

estat sbknown — Test for a structural break with a known break date

Description	Quick start	Menu for estat	Syntax
Options	Remarks and examples	Stored results	Methods and formulas
References	Also see		

Description

`estat sbknown` performs a Wald or a likelihood-ratio (LR) test of whether the coefficients in a time-series regression vary over the periods defined by known break dates.

`estat sbknown` requires that the current estimation results be from [regress](#) or [ivregress 2sls](#).

Quick start

Test for a structural break at January 1983 for current estimation results

```
estat sbknown, break(tm(1983m1))
```

As above, but for the first quarter of 1997

```
estat sbknown, break(tq(1997q1))
```

As above, but perform an LR test instead of a Wald test

```
estat sbknown, break(tq(1997q1)) lr
```

Perform a Wald test for multiple breaks at dates 1997q1 and 2005q1

```
estat sbknown, break(tq(1997q1) tq(2005q1))
```

Menu for estat

Statistics > Postestimation

Syntax

```
estat sbknown, break(time_constant_list) [options]
```

<i>options</i>	Description
* break (<i>time_constant_list</i>)	specify one or more break dates
breakvars ([<i>varlist</i>] [, <u>constant</u>])	specify variables to be included in the test; by default, all coefficients are tested
wald	request a Wald test; the default
lr	request an LR test

***break**() is required.

You must **tsset** your data before using **estat sbknown**; see [TS] **tsset**.

collect is allowed; see [U] **11.1.10 Prefix commands**.

Options

break(*time_constant_list*) specifies a list of one or more hypothesized break dates. **break**() is required with at least one break date.

time_constant_list is a list of one or more time-constant elements. These elements can be specified as time literals or as the underlying numeric values that correspond to the hypothesized break dates. If you specify *time_constant_list* using time literals, you must use one of the **datetime** pseudofunctions; see [D] **Datetime**.

breakvars([*varlist*] [, constant]) specifies variables to be included in the test. By default, all the coefficients are tested.

constant specifies that a constant be included in the list of variables to be tested. **constant** may be specified only if the original model was fit with a constant term.

wald requests that a Wald test be performed. This is the default.

lr requests that an LR test be performed instead of a Wald test.

Remarks and examples

[stata.com](http://www.stata.com)

estat sbknown performs a test of the null hypothesis that the coefficients do not vary over the subsamples defined by the specified known break dates. The null hypothesis of no structural break can be tested using a Wald or an LR test.

Consider the linear regression

$$y_t = \mathbf{x}_t\boldsymbol{\beta} + \epsilon_t$$

A model with a structural break allows the coefficients to change after a break date. If b is the break date, the model is

$$y_t = \begin{cases} \mathbf{x}_t\boldsymbol{\beta} + \epsilon_t & \text{if } t \leq b \\ \mathbf{x}_t(\boldsymbol{\beta} + \boldsymbol{\delta}) + \epsilon_t & \text{if } t > b \end{cases}$$

For this model, the null and alternative hypotheses are $H_0: \boldsymbol{\delta} = 0$ and $H_a: \boldsymbol{\delta} \neq 0$.

For a classical linear model where ϵ_t is independent and identically distributed, this test is known as the **Chow (1960)** test.

▷ Example 1: Test for a single known break date

In `usmacro.dta`, we have data for the `fedfunds` series from the third quarter of 1954 to the fourth quarter of 2010 from the Federal Reserve Economic Database (FRED), a macroeconomic database provided by the Federal Reserve Bank of Saint Louis.

We want to model the federal funds rate as a function of its first lag, but we are concerned that there may be a structural break after 1981q2. We fit the model parameters using `regress`, and then we use `estat sbknown` to test for a structural break.

```
. use https://www.stata-press.com/data/r17/usmacro
(Federal Reserve Economic Data - St. Louis Fed)

. regress fedfunds L.fedfunds
(output omitted)

. estat sbknown, break(tq(1981q2))

Wald test for a structural break: Known break date
Sample: 1954q4 thru 2010q4
Break date: 1981q2
H0: No structural break
Number of obs =    225
      chi2(2) = 6.4147
      Prob > chi2 = 0.0405

Exogenous variables: L.fedfunds
Coefficients included in test: L.fedfunds _cons
```

We reject the null hypothesis of no structural break at the 5% level.

◀

▷ Example 2: Test for multiple known breaks

Suppose we divide the data into three subsamples for the periods 1954q4 to 1970q1, 1970q2 to 1995q1, and 1995q2 to 2010q4, specified by the break dates at 1970q1 and 1995q1. We would like to test whether the coefficients are the same in these subsamples. We do this by specifying multiple dates in the `break()` option.

```
. estat sbknown, break(tq(1970q1) tq(1995q1))

Wald test for a structural break: Known break dates
Sample: 1954q4 thru 2010q4
Break dates: 1970q1 1995q1
H0: No structural break
Number of obs =    225
      chi2(4) = 4.6739
      Prob > chi2 = 0.3224

Exogenous variables: L.fedfunds
Coefficients included in test: L.fedfunds _cons
```

We fail to reject the null hypothesis of no structural break for the specified dates.

◀

Stored results

`estat sbknown` stores the following in `r()`:

Scalars

<code>r(chi2)</code>	χ^2 test statistic
<code>r(p)</code>	p -value for χ^2 test
<code>r(df)</code>	degrees of freedom

Macros

<code>r(breakdate)</code>	list of break dates
<code>r(breakvars)</code>	list of variables whose coefficients are included in the test
<code>r(test)</code>	type of test

Methods and formulas

A test for a structural break with a known break date can be constructed by fitting a linear regression with an indicator variable as

$$y_t = \mathbf{x}_t\boldsymbol{\beta} + (b \leq t)\mathbf{x}_t\boldsymbol{\delta} + \epsilon_t$$

The null hypothesis of no structural break is $H_0: \boldsymbol{\delta} = \mathbf{0}$. This can be tested by constructing a Wald statistic or an LR statistic, both with $\chi^2(k)$ as the limiting distribution, where k is the number of parameters in the model.

A regression model with multiple breaks may be expressed as

$$y_t = \mathbf{x}_t\boldsymbol{\beta} + \mathbf{x}_t\{(b_1 \leq t < b_2)\boldsymbol{\delta}_1 + (b_2 \leq t < b_3)\boldsymbol{\delta}_2 + \dots + (b_m \leq t)\boldsymbol{\delta}_m\} + \epsilon_t$$

where b_1, \dots, b_m are $m \geq 2$ break dates. The null hypothesis of no structural break is a joint test given by $H_0: \boldsymbol{\delta}_1 = \dots = \boldsymbol{\delta}_m = \mathbf{0}$.

References

- Chow, G. C. 1960. Tests of equality between sets of coefficients in two linear regressions. *Econometrica* 28: 591–605. <https://doi.org/10.2307/1910133>.
- Linden, A. 2015. Conducting interrupted time-series analysis for single- and multiple-group comparisons. *Stata Journal* 15: 480–500.
- . 2017. A comprehensive set of postestimation measures to enrich interrupted time-series analysis. *Stata Journal* 17: 73–88.

Also see

- [TS] [estat sbcusum](#) — Cumulative sum test for parameter stability
- [TS] [estat sbsingle](#) — Test for a structural break with an unknown break date
- [TS] [tsset](#) — Declare data to be time-series data
- [R] [ivregress](#) — Single-equation instrumental-variables regression
- [R] [regress](#) — Linear regression