The Stata Journal (2018)
18, Number 1, pp. 22–28

Unit-root tests based on forward and reverse Dickey–Fuller regressions

Jesús Otero
Universidad del Rosario
Bogotá, Colombia
jesus.otero@urosario.edu.co

Christopher F. Baum
Boston College
Chestnut Hill, MA
baum@bc.edu

Abstract. In this article, we present the command `adfmaxur`, which computes the Leybourne (1995, *Oxford Bulletin of Economics and Statistics* 57: 559–571) unit-root statistic for different numbers of observations and the number of lags of the dependent variable in the test regressions. The latter can be either specified by the user or endogenously determined. We illustrate the use of `adfmaxur` with an empirical example.

Keywords: st0511, adfmaxur, unit-root test, critical values, lag length, *p*-values

1 Introduction

During the last three decades or so, testing for the presence of a unit root has become a key step in the empirical analysis of time series in economics and finance. Among the several tests that have been proposed in the literature over the years, perhaps the most commonly applied is the augmented Dickey–Fuller (ADF) test proposed by Said and Dickey (1984), which follows the work of Dickey and Fuller (1979). This test has had long-lasting appeal because it is regression based; thus, it is straightforward to calculate. Despite this clear advantage, a common criticism is that the ADF test exhibits poor power properties as documented, for example, by DeJong et al. (1992). To improve the power of the test, Leybourne (1995) recommends an approach based on the maximum ADF t statistic that results from estimating Dickey–Fuller type regressions using forward and reverse realizations of the data. According to Leybourne (1995), the resulting statistic, commonly referred to as the ADF_{max} test, is not only easy to compute but also exhibits greater power, so it is more likely to reject a false unit-root hypothesis than the standard ADF test.

To implement the ADF_{max} test, Leybourne (1995) tabulates critical values for the two model specifications that are mainly used in practical empirical work: a model with an intercept and a model with an intercept and trend. In both cases, the tabulated critical values are based on 20,000 replications for $T = 25$, 50, 100, 200, and 400 observations. However, these critical values do not account for their possible dependence on the number of lags of the dependent variable that may be needed to account for serial correlation, nor do they account for the information criterion used to determine the optimal number of lags. To improve inference, Otero and Smith (2012) performed an extensive set of Monte Carlo simulation experiments that were summarized through response surface regressions, from which critical values of the Leybourne (1995) tests were

 st0511

subsequently calculated for different specifications of deterministic components, number of observations, and lag order, where the last can be determined either exogenously by the user or endogenously through a data-dependent procedure. For any given $\text{ADF}_{\max}$ statistic, Otero and Smith (2012) provided an Excel spreadsheet to calculate critical values at the 1%, 5%, and 10% significance levels as well as the associated p-value of the statistic.[1]

In this article, we present the command `adfmaxur`, which computes the $\text{ADF}_{\max}$ test statistic jointly with the 1%, 5%, and 10% significance levels and associated p-value. The command allows for the presence of stochastic processes with nonzero mean and nonzero trend and for the lag order to be determined either exogenously or endogenously using a data-dependent procedure.

The article is organized as follows: Section 2 provides an overview of the $\text{ADF}_{\max}$ unit-root test. Section 3 describes the command `adfmaxur`. Section 4 illustrates the use of `adfmaxur` with an empirical example. Section 5 concludes.

2 The Leybourne $\text{ADF}_{\max}$ unit-root test

Leybourne (1995) proposes a unit-root test that involves ordinary least-squares estimation of ADF-style regressions based on the forward and reverse realizations of the time series of interest. Thus, letting y_t indicate the forward realization of a time series, consider an ADF-type regression for the model that includes a constant and a trend as deterministic components,

$$\Delta y_t = \alpha + \gamma t + \beta y_{t-1} + \sum_{j=1}^{p} \delta_j \Delta y_{t-j} + \varepsilon_t \qquad (1)$$

where Δ is the first-difference operator, p is the number of lags of the dependent variable that are included as additional regressors to allow for residual serial correlation, and $t = 1, \ldots, T$ time observations. In this setting, the unit-root null hypothesis is given by $H_0 : \beta = 0$, against the alternative that the series of interest is stationary around a linear trend term; that is, $H_1 : \beta < 0$. Let ADF_f denote the regression t statistic for $\beta = 0$ in (1).[2]

Now, let us consider the reverse realization of y_t, which is given by $z_t = y_{T-t+1}$. Thus, $z_1 = y_T$, $z_2 = y_{T-1}$, and so on until $z_T = y_1$. The corresponding ADF type regression applied to z_t is

$$\Delta z_t = \alpha^* + \gamma^* t + \beta^* z_{t-1} + \sum_{j=1}^{p} \delta_j^* \Delta z_{t-j} + \varepsilon_t^*$$

1. See http://www2.warwick.ac.uk/fac/soc/economics/staff/academic/jeremysmith/research.
2. Dickey and Fuller (1981) construct the likelihood-ratio statistics to test the joint null hypotheses that the true model is a random walk with and without drift, which in terms of (1) are represented as $H_0 : \gamma = \beta = 0$ and $H_0 : \alpha = \gamma = \beta = 0$, respectively. The distributions under the null hypothesis of these additional statistics were not tabulated by Leybourne (1995) and therefore are not considered in this article.

where we are interested in ADF_r, the regression t statistic for $\beta^* = 0$. Finally, the $\text{ADF}_{\max}$ statistic is simply given by the maximum between ADF_f and ADF_r.

3 The adfmaxur command

The command **adfmaxur** calculates the $\text{ADF}_{\max}$ test statistic, the statistic's associated 1%, 5%, and 10% finite sample critical values, and the statistic's approximate p-value. The estimation of critical values and approximate p-value permits different combinations of the number of observations, T, and lags of the dependent variable in the test regression, p, where the latter can be either specified by the user or optimally selected using a data-dependent procedure.[3]

3.1 Syntax

`adfmaxur` *varname* [*if*] [*in*] [`, noprint` `maxlag(`*integer*`)` `trend`]

Note that *varname* may not contain gaps. *varname* may contain time-series operators. **adfmaxur** may be applied to one unit of a panel.

3.2 Options

`noprint` specifies that the results be returned but not printed.

`maxlag(`*integer*`)` sets the number of lags to be included in the test regression to account for residual serial correlation. By default, **adfmaxur** sets the number of lags following Schwert (1989), with the formula `maxlag()` $= \text{int}\{12(T/100)^{0.25}\}$, where T is the total number of observations. In either case, the number of lags appears in the row labeled `FIXED` of the output table.

In addition, **adfmaxur** determines the optimal number of lags according to the Akaike and Schwarz information criteria, denoted `AIC` and `SIC`, respectively, and also following the general-to-specific (GTS) algorithm advocated by Hall (1994) and Ng and Perron (1995). The idea of the GTS algorithm is to start by setting an upper bound on p, denoted $p_{\max}$, estimating (1) with $p = p_{\max}$, and testing the statistical significance of $\delta_{p_{\max}}$. If this coefficient is statistically significant, using, for example, significance levels of 5% (referred to as $\text{GTS}_{0.05}$) or 10% (referred to as $\text{GTS}_{0.10}$), one selects $p = p_{\max}$. Otherwise, the order of the estimated autoregression in (1) is reduced by one until the coefficient on the last included lag is found to be statistically different from zero.

3. The optimally selected number of lags yields the global minimizer of an information criterion; see, for example, Ng and Perron (2005).

`trend` specifies the modeling of intercepts and trends. By default, `adfmaxur` assumes *varname* is a nonzero mean stochastic process, so a constant is included in the test regression. If, on the other hand, the `trend` option is specified, *varname* is assumed to be a nonzero trend stochastic process, in which case a constant and a trend are included in the test regression.

3.3 Stored results

`adfmaxur` stores the following results in `r()`:

Scalars	
`r(N)`	number of observations in the test regression
`r(minp)`	first time period used in the test regression
`r(maxp)`	last time period used in the test regression
Macros	
`r(varname)`	variable name
`r(treat)`	either `constant` or `constant and trend`
`r(tsfmt)`	time-series format of the time variable
Matrices	
`r(results)`	results matrix, 5×6

The rows of the results matrix indicate which method of lag length was used: `FIXED` (lag selected by user, or using Schwert's formula); `AIC`; `SIC`; `GTS05`; or `GTS10`.

The columns of the results matrix contain, for each method, the number of lags used; the $\mathrm{ADF}_{\mathrm{max}}$ statistic; its p-value; and the critical values at 1%, 5%, and 10%, respectively.

4 Empirical application

Examining the time-series properties of the unemployment rate is important to understand the way this variable responds to shocks. On the one hand, under the natural rate hypothesis, shocks to unemployment have a temporary effect because unemployment will return to its natural rate in the long run. On the other hand, under the hysteresis hypothesis, the same shocks will have a permanent effect because of rigidities in the labor market.

This section illustrates the use of `adfmaxur` to assess these two hypotheses. To this end, we use monthly seasonally adjusted data on unemployment rates for the 48 contiguous U.S. states. The data, which are freely available from the Federal Reserve Economic Data of the Federal Reserve Bank of St. Louis, cover the period between `1976m1` and `2017m2`, for a total of $T = 494$ time observations for each state.[4]

4. The task of downloading the time series from the Federal Reserve Economic Data database was greatly simplified using the command `freduse`; see Drukker (2006).

We begin by loading the dataset and declaring it as a time series:

```
. use usurates
. tsset date, monthly
        time variable:  date, 1976m1 to 2017m2
                delta:  1 month
```

We would like to test whether the unemployment rate in each state contains a unit root against the alternative that it is a stationary process. For practical purposes, visual inspection of the time plot of the variable of interest often provides useful guidelines as to whether a linear trend term should be included in the test regressions. For our purposes, given that each unemployment series has a nonzero mean, but not trending behavior, the relevant ADF_{max} statistic is that based on a test regression that includes constant but not trend, which is the default option for `adfmaxur`.

Setting a `maxlag()` of $p = 12$, we see that the results of applying `adfmaxur` to the unemployment rate in the state of Alabama, denoted `ALUR`, are

```
. adfmaxur ALUR, maxlag(12)
Leybourne (1995) test results for 1977m2 - 2017m2
Variable name: ALUR
Ho: Unit root
Ha: Stationarity
Model includes constant
--------------------------------------------------------------------------
Criteria   Lags    ADFmax stat. p-value     1% cv       5% cv     10% cv
--------------------------------------------------------------------------
FIXED       12      -2.165        0.088    -3.006      -2.412     -2.105
  AIC       9       -2.234        0.085    -3.103      -2.475     -2.153
  SIC       6       -2.673        0.027    -3.039      -2.430     -2.118
GTS05       9       -2.234        0.086    -3.114      -2.483     -2.161
GTS10       9       -2.234        0.088    -3.133      -2.498     -2.172
--------------------------------------------------------------------------
```

`adfmaxur` runs in a fraction of a second, and the `maxlag()` option, if not specified, is determined following Schwert (1989), as it is in the official `dfgls` command.

Table 1 summarizes the results of applying `adfmaxur` to the first 20 unemployment rates that are included in `usurates.dta`. Interestingly, the results reveal that, in some cases, inference may depend upon the procedure used to select the augmentation order of the test regressions. For example, in the cases of Kansas (`KSUR`) and Massachusetts (`MAUR`), we fail to reject the unit-root null at the 5% significance level when using `FIXED` but not with the other criteria. In contrast, in the case of Idaho (`IDUR`), the unit-root null is rejected (at the same significance level) using `FIXED` but not with the other criteria. In summary, the output reported by `adfmaxur` can be viewed as serving the purpose of assessing the sensitivity of the results to the selection of the method for lag determination in the test regressions.

Table 1. Application of the $\text{ADF}_{\max}$ test to unemployment rates in selected U.S. states

State	FIXED		AIC		SIC		GTS05		GTS10	
	Stat.	p.val.	Stat.	p.val.	Stat.	p.val.	Stat.	p.val.	Stat.	p.val
ALUR	−2.165	0.088	−2.234	0.085	−2.673	0.027	−2.234	0.086	−2.234	0.088
ARUR	−1.403	0.336	−1.631	0.252	−1.399	0.343	−1.179	0.458	−1.469	0.324
AZUR	−2.057	0.110	−2.057	0.121	−2.736	0.023	−2.057	0.122	−2.057	0.125
CAUR	−2.906	0.013	−2.851	0.020	−3.005	0.011	−3.152	0.009	−2.851	0.021
COUR	−2.615	0.030	−2.581	0.039	−2.581	0.035	−2.581	0.040	−2.581	0.041
CTUR	−2.406	0.051	−2.152	0.100	−1.723	0.210	−1.723	0.221	−2.406	0.061
DEUR	−2.251	0.073	−2.185	0.094	−2.185	0.087	−2.290	0.077	−2.251	0.085
FLUR	−2.561	0.034	−3.050	0.012	−2.693	0.026	−3.050	0.012	−3.050	0.013
GAUR	−1.992	0.126	−1.959	0.145	−2.314	0.065	−1.799	0.195	−1.959	0.150
IAUR	−1.810	0.177	−2.010	0.132	−1.831	0.174	−2.010	0.134	−1.830	0.187
IDUR	−2.623	0.029	−2.278	0.077	−1.983	0.131	−1.983	0.141	−2.278	0.081
ILUR	−2.061	0.110	−2.085	0.114	−2.975	0.012	−2.085	0.116	−2.085	0.119
INUR	−2.134	0.094	−2.338	0.068	−2.157	0.092	−2.173	0.098	−2.338	0.071
KSUR	−2.379	0.054	−2.853	0.020	−3.109	0.008	−3.109	0.010	−3.109	0.011
KYUR	−2.199	0.082	−2.077	0.116	−2.073	0.110	−2.077	0.118	−2.077	0.120
LAUR	−1.815	0.176	−1.763	0.204	−2.180	0.088	−1.763	0.207	−1.763	0.209
MAUR	−2.268	0.070	−2.494	0.048	−3.467	0.003	−2.818	0.022	−2.494	0.050
MDUR	−2.924	0.013	−3.040	0.012	−3.040	0.010	−3.040	0.012	−3.040	0.013
MEUR	−2.372	0.055	−2.367	0.064	−2.367	0.058	−2.367	0.065	−2.367	0.067
MIUR	−2.015	0.120	−2.066	0.119	−2.077	0.109	−2.247	0.084	−2.247	0.086

5 Concluding remarks

In this article, we presented the command **adfmaxur**, which calculates the Leybourne (1995) $\text{ADF}_{\max}$ unit-root test, the test's 1%, 5%, and 10% finite-sample critical values, and the test's approximate p-value. The critical values and approximate p-value are estimated as a function of the number of observations, T, lags of the dependent variable in the test regressions, p, and also allow for the lag length to be determined either exogenously by the user or endogenously using a data-dependent procedure.

6 Acknowledgments

Jesús Otero acknowledges financial support received from the Universidad del Rosario through Fondo de Investigaciones de la Universidad del Rosario. We are also grateful to an anonymous referee for useful comments and suggestions. The usual disclaimer applies.

7 References

DeJong, D. N., J. C. Nankervis, N. E. Savin, and C. H. Whiteman. 1992. The power problems of unit root test in time series with autoregressive errors. *Journal of Econometrics* 53: 323–343.

Dickey, D. A., and W. A. Fuller. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74: 427–431.

———. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* 49: 1057–1072.

Drukker, D. M. 2006. Importing Federal Reserve economic data. *Stata Journal* 6: 384–386.

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.

Leybourne, S. J. 1995. Testing for unit roots using forward and reverse Dickey–Fuller regressions. *Oxford Bulletin of Economics and Statistics* 57: 559–571.

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.

———. 2005. A Note on the Selection of Time Series Models. *Oxford Bulletin of Economics and Statistics* 67: 115–134.

Otero, J., and J. Smith. 2012. Response surface models for the Leybourne unit root tests and lag order dependence. *Computational Statistics* 27: 473–486.

Said, S. E., and D. A. Dickey. 1984. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* 71: 599–607.

Schwert, G. W. 1989. Tests for unit roots: A Monte Carlo investigation. *Journal of Business and Economic Statistics* 7: 147–159.

About the authors

Jesús Otero is a professor of economics at the Universidad del Rosario in Colombia. He received his PhD in economics from Warwick University in the United Kingdom. His field of study is applied time-series econometrics.

Christopher F. Baum is a professor of economics and social work at Boston College. He is an associate editor of the *Stata Journal* and the *Journal of Statistical Software*. Baum founded and manages the Boston College Statistical Software Components (SSC) archive at RePEc (http://repec.org). His recent research has addressed issues in social epidemiology and the progress of refugees in developed economies.