sts15.1	Tests for stationarity of a time series: update

Christopher F. Baum, Boston College, baum@bc.edu
Richard Sperling, The Ohio State University, rsperling@boo.net

Abstract: Enhances the Elliott–Rothenberg-Stock DF-GLS test and the Kwiatkowski–Phillips–Schmidt–Shin KPSS tests for stationarity of a time series introduced in Baum (2000), and corrects an error in both routines.

Keywords: stationarity, unit root, time series.

Changes to dfgls

`dfgls` did not handle missing initial values properly. That is, if the time series variable specified had initial values not excluded by `if` or `in` conditions, those values were improperly considered in the construction of the sample size. This would apply as well to the consideration of variables with time series operators, such as `D.gdp`, since those variables will have at least one missing observation at the outset. This has been corrected.

The `dfgls` routine has been enhanced to add a very powerful lag selection criterion, the "modified AIC" (MAIC) criterion proposed by Ng and Perron (2000). They have established that use of this MAIC criterion may provide "huge size improvements" 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.

New syntax for kpss

`kpss` *varname* [`if` *exp*] [`in` *range*] [`, maxlag(`*#*`) notrend qs auto`]

`kpss` did not make use of all available observations in the computation of the autocovariance function. This has been corrected. The online help file now provides instructions for reproducing the statistics reported in `kpss` (Table 5) from a dataset available online.

The `kpss` routine has been enhanced to add 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 combination. It is in combination 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, 14).

Options

`qs` 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, 6).

`auto` specifies that the automatic bandwidth selection procedure proposed by Newey and West (1994) as described by Hobijn et al. (1998, 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, $\widehat{m}_T$, is then calculated as $\min\left[T, \mathrm{int}\left[\hat{\gamma}T^\theta\right]\right]$ where $\theta = 1/3$ for the Bartlett kernel and $\theta = 1/5$ for the quadratic spectral kernel.

Additional saved results

`dfgls` saves the modified AIC at lag n in `r(maic`*n*`)`.

References

Andrews, D. W. K. 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59: 817–858.

Baum, K. 2000. sts15: Test for stationarity of a time series. *Stata Technical Bulletin* 57: 36–39.

Hobijn, B., P. H. Franses, and M. 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`.

Newey, W. K. and K. D. West. 1994. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61: 631–653.

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.

——. 2000. Lag length selection and the construction of unit root tests with good size and power. *Econometrica*, in press.