EC821: Time Series Econometrics, Spring 2003

Notes Section 5

Unit root tests

Given the distinction between trend-stationary and unit root processes, it would seem to be very important to be able to determine whether a particular timeseries which, for instance, generally increases in value is being driven by some underlying trend, or whether its evolution reflects a unit root in its data generating process. Those who study macroeconomic phenomena will want to know whether economic recessions have permanent consequences for the level of future GDP (as they would if GDP exhibits a unit root), or whether they are merely deviations from a trend rate of growth, temporary downturns that will be offset by the following recovery. Those who are concerned with the stock market want to know whether stock prices really do follow a random walk—i.e. exhibit unit root behavior—rather than some complicated combination of trends and cycles. If stock prices' behavior reflect a unit root, then "technical analysis" or "charting" is no more useful than astrology. On the other hand, if there are no unit roots in stock prices, all of the effort applied by stock analysts to studying the behavior of these series should have a reward.

This concern has given rise to a battery of unit root tests: statistical procedures that are designed to render a verdict as to whether a given sample of timeseries data appears to imply the presence of a unit root in that timeseries, or whether the series may be considered stationary. In terms of our prior terminology, we are trying to discern whether the series exhibits $I(1)$ (unit root) or $I(0)$ (stationary) behavior. It turns out that this is a fairly difficult problem, from a statistical perspective. It might appear sufficient to merely estimate an equation such as $y_t = \phi_1 y_{t-1} + \epsilon_t$, modified to the form

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

using the available sample of data, and test the null hypothesis that $\gamma = (\phi_1 - 1) = 1$. For various reasons, that turns out to be problematic, in the sense that the distribution of the test statistic is nonstandard under that null. The t test for $\gamma = 0$ does not have a t distribution under the null hypothesis; even as $T \to \infty$, the distribution of this t statistic will not be $N(0, 1)$. Under the alternative hypothesis, the test statistic is well behaved, but under the null—the point of interest—it follows the "Dickey–Fuller" distribution rather than the Normal or t. The critical points on the D–F distribution, as established by simulation, are considerably larger than those of the equivalent t; whereas a value of –1.645 would be on the borderline of rejection at the 95% level for a one–tailed t test, the D-F critical value would be –1.961 for $T = 1000$.

Of course, the model (1) may not be the appropriate special case of the autoregressive distributed lag model; we may want to allow for an additional term which would become a constant term in a stable autoregressive process, or a drift term in a random walk process. Otherwise, we are specifying a stable autoregressive process with a zero mean under the alternative hypothesis, which may not be sensible (unless the series has already been demeaned). With that modification, we would test

$$\Delta y_t = \mu + \gamma y_{t-1} + \epsilon_t \tag{2}$$

which would then allow both a test for a unit root $(\gamma = 0)$ and a joint test for a white noise process (an F test for $\gamma = 0$ and $\mu = 0$). Note that the critical values for the t test are not the same as those that would be used in (1); for instance, the D–F critical value for $T = 1000$ in this test is –2.864. One must also note that this model would not be appropriate if there was an obvious trend in the series, since the model under the alternative has no mechanism to generate such a trend (as the RW–with–drift model does under the null).

The most general form of the standard D–F test allows for both a constant in the relationship and a deterministic trend:

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

Such a model will allow for both a nonzero mean for y (with $\mu \neq 0)$ and trending behavior (with $\beta \neq 0)$ under the alternative hypothesis, where $\gamma < 0$. The most likely null hypothesis is then that of a RW–with–drift, so that under H_0 $\gamma = 0$ and $\beta = 0$ (no deterministic trend). This null could be rejected for three reasons: (a) there could be no unit root, but a deterministic trend; (b) there could be a unit root, but with a deterministic trend; or (c) there might be no unit root nor deterministic trend. The most general alternative is (a), for which an F test is required (since two restrictions on the parameter vector are implied under the null). The F statistic is calculated in the normal way, but the distribution is again nonstandard, and tabulated values for the "D–F F distribution" must be consulted. More commonly, we consider a t test on γ; once again, the critical values are specific to model (3a). For instance, the D–F critical value for $T = 1000$ in this test is –3.408: larger yet than the critical values in the constant–only model, which in turn exceed those for the original white noise model.

Any of the forms of this test presume the existence of white noise errors in the regression. If that is implausible, the test will lose significant power. To cope with this issue, any of the 'Dickey–Fuller" tests in practice are usually employed as the "augmented Dickey–Fuller" test, or ADF test, in which a number of lags of the dependent variable are added to the regression to whiten the errors:

$$\Delta y_t = \mu + \gamma y_{t-1} + \vartheta_1 \Delta y_{t-1} + \vartheta_2 \Delta y_{t-2} + \ldots + \beta t + \epsilon_t \quad (4)$$

In this formulation, we consider an $AR(p)$ model as the baseline model, rather than the $AR(1)$ model of the simple Dickey–Fuller framework. The choice of appropriate lag length is likely to depend on the frequency of the data; a general–to–specific strategy (analogous to the Ng–Perron sequential t procedure) or an information criterion may also be used. We will discuss the use of a modified AIC below.

Phillips–Perron tests

The augmentation of the original D–F regression with lags of the dependent variable is motivated by the need to generate iid errors in that model, since an OLS estimator of the covariance matrix is being employed. An alternative strategy for allowing for errors that are not iid is that of Phillips (1987) and Phillips and Perron (1988), known as the Phillips–Perron (PP) unit root test. The PP test deals with potential serial correlation in the errors by employing a correction factor that estimates the long–run variance of the error process with a variant of the Newey–West formula. Like the ADF test, use of the PP test requires specification of a lag order; in the latter case, the lag order designates the number of lags to be included in the long–run variance estimate. The PP test allows for dependence among disturbances of either AR or MA form, but have been shown to exhibit serious size distortions in the presence of negative autocorrelations. In principle, the PP tests should be more powerful than the ADF alternative. The same critical values are used for the ADF and PP tests.

The DF–GLS test

Conventional unit root tests are known to lose power dramatically against stationary alternatives with a low order MA process: a characterization that fits well to a number of macroeconomic time series. Consequently, these original tests have been largely supplanted in many researchers' toolkits by improved alternatives. Along the lines of the ADF test, a more powerful variant is the DFGLS test proposed by Elliott, Rothenberg and Stock (ERS, 1996), described in Baum (2000, 2001), and implemented in Stata[1] as command `dfgls`. `dfgls` performs the ERS efficient test for an autoregressive unit root. This test is similar to an (augmented) Dickey-Fuller *t* test, as performed by `dfuller`, but has the best overall performance in

[1] This command is not built in to Stata version 7, but can be readily installed by using the command "ssc install dfgls". The command is already installed in the version of Stata running on fmrisc.bc.edu. It is available as part of official Stata in version 8.

terms of small-sample size and power, dominating the ordinary Dickey-Fuller test. The `dfgls` test "has substantially improved power when an unknown mean or trend is present" (ERS, p.813).

`dfgls` applies a generalized least squares (GLS) detrending (demeaning) step to the `varname`:

$y_t^d = y_t - \hat{\beta}' z_t$

For detrending, $z_t = (1, t)'$ and $\hat{\beta}_0, \hat{\beta}_1$ are calculated by regressing $[y_1, (1 - \bar{\alpha}L)\, y_2, ..., (1 - \bar{\alpha}L)\, y_T]$ onto $[z_1, (1 - \bar{\alpha}L)\, z_2, ..., (1 - \bar{\alpha}L)\, z_T]$ where $\bar{\alpha} = 1 + \bar{c}/T$ with $\bar{c} = -13.5$, and L is the lag operator. For demeaning, $z_t = (1)'$ and the same regression is run with $\bar{c} = -7.0$. The values of $\bar{c}$ are chosen so that "the test achieves the power envelope against stationary alternatives (is asymptotically MPI *(most powerful invariant))* at 50 percent power" (Stock, 1994, p.2769; emphasis added). The augmented Dickey-Fuller regression is then computed using the y_t^d series:

$\Delta y_t^d = \alpha + \gamma t + \rho y_{t-1}^d + \sum_{i=1}^{m} \delta_i \Delta y_{t-i}^d + \epsilon_t$

where m=`maxlag`.The `notrend` option suppresses the time trend in this regression.

Approximate critical values for the GLS detrended test are taken from ERS, Table 1 (p.825). Approximate critical values for the GLS demeaned test are identical to those applicable to the no–constant, no–trend Dickey–Fuller test, and are computed using the `dfuller` code.

The `dfgls` routine includes a very powerful lag selection criterion, the "modified AIC" (MAIC) criterion proposed by Ng and Perron (2001). They have established that use of this MAIC criterion may provide "huge size improvements" (2001, abstract) in the dfgls test. The criterion, indicating the appropriate lag order, is printed on dfgls' output, and may be used to select the test statistic from which inference is to be drawn.

It should be noted that all of the lag length criteria employed by dfgls (the sequential *t* test of Ng and Perron (1995), the SC, and the MAIC) are calculated, for various lags, by holding the sample size fixed at that defined for the longest lag. These criteria cannot be meaningfully compared over lag lengths if the underlying sample is altered to use all available observations. That said, if the optimal lag length (by whatever criterion) is found to be much less than that picked by the Schwert criterion, it would be advisable to rerun the test with the maxlag option specifying that optimal lag length, especially when using samples of modest size.

The KPSS test

An alternative test is that proposed by Kwiatkowski et al. (1992), the so-called KPSS test, which has a null hypothesis of stationarity (that is, $H_0 : y \sim I(0)$). It is also described in Baum (2000) and implemented in Stata[2] as command kpss. kpss performs the Kwiatkowski, Phillips, Schmidt, Shin (KPSS, 1992) test for stationarity of a time series. The test may be conducted under the null of either trend stationarity (the default) or level stationarity. Inference from this test is complementary to that derived from those based on the Dickey–Fuller distribution (such

[2] This command is not built in to Stata, but can be readily installed in any version of Stata with access to the Web by using the "ssc install kpss" command. The command is already installed in the version of Stata running on fmrisc.bc.edu.

as `dfgls`, `dfuller` and `pperron`). The KPSS test is often used in conjunction with those tests to investigate the possibility that a series is fractionally integrated (that is, neither *I(1)* nor *I(0)*): see Lee and Schmidt (1996).

The series is detrended (demeaned) by regressing y on $z_t = (1, t)^{'} \left(z_t = (1)'\right)$, yielding residuals e_t. Let the partial sum series of e_t be s_t. Then the zero-order KPSS statistic $k_0 = T^{-2} \sum_{t=1}^{T} s_t^2 \,/\, T^{-1} \sum_{t=1}^{T} e_t^2$. For `maxlag`$> 0$, the denominator is computed as the Newey-West estimate of the long run variance of the series; see `[R] newey.`

Approximate critical values for the KPSS test are taken from KPSS (1992). The `kpss` routine includes two options recommended by the work of Hobijn et al. (1998). An automatic bandwidth selection routine has been added, rendering it unnecessary to evaluate a range of test statistics for various lags. An option to weight the empirical autocovariance function by the Quadratic Spectral kernel, rather than the Bartlett kernel employed by KPSS, has also been introduced. These options may be used separately or in conjunction. It is in conjunction that Hobijn et al. found the greatest improvement in the test: "Our Monte Carlo simulations show that the best small sample results of the test in case the process exhibits a high degree of persistence are obtained using both the automatic bandwidth selection procedure and the Quadratic Spectral kernel." (1998, p.14) The `qs` option specifies that the autocovariance function is to be weighted by the Quadratic Spectral kernel, rather than the Bartlett kernel. Andrews (1991) and Newey and West (1994) "indicate that it yields more accurate estimates of σ_ϵ^2 than other kernels in finite samples." (Hobijn et al., 1998, p.6) The `auto` option specifies that the automatic bandwidth selection procedure proposed by Newey and West (1994) as described by Hobijn et al. (1998, p.7) is used to determine `maxlag`, in two stages. First, the "a priori nonstochastic bandwidth parameter" n_T is chosen as a function of the sample size and the specified kernel. The autocovariance function of the estimated residuals is calculated, and used to generate γ as a function of sums of autocorrelations. The `maxlag` to be used in computing the long-run variance, $\hat{m}_T$, is then calculated as $\min\left[T, int\left[\hat{\gamma}T^{\theta}\right]\right]$ where θ=1/3 for the Bartlett kernel and 1/5 for the Quadratic Spectral kernel.

The Leybourne–McCabe test

Like the KPSS test, the test proposed by Leybourne and McCabe (LMc, 1994, 1999) has a null of stationarity, and a unit root alternative hypothesis. The difference lies in the specification. The LMc test, like the ADF test, is parametric; it posits a null of $ARIMA(p, 0, 0)$ (that is, $AR(p)$ with deterministic trend) with an alternative of $ARIMA(p, 1, 1)$. The LMc test is more cumbersome, as it requires estimation of this nonlinear model under the null hypothesis, but it has been shown to be more powerful than the KPSS test. Just as the PP test is a semiparametric alternative to the ADF—that is, the PP test uses a Newey–West long run estimate to deal with dependence in the error process, rather than an explicit $AR(p)$ specification—the KPSS test may be considered as a semiparametric alternative to the LMc test. The tests differ "in how they take account of autocorrelation in y_t under H0" (1994, p.160): the LMc test accounts for autocorrelation in a parametric manner by including lagged terms in the first difference in the series, as does the augmented Dickey-Fuller test. In contrast, the KPSS test modifies the test nonparametrically, "in a manner similar to that in which the Phillips-Perron test is a nonparametric adjustment of the simple DF test." (op.cit.) The authors find that "once y_t does not resemble white noise, the size of [the KPSS test] is likely to be quite badly approximated by its asymptotic distribution, even when the lag length l is relatively high." (1994, p.161) The critical values for the LMc test are identical to those used for KPSS.

Combining inference from I(1) and I(0) tests

The two families of unit root tests may be used in conjunction to establish the nature of the data generating process for a given timeseries, and in particular to signal the presence of fractional integration in the series. If inference from the DFGLS test rejects its null hypothesis of unit root behavior, or nonstationarity, while the KPSS test also rejects its null, then we might conclude that both $I(1)$ and $I(0)$ are rejected by the data. That sets the stage for an alternative explanation of the timeseries' behavior: that of fractional integration, or long-range dependence, in which the series may be characterized as $I(d)$, $0 < d < 1$, neither $I(0)$ nor $I(1)$.

Seasonal unit root tests

The implicit assumption in applying unit root tests to data which have been deseasonalized is that the adjustment method does not affect inference on the stationarity of the time series. However, several authors have called that assumption into question in the case where the SA data have been generated by a moving average filter. The most popular SA technique is the US Census Bureau's X-11 seasonal adjustment program, which passes the data through a sequence of moving average filters. Monte Carlo simulations show that the power of standard unit root tests applied to SA data generated in this manner is reduced, so that the null of nonstationarity is not rejected frequently enough. Although one could deal with this issue by testing NSA data, they are not always available. One can modify the standard AR model that gives rise to the unit root test to

$$\begin{aligned} y_t &= \phi_{1s} y_{t-s} + \epsilon_t \\ (1 - \phi_{1s} L^s) y_t &= \epsilon_t \end{aligned}$$

where $s = 4$ for quarterly data, $s = 12$ for monthly data, etc. If $\phi_{1s} = 1$, then there is a unit root at the s^{th} seasonal frequency, and the s^{th} difference of y_t will remove it. Although this model could be directly tested with the D–F methodology, it is likely to be too simple for most applications, since it restricts the dynamics of y to depend only on seasonal differences. A natural extension of this model would be

$$\phi(L)\phi_s(L^s)y_t = \epsilon_t \tag{5}$$

where $\phi(L)$ is a standard autoregressive polynomial, and $\phi_s(L^s) = (1 - \phi_{1s}L^s - ... - \phi_{rs}L^{rs})$ is the seasonal polynomial. One could have unit roots in either, both or neither of the polynomials, and if they exist in both of the polynomials, both an ordinary difference and a seasonal difference would have to be applied to render the resulting series stationary. The test for a unit root in this context is that of Hylleberg et al. (HEGY, 1990). For quarterly data, the test may be implemented within Stata via the routine `hegy4` of Baum and Sperling (`findit hegy4`).

In this representation, the unit root at the quarterly frequency can be written as

$$\begin{aligned}(1-L^4) &= (1-L)(1+L+L^2+L^3) \\ &= (1-L)(1+L)(1-iL)(1+iL)\end{aligned}$$

and the composite polynomial in (5) can be expanded about its roots with a remainder term:

$$\begin{aligned}\phi(L)\phi_4(L^4) &= -\gamma_1 L(1+L+L^2+L^3) \\ &\quad -\gamma_2(-L)(1-L+L^2-L^3) \\ &\quad -(\gamma_3 L+\gamma_4)(-L)(1-L^2)+\phi^*(L)(1-L^4) \\ \phi^*(L)\Delta^4 y_t &= \gamma_1 y_{1,t-1}+\gamma_2 y_{2,t-1}+\gamma_3 y_{3,t-2}+\gamma_4 y_{3,t-1}+\epsilon_t \\ y_{1t} &= (1+L+L^2+L^3)y_t \\ y_{2t} &= -(1-L+L^2-L^3)y_t \\ y_{3t} &= -(1-L^2)y_t\end{aligned}$$

and this regression may be run as the equivalent of the D–F regression. A test of $\gamma_1 = 0$ corresponds to a test of the standard unit root hypothesis against the alternative of stationarity. A test of $\gamma_2 = 0$ allows for the semi–annual root of -1 versus the stationary alternative. Seasonal unit roots at the quarterly frequency correspond to $\gamma_3 = \gamma_4 = 0$. Thus there will ne no seasonal unit roots in the series if $\gamma_2 \neq 0$ and either $\gamma_3 \neq 0$ or $\gamma_4 \neq 0$, corresponding to a rejection of the null that $\gamma_2 \neq 0$ and the joint null that $\gamma_3 = \gamma_4 = 0$.

The Stata routine `hegy4` performs the test for seasonal unit roots by estimating the four roots of the timeseries representation and presents estimates of these roots as $\pi_1 ... \pi_4$. A joint test for $\pi_3 = \pi_4 = 0$ is also presented. Critical values are those appropriate for T=100, taken from HEGY Table 1. Joint tests for $\pi_2 = \pi_3 = \pi_4 = 0$ and $\pi_1 = \pi_2 = \pi_3 = \pi_4 = 0$, with critical values, are those presented by Ghysels et al. (1994). Critical values for the case of muliplicative seasonality (see below) are from tables 1a-c in Smith and Taylor (1998). Critical values are linearly interpolated for sample sizes in the ranges (48,100) and (100,200).

L:ike the standard D–F test, it may be necessary to augment the HEGY test with additional lags of the dependent variable. The option `Lags(numlist)` specifies the lag orders to be used in augmenting the model with lags of the fourth difference of the timeseries. Its default is zero. If sequential lags are specified starting with 1, HEGY4 automatically conducts a sequential t-test to determine the optimal lag length and optimal lags to be included in the auxiliary regression. It may also be desirable to include deterministic terms such as a constant, trend and in this case seasonal dummy variables in the model. In the `hegy4` routine, the option `det` may take on values `none, const, seas, trend, strend` or `mult`, specifying the process to be tested. The default, as suggested by HEGY and Ghysels et al. (1994), is `seas`, indicating that a set of 3 seasonal dummies plus constant are to be included in the regression. `none` specifies that no deterministic variables are to be included; `const` specifies only a constant. `trend` specifies that a trend is to be included along with a constant term. `strend` specifies that a trend is to be included along with seasonal dummies and a constant term. `mult` specifies that seasonal intercepts (the case of multiplicative seasonality, recommended by Smith and Taylor (1998)) are to be included along with seasonal dummies and a constant term.

The HEGY model may also be defined for monthly data (although the algebra in that case is rather menacing); a Stata routine for that purpose is under development.

Testing for unit roots with structural breaks

A well known problem in the unit root literature is the potential for a series which exhibits structural shifts to fail to reject the unit root null. In the simplest case, a series which undergoes a mean shift is not covariance stationary, but could be made so if regressed on a dummy that identified the shift period (zero before, one after). Early work along these lines was that of Perron (1989) and Perron and Vogelsang (1992). If there is a known break in a sample of T observations at point T_b, we may consider three extensions of the random–walk–with–drift model, where y is considered to be in logs, and no further dynamics are present:

$$\begin{aligned} y_t &= \mu + \delta_1 DVTB_t + y_{t-1} + \epsilon_t \\ y_t &= \mu + \delta_2 DVU_t + y_{t-1} + \epsilon_t \\ y_t &= \mu + \delta_1 DVTB_t + \delta_2 DVU_t + y_{t-1} + \epsilon_t \end{aligned}$$

where $DVTB_t = 1$ in the period $T_b + 1$, and $DVU_t = 1$ for $t > T_b$. The first model considers a level shift (jump) at time $T_b + 1$; the second model considers a change in the growth rate of the series effective at time $T_b + 1$; and the third model considers the possibility that both occur. The $DVTB_t$ dummy is an "impulse" dummy, picking out the single period of the shift, while DVU_t is a shift dummy, which changes the underlying slope of the stochastic trend. The alternative hypotheses to each of these models are, respectively:

$$\begin{aligned} y_t &= \mu + \beta t + \delta_2 DVU_t + \epsilon_t \\ y_t &= \mu + \beta t + \delta_3 DVT_t^* + \epsilon_t \\ y_t &= \mu + \beta t + \delta_2 DVU_t + \delta_3 DVT_t + \epsilon_t \end{aligned}$$

where DVT_t^* and DVT_t both equal zero if $t \leq T_b$. $DVT_t^* = t - T_b$ if $t > T_b$, and $DVT_t = t$ if $t > T_b$. Since there is drift in the first and third models under the null hypothesis, the alternative includes a deterministic trend. In the first model, the null is a unit root with level change; under the alternative, we have a trend stationary series with a change in the intercept to $\mu + \delta_2$. In the second model, the null hypothesis is a unit root with a change in the drift, and the alternative is a trend stationary series with a change in the slope to $\beta + \delta_3$. In the third model, the null is a unit root series with change in both level and drift, and the alternative is a trend stationary series with changes in the intercept and slope. In practice, further dynamics may be necessary to whiten the residuals.

Perron suggests the use of two procedures for modeling this process, depending on whether adjustment following the break is assumed to be instantaneous or gradual. The former is known as the AO (additive outlier) case, where there is a single effect at the breakpoint. Alternatively, the IO (innovational outlier) model may be applied, which allows for a gradual adjustment of the series following the break.

The original derivation of these tests was performed conditional on a known breakpoint. More realistically, we may not know when (or even if) such a breakpoint exists. If the tests are performed conditional on an a priori breakpoint, they may not have maximal power. More recent derivations of unit root tests in the presence of structural change have focused on unknown breakpoints, and in some cases on multiple structural breaks and the methods needed to consistently detect them. For instance, Perron (1997) presents a procedure for locating the single breakpoint with highest likelihood by considering all possible breakpoints in the interior of the sample, selecting that point which maximizes the absolute value of the $t-$statistic for the structural change term. That article presents asymptotic critical values for the unit root test statistic in the presence of a single breakpoint.

A paper using this methodology to examine the impact of structural breaks on unit root testing is Baum et al. (1999), which considers the possibility that common findings of nonstationarity in real exchange rates may be an artifact of structural breaks in the series. In the end, they conclude that unit roots are present even when structural breaks (and the potential for fractional integration) are accounted for. The unit-root test statistics forthcoming from the AO and IO models will account for one-time level shifts which might otherwise be identified as departures from stationarity. However, the behavior of real exchange rate series over our sample period may not be adequately characterized by a single shift; as Lothian (1998) has noted, US dollar-based real exchange rates appear to have exhibited two shifts in mean over the 1980-1987 period, approximately reverting to their pre-1980 level after 1987. In these circumstances, allowing for a single level shift will not suffice. The Perron-Vogelsang methodology has been extended to double mean shifts by Clemente et al. (1988), who demonstrate that a two-dimensional grid search for breakpoints (T_{b1} and T_{b2}) may be used for either the AO or IO models, and provide critical values for the tests. In this context, the AO model involves the estimation of:

$$y_t = \mu + \delta_1 DU_{1t} + \delta_2 DU_{2t} + \tilde{y}_t \qquad (6)$$

and subsequently searching for the minimal $t-$ratio for the hypothesis $\alpha = 1$ in the model:

$$\begin{aligned} \tilde{y}_t \ = \ & \sum_{i=0}^{k} \omega_i DT_{b1,t-i} \sum_{i=0}^{k} \omega_i DT_{b2,t-i} + \alpha \tilde{y}_{t-1} + \\ & \sum_{i=1}^{k} \theta_i \Delta \tilde{y}_{t-i} + e_t, \end{aligned}$$

for $t = k + 2, ..., T$.

For the IO model, the modified equation to be estimated becomes:

$$y_t \ = \ \mu + \delta_1 DU_{1t} + \delta_2 DU_{2t} + \vartheta_1 DT_{b1,t} + \vartheta_2 DT_{b2,t} \qquad (7)$$

$$+\alpha y_{t-1} + \sum_{i=1}^{k} \theta_i \Delta y_{t-i} + e_t, \qquad (8)$$

for $t = k + 2, ..., T$, with a search for the minimal $t-$ratio for the hypothesis $\alpha = 1$.[3]

Code to estimate unit root tests, allowing for one or two structural breaks in either an AO or IO context, is available for Stata as routines `clemao1, clemao2, clemio1`, and `clemio2`. As it has not yet been documented, it is not in the SSC archive, but it is available from the instructor on request.

References

[1] Andrews, D.W.K. 1991. Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. Econometrica, 59, 817-858.

[2] Baum, Christopher F, 2000. Tests for stationarity of a timeseries. Stata Technical Bulletin 57. Available from the course homepage.

[3] Baum, Christopher F., John T. Barkoulas, and Mustafa Caglayan. 1999. Long memory or structural breaks: Can either explain nonstationary real exchange rates under the current float? Journal of International Financial Markets, Institutions, and Money, 9, 359-376. Available as BC EC WP 380.

[4] Baum, Christopher F. and Vince Wiggins, 2001. Tests for stationarity of a timeseries: Update. Stata Technical Bulletin 58. Available from the course homepage.

[5] Clemente, J., Montañes, A. and M. Reyes. 1998. Testing for a unit root in variables with a double change in the mean, Economics Letters, 59, 175-182.

[6] Elliott, Graham, Rothenberg, Thomas J., and James H. Stock. 1996. Efficient Tests for an Autoregressive Unit Root. Econometrica, 64:4, 813-836.

[7] Ghysels, E., Lee, H.S., 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.

[3] These tests customarily are applied to a trimmed sample; we trimmed 5% of the sample from each end when searching for the breakpoints.

[8] Hylleberg, S., Engle, R.F., Granger, C.W.J. and B.S. Yoo. 1990. Seasonal Integration and Cointegration. Journal of Econometrics, 44, 215-238.

[9] Hobijn, Bart, Franses, Philip Hans, and Marius Ooms. 1998. Generalizations of the KPSS-test for Stationarity. Econometric Institute Report 9802/A, Econometric Institute, Erasmus University Rotterdam. http://www.eur.nl/few/ei/papers

[10] Lee, D. and P. Schmidt. 1996. On the power of the KPSS test of stationarity against fractionally-integrated alternatives. Journal of Econometrics, 73, 285-302.

[11] Leybourne, S.J. and B.P.M. McCabe. A consistent test for a unit root. Journal of Business and Economic Statistics, 12, 1994, 157-166.

[12] Leybourne, S.J. and B.P.M. McCabe. Modified stationarity tests with data-dependent model-selection rules. Journal of Business and Economic Statistics, 17, 1999, 264-270.

[13] Lothian, J. 1998. Some new stylized facts of floating exchange rates, Journal of International Money and Finance, 17, 29-39.

[14] Newey, W.K. and K.D. West. 1994. Automatic Lag Selection in Covariance Matrix Estimation. Review of Economic Studies, 61, 631-653.

[15] Ng, Serena, and Pierre 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.

[16] Ng, Serena, and Pierre Perron. 2001. Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power. Econometrica, 69, 1519–1554.

[17] Perron, Pierre. 1989. The Great Crash, The Oil Price Shock and the Unit Root Hypothesis. Econometrica, 57, 1361–1401.

[18] Perron, P. 1990. Testing for a unit root in a time series with a changing mean, Journal of Business and Economic Statistics, 8:2, 153-162.

[19] Perron, Pierre. 1997. Further Evidence on Breaking Trend Functions in Macroeconomic Variables. Journal of Econometrics, 80, 355–385.

[20] Perron, P. and T. Vogelsang. 1992. Nonstationarity and level shifts with an application to purchasing power parity, Journal of Business and Economic Statistics, 10:3, 301-320.

[21] Smith, R. J., and A. M. R. Taylor. 1998. Additional Critical Values and Asymptotic Representations for Seasonal Unit

Root Tests, Journal of Econometrics, 85, 269-88.

[22] Stock, James H. 1994. Unit Roots, Structural Breaks and Trends. Chapter 46 in Handbook of Econometrics, Volume IV, R.F. Engle and D.L. McFadden, eds. Amsterdam: Elsevier.