The `ZivotAndrews` package

Riccardo (Jack) Lucchetti

version 1.0

Abstract

This package implements the Zivot and Andrews (1992) unit-root test for series with possible structural breaks, using a novel approach for calculating p-values and critical values. A graphical interface is also available.

1 Description

The issue of testing for unit roots in time series potentially subject to structural breaks was first thoroughly investigated by Perron (1989), where the break date was assumed known; shortly afterwards, this limitation was overcome by Zivot and Andrews (1992), who proposed a simple adaptation of Perron's setup to accommodate cases when the break date is unknown.

Like Perron's test, the Zivot-Andrews unit-root test comes in three different flavours: model (A), in which the series may be subject to an abrupt change in level (the so-called "crash" model), model (B), in which the structural break affects the slope of a deterministic linear trend and (C), which combines the two previous ones. Here I follow the notation in Perron (1989) and Zivot and Andrews (1992) and indicate with T_B the observation at which the structural break occurs. The test statistic for the three models (A), (B) and (C) are simple variations on the Augmented Dickey-Fuller theme, and use an axiliary regression of the following form:

$$\Delta y_t = \mu + \theta\ DU_t + \beta t + d\ D(T_B)_t + \alpha y_{t-1} + \sum_{j=1}^{k} c_j \Delta y_{t-j} + \varepsilon_t \quad (1)$$

$$\Delta y_t = \mu + \beta t + \gamma\ DT^*_t + \alpha y_{t-1} + \sum_{j=1}^{k} c_j \Delta y_{t-j} + \varepsilon_t \quad (2)$$

$$\Delta y_t = \mu + \theta\ DU_t + \gamma\ DT^*_t + \beta t + d\ D(T_B)_t + \alpha y_{t-1} + \sum_{j=1}^{k} c_j \Delta y_{t-j} + \varepsilon_t \quad (3)$$

where $D(T_B)_t = \mathbf{1}[t = T_B + 1]$, $DU_t = \mathbf{1}[t > T_B]$, $DT^*_t = \mathbf{1}[t > T_B](t - T_B)$, and $\mathbf{1}[\cdot]$ is the indicator function. As the in ADF test, the relevant test statistic is the t-ratio for the parameter α, and terms containing the lagged dependent variables in differences are used to mop up short-run dynamics; the choice of the maximal lag k is usually data-driven (more on this later). The test is performed by computing the regression above for all possible values of T_B and choosing the one that minimises the test statistic.

For each of the three models, the limit distribution of the test statistics (1)—(3) is a functional of suitably modified Brownian motions and lacks a closed-form representation. In this package, p-values and critical values are computed using the approximation described in Lucchetti (2026).

2 Syntax

The main function that the package provides is called `ZA_test`, and takes three arguments: the series name, the model type as an integer (1=A, 2=B, 3=C) and an optional bundle to inflect the behaviour of the function.

The accepted values for the option bundle are `maxlag`, to set a maximum value for k (default=12) and `method`, for choosing the k actually used in the auxiliary regression. The `method` key can take 5 different values: with `"fixed"` the regression will set k to the `maxlag` key. Otherwise, an automatic procedure will be followed: if `method` equals `"aic"`, `"bic"` or `"hqc"`, the corresponding information criterion will be used. Otherwise, one can use `"tstat"` (the default), where decreasing values of k, starting from `maxlag`, will be used, stopping as soon as the last c_j coefficient is significant at a 10% level.

The complete list of the items in he bundle returned by `ZA_test` is in Section 5. The most useful are probably `mintest`, the actual test statistic, and `pvalue`, its p-value calculated according to Lucchetti (2026). The ordinal number of the break point T_B is contained in `mintime`, with `mintlab` containing its translation as a string.

Another public function is available, called `ZA_plot`, that plots the t-test for $\alpha = 0$ for all values of T_B, together with the 10% and 5% critical values as horizontal lines. It takes as its argument a bundle created by the `ZA_test` function. Optionally, a second parameter (a string) can be given, to redirect the output to a graphical file.

3 An example

In this example, we replicate one of the tests presented in Zivot and Andrews (1992) using the famous dataset provided in Nelson and Plosser (1982) and available in gretl as `np.gdt`. We'll use the log real per capita

GNP, for which the variant (A) of the test was chosen in the original article. The code fragment

```
set verbose off
include ZivotAndrews.gfn
open np

series y = log(rpcgnp)
bun0 = ZA_test(y, 1)
```

performs the test with the default parameters, with the following output

```
Zivot-Andrews test on y, Model (A)
Sample: 1909 - 1970 (55 observations)
Min. t-test at obs 70 (1929): -4.53529 [pval = 0.085253]
Number of lags of Δy in regression = 6: (max = 12, method = tstat)
```

In the original article, however, the number of lags k was chosen via a different criterion and was set to 7. This can be obtained by using the `fixed` method as follows:

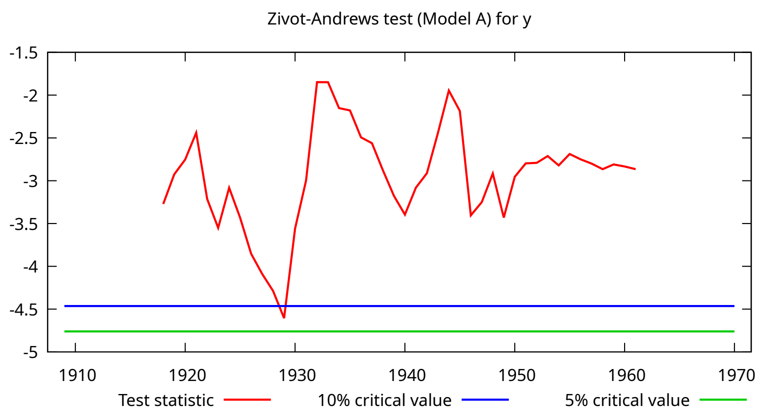
```
bun1 = ZA_test(y, 1, _(method="fixed", maxlag=7))
```

which yields

```
Zivot-Andrews test on y, Model (A)
Sample: 1909 - 1970 (54 observations)
Min. t-test at obs 70 (1929): -4.60582 [pval = 0.0722364]
Number of lags of Δy in regression = 7: (max = 7, method = fixed)
```

and corresponds to the published result.

Figure 1: Plot of the test statistic versus T_B



Invoking the `ZA_plot` function on `bun1`, as in

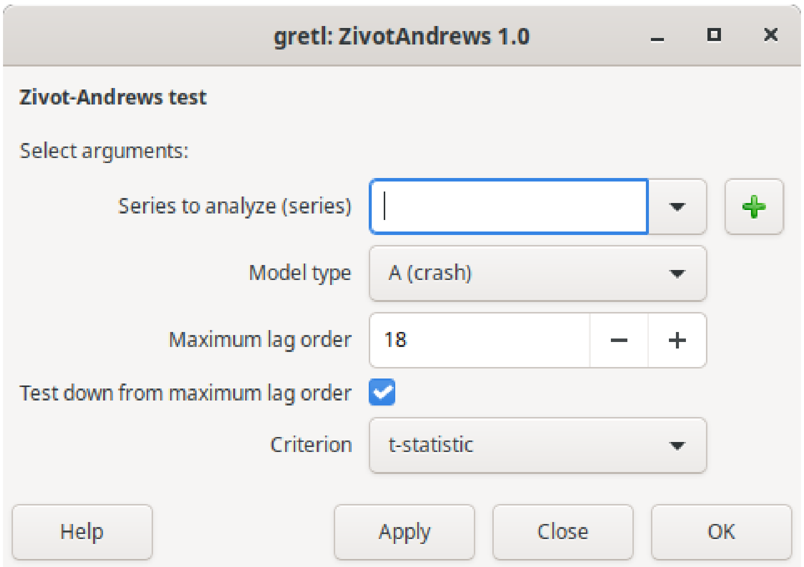
```
ZA_plot(bun1)
```

a plot like the one shown in Figure 1 is produced.

4 The graphical interface

The functions described in the previous section are also accessible via a graphical interface, similar to the one in Figure 2, whose usage should be self-explanatory. When you install the package, this can be attached to the *Variable > Unit root tests* menu of the gretl program.[1] On successful completion, an output text window will be produced; this window will also contain an icon bar on top that makes it possible to save the results and/or generate the plot.

Figure 2: GUI Interface



5 Public functions

```
function bundle ZA_test(series y, scalar type, bundle opts)
```

Compute the test and returns a bundle. The arguments are

1. `y`, the series to test
2. `type`, the model variant (see eqns. (1)—(3))

[1] If you're running gretl with a language setting different from English, this will be translated, but you get the idea.

3. `opts`, an options bundle (optional).

The options bundle can contain the following keys (see section 2):

1. `method`, a string among `fixed`, `tstat` (the default), `aic`, `bic` and `hqc`;
2. `maxlag`, an integer
3. `verbosity`, scalar, amount of output. 0=no output; 1=test details (default), 2=test details + auxiliary regression at T_B

The returned bundle contains

Key	**Value**
`type`	model type (1:A, 2:B, 3:C)
`mintime`	the value of T_B
`mintest`	the value of the test statistic
`mintlab`	a label with T_B in text form
`pvalue`	the test p-value
`t1`	beginning of sample
`t2`	end of sample
`T`	sample size
	Items from the test regression
`coeff`	coefficient vector of the test regression
`depname`	string, the name of the series y_t
`vcv`	the coefficient cov. Matrix of the test regression
`stderr`	standard errors for the coefficients of the test regression
`maxk`	maximum number of lags for the test regression
`bestlag`	value of k actually used
`zatest`	series, value of the test for each T_B

```
function void ZA_plot(bundle tbun, string dest)
```

Plots the values of the test statistic as a function of the breakpoint T_B. Its arguments are:

1. `tbun`, a bundle generated via `ZA_test`.
2. `dest`, an optional string with a filename to store the output, such as `myplot.pdf` or `foobar.eps`. If omitted or `display`, the plot will be produced on screen.

The function doesn't return anything.

6 Changelog

1.0 : Initial release.

References

Lucchetti, Riccardo. 2026. An improved density approximation for the Zivot-Andrews test. *Economics Letters* https://www.sciencedirect.com/science/article/pii/S0165176526000273.

Nelson, Charles R and Charles R Plosser. 1982. Trends and random walks in macroeconomic time series: some evidence and implications. *Journal of Monetary Economics* 10 (2):139–162.

Perron, Pierre. 1989. The great crash, the oil price shock, and the unit root hypothesis. *Econometrica* :1361–1401.

Zivot, Eric and Donald W. K. Andrews. 1992. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business and Economic Statistics* 20 (1):25–44.