit root at $\omega = \pi/2$.

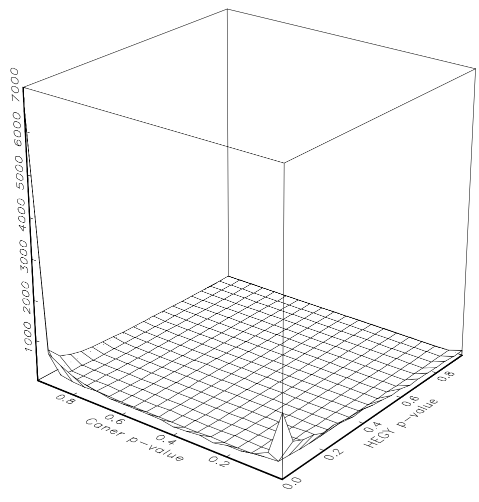

Figure 8: p–values for HEGY and Caner test procedure for a unit root at $\omega = \pi$ when the generating process is the CH model with a unit root at $\omega = \pi/2$ and no unit root at $\omega = \pi$.

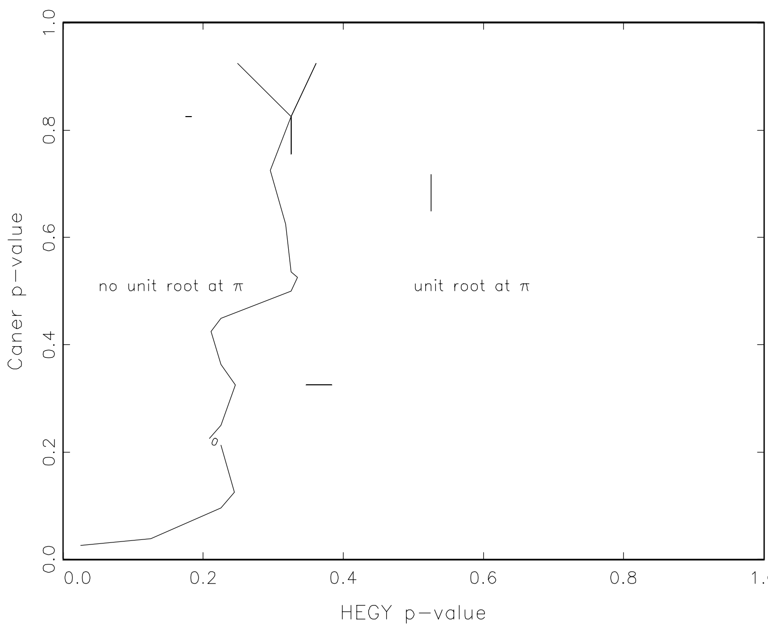

Figure 9: Optimal decision for the hypothesis of a unit root at π.

3.2 Caner and HEGY: Testing for a unit root at $\pi/2$

If a unit root at $\pi/2$ is being tested for, it is convenient to note that the sample contains less information with regard to this hypothesis, as the number of fully observed cycles of length 4 is exactly half the number of cycles of length 2. One may conjecture that this fact should be reflected in the decision procedure in such a way that *weaker* evidence against a 'null hypothesis' and hence a larger p–value might already lead the decision maker to decide against the 'null hypothesis'. Figure 10 displays a summary contour plot and is the counterpart to the Figure 9 in the previous subsection. Although the preference for the HEGY decision relative to the CANER–test decision is similar to the previous case, we note that the boundary has actually shifted left and not right, as one would have presumed. A much *stronger* evidence against the unit root at $\pi/2$ is necessary to recommend a decision for a deterministic 'annual' cycle, corresponding to p–values in the range of 0.05 to 0.1, i.e., in the classical and traditional range.

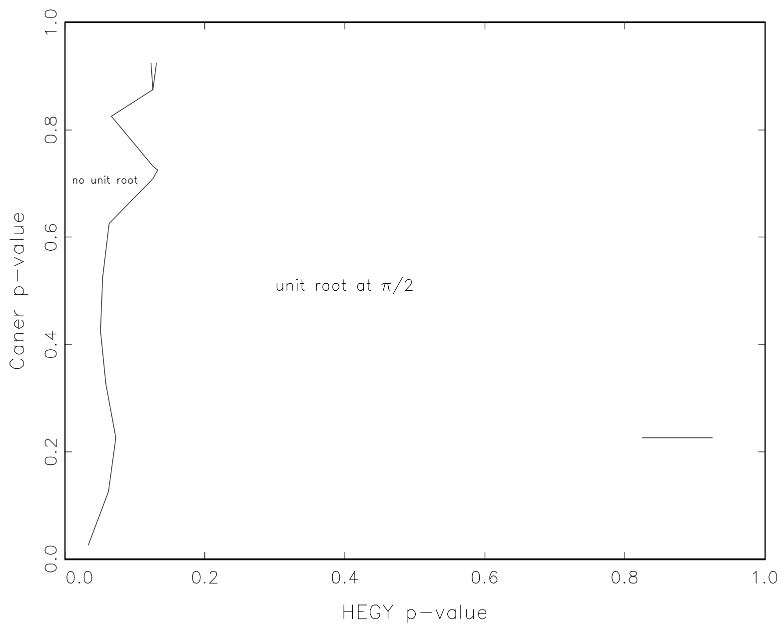

Figure 10: Optimal decision for the hypothesis of a unit root at $\pi/2$.

Although we cannot give a simple explanation for this unexpected phenomenon, we find that it is mainly caused by a rather high frequency of 'almost significant' p–values for the HEGY test, even when the generating mechanism is drawn from any of the cases with a unit root at $\pi/2$. We conjecture that the roles of the two seasonal frequencies may be reversed in larger samples and that the $\omega = \pi$ decision boundary will 'overtake' the $\omega = \pi/2$ decision boundary. We note that the frame may also be responsible for the phenomenon, as the design of purely real autoregressive roots comes arbitrarily close to the unit root of -1 but not to the unit root at $\pm i$ and hence may simplify discrimination at $\omega = \pi/2$. A variation of the frame in the direction of complex roots will be a task for future research.

3.3 Canova-Hansen and HEGY: Testing for a unit root at π

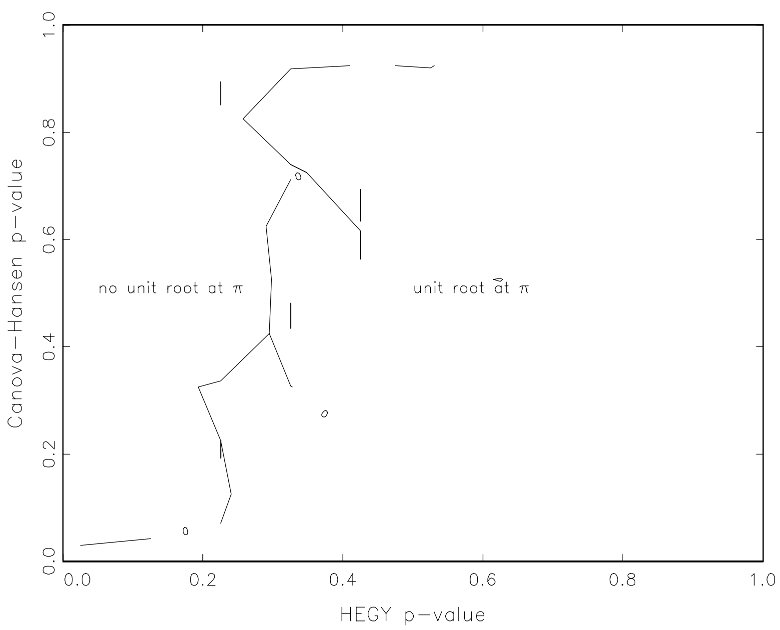

Figure 11: Optimal decision for the hypothesis of a unit root at π.

If the Caner test statistic is replaced by the Canova-Hansen (CH) test statistic, one may conjecture that the results will not be affected greatly by the modification, as both tests rely on a similar approach. However, one may expect a situation that is slightly less favorable to the HEGY test as compared to its rival, as the frame contains CH–type processes rather than Caner-type ones and the test should find it easier to make correct decisions in the framework where it was originally developed.

Figure 11 shows that there is in fact little change relative to the joint test of Caner and HEGY for $\omega = \pi$. The boundary between decision areas is roughly parallel to the ordinate axis at a HEGY p–value of 0.3 and bends westward close to the origin, i.e., for highly significant CH statistics. In summary, highly significant Caner or CH statistics tell that a unit root at π is very likely. If the Caner or CH statistic yields a p–value of over 5%, the information in the statistic can be ignored altogether and unit-root decision should be based on a HEGY statistic only. If the HEGY statistic yields a p–value of less than 20%, the unit-root hypothesis is recommended to be rejected.

3.4 Canova-Hansen and HEGY: Testing for a unit root at $\pi/2$

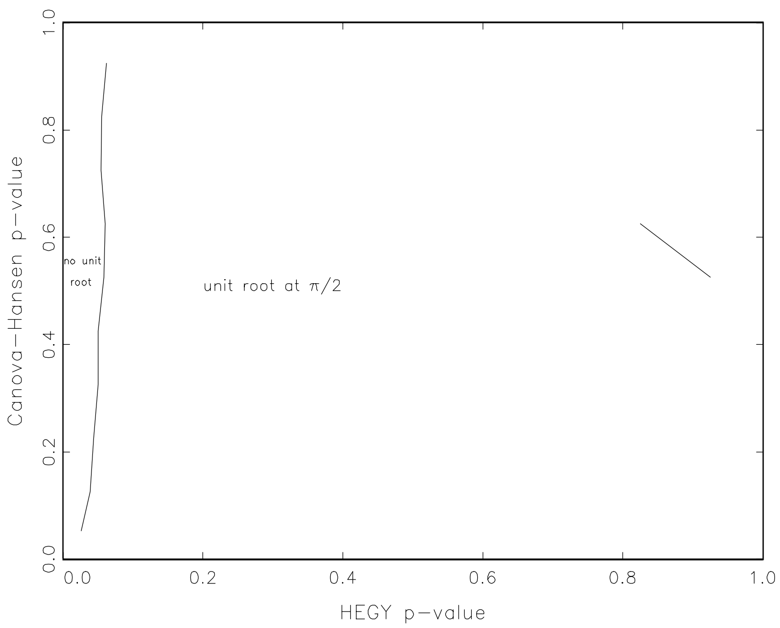

Figure 12: Optimal decision for the hypothesis of a unit root at $\pi/2$.

This is the last of the four cases, and again the boundary shows a marked leftward shift from the $\omega = \pi$ frequency. We note that the boundary is even left from the joint test based on a Caner and a HEGY statistic. It indicates that decision on the unit root at $\omega = \pi/2$ is to be based solely on the HEGY test. Whenever the HEGY test rejects at a significance level of around 5%, one opts for a model without a unit root at $\pi/2$, otherwise the unit root at the annual frequency is confirmed. In summary, for the problem of testing a unit root of ± 1, the information provided by the CH test statistic is redundant, given the value of the HEGY test statistic.

4 Summary and conclusions

Unit-root tests are conducted on one out of two basic principles: on autoregressive representations or on over-differencing and testing for a moving-average unit root. Occasionally, undue relevance has been attributed to the fact that, formally, null and alternative hypothesis change place across the two principles. More importantly, a researcher may try to reach a more conclusive decision by exploiting the information in two test statistics, one out of each class.

We consider two such test combinations. Firstly, seasonal unit roots at $\omega = \pi$ and at $\omega = \pi/2$ are checked by a pair of test statistics that is formed by a HEGY statistic and a statistic that has recently been recommended by Caner (1998). Secondly, we consider a joint test formed from the HEGY statistic and the statistic suggested by Canova and Hansen (1995). In order to evaluate the power of a joint decision, we rely on a scheme that was suggested by Hatanaka (1996) who uses a square formed from the p–values of two test statistics for the problem of testing for a unit root at $\omega = 0$.

Our evaluation scheme deviates from the cited work of Hatanaka (1996) or Hylleberg (1995), as we employ a pseudo-Bayesian framework and assume that, *a priori*, the same probability is allotted to the unit-root hypothesis and to the no-unit-roots hypothesis. We further employ a complex non-informative prior scheme within either hypothesis. We conceded that variations in this prior scheme may influence our results but we doubt whether they do so in an important qualitative way.

Our main result is that it appears that the HEGY test is more important in reaching a decision on the hypotheses of interest than its competitors that are derived from an over-differencing paradigm.

For the semi-annual cycle at $\omega = \pi$, we find that the over-differencing tests can be used, in a first round, to sort out trajectories that clearly point to a unit root. Among the remaining trajectories, the HEGY statistic should be applied with a loose significance level. In other words, unless the over-differencing (Caner or CH) tests clearly recommend a unit root of -1, such a unit root is to be rejected on the basis of even slight indications of its non-existence in a HEGY test.

For the annual cycle at $\omega = \pi/2$, we find that the over-differencing tests provide little guidance and that the decision on a unit root at $\pm i$ should preferably be based solely on a HEGY test at a traditional significance level. We conjecture that this result is robust with respect to replacing the HEGY test by another test from the autoregressive class, such as a seasonal variant of the semi-parametric test of Saikkonen and Luukkonen (1993).

Our results do not contradict the simulations of Hylleberg (1995) who recommends to use either test class—his study is restricted to the HEGY and CH specifications—in their 'natural habitat'. Because the generating law is not known to the researcher, this recommendation would require an additional selection procedure as a first step. It is conceivable that such a refined testing procedure would be more powerful than our global approach that solely uses the two statistics at hand. Such a comparison may constitute a further area for future research.

References

[1] Abadir, K.M., and Taylor, A.M.R. (1999) 'On the definitions of (co-)integration'. *Journal of Time Series Analysis* **20**, 129–138.

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

[3] 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.

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

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

[6] ——— and Kunst, R.M.(1999) 'On the Role of Seasonal Intercepts in Seasonal Cointegration'. *Oxford Bulletin of Economics and Statistics* **61**, 409–434.

[7] Ghysels, E., and Perron, P. (1993) 'The Effect of Seasonal Adjustment Filters on Tests for a Unit Root', *Journal of Econometrics* 55, 57–98.

[8] Hannan, E.J. (1970) *Multiple Time Series*. Wiley.

[9] Hatanaka, M. (1996) *Time-Series-Based Econometrics: Unit Roots and Co-Integration*. Oxford University Press.

[10] Hatanaka, M., and Koto, Y. (1995) 'Are There Unit Roots in Real Economic Variables? An Encompassing Analysis of Difference and Trend Stationarity', *Japanese Economic Review* **46**, 166–190.

[11] Hylleberg, S. (1992) *Modelling Seasonality*. Oxford University Press.

[12] ——— (1995) 'Tests for seasonal unit roots: General to specific or specific to general?', *Journal of Econometrics* **69**, 5–25.

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

[14] Joyeux, R. (1992) 'Tests for Seasonal Cointegration Using Principal Components'. *Journal of Time Series Analysis* **13**, 109–118.

[15] Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., and Shin, Y. (1992) 'Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure are We that Economic Time Series Have a Unit Root?', *Journal of Econometrics* **54**, 159–178.

[16] Lee, H.S. (1992). 'Maximum likelihood inference on cointegration and seasonal cointegration', *Journal of Econometrics* **54**, 1–47.

[17] Marriott, J., and Newbold, P. (1998) 'Bayesian Comparison of ARIMA and Stationary ARMA Models'. *International Statistical Review* **66**, 323–336.

[18] Saikkonen, P., and Luukkonen, R. (1993) 'Testing for a Moving Average Unit Root in Autoregressive Integrated Moving Average Models', *Journal of the American Statistical Association* **88**, 596–601.

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