Implementing the Leybourne–Taylor test for seasonal unit roots in Stata

Christopher F Baum Jesús Otero

Stata London Users' Meeting, September 2018

Introduction

Testing for unit roots, or nonstationary behavior, in economic time series has been a prominent component of time series econometrics since Granger and Newbold (*J. Econometrics*, 1974) introduced the concept of spurious regressions, and Nelson and Plosser (*J. Monetary Econ.*, 1982) presented evidence of its relevance for a large set of macroeconomic series.

The underlying concern in determining whether a time series exhibits stationary ($I(0)$) behavior or nonstationary, unit root ($I(1)$) behavior can be expressed in terms of deterministic versus stochastic trends.

Introduction

Testing for unit roots, or nonstationary behavior, in economic time series has been a prominent component of time series econometrics since Granger and Newbold (*J. Econometrics*, 1974) introduced the concept of spurious regressions, and Nelson and Plosser (*J. Monetary Econ.*, 1982) presented evidence of its relevance for a large set of macroeconomic series.

The underlying concept in determining whether a time series exhibits stationary ($I(0)$) behavior or nonstationary, unit root ($I(1)$) behavior can be expressed in terms of deterministic versus stochastic trends.

Consider the time series model

$$y_t = \alpha + \rho y_{t-1} + \gamma t + \epsilon_t \tag{1}$$

Depending on the value of ρ, this could be a model of a stochastic and a deterministic trend. If $\rho = 1$, the process contains a stochastic trend.

If ρ lies within the unit circle, y could be rendered covariance stationary by detrending: that is, regressing y on trend t and saving the residuals from that regression as y^*. The y^* series will have a constant mean, and will no longer contain a trend.

Consider the time series model

$$y_t = \alpha + \rho y_{t-1} + \gamma t + \epsilon_t \quad (1)$$

Depending on the value of ρ, this could be a model of a stochastic and a deterministic trend. If $\rho = 1$, the process contains a stochastic trend.

If ρ lies within the unit circle, y could be rendered covariance stationary by detrending: that is, regressing y on trend t and saving the residuals from that regression as y^*. The y^* series will have a constant mean, and will no longer contain a trend.

However, if $\rho = 1$, the series contains a unit root, and that detrending regression will not remove a stochastic trend from the series. We can rewrite the model as

$$\Delta y_t = \alpha + \gamma t + \epsilon_t \tag{2}$$

In that case, if $\alpha = \gamma = 0$, the y series follows a pure random walk, and is therefore a nonstationary (or $I(1)$) process. If $\alpha \neq 0$, the series follows a random walk with drift. By definition, the level series contains a stochastic trend, which can only be removed by first differencing the y series.

If $\gamma \neq 0$, the y series follows a random walk with a quadratic trend. The proper transformation to remove the stochastic trend is the regression of Δy on t. The residuals from that series, Δy^*, will be stationary, and will no longer contain a trend.

However, if $\rho = 1$, the series contains a unit root, and that detrending regression will not remove a stochastic trend from the series. We can rewrite the model as

$$\Delta y_t = \alpha + \gamma t + \epsilon_t \tag{2}$$

In that case, if $\alpha = \gamma = 0$, the y series follows a pure random walk, and is therefore a nonstationary (or $I(1)$) process. If $\alpha \neq 0$, the series follows a random walk with drift. By definition, the level series contains a stochastic trend, which can only be removed by first differencing the y series.

If $\gamma \neq 0$, the y series follows a random walk with a quadratic trend. The proper transformation to remove the stochastic trend is the regression of Δy on t. The residuals from that series, Δy^*, will be stationary, and will no longer contain a trend.

On the other hand, if ρ lies within the unit circle, differencing the level equation will remove the constant α:

$$\Delta y_t = \rho \Delta y_{t-1} + \gamma + \Delta \epsilon_t \tag{3}$$

but the constant term in the differenced series is the trend coefficient in the level series, so that the trend has not been removed. Furthermore, if ϵ_t is *i.i.d.*, the $\Delta\epsilon_t$ process is now a first-order moving average ($MA(1)$), and its elements are no longer independent.

On the other hand, if ρ lies within the unit circle, differencing the level equation will remove the constant α:

$$\Delta y_t = \rho \Delta y_{t-1} + \gamma + \Delta \epsilon_t \tag{3}$$

but the constant term in the differenced series is the trend coefficient in the level series, so that the trend has not been removed. Furthermore, if ϵ_t is *i*.*i*.*d*., the $\Delta \epsilon_t$ process is now a first-order moving average ($MA(1)$), and its elements are no longer independent.

It follows, then, that our concern over the order of integration of a time series—whether it is $I(0)$ or $I(1)$—relates to how we should work with that series in order to render it covariance stationary.

- If it contains a deterministic trend, it should be detrended, as differencing will not remove the trend in the level series.
- If it contains a stochastic trend, it should be differenced, as detrending will not make the series stationary.

In order to use the series in an estimated model, we must be able to determine its order of integration.

It follows, then, that our concern over the order of integration of a time series—whether it is $I(0)$ or $I(1)$—relates to how we should work with that series in order to render it covariance stationary.

- If it contains a deterministic trend, it should be detrended, as differencing will not remove the trend in the level series.
- If it contains a stochastic trend, it should be differenced, as detrending will not make the series stationary.

In order to use the series in an estimated model, we must be able to determine its order of integration.

It follows, then, that our concern over the order of integration of a time series—whether it is $I(0)$ or $I(1)$—relates to how we should work with that series in order to render it covariance stationary.

- If it contains a deterministic trend, it should be detrended, as differencing will not remove the trend in the level series.
- If it contains a stochastic trend, it should be differenced, as detrending will not make the series stationary.

In order to use the series in an estimated model, we must be able to determine its order of integration.

It follows, then, that our concern over the order of integration of a time series—whether it is $I(0)$ or $I(1)$—relates to how we should work with that series in order to render it covariance stationary.

- If it contains a deterministic trend, it should be detrended, as differencing will not remove the trend in the level series.
- If it contains a stochastic trend, it should be differenced, as detrending will not make the series stationary.

In order to use the series in an estimated model, we must be able to determine its order of integration.

It follows, then, that our concern over the order of integration of a time series—whether it is $I(0)$ or $I(1)$—relates to how we should work with that series in order to render it covariance stationary.

- If it contains a deterministic trend, it should be detrended, as differencing will not remove the trend in the level series.
- If it contains a stochastic trend, it should be differenced, as detrending will not make the series stationary.

In order to use the series in an estimated model, we must be able to determine its order of integration.

This elementary discussion of stationarity considers how we might treat one common feature of a time series, the trend, as deterministic or stochastic.

The standard Unobserved Components decomposition of a time series (see [TS] **ucm**) specifies that the series contains four components: Trend, Seasonal, Cyclical, and Irregular.

Just as we may be concerned about the identification of a trend in the model as deterministic or stochastic, we may need to consider the possibility that a seasonal component of a time series may be either deterministic or stochastic: that is, the series may exhibit *seasonal unit roots*. We now turn to consideration of that feature of the series.

This elementary discussion of stationarity considers how we might treat one common feature of a time series, the trend, as deterministic or stochastic.

The standard Unobserved Components decomposition of a time series (see [TS] **ucm**) specifies that the series contains four components: Trend, Seasonal, Cyclical, and Irregular.

Just as we may be concerned about the identification of a trend in the model as deterministic or stochastic, we may need to consider the possibility that a seasonal component of a time series may be either deterministic or stochastic: that is, the series may exhibit *seasonal unit roots*. We now turn to consideration of that feature of the series.

This elementary discussion of stationarity considers how we might treat one common feature of a time series, the trend, as deterministic or stochastic.

The standard Unobserved Components decomposition of a time series (see [TS] **ucm**) specifies that the series contains four components: Trend, Seasonal, Cyclical, and Irregular.

Just as we may be concerned about the identification of a trend in the model as deterministic or stochastic, we may need to consider the possibility that a seasonal component of a time series may be either deterministic or stochastic: that is, the series may exhibit *seasonal unit roots*. We now turn to consideration of that feature of the series.

Seasonal unit root tests

A key contribution to the literature on seasonal unit root tests is the paper by Hylleberg, Engle, Granger, and Yoo (*J. Econometrics*, 1990), commonly known as HEGY. As del Barrio Castro et al. (*Stata J.*, 2016) point out, “The HEGY approach has become the most popular one to test for the presence of seasonal unit roots.”

If seasonality is considered as a deterministic component of a time series y, the series may be deseasonalized by regressing y on a set of seasonal indicator variables: 3 for quarterly data, or 11 for monthly data. If the y series is trending, a time trend can also be included. The residuals from this regression, $y*$, are a deseasonalized (or deseasonalized and detrended) series.

Seasonal unit root tests

A key contribution to the literature on seasonal unit root tests is the paper by Hylleberg, Engle, Granger, and Yoo (*J. Econometrics*, 1990), commonly known as HEGY. As del Barrio Castro et al. (*Stata J.*, 2016) point out, “The HEGY approach has become the most popular one to test for the presence of seasonal unit roots.”

If seasonality is considered as a deterministic component of a time series y, the series may be deseasonalized by regressing y on a set of seasonal indicator variables: 3 for quarterly data, or 11 for monthly data. If the y series is trending, a time trend can also be included. The residuals from this regression, $y*$, are a deseasonalized (or deseasonalized and detrended) series.

However, if the seasonality is considered stochastic rather than deterministic, a procedure such as HEGY must be used to test for seasonal unit roots.

As with other unit root tests, you must choose the deterministic components to be included in the HEGY regression. The authors recommend that the model contain a set of seasonal indicators and constant. A trend could also be included. Analogous to the augmented Dickey–Fuller test or the [TS] **dfgls** test of Elliott, Rothenberg, Stock, a set of lagged fourth differences of the series may be included in the HEGY regression.

However, if the seasonality is considered stochastic rather than deterministic, a procedure such as HEGY must be used to test for seasonal unit roots.

As with other unit root tests, you must choose the deterministic components to be included in the HEGY regression. The authors recommend that the model contain a set of seasonal indicators and constant. A trend could also be included. Analogous to the augmented Dickey–Fuller test or the [TS] **dfgls** test of Elliott, Rothenberg, Stock, a set of lagged fourth differences of the series may be included in the HEGY regression.

The HEGY model is based on a potentially infinite autoregression

$$\phi(B)x_t = \epsilon_t \tag{4}$$

where B is the backshift operator. All of the roots of $\phi(B) = 0$ lie outside the unit circle; some may be complex conjugate pairs.

For quarterly data, with four seasons per year, this may be written as

$$(1 - B^4)x_t = \epsilon_t \tag{5}$$

and the polynomial $(1 - B^4)$ can be expressed as

$$(1 - B^4) = (1 - B)(1 + b)(1 - iB)(1 + iB) \tag{6}$$

where in the presence of unit roots, the roots of $1, -1, i, -i$ correspond to the zero frequency, 1/2 cycle per quarter (2 cycles per year) and two instances of 1/4 cycle per quarter (one cycle per year).

Testing for seasonal unit roots

The HEGY testing strategy is then implemented by running a simple regression of $(1 - B^4)x_t$ on the lagged values of four terms which are combinations of the four lags of x:

$$
\begin{aligned}
y_{1t} &= x_{t-1} + x_{t-2} + x_{t-3} + x_{t-4} \\
y_{2t} &= -x_{t-1} + x_{t-2} - x_{t-3} + x_{t-4} \\
y_{3t} &= -y_{t-2} + y_{t-4} \\
y_{4t} &= -y_{t-1} + y_{t-3}
\end{aligned}
\tag{7}
$$

This regression can be augmented by deterministic components as well.

The regression coefficients in

$$(1 - B^4)x_t = \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-1} + \pi_4 y_{4,t-1} \tag{8}$$

can then be used to carry out the HEGY tests. To test for a root of 1 at the zero frequency, we can merely test $\pi_1 = 0$. To test for a root of -1 at the Nyquist frequency, we can test $\pi_2 = 0$. For the complex conjugate pair, we can do a joint test on π_3 and π_4.

There will be no seasonal unit roots if π_2 and either π_3 or π_4 are different from zero, implying that we need the rejection of a both a test for π_2 and a joint test for π_3 and π_4.

The regression coefficients in

$$(1 - B^4)x_t = \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-1} + \pi_4 y_{4,t-1} \tag{8}$$

can then be used to carry out the HEGY tests. To test for a root of 1 at the zero frequency, we can merely test $\pi_1 = 0$. To test for a root of -1 at the Nyquist frequency, we can test $\pi_2 = 0$. For the complex conjugate pair, we can do a joint test on π_3 and π_4.

There will be no seasonal unit roots if π_2 and either π_3 or π_4 are different from zero, implying that we need the rejection of a both a test for π_2 and a joint test for π_3 and π_4.

The first implementation of HEGY in Stata was provided by Baum and Sperling (SSC, 2001); their `hegy4` routine computed HEGY tests, as described above, for quarterly data.

Depalo provided the `sroot` routine for quarterly data in a 2009 *Stata Journal* article. Unit roots are allowed to be at seasonal frequencies rather than only at frequency zero. The null hypothesis is that the variable contains a unit root at that frequency, and the alternative is that the variable was generated by a stationary process. If the constant is excluded, the null defines a pure random walk. When a constant and trend are included, the null is a random walk with drift.

The first implementation of HEGY in Stata was provided by Baum and Sperling (SSC, 2001); their `hegy4` routine computed HEGY tests, as described above, for quarterly data.

Depalo provided the `sroot` routine for quarterly data in a 2009 *Stata Journal* article. Unit roots are allowed to be at seasonal frequencies rather than only at frequency zero. The null hypothesis is that the variable contains a unit root at that frequency, and the alternative is that the variable was generated by a stationary process. If the constant is excluded, the null defines a pure random walk. When a constant and trend are included, the null is a random walk with drift.

The most recent published implementation of the HEGY test for Stata was provided by del Barrio Castro, Bodnar and Sansó in a 2016 *Stata Journal* article. Their `hegy` routine, handles both quarterly and monthly seasonality and allows for detrending via OLS and GLS, analogous to [TS] **dfgls** for standard unit root tests as proposed by Rodrigues and Taylor (*J. Econometrics*, 2007).

In their 2016 article, del Barrio Castro et al. define the data generating process of a seasonal time series as

$$\begin{aligned} y_{St+s} &= \mu_{St+s} + x_{St+s} \\ \alpha(L)x_{St+s} &= u_{St+s} \end{aligned} \tag{9}$$

where S is the number of seasons (4 or 12), $s = (1 - S), \ldots, 0$ and $T = 1, \ldots N$, the number of years of data.

The time series may be decomposed into the deterministic part μ and the stochastic part x. $\alpha(L)$ is a polynomial of order S in the lag operator. The object of HEGY is to test for the presence of unit roots in that polynomial.

In their 2016 article, del Barrio Castro et al. define the data generating process of a seasonal time series as

$$\begin{aligned} y_{St+s} &= \mu_{St+s} + x_{St+s} \\ \alpha(L)x_{St+s} &= u_{St+s} \end{aligned} \tag{9}$$

where S is the number of seasons (4 or 12), $s = (1 - S), \ldots, 0$ and $T = 1, \ldots N$, the number of years of data.

The time series may be decomposed into the deterministic part μ and the stochastic part x. $\alpha(L)$ is a polynomial of order S in the lag operator. The object of HEGY is to test for the presence of unit roots in that polynomial.

The Leybourne–Taylor seasonal unit root test

The Leybourne–Taylor test, presented in *Journal of Time Series Analysis* (2003), leverages the HEGY test by employing both forward and reverse regressions to derive the test statistics.

This innovation parallels the development of Leybourne's (*OBES*, 1995) ADFmax unit root test. That test, which we implemented as the Stata routine `adfmaxur`, is described in our 2018 *Stata Journal* article.

The `adfmaxur` test involves running Dickey–Fuller regressions using forward and reverse realizations of the time series. According to Leybourne, this test exhibits greater power than the standard ADF test, so it is more likely to reject a false unit root hypothesis.

The Leybourne–Taylor seasonal unit root test

The Leybourne–Taylor test, presented in *Journal of Time Series Analysis* (2003), leverages the HEGY test by employing both forward and reverse regressions to derive the test statistics.

This innovation parallels the development of Leybourne's (*OBES*, 1995) ADFmax unit root test. That test, which we implemented as the Stata routine `adfmaxur`, is described in our 2018 *Stata Journal* article.

The `adfmaxur` test involves running Dickey–Fuller regressions using forward and reverse realizations of the time series. According to Leybourne, this test exhibits greater power than the standard ADF test, so it is more likely to reject a false unit root hypothesis.

The Leybourne–Taylor seasonal unit root test

The Leybourne–Taylor test, presented in *Journal of Time Series Analysis* (2003), leverages the HEGY test by employing both forward and reverse regressions to derive the test statistics.

This innovation parallels the development of Leybourne's (*OBES*, 1995) ADFmax unit root test. That test, which we implemented as the Stata routine `adfmaxur`, is described in our 2018 *Stata Journal* article.

The `adfmaxur` test involves running Dickey–Fuller regressions using forward and reverse realizations of the time series. According to Leybourne, this test exhibits greater power than the standard ADF test, so it is more likely to reject a false unit root hypothesis.

Leybourne and Taylor (2003) claim that a similar improvement in power of the HEGY test can be achieved by basing the seasonal unit root test statistics on forward and reverse realizations of the time series.

Their article presents finite-sample critical values for quarterly data as well as asymptotic critical values. They state that “Monte Carlo simulation of the finite-sample size and power properties of the new tests reveals that, overall, they perform rather better than extant tests of the seasonal unit root hypothesis.”

Leybourne and Taylor (2003) claim that a similar improvement in power of the HEGY test can be achieved by basing the seasonal unit root test statistics on forward and reverse realizations of the time series.

Their article presents finite-sample critical values for quarterly data as well as asymptotic critical values. They state that "Monte Carlo simulation of the finite-sample size and power properties of the new tests reveals that, overall, they perform rather better than extant tests of the seasonal unit root hypothesis."

In their formulation, Leybourne and Taylor consider a pure AR process for data containing S seasons:

$$\begin{aligned} \alpha(L)[x_t - \mu_t] &= u_t, \quad t = S+1, \ldots, ST \\ \mu_t &= \sum_{j=1}^{S} (\gamma_j D_{j,t} + \delta_j [D_{j,t}] t]) \end{aligned} \tag{10}$$

where $D_{j,t}, j = 1, \ldots, S$ are seasonal indicator variables, $\alpha(L) = 1 - \sum_{j=1}^{S} \alpha_j L^j$ and the error u_t is assumed to follow an $AR(p)$ process.

This defines three cases of interest:

1. Seasonal intercepts, no trend: $\delta_j = 0 \quad \forall j$
2. Seasonal intercepts, constant trend: $\delta_j = \delta \quad \forall j$
3. Seasonal intercepts and trends: γ_j, δ_j unrestricted $\forall j$.

This defines three cases of interest:

1. Seasonal intercepts, no trend: $\delta_j = 0 \quad \forall j$
2. Seasonal intercepts, constant trend: $\delta_j = \delta \quad \forall j$
3. Seasonal intercepts and trends: γ_j, δ_j unrestricted $\forall j$.

This defines three cases of interest:

1. Seasonal intercepts, no trend: $\delta_j = 0 \quad \forall j$
2. Seasonal intercepts, constant trend: $\delta_j = \delta \quad \forall j$
3. Seasonal intercepts and trends: γ_j, δ_j unrestricted $\forall j$.

As presented above, the HEGY tests involve the computation of both *t*-statistics for some of the estimated coefficients and *F*-statistics for joint tests on one or more pairs of coefficients.

In the Leybourne–Taylor strategy, where both forward and reverse regressions are estimated, the resulting test statistic are based on the *maximum t*-statistics and the *minimum F*-statistics from the forward and reverse regression estimates.

The *t* statistic rejects for large negative values of the statistic, analogous to Dickey–Fuller statistics, while the *F* statistic rejects for large positive values.

As presented above, the HEGY tests involve the computation of both t-statistics for some of the estimated coefficients and F-statistics for joint tests on one or more pairs of coefficients.

In the Leybourne–Taylor strategy, where both forward and reverse regressions are estimated, the resulting test statistic are based on the *maximum* t-statistics and the *minimum* F-statistics from the forward and reverse regression estimates.

The t statistic rejects for large negative values of the statistic, analogous to Dickey–Fuller statistics, while the F statistic rejects for large positive values.

As presented above, the HEGY tests involve the computation of both t-statistics for some of the estimated coefficients and F-statistics for joint tests on one or more pairs of coefficients.

In the Leybourne–Taylor strategy, where both forward and reverse regressions are estimated, the resulting test statistic are based on the *maximum* t-statistics and the *minimum* F-statistics from the forward and reverse regression estimates.

The t statistic rejects for large negative values of the statistic, analogous to Dickey–Fuller statistics, while the F statistic rejects for large positive values.

Dealing with lag order selection

Seasonal unit root tests are customarily augmented with lags of the dependent variable, just as Dickey–Fuller or [TS] **dfgls** tests are customarily augmented.

We consider that the selection of lag order may be an important element of unit root testing, as the method used for lag order selection may affect the finite-sample critical values.

Dealing with lag order selection

Seasonal unit root tests are customarily augmented with lags of the dependent variable, just as Dickey–Fuller or [TS] **dfgls** tests are customarily augmented.

We consider that the selection of lag order may be an important element of unit root testing, as the method used for lag order selection may affect the finite-sample critical values.

We have presented evidence of this sensitivity for conventional unit root tests: the Elliot–Rothenberg–Stock [TS] **dfgls** test, using our `ersur` routine (*Stata J.*, 2017) and the Leybourne ADFmax test, using our `adfmaxur` routine (*Stata J.*,, 2018).

In the current study, we extend the analysis of lag order selection and how it affects inference to the seasonal unit root test of Leybourne and Taylor, complementing the efforts of del Barrio Castro et al.. in their `hegy` routine.

We have presented evidence of this sensitivity for conventional unit root tests: the Elliot–Rothenberg–Stock [TS] **dfgls** test, using our `ersur` routine (*Stata J.*, 2017) and the Leybourne ADFmax test, using our `adfmaxur` routine (*Stata J.*,, 2018).

In the current study, we extend the analysis of lag order selection and how it affects inference to the seasonal unit root test of Leybourne and Taylor, complementing the efforts of del Barrio Castro et al.. in their `hegy` routine.

The Monte Carlo experiment

- Response surface estimates are generated with a Monte Carlo simulation experiment similar to that used by Otero and Smith (*Comp.Stat.*, 2012).
- Assume that y_t is a unit root process with standard Normal errors and a sample of $T + 4$ observations, with T ranging from 40 to 5000 (37 sample sizes).
- The number of lagged differences of y_t, p, varies between 0 and 8.
- The experiment defines 296 combinations of T and p for each of the three specifications of the test, and involves 100,000 Monte Carlo replications.

The Monte Carlo experiment

- Response surface estimates are generated with a Monte Carlo simulation experiment similar to that used by Otero and Smith (*Comp.Stat.*, 2012).
- Assume that y_t is a unit root process with standard Normal errors and a sample of $T + 4$ observations, with T ranging from 40 to 5000 (37 sample sizes).
- The number of lagged differences of y_t, p, varies between 0 and 8.
- The experiment defines 296 combinations of T and p for each of the three specifications of the test, and involves 100,000 Monte Carlo replications.

The Monte Carlo experiment

- Response surface estimates are generated with a Monte Carlo simulation experiment similar to that used by Otero and Smith (*Comp.Stat.*, 2012).
- Assume that y_t is a unit root process with standard Normal errors and a sample of $T + 4$ observations, with T ranging from 40 to 5000 (37 sample sizes).
- The number of lagged differences of y_t, p, varies between 0 and 8.
- The experiment defines 296 combinations of T and p for each of the three specifications of the test, and involves 100,000 Monte Carlo replications.

The Monte Carlo experiment

- Response surface estimates are generated with a Monte Carlo simulation experiment similar to that used by Otero and Smith (*Comp.Stat.*, 2012).
- Assume that y_t is a unit root process with standard Normal errors and a sample of $T + 4$ observations, with T ranging from 40 to 5000 (37 sample sizes).
- The number of lagged differences of y_t, p, varies between 0 and 8.
- The experiment defines 296 combinations of T and p for each of the three specifications of the test, and involves 100,000 Monte Carlo replications.

- Critical values are computed for each of 221 significance levels: $0.0001, 0.0002, \ldots, 0.9998, 0.9999$ for both the detrended and demeaned cases.
- Response surface models are then estimated for each significance level.
- The functional form of these models follows MacKinnon (1991), Cheung and Lai (*JBES*, 1995; *OBES*, 1995) and Harvey and Van Dijk (*CSDA*, 2006), in which the critical values are regressed on an intercept term and power functions of $\frac{1}{T}$ and $\frac{p}{T}$.

- Critical values are computed for each of 221 significance levels: $0.0001, 0.0002, \ldots, 0.9998, 0.9999$ for both the detrended and demeaned cases.
- Response surface models are then estimated for each significance level.
- The functional form of these models follows MacKinnon (1991), Cheung and Lai (*JBES*, 1995; *OBES*, 1995) and Harvey and Van Dijk (*CSDA*, 2006), in which the critical values are regressed on an intercept term and power functions of $\frac{1}{T}$ and $\frac{p}{T}$.

- Critical values are computed for each of 221 significance levels: $0.0001, 0.0002, \ldots, 0.9998, 0.9999$ for both the detrended and demeaned cases.
- Response surface models are then estimated for each significance level.
- The functional form of these models follows MacKinnon (1991), Cheung and Lai (*JBES*, 1995; *OBES*, 1995) and Harvey and Van Dijk (*CSDA*, 2006), in which the critical values are regressed on an intercept term and power functions of $\frac{1}{T}$ and $\frac{p}{T}$.

The chosen functional form is:

$$CV^{l}_{T,p} = \theta^{l}_{\infty} + \sum_{i=1}^{4} \theta^{l}_{i} \left(\frac{1}{T}\right)^{i} + \sum_{i=1}^{4} \phi^{l}_{i} \left(\frac{p^{i}}{T}\right) + \epsilon^{l}, \tag{11}$$

where $CV^{l}_{T,p}$ is the critical value estimate at significance level l, T refers to the number of observations on Δy_t, which is one less than the total number of available observations, and p is the number of lags of the dependent variable that are included to account for residual serial correlation.

Note that the larger the number of observations, T, the weaker is the dependence of the critical values on the lag truncation p. Also, as $T \rightarrow \infty$ the intercept term, θ^{l}_{∞}, can be thought of as an estimate of the corresponding asymptotic critical value.

The chosen functional form is:

$$CV^{l}_{T,p} = \theta^{l}_{\infty} + \sum_{i=1}^{4} \theta^{l}_{i} \left(\frac{1}{T} \right)^{i} + \sum_{i=1}^{4} \phi^{l}_{i} \left(\frac{p^{i}}{T} \right) + \epsilon^{l}, \tag{11}$$

where $CV^{l}_{T,p}$ is the critical value estimate at significance level l, T refers to the number of observations on Δy_t, which is one less than the total number of available observations, and p is the number of lags of the dependent variable that are included to account for residual serial correlation.

Note that the larger the number of observations, T, the weaker is the dependence of the critical values on the lag truncation p. Also, as $T \to \infty$ the intercept term, θ^{l}_{∞}, can be thought of as an estimate of the corresponding asymptotic critical value.

Dealing with endogenous lag order

- The tabulated response surface values can be used to obtain critical values for any given T and fixed lag order.
- In practice the lag order, p, is rarely fixed by the user, but rather chosen endogenously using a data-dependent procedure such as the information criteria of Akaike and Schwarz, AIC and SIC respectively.
- The optimal number of lags is determined by varying p, the number of augmented lags of the dependent variable, between $p_{\max}$ and 0 lags, and choosing the best model according to the information criterion that is being employed.

Dealing with endogenous lag order

- The tabulated response surface values can be used to obtain critical values for any given T and fixed lag order.
- In practice the lag order, p, is rarely fixed by the user, but rather chosen endogenously using a data-dependent procedure such as the information criteria of Akaike and Schwarz, AIC and SIC respectively.
- The optimal number of lags is determined by varying p, the number of augmented lags of the dependent variable, between $p_{\max}$ and 0 lags, and choosing the best model according to the information criterion that is being employed.

Dealing with endogenous lag order

- The tabulated response surface values can be used to obtain critical values for any given T and fixed lag order.
- In practice the lag order, p, is rarely fixed by the user, but rather chosen endogenously using a data-dependent procedure such as the information criteria of Akaike and Schwarz, AIC and SIC respectively.
- The optimal number of lags is determined by varying p, the number of augmented lags of the dependent variable, between $p_{\max}$ and 0 lags, and choosing the best model according to the information criterion that is being employed.

- We also consider another data-dependent procedure to optimally select p, which is commonly referred to as the general-to-specific (GTS) algorithm of Campbell and Perron (1991), Hall (*JBES*, 1994) and Ng and Perron (*JASA*, 1995).
- This algorithm starts by setting some upper bound on p, let us say $p_{\max}$, where $p_{\max} = 0, 1, 2, ..., 8$, estimating the equation with $p = p_{\max}$, and testing the statistical significance of $b_{p_{\max}}$.
- If this coefficient is statistically significant, for instance using a significance level of 5% (denoted GTS_5) or 10% (denoted GTS_{10}), one chooses $p = p_{\max}$. Otherwise, the order of the estimated autoregression is reduced by one until the coefficient on the last included lag is statistically different from zero.

- We also consider another data-dependent procedure to optimally select p, which is commonly referred to as the general-to-specific (GTS) algorithm of Campbell and Perron (1991), Hall (*JBES*, 1994) and Ng and Perron (*JASA*, 1995).
- This algorithm starts by setting some upper bound on p, let us say $p_{\max}$, where $p_{\max} = 0, 1, 2, ..., 8$, estimating the equation with $p = p_{\max}$, and testing the statistical significance of $b_{p_{\max}}$.
- If this coefficient is statistically significant, for instance using a significance level of 5% (denoted GTS_5) or 10% (denoted GTS_{10}), one chooses $p = p_{\max}$. Otherwise, the order of the estimated autoregression is reduced by one until the coefficient on the last included lag is statistically different from zero.

- We also consider another data-dependent procedure to optimally select p, which is commonly referred to as the general-to-specific (GTS) algorithm of Campbell and Perron (1991), Hall (*JBES*, 1994) and Ng and Perron (*JASA*, 1995).
- This algorithm starts by setting some upper bound on p, let us say $p_{\max}$, where $p_{\max} = 0, 1, 2, ..., 8$, estimating the equation with $p = p_{\max}$, and testing the statistical significance of $b_{p_{\max}}$.
- If this coefficient is statistically significant, for instance using a significance level of 5% (denoted GTS_5) or 10% (denoted GTS_{10}), one chooses $p = p_{\max}$. Otherwise, the order of the estimated autoregression is reduced by one until the coefficient on the last included lag is statistically different from zero.

To obtain *p*-values of the Leybourne–Taylor test statistic, we follow MacKinnon (*JBES*, 1994; *App. Econometrics*, 1996) by estimating the regression:

$$\Phi^{-1}(l) = \gamma_0^l + \gamma_1^l \widehat{CV^l} + \gamma_2^l \left(\widehat{CV^l}\right)^2 + \upsilon^l, \tag{12}$$

where Φ^{-1} is the inverse of the cumulative standard normal distribution at each of the 221 quantiles, and $\widehat{CV^l}$ is the fitted value from (11) at the l quantile. Following Harvey and van Dijk (*CSDA*, 2006), equation (12) is estimated by OLS using 15 observations. Approximate *p*-values of the Leybourne–Taylor test statistic can then be obtained as:

$$pvalue = \Phi\left(\widehat{\gamma}_0^l + \widehat{\gamma}_1^l ERS(p) + \widehat{\gamma}_2^l (ERS(p))^2\right), \tag{13}$$

where $\widehat{\gamma}_0^l$, $\widehat{\gamma}_1^l$ and $\widehat{\gamma}_2^l$ are the OLS parameter estimates from (12).

Implementation issues

- The result of the Monte Carlo experiment is a 221 $\times$ 136 matrix, with the rows indexed by the quantile and the columns representing combinations of lag order selection method, model specification, and sample size.
- Although it would be possible to include this as a Stata matrix coded into the ado-file, that appeared to be a very inelegant solution.
- Accordingly, the matrix was stored as a binary matrix using Mata's `fputmatrix()` function, and the ado-file uses Mata's `fopen()` and `fgetmatrix()` functions to retrieve it from the `PLUS` directory.

Implementation issues

- The result of the Monte Carlo experiment is a 221 $\times$ 136 matrix, with the rows indexed by the quantile and the columns representing combinations of lag order selection method, model specification, and sample size.
- Although it would be possible to include this as a Stata matrix coded into the ado-file, that appeared to be a very inelegant solution.
- Accordingly, the matrix was stored as a binary matrix using Mata's `fputmatrix()` function, and the ado-file uses Mata's `fopen()` and `fgetmatrix()` functions to retrieve it from the `PLUS` directory.

Implementation issues

- The result of the Monte Carlo experiment is a 221×136 matrix, with the rows indexed by the quantile and the columns representing combinations of lag order selection method, model specification, and sample size.
- Although it would be possible to include this as a Stata matrix coded into the ado-file, that appeared to be a very inelegant solution.
- Accordingly, the matrix was stored as a binary matrix using Mata's `fputmatrix()` function, and the ado-file uses Mata's `fopen()` and `fgetmatrix()` functions to retrieve it from the `PLUS` directory.

- The table lookup routine, and the regression referenced in equation 13, is implemented as a Mata function contained in `ltsur.ado`.
- Our `ltsur` routine is still under development and validation. At present, it only provides the Leybourne–Taylor seasonal unit root test for quarterly data. Just as HEGY can be applied to monthly data, as available in del Barrio Castro et al., our analysis can be extended to analysis of monthly seasonals.

- The table lookup routine, and the regression referenced in equation 13, is implemented as a Mata function contained in `ltsur.ado`.
- Our `ltsur` routine is still under development and validation. At present, it only provides the Leybourne–Taylor seasonal unit root test for quarterly data. Just as HEGY can be applied to monthly data, as available in del Barrio Castro et al., our analysis can be extended to analysis of monthly seasonals.

Summary discussion

- The Leybourne–Taylor seasonal unit root test is claimed to provide improvements in size and power over the HEGY test.
- As we demonstrated for the Elliott–Rothenberg–Stock [TS] **dfgls** test and the Leybourne ADFmax test, the choice of lag order selection method can have a considerable impact on significance levels of the test.
- Our implementation of the Leybourne–Taylor test, soon to be available from SSC, will make it possible to contrast its findings with that of the HEGY test.

Summary discussion

- The Leybourne–Taylor seasonal unit root test is claimed to provide improvements in size and power over the HEGY test.
- As we demonstrated for the Elliott–Rothenberg–Stock [TS] **dfgls** test and the Leybourne ADFmax test, the choice of lag order selection method can have a considerable impact on significance levels of the test.
- Our implementation of the Leybourne–Taylor test, soon to be available from SSC, will make it possible to contrast its findings with that of the HEGY test.

Summary discussion

- The Leybourne–Taylor seasonal unit root test is claimed to provide improvements in size and power over the HEGY test.
- As we demonstrated for the Elliott–Rothenberg–Stock [TS] **dfgls** test and the Leybourne ADFmax test, the choice of lag order selection method can have a considerable impact on significance levels of the test.
- Our implementation of the Leybourne–Taylor test, soon to be available from SSC, will make it possible to contrast its findings with that of the HEGY test.