AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

Check for updates

The Stata Journal (2016)
16, Number 3, pp. 740–760

The lag-length selection and detrending methods for HEGY seasonal unit-root tests using Stata

Tomás del Barrio Castro
Department of Applied Economics
University of the Balearic Islands
Palma, Spain
tomas.barrio@uib.es

Andrii Bodnar
andrii.bodnar@gmail.com

Andreu Sansó
Department of Applied Economics
University of the Balearic Islands
Palma, Spain
Andreu.sanso@uib.es

Abstract. The article extends the previous Hylleberg, Engle, Granger, and Yoo seasonal unit-root test commands (Baum and Sperling, 2001, https://ideas.repec.org/c/boc/bocode/s416502.html; Depalo, 2009, *Stata Journal* 9: 422–438), which allow for the use of both quarterly and monthly data. It is also possible to choose between ordinary least-squares and generalized least-squares detrending (Rodrigues and Taylor, 2007, *Journal of Econometrics* 141: 548–573) procedures to deal with the deterministic part of the process. The command allows for the use of the sequential method proposed by Hall (1994, *Journal of Business and Economic Statistics* 12: 461–470) and Ng and Perron (1995, *Journal of the American Statistical Association* 90: 268–281), the adaptation of the modified Akaike information criteria to the case of seasonal unit-root tests (del Barrio Castro, Osborn, and Taylor, 2016, *Econometric Reviews* 35: 122–168) as well as the inclusion of Akaike information and Bayesian information criteria to determine the order of augmentation of the serial correlation in the augmented Hylleberg, Engle, Granger, and Yoo regression. Finally, the use of the command is illustrated with an empirical application to the case of monthly passenger airport arrivals to Palma de Mallorca.

Keywords: st0453, hegy, HEGY test, GLS detrending, optimal augmentation lag

1 Introduction

Seasonality is important in any study involving time-series data, such as studies on hospitalizations, bouts of depression, influenza, tourism arrivals, or economic variables. In the case of economic time series, seasonal adjustment techniques such as X-11 and autoregressive (AR) integrated moving-average-based methods are widely used by practitioners to eliminate seasonal fluctuations. But the use of seasonally adjusted data in

© 2016 StataCorp LP

applied work, instead of simplifying the data, may actually produce undesirable features in the filtered data, for example, noninvertible moving-average processes and complicated dynamics; see Maravall (1993) for details. Hence, the use of raw (nonseasonally adjusted) data when using econometric tools based on fitting AR processes or vector AR models is recommended.

Over the last three decades, there has been debate in the literature on whether seasonality is deterministic or attributable to one or multiple unit roots at seasonal frequencies. Over the years, this has led to the development of a large number of seasonal unit-root testing procedures; see, for example, Dickey, Hasza and Fuller (1984), Osborn et al. (1988), Hylleberg, Engle, Granger, and Yoo (1990) (henceforth, HEGY), and Rodrigues and Taylor (2004, 2007). The HEGY approach has become the most popular one to test for the presence of seasonal unit roots.

Recently, Rodrigues and Taylor (2007) extended the generalized least-squares (GLS) detrending procedure proposed by Elliott, Rothenberg, and Stock (1996) for the zero-frequency augmented Dickey–Fuller test to the augmented HEGY tests. Rodrigues and Taylor (2007) show that GLS-detrended augmented HEGY tests have asymptotic local power functions that lie arbitrarily close to the Gaussian power envelopes. Also, they find power gains over the standard ordinary least-squares (OLS)-detrended HEGY tests when using GLS-detrended HEGY tests.

Dealing with serial correlation and determining the order of augmentation in augmented HEGY regression has been an important issue in seasonal unit-root testing procedures because their performance depends critically on handling this feature of the data. In the context of the standard unit-root tests, the most popular way of determining the order of augmentation in the augmented Dickey–Fuller tests is with the modified Akaike information criterion (MAIC) criteria proposed by Ng and Perron (2001); recently, the MAIC criteria have been extended by del Barrio Castro, Osborn, and Taylor (2016) to augmented HEGY tests.

The HEGY seasonal unit-root testing routines were developed for Stata to consider quarterly data and OLS detrending. In particular, the `hegy4` command, developed by Baum and Sperling (2001), performs the HEGY procedure to test for seasonal unit roots with quarterly data. It provides the t statistics and F statistics to test for the presence of unit roots at the zero, semiannual, and annual frequencies. The statistics are reported jointly with the critical values. The program can automatically conduct a sequential t test to determine the optimal lag length to be included in the auxiliary regression, which is in line with the proposal made by Hall (1994) and Ng and Perron (1995). Additionally, Depalo (2009) continues to develop the seasonal unit-root tests under Stata. The `sroot` command (Depalo 2009) has a more diverse output: the `generate()` option creates time series of HEGY auxiliary variables and residuals. According to Engle et al. (1993), this information is important, and seasonal integration can be studied from the transformed variables.

In this article, we expand on various aspects of the previous HEGY test commands in Stata. First, our `hegy` command allows for the use of quarterly and monthly data. Also, the user can choose between OLS detrending and GLS detrending to deal with the specified deterministic part of the process. For all the possible specifications of the deterministic part (see section 2 for details) and for both OLS and GLS detrending, the command provides critical values, obtained from response surfaces according to the proposal in MacKinnon (1996) (see del Barrio Castro, Bodnar, and Sansó [2015] for details), at the 1%, 5%, and 10% levels of significance. With our `hegy` command, a wider range of possibilities for determining the order of augmentation in the HEGY regression is available. As with the `hegy4` command, one can use the sequential method proposed by Hall (1994) and Ng and Perron (1995). One can also use this method jointly with the conventional method based on the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) criteria for determining the lag length in the augmented HEGY regression as well as the recently extended seasonal version of the MAIC criteria suggested by del Barrio Castro, Osborn, and Taylor (2016). Finally, the `hegy` command reports the autocorrelation function (ACF) and partial ACF, along with the Ljung–Box Q statistics of the residuals. In developing the `hegy` command, we use the base structure of the `hegy4` command from Baum and Sperling (2001).

This article is organized as follows: Section 2 discusses all the methodological details of `hegy` and describes its implementation. Section 3 presents an empirical application of the command using the monthly time series of tourism arrivals to Palma de Mallorca from the UK. Section 4 concludes.

2 Methodology

2.1 The seasonal model

Here we use the conventional HEGY methodology to test for the presence of seasonal unit roots. We use the augmentation lag selection and detrending methods suggested by del Barrio Castro, Osborn, and Taylor (2012) and Rodrigues and Taylor (2007).

The data-generating process of a univariate seasonal time series is assumed to be as follows,

$$\begin{aligned} y_{St+s} &= \mu_{St+s} + x_{St+s} \\ \alpha(L)x_{St+s} &= u_{St+s}, \quad s = 1-S, \ldots, 0, \quad t = 1, 2, \ldots, N \end{aligned} \tag{1}$$

where S denotes the number of seasons. Because the Stata commands deal with monthly ($S = 12$) and quarterly ($S = 4$) data, in the rest of the article, we assume that S is even. N represents the number of years. We assume that the observed time series y_{St+s} can be decomposed into two parts: the deterministic part μ_{St+s} and the stochastic part x_{St+s}. $\alpha(L)$ is an AR(S) polynomial $\alpha(L) = (1 - \alpha_1^* L - \alpha_2^* L^2 - \cdots - \alpha_S^* L^S)$, where L is the usual lag operator. This polynomial can be factorized as

$$\alpha(L) = (1-\alpha_0 L)(1+\alpha_{S/2}L)\prod_{j=1}^{S^*}\left[1 - 2\left\{\alpha_j \cos\left(\frac{2\pi j}{S}\right) - \beta_j \sin\left(\frac{2\pi j}{S}\right)\right\} L \right. \\ \left. + \left(\alpha_j^2 + \beta_j^2\right) L^2\right]$$

with $S^* = S/2 - 1$. Our focus is to test for the presence of unit roots in the polynomial $\alpha(L)$. Note that the parameter α_0 of $(1-\alpha_0 L)$ is associated with the zero frequency; that the parameter $\alpha_{S/2}$ of $(1+\alpha_{S/2}L)$ is associated with the Nyquist frequency (π); and that the parameters α_j and β_j of $(1-2[\alpha_j \cos\{(2\pi j)/S\} - \beta_j \sin\{(2\pi j)/S\}]L + (\alpha_j^2+\beta_j^2)L^2)$ are associated with the conjugate (harmonic) seasonal frequencies $(2\pi j)/S$ and $2\pi - (2\pi j)/S$ for $j = 1, \ldots, S^* = S/2 - 1$.

Following Smith and Taylor (1998), Rodrigues and Taylor (2007), and Smith, Taylor, and del Barrio Castro (2009), we can define six scenarios for the deterministic part μ_{St+s}: no deterministic terms, zero-frequency intercept (one intercept), zero-frequency intercept and a trend, seasonal intercepts, seasonal intercepts and a zero-frequency trend, and seasonal intercepts and trends. Hence, in terms of $\mu_{St+s} = \delta' z_{St+s,\zeta}$, we have the following:

Case 0: No deterministic terms:

$$\mu_{St+s} = 0$$

Case 1: Only a constant:

$$z_{St+s,1} = [1] \quad \text{with} \quad \delta = (\delta_0)$$

Case 2: Constant and a zero-frequency trend:

$$z_{St+s,2} = [1, St+s]' \quad \text{with} \quad \delta = \left(\delta_0, \overline{\delta}_0\right)'$$

Case 3: Seasonal intercepts:

$$z_{St+s,3} = [1, \cos(2\pi(St+s)/S), \sin(2\pi(St+s)/S), \ldots, \\ \cos(2\pi S^*(St+s)/S), \sin(2\pi S^*(St+s)/S), (-1)^{St+2}]'$$

with $\delta = (\delta_0, \delta_1', \ldots, \delta_{S^*}', \delta_{S/2})'$, $\delta_k = (\delta_{k,1}, \delta_{k,2})'$, $k = 1, \ldots, S^*$.

Case 4: Seasonal intercepts and a zero-frequency trend:

$$z_{St+s,4} = \left[z_{St+s,3}', St+s\right]'$$

with $\delta = \left(\delta_0, \delta_1', \ldots, \delta_{S^*}', \delta_{S/2}, \overline{\delta}_0\right)'$, $\delta_k = (\delta_{k,1}, \delta_{k,2})'$, $k = 1, \ldots, S^*$

Case 5: Seasonal intercepts and trends:

$$z_{St+s,5} = \left[z'_{St+s,3}, (St+s)z'_{St+s,3}\right]'$$

with $\delta = (\delta_0, \delta'_1, \dots, \delta'_{S^*}, \delta_{S/2}, \overline{\delta}_0, \overline{\delta}'_1, \dots, \overline{\delta}'_{S^*}, \overline{\delta}_{S/2})'$, $\delta_k = (\delta_{k,1}, \delta_{k,2})'$, $k = 1, \dots, S^*$.

As shown by Smith and Taylor (1998) and Smith, Taylor, and del Barrio Castro (2009), the inclusion of seasonal intercepts allows for tests invariant to the presence of nonzero initial conditions to be obtained under the null hypothesis of seasonal integration; the inclusion of both seasonal intercepts and trends allows tests invariant to the presence of both nonzero initial values and seasonal drifts to be obtained. As we will discuss later, the deterministic part considered in the seasonal unit-root procedures plays an important role in the distribution of the tests.

The overall null hypothesis of the presence of all unit roots is $H_0 \colon \alpha(L) = 1 - L^S = \Delta_S$; hence, the time series y_{St+s} is seasonally integrated. This can be partitioned into the following null hypotheses:

$$\begin{aligned} H_{0,0} \colon \alpha_0 &= 1 \\ H_{0,S/2} \colon \alpha_{S/2} &= 1 \\ H_{0,k} \colon \alpha_k &= 1, \quad \beta_k = 0 \quad k = 1, \dots, S/2 - 1 \end{aligned}$$

Under $H_{0,0}$, a unit root is associated with the zero frequency; under $H_{0,S/2}$, a unit root is associated with the Nyquist frequency (π). And under $H_{0,k}$, we have a pair of complex conjugate roots associated with the seasonal harmonic frequencies $(2\pi k)/S$ for $k = 1, \dots, S^* = S/2 - 1$. The alternative hypothesis is of stationarity at one or more of the zero or seasonal frequencies; that is, $H_1 = U_{j=0}^{S/2} H_{1,j}$, where

$$\begin{aligned} H_{1,0} \colon \alpha_0 &< 1 \\ H_{1,S/2} \colon \alpha_{S/2} &< 1 \\ H_{1,k} \colon \alpha_k^2 + \beta_k^2 &< 1 \quad k = 1, \dots, S/2 - 1 \end{aligned}$$

The filters that remove possible unit roots at the zero $\{\Delta_0^0(L)\}$, Nyquist $\{\Delta_{S/2}^0(L)\}$, and seasonal harmonic $\{\Delta_k^0(L)\}$ frequencies $(2\pi k)/S$ for $k = 1, \dots, S^* = S/2 - 1$ are defined as follows,

$$
\begin{aligned}
\Delta_0^0(L) &= \frac{1-L^S}{1-L} = (1+L+L^2+\cdots+L^{S-1}) \\
\Delta_{S/2}^0(L) &= -\frac{1-L^S}{1+L} = -(1-L+L^2-\cdots-L^{S-1}) \\
\Delta_k^0(L) &= -\frac{1-L^S}{\{1-2\cos(\omega_k)L+L^2\}} = -\frac{\sum_{j=0}^{S-1}\sin\{(j+1)\,\omega_k\}L^j}{\sin(\omega_k)} \\
&= -(1-L^2)\sum_{j\neq k, j=1}^{S^*}\left\{1-2\cos(\omega_j)L+L^2\right\} \\
&\quad \text{for} \quad k=1,\ldots,S/2-1
\end{aligned}
$$

As can be seen in the next section, these filters are closely related to the auxiliary variables used in the HEGY procedure.

2.2 The HEGY tests

Following Hylleberg et al. (1990) and Smith, Taylor, and del Barrio Castro (2009), we carry out the regression-based approach for testing for unit roots in $\alpha(L)$ in two steps. First, we detrend the data to obtain tests that will be invariant to the parameters that characterize the deterministic part: μ_{St+s}. The most popular methods use OLS detrending (see, for example, Hylleberg et al. [1990] and Smith, Taylor, and del Barrio Castro [2009]) or GLS detrending (see Rodrigues and Taylor [2007]). With OLS detrending, the resulting detrended time series is obtained from $y_{St+s}^{\xi} = y_{St+s} - \widehat{\delta}' z_{St+s,\xi}$, where $\widehat{\delta}'$ comes from the OLS regression of $\mathbf{y}$ on $\mathbf{z}_\xi$, where $\mathbf{y}$ is a vector with the generic element y_{St+s} and $\mathbf{z}_\xi$ is a matrix with the generic row element $z_{St+s,\xi}$. Then, ξ corresponds to the deterministic part being considered. With GLS detrending, the resulting detrended time series is defined as $y_{St+s}^{\xi} = y_{St+s} - \widehat{\delta}' z_{St+s,\xi}$, and $\widehat{\delta}'$ is obtained from the OLS regression of y_c on $z_{c,\xi}$ with

$$
\begin{aligned}
y_c = (&y_{1-S}, y_{2-S} - \alpha_1^c y_{1-S}, y_{3-S} - \alpha_1^c y_{2-S} - \alpha_2^c y_{1-S}, \ldots, y_0 - \alpha_1^c y_{-1} - \cdots \\
&- \alpha_S^c y_{1-S}, \Delta_c y_1, \ldots, \Delta_c y_T)' \\
z_{c,\xi} = (&z_{1-S,\xi}, z_{2-S,\xi} - \alpha_1^c z_{1-S,\xi}, z_{3-S,\xi} - \alpha_1^c z_{2-S,\xi} - \alpha_2^c z_{1-S,\xi}, \ldots, z_{0,\xi} - \alpha_1^c z_{1,\xi} \\
&- \cdots - \alpha_S^c z_{1-S,\xi}, \Delta_c z_{1,\xi}, \ldots, \Delta_c z_{T,\xi})'
\end{aligned}
$$

and

$$
\begin{aligned}
\Delta_c &= (1-\overline{\alpha}_0 L)\left(1-\overline{\alpha}_{S/2}L\right)\sum_{j=1}^{S/2-1}\left[1-2\left\{\overline{\alpha}_j\cos\left(\frac{2\pi j}{S}\right)\right\}L+\overline{\alpha}_j^2 L^2\right] \\
&= \left(1-\sum_{j=1}^{S}\alpha_j^c L^j\right)
\end{aligned}
$$

where

$$\overline{\alpha}_0 = 1 + \frac{c_0}{ST}, \quad \overline{\alpha}_{S/2} = 1 + \frac{c_{S/2}}{ST}, \quad \overline{\alpha}_j = 1 + \frac{c_j}{ST} \quad j = 1, 2, \ldots, S/2 - 1$$

Table 1 presents the values for the parameters c_j, $j = 0, 1, 2, \ldots, S/2$, suggested by Elliott, Rothenberg, and Stock (1996), Gregoir (2006), and Rodrigues and Taylor (2007) and that are used in our command.

Table 1. The quasidifferencing detrending parameters

Parameter	Only a constant (Case 1)	Constant and zero-frequency trend (Case 2)	Seasonal intercepts	Seasonal intercepts and zero-frequency trends	Seasonal intercepts and trends
c_0 [(a)]	−7.00	−13.5	−7.00	−13.5	−13.5
$c_j \quad j = 1, 2, \ldots, S/2 - 1$ [(b)]	0.00	0.00	−3.75	−3.75	−8.65
$c_{S/2}$ [(a)]	0.00	0.00	−7.00	−7.00	−13.5

Source: [(a)] Elliott, Rothenberg, and Stock (1996) and [(b)] Gregoir (2006)

Second, with the detrended data obtained by OLS or GLS detrending, we use the approach of Hylleberg et al. (1990), which is based on expanding $\alpha(L)$ around the zero and seasonal frequency unit roots $\{\exp(\pm i2\pi j/S);\, j = 0, \ldots, S/2\}$; hence, we can write the testing equation of the augmented HEGY approach as

$$\Delta_S y^{\xi}_{St+s} = \pi_0 y^{\xi}_{0,St+s} + \pi_{S/2} y^{\xi}_{S/2,St+s} + \sum_{j=1}^{S/2-1} (\pi_{1j} y^{\xi}_{1j,St+s} + \pi_{2j} y^{\xi}_{2j,St+s}) + \sum_{j=1}^{k} d_j \Delta_S y^{\xi}_{St+s-j} + e^{\xi}_{St+s,k} \tag{2}$$

where

$$y^{\xi}_{0,St+s} = \Delta^0_0(L) y^{\xi}_{St+s-1} = \sum_{i=0}^{S-1} y^{\xi}_{St+s-i-1} \tag{3}$$

$$y^{\xi}_{S/2,St+s} = \Delta^0_{S/2}(L) y^{\xi}_{Sn+s-1} = \sum_{i=0}^{S-1} \cos\{(i+1)\pi\} y^{\xi}_{St+s-i-1}$$

$$y^{\xi}_{1j,St+s} = -\{\cos(\omega_j) - L\} \Delta^0_j(L) y^{\xi}_{Sn+s-i-1} = \sum_{q=0}^{S-1} \cos\{(q+1)\omega_j\} y^{\xi}_{St+s-q-1}$$

$$y^{\xi}_{2j,St+s} = \sin(\omega_j) \Delta^0_j(L) y^{\xi}_{St+s-1} = -\sum_{q=0}^{S-1} \sin\{(q+1)\omega_j\} y^{\xi}_{St+s-q-1}$$

$$j = 1, \ldots, S/2 - 1$$

Under the HEGY approach, the possible presence of serial correlation in the innovation u_{St+s} in (1) is handled via augmenting regression (2.3) by adding lags of $\Delta_S y^{\xi}_{St+s}$, which approximates the possible serial correlation in u_{St+s} by a finite AR(k) process. As del Barrio Castro, Osborn, and Taylor (2012) show, this approach is valid for innovations that are allowed to follow a general linear process; hence, u_{St+s} allows for causal and invertible AR moving-average (p, q) representation. See del Barrio Castro, Osborn, and Taylor (2012) for details regarding assumptions.

As Hylleberg et al. (1990) and Smith, Taylor, and del Barrio Castro (2009) show, testing $H_{0,0} : \alpha_0 = 1$ and $H_{0,S/2} : \alpha_{S/2} = 1$ is equivalent to testing $H_{0,0} : \pi_0 = 0$ and $H_{0,S/2} : \pi_{S/2} = 0$, respectively. Note that the coefficients π_0 and $\pi_{S/2}$ in (2.3) are associated with the corresponding auxiliary variables $y^{\xi}_{0,St+s}$ and $y^{\xi}_{S/2,St+s}$, which refer to the unit roots at the zero and Nyquist frequencies, respectively. In both cases, the test is carried out using lower-tailed regression t-test statistics.

Testing for the pairs of complex conjugate unit roots ($H_{0,k} : \alpha_k = 1$, $\beta_k = 0$, $k = 1, \ldots, S/2 - 1$) is equivalent to testing $H_{0,k} : \pi_{1k} = 0, \pi_{2k} = 0$ associated with the auxiliary variables $y^{\xi}_{1j,St+s}$ and $y^{\xi}_{2j,St+s}$. For this purpose, a lower-tailed regression t-test statistic for $\pi_{1k} = 0$ and a two-tailed regression t-test statistic for $\pi_{2k} = 0$ are proposed in the original HEGY article. Also, an upper-tailed regression F-type test is suggested to test the joint null $H_{0,k} : \pi_{1k} = 0$, $\pi_{2k} = 0$. Further, Ghysels, Lee, and Noh (1994) and Smith, Taylor, and del Barrio Castro (2009) consider joint frequency tests, in particular, the F-type test to control for the presence of any seasonal unit roots by checking the hypotheses $H_{0,S/2} : \pi_{S/2} = 0$ and $H_{0,k} : \pi_{1k} = 0, \pi_{2k} = 0$. Finally, the presence of any unit root is tested jointly by the hypotheses $H_{0,0} : \pi_0 = 0$, $H_{0,S/2} : \pi_{S/2} = 0$, and $H_{0,k} : \pi_{1k} = 0, \pi_{2k} = 0$.

Burridge and Taylor (2001) and Smith, Taylor, and del Barrio Castro (2009) for AR innovations, del Barrio Castro and Osborn (2011) for moving-average innovations, and del Barrio Castro, Osborn, and Taylor (2012) for general linear processes show that if regression (2.3) is properly augmented, the limiting null distributions of the t statistics for unit roots at the zero and Nyquist frequencies and joint F-type statistics are pivotal, while those of the t statistics at the harmonic seasonal frequencies depend on nuisance parameters, which are functions of the parameters associated with the process followed by the innovation. So in practice, we recommend using only the t statistics for unit roots at the zero and Nyquist frequencies and joint F-type statistics for all other unit roots. Hence, the `hegy` command reports results from the left-tail t-test statistics for zero and Nyquist frequencies as well as all the F-type test statistics described above.

As Smith and Taylor (1998) and Smith, Taylor, and del Barrio Castro (2009) show, when there is no deterministic part ($\mu_{St+s} = 0$), the distribution of the tests is a function of standard Brownian motions. With OLS detrending, when seasonal intercepts are considered (case 1), the distribution of the tests is a function of demeaned Brownian motions. With a zero-frequency intercept, only the distribution of the tests associated with the zero frequency is a function of demeaned Brownian motion. When seasonal intercepts and trends are included, the distribution of the tests is a function of demeaned and detrended Brownian motions. Finally, if seasonal intercepts and a zero-frequency trend are considered, the distribution of all the tests is a function of demeaned Brownian motions, except for the zero-frequency test, which is a function of demeaned and detrended Brownian motion. With GLS detrending, when considering the inclusion of seasonal intercepts and seasonal intercepts with trends, one can obtain the limit distribution of the statistics by replacing the detrended Brownian motions with their relevant local GLS detrended analogues; see theorem 5.1 of Rodrigues and Taylor (2007, 559–560).

2.3 Lag-length selection methods

With the hegy command, users can choose between five different methods to determine the order of augmentation of (2.3). They can provide a desired order of augmentation for the augmented HEGY regression, use information criteria (such as the AIC, BIC, and MAIC), or use the Hall (1994) and Ng and Perron (1995) sequential method to determine the lag length to be used in (2.3). In the last case, users can specify a maximum lag length ($k_{\max}$) or use the following default rule, $k_{\max} = \text{Int}\lfloor 12(T/100)^{1/4} \rfloor$, where $\text{Int}[\cdot]$ is the integer part and T is the total sample size ($T = SN$). When the sequential method is chosen, the procedure starts by fitting model (2.3) with $k = k_{\max}$ and tests sequentially for the significance of the coefficient associated with Δy^{ξ}_{St+s-k} until the null is rejected. To determine the individual significance of the lags of Δy^{ξ}_{St+s}, one uses the standard normal distribution critical values and selects a 10% level of significance as the default value following the suggestion of Ng and Perron (1995). When using the sequential method, one can change the default 10% for another significance level (say, 5% or 1%), but the results reported in Ng and Perron (1995) show that with the 10% level, the sequential method achieves better results. In the method based on the information criteria, the process chooses the order of augmentation from $k_{\max}$ to 0 that obtains the lowest value for the AIC, BIC, or the MAIC criteria. With the MAIC criteria, we must first consider the usual AIC,

$$\text{AIC} = \ln\left(\widehat{\sigma}^2_k\right) + \frac{2k}{T}$$

where $\widehat{\sigma}^2_k = \text{RSS}_k/(T-k)$ and RSS_k is the residual sum of squares (RSS) obtained from testing the regression with k lags. Then, the MAIC has an additional penalization term, $\tau_T(k)$, added where T is the number of observations. The optimal augmentation lag corresponds to the lowest AIC value,

$$\text{MAIC} = \ln\left(\widehat{\sigma^2_k}\right) + \frac{2\{\tau_T(k)+k\}}{T-k_{\max}}$$

$$\tau_T(k) = \left(\widehat{\sigma}^2_k\right)^{-1}\left\{\widehat{\pi}^2_0 \sum_t \sum_s \left(y^{\xi}_{0,St+s}\right)^2 + \widehat{\pi}^2_{S/2} \sum_t \sum_s \left(y^{\xi}_{S/2,St+s}\right)^2 \right.$$
$$\left. + \sum_{j=1}^{S/2-1} \widehat{\pi}^2_{1j} \sum_t \sum_s \left(y^{\xi}_{1j,St+s}\right)^2 + \widehat{\pi}^2_{2j} \sum_t \sum_s \left(y^{\xi}_{2j,St+s}\right)^2 \right\}$$

where $\tau_T(k)$, as part of the penalty function, considers the possible nonstationarity of the regressors $(y^{\xi}_{0,St+s}, y^{\xi}_{S/2,St+s}, y^{\xi}_{1j,St+s}, y^{\xi}_{2j,St+s})$ in (2.3) (see Ng and Perron [2001] and del Barrio Castro, Osborn, and Taylor [2016] for details). Following Perron and Qu (2007) and del Barrio Castro, Osborn, and Taylor (2016), the MAIC is computed based on (2.3) and (2.3) with OLS detrending to determine the order of augmentation of the augmented HEGY regression even if GLS detrending is considered.

2.4 Critical values

The `hegy` command uses the results of del Barrio Castro, Bodnar, and Sansó (2015). The critical values at 1%, 5%, and 10% significance levels are calculated from the quantile functions, which are estimated as follows,

$$q^p(T_i) = \theta^p_\infty + \theta^p_1 T_i^{-1} + \theta^p_2 T_i^{-2} + \theta^p_3 T_i^{-3} + \epsilon_i$$

where q_p is a quantile value at p significance level and T_i is a corresponding sample size. The quantile coefficients are computed based on 9.6 million Monte Carlo simulations for each of the 27 sample sizes and 6 possible deterministic cases for OLS detrending and 5 possible cases for GLS detrending. Monthly and quarterly frequencies are considered separately.

2.5 The hegy command

`hegy` *varname* [*if*] [*in*] [, `maxlag(`*integer*`)` `det(`*string*`)` `level(`*integer*`)` `mode(`*string*`)` `residuals(`*string*`)` `noac` `noreg` `gls`]

You must `tsset` your data before using `hegy`, and you must specify either `quarterly` or `monthly` with `tsset`; see [TS] **tsset**.

Options

`maxlag(`*integer*`)` specifies the maximum lag order to be included when augmenting the model with AR terms. Its default value is obtained from the following expression, $\text{Int}\lfloor 12(T/100)^{1/4}\rfloor$, where $\text{Int}[\cdot]$ is the integer part and T is the total sample size ($T = SN$). This option is used for all methods: AIC, MAIC, BIC, sequential t test, and fixed lag.

`det(`*string*`)` controls for the deterministic terms present in the time series. *string* may take on values `none`, `const`, `seas`, `trend`, `strend`, or `mult` to specify the deterministic part of the process to be tested. The default is `det(seas)`, as suggested by Hylleberg et al. (1990) and Ghysels, Lee, and Noh (1994). It indicates that a set of seasonal intercepts be included in the regression; `none` specifies that no deterministic variables be included; `const` specifies only a constant; `trend` specifies that a linear trend be included along with a constant term; `strend` specifies that a linear trend be included along with seasonal intercepts; `mult` specifies that seasonal intercepts along with seasonal trends be included (the case of multiplicative seasonality recommended by Smith and Taylor [1998]). If the `gls` option is selected, `det()` cannot be `none`.

`level(`*integer*`)` indicates the significance level for the sequential t test as a percentage. This option must be used with the `mode(seq)` option. The default is `level(10)`, which corresponds to a 10% significance level.

`mode(`*string*`)` specifies the method for selecting the augmentation lag. *string* may take the values `aic`, `maic`, `bic`, `seq`, or `fix`. `aic` corresponds to the AIC, `maic` to the MAIC, `bic` to the BIC, `seq` to the sequential t test method, and `fix` to a user-specified lag length. The default is `mode(maic)`.

`residuals(`*string*`)` generates a variable containing the residual terms.

`noac` suppresses the ACF, the partial ACF, and the Ljung–Box Q statistics of the residuals. The default value of ACF lags is equal to `maxlag()`.

`noreg` suppresses the corresponding regression table, which is reported by default.

`gls` introduces the GLS detrending procedure proposed by Rodrigues and Taylor (2007) before applying the HEGY test. This option requires deterministic terms to be specified as `det(const)`, `det(trend)`, `det(seas)`, `det(strend)`, or `det(mult)`.

Execution of the command

When we use monthly data, (2.3) and (2.3) become

$$\Delta_{12} y^{\xi}_{12t+s} = \pi_0 y^{\xi}_{0,12t+s} + \pi_6 y^{\xi}_{6,12t+s} + \sum_{j=1}^{5} \left(\pi_{1j} y^{\xi}_{1j,12t+s} + \pi_{2j} y^{\xi}_{2j,12t+s} \right) + \sum_{j=1}^{k} d_j \Delta_S y^{\xi}_{12t+s-j} + e^{\xi}_{12t+s,k}$$

where

$$\begin{aligned} y^{\xi}_{0,12t+s} &= y^{\xi}_{12t+s-1} + y^{\xi}_{12t+s-2} + y^{\xi}_{12t+s-3} + y^{\xi}_{12t+s-4} + y^{\xi}_{12t+s-5} + y^{\xi}_{12t+s-6} \\ &\quad + y^{\xi}_{12t+s-7} + y^{\xi}_{12t+s-8} + y^{\xi}_{12t+s-9} + y^{\xi}_{12t+s-10} + y^{\xi}_{12t+s-11} + y^{\xi}_{12(t-1)+s} \\ y^{\xi}_{6,12t+s} &= -y^{\xi}_{12t+s-1} + y^{\xi}_{12t+s-2} - y^{\xi}_{12t+s-3} + y^{\xi}_{12t+s-4} - y^{\xi}_{12t+s-5} + y^{\xi}_{12t+s-6} \\ &\quad - y^{\xi}_{12t+s-7} + y^{\xi}_{12t+s-8} - y^{\xi}_{12t+s-9} + y^{\xi}_{12t+s-10} - y^{\xi}_{12t+s-11} + y^{\xi}_{12(t-1)+s} \\ x^{\xi}_{1j,12t+s} &= \sum_{q=0}^{11} \cos\left\{ (q+1)\frac{2\pi j}{S} \right\} y^{\xi}_{12t+s-q-1} \quad j = 1,2,3,4,5 \\ x^{\xi}_{2j,12t+s} &= -\sum_{q=0}^{11} \sin\left\{ (q+1)\frac{2\pi j}{S} \right\} y^{\xi}_{12t+s-q-1} \quad j = 1,2,3,4,5 \\ t &= 1,2,\ldots,N \quad s = -11,-10,-9,-8,\ldots,-1,0 \end{aligned}$$

In this case, the `hegy` command reports the following tests:

H_0	Roots	Frequency ω_j	$2\pi/\omega_j$ [a]	Test	Tail
$\pi_0 = 0$	One real root	0	∞	`t[0]`	Left tail
$\pi_6 = 0$	One real root	π	2	`t[Pi]`	Left tail
$\pi_{11} = \pi_{21} = 0$	Two complex conjugate roots	$\frac{\pi}{6}$	12	`F[Pi/6]`	Upper tail
$\pi_{12} = \pi_{22} = 0$	Two complex conjugate roots	$\frac{\pi}{3}$	6	`F[Pi/3]`	Upper tail
$\pi_{13} = \pi_{23} = 0$	Two complex conjugate roots	$\frac{\pi}{2}$	4	`F[Pi/2]`	Upper tail
$\pi_{14} = \pi_{24} = 0$	Two complex conjugate roots	$\frac{2\pi}{3}$	3	`F[2Pi/3]`	Upper tail
$\pi_{15} = \pi_{25} = 0$	Two complex conjugate roots	$\frac{5\pi}{6}$	$\frac{12}{5}$	`F[5Pi/6]`	Upper tail

[a] Number of periods to complete a full cycle

Additionally, the command reports the following joint tests:

H_0	Roots	Frequencies ω_j	Test	Tail
$\pi_{11} = \pi_{21} = \pi_{12} = \pi_{22} = \pi_{13} = \pi_{23} = \pi_{14} = \pi_{24} = \pi_{15} = \pi_{25} = \pi_6 = 0$	All seasonal roots	$\frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{5\pi}{6}, \pi$	`F[seas]`	Upper tail
$\pi_0 = \pi_{11} = \pi_{21} = \pi_{12} = \pi_{22} = \pi_{13} = \pi_{23} = \pi_{14} = \pi_{24} = \pi_{15} = \pi_{25} = \pi_6 = 0$	All roots in $(1 - L^{12})$	$0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{5\pi}{6}, \pi$	`F[All]`	Upper tail

Finally, when we use quarterly data, (2.3) and (2.3) become

$$\begin{aligned}\Delta_4 y^{\xi}_{4t+s} = \pi_0 y^{\xi}_{0,4t+s} + \pi_2 y^{\xi}_{2,4t+s} + \pi_{11} y^{\xi}_{1j,12t+s} + \pi_{21} y^{\xi}_{2j,12t+s} \\ + \sum_{j=1}^{k} d_j \Delta_s y^{\xi}_{12t+s-j} + e^{\xi}_{12t+s,k}\end{aligned}$$

where

$$\begin{aligned}
y^{\xi}_{0,4t+s} &= y^{\xi}_{4t+s-1} + y^{\xi}_{4t+s-2} + y^{\xi}_{4t+s-3} + y^{\xi}_{4(t-1)+s} \\
y^{\xi}_{2,4t+s} &= -y^{\xi}_{4t+s-1} + y^{\xi}_{4t+s-2} - y^{\xi}_{4t+s-3} + y^{\xi}_{4(t-1)+s} \\
x^{\xi}_{11,4t+s} &= -y^{\xi}_{4t+s-2} + y^{\xi}_{4(t-1)+s} \\
x^{\xi}_{2j,12t+s} &= -y^{\xi}_{4t+s-1} + y^{\xi}_{4t+s-3} \\
& t = 1, 2, \ldots, N \quad s = -3, -2, -1, 0
\end{aligned}$$

And the tests reported are the following:

H_0	Roots	Frequency ω_j	$2\pi/\omega_j$ [a]	Test	Tail
$\pi_0 = 0$	One real root	0	∞	`t[0]`	Left tail
$\pi_2 = 0$	One real root	π	2	`t[Pi]`	Left tail
$\pi_{11} = \pi_{21} = 0$	Two complex conjugate roots	$\frac{\pi}{2}$	4	`F[Pi/2]`	Upper tail

[a] Number of periods to complete a full cycle

H_0	Roots	Frequencies ω_j	Test	Tail
$\pi_{11} = \pi_{21} = \pi_2 = 0$	All seasonal roots	$\frac{\pi}{2}, \pi$	`F[seas]`	Upper tail
$\pi_0 = \pi_{11} = \pi_{21} = \pi_2 = 0$	All roots in $(1 - L^4)$	$0, \frac{\pi}{2}, \pi$	`F[All]`	Upper tail

3 Empirical application

In this section, we use the monthly data of airport passenger arrivals to Mallorca from the UK as an illustrative example. The island of Mallorca is a summer tourist destination with peak activity in July and August. As such, the data in hand show strong seasonal behavior. The results are based on a sample size that covers the period 1980M1 to 2014M4, which includes 412 observations.

Tourism activity is highly influenced by seasonality. The reasons are obvious, and arrivals to most "sun and sea" destinations demonstrate high peaks during the summer season and low troughs in the winter. Figure 1 below depicts the evolution of the analyzed time series in total [figure 1(a)] and in natural logs [figure 1(b)]. It is interesting to try to determine whether the seasonality observed in this time series shows only a deterministic nature or whether it is also caused by the presence of seasonal unit roots. That is, is the clear seasonal pattern observed in figure 1 only deterministic in nature, or is it also stochastic?

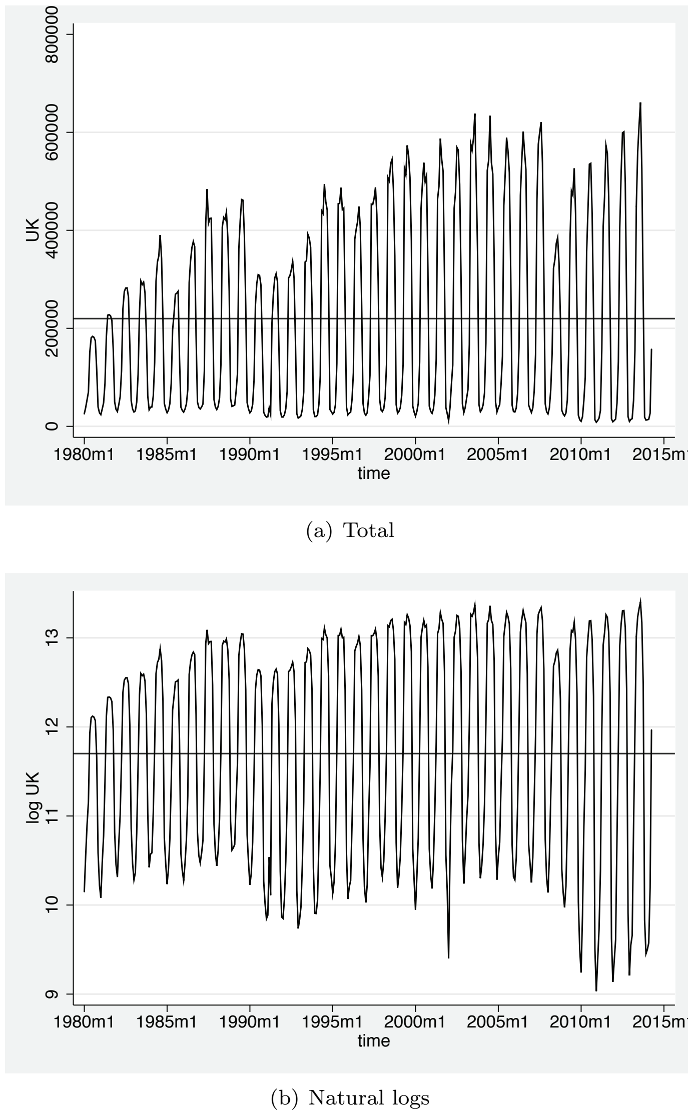

Figure 1. Total and natural log of UK arrivals to Palma de Mallorca, Spain

Based on figure 1, we decided to perform the analysis for our time series using natural logs of the data. Note that the evolution of figure 1(b) shows a less volatile evolution than the one depicted in figure 1(a); hence, to avoid negative effects on the performance of the seasonal unit-root tests stemming from the changing variance in our sample, we chose to use natural logs. Tables 2–5 report the results obtained from our `hegy` command using OLS and GLS detrending considering the case of seasonal intercepts with seasonal trends for the deterministic part of the process. We report the results that were obtained when the order of augmentation was based on the MAIC criteria and then with the sequential method. In both cases, we use the default option to determine the maximum lag length with which to start; that is, ($k_{\max} = \text{Int}\lfloor 12(T/100)^{1/4}\rfloor$). In our case, it is 17 lags.

Figure 2 depicts the evolution of the periodogram of the time series in natural logs. Note that we observe important peaks associated with 0, 1/12, 1/6, 1/4, 1/3, 5/12, and 1/2 along the frequency axis of the figure. These correspond to frequencies (ω_j) 0, $\pi/6$, $\pi/3$, $\pi/2$, $2\pi/3$, $5\pi/6$, and π, which complete a full cycle after ∞, 12, 6, 4, 3, 12/5, and 2 months (periods $2\pi/\omega_j$), respectively. Based on figures 1 and 2, we decided to use seasonal intercepts and trends (case 5) because we observe clear seasonal behavior in our data and important peaks in the periodogram at seasonal frequencies that could be caused by seasonal trends.

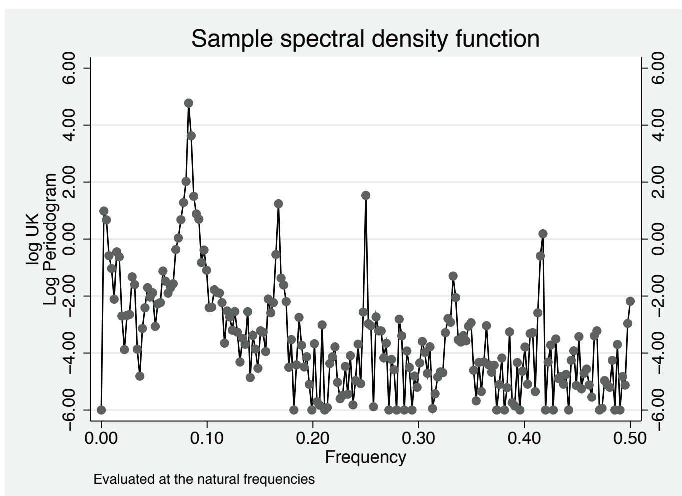

Figure 2. Periodogram of natural log of UK arrivals

Table 2 collects the results obtained with OLS detrending and with the order of augmentation determined with the MAIC criteria. Table 3 collects the results with GLS detrending and the MAIC criteria. Finally, tables 4 and 5 collect the results where the order of augmentation is determined with the sequential method and with OLS and GLS detrending, respectively.

Table 2. OLS MAIC results of HEGY seasonal unit-root test

```
. hegy luk, mode(maic) det(mult) noreg

HEGY Monthly seasonal unit root test for luk

Number of observations  :  399
Deterministic variables : Seasonal dummies and seasonal trends
Optimal lag selection method: Modified AIC
Lags tested: 17
Augmented by lags : 1

               Stat    1% critical  5% critical  10% critical
------------------------------------------------------------------------------
t[0]          -2.263     -3.896       -3.347       -3.065
t[Pi]         -4.643     -3.897       -3.347       -3.065

F[Pi/6]        9.696     11.798        9.356        8.206
F[Pi/3]       25.254     11.798        9.356        8.206
F[Pi/2]       15.996     11.798        9.356        8.206
F[2*Pi/3]     23.226     11.798        9.356        8.206
F[5*Pi/6]     19.944     11.798        9.356        8.206
F[All seas]   21.444      8.173        7.219        6.744
F[All]        20.136      8.076        7.160        6.703
  (output omitted)
```

Table 3. GLS MAIC results of HEGY seasonal unit-root test

```
. hegy luk, mode(maic) det(mult) noreg gls

HEGY Monthly seasonal unit root test with GLS detrending for luk

Number of observations  :  399
Deterministic variables : Seasonal dummies and seasonal trends
Optimal lag selection method: Modified AIC
Lags tested: 17
Augmented by lags : 1

               Stat    1% critical  5% critical  10% critical
------------------------------------------------------------------------------
t[0]          -1.914     -3.691       -3.143       -2.865
t[Pi]         -4.672     -3.691       -3.143       -2.866

F[Pi/6]        8.858      9.740        7.578        6.583
F[Pi/3]       21.138      9.740        7.578        6.583
F[Pi/2]       13.275      9.740        7.578        6.583
F[2*Pi/3]     23.759      9.740        7.578        6.583
F[5*Pi/6]     19.775      9.740        7.578        6.583
F[All seas]   19.097      6.507        5.734        5.353
F[All]        17.836      6.455        5.714        5.348
  (output omitted)
```

Table 4. OLS sequential results of HEGY seasonal unit-root test

```
. hegy luk, mode(seq) det(mult) noreg

HEGY Monthly seasonal unit root test for luk

Number of observations  :  387
Deterministic variables : Seasonal dummies and seasonal trends
Optimal lag selection method: Sequential at 10% level
Lags tested: 17
Augmented by lags : 13

               Stat    1% critical  5% critical  10% critical
------------------------------------------------------------------------------
t[0]          -2.159     -3.895       -3.345       -3.063
t[Pi]         -3.299     -3.895       -3.345       -3.064

F[Pi/6]        3.921     11.786        9.344        8.194
F[Pi/3]       10.335     11.786        9.344        8.194
F[Pi/2]        8.166     11.786        9.344        8.194
F[2*Pi/3]     10.895     11.786        9.344        8.194
F[5*Pi/6]      5.804     11.786        9.344        8.194
F[All seas]    8.763      8.179        7.221        6.745
F[All]         8.414      8.083        7.163        6.704
```

(*output omitted*)

Table 5. GLS sequential results of HEGY seasonal unit-root test

```
. hegy luk, mode(seq) det(mult) noreg gls

HEGY Monthly seasonal unit root test with GLS detrending for luk

Number of observations  :  387
Deterministic variables : Seasonal dummies and seasonal trends
Optimal lag selection method: Sequential at 10% level
Lags tested: 17
Augmented by lags : 13

               Stat    1% critical  5% critical  10% critical
------------------------------------------------------------------------------
t[0]          -1.950     -3.697       -3.149       -2.872
t[Pi]         -3.313     -3.697       -3.149       -2.872

F[Pi/6]        3.441      9.769        7.603        6.606
F[Pi/3]        7.289      9.769        7.603        6.606
F[Pi/2]        5.913      9.769        7.603        6.606
F[2*Pi/3]     11.046      9.769        7.603        6.606
F[5*Pi/6]      5.885      9.769        7.603        6.606
F[All seas]    7.333      6.541        5.764        5.382
F[All]         7.045      6.490        5.745        5.378
```

(*output omitted*)

As mentioned in section 2, based on the proposal of Perron and Qu (2007) and del Barrio Castro, Osborn, and Taylor (2016), the order of augmentation for the OLS and GLS detrending results is based on the same OLS-detrended augmented HEGY regression. So the order of augmentation in tables 2–5 with the MAIC criteria is the same as it is for both OLS and GLS detrending.

Note that the order of augmentation obtained with the MAIC criteria is one lag, but with the sequential method, the order of augmentation is set to 13 lags. To decide whether one lag is enough to purge the presence of serial correlation in the residuals of the HEGY regression, we do not use the **hegy** command's **noac()** option; thus the command reports the ACF, partial ACF, and Ljung–Box Q statistic associated with the residuals of the augmented HEGY regression. These results can be found in table 6. The sample autocorrelation and partial autocorrelation coefficients do not show evidence in favor of the presence of serial correlation in the residuals. Looking at the Ljung–Box Q statistic, we see that we do not reject the null hypothesis of white noise behavior, because the p-values reported in the last column of table 6 clearly show that the null hypothesis is not rejected at the 1%, 5%, and 10% levels. Also note that in the case of say, 12 lags, the p-value associated with the Q statistic is 0.9601; hence, we need to work with a significance level of 96% to reject the null that the autocorrelation coefficients from lag 1 to 12 are equal to 0. So, we could conclude that one lag in the augmented HEGY regression is enough to deal with serial correlation. In the part that follows, we focus on the results obtained when using the MAIC criteria results (tables 2 and 3).

Table 6. Correlogram of residuals

Correlogram of residuals:

LAG	AC	PAC	Q	Prob>Q	-1 0 1 [Autocorrelation]	-1 0 1 [Partial Autocor]
1	-0.0248	-0.0252	.2474	0.6189		
2	-0.0067	-0.0072	.26574	0.8756		
3	0.0246	0.0254	.51098	0.9165		
4	0.0513	0.0543	1.5775	0.8128		
5	0.0228	0.0263	1.7887	0.8775		
6	0.0448	0.0476	2.6051	0.8565		
7	0.0492	0.0511	3.5917	0.8254		
8	0.0176	0.0177	3.718	0.8816		
9	0.0156	0.0130	3.818	0.9230		
10	0.0124	0.0063	3.8817	0.9525		
11	0.0462	0.0411	4.7627	0.9421		
12	-0.0204	-0.0253	4.9345	0.9601		
13	0.0543	0.0500	6.156	0.9403		
14	0.0083	0.0033	6.1844	0.9616		
15	0.0019	-0.0035	6.1858	0.9765		
16	-0.0846	-0.0948	9.1731	0.9061		
17	0.0183	0.0057	9.314	0.9299		

Based on tables 2 and 3, one can check that the null hypothesis of the presence of a unit root at frequency zero is not rejected at all significance levels. Note that the `t[0]` statistic is a left-tail test and that in both cases (OLS and GLS detrending), the reported statistics are less negative than the critical values associated with the 1%, 5%, and 10% levels. Therefore, the statistics do not fall in the rejection region of the null hypothesis. And clearly, the null hypothesis of a unit root at frequency zero is not rejected. In the case of the `t[Pi]` test (the other left-tail test), the situation is the opposite. The test statistics with OLS and GLS detrending are smaller than the critical values; hence, we clearly reject the null of the presence of a unit root at frequency π. All the remaining tests (`F[Pi/6]` to `F[All]`) are upper-tail tests, and as such, the null hypothesis is rejected when the test statistic is bigger than the critical value. For the remaining frequencies, the null hypothesis is rejected at all significance levels except in the case of the two complex roots associated with frequency $\pi/6$ (test `F[Pi/6]`), where the null is rejected at the 5% and 10% levels. Hence, we find evidence of one unit root only at frequency zero, associated with pure trend behavior (oscillations that need infinity periods to complete a cycle), and the two complex conjugate unit roots at frequency $\pi/6$, associated with annual seasonality (oscillations that need 12 months to complete a full cycle). Therefore, based on the augmented HEGY procedure when the lag length is determined by the MAIC criteria, the UK tourist arrivals to Mallorca have nonstationary behavior associated with the zero frequency and with the frequency (associated with seasonal oscillations that complete a full cycle after 12 months).

4 Conclusions

In this article, we presented the `hegy` command. This command allows for the HEGY procedure to be used to test for the presence of seasonal unit roots in quarterly and monthly data. The order of augmentation of the HEGY regression can be determined using the MAIC criteria (as was recently applied to the case of seasonal unit roots by del Barrio Castro, Osborn, and Taylor [2016]), the sequential method proposed by Hall (1994) and Ng and Perron (1995), or the AIC and BIC criteria. In terms of the deterministic part of the process, we implemented the relevant cases of seasonal intercepts, seasonal intercepts with zero-frequency trend, and seasonal intercepts with trends. In addition to the usual OLS detrending, one can also use GLS detrending, as was recently proposed by Rodrigues and Taylor (2007). Finally, we reported 1%, 5%, and 10% critical values by means of the response surfaces implemented in del Barrio Castro, Bodnar, and Sansó (2015). We illustrated the command with an empirical application using monthly UK airport arrivals to Mallorca.

5 Acknowledgments

Tomás del Barrio Castro and Andreu Sansó acknowledge financial support from Spanish Ministerio de Educación, Cultura y Deporte under grant ECO2014-58991-C3-3-R. The authors also gratefully acknowledge valuable comments from an anonymous referee.

6 References

Baum, C. F., and R. Sperling. 2001. hegy4: Stata module to compute Hylleberg et al. seasonal unit root test. Statistical Software Components S416502, Department of Economics, Boston College. https://ideas.repec.org/c/boc/bocode/s416502.html.

Burridge, P., and A. M. R. Taylor. 2001. On the properties of regression-based tests for seasonal unit roots in the presence of higher-order serial correlation. *Journal of Business and Economic Statistics* 19: 374–379.

del Barrio Castro, T., A. Bodnar, and A. Sansó. 2015. Numerical distribution functions for seasonal unit root tests with OLS and GLS detrending. Working Paper 73, DEA. http://dea.uib.es/digitalAssets/353/353054_w73.pdf.

del Barrio Castro, T., and D. R. Osborn. 2011. HEGY tests in the presence of moving averages. *Oxford Bulletin of Economics and Statistics* 73: 691–704.

del Barrio Castro, T., D. R. Osborn, and A. M. R. Taylor. 2012. On augmented HEGY tests for seasonal unit roots. *Econometric Theory* 28: 1121–1143.

———. 2016. The performance of lag selection and detrending methods for HEGY seasonal unit root tests. *Econometric Reviews* 35: 122–168.

Depalo, D. 2009. A seasonal unit-root test with Stata. *Stata Journal* 9: 422–438.

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

Elliott, G., T. J. Rothenberg, and J. H. Stock. 1996. Efficient tests for an autoregressive unit root. *Econometrica* 64: 813–836.

Engle, R. F., C. W. J. Granger, S. Hylleberg, and H. S. Lee. 1993. Seasonal cointegration: The Japanese consumption function. *Journal of Econometrics* 55: 275–298.

Ghysels, E., H. S. Lee, and J. Noh. 1994. Testing for unit roots in seasonal time series: Some theoretical extensions and a Monte Carlo investigation. *Journal of Econometrics* 62: 415–442.

Gregoir, S. 2006. Efficient tests for the presence of a pair of complex conjugate unit roots in real time series. *Journal of Econometrics* 130: 45–100.

Hall, A. 1994. Testing for a unit root in time series with pretest data-based model selection. *Journal of Business and Economic Statistics* 12: 461–470.

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

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

Maravall, A. 1993. Stochastic linear trends: Models and estimators. *Journal of Econometrics* 56: 5–37.

Ng, S., and P. Perron. 1995. Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association* 90: 268–281.

———. 2001. Lag length selection and the construction of unit root tests with good size and power. *Econometrica* 69: 1519–1554.

Osborn, D. R., A. P. L. Chui, J. P. Smith, and C. R. Birchenhall. 1988. Seasonality and the order of integration for consumption. *Oxford Bulletin of Economics and Statistics* 50: 361–377.

Perron, P., and Z. Qu. 2007. A simple modification to improve the finite sample properties of Ng and Perron's unit root tests. *Economics Letters* 94: 12–19.

Rodrigues, P. M. M., and A. M. R. Taylor. 2004. Asymptotic distributions for regression-based seasonal unit root test statistics in a near-integrated model. *Econometric Theory* 20: 645–670.

———. 2007. Efficient tests of the seasonal unit root hypothesis. *Journal of Econometrics* 141: 548–573.

Smith, R. J., and A. M. R. Taylor. 1998. Additional critical values and asymptotic representations for seasonal unit root tests. *Journal of Econometrics* 85: 269–288.

Smith, R. J., A. M. R. Taylor, and T. del Barrio Castro. 2009. Regression-based seasonal unit root tests. *Econometric Theory* 25: 527–560.

About the authors

Tomás del Barrio Castro is a senior lecturer in the Applied Economics Department at the University of the Balearic Islands. His main research field is time-series econometrics. He has published in *Econometric Theory*, *Oxford Bulletin of Economics and Statistics*, *The Econometrics Journal*, *Journal of Time Series Analysis*, *Journal of Time Series Econometrics*, and *Economic Letters*. In 2014, he won the "Multa Scripsit" Econometric Theory Award.

Andrii Bodnar is a PhD student in the Master and PhD program in Tourism and Environmental Economics at the University of the Balearic Islands.

Andreu Sansó is a professor in the Applied Economics Department at the University of the Balearic Islands. His main research field is time-series econometrics. He has published in *Journal of Econometrics*, *Journal of Business and Economics Statistics*, *Oxford Bulletin of Economics and Statistics*, *Econometric Theory*, *The Econometrics Journal*, and *Economic Letters*.