DOCUMENTOS DE TRABAJO

BILTOKI

D.T. 2013.03

Seasonal Stability Tests in *gretl*
An Application to International Tourism Data

Ignacio Díaz-Emparanza y M. Paz Moral

eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea

Facultad de Ciencias Económicas.
Avda. Lehendakari Aguirre, 83
48015 BILBAO.

Documento de Trabajo BILTOKI DT2013.03
Editado por el Departamento de Economía Aplicada III (Econometría y Estadística) de la Universidad del País Vasco.

ISSN: 1134-8984

Seasonal Stability Tests in *gretl*. An Application to International Tourism Data

Ignacio Díaz-Emparanza*

Mª Paz Moral†

University of the Basque Country UPV/EHU

September 10, 2013

Abstract

The seasonal stability tests of Canova & Hansen (1995) (CH) provide a method complementary to that of Hylleberg et al. (1990) for testing for seasonal unit roots. But the distribution of the CH tests are unknown in small samples. We present a method to compute numerically critical values and P-values for the CH tests for any sample size and any seasonal periodicity. In fact this method is applicable to the types of seasonality which are commonly in use, but also to any other.

Keywords: seasonality, unit roots, surface response analysis

*ignacio.diaz-emparanza@ehu.es (Corresponding author) Address: Avenida Lehendakari Aguirre, 83. E48015 Bilbao, Spain. Financial support from research project ECO2010-15332 from Ministerio de Ciencia e Innovación, and from UPV/EHU *Econometrics Research Group*, Basque Government grant IT-642-13, is gratefully acknowledged. The SGI/IZO-SGIker UPV/EHU is gratefully aknowledged for its generous allocation of computational resources.

†mpaz.moral@ehu.es. Financial support from research project SAIOTEK-2012 S-PE12UN088 and grant IT869-13 from Basque Government is gratefully acknowledged.

1 Introduction

Seasonality is one of the key characteristics of economic time series. When seasonal time series are included in a regression, seasonal unit roots may affect the properties of estimators in the same way as unit roots in the trend (zero frequency). The structure of the seasonal pattern is important for modelling and prediction. In the specification of a model it is necessary to properly identify the nature of the seasonal component: deterministic, stochastic dominated by transitory features or by permanent stochastic cycles (seasonal unit roots).

Several tests for seasonal unit roots and seasonal stability may be found in the literature. Among the methods for testing seasonal unit roots one of the more used is that proposed by Hylleberg, Engle, Granger & Yoo (1990), which test for seasonal unit roots separately for each frequency. Canova & Hansen (1995) (hereinafter CH) proposed a method to test the stability of the seasonal pattern of a series against an alternative of one or several seasonal unit roots. This method is based on the representation of the series in the state-space. As mentioned in Canova & Hansen (1995, page 238), their tests are based on a Lagrange multiplier as a tool to test the significance of the variance in the state equation.

The distribution of the CH tests are unknown in small samples. We present a method to compute numerically critical values and P-values for the tests for any sample size and any seasonal periodicity.

2 Seasonal stability tests

Let y_t be a real valued variable observed S times per year. The Canova-Hansen tests for stability are based on the residuals of the following auxiliary regression:

$$y_t = \mu + x_t'\beta + Seas_t + \varepsilon_t, \quad t = 1, 2, \ldots, T \tag{1}$$

where T is the number of observations or sample size, x_t is a $k \times 1$ vector of explanatory variables, $Seas_t$ is a deterministic seasonal component and $\varepsilon_t \sim (0, \sigma^2)$ is an error uncorrelated with x_t and $Seas_t$. The dependent variable y_t must bee free of unit roots at the zero frequency, so Canova and Hansen suggest to difference the observed series in order to eliminate the zero frequency unit root.

The deterministic seasonal component may be specified into two different ways: seasonal dummies or trigonometric terms. In the first case the auxiliary regression (1) is:

$$y_t \quad = \quad x_t'\beta + d_t'\alpha + \varepsilon_t \tag{2}$$

where d_t is an $S \times 1$ vector of seasonal dummy indicators and α is an $S \times 1$ parameter vector, being S the seasonal periodicity ($S = 4$ for quarterly data, $S = 12$ for monthly data,...). To study whether the seasonal intercepts α change over time, Canova and Hansen consider stochastic variation of a martingale form: $A'\alpha_t = A'\alpha_{t-1} + u_t$ where α_0 is fixed and u_t is a martingale difference sequence (MDS) with covariance matrix $E(u_t u_t') = \tau^2 G$. A is an $S \times a$ selection matrix which serve to select the elements of α that we allow to stochastically change under the alternative. Testing stability of the jth intercept can be achieved by choosing A to be the unit vector with a 1 in the jth element and zeros elsewhere.

The equation (1) may be also specified in a form equivalent to (2) defining the seasonal component by means of trigonometric terms. In this case, the auxiliary regression may be written as:

$$y_t = \mu + x_t'\beta + \sum_{j=1}^{q} f_{jt}'\gamma_j + \varepsilon_t, \tag{3}$$

where $q = [S/2]$ is the integer part of $S/2$ and f_{jt} is the deterministic cyclical process at the seasonal frequency $\theta_j = \frac{2\pi j}{S}$, $j = 1, \ldots, q$, which, defining S^* as $S/2 - 1$ if S is even and

$(S-1)/2$ if S is odd, may be defined as:

$$f'_{jt} \quad = \quad \begin{cases} (\cos(\theta_j t), \ \sin(\theta_j t)) & j = 1, \ldots, S^* \\ \cos(\theta_j t) & j = S/2 \quad \text{(only for } S \ even) \end{cases} \tag{4}$$

Stacking the q elements of the previous sum in a vector $\gamma = (\gamma_1, \ldots \gamma_q)'$, $f_t = (f_{1t}, \ldots, f_{qt})'$ the equation (3) may be expressed as:

$$y_t = \mu + x_t'\beta + f_t'\gamma + \varepsilon_t. \tag{5}$$

For example for quarterly data, $S = 4$, $q = 2$, $S^* = 1$ and $f_t' = \left(\cos(\frac{\pi}{2}t), \quad \sin(\frac{\pi}{2}t), \quad \cos(\pi t)\right)$.

Under (5), the seasonal pattern is stable and γ is a vector of $S-1$ seasonal coefficients which are constant over time. To setup the tests of seasonal stability against seasonal unit roots Canova and Hansen propose for γ_t the process $A'\gamma_t = A'\gamma_{t-1} + u_t$ with γ_0 fixed and u_t a MDS. Where A is a $(S-1) \times a$ selection matrix such that, for example, $A = I_{S-1}$ may be used to test whether the entire vector is stable, $A = (\tilde{0} \quad I_2 \quad \tilde{0})'$ (commensurate with γ_i) to test for a unit root only at a specific frequency j $(\neq \pi)$ and $A = (\tilde{0} \quad 1)'$ serve for testing a unit root at frequency π.

The variance-covariance matrix of u_t is $E(u_t u_t') = \tau^2 G$ where G is a full rank $a \times a$ matrix and $\tau^2 \geq 0$ is real valued. When $\tau^2 = 0$ $\gamma_t = \gamma_0$ and the model has no seasonal unit roots. When $\tau^2 > 0$, y_t has a unit root at the seasonal frequencies determined by A.

Considering the hypothesis test of $H_0 : \tau^2 = 0$ against $H_1 : \tau^2 > 0$ they propose a LM test statistic which takes the form:

$$L = \frac{1}{T^2} tr \left((A'\hat{\Omega}^f A)^{-1} A' \sum_{t=1}^{T} \hat{F}_t \hat{F}_t' A \right), \tag{6}$$

where $\hat{F}_t = \sum_{j=1}^{t} f_j \hat{\varepsilon}_j$, $\hat{\Omega}^f$ is a consistent estimate of the long run covariance matrix of $f_t \varepsilon_t$,

$$\widehat{\Omega}^f = \sum_{k=-m}^{k=m} w\left(\frac{k}{m}\right) \frac{1}{T} \sum_{t=1}^{T} f_{t+k} \hat{\varepsilon}_{t+k} f_t' \hat{\varepsilon}_t \tag{7}$$

and $w()$ is a kernel which gives a positive semidefinite matrix, such as the Bartlett kernel.

When testing stability of the jth intercept (in the model with dummy variables) f_t should be substituted by d_t in (6) and (7).

The large-sample distribution of L was studied by Nyblom(1989) and Hansen(1990,1992), establishing that under $H_0,\ L \xrightarrow{d} VM(a)$, the *generalized Von Mises* distribution with a degrees of freedom which is tabulated in the Canova-Hansen paper. In spite of new studies of these asymptotic distributions, as Harvey (2005) for example, not much has been done in obtaining tools to be used for small samples.

3 Numerical distribution functions

The finite sample distribution functions of four different CH statistics for testing seasonal stability are studied here.

1. In model (5), and for even values of S, the statistic L for the individual stability test of the cycle with frequency π, which will be denoted L_π. In this case $A = (\tilde{0} \quad 1)'$, $a = 1$ and the asymptotic distribution of L_π is a VM(1) (which corresponds to the first row of Table 1 in Canova & Hansen (1995)).

2. In the same model, for any S, the statistic L for testing stability at any other individual frequency $\theta_j \neq \pi$, which will be denoted as L_j. Now $A = (\tilde{0} \quad I_2 \quad \tilde{0})'$ has two columns, so $a = 2$ and L_j follows asymptotically a VM(2) distribution whose values can be found

at the second row of the mentioned table.

3. In model (2) (with dummy variables), we will study also the distribution of the L statistic for testing stability in an individual season, i.e. stability of the coefficient of a specific season. This statistic will be denoted L_D and as stated in CH (Theorem 4) has also a VM(1) asymptotic distribution.

4. The statistic L for a joint test of stability at all the seasonal frequencies may be created from the model (5) with trigonometric terms. This statistic will be denoted L_f. In this case $A = I_{S-1}$ and the large sample distribution of this statistic is VM($S-1$). Table 1 in CH contains critical values for this distribution for $S - 1 = 1 \ldots, 13$.

We present here a procedure, based on *response surface regressions*, that allows to obtain critical values and P-values for these tests for any sample size and any seasonal periodicity.

The first step in implementing the response surface regressions is to estimate the relevant quantiles of the distributions of the statistics for several combinations of T and S from a large set of Monte Carlo simulations. Following MacKinnon (2002), the process is then repeated M = 100 times for each value of T to obtain more accurate results. Each experiment consists of a great number of replications (in particular N = 100,000 are used here), where a series y_t is generated by the data generation process $y_t = \varepsilon_t$ with $\varepsilon_t \sim nid(0, 1)$ and the equation estimated is the auxiliary regression (2) or (5).

For each set of replications the quantiles of the relevant statistics are calculated for two alternative deterministic seasonal terms: trigonometric and seasonal dummies and it is convenient also to consider the possibility of a linear trend in the auxiliary regression and a lag of the dependent variable. This implies that for each pair of T and S four different deterministic components will be considered and eight different DGPs will be simulated and estimated.

There is another parameter that is also involved in the calculation of the CH statistics: the truncation lag, m, for the long-run covariance matrix. For determining m we used two alternative automatic criteria[1] that satisfy the condition recommended that $m \to \infty$ as $T \to \infty$ such that $m^5/T = O(1)$: $m_1 = [0.75\sqrt{T}]$ and $m_2 = [4(T/100)^{2/9}]$.

The quantiles obtained from the simulations are then used as values for the dependent variable in a regression depending on T and S, the *surface response regression*. This estimated regression is used to calculate forecasts of a quantile for given values of T and S and, based on these forecasts, estimations of the critical values and P-values may be obtained.

3.1 Design of the experiments

We assume that the distribution functions of the CH test statistics in small samples depend on the sample size T and the seasonal periodicity S. For estimating the surface responses we need to carry out a set of Monte Carlo simulations that covers a wide range of the values normally used in the econometric applications for these parameters. We consider these values:

$$S = \quad 4, 5, 7, 12, 24, 48, 52$$
$$T = \quad 54, 104, 154, 250, 500, 750, 2000, 5000$$

A full factorial design may be setup for this procedure, and this means to estimate the auxiliary regressions (2) and (5) for every alternative of the deterministic term and repeat them for 53 different combinations of T and S (Note than three of the 56 possible combinations above cannot be made because of lack of degrees of freedom).

Taking into account the necessity of a great number of replications for each simulated model and the already mentioned repetitions suggested by MacKinnon, it was necessary to evaluate

[1]Although the Monte Carlo simulations were made with the two formulas, in practice we found a very similar results for the estimated P-values. Then for simplifying the interface for the user, in the final Gretl algorithm we decided to use only m_1.

Table 1: D-optimal design

Point:	1	2	3	4	5	6	7	8	9	10	11	12
S:	24	12	7	4	24	7	4	24	24	52	48	7
T:	250	500	250	154	54	154	500	750	2000	250	750	5000

Point:	13	14	15	16	17	18	19	20	21	22	23
S:	52	48	52	5	5	12	52	4	24	12	5
T:	500	104	154	5000	250	54	750	104	154	2000	2000

carefully the computing costs of the entire project (in terms of computing time and money as we were going to use a computing service that has an economic cost for us). We launched a pilot project to assess such costs and after analyzing the results, we decided to use a fractional design with no more than 23 points. For the calculus of the design the response-surface equation

$$q_i^\alpha(T,S) = \theta_\infty^\alpha + \theta_1^\alpha \frac{1}{T} + \theta_2^\alpha \frac{1}{T^2} + \theta_3^\alpha \frac{S}{T} + \theta_4^\alpha \frac{S}{T^2} + e_i \tag{8}$$

was taken as reference model for the distribution of the CH tests. Here $q_i^\alpha(T,S)$ denotes quantile α obtained from the i-th experiment with given values of T and S. A D-optimal fractional design over 15000 runs was calculated making use of the RcmdrPlugin.DoE package of R (see http://prof.beuth-hochschule.de/groemping/DoE). The design finally obtained is in Table 1.

3.2 Quantile regressions, P-values and critical values

The Monte Carlo simulations were programmed in Gretl 1.9.11 (See Cottrell & Lucchetti 2013). From each Monte Carlo experiment a record is made of the 221 quantiles estimated of the statistics L_j, L_π, L_f, and L_D for probabilities $\alpha = 0.0001, 0.0002, 0.0005, 0.001, 0.002, \ldots,$ $0.01, 0.015, \ldots, 0.99, 0.991, \ldots, 0.999, 0.9995, 0.9998, 0.9999$ and the quantiles estimated are used as dependent variables in response surface regressions. After some experimentation we consider for L_j and L_π rather apropriate the structure of (8) . For L_D it seems necessary to include third degree terms so the best equation in the set we analyzed is

$$q_i^\alpha(T,S) = \theta_\infty^\alpha + \theta_1^\alpha \frac{1}{T} + \theta_2^\alpha \frac{1}{T^2} + \theta_3^\alpha \frac{1}{T^3} + \theta_4^\alpha \frac{S}{T} + \theta_5^\alpha \frac{S}{T^2} + \theta_6^\alpha \frac{S}{T^3} + e_i \tag{9}$$

Given that the asymptotic distribution of statistic L_f depends on S, it needs a different structure for the surface-regression. Between the possible alternatives the following one has been proved to have good properties

$$q_i^\alpha(T,S) = \theta_0^\alpha + \theta_1^\alpha \frac{1}{T} + \theta_2^\alpha \frac{1}{T^2} + \theta_3^\alpha \frac{S}{T} + \theta_4^\alpha \frac{S}{T^2} + \theta_5^\alpha (S-1) + \theta_6 (S-1)^2 + e_i \tag{10}$$

In equations (8) and (9) parameter θ_∞^α represents quantile α of the asymptotic distribution when $T \to \infty$. In the case of statistic L_f, the surface-response regression (10) shows that the expectation of $q^\alpha(T,S)$ when $T \to \infty$ is a function depending on S: $\theta_0 + \theta_5(S-1) + \theta_6(S-1)^2$ whose estimation is in good agreement with what we see in Table 1 of Canova-Hansen.

When the parameters of these surface-regression equations are estimated by ordinary least squares the errors are heteroscedastic with variance depending on T and S, so a weighted least squares method is recommended here. (Tables in `http://bit.ly/CHpval` in file `tablas.pdf` show the coefficients estimated by weighted least squares for equations (8), (9) and (10) for the quantiles of probability 0.10, 0.05 and 0.01).

After the surface regression is estimated for the 221 quantiles for every statistic, an interpolation between these values may be made using the method by MacKinnon (1996), which is also used for example in Harvey & van Dijk (2006) and Diaz-Emparanza (2013). Consider the regression

$$\Phi^{-1}(\alpha) = \gamma_0 + \gamma_1 \hat{q}(\alpha) + \gamma_2 \hat{q}^2(\alpha) + \gamma_3 \hat{q}^3(\alpha) + e_\alpha \tag{11}$$

where α denotes one of the 221 points at which the quantiles are estimated, with $0 < \alpha < 1$, $\hat{q}(\alpha)$ denotes the estimate of q^α and $\Phi^{-1}(\alpha)$ is the inverse of the $\chi^2(2)$ distribution. The surface-regression equation is usually estimated with a small, odd number of points, ℓ, around the specified significance level, in particular, $\ell =$ 9, 11, 13 or 15 points are considered reasonable. To account for heteroscedasticity and serial correlation a feasible GLS estimator may be used, with a symmetric covariance matrix with elements

$$\hat{\omega}_{ij} = s.e\left(\hat{\theta}_\infty^{\alpha_i}\right) s.e\left(\hat{\theta}_\infty^{\alpha_j}\right) \sqrt{\frac{\alpha_i(1-\alpha_j)}{\alpha_j(1-\alpha_i)}}, \quad i < j, \tag{12}$$

where the standard errors of $\hat{\theta}_\infty^{\alpha_i}$ are obtained from the OLS estimation of equation (11). In order to calculate the P-value for an observed test statistic, τ_*, it is possible simply to estimate equation (11) for an small set of values of $\hat{q}(\alpha)$ near τ_* and then compute $P^* = \Phi(\hat{\gamma}_0 + \hat{\gamma}_1\tau_* + \hat{\gamma}_2\tau_*^2 + \hat{\gamma}_3\tau_*^3)$. An algorithm prepared by the authors in the gretl scripting language (now called *Hansl*) is available for the users in `http://bit.ly/CHpval`.

To calculate the critical values of the tests the following equation may be used

$$\hat{q}(p) = \delta_0 + \delta_1\Phi^{-1}(p) + \delta_2(\Phi^{-1}(p))^2 + \delta_3(\Phi^{-1}(p))^3 + e_p^* \tag{13}$$

The method consists of first finding the quantile p^* from the set of 221 mentioned above that is closest to the probability p whose critical value is to be obtained, then estimating the δ coefficients in (13) with the $(\ell-1)/2$ quantiles above and the $(\ell-1)/2$ quantiles below p^* and finally evaluating the right hand side of the regression estimated at p to obtain the desired critical value.

4 Example: International tourism demand

Two quarterly indicators of international tourism demand in Spain are analysed here: foreign *same-day visitors* (i.e. people who stay in Spain less than 24 hours) and *international tourists* (those who stay at least 24 hours). The sample goes from 1995:1 to 2013:1 (73 observations, see Figure 1). The data are taken from the FRONTUR survey carried out by the Institute of Tourist Studies `http://www.iet.tourspain.es`.

Figure 1: Quarterly visitor entries to Spain

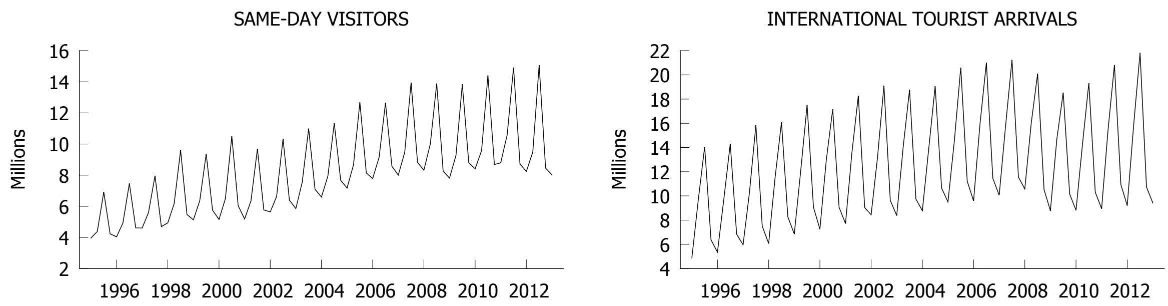

Table 4 shows the results of the Canova-Hansen tests. The auxiliary model is estimated in log-differences with a dummy variable for Easter effect. Both series display significant seasonal patterns but their structure seems to be rather different. The joint test does not reject at a 5% significance level the presence of an stable seasonal pattern for the *same-day visitors* series. But the evidence of instability is strong in the series of *tourists*, for which the null joint hypothesis of stability at all seasonal frequencies is rejected. Analysing each frequency individually, we may see that the tests do not reject the null hypotheses of stability at the biannual (π) frequency for both of the series. The rejection of the joint test in the *tourists* series is due to the presence of a unit root at the annual cycle ($\pi/2$ frequency). The test of stability of individual seasonal intercepts also reinforces this conclusion, showing significant time variations for the *tourists* series in three quarters.

Table 2: Canova-Hansen tests of stability at seasonal frequencies for foreign entries. Degrees of freedom (T-k) = 67, lag order = 4

	Seasonal dummies				Seasonal frequencies		
	Q1	Q2	Q3	Q4	$\pi/2$	π	Joint
International Tourists							
Statistic	0.3072	0.9003*	0.7843*	0.6767*	1.0528*	0.1344	1.1336*
P-value	0.1512	0.0000	0.0000	0.0035	0.0000	0.5016	0.0002
Same-day visitors							
Statistic	0.1240	0.3183	0.2974	0.2224	0.3557	0.2621	0.5761
P-value	0.5459	0.1409	0.1606	0.2595	0.2621	0.1898	0.2473

* indicates significance at the 5% level

This instability in the seasonal pattern of the *tourists* series may be caused by changes in the consumer behaviour, who make more trips along the year but shorter (they tend to have several shorter holiday periods rather than concentrating them in one month). While the visits shorter than a day are associated with other factors, for example border proximity, which may be more stable.

5 Concluding remarks

In this paper a numerical method for calculating the small sample distributions of four statistics for analyzing the stability of seasonal patterns, proposed by Canova & Hansen (1995), is presented. The numerical method is based on response surface regressions, following a procedure similar to that established by MacKinnon (1996).

The main contributions of this paper are two: first, we describe the numerical approximation and provide a program written in the Gretl scripting language which calculates the Canova-Hansen statistics and their associated P-values and second, for users of any other software package, we offer tables of estimated coefficients of the response surface regression for the quantiles of probability $\alpha = 0.10, 0.05$ and 0.01. Both the computer program and the tables (in pdf format) are available online in `http://bit.ly/CHpval`.

The main advantage of this procedure is that it provides the P-value of the CH statistics whatever the periodicity and sample size of the data. This opens the range of applicability of these tests to a wide range of time series, for example, to high frequency data.

References

Canova, F. & Hansen, B. E. (1995), 'Are seasonal patterns constant over time? a test for seasonal stability', *Journal of Business and Economic Statistics* **13**, 237–252.

Cottrell, A. & Lucchetti, R. (2013), *Gretl User's Guide*, Department of Economics, Wake Forest University. http://gretl.sourceforge.net.

Diaz-Emparanza, I. (2013), 'Numerical distribution functions for seasonal unit root tests', *Computational Statistics & Data Analysis*. (forthcoming) Available online at http://www.sciencedirect.com.

Harvey, A. C. (2005), *A Unified Approach to Testing for Stationarity and Unit Roots*, Cambridge University Press.

Harvey, D. I. & van Dijk, D. (2006), 'Sample size, lag order and critical values of seasonal unit root tests', *Computational Statistics & Data Analysis* **50**, 2734–2751. Available at http://www.sciencedirect.com.

Hylleberg, S., Engle, R. F., Granger, C. W. J. & Yoo, B. S. (1990), 'Seasonal integration and cointegration', *Journal of Econometrics* **44**, 215–38.

MacKinnon, J. (1996), 'Numerical distribution functions for unit root and cointegration tests', *Journal of Applied Econometrics* **11**, 601–618.

MacKinnon, J. (2002), Computing numerical distribution functions in econometrics, *in* A. Pollard, D. J. K. Mewhort & D. F. Weaver, eds, 'High Performance Computing Systems and Applications', Vol. 541 of *The Kluwer International Series in Engineering and Computer Science*, Springer US, pp. 455–471. 10.1007/0-306-47015-2_45.