

Package ‘pdR’

May 9, 2026

Type Package

Title Threshold Model and Unit Root Tests in Cross-Section and Time Series Data

Version 1.9.4

Date 2025-08-19

Author Ho Tsung-wu [aut, cre]

Maintainer Ho Tsung-wu <tsungwu@ntnu.edu.tw>

Description

Threshold model, panel version of Hylleberg et al. (1990) <[DOI:10.1016/0304-4076\(90\)90080-D](https://doi.org/10.1016/0304-4076(90)90080-D)> seasonal unit root tests, and panel unit root test of Chang (2002) <[DOI:10.1016/S0304-4076\(02\)00095-7](https://doi.org/10.1016/S0304-4076(02)00095-7)>.

License GPL (>= 2)

LazyData TRUE

LazyLoad yes

Depends R (>= 3.5.0)

Imports plm, pglm, lme4, glmmTMB, uroot

NeedsCompilation no

Repository CRAN

Date/Publication 2025-08-19 06:00:02 UTC

Contents

pdR-package	2
bank_income	3
cigaretts	4
crime	4
dur_john	6
htest_pglm	7
IGF	8
inf19	9
inf_M	10

inf_Q	11
invest	12
ipsHEGY	13
lagSelect	14
model	15
pIGF	16
productivity	17
ptm	18
r_est	19
ship	20
SMPLSplit_est	21
SMPLSplit_example	22
SMPLSplit_het	23
sse_calc	24
tbar	25
thr_sse	25
tr	26
Index	27

pdR-package

Panel Data Regression: Threshold Model and Unit Root Tests

Description

Functions for analysis of panel data, including the panel threshold model of Hansen (1999,JE), panel unit root test of Chang(2002,JE) based upon instruments generating functions (IGF), and panel seasonal unit root test based upon Hylleberg et al.(1990,JE).

Details

This version offers formatted output. This package designs a specification function `ptm()` to estimate the panel threshold model of Hansen(1999). The key feature of `ptm()` is to generalize Hansen's original code to allow multiple (more-than-one) regime-dependent right-hand-side independent variables; Dr. Hansen's original code admits only 1 regime-dependent right-hand-side independent variable. This version also includes panel unit root tests based on the instrument generating functions(IGF), proposed by Chang (2002, J. of Econometrics), and the panel version of Hylleberg et al.(1990) seasonal unit root test, proposed by Otero, et al. (2005, 2007).

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, College of Management, National Taiwan Normal University.

Maintainer: Ho Tsung-wu <tsungwu@ntnu.edu.tw>

References

- Chang, Y. (2002) Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency. *Journal of Econometrics*, 110, 261-292.
- Hansen B. E. (1999) Threshold effects in non-dynamic panels: Estimation, testing and inference. *Journal of Econometrics*, 93, 345-368.
- Hylleberg, S., Engle, R.F., Granger, C.W.J., and Yoo, B.S.(1990) Seasonal integration and cointegration. *Journal of Econometrics*, 44, 215-238.
- Otero, J., Smith, J., and Giuliatti, M. (2005) Testing for seasonal unit roots in heterogeneous panels. *Economics Letters*, 86, 229-235.
- Otero, J., Smith, J., and Giuliatti, M. (2007) Testing for seasonal unit roots in heterogeneous panels in the presence of cross section dependence. *Economics Letters*, 86, 179-184.
- Pesaran M. Hashem (2007) A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, 22, 265-312.

bank_income

Panel data of bank, 2001Q1~2010Q1

Description

A quarterly panel data frame with 1000 observations on the following 7 variables, unbalanced panel data ranges from 2001Q1~2010Q1.

Usage

```
data("bank_income")
```

Format

ID a numeric vector
Qtr a numeric vector
preTax_Income a numeric vector
shortRatio a numeric vector
longRatio a numeric vector
Current_ratio a numeric vector
LoanDeposit_ratio a numeric vector

Examples

```
data(bank_income)
```

cigaretts

Cigaretts consumption of US states

Description

Cigaretts consumption of US states

Usage

```
data(cigaretts)
```

Format

A data frame of 48 US states' cigarette consumption

State State abbreviation, N

Year Year, t

Y_SALES Cigarette sales in packs per capita, deflated by population

X1_PRICE P=Real price per pack of cigarettes, deflated by 1983 CPI.

X2_PIMIN Real minimum price in adjoining states per pack of cigarettes, deflated by CPI

X3_NDI Per capita disposable income

References

Baltagi Badi H. (2005) *Econometric Analysis of Panel Data*. John Wiley.

Examples

```
data(cigaretts)
head(cigaretts)
```

crime

Annual crime dataset of US counties

Description

Annual crime dataset of US counties

Usage

```
data(crime)
```

Format

A data frame of US counties

county counties index, N

year Year, t

crmrte crime rate(crime/population)

prbarr probability of arrest (arrests/offenses)

prbconv probability of conviction, given arrest

prbpris probability of a prison, given conviction

avgsen sanction severity(average prison sentence in days)

polpc ability of police force to detect crime(# of police per capita)

density population density(POP/area)

taxpc Taxpayment per capita

region region index of county

smsa =1 if SAMA, POP>50000; =0 else

pctmin See Baltagi(2006) for details

wcon See Baltagi(2006) for details

wtuc See Baltagi(2006) for details

wtrd See Baltagi(2006) for details

wfir See Baltagi(2006) for details

wser See Baltagi(2006) for details

wmfg See Baltagi(2006) for details

wfed See Baltagi(2006) for details

wsta See Baltagi(2006) for details

wloc See Baltagi(2006) for details

mix See Baltagi(2006) for details

pctymle See Baltagi(2006) for details

References

Baltagi Badi H. (2005) *Econometric Analysis of Panel Data*. John Wiley. Baltagi Badi H. (2006) Estimating an Economic Model of Crime Using Panel Data from North Carolina. *J.of Applied Econometrics* 21: 543;V547.

dur_john

The cross-country growth data in Durlauf and Johnson(1995)

Description

The Durlauf-Johnson data manipulated by Hansen(2000),excluding missing variables and oil states

Usage

```
data(dur_john)
```

Format

A data frame with 19 countries

gdpGrowth Economic growth measured by GDP of 1960 and 1985

logGDP60 log Per capita GDP in 1960

Inv_GDP Average ratio of investment (including Government Investment) to GDP from 1960 to 1985

popGrowth Average growth rate of working-age population 1960 to 1985

School Average fraction of working-age population enrolled in secondary school from 1960 to 1985

GDP60 Per capita GDP in 1960

Literacy fraction of the population over 15 years old that is able to read and write in 1960

Details

Steven N. Durlauf and Paul A. Johnson, "Multiple Regimes and Cross-Country Growth Behavior," Journal of Applied Econometrics, Vol. 10, No. 4, 1995, 365-384.

Examples

```
data(dur_john)
head(dur_john)
```

htest_pglm	<i>Specification test for panel glm models</i>
------------	--

Description

This function performs Hausman specification test for panel glm.

Usage

```
htest_pglm(RE, FE, re.method)
```

Arguments

RE	Random effect objects. Support pglm, glmer, glmmTMB
FE	Fixed effect objects.
re.method	Method that used to estimate the random effect estimation, in addition to "pglm", it also supports "glmmTMB" of package glmmTMB, and "glmer" of package lme4.

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, modified from phtest() of plm

References

Hausman J.A. (1978). Specification Tests in Econometrics. *Econometrica*, 46, 1251-1271.

Examples

```
data(ship)
library(pglm)
Eq1="accident ~ op_75_79+co_65_69+co_70_74+co_75_79"
FE.pois <- pglm(Eq1,data=ship,family = "poisson",model = "within",index = 'ship',R=10)

RE.pois <- pglm::pglm(Eq1,data=ship,family = "poisson", model = "random", index = 'ship')

## Hausman test
htest_pglm(RE=RE.pois, FE=FE.pois, re.method="pglm")

Eq2=accident ~ op_75_79+co_65_69+co_70_74+co_75_79 + (1 | ship)
re.glmmTMB=glmmTMB::glmmTMB(Eq2,data=ship, family="poisson")

## Hausman test
htest_pglm(RE=re.glmmTMB, FE=FE.pois, re.method="glmmTMB")
```

IGF

*Unit root test based on Change(2002)***Description**

This function estimates the unit root regression based on instrument generating function of Change(2002) and returns useful outputs.

Usage

```
IGF(y, maxp, ic, spec)
```

Arguments

y	A univariate time series data
maxp	the max number of lags
ic	Information criteria, either "AIC" or "BIC"
spec	regression model specification. =0, no intercept and trend. =1, intercept only. =2, intercept and trend.

Details

Estimate univariate unit root test of Chang(2002).

Value

tstat.IGF	IGF unit root test
beta	regression coefficients. The first one is the AR(1) coefficient of unit root, and the last one is the intercept or trend
sdev	The IGF standard error for unit root coefficient
cV	The scalar C in IGF equation
p	The optimal number of lag

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, College of Management, National Taiwan Normal University.

References

Chang, Y. (2002) Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency. *Journal of Econometrics*, 110, 261-292.

Examples

```
data(inf19)
y <- inf19[,1]
IGF(y,maxp=35,ic="BIC",spec=2)$tstat.IGF
```

inf19

Monthly inflation time series of 19 countries

Description

Monthly inflation time series of 19 countries, 1984.1~2011.3

Usage

```
data(inf19)
```

Format

A data frame with 19 countries

AUSTRIA inflation of Austria

BELGIUM inflation of Belgium

CANADA inflation of Canada

DENMARK inflation of Denmark

FINLAND inflation of Finland

FRANCE inflation of France

GREECE inflation of Greece

ICELAND inflation of Iceland

ITALY inflation of Italy

JAPAN inflation of Japan

LUXEMBOURG inflation of Luxembourg

NETHERLANDS inflation of Netherlands

NORWAY inflation of Norway

PORTUGAL inflation of Portugal

SPAIN inflation of Spain

SWEDEN inflation of Sweden

SWITZERLAND inflation of Switzerland

UK inflation of UK

USA inflation of USA

Details

Monthly CIP, seasonally differenced of log CPI of 19 countries

Examples

```
data(inf19)
head(inf19)
```

inf_M	<i>Monthly inflation time series of 20 countries</i>
-------	--

Description

Monthly inflation time series of 19 countries, 1971.1~2011.12

Usage

```
data(inf_M)
```

Format

A data frame with 20 countries

- AUSTRALIA inflation of Austrlia
- AUSTRIA inflation of Austria
- BELGIUM inflation of Belgium
- CANADA inflation of Canada
- DENMARK inflation of Denmark
- FINLAND inflation of Finland
- FRANCE inflation of France
- GREECE inflation of Greece
- ICELAND inflation of Iceland
- ITALY inflation of Italy
- JAPAN inflation of Japan
- LUXEMBOURG inflation of Luxembourg
- NETHERLANDS inflation of Netherlands
- NORWAY inflation of Norway
- PORTUGAL inflation of Portugal
- SPAIN inflation of Spain
- SWEDEN inflation of Sweden
- SWITZERLAND inflation of Switzerland
- UK inflation of UK
- USA inflation of USA

Details

Monthly CIP, seasonally differenced of log CPI of 20 countries

Examples

```
data(inf_M)
head(inf_M)
```

inf_Q	<i>Quarterly inflation time series of 20 countries</i>
-------	--

Description

Quarterly inflation time series of 19 countries, 1971Q1~2014Q4

Usage

```
data(inf_Q)
```

Format

A data frame with 19 countries

- AUSTRALIA inflation of Australia
- AUSTRIA inflation of Austria
- BELGIUM inflation of Belgium
- CANADA inflation of Canada
- DENMARK inflation of Denmark
- FINLAND inflation of Finland
- FRANCE inflation of France
- GREECE inflation of Greece
- ICELAND inflation of Iceland
- ITALY inflation of Italy
- JAPAN inflation of Japan
- LUXEMBOURG inflation of Luxembourg
- NETHERLANDS inflation of Netherlands
- NORWAY inflation of Norway
- PORTUGAL inflation of Portugal
- SPAIN inflation of Spain
- SWEDEN inflation of Sweden
- SWITZERLAND inflation of Switzerland
- UK inflation of UK
- USA inflation of USA

Details

Quarterly CIP, seasonally differenced of log CPI of 20 countries

Examples

```
data(inf_Q)
head(inf_Q)
```

invest	<i>investment data of 565 listed companies, 1973-1987</i>
--------	---

Description

investment data of 565 listed companies, 1973-1987, from Hansen's example

Usage

```
data(invest)
```

Format

A pooled data frame

invest[,1] investment/assets

invest[,2] Tobin's Q

invest[,3] cash-flow/assets

invest[,4] debt/assets

Details

This is a pooled data frame, without date (T) and cross-section(N) ID columns

Examples

```
#data(invest)
#head(invest)
```

ipsHEGY *IPS-HEGY seasonal unit root test in panel data, Otero et al.(2007).*

Description

This function performs panel data-based HEGY seasonal unit root test, the asymptotics is based upon Otero et al.(2007).

Usage

```
ipsHEGY(data,deterministic =c(1,0,0),lag.method="AIC",maxlag=12,CIPS = TRUE)
```

Arguments

data	Panel data, T by N
deterministic	Options for deterministic components. c(1,0,0), intercept only. c(0,1,0), trend only. c(0,0,1),, season dummy only.
lag.method	Selection of optimal lags, having for options: "fixed","AIC", "BIC", "AICc".
maxlag	Maximum number of lags for searching optimal criteria.
CIPS	Logical. If TRUE, using Pesaran(2007) to account for cross-section correlation. The default is TRUE.

Details

Mode for selectlags has four options, AIC,AICc and BIC use R built-in functions for linear model and their meanings are popular and straightforward. "fixed" fixes maxlag as the number of lags.

The critical values for panel HEGY are standard normal for individual t-ratios, however,you need to perform simulation for the critical values of F joint test, at pdR 1.3. To this end, you are encouraged to work this out for yourself: using arima.sim() to sample seasonal time series with unit root (1-order difference) and obtain their statistics under the null using ipsHEGY(), then it is straightforward to obtain critical values.

Otero et al. (2007) provide critical values for quarterly frequency.

The univariate HEGY is based on R package uroot, simply modified to include cross-sectionaal average.

Value

P_HEGY	Panel HEGY statistics.
U_HEGY	Individual HEGY statistics of N units.

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, College of Management, National Taiwan Normal University.

References

Otero, J., Smith, J., and Giuliatti, M. (2005) Testing for seasonal unit roots in heterogeneous panels. *Economics Letters*, 86, 229-235.

Otero, J., Smith, J., and Giuliatti, M. (2007) Testing for seasonal unit roots in heterogeneous panels in the presence of cross section dependence. *Economics Letters*, 86, 179-184.

Pesaran M. Hashem (2007) A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, 22, 265-312.

Examples

```
data(inf_M)

#Seasonal dummy only takes quarters 1:3,
#becasue of the presence of common intercept.

OUT<-ipsHEGY(inf_M, deterministic = c(1,1,1), lag.method = "AIC", maxlag = 12,CIPS=TRUE)
OUT$P_HEGY
OUT$U_HEGY

# Simulation of critical values
```

lagSelect

Select the optimal number of lags, given criteria

Description

Determine the optimal number of lags for dynamic regression

Usage

```
lagSelect(y, maxp, ic)
```

Arguments

y	A univariate time series data
maxp	the max number of lags
ic	Information criteria, either "AIC" or "BIC"

Details

Information criteria "AIC" and "BIC" use the R built-in functions.

Value

It returns an integer, indicating the optimal lags

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, College of Management, National Taiwan Normal University.

Examples

```
#library(pdR)
#data(inf19)
#y<-inf19[,1]
#lagSelect(y,maxp=25,ic="BIC")
```

 model

Estimate specified panel threshold model

Description

This function is the main function estimating threshold regression for function ptm()

Usage

```
model(r, trim, rep, it, qq1, cf, xt, ct, thresh, tt, qn1, n, qn, cc, yt, ty, k)
```

Arguments

r	vector of threshold estimate(s).
trim	value of trimmed percentage.
rep	number bootstrap repetition.
it	number of regime during computation, used in a for loop.
qq1	defined parameter.
cf	special declaration, e.g. lag().
xt	regime independent variables.
ct	trace of regime dependent variables.
thresh	threshold variable.
tt	length of time period.
qn1	as defined by nrow(qq1).
n	number of cross-section units.

qn	number of quantiles to examine.
cc	as defined by $2*\log(1-\sqrt{\text{conf_lev}})$.
yt	vectorized dependent variable.
ty	trace of yt.
k	number of regime-independent independent variables.

Note

Original code offered by Dr. B. E.Hansen (<http://www.ssc.wisc.edu/~bhansen/>).

References

Hanse B. E. (1999) Threshold effects in non-dynamic panels: Estimation, testing and inference. *Journal of Econometrics*,93, 345-368.

pIGF *Panel unit root test of Chang(2002)*

Description

Compute the panel unit root test statistic of Chang(2002).

Usage

pIGF(datamat, maxp, ic, spec)

Arguments

datamat	T by N panel data.T is the time length,N is the number of cross-section units.
maxp	the max number of lags
ic	Information criteria, either "AIC" or "BIC".
spec	model specification. =0, no intercept and trend. =1, intercept only. =2, intercept and trend.

Details

This function estimates the panel unit root test based on univariate instrument generating function of (Chang,2002).

Value

panel.tstat	panel IGF test statistics
pvalue	P-value of the panel.tstat

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, College of Management, National Taiwan Normal University.

References

Chang, Y. (2002) Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency. *Journal of Econometrics*, 110, 261-292.

Examples

```
data(inf19)
datam <- inf19
pIGF(datam, maxp=25, ic="BIC", spec=2)
```

productivity

Productivity data of 48 US state, 1970-1986

Description

Gross state production data

Usage

```
data(productivity)
```

Format

A data frame with US production

state US state index, 1-48

year Year index

y_gsp Gross state product

x1_hwy Expenditure of public utility- highway construction

x2_water Expenditure of public utility- water

x3_other Expenditure of others

x4_private Private consumption of each state

x5_emp Employment rate of each state

x6_unemp Unemployment rate of each state

Examples

```
data(productivity)
head(productivity)
```

ptm *Threshold specification of panel data*

Description

A generalized specification for estimating panel threshold model.

Usage

```
ptm(dep, ind1, ind2, d, bootn, trimn, qn, conf_lev, t, n)
```

Arguments

dep	Dependent variable
ind1	Independent variables: regime dependent
ind2	Independent variables: regime independent
d	Threshold variable
bootn	Vector of bootstrap repetition
trimn	Vector of trimmed percentage
qn	Number of quantiles to examine
conf_lev	Confidence level
t	Length of time period
n	Number of cross-section units

Details

This code fits only balanced panel data. It generalizes the simple code of Dr. Hansen (<http://www.ssc.wisc.edu/~bhansen/>), allowing multiple (more-than-one) regime-dependent (ind1) variables. We generalize the original code to better fit general need of threshold modeling in panel data.

bootn and trimn are vector of 3 by 1, indicating numbers of three corresponding regimes.

This version corrects a slight error incurred by argument max_lag, which is used by Hansen to arrange investment data via lags. In this package, users manipulate data to fit personal research to ptm(), hence this argument is omitted lest should degree of freedom will suffer a loss of N.

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, College of Management, National Taiwan Normal University.

References

Hansen B. E. (1999) Threshold effects in non-dynamic panels: Estimation, testing and inference. *Journal of Econometrics*, 93, 345-368.

Examples

```

# library(pdR)
#data(invest)
#dat<-invest[1:1500,] # subsetting the first 1500 obs., #for simplicity
#t <- 15 #Length of time period
#nt <- nrow(dat)
#n <- nt/t # number of cross-section units

#dep<- as.matrix(dat[,1]) # investment/assets
#th1<- as.matrix(dat[,2]) #Tobin's Q
#th2<- as.matrix(dat[,3]) # cash-flow/assets
#ind1<- cbind(th1,th2) #regime-dep covariates
#d <- as.matrix(dat[,4]) # Threshold variable
#ind2 <- cbind((th1^2),(th1^3),(th1*d)) # regime-indep covariates:
#bootn<-c(100,200,300) # bootstrapping replications for each threshold esitmaton
#trimn<-c(0.05,0.05,0.05) #trimmed percentage for each threshold esitmaton

#qn<-400
#conf_lev<-0.95

#Output=ptm(dep,ind1,ind2,d,bootn,trimn,qn,conf_lev,t,n)
#Output[[1]] #Formatted output of 1st threshold, 2 regimes
#Output[[2]] #Formatted output of 2nd threshold, 3 regimes
#Output[[3]] #Formatted output of 3rd threshold, 4 regimes

# In the output, the Regime-dependent Coefficients matrix
# is, from top to bottom, regime-wise.

```

r_est

A subroutine for model()

Description

This function is a subroutine for model(), estimation procedure.

Usage

```
r_est(y, r, trim, tt, qq1, qn1, qn, n, cf, xt, ct, thresh)
```

Arguments

y	vector of dependent variable.
r	numer of regime.
trim	value of trimmed percentage.
tt	length of time period.
qq1	parameter defined by <code>as.matrix(unique(thresh)[floor(sq*nrow(as.matrix(sort(unique(thresh))))))]</code> .
qn1	as defined by <code>nrow(qq1)</code> .

qn	number of quantiles to examine.
n	parameter of cross-section units.
cf	special declaration, e.g. lag().
xt	regime independent variables.
ct	trace of regime dependent variables.
thresh	threshold variable.

References

Hanse B. E. (1999) Threshold effects in non-dynamic panels: Estimation, testing and inference. *Journal of Econometrics*, 93, 345-368.

Original code from Dr. Hansen (<http://www.ssc.wisc.edu/~bhansen/>).

ship *Panel data on the number of ship accidents*

Description

Panel data on the number of ship accidents, McCullagh and Nelder(1989)

Usage

```
data("ship")
```

Format

```
accident the number of ship accidents
ship Ship iD
service the number of months in service
op_75_79 the operating period between 1975 and 1979
co_65_69 consecutive construction periods of 5 years
co_70_74 consecutive construction periods of 5 years
co_75_79 consecutive construction periods of 5 years
yr_con years of construction
yr_op years of operation
```

References

McCullagh, P., and J. A. Nelder (1989) *Generalized Linear Models*. 2nd ed. London: Chapman and Hall/CRC.

Examples

```
data(ship)
```

SMPLSplit_est	<i>Estimation of sub-sampled data</i>
---------------	---------------------------------------

Description

A function for estimating the subsampled data.

Usage

```
SMPLSplit_est(data,dep,indep,th,plot,h=1,nonpar=2)
```

Arguments

data	the data in either data.frame or matrix.
dep	the name of dependent variable.
indep	the name(s) of independent variable(s).
th	the name of threshold variable.
plot	=1, plot; =0, do not plot.
h	h=1, heteroskedasticity-consistent covariance; h=0, homoskedastic case.
nonpar	Indicator for non-parametric method used to estimate nuisance scale in the presence of heteroskedasticity (only relevant if h=1).Set nonpar=1 to estimate regressions using a quadratic.Set nonpar=2 to estimate regressions using an Epanechnikov kernel with automatic bandwidth.

Details

This code estimates the parameters of sub-sampled data. It generalizes the simple code of Dr. Hansen, allowing White Corrected Heteroskedastic Errors.

Value

threshold	values of threshold estimates.
est0	coefficient estimates of global data.
est.low	coefficient estimates of low regime.
est.high	coefficient estimates of high regime.
est0.info	additional information of global data.
est.joint.info	additional information of joint threshods.
est.low.info	additional information of est.low.
est.high.info	additional information of est.high.

Note

Original code offered by Dr. B. E.Hansen (<http://www.ssc.wisc.edu/~bhansen/>).

References

Hansen B. E. (2000) Sample Splitting and Threshold Estimation. *Econometrica*, 68, 575-603.

Examples

```
## Not run, because of bootstrap replication takes time. Users may unmark # and run.
data("dur_john")
rep <- 500
trim_per <- 0.15
dep <- "gdpGrowth"
indep <- colnames(dur_john)[c(2,3,4,5)]

SMPLSplit_est(data=dur_john,dep,indep,th="GDP60",plot=0,h=1,nonpar=2)
```

SMPLSplit_example *Example code for sample splitting*

Description

A sample code for learning sample splitting.

Usage

```
SMPLSplit_example(data,dep,indep,th1,th2,trim_per,rep,plot)
```

Arguments

data	the data in either data.frame or matrix.
dep	the name of dependent variable.
indep	the name(s) of independent variable(s)
th1	the first threshold variable.
th2	the second threshold variable.
trim_per	trimmed percentage.
rep	nNumber of bootstrap repetitions.
plot	=1, plot; =0, do not plot.

Details

This code is the learning example for learning Hansen's econometric sample splitting. I detailed the description of each threshold stage.

Note

Original code offered by Dr. B. E.Hansen (<http://www.ssc.wisc.edu/~bhansen/>).

References

Hansen B. E. (2000) Sample Splitting and Threshold Estimation. *Econometrica*, 68, 575-603.

Examples

```
## Not run, because of bootstrap replication takes time. Users may unmark # and run.
data("dur_john")
#rep <- 500
#trim_per <- 0.15
#dep <- "gdpGrowth"
#indep <- colnames(dur_john)[c(2,3,4,5)]
#th1 <- "GDP60"
#th2 <- "Literacy"
#OUT=SMPLSplit_est(data=dur_john,dep,indep,th=th1,plot=0,h=1,nonpar=2)
#OUT$TEST
#OUT$Hypothesis
#OUT$Threshold
#stat=matrix(as.numeric(OUT$TEST),byrow = TRUE,8,2)
#colnames(stat)=c("F-Stat","P-value")
#rownames(stat)=OUT$Hypothesis
#stat
```

SMPLSplit_het

Testing for sample splitting

Description

A function for testing sample split given subsampled data.

Usage

```
SMPLSplit_het(data,dep,indep,th,trim_per,rep,plot)
```

Arguments

data	the data in either data.frame or matrix
dep	the name of dependent variable.
indep	the name(s) of independent variable(s).
th	the name of threshold variable.
trim_per	trimmed percentage.
rep	number of bootstrap repetition.
plot	=1, plot; =0, do not plot.

Details

This code tests for the presence of threshold. It generalizes the simple code of Dr. Hansen, allowing Heteroskedastic Errors (White Corrected).

Value

fstat LM-test for no threshold.
pvalue bootstrap P-Value.

Note

Original code offered by Dr. B. E.Hansen (<http://www.ssc.wisc.edu/~bhansen/>).

References

Hanse B. E. (2000) Sample Splitting and Threshold Estimation. *Econometrica*, 68, 575-603.

sse_calc *a subroutine of model()*

Description

SSE calculation

Usage

sse_calc(y, x)

Arguments

This function is a sub-routine for model(), calculating SSE of each regression

y vector of dependent variable.
x matrix of independent variables.

References

Hanse B. E. (1999) Threshold effects in non-dynamic panels: Estimation, testing and inference. *Journal of Econometrics*, 93, 345-368.

Original code from Dr. Hansen (<http://www.ssc.wisc.edu/~bhansen/>).

tbar	<i>Compute the resursive mean</i>
------	-----------------------------------

Description

Compute the resursive mean of each series

Usage

```
tbar(x)
```

Arguments

x A univariate time series data

Details

This function computes the resursive mean

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>

Examples

```
data(inf19)
y <- inf19[,1]
tbar(y)
```

thr_sse	<i>a subroutine calculating SSE</i>
---------	-------------------------------------

Description

This function is a sub-routine for model(), calculating SSE of each threshold regression.

Usage

```
thr_sse(y, q, r, cf, xt, ct, thresh, tt, n)
```

Arguments

y	parameter.
q	qq1 in model().
r	parameter.
cf	as defined in model().
xt	as defined in model().
ct	as defined in model().
thresh	as defined in model().
tt	as defined in model().
n	as defined in model().

References

Hanse B. E. (1999) Threshold effects in non-dynamic panels: Estimation, testing and inference. *Journal of Econometrics*,93, 345-368.

Original code from Dr. Hansen (<http://www.ssc.wisc.edu/~bhansen/>).

tr

A sub-routine calculating trace

Description

Estimation of trace.

Usage

```
tr(y, tt, n)
```

Arguments

This function is a sub-routine for model(), calculating trace of matrix

y	data vector.
tt	time period length.
n	number of cross-section units.

References

Hanse B. E. (1999) Threshold effects in non-dynamic panels: Estimation, testing and inference. *Journal of Econometrics*,93, 345-368.

Original code from Dr. Hansen (<http://www.ssc.wisc.edu/~bhansen/>).

Index

*** instrument generating functions**

IGF, 8

*** unit root**

IGF, 8

bank_income, 3

cigaretts, 4

crime, 4

dur_john, 6

htest_pglm, 7

IGF, 8

inf19, 9

inf_M, 10

inf_Q, 11

invest, 12

ipsHEGY, 13

lagSelect, 14

model, 15

pdR (pdR-package), 2

pdR-package, 2

pIGF, 16

productivity, 17

ptm, 18

r_est, 19

ship, 20

SMPLSplit_est, 21

SMPLSplit_example, 22

SMPLSplit_het, 23

sse_calc, 24

tbar, 25

thr_sse, 25

tr, 26