Freie Universität Bozen
Libera Università di Bolzano
Università Liedia de Bulsan

xtbreak: Testing for structural breaks in Stata

Jan Ditzen[1], Yiannis Karavias[2], Joakim Westerlund[3]

[1]Free University of Bozen-Bolzano, Bozen, Italy
[2]Brunel University, UK
[3]Lund University, Sweden

September 11, 2025

Motivation I

Motivation

- In time series or panel time series structural breaks (or change points) in the relationships between key variables can occur.
- Estimations and forecasts depend on knowledge about structural breaks.
- Structural breaks might influence interpretations and policy recommendations.
- Break can be unknown or known and single and multiple breaks can occur.
- Examples: Financial Crisis, oil price shock, Brexit Referendum, COVID19,...
- Question: Can we estimate when the breaks occur and test them?

Motivation

Literature

- Time Series:
  - Andrews (1993) test for parameter instability and structure change with unknown change point.
  - Bai and Perron (1998) propose three tests for and estimation of multiple change points.
- Panel (Time) Series
  - Karavias et al. (2023); Ditzen et al. (2025b) single and multiple breaks in panel data with cross-section dependence.
- `xtbreak` introduces tests for multiple structural breaks in time series and panel data based on Bai and Perron (1998) and Karavias et al. (2023); Ditzen et al. (2025b).

Econometric Model I

- Static linear panel regression model with s breaks:

$$y_{i,t} = x'_{i,t}\beta + w'_{i,t}\delta_j + e_{i,t} \tag{1}$$

- t denotes the time dimension, t and i the cross-sectional dimension, $n = 1, ..., N$.
- $w_{i,t}$ is a variable affected by a structural break, δ_j is the coefficient of the j-th segment.
- $x_{i,t}$ is a variable unaffected by breaks.
- $e_{i,t}$ is a term of unobservables, including the error term.

Econometric Model II

- We can write the model with s breaks as:

$$
\begin{aligned}
y_{i,t} &= x_{i,t}'\beta + w_{i,t}'\delta_1 + e_{i,t} \text{ for } t = T_0, ..., T_1,\\
y_{i,t} &= x_{i,t}'\beta + w_{i,t}'\delta_2 + e_{i,t} \text{ for } t = T_1, ..., T_2,\\
&\vdots\\
y_{i,t} &= x_{i,t}'\beta + w_{i,t}'\delta_{s+1} + e_{i,t} \text{ for } t = T_s, ..., T_{s+1}.
\end{aligned}
$$

- We can further assume fixed effects or a common factor structure in $e_{i,t}$, $x_{i,t}$ and $w_{i,t}$
- The aim is to i) estimate the number of breaks $\hat{s}$, ii) the location of the breaks, $\hat{T}_j$ and iii) test for breaks.

Testing for Structural Breaks

- Three hypothesis:
  (1) H_0: no breaks versus H_1: s breaks, where the number of breaks under H_1, s is known.
  (2) H_0: no breaks versus H_1: $1 \leq s \leq s_{max}$ breaks, where the maximum number of breaks under H_1, s_{max}.
  (3) H_0: s breaks versus H_1: $s+1$ breaks.
- What to do if number of breaks is unknown?
- Main idea: if the model has the true number of breaks and the true point in time, then the SSR should be smaller than for a model with a larger or smaller number of breaks.
- No further knowledge of the break points required.

Testing for Structural Breaks I

Hypothesis 1: No breaks vs. s breaks

- H_0: no breaks versus H_1: s breaks,
- Simple F-Test:

$$F(\tau_s) = df \, \hat{\delta}' R' \left(R \hat{V}(\hat{\delta}) R' \right)^{-1} \hat{\delta} R$$

- df is the degree of freedom.
- $\hat{\delta}$ is the OLS estimator for δ with s breaks at dates τ_s
- R imposes restrictions such that $R\delta' = \delta_1' - \delta_2', ..., \delta_s' - \delta_{s-1}'$.
- $\hat{V}(\hat{\delta})$ is a variance estimate.
- If tests are known, then $F(\tau_s) \sim F_{s,df}$.

Testing for Structural Breaks II

Hypothesis 1: No breaks vs. s breaks

- If the breakdates τ_s are not known, then a supremum statistic can be used:

$$\text{sup}F(s) = \sup_{\mathcal{T}_s \in \mathcal{T}_{s,\epsilon}} F(\mathcal{T}_s).$$

- with $\mathcal{T}_{s,\epsilon} = \{(T_1, ..., T_s) : T_{j+1} - T_j \geq \epsilon T, T_1 \geq \epsilon T, T_s \leq (1-\epsilon)T\}$. is the set of permissible break dates with ϵ being a user-defined trimming parameter.
- Asymptotic critical values depending on the number of breaks s and regressors q are given in Bai and Perron (1998, Table 1).
- The test statistic is essentially the maximum SSR of all possible SSR combination with

Testing for Structural Breaks

Hypothesis 2: no breaks versus $1 \leq s \leq s_{max}$ breaks

- Test if a maximum of s breaks occurs.
- "Double Maximum" test, where the maximum of the test using hypothesis 1 for the number of breaks between 1 and s is taken.

$$\mathrm{WDmax}F(s_{max}) = \max_{1 \leq s \leq s_{max}} \frac{c_{\alpha,1}}{c_{\alpha,s}} \mathrm{sup}F(s),$$

- $c_{\alpha,s}$ is the critical value from Bai and Perron (1998, Table 1).

Testing for Structural Breaks

Hypothesis 3: s breaks vs. $s+1$ breaks

- Idea: Test if an additional break occurs.

$$F(s+1|s) = \sup_{1 \leq j \leq s+1} \sup_{\tau \in \hat{\mathcal{T}}_{j,\epsilon}} F(\tau|\hat{\mathcal{T}}_s).$$

- It is essentially the difference of the minimum of combinations of the SSR with s and s + 1 breaks.
- Asymptotic critical values depending on the number of breaks s and regressors q are given in Bai and Perron (1998, Table 1).
- Test can be repeated sequentially until hypothesis cannot be rejected any longer to find number of breaks.

Minimising SSR and Dynamic Programming Approach

- Tests are based on minimising the SSR, but how to find the minimal SSR?
- Calculate SSR for all necessary subsamples (Bai and Perron, 2003).
- The estimation for the breakdates is:
  $\hat{\mathcal{T}}_s = \arg\min_{\mathcal{T}_s \in \mathcal{T}_{s,\varepsilon}} SSR(\mathcal{T}_s)$
- Number of possible segments for m breaks and $h = \epsilon T$ is:
  $T(T+1)/2 - (h-1)T + (h-2)(h-1)/2 - h^2 m(m+1)/2$

- For example: Break in period 2 ($T_1 = 2$), then $\text{SSR} = \text{SSR}(1,2) + \text{SSR}(2,T)$.

		End				
		1	2	3	...	T
Start	1	⋱	SSR(1,2)	SSR(1,3)	...	SSR(1,T)
	2		⋱	SSR(2,3)	...	SSR(2,T)
	3			⋱		SSR(3,T)
	⋮				⋱	
	T					⋱

Syntax

- Automatic estimation of number of breaks and break dates:

  xtbreak *depvar* [*indepvars*] [*if*] [, *options1 options2 options3 options5 options6*]

- Known Breaks:

  xtbreak test *depvar* [*indepvars*] [*if*] , *breakpoints(numlist—datelist* [,index|fmt(*string*)] *)* [*options1 options5*]

- Unknown Breaks

  xtbreak test *depvar* [*indepvars*] [*if*] [, hypothesis(1|2|3) breaks(real) *options1 options2 options3 options4 options5*]

Empirical Example

Leader Approval Rating

- Relationship between consumer confidence (CCI) and the approval rating of a country's leader.
- Likely exposed to structural breaks over time.
- Dataset from Ditzen et al. (2025a):
  - 8 countries with monthly data from Jan 1990 - Dec 2021, T = 383
  - Approval Rating from EAP 3.0 Database (Carlin et al., 2023).
  - Consumer confidence from OECD
- Aim: estimate number of breaks and breakpoints in effect of CCI on approval rating ($appr_{i,t}$).
- Add dummy ($d_{i,t}$) for month before, during and past an election. Dummy is non-breaking.

$$appr_{i,t} = \alpha_i + \delta_{i,s} CCI_{i,t} + \beta el_{i,t} + \epsilon_{i,t}$$

```
. xtbreak  d.approval d.CCI , trim(0.05) nobreakvar(ElectionQ)  strict maxbreaks(5) python
Test for multiple breaks at unknown breakdates
(Ditzen, Karavias & Westerlund. 2025)
H0: no break(s) vs. H1: 1 <= s <= 5 break(s)
                 ------------------ Bai & Perron Critical Values ------------------
                      Test       1% Critical      5% Critical     10% Critical
                   Statistic        Value            Value            Value
-------------------------------------------------------------------------------
 UDmax               13.91          13.74            10.17             8.78
-------------------------------------------------------------------------------

Sequential test for multiple breaks at unknown breakpoints
(Ditzen, Karavias & Westerlund. 2025)
                 ------------------ Bai & Perron Critical Values ------------------
                      Test       1% Critical      5% Critical     10% Critical
                   Statistic        Value            Value            Value
-------------------------------------------------------------------------------
 F(1|0)                12.63          13.58             9.63             8.02
 F(2|1)                26.99          15.03            11.14             9.56
 F(3|2)                 5.32          15.62            12.16            10.45
-------------------------------------------------------------------------------
Detected number of breaks:                2                2                2
-------------------------------------------------------------------------------
The detected number of breaks indicates the highest number of
 breaks for which the null hypothesis is rejected.
Estimation of break points
                                           Number of obs        =    3064
                                           Number of Groups     =       8
                                           Obs per group        =     383
                                           SSR                  =  34584.14
                                           Trimming             =      0.05
-------------------------------------------------------------------------------
  #       Index     Date                          [95% Conf. Interval]
-------------------------------------------------------------------------------
  1         344     2018m9                        2017m3          2020m3
  2         363     2020m4                        2019m2          2021m6
-------------------------------------------------------------------------------
```

Empirical Example

Leader Approval Rating

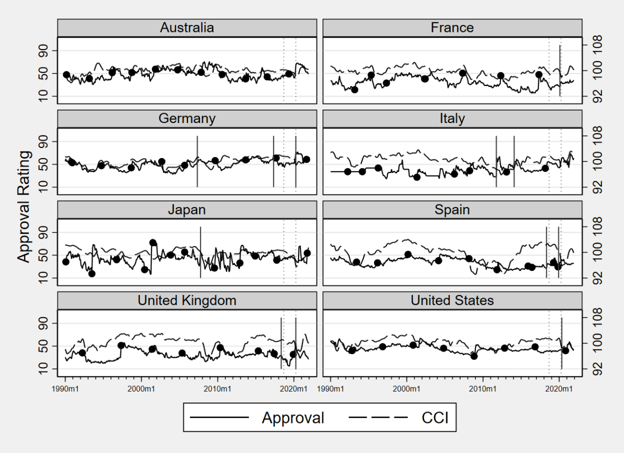

Figure: Leader Confidence. Dashed lines indicate break point estimates on country level, dotted on panel level and dots indicate elections.

Empirical Example

Leader Approval Rating - SSR

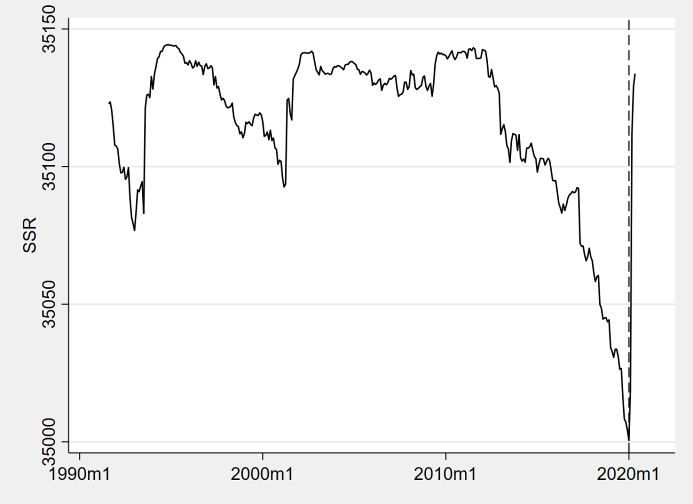

Figure: SSR for a single break over time

Speed

`xtbreak` is slow....

- Estimation and testing of unknown breaks depend on calculation of SSRs for subsamples (segments).
- Number of segments increases with T, trimming ϵ and number of breaks:

		End				
		1	2	3	...	T
Start	1	$\ddots$	SSR(1,2)	SSR(1,3)	...	SSR(1,T)
	2		$\ddots$	SSR(2,3)	...	SSR(2,T)
	3			$\ddots$		SSR(3,T)
	$\vdots$				$\ddots$	
	T					$\ddots$

		T			
	ϵ	20	50	100	500
m=1	0.15	162	912	3,516	85,326
	0.1	186	1,056	4,086	99,426
	0.05	209	1195	4631	112901
m=2	0.15	144	800	3,066	74,076
	0.1	178	1,006	3,886	94,426
	0.05	207	1,182	4,581	111,651

- Even for a moderate $T = 100$ and 2 breaks, more than 3,000 SSRs need to be computed.

Speed

SSR - a closer look

- The main bottle neck in terms of speed are the cacluations of the SSRs.
- For N=50, T=50, `xtbreak` takes 10.6 seconds, the SSRs take 10.5 seconds.
- SSR for segment τ_j can be written as $SSR(\tau_j) = F\left((\sum_{i=1}^{N} \tilde{\mathbf{X}}'_{\tau_j,i} \tilde{\mathbf{X}}_{\tau_j,i})^{-1}, \sum_{i=1}^{N} \tilde{\mathbf{X}}'_{\tau_j,i} \tilde{\mathbf{Y}}_{\tau_j,i} \right)$, where $\tilde{\mathbf{X}}_{\tau_j,i} = \mathbf{X}_{\tau_j,i} \mathbf{M}_{\tau_j,i}$, and $\mathbf{M}_{\tau_j,i}$ is a projection matrix to partial out fixed effects, cross-section averages, etc.
- Four main components:
  1. Inversion of $\tilde{\mathbf{X}}'_{\tau_j,i} \tilde{\mathbf{X}}_{\tau_j,i}$.
  2. Selecting subsamples.
  3. Loop over number of segments.
  4. Loop over number of cross-sections.

Speed

Speed improvements

- On 1): Matrix inversion in mata fast.
- On 2): mix of `panelsetup` and range subscripts *relatively* fast.
- On 3) and 4): `xtbreak` has an option `python` which invokes Python:
    - Vecorisation of $\mathbf{X}$ and $\mathbf{Y}$ as a 3D Arrays, $T \times K \times N$ matrices.
    - Loop over segments can be parallised.
    - "Fixed Costs" invoking Python disadvantage

Speed

Speed improvements - benchmarks

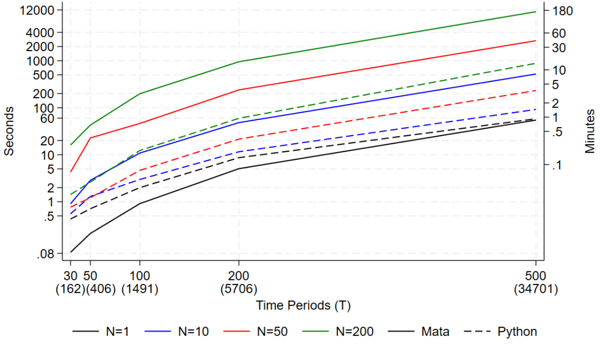

Figure: Timings of SSR calculations. Hypothesis (2) with no break vs. up to 4 breaks, two breaking variables and one non-breaking variable. Number of possible segments in parenthesis. System: Stata 19 SE; Core i7-10610U; 4 Cores.

- Time series (N=1) Mata is *marginally* faster.
- Panel Data ($N > 1$) vectorization in Python reduces computation time from 180 minutes to 17 minutes, representing a 10.6-fold speed improvement.

Conclusion

- `xtbreak` allows for the estimation of number and location of breaks and 3 tests.
- Supports time series and (unbalanced) panel data.
- Allows for
  - fixed effects, unobserved common
  - different variance estimator and trimmings
  - non-breaking variables
  - post estimation functions
- Speed improvements using Python.
- Paper (open access!) in Stata Journal 25(3).

https://janditzen.github.io/xtbreak/

The Stata Journal 25(3)

References I

Andrews, D. W. K. 1993. Tests for Parameter Instability and Structural Change With Unknown Change Point. Econometrica 61(4): 821–856.

Bai, B. Y. J., and P. Perron. 1998. Estimating and Testing Linear Models with Multiple Structural Changes. Econometrica, 66(1): 47–78.

Bai, J., and P. Perron. 2003. Computation and analysis of multiple structural change models. Journal of Applied Econometrics 18(1): 1–22.

Carlin, R., J. Hartlyn, T. Hellwig, G. Love, C. Martinez-Gallardo, and M. Singer. 2023. Executive Approval Database 3.0. Available for download at `www.executiveapproval.org`.

Ditzen, J., Y. Karavias, and J. Westerlund. 2025a. Testing and estimating structural breaks in time series and panel data in Stata. The Stata Journal 25(3): 526–560.

References II

———. 2025b. Multiple Structural Breaks in Interactive Effects Panel Data Models. Journal of Applied Econometrics 40(1): 74–88. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.3097.

Karavias, Y., P. K. Narayan, and J. Westerlund. 2023. Structural Breaks in Interactive Effects Panels and the Stock Market Reaction to COVID-19. Journal of Business & Economic Statistics 41(3): 653–666.