Package 'mbreaks'

November 11, 2024

Type Package

Title Estimation and Inference for Structural Breaks in Linear Regression Models

Version 1.0.1

Description
Functions provide comprehensive treatments for estimating, inferring, testing and model selecting in linear regression models with structural breaks. The tests, estimation methods, inference and information criteria implemented are discussed in Bai and Perron (1998) ``Estimating and Testing Linear Models with Multiple Structural Changes" <doi:10.2307/2998540>.

URL https://github.com/RoDivinity/mbreaks

BugReports https://github.com/RoDivinity/mbreaks/issues

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

Depends R (>= 3.3)

Imports ggplot2

Suggests knitr, rmarkdown, graphics, testthat

VignetteBuilder knitr, rmarkdown

NeedsCompilation no

Author Linh Nguyen [aut, cre],
Yohei Yamamoto [aut],
Pierre Perron [aut]

Maintainer Linh Nguyen <nguye535@purdue.edu>

Repository CRAN

Date/Publication 2024-11-11 11:40:02 UTC

Contents

`compile_model` *Format output of n break model*

Description

`compile_model()` compiles the information of `model` class object `x` into three main tables:

date_tab Table for estimated break date in the model with 90% and 95% confidence intervals based on `robust`,`hetomega`, `hetq` options for errors and `prewhit` option.

RS_tab Table for estimated coefficients for `z` regressors with corrected standard errors based on `robust`,`hetdat`,`hetvar` options for errors and `prewhit` option.

FS_tab Table for estimated coefficients for `x` regressors with corrected standard errors based on `robust`, `hetdat`, and `hetvar` options for errors and `prewhit` option.

Usage

```
compile_model(x, digits = 3)
```

Arguments

x	the model class to format
digits	number of digits displayed in console. Default value is 3

Value

Formatted x with the following appended tables:

'date_tab' A data frame storing the break date estimated by the model, and their corresponding confidence intervals.

'RS_tab' A data frame storing the estimated coefficients which allowed to change across regimes with corrected standard errors.

'FS_tab' A data frame storing the estimated coefficients which is constant across regimes with corrected standard errors.

Note

- If x returns 0 number of estimated break, the function will return NULL value instead of the list in Value.
- If x is a pure structural break, the 'FS_tab' will return NULL in Value.

---

compile_sbtests *Compile the Output of Sup Wald Test*

Description

'compile_sbtests' formats the output of 'sbtests' into two tables.

Usage

```
compile_sbtests(x, digits = 3)
```

Arguments

x	An 'sbtests' class object.
digits	The number of decimal places displayed.

Value

A modified ‘sbtests‘ object, ‘x‘, with two appended data frames:

supF1 A data frame containing SupF test statistics for testing 0 versus m breaks, where m is the maximum number of breaks considered in ‘x‘. It includes critical values at the *10%, 5%, 2.5%, and 1%* levels.

UDMax A data frame containing Double Max test statistics with critical values at the *10%, 5%, 2.5%, and 1%* levels.

`compile_seqtests` *Compile the output of sequential Sup Wald test*

Description

‘compile_seqtests‘ formats the output of the ‘seqtests‘ class object to 1 table

‘sfl‘ A table containing sequential sup F tests statistics of ‘l‘ versus ‘l+1‘ for ‘l‘ in ‘1‘ to ‘m‘ breaks with critical values of the corresponding tests at *1%, 2.5%, 5%, and 10%* significance levels.

Usage

```
compile_seqtests(x)
```

Arguments

`x`	‘seqtests‘ class object

Value

class ‘seqtests‘ list ‘x‘ with appended data frame ‘sfl‘ containing the sequential SupF test statistics with critical values at *10%, 5%, 2.5%, and 1%* level.

`correct` *Heteroskedasticy and autocorrelation consistency correction for residuals*

Description

‘hac()‘ corrects the estimated errors based on options of prewhitening using a AR(1) process estimation of error terms to obtain heteroskedasticy and autocorrelation consistency (HAC) errors with automatic bandwith and kernel similar to Andrews, 1994

Usage

```
correct(reg, res, prewhit)
```

Arguments

reg	matrix of regressors
res	matrix of estimated residuals
prewhit	Option of using prewhitening process. If 1, an AR(1) process will be used to filter. If 0, skipped the filtering process

Value

hac Heteroskedasticy and autocorrelation consistent errors

---

dating *Computation of global minimizer for pure structural change model*

Description

'dating()' computes break points that globally minimizes SSR via dynamic programming approach. To avoid recursion depth increases as number of breaks in the model increases, a temporary array is used to store optimal partition with corresponding SSR for all permissible subsamples for all 1:m-1 breaks. For the m-th break, the problem becomes finding where to insert the last feasible m+1-th segment into the sample partitioned by m-1 breaks to obtain minimum SSR over the sample

Usage

```
dating(y, z, h, m, q, bigT)
```

Arguments

y	matrix of dependent variable
z	matrix of regressors with coefficients allowed to change across regimes
h	minimum length of segment
m	maximum number of breaks
q	number of 'z' regressors
bigT	sample period T

Value

A list containing the following components:

glb	minimum global SSR
datevec	Vector of dates (optimal minimizers)
bigvec	Associated SSRs

diag_par *Diagonal partition given break dates*

Description

'diag_par()' partition the matrix of 'z' regressors which coefficients are changed based on the provided break dates

Usage

```
diag_par(input, m, date)
```

Arguments

input	matrix of independent variables z with coefficients allowed to change overtime
m	number of breaks in the series
date	vector of break dates

Value

output: matrix of partitioned variables corresponds to break dates

Examples

```
z = matrix(c(1:100),50,2)
m = 2  #2 breaks
date = matrix(c(15,30),2,1) #first break at t = 15; second break at t = 30
diag_par(z,m,date)
```

dofix *Estimate a model with pre-specified number of breaks*

Description

'dofix()' compute a structural change model with pre-specified number of breaks.

Usage

```
dofix(
  y_name,
  z_name = NULL,
  x_name = NULL,
  data,
  fixn = 5,
  eps = 1e-05,
```

```
  eps1 = 0.15,
  maxi = 10,
  fixb = 0,
  betaini = 0,
  printd = 0,
  prewhit = 1,
  robust = 1,
  hetdat = 1,
  hetvar = 1,
  hetq = 1,
  hetomega = 1,
  const = 1,
  h = NULL
)
```

Arguments

y_name	name of dependent variable in the data set
z_name	name of independent variables in the data set which coefficients are allowed to change across regimes. default is vector of 1 (Mean-shift model)
x_name	name of independent variables in the data set which coefficients are constant across regimes. default is 'NULL'
data	name of data set used
fixn	number of breaks specified
eps	convergence criterion for iterative recursive computation
eps1	value of trimming (in percentage) for the construction and critical values. Minimal segment length 'h' will be set at default = int(eps1*T) (T is total sample size). • eps1 = 0.05 Maximal value of m = 10. • eps1 = 0.10 Maximal value of m = 8. • eps1 = 0.15 Maximal value of m = 5. • eps1 = 0.20 Maximal value of m = 3. • eps1 = 0.25 Maximal value of m = 2. • eps1=0 This option allows users to explicitly specify minimum segment length 'h' parameters.
maxi	maximum number of iterations
fixb	option to use fixed initial input β. If 1, the model will use values given in betaini. If 0, betaini is skipped
betaini	Initial $beta_0$ to use in estimation
printd	Print option for model estimation. default = 0, to suppress intermediate outputs printing to console
prewhit	set to 1 to apply AR(1) prewhitening prior to estimating the long run covariance matrix.

robust	set to 1 to allow for heterogeneity and autocorrelation in the residuals, 0 otherwise. The method used is *Andrews(1991)* automatic bandwidth with AR(1) approximation with quadratic kernel. Note: Do not set to 1 if lagged dependent variables are included as regressors.
hetdat	option for the construction of the F tests. Set to 1 if want to allow different moment matrices of the regressors across segments. If hetdat = 0, the same moment matrices are assumed for each segment and estimated from the ful sample. It is recommended to set hetdat=1 if number of regressors x > 0.
hetvar	option for the construction of the F tests.Set to 1 if users want to allow for the variance of the residuals to be different across segments. If hetvar=0, the variance of the residuals is assumed constant across segments and constructed from the full sample. hetvar=1 when robust =1)
hetq	used in the construction of the confidence intervals for the break dates. If hetq=0, the moment matrix of the data is assumed identical across segments
hetomega	used in the construction of the confidence intervals for the break dates. If hetomega=0, the long run covariance matrix of zu is assumed identical across segments (the variance of the errors u if robust=0).
const	indicates whether the regression model include an intercept changing across regimes. Default value is 1
h	Minimum segment length of regime considered in estimation. If users want to specify a particular value, please set 'eps1=0'

Value

out A list of class 'model' contains all information about the estimated structural change model with 'fixn' breaks

Examples

```
dofix('rate',data=real,fixn=3)
```

---

doglob *Global SSR minimizer for structural change model*

Description

'doglob()' identify if the structural change model is i) pure or ii) partial change model. The procedure then calls appropriate functions dating to estimate the pure change model and nldat to estimate the partial change model.

Usage

```
doglob(y, z, x, m, eps, h, maxi, fixb, betaini, printd, eps1)
```

Arguments

y	matrix of dependent variable
z	matrix of independent variables with coefficients allowed to change across regimes
x	matrix of independent variables with coefficients constant across regimes
m	number of breaks in the structural change model
eps	convergence criterion for iterative recursive computation. (For partial change model ONLY)
h	Minimum segment length of regime considered in estimation. If users want to specify a particular value, please set 'eps1=0'
maxi	maximum number of iterations. (For partial change model ONLY)
fixb	option to use fixed initial input β. If 1, the model will use values given in betaini. If 0, betaini is skipped
betaini	Initial $beta_0$ to use in estimation (Must be a 'p x 1' matrix, where p is number of x variables)
printd	Print option for model estimation. default = 0, to suppress intermediate outputs printing to console
eps1	trimming level

Value

A list containing the following components:

glb Minimum global SSR.

datevec Vector of dates (optimal minimizers).

bigvec Associated SSRs with possible break dates combination.

---

doorder *Estimating number of breaks via information criterion*

Description

'doorder()' estimates the number of breaks using one of the following information criteria:

- modified Bayesian information criterion by Kurozumi and Tuvaandorj, 2011,
- modified Schwarz information criterion by Liu, Wu and Zidek, 1997,
- Bayesian information criterion by Yao, 1988

and the structural break model corresponding to estimated number of breaks.

Usage

```
doorder(
  y_name,
  z_name = NULL,
  x_name = NULL,
  data,
  m = 5,
  eps = 1e-05,
  eps1 = 0.15,
  maxi = 10,
  fixb = 0,
  betaini = 0,
  printd = 0,
  ic = "KT",
  const = 1,
  h = NULL,
  prewhit = 1,
  hetdat = 1,
  hetq = 1,
  hetomega = 1,
  hetvar = 1,
  robust = 1
)
```

Arguments

y_name	name of dependent variable in the data set
z_name	name of independent variables in the data set which coefficients are allowed to change across regimes. default is vector of 1 (Mean-shift model)
x_name	name of independent variables in the data set which coefficients are constant across regimes. default is 'NULL'
data	name of data set used
m	maximum number of breaks
eps	convergence criterion for iterative recursive computation
eps1	value of trimming (in percentage) for the construction and critical values. Minimal segment length 'h' will be set at default = int(eps1*T) (T is total sample size). There are five options: • 'eps1=0.05' Maximal value of 'm' = 10. • 'eps1=0.10' Maximal value of 'm' = 8. • 'eps1=.15' Maximal value of 'm' = 5. • 'eps1=.20' Maximal value of 'm' = 3. • 'eps1=.25' Maximal value of 'm' = 2. • 'eps1=0' This option allows users to explicitly specify minimum segment length 'h' parameters
maxi	maximum number of iterations

fixb	option to use fixed initial input β. If 1, the model will use values given in betaini. If 0, betaini is skipped
betaini	Initial $beta_0$ to use in estimation
printd	Print option for model estimation. default = 0, to suppress intermediate outputs printing to console
ic	indicator which information criterion is used in selecting number of breaks: • KT • BIC • LWZ The default value is KT
const	indicates whether the regression model include an intercept changing across regimes. Default value is 1
h	Minimum segment length of regime considered in estimation. If users want to specify a particular value, please set 'eps1=0'
prewhit	set to 1 to apply AR(1) prewhitening prior to estimating the long run covariance matrix.
hetdat	option for the construction of the F tests. Set to 1 if want to allow different moment matrices of the regressors across segments. If hetdat = 0, the same moment matrices are assumed for each segment and estimated from the ful sample. It is recommended to set hetdat=1 if number of regressors x > 0.
hetq	used in the construction of the confidence intervals for the break dates. If hetq=0, the moment matrix of the data is assumed identical across segments
hetomega	used in the construction of the confidence intervals for the break dates. If hetomega=0, the long run covariance matrix of zu is assumed identical across segments (the variance of the errors u if robust=0)
hetvar	option for the construction of the F tests.Set to 1 if users want to allow for the variance of the residuals to be different across segments. If hetvar=0, the variance of the residuals is assumed constant across segments and constructed from the full sample. hetvar=1 when robust =1)
robust	set to 1 to allow for heterogeneity and autocorrelation in the residuals, 0 otherwise. The method used is *Andrews(1991)* automatic bandwidth with AR(1) approximation with quadratic kernel. Note: Do not set to 1 if lagged dependent variables are included as regressors.

Value

A list of class 'model' that contains one of the following:

mBIC change model with number of breaks selected by BIC

mLWZ change model with number of breaks selected by LWZ

mKT change model with number of breaks selected by KT

References

Liu J, Wu S, Zidek JV (1997). *"On Segmented Multivariate Regressions"*, Statistica Sinica, 7, 497-525. Yao YC (1988). *"Estimating the Number of Change-points via Schwartz Criterion"*, Statistics and Probability Letters, 6, 181-189. Kurozumi E, Tuvaandorj P (2011). *"Model Selection Criteria in Multivariate Models with Multiple Structural Changes"*, Journal of Econometrics 164, 218-238.

Examples

```
doorder('rate',data=real,ic=c('BIC'))
```

`dorepart` *Estimating number of breaks using repartition procedure*

Description

'dorepart()' computes the repartition estimates of the breaks obtained by the sequential method by Bai, 1995. It allows estimates that have the same asymptotic distribution as those obtained by global minimization. Otherwise, the output from the procedure "estim" below does not deliver asymptotically correct confidence intervals for the break dates.

Usage

```
dorepart(
  y_name,
  z_name = NULL,
  x_name = NULL,
  data,
  m = 5,
  eps = 1e-05,
  eps1 = 0.15,
  maxi = 10,
  fixb = 0,
  betaini = 0,
  printd = 0,
  prewhit = 1,
  robust = 1,
  hetdat = 1,
  hetvar = 1,
  const = 1,
  signif = 2
)
```

Arguments

y_name	name of dependent variable in the data set
z_name	name of independent variables in the data set, whose coefficients are allowed to change across regimes. default is a vector of 1 (Mean-shift model).
x_name	name of independent variables in the data set whose coefficients are constant across regimes. default is NULL.
data	name of the data set used
m	Maximum number of structural changes allowed. If not specified, m will be set to default value matching the eps1 input
eps	convergence criterion for iterative recursive computation
eps1	value of trimming (in percentage) for the construction and critical values. Minimal segment length h will be set at default = int(eps1 * T) (T is total sample size). There are five options: • eps1 = 0.05 Maximal value of m = 10. • eps1 = 0.10 Maximal value of m = 8. • eps1 = 0.15 Maximal value of m = 5. • eps1 = 0.20 Maximal value of m = 3. • eps1 = 0.25 Maximal value of m = 2. • eps1 = 0 This option is not allowed.
maxi	maximum number of iterations
fixb	option to use fixed initial input β. If 1, the model will use values given in betaini. If 0, betaini is skipped
betaini	Initial β_0 to use in estimation
printd	Print option for model estimation. default = 0, to suppress intermediate outputs printing to console
prewhit	set to 1 to apply AR(1) prewhitening prior to estimating the long run covariance matrix.
robust	set to 1 to allow for heterogeneity and autocorrelation in the residuals, 0 otherwise. The method used is *Andrews(1991)* automatic bandwidth with AR(1) approximation with quadratic kernel. Note: Do not set to 1 if lagged dependent variables are included as regressors.
hetdat	option for the construction of the F tests. Set to 1 if you want to allow different moment matrices of the regressors across segments. If hetdat = 0, the same moment matrices are assumed for each segment and estimated from the full sample. It is recommended to set hetdat = 1 if number of regressors x > 0.
hetvar	option for the construction of the F tests. Set to 1 if users want to allow for the variance of the residuals to be different across segments. If hetvar = 0, the variance of the residuals is assumed constant across segments and constructed from the full sample. hetvar = 1 when robust = 1
const	indicates whether the regression model includes an intercept changing across regimes. Default value is 1
signif	significance level used to sequential test to select number of breaks.

- 4: 1% level
- 3: 2.5% level
- 2: 5% level
- 1: 10% level

Value

A list of class `model` for the structural break model estimated by the repartition procedure.

References

Bai, J. 1995, *"Estimating Breaks One at a Time"*, Econometric Theory, 13, 315-352

Examples

```
dorepart('inf', 'inflag', 'inffut', data = nkpc)
```

`doseqtests` *Sequential Sup F tests*

Description

'doseqtests()' computes the sequential sup F tests of l versus l+1 for l from 1 to m with each corresponding null hypothesis of maximum number of break is l and alternative hypothesis is l+1. The l breaks under the null hypothesis are taken from the global minimization estimation

Usage

```
doseqtests(
  y_name,
  z_name = NULL,
  x_name = NULL,
  data,
  m = 5,
  eps = 1e-05,
  eps1 = 0.15,
  maxi = 10,
  fixb = 0,
  betaini = 0,
  printd = 0,
  prewhit = 1,
  robust = 1,
  hetdat = 1,
  hetvar = 1,
  hetq = 1,
  hetomega = 1,
  const = 1
)
```

Arguments

y_name	name of dependent variable in the data set
z_name	name of independent variables in the data set which coefficients are allowed to change across regimes. default is a vector of 1 (Mean-shift model)
x_name	name of independent variables in the data set which coefficients are constant across regimes. default is 'NULL'
data	name of data set used
m	maximum number of breaks
eps	convergence criterion for recursive calculations (For partial change model ONLY)
eps1	value of trimming (in percentage) for the construction and critical values. Minimal segment length 'h' will be set at default = int(eps1*T) (T is total sample size). There are five options: • 'eps1=0.05' Maximal value of 'm' = 10. • 'eps1=0.10' Maximal value of 'm' = 8. • 'eps1=.15' Maximal value of 'm' = 5. • 'eps1=.20' Maximal value of 'm' = 3. • 'eps1=.25' Maximal value of 'm' = 2. • 'eps1=0' is not allowed. The test is undefined for no trimming level.
maxi	number of maximum iterations for recursive calculations of finding global minimizers.default = 10 (For partial change model ONLY)
fixb	option to use fixed initial input β. If 1, the model will use values given in betaini. If 0, betaini is skipped
betaini	Initial $beta_0$ to use in estimation (Must be a 'p x 1' matrix, where p is number of x variables)
printd	Print option for model estimation. default = 0, to suppress intermediate outputs printing to console
prewhit	set to 1 to apply AR(1) prewhitening prior to estimating the long run covariance matrix.
robust	set to 1 to allow for heterogeneity and autocorrelation in the residuals, 0 otherwise. The method used is *Andrews(1991)* automatic bandwidth with AR(1) approximation with quadratic kernel. Note: Do not set to 1 if lagged dependent variables are included as regressors.
hetdat	option for the construction of the F tests. Set to 1 if want to allow different moment matrices of the regressors across segments. If hetdat = 0, the same moment matrices are assumed for each segment and estimated from the ful sample. It is recommended to set hetdat=1 if number of regressors x > 0.
hetvar	option for the construction of the F tests.Set to 1 if users want to allow for the variance of the residuals to be different across segments. If hetvar=0, the variance of the residuals is assumed constant across segments and constructed from the full sample. hetvar=1 when robust =1)
hetq	used in the construction of the confidence intervals for the break dates. If hetq=0, the moment matrix of the data is assumed identical across segments

`hetomega`	used in the construction of the confidence intervals for the break dates. If `hetomega=0`, the long run covariance matrix of zu is assumed identical across segments (the variance of the errors u if `robust=0`)
`const`	indicates whether the regression model include an intercept changing across regimes. Default value is 1

Value

A list that contains following:

supfl SupF(l+1|l) test statistics.

cv Critical values for SupF(l+1|l) test.

Examples

```
doseqtests('inf',c('inflag','lbs','inffut'),data=nkpc,prewhit=0)
```

`dosequa` *Estimating number of breaks using sequential tests*

Description

'dosequa()' sequentially increases the number of breaks from '1' to 'm' until the sequential tests reject and estimate the structural change model with corresponding estimated breaks. The procedure is proposed by Bai and Perron, 1998.

Usage

```
dosequa(
  y_name,
  z_name = NULL,
  x_name = NULL,
  data,
  m = 5,
  eps = 1e-05,
  eps1 = 0.15,
  maxi = 10,
  fixb = 0,
  betaini = 0,
  printd = 0,
  prewhit = 1,
  robust = 1,
  hetdat = 1,
  hetvar = 1,
  hetq = 1,
  hetomega = 1,
```

```
  const = 1,
  signif = 2
)
```

Arguments

y_name	name of dependent variable in the data set
z_name	name of independent variables in the data set, which coefficients are allowed to change across regimes. Default value is vector of 1 (Mean-shift model).
x_name	name of independent variables in the data set, which coefficients are constant across regimes. Default value is NULL.
data	name of the data set used
m	maximum number of breaks
eps	convergence criterion for iterative recursive computation
eps1	value of trimming (in percentage) for the construction and critical values. Minimal segment length 'h' will be set at default value = int(eps1 * T) (T is total sample size). There are five options: • eps1 = 0.05 Maximal value of m = 10. • eps1 = 0.10 Maximal value of m = 8. • eps1 = 0.15 Maximal value of m = 5. • eps1 = 0.20 Maximal value of m = 3. • eps1 = 0.25 Maximal value of m = 2. • eps1 = 0 This option is not allowed.
maxi	maximum number of iterations
fixb	option to use fixed initial input β. If 1, the model will use values given in betaini. If 0, betaini is skipped
betaini	Initial β_0 to use in estimation
printd	Print option for model estimation. default = 0, to suppress intermediate outputs printing to console
prewhit	set to 1 to apply AR(1) prewhitening prior to estimating the long run covariance matrix.
robust	set to 1 to allow for heterogeneity and autocorrelation in the residuals, 0 otherwise. The method used is *Andrews(1991)* automatic bandwidth with AR(1) approximation with quadratic kernel. Note: Do not set to 1 if lagged dependent variables are included as regressors.
hetdat	option for the construction of the F tests. Set to 1 if you want to allow different moment matrices of the regressors across segments. If hetdat = 0, the same moment matrices are assumed for each segment and estimated from the full sample. It is recommended to set hetdat = 1 if number of regressors x > 0.
hetvar	option for the construction of the F tests. Set to 1 if users want to allow for the variance of the residuals to be different across segments. If hetvar = 0, the variance of the residuals is assumed constant across segments and constructed from the full sample. hetvar = 1 when robust = 1

`hetq`	used in the construction of the confidence intervals for the break dates. If `hetq = 0`, the moment matrix of the data is assumed identical across segments.
`hetomega`	used in the construction of the confidence intervals for the break dates. If `hetomega = 0`, the long run covariance matrix of zu is assumed identical across segments (the variance of the errors u if `robust = 0`).
`const`	indicates whether the regression model includes an intercept changing across regimes. Default value is 1
`signif`	significance level used in the sequential test to select number of breaks. • 4: 1% level • 3: 2.5% level • 2: 5% level • 1: 10% level

Value

A list of 'model' class with the number of breaks selected by sequential tests.

Examples

```
dosequa('rate', data = real, signif = 1)
```

`dotest` *SupF, UDMax & WDMax testing procedure*

Description

'dotest()' compute the test statistics and report the critical values of the 2 main supF tests below:

- SupF test of 0 vs m breaks
- Double Max test proposed by Perron and Bai, 1998

Usage

```
dotest(
  y_name,
  z_name = NULL,
  x_name = NULL,
  data,
  m = 5,
  eps = 1e-05,
  eps1 = 0.15,
  maxi = 10,
  fixb = 0,
  betaini = 0,
  printd = 0,
```

```
  prewhit = 1,
  robust = 1,
  hetdat = 1,
  hetvar = 1,
  hetq = 1,
  hetomega = 1,
  const = 1
)
```

Arguments

y_name	matrix of dependent variable
z_name	matrix of regressors which coefficients are allowed to change across regimes.
x_name	matrix of regressors which coefficients are constant across regimes.
data	the data set for estimation
m	maximum number of breaks
eps	convergence criterion for iterative recursive computation
eps1	trimming level
maxi	maximum number of iterations
fixb	option to use fixed initial input β. If 1, the model will use values given in betaini. If 0, betaini is skipped
betaini	Initial $beta_0$ to use in estimation (Must be a 'p x 1' matrix, where p is number of x variables)
printd	Print option for model estimation. default = 0, to suppress intermediate outputs printing to console
prewhit	option to use AR(1) for prewhitening
robust	set to 1 to allow for heterogeneity and autocorrelation in the residuals, 0 otherwise. The method used is *Andrews(1991)* automatic bandwidth with AR(1) approximation with quadratic kernel. Note: Do not set to 1 if lagged dependent variables are included as regressors.
hetdat	option for the construction of the F tests. Set to 1 if want to allow different moment matrices of the regressors across segments. If hetdat = 0, the same moment matrices are assumed for each segment and estimated from the ful sample. It is recommended to set hetdat=1 if number of regressors x > 0.
hetvar	option for the construction of the F tests.Set to 1 if users want to allow for the variance of the residuals to be different across segments. If hetvar=0, the variance of the residuals is assumed constant across segments and constructed from the full sample. hetvar=1 when robust =1)
hetq	used in the construction of the confidence intervals for the break dates. If hetq=0, the moment matrix of the data is assumed identical across segments
hetomega	used in the construction of the confidence intervals for the break dates. If hetomega=0, the long run covariance matrix of zu is assumed identical across segments (the variance of the errors u if robust=0)
const	indicates whether the regression model include an intercept changing across regimes. Default value is 1

Value

A list that contains following:

ftest SupF test of 0 vs m (1 to maximum) breaks statistics

cv_supF Critical values for Sup F test

cv_Dmax Critical values for Double Max test

supF1 table summarizing the SupF test (for viewing purposes)

UDMax table summarizing the Double Max test (including UDMax statistics and CVs)

`estim` *Structural change model estimation*

Description

'estim()' estimates the structural change model by OLS given specified vector of break dates It also computes and reports confidence intervals for the break dates based on asymptotic distributions of break date and corrected standard errors of coefficients estimates given the structure of covariance matrix for model errors by specifying error options 'robust', 'hetomega', 'hetq', 'hetdat' and 'hetvar'

Usage

```
estim(m, q, z, y, b, robust, prewhit, hetomega, hetq, x, p, hetdat, hetvar)
```

Arguments

`m`	number of breaks
`q`	number of 'z' regressors z
`z`	matrix of regressors with coefficients are allowed to change across regimes
`y`	matrix of dependent variable
`b`	vector of break dates
`robust`, `hetomega`, `hetq`, `hetdat`, `hetvar`	options for assumptions on the error terms. For more details, please refer to `mdl`.
`prewhit`	option to use prewhitening process based on AR(1) approximation
`x`	matrix of regressors with coefficients are constant across regimes
`p`	number of regressors x

Value

A list containing the following components:

date List of estimated breaks.

CI List of Confidence Intervals for each corresponding break.

beta Estimated coefficients of the regression. The first (m+1)*q are coefficients of q variables z that change across regimes. The last p are coefficients of p variables x that are constant across regimes.

SE Corrected standard errors for the coefficients' estimates

`interval` *Estimatd break confidence interval*

Description

'interval()' computes confidence intervals for the break dates based on approximating the limiting distribution of the break date following the "shrinking shifts" asymptotic framework

Usage

```
interval(y, z, zbar, b, q, m, robust, prewhit, hetomega, hetq, x, p)
```

Arguments

`y`	matrix of dependent variable
`z`	matrix of independent variables with coefficients allowed to change across regimes
`zbar`	partitioned matrix of independent variables with coefficients allowed to change across regimes according to break date vector 'b'
`b`	vector of break breaks
`q`	number of 'z' regressors
`m`	maximum number of breaks
`robust`	set to `1` to allow for heterogeneity and autocorrelation in the residuals, `0` otherwise. The method used is Andrews(1991) automatic bandwidth with AR(1) approximation with quadratic kernel. Note: Do not set to `1` if lagged dependent variables are included as regressors.
`prewhit`	Option of using prewhitening process. If `1`, an AR(1) process will be used to filter. If `0`, skipped the filtering process
`hetomega`, `hetq`	options for assumptions of error terms structure. For more details, refers to [mdl()]
`x`	matrix of independent variables with coefficients constant across regimes
`p`	number of 'x' regressors

Value

bound Confidence intervals of break date in `90%` and `95%` significant level

mdl *Comprehensive structural change estimation and testing*

Description

'mdl()' calls main functions of the 'mbreaks' package to execute the following estimation procedures:

'dotest()' Function `dotest` conducts Sup F tests of '0' versus 'm' breaks and Double Max tests.

'doseqtests()' Function `doseqtests` conducts the sequential Sup F tests of 'l' versus 'l+1' breaks.

'doorder()' Function `doorder` conducts the number of breaks selection from '1' to 'm' breaks using the following information critera: `KT`,`BIC`, and `LWZ`.

'dosequa()' Function `dosequa` conducts the number of breaks selection by sequential tests from '1' to 'm' breaks using sequential Sup F tests.

'dofix()' Function `dofix` conducts structural break model estimation with 'fixn' breaks.

All the procedures automatically identify if the 'model' is either i) pure structural breaks model or ii) partial structural breaks model

Usage

```
mdl(
  y_name,
  z_name = NULL,
  x_name = NULL,
  data,
  eps1 = 0.15,
  m = 5,
  prewhit = 1,
  robust = 1,
  hetdat = 1,
  hetvar = 1,
  hetomega = 1,
  hetq = 1,
  maxi = 10,
  eps = 1e-05,
  fixn = -1,
  fixb = 0,
  betaini = 0,
  printd = 0,
  const = 1,
  signif = 2,
  h = NULL
)
```

Arguments

y_name	name of dependent variable in the data set.
z_name	name of independent variables in the data set which coefficients are allowed to change across regimes. default is vector of 1 (Mean-shift model).
x_name	name of independent variables in the data set which coefficients are constant across regimes. default is NULL.
data	the data set for estimation.
eps1	value of trimming (in percentage) for the construction and critical values. Minimal segment length 'h' will be set at default = int(eps1*T) (T is total sample size). • eps1 = 0.05 Maximal value of m = 10. • eps1 = 0.10 Maximal value of m = 8. • eps1 = 0.15 Maximal value of m = 5. • eps1 = 0.20 Maximal value of m = 3. • eps1 = 0.25 Maximal value of m = 2. • eps1 = 0 This option allows users to explicitly specify minimum segment length 'h' parameters. However, this option will not be allowed for testing and testing related functions. The default value is set at eps1 = 0.15.
m	Maximum number of structural changes allowed. If not specify, m will be set to default value matching 'eps1' input.
prewhit	set to 1 to apply AR(1) prewhitening prior to estimating the long run covariance matrix.
robust	set to 1 to allow for heterogeneity and autocorrelation in the residuals, 0 otherwise. The method used is Andrews(1991) automatic bandwidth with AR(1) approximation with quadratic kernel. Note: Do not set to 1 if lagged dependent variables are included as regressors.
hetdat	option for the construction of the F tests. Set to 1 if want to allow different moment matrices of the regressors across segments. If hetdat = 0, the same moment matrices are assumed for each segment and estimated from the ful sample. It is recommended to set hetdat=1 if number of regressors x > 0.
hetvar	option for the construction of the F tests.Set to 1 if users want to allow for the variance of the residuals to be different across segments. If hetvar=0, the variance of the residuals is assumed constant across segments and constructed from the full sample. hetvar=1 when robust =1)
hetomega	used in the construction of the confidence intervals for the break dates. If hetomega=0, the long run covariance matrix of zu is assumed identical across segments (the variance of the errors u if robust=0)
hetq	used in the construction of the confidence intervals for the break dates. If hetq=0, the moment matrix of the data is assumed identical across segments
maxi	number of maximum iterations for recursive calculations of finding global minimizers.default = 10 (For partial change model ONLY).
eps	convergence criterion for recursive calculations (For partial change model ONLY)

`fixn`	number of pre-specified breaks. `default` = -1. It will be replaced automatically to 2 if no specification is given (For partial change model ONLY)
`fixb`	option to use fixed initial input β. If `1`, the model will use values given in `betaini`. If `0`, betaini is skipped
`betaini`	Initial β_0 to use in estimation (Must be a 'p x 1' matrix, where 'p' is number of x variables)
`printd`	Print option for model estimation. `default` = 0, to suppress intermediate outputs printing to console
`const`	indicates whether the regression model include an intercept changing across regimes. Default value is 1.
`signif`	significance level used to sequential test to select number of breaks. • 4: 1% level • 3: 2.5% level • 2: 5% level • 1: 10% level
`h`	Minimum segment length of regime considered in estimation. If users want to specify a particular value, please set 'eps1=0'

Value

A list that contains the following:

sbtests A list of class 'sbtests' representing Sup F tests of 0 versus m breaks and Double Max tests.

seqtests A list of class 'seqtests' representing sequential Sup F test of l versus l+1 breaks.

BIC A list of class 'model' with structural break model estimated by number of breaks by `BIC` criterion.

LWZ A list of class 'model' with structural break model estimated by number of breaks by `LWZ` criterion.

KT A class 'model' with structural break model estimated by number of breaks by `KT` criterion.

sequa A class 'model' with structural break model estimated by number of breaks by sequential tests.

fix A class 'model' with structural break model estimated by pre-specified 'fixn' number of breaks.

Note: All `default` values of error assumptions (`robust`, `hetdat`, `hetvar`, `hetq`) are set to 1. The implications on the structure of model\'s errors related to individual settings are explained within the arguments section for each option.

See Also

`dotest`, `doseqtests`, `doorder`, `dosequa`, and `dofix` which are functions called by 'mdl()'.

Examples

```
US_rate = mdl('rate',data=real)
nkpc_lbs = mdl('inf',c('inflag','lbs','inffut'),data=nkpc,prewhit = 0)
```

nkpc *New Keynesian Phillips curve data*

Description

Data set from inflation and other macroeconomic variables

Usage

```
nkpc
```

Format

‘nkpc‘ A data frame with 151 rows and 12 columns:

year Current period year

quarter Quarter in current period year

inf Inflation rate

inflag Inflation rate in previous period

inffut Expected inflation rate, taken as value of inflation rate of next period

ygap Productivity output gap

lbs

lbslag

spreadlag

dwlag

dcplag

Source

Perron, P. and Yamamoto, Y., 2015. "Using ols to estimate and test for structural changes in models with endogenous regressors." Journal of Applied Econometrics 30, 119–144.

nldat *Computation of global minimizer for partial structural change model*

Description

‘nldat()‘ computes the break dates of a partial structural change model for a pre-specified number of breaks ‘m‘. The procedure iterates between estimating the invariant and changing coefficients of ‘x‘ and ‘z‘ regressors until convergence, by noting that the residuals from linear regression model between ‘y‘ and ‘x‘ regressors is a pure structural change model, while the residuals from pure structural change model between ‘y‘ and ‘z‘ regressors is a linear regression

Usage

```
nldat(y, z, x, h, m, p, q, bigT, fixb, eps, maxi, betaini, printd)
```

Arguments

y	dependent variable in matrix form
z	matrix of regressors which coefficients are allowed to change across regimes
x	matrix of regressors which coefficients are constant across regime
h	minimum segment length
m	number of breaks
p	number of 'z' regressors
q	number of 'x' regressors
bigT	the sample size T
fixb	option to use initial β If 1, procedure requires betaini. If 0, procedure will not use initial beta values
eps	Convergence criterion (For partial change model ONLY)
maxi	Maximum number of iterations (For partial change model ONLY)
betaini	initial beta values. Required when use with option fixb
printd	option to print output from iterated estimations. If 1, the results for each iteration will be printed in console log. If 0, no output will be printed

Value

A list containing the following components:

glb	minimum global SSR
datevec	Vector of dates (optimal minimizers)
bigvec	Associated SSRs

References

Bai J, Perron P (1998). *"Estimating and Testing Linear Models with Multiple Structural Changes"* Econometrica, 66, 47-78. Bai J, Perron P (2003). *"Computation and Analysis of Multiple Structural Change Models"* Journal of Applied Econometrics 18, 1-22

pftest *SupF test for 0 vs i breaks*

Description

Function compute the supF test statistics of testing procedure with null hypothesis: no break versus alternative hypothesis: `i` breaks.

Usage

```
pftest(y, z, i, q, bigT, datevec, prewhit, robust, x, p, hetdat, hetvar)
```

Arguments

`y`	dependent variables
`z`	independent variables with coefficients are allowed to change across regimes
`i`	number of breaks in the model
`q`	number of `z` regressors
`bigT`	sample period T
`datevec`	`i` estimated dates from the model
`prewhit`	Options for prewhitening process
`robust`, `hetdat`, `hetvar`	options for assumptions on error terms
`x`	independent variables with constant coefficients across regimes
`p`	number of `x` regressors

Value

ftest SupF test results

plambda *Construct diagonal matrix according to break date*

Description

Function constructs a diagonal matrix of dimension (m+1) by (m+1) matrix with i-th entry $\frac{T_i - T_{i-1}}{T}$

Usage

```
plambda(b, m, bigT)
```

Arguments

b	Estimated date of changes
m	Number of breaks
bigT	The sample size T

Value

lambda (m+1) by (m+1) diagonal matrix with i-th entry $\frac{T_i - T_{i-1}}{T}$

---

plot_model *Plot structural change model*

Description

'plot_model()' visualizes any object of class 'model' with comparison between real, fitted values between model of 'm' breaks and null model of '0' breaks with options for confidence interval of break date.

Usage

```
plot_model(model, CI = 0.95, title = NULL)
```

Arguments

model	object of class 'model' in 'mbreaks' package
CI	confidence intervals for break date and coefficient estimates visualize in terms of fitted values
title	title of the graph

Value

No return value, called for plotting class 'model' object. For more details on 'model' class, see [compile_model]

Examples

```
rate = dofix('rate',data=real,fixn=2)
plot_model(rate,title='Ex-post US exchange rate')
```

print.mdl *Print information of 'mbreaks' comprehensive procedure*

Description

'print' prints the class 'mdl' object with default showing only certain procedures called by 'mdl()' function including: 'seqtests' class object, 'sbtests' class object, and 'model' class object using KT information criterion

Usage

```
## S3 method for class 'mdl'
print(x, ...)
```

Arguments

x	class 'mdl' object
...	further arguments passed to or from other methods

Value

No return value, only for printing 'model', 'sbtests' and 'seqtests' class objects invoked during 'mdl()'.

Examples

```
rate = mdl('rate',data=real)
print(rate)
```

print.model *Summary output of a structural breaks 'model'*

Description

'print' the output of the S3 class 'model' with all relevant information:

- name of procedure used to obtain number of breaks in the model
- print a table summarizing the break date estimation (including confidence interval for the estimated date)
- print a table summarizing the estimated coefficients for 'z' regressors
- print a table summarizing the estimated coefficients for 'x' regressors (if any)

Usage

```
## S3 method for class 'model'
print(x, ...)
```

Arguments

x	object of S3 class 'model'
...	further arguments passed to or from other methods.

Value

No return value, called for printing to console the following information in 'x':

- Basic details of the model: name of prodecures invoked, number of estimated breaks, pure/partial structural change model,global min SSR
- 'date_tab' summarizes estimated break dates, see `compile_model`
- 'RS_tab' summarizes estimated coefficients allowed to change across regimes, see `compile_model`
- 'FS_tab' summarizes estimated coefficients constant across regimes, see `compile_model`

---

print.sbtests *Print Sup F and UDMax tests*

Description

'print' prints the following information from a 'sbtests' class object:

'supF1' A table reports sup F tests of 0 versus '1' upto 'm' breaks with critical values for *1%, 2.5%, 5%, and 10%* significance levels.

'UDmax' A table reporting Double Max tests with critical values for *1%, 2.5%, 5%, and 10%* significance levels.

Usage

```
## S3 method for class 'sbtests'
print(x, ...)
```

Arguments

x	class 'sbtests' object
...	further arguments passed to or from other methods

Value

No return value, only for printing formatted 'sbtests' class object to console

Examples

```
supF = dotest('inf','inflag',data=nkpc)
print(supF)
```

print.seqtests *Print sequential SupF tests*

Description

'print' prints the object of class 'seqtests' with the following information

- Maximum number of breaks 'm' in the tests
- 'sfl' table with sequential sup F tests statistics of l versus l+1 breaks up to 'm' breaks

Usage

```
## S3 method for class 'seqtests'
print(x, ...)
```

Arguments

x	'seqtests' class object.
...	further arguments passed to or from other methods.

Value

No return value, only for printing formatted 'seqtests' class object to console

Examples

```
seq_supF = doseqtests('inf','inflag',data=nkpc)
print(seq_supF)
```

psigmq *Construct diagonal matrix of estimated variance*

Description

Function computes a diagonal matrix of dimension m+1 by m+1 with i-th entry is the estimated variance of residuals of segment i

Usage

```
psigmq(res, b, q, m, nt)
```

Arguments

res	big residual vector of the model
b	Estimated date of changes
q	Number of ‘z‘ regressors
m	Number of breaks
nt	The size of ‘z‘ regressors

Value

sigmat (‘m‘+1)x(‘m‘+1) diagonal matrix with i-th entry equal to estimated variance of regime i

---

real *World Health Organization TB data*

Description

Data set from the Garcia and Perron study’s of ex-post exchange rate.

Usage

```
real
```

Format

‘real‘ A data frame with 103 rows and 1 column:

rate Real exchange rate

Source

Garcia, R. and Perron, P., 1996. "An analysis of the real interest rate under regime shifts." Review of Economics and Statistics 111–125.

spflp1 *SupF(l+1|l) test*

Description

'spflp1' computes the test statistics of supF(l+1|l) test with null hypothesis is l='nseg-1' and alternative hypothesis is l+1. The l breaks under the null hypothesis are taken from the global minimization.

Usage

```
spflp1(bigvec, dt, nseg, y, z, h, q, prewhit, robust, x, p, hetdat, hetvar)
```

Arguments

bigvec	associated SSR of estimated break date under H0
dt	vector of estimated date under H0
nseg	number of segment under H1
y	matrix of dependent variable
z	matrix of variables with coefficients are allowed to change across regimes
h	minimum segment length
q	number of 'z' regressors
prewhit, robust, hetdat, hetvar	options on residuals/errors. For more details, please refer to [mdl()]
x	matrix of variables with constant coefficients across regimes
p	number of 'x' regressors

Value

A list that contains the following:

maxf Maximum value of test

newd Additional date in alternative hypothesis

Index