Package 'pder'

January 26, 2022

Version 1.0-2

Date 2022-01-25

Title Panel Data Econometrics with R

Depends R (>= 3.5.0), plm

Suggests car, dplyr, ggplot2, lmtest, msm, pglm, splm, survival, texreg

Description Data sets for the Panel Data Econometrics with R <doi:10.1002/9781119504641> book.

License GPL (>= 2)

URL https://cran.r-project.org/package=pder

Encoding UTF-8

NeedsCompilation no

Author Yves Croissant [aut, cre] (<https://orcid.org/0000-0002-4857-7736>), Giovanni Millo [aut]

Maintainer Yves Croissant <yves.croissant@univ-reunion.fr>

Repository CRAN

Date/Publication 2022-01-26 20:02:25 UTC

R topics documented:

`CallBacks` *Callbacks to Job Applications*

Description

a pseudo-panel of 1518 resumes from 2014

number of observations : 6072

number of individual observations : 4

country : United States

package : binomial

JEL codes: E24, E32, J14, J22, J23, J64

Chapter : 08

Usage

```
data(CallBacks)
```

Format

A dataframe containing:

jobid the job index

unempdur unemployment duration in month

interim a dummy for interim experience

callback a dummy for call backs

old a dummy for age 57-58

Source

American Economic Association Data Archive : https://www.aeaweb.org/aer/

References

Farber, Henry S.; Silverman, Dan and Till von Wachter (2016) "Determinants of Callbacks to Job Applications: An Audit Study", *American Economic Review*, **106(5)**, 314-318, doi: 10.1257/aer.p20161010 .

CoordFailure *How to Overcome Organization Failure in Organization*

Description

a pseudo-panel of 240 individuals

number of observations : 7168

number of individual observations : 30

country : United States and Spain

package : ordinalpanelexpe

JEL codes: C92, D23

Chapter : 08

Usage

```
data(CoordFailure)
```

Format

A dataframe containing:

firm the firm index

id the individual index

period the period

place either Cleveland or Barcelona

bonus1 the bonus for the first block of 10 rounds

bonus2 the bonus for the second block of 10 rounds

bonus3 the bonus for the third block of 10 rounds

effort the level of effort of the employee

Source

American Economic Association Data Archive : https://www.aeaweb.org/aer/

References

Brandts, Jordi and David J. Cooper (2006) "A Change Would Do You Good... An Experimental Study on How to Overcome Coordination Failure in Organizations", *American Economic Review*, **96(3)**, 669-693, doi: 10.1257/aer.96.3.669 .

`DemocracyIncome` *The Relation Between Democraty and Income*

Description

5-yearly observations of 211 countries from 1950 to 2000

number of observations : 2321

number of time-series : 11

country : world

package : panel

JEL codes: D72, O47

Chapter : 02, 07

Usage

```
data(DemocracyIncome)
```

Format

A dataframe containing:

country country

year the starting year of the 5-years period

democracy democracy index

income the log of the gdp per capita

sample a dummy variable to select the subset used in the original article

Source

American Economic Association Data Archive : https://www.aeaweb.org/aer/

References

Daron Acemoglu, Simon Johnson, James A. Robinson and Pierre Yared (2008) "Income and Democracy", *American Economic Review*, **98(3)**, 808-842, doi: 10.1257/aer.98.3.808 .

Examples

```
#### Example 7-1

## ------------------------------------------------------------------------

## Not run:
data("DemocracyIncome", package = "pder")

## ------------------------------------------------------------------------
data("DemocracyIncome", package="pder")
set.seed(1)
di2000 <- subset(DemocracyIncome, year == 2000,
                 select = c("democracy", "income", "country"))
di2000 <- na.omit(di2000)
di2000$country <- as.character(di2000$country)
di2000$country[- c(2,5, 23, 16, 17, 22, 71,  125, 37, 43, 44,
                   79, 98, 105, 50, 120,  81, 129, 57, 58,99)] <- NA

if(requireNamespace("ggplot2")){
    library("ggplot2")
    ggplot(di2000, aes(income, democracy, label = country)) +
        geom_point(size = 0.4) +
        geom_text(aes(y= democracy + sample(0.03 * c(-1, 1),
                                            nrow(di2000), replace = TRUE)),
                  size = 2) +
        theme(legend.text = element_text(size = 6),
              legend.title= element_text(size = 8),
              axis.title = element_text(size = 8),
              axis.text = element_text(size = 6))
}


## ------------------------------------------------------------------------
library("plm")
pdim(DemocracyIncome)
head(DemocracyIncome, 4)


#### Example 7-2

## ------------------------------------------------------------------------
mco <- plm(democracy ~ lag(democracy) + lag(income) + year - 1,
           DemocracyIncome, index = c("country", "year"),
           model = "pooling", subset = sample == 1)

## ------------------------------------------------------------------------
mco <- plm(democracy ~ lag(democracy) + lag(income),
```

```
           DemocracyIncome, index = c("country", "year"),
           model = "within", effect = "time",
           subset = sample == 1)
coef(summary(mco))


#### Example 7-3

## ------------------------------------------------------------------------
within <- update(mco, effect = "twoways")
coef(summary(within))


#### Example 7-4

## ------------------------------------------------------------------------
ahsiao <- plm(diff(democracy) ~ lag(diff(democracy)) +
              lag(diff(income)) + year - 1  |
              lag(democracy, 2) + lag(income, 2) + year - 1,
              DemocracyIncome, index = c("country", "year"),
              model = "pooling", subset = sample == 1)
coef(summary(ahsiao))[1:2, ]


#### Example 7-5

## ------------------------------------------------------------------------
diff1 <- pgmm(democracy ~ lag(democracy) + lag(income) |
              lag(democracy, 2:99)| lag(income, 2),
              DemocracyIncome, index=c("country", "year"),
              model="onestep", effect="twoways", subset = sample == 1)
coef(summary(diff1))

## ------------------------------------------------------------------------
diff2 <- update(diff1, model = "twosteps")
coef(summary(diff2))


#### Example 7-7

## ------------------------------------------------------------------------
sys2 <- pgmm(democracy ~ lag(democracy) + lag(income) |
             lag(democracy, 2:99)| lag(income, 2),
             DemocracyIncome, index = c("country", "year"),
             model = "twosteps", effect = "twoways",
             transformation = "ld")
coef(summary(sys2))


#### Example 7-8

## ------------------------------------------------------------------------
sqrt(diag(vcov(diff2)))[1:2]
```

```
sqrt(diag(vcovHC(diff2)))[1:2]


#### Example 7-10

## ------------------------------------------------------------------------
mtest(diff2, order = 2)


#### Example 7-9

## ------------------------------------------------------------------------
sargan(diff2)
sargan(sys2)

## End(Not run)
```

`DemocracyIncome25` *The Relation Between Democraty and Income*

Description

25-yearly observations of 25 countries from 1850 to 2000

number of observations : 175

number of time-series : 7

country : world

package : panel

JEL codes: D72, O47

Chapter : 02, 07

Usage

```
data(DemocracyIncome25)
```

Format

A dataframe containing:

country country

year the starting year of the 5-years period

democracy democracy index

income the log of the gdp per capita

Source

American Economic Association Data Archive : https://www.aeaweb.org/aer/

References

Daron Acemoglu, Simon Johnson, James A. Robinson and Pierre Yared (2008) "Income and Democracy", *American Economic Review*, **98(3)**, 808-842, doi: 10.1257/aer.98.3.808 .

Examples

```
#### Example 2-7

## ------------------------------------------------------------------------
library("plm")
data("DemocracyIncome25", package = "pder")
DI <- pdata.frame(DemocracyIncome25)
summary(lag(DI$income))
ercomp(democracy ~ lag(income), DI)
models <- c("within", "random", "pooling", "between")
sapply(models, function(x)
       coef(plm(democracy ~ lag(income), DI, model = x))["lag(income)"])


#### Example 7-6

## ------------------------------------------------------------------------
data("DemocracyIncome25", package = "pder")
pdim(DemocracyIncome25)

## ------------------------------------------------------------------------
diff25 <- pgmm(democracy ~ lag(democracy) + lag(income) |
               lag(democracy, 2:99) + lag(income, 2:99),
               DemocracyIncome25, model = "twosteps")

## ------------------------------------------------------------------------
diff25lim <- pgmm(democracy ~ lag(democracy) + lag(income) |
                  lag(democracy, 2:4)+ lag(income, 2:4),
                  DemocracyIncome25, index=c("country", "year"),
                  model="twosteps", effect="twoways", subset = sample == 1)
diff25coll <- pgmm(democracy ~ lag(democracy) + lag(income) |
                   lag(democracy, 2:99)+ lag(income, 2:99),
                   DemocracyIncome25, index=c("country", "year"),
                   model="twosteps", effect="twoways", subset = sample == 1,
                   collapse = TRUE)
sapply(list(diff25, diff25lim, diff25coll), function(x) coef(x)[1:2])

#### Example 7-9

## ------------------------------------------------------------------------
sapply(list(diff25, diff25lim, diff25coll),
       function(x) sargan(x)[["p.value"]])
```

`Dialysis` *Diffusion of Haemodialysis Technology*

Description

yearly observations of 50 states from 1977 to 1990

number of observations : 700

number of time-series : 14

country : United States

package : panel

JEL codes: I18, O31

Chapter : 09

Usage

```
data(Dialysis)
```

Format

A dataframe containing:

state the state id

time the year of observation

diffusion the number of equipment divided by the number of the equipment in the given state for the most recent period

trend a linear trend

regulation a dummy variable for the presence of a certificate of need regulation for the given state and the given period

Source

Journal of Applied Econometrics Data Archive : http://qed.econ.queensu.ca/jae/

References

Steven B. Caudill, Jon M. Ford and David L. Kaserman (1995) "Certificate of Need Regulation and the Diffusion of Innovations : a Random Coefficient Model", *Journal of Applied Econometrics*, **10**, 73–78., doi: 10.1002/jae.3950100107 .

Examples

```
#### Example 9-1

## ------------------------------------------------------------------------
library("plm")

## ------------------------------------------------------------------------
data("Dialysis", package = "pder")
rndcoef <- pvcm(log(diffusion / (1 - diffusion)) ~ trend + trend:regulation,
                Dialysis, model="random")
summary(rndcoef)

## ------------------------------------------------------------------------
cbind(coef(rndcoef), stdev = sqrt(diag(rndcoef$Delta)))
```

Donors *Dynamics of Charitable Giving*

Description

a pseudo-panel of 32 individuals from 2006

number of observations : 1039

number of individual observations : 4-80

country : United States

package : limdeppanel

JEL codes: C93, D64, D82, H41, L31, Z12

Chapter : 08

Usage

```
data(Donors)
```

Format

A dataframe containing:

id the id of the sollicitor

solsex the sex of the sollicitor

solmin does the sollicitor belongs to a minority ?

beauty beauty rating for the sollicitor

assertive assertive rating for the sollicitor

social social rating for the sollicitor

efficacy efficacy rating for the sollicitor

performance performance rating for the sollicitor

confidence confidence rating for the sollicitor

age age of the individual

sex sex of the individual

min does the individual belongs to a minority

treatment the treatment, one of "vcm", "sgift" and "lgift"

refgift has the individual refused the gift ?

donation the amount of the donation

prior has the individual been visited during the previous campaign ?

prtreat the treatment during the previous campaign, one of "none", "vcm", and "lottery"

prcontr has the individual made a donation during the previous campaign ?

prdonation the amount of the donation during the previous campaign

prsolsex the sex of the sollicitor during the previous campaign

prsolmin did the sollicitor of the previous campaign belong to a minority ?

prbeauty beauty rating for the sollicitor of the previous campaign

Source

American Economic Association Data Archive : https://www.aeaweb.org/aer/

References

Landry, Craig E.; Lange, Andreas; List, John A.; Price, Michael K. and Nicholas G. Rupp (2010) "Is a Donor in Hand Better Than Two in the Bush ? Evidence From a Natural Field Experiment", *American Economic Review*, **100(3)**, 958–983, doi: 10.1257/aer.100.3.958 .

Examples

```
#### Example 8-5

## ------------------------------------------------------------------------
## Not run:
data("Donors", package = "pder")
library("plm")
T3.1 <- plm(donation ~ treatment +  prcontr, Donors, index = "id")
T3.2 <- plm(donation ~ treatment * prcontr - prcontr, Donors, index = "id")
T5.A <- pldv(donation ~ treatment +  prcontr, Donors, index = "id",
             model = "random", method = "bfgs")
T5.B <- pldv(donation ~ treatment * prcontr - prcontr, Donors, index = "id",
             model = "random", method = "bfgs")

## End(Not run)
```

etw *Spatial weights matrix for EvapoTransp*

Description

Spatial weights matrix for the EvapoTransp data frame

Usage

```
data(etw)
```

Format

A 86x86 matrix with elements different from zero if area i and j are neighbours. Weights are row standardized.

Author(s)

Giovanni Millo

EvapoTransp *Evapotranspiration*

Description

a pseudo-panel of 86 areas from 2008

number of observations : 430

number of individual observations : 5

country : France

package : panel

Chapter : 10

Usage

```
data(EvapoTransp)
```

Format

A dataframe containing:

id observation site

period measuring period

et evapotranspiration

prec precipitation

meansmd mean soil moisture deficit

potet potential evapotranspiration

infil infiltration rate

biomass biomass

biomassp1 biomass in early growing season

biomassp2 biomass in main growth period

biomassp3 peak biomass

biomassp4 peak biomass after clipping

biomassp5 biomass in autumn

plantcover plant cover

softforbs soft-leaved forbs

tallgrass tall grass

diversity species diversity

matgram mat-forming graminoids

dwarfshrubs dwarf shrubs

legumes abundance of legumes

Source

kindly provided by the authors

References

Obojes, N.; Bahn, M.; Tasser, E.; Walde, J.; Inauen, N.; Hiltbrunner, E.; Saccone, P.; Lochet, J.; Clément, J. and S. Lavorel (2015) "Vegetation Effects on the Water Balance of Mountain Grasslands Depend on Climatic Conditions", *Ecohydrology*, **8(4)**, 552-569, doi: 10.1002/eco.1524 .

Examples

```
#### Example 10-14

## ------------------------------------------------------------------------
## Not run:
data("EvapoTransp", package = "pder")
data("etw", package = "pder")
if (requireNamespace("splm")){
```

```
    library("splm")
    evapo <- et ~ prec + meansmd + potet + infil + biomass + plantcover +
        softforbs + tallgrass + diversity + matgram + dwarfshrubs + legumes
    semsr.evapo <- spreml(evapo, data=EvapoTransp, w=etw,
                          lag=FALSE, errors="semsr")
    summary(semsr.evapo)
}

## ------------------------------------------------------------------------
library("plm")
if (requireNamespace("lmtest")){
    coeftest(plm(evapo, EvapoTransp, model="pooling"))
}

## ------------------------------------------------------------------------

if (requireNamespace("lmtest") & requireNamespace("splm")){
    coeftest(spreml(evapo, EvapoTransp, w=etw, errors="sem"))
}



#### Example 10-17

## ------------------------------------------------------------------------

if (requireNamespace("lmtest")){
    saremsrre.evapo <- spreml(evapo, data = EvapoTransp,
                              w = etw, lag = TRUE, errors = "semsr")
    summary(saremsrre.evapo)$ARCoefTable
    round(summary(saremsrre.evapo)$ErrCompTable, 6)
}

## End(Not run)
```

FinanceGrowth *Financial Institutions and Growth*

Description

5-yearly observations of 78 countries from 1960 to 1995

number of observations : 546

number of time-series : 7

country : world

package : panel

JEL codes: G20, O16, O47, C23, C33, O15

Chapter : 07

Usage

```
data(FinanceGrowth)
```

Format

A dataframe containing:

country country name

period period

growth growth rate * 100

privo log private credit / GDP

lly log liquid liabilities / GDP

btot log bank credit/total credit

lgdp log initial gdp per capita (PPP)

sec mean years of secondary schooling

gov log government spending / GDP

lbmp log(1 black market premium)

lpi log(1 + inflation rate)

trade log (imports + exports)/GDP

Source

http://www.cgdev.org/content/publications/detail/14256

References

Levine, Ross; Loayza, Norman and Thorsten Beck (2000) "Financial Intermediation and Growth: Causality and Causes", *Journal of Monetary Economics*, **46**, 31-77, doi: 10.1016/S03043932(00)00017-9 .

Roodman, David (2009) "A Note on the Theme of Two Many Instruments", *Oxford Bulletin of Economics An Statistics*, **71(1)**, 135–158, doi: 10.1111/j.14680084.2008.00542.x .

ForeignTrade *Foreign Trade of Developing Countries*

Description

yearly observations of 31 countries from 1963 to 1986

number of observations : 744

number of time-series : 24

country : developing countries

package : panelivreg

JEL codes: O19, C51, F17

Chapter : 02, 06

Usage

```
data(ForeignTrade)
```

Format

A dataframe containing:

country country name

year year

exports nominal exports deflated by the unit value of exports per capita

imports nominal imports deflated by the unit value of exports per capita

resimp official foreing reserves (in US dollars) divided by nominal imports (in US dollars)

gnp real GNP per capita

pgnp trend real GNP per capita calculated by fitting linear trend yit*=y0iexp(gi t), where y0i is the initial value of real gnp per capita for country i and gi is the ith country's average growth rate over 1964-1986

gnpw real genp for USA per capita

pm unit value of imports (in US dollars), 1980 = 100

px unit value of exports (in US dollars), 1980 = 100

cpi domestic CPI, 1980 = 100

pw US producer's price index, 1980 = 100

exrate exchange rate (price of US dollars in local currency), 1980 = 1

consump domestic consumption per capita,

invest domestic fixed gross investment per capita

income domestic disposable income per capita

pop population

reserves official foreing reserves (in US dollars)

money domestic money supply per capita

trend trend dummy, 1964 = 1

pwcpi log of us producer price index divided by domestic cpi

importspmpx log of nominal imports divided by export prices

pmcpi log of imports price divided by domestic cpi

pxpw log of exports price divided by domestic cpi

Source

Journal of Applied Econometrics Data Archive : http://qed.econ.queensu.ca/jae/

References

Kinal, T. and K. Lahiri (1993) "On the Estimation of Simultaneous-equations Error-components Models with An Application to a Model of Developing Country Foreign Trade", *Journal of Applied Economics*, **8**, 81-92, doi: 10.1002/jae.3950080107 .

Examples

```
#### Example 2-4

## ------------------------------------------------------------------------
library("plm")
data("ForeignTrade", package = "pder")
FT <- pdata.frame(ForeignTrade)
summary(FT$gnp)
ercomp(imports ~ gnp, FT)
models <- c("within", "random", "pooling", "between")
sapply(models, function(x) coef(plm(imports ~ gnp, FT, model = x))["gnp"])


#### Example 6-2

## ------------------------------------------------------------------------
data("ForeignTrade", package = "pder")
w1 <- plm(imports~pmcpi + gnp + lag(imports) + lag(resimp)  |
          lag(consump) + lag(cpi) + lag(income) + lag(gnp) + pm +
          lag(invest) + lag(money) + gnpw + pw + lag(reserves) +
          lag(exports) + trend + pgnp + lag(px),
          ForeignTrade, model = "within")
r1 <- update(w1, model = "random", random.method = "nerlove",
             random.dfcor = c(1, 1), inst.method = "baltagi")

## ------------------------------------------------------------------------
phtest(r1, w1)

## ------------------------------------------------------------------------
r1b <- plm(imports ~ pmcpi + gnp + lag(imports) + lag(resimp) |
            lag(consump) + lag(cpi) + lag(income) + lag(px) +
            lag(reserves) + lag(exports) | lag(gnp) + pm +
            lag(invest) + lag(money) + gnpw + pw  + trend + pgnp,
            ForeignTrade, model = "random", inst.method = "baltagi",
            random.method = "nerlove", random.dfcor = c(1, 1))

phtest(w1, r1b)

## ------------------------------------------------------------------------
rbind(within = coef(w1), ec2sls = coef(r1b)[-1])

## ------------------------------------------------------------------------
elast <- sapply(list(w1, r1, r1b),
                function(x) c(coef(x)["pmcpi"],
                              coef(x)["pmcpi"] / (1 - coef(x)["lag(imports)"])))
dimnames(elast) <- list(c("ST", "LT"), c("w1", "r1", "r1b"))
elast

## ------------------------------------------------------------------------
rbind(within = coef(summary(w1))[, 2],
      ec2sls = coef(summary(r1b))[-1, 2])
```

```
#### Example 6-4

## ------------------------------------------------------------------------
eqimp <- imports ~ pmcpi + gnp + lag(imports) +
                lag(resimp) | lag(consump) + lag(cpi) + lag(income) +
                lag(px) + lag(reserves) + lag(exports) | lag(gnp) + pm +
                lag(invest) + lag(money) + gnpw + pw  + trend + pgnp
eqexp <- exports ~ pxpw + gnpw + lag(exports) |
                lag(gnp) + pw + lag(consump) + pm + lag(px) + lag(cpi) |
                lag(money) + gnpw +  pgnp + pop + lag(invest) +
                lag(income) + lag(reserves) + exrate
r12 <- plm(list(import.demand = eqimp,
                export.demand = eqexp),
           data = ForeignTrade, index = 31, model = "random",
           inst.method = "baltagi", random.method = "nerlove",
           random.dfcor = c(1, 1))
summary(r12)

## ------------------------------------------------------------------------
rbind(ec2sls = coef(summary(r1b))[-1, 2],
      ec3sls = coef(summary(r12), "import.demand")[-1, 2])
```

`GiantsShoulders` *Impact of Institutions on Cumulative Research*

Description

yearly observations of 216 articles from 1970 to 2001

number of observations : 4880

number of time-series : 32

country : United States

package : countpanel

JEL codes: D02, D83, I23, O30

Chapter : 08

Usage

```
data(GiantsShoulders)
```

Format

A dataframe containing:

pair the pair article index

article the article index

brc material of the article is deposit on a Biological Ressource Center

pubyear publication year of the article

brcyear year of the deposit in brc of the material related to the article

year the year index

citations the number of citations

Source

American Economic Association Data Archive : https://www.aeaweb.org/aer/

References

Furman, Jeffrey L. and Scott Stern (2011) "Climbing Atop the Shoulders of Giants: the Impact of Institutions on Cumulative Research", *American Economic Review*, **101(5)**, 1933-1963, doi: 10.1257/aer.101.5.1933 .

Examples

```
#### Example 8-6

## ------------------------------------------------------------------------
## Not run:
data("GiantsShoulders", package = "pder")
head(GiantsShoulders)

## ------------------------------------------------------------------------

if (requireNamespace("dplyr")){
    library("dplyr")
    GiantsShoulders <- mutate(GiantsShoulders, age = year - pubyear)
    cityear <- summarise(group_by(GiantsShoulders, brc, age),
                         cit = mean(citations, na.rm = TRUE))
    GiantsShoulders <- mutate(GiantsShoulders,
                              window = as.numeric( (brc == "yes") &
                                                   abs(brcyear - year) <= 1),
                              post_brc = as.numeric( (brc == "yes") &
                                                     year - brcyear > 1),
                              age = year - pubyear)
    GiantsShoulders$age[GiantsShoulders$age == 31] <- 0
    #GiantsShoulders$year[GiantsShoulders$year
    #GiantsShoulders$year[GiantsShoulders$year
    GiantsShoulders$year[GiantsShoulders$year < 1975] <- 1970
  GiantsShoulders$year[GiantsShoulders$year >= 1975 & GiantsShoulders$year < 1980] <- 1975

    if (requireNamespace("pglm")){
        library("pglm")
        t3c1 <- lm(log(1 + citations) ~ brc + window + post_brc + factor(age),
                   data = GiantsShoulders)
        t3c2 <- update(t3c1, . ~ .+  factor(pair) + factor(year))
        t3c3 <- pglm(citations ~ brc + window + post_brc + factor(age) + factor(year),
                     data = GiantsShoulders, index = "pair",
```

```
                    effect = "individual", model = "within", family = negbin)
        t3c4 <- pglm(citations ~ window + post_brc + factor(age) + factor(year),
                    data = GiantsShoulders, index = "article",
                    effect = "individual", model = "within", family = negbin)
        ## screenreg(list(t3c2, t3c3, t3c4),
        ##           custom.model.names = c("ols: age/year/pair-FE",
       ##                                  "NB:age/year/pair-FE", "NB: age/year/article-FE"),
        ##           omit.coef="(factor)|(Intercept)", digits = 3)
    }
}

## End(Not run)
```

HousePricesUS *House Prices Data*

Description

yearly observations of 49 regions from 1976 to 2003

number of observations : 1421

number of time-series : 29

country : United States

package : hedprice

JEL codes: C51, R31

Chapter : 09, 10

Usage

```
data(HousePricesUS)
```

Format

A dataframe containing:

state state index

year year

names state name

plate state number plate index

region region index

region.name region name

price real house price index, 1980=100

income real per-capita income

pop total population

intrate real interest rate on borrowing

Source

Journal of Applied Econometrics Data Archive : http://qed.econ.queensu.ca/jae/

References

Holly, S.; Pesaran, M.G. and T. Yamagata (2010) "A Spatio-temporal Model of House Prices in the USA", *Journal of Econometrics*, **158(1)**, 160–173, doi: 10.1016/j.jeconom.2010.03.040 .

Millo, Giovanni (2015) "Narrow Replication of 'spatio-temporal Model of House Prices in the Usa', Using R", *Journal of Applied Econometrics*, **30(4)**, 703–704, doi: 10.1002/jae.2424 .

Examples

```
#### Example 4-11

## ------------------------------------------------------------------------
## Not run:
data("HousePricesUS", package = "pder")
library("plm")
php <- pdata.frame(HousePricesUS)

## ------------------------------------------------------------------------
cbind("rho"   = pcdtest(diff(log(php$price)), test = "rho")$statistic,
      "|rho|" = pcdtest(diff(log(php$price)), test = "absrho")$statistic)

## ------------------------------------------------------------------------
regions.names <- c("New Engl", "Mideast", "Southeast", "Great Lks",
                   "Plains", "Southwest", "Rocky Mnt", "Far West")
corr.table.hp <- cortab(diff(log(php$price)), grouping = php$region,
                        groupnames = regions.names)
colnames(corr.table.hp) <- substr(rownames(corr.table.hp), 1, 5)
round(corr.table.hp, 2)

## ------------------------------------------------------------------------
pcdtest(diff(log(price)) ~ diff(lag(log(price))) + diff(lag(log(price), 2)),
        data = php)

#### Example 9-2

## ------------------------------------------------------------------------
data("HousePricesUS", package = "pder")
swmod <- pvcm(log(price) ~ log(income), data = HousePricesUS, model= "random")
mgmod <- pmg(log(price) ~ log(income), data = HousePricesUS, model = "mg")
coefs <- cbind(coef(swmod), coef(mgmod))
dimnames(coefs)[[2]] <- c("Swamy", "MG")
coefs

#### Example 9-3

## ------------------------------------------------------------------------

if (requireNamespace("texreg")){
```

```
    library("texreg")
    data("RDSpillovers", package = "pder")
    fm.rds <- lny ~ lnl + lnk + lnrd
    mg.rds <- pmg(fm.rds, RDSpillovers, trend = TRUE)
    dmg.rds <- update(mg.rds, . ~ lag(lny) + .)
    screenreg(list('Static MG' = mg.rds, 'Dynamic MG'= dmg.rds), digits = 3)
    if (requireNamespace("msm")){
        library("msm")
        b.lr <- coef(dmg.rds)["lnrd"]/(1 - coef(dmg.rds)["lag(lny)"])
        SEb.lr <- deltamethod(~ x5 / (1 - x2),
                              mean = coef(dmg.rds), cov = vcov(dmg.rds))
        z.lr <- b.lr / SEb.lr
        pval.lr <- 2 * pnorm(abs(z.lr), lower.tail = FALSE)
        lr.lnrd <- matrix(c(b.lr, SEb.lr, z.lr, pval.lr), nrow=1)
        dimnames(lr.lnrd) <- list("lnrd (long run)", c("Est.", "SE", "z", "p.val"))
        round(lr.lnrd, 3)
    }
}


#### Example 9-4

## ------------------------------------------------------------------------
housep.np <- pvcm(log(price) ~ log(income), data = HousePricesUS, model = "within")
housep.pool <- plm(log(price) ~ log(income), data = HousePricesUS, model = "pooling")
housep.within <- plm(log(price) ~ log(income), data = HousePricesUS, model = "within")

d <- data.frame(x = c(coef(housep.np)[[1]], coef(housep.np)[[2]]),
                coef = rep(c("intercept", "log(income)"),
                           each = nrow(coef(housep.np))))
if (requireNamespace("ggplot2")){
    library("ggplot2")
    ggplot(d, aes(x)) + geom_histogram(col = "black", fill = "white", bins = 8) +
        facet_wrap(~ coef, scales = "free") + xlab("") + ylab("")
}


## ------------------------------------------------------------------------
summary(housep.np)

## ------------------------------------------------------------------------
pooltest(housep.pool, housep.np)
pooltest(housep.within, housep.np)


#### Example 9-5

## ------------------------------------------------------------------------
library("texreg")
cmgmod <- pmg(log(price) ~ log(income), data = HousePricesUS, model = "cmg")
screenreg(list(mg = mgmod, ccemg = cmgmod), digits = 3)

#### Example 9-6
```

```
## ------------------------------------------------------------------------
ccemgmod <- pcce(log(price) ~ log(income), data=HousePricesUS, model="mg")
summary(ccemgmod)

## ------------------------------------------------------------------------
ccepmod <- pcce(log(price) ~ log(income), data=HousePricesUS, model="p")
summary(ccepmod)



#### Example 9-8

## ------------------------------------------------------------------------
data("HousePricesUS", package = "pder")
price <- pdata.frame(HousePricesUS)$price
purtest(log(price), test = "levinlin", lags = 2, exo = "trend")
purtest(log(price), test = "madwu", lags = 2, exo = "trend")
purtest(log(price), test = "ips", lags = 2, exo = "trend")


#### Example 9-9

## ------------------------------------------------------------------------
tab5a <- matrix(NA, ncol = 4, nrow = 2)
tab5b <- matrix(NA, ncol = 4, nrow = 2)

for(i in 1:4) {
    mymod <- pmg(diff(log(income)) ~ lag(log(income)) +
                 lag(diff(log(income)), 1:i),
                 data = HousePricesUS,
                 model = "mg", trend = TRUE)
    tab5a[1, i] <- pcdtest(mymod, test = "rho")$statistic
    tab5b[1, i] <- pcdtest(mymod, test =  "cd")$statistic
}

for(i in 1:4) {
    mymod <- pmg(diff(log(price)) ~ lag(log(price)) +
                 lag(diff(log(price)), 1:i),
                 data=HousePricesUS,
                 model="mg", trend = TRUE)
    tab5a[2, i] <- pcdtest(mymod, test = "rho")$statistic
    tab5b[2, i] <- pcdtest(mymod, test =  "cd")$statistic
}

tab5a <- round(tab5a, 3)
tab5b <- round(tab5b, 2)
dimnames(tab5a) <- list(c("income", "price"),
                        paste("ADF(", 1:4, ")", sep=""))
dimnames(tab5b) <- dimnames(tab5a)

tab5a
tab5b
```

```
## ------------------------------------------------------------------------
php <- pdata.frame(HousePricesUS)
cipstest(log(php$price), type = "drift")
cipstest(diff(log(php$price)), type = "none")

## ------------------------------------------------------------------------
cipstest(resid(ccemgmod), type="none")
cipstest(resid(ccepmod), type="none")


#### Example 10-2

## ------------------------------------------------------------------------
data("usaw49", package="pder")
library("plm")
php <- pdata.frame(HousePricesUS)
pcdtest(php$price, w = usaw49)

## ------------------------------------------------------------------------

if (requireNamespace("splm")){
    library("splm")
    rwtest(php$price, w = usaw49, replications = 999)
}

## ------------------------------------------------------------------------
mgmod <- pmg(log(price) ~ log(income), data = HousePricesUS)
ccemgmod <- pmg(log(price) ~ log(income), data = HousePricesUS, model = "cmg")
pcdtest(resid(ccemgmod), w = usaw49)
rwtest(resid(mgmod), w = usaw49, replications = 999)

## End(Not run)
```

`IncomeMigrationH` *Income and Migration, Household Data*

Description

yearly observations of 317 households from 2000 to 2006

number of observations : 2219

number of time-series : 7

country : Indonesia

package : limdeppanel

JEL codes: F22, J43, O13, O15, Q11, Q12, R23

Chapter : 08

Usage

```
data(IncomeMigrationH)
```

Format

A dataframe containing:

household household index

year the year

migration a dummy indicating whether a household has any migrant departing in year t+1

price rice price shock

rain rain shock

land landholdings (ha)

Source

American Economic Association Data Archive : https://www.aeaweb.org/aer/

References

Bazzi, Samuel (2017) "Wealth Heterogeneity and the Income Elasticity of Migration", *American Economic Journal, Applied Economics*, **9(2)**, 219–255, doi: 10.1257/app.20150548 .

IncomeMigrationV *Income and Migration, Village Data*

Description

3-yearly observations of 44674 villages from 2005 to 2008

number of observations : 89348

number of time-series : 2

country : Indonesia

package : panellimdep

JEL codes: F22, J43, O13, O15, Q11, Q12, R23

Chapter : 08

Usage

```
data(IncomeMigrationV)
```

Format

A dataframe containing:

village village index

year the year

emigration share of the emigrants in the total population

district the district of the village

price rice price shock

rain rain shock

pareto Pareto parameter of the landholdings distribution

Source

American Economic Association Data Archive : https://www.aeaweb.org/aer/

References

Bazzi, Samuel (2017) "Wealth Heterogeneity and the Income Elasticity of Migration", *American Economic Journal, Applied Economics*, **9(2)**, 219–255, doi: 10.1257/app.20150548 .

`Index.jel` *JEL codes*

Description

- **C13** : Estimation: General
  - `TexasElectr` : Production of electricity in Texas
  - `Tileries` : Production of tileries in Egypt
- **C23** : Single Equation Models; Single Variables: Panel Data Models; Spatio-temporal Models
  - `FinanceGrowth` : Financial institutions and growth
  - `IneqGrowth` : Inequality and growth
  - `TexasElectr` : Production of electricity in Texas
  - `Tileries` : Production of tileries in Egypt
- **C33** : Multiple or Simultaneous Equation Models: Panel Data Models; Spatio-temporal Models
  - `FinanceGrowth` : Financial institutions and growth
  - `IneqGrowth` : Inequality and growth
- **C51** : Model Construction and Estimation
  - `ForeignTrade` : Foreign Trade of Developing countries
  - `HousePricesUS` : House Prices data
  - `RDPerfComp` : R and D performing companies

  - RDSpillovers : Research and development spillovers data
  - TexasElectr : Production of electricity in Texas
  - Tileries : Production of tileries in Egypt
  - TradeEU : Trade in the European Union
- **C78** : Bargaining Theory; Matching Theory
  - LateBudgets : Late Budgets
- **C90** : Design of Experiments: General
  - Seniors : Intergenerationals experiments
- **C92** : Design of Experiments: Laboratory, Group Behavior
  - CoordFailure : How to overcome organization failure in organization
- **C93** : Field Experiments
  - Donors : Dynamics of charitable giving
- **D02** : Institutions: Design, Formation, Operations, and Impact
  - GiantsShoulders : Impact of institutions on cumulative research
- **D23** : Organizational Behavior; Transaction Costs; Property Rights
  - CoordFailure : How to overcome organization failure in organization
- **D24** : Production; Cost; Capital; Capital, Total Factor, and Multifactor Productivity; Capacity
  - RDPerfComp : R and D performing companies
  - RDSpillovers : Research and development spillovers data
  - TexasElectr : Production of electricity in Texas
  - Tileries : Production of tileries in Egypt
  - TurkishBanks : Turkish Banks
- **D64** : Altruism; Philanthropy; Intergenerational Transfers
  - Donors : Dynamics of charitable giving
- **D72** : Political Processes: Rent-seeking, Lobbying, Elections, Legislatures, and Voting Behavior
  - DemocracyIncome : The relation between democraty and income
  - DemocracyIncome25 : The relation between democraty and income
  - LandReform : Politics and land reforms in India
  - LateBudgets : Late Budgets
  - Mafia : Mafia and Public Spending
  - Reelection : Deficits and reelection
  - RegIneq : Interregional redistribution and inequalities
  - ScrambleAfrica : The long-run effects of the scramble for Africa
- **D74** : Conflict; Conflict Resolution; Alliances; Revolutions
  - ScrambleAfrica : The long-run effects of the scramble for Africa
- **D82** : Asymmetric and Private Information; Mechanism Design
  - Donors : Dynamics of charitable giving
- **D83** : Search; Learning; Information and Knowledge; Communication; Belief; Unawareness

  - `Donors` : Dynamics of charitable giving
- **H61** : National Budget; Budget Systems
  - `LateBudgets` : Late Budgets
- **H62** : National Deficit; Surplus
  - `Reelection` : Deficits and reelection
- **H71** : State and Local Taxation, Subsidies, and Revenue
  - `Mafia` : Mafia and Public Spending
  - `RegIneq` : Interregional redistribution and inequalities
- **H72** : State and Local Budget and Expenditures
  - `LateBudgets` : Late Budgets
- **H73** : State and Local Government; Intergovernmental Relations: Interjurisdictional Differentials and Their Effects
  - `RegIneq` : Interregional redistribution and inequalities
- **H77** : Intergovernmental Relations; Federalism; Secession
  - `RegIneq` : Interregional redistribution and inequalities
- **I18** : Health: Government Policy; Regulation; Public Health
  - `Dialysis` : Diffusion of haemodialysis technology
- **I23** : Higher Education; Research Institutions
  - `GiantsShoulders` : Impact of institutions on cumulative research
- **J14** : Economics of the Elderly; Economics of the Handicapped; Non-labor Market Discrimination
  - `CallBacks` : Callbacks to job applications
  - `Seniors` : Intergenerationals experiments
- **J15** : Economics of Minorities, Races, Indigenous Peoples, and Immigrants; Non-labor Discrimination
  - `ScrambleAfrica` : The long-run effects of the scramble for Africa
- **J22** : Time Allocation and Labor Supply
  - `CallBacks` : Callbacks to job applications
- **J23** : Labor Demand
  - `CallBacks` : Callbacks to job applications
- **J26** : Retirement; Retirement Policies
  - `Seniors` : Intergenerationals experiments
- **J31** : Wage Level and Structure; Wage Differentials
  - `TexasElectr` : Production of electricity in Texas
  - `Tileries` : Production of tileries in Egypt
- **J43** : Agricultural Labor Markets
  - `IncomeMigrationH` : Income and Migration, household data
  - `IncomeMigrationV` : Income and Migration, village data

- **J64** : Unemployment: Models, Duration, Incidence, and Job Search
  - `CallBacks` : Callbacks to job applications
- **K42** : Illegal Behavior and the Enforcement of Law
  - `Mafia` : Mafia and Public Spending
  - `SeatBelt` : Seat belt usage and traffic fatalities
- **L31** : Nonprofit Institutions; NGOs; Social Entrepreneurship
  - `Donors` : Dynamics of charitable giving
- **L33** : Comparison of Public and Private Enterprises and Nonprofit Institutions; Privatization; Contracting Out
  - `TurkishBanks` : Turkish Banks
- **L82** : Entertainment; Media
  - `MagazinePrices` : Magazine prices
- **M12** : Personnel Management; Executives; Executive Compensation
  - `Seniors` : Intergenerationals experiments
- **M51** : Personnel Economics: Firm Employment Decisions; Promotions
  - `Seniors` : Intergenerationals experiments
- **O13** : Economic Development: Agriculture; Natural Resources; Energy; Environment; Other Primary Products
  - `IncomeMigrationH` : Income and Migration, household data
  - `IncomeMigrationV` : Income and Migration, village data
  - `LandReform` : Politics and land reforms in India
- **O15** : Economic Development: Human Resources; Human Development; Income Distribution; Migration
  - `FinanceGrowth` : Financial institutions and growth
  - `IncomeMigrationH` : Income and Migration, household data
  - `IncomeMigrationV` : Income and Migration, village data
  - `IneqGrowth` : Inequality and growth
  - `ScrambleAfrica` : The long-run effects of the scramble for Africa
- **O16** : Economic Development: Financial Markets; Saving and Capital Investment; Corporate Finance and Governance
  - `FinanceGrowth` : Financial institutions and growth
  - `IneqGrowth` : Inequality and growth
  - `TwinCrises` : Costs of currency and banking crises
- **O17** : Formal and Informal Sectors; Shadow Economy; Institutional Arrangements
  - `LandReform` : Politics and land reforms in India
  - `ScrambleAfrica` : The long-run effects of the scramble for Africa
- **O19** : International Linkages to Development; Role of International Organizations
  - `ForeignTrade` : Foreign Trade of Developing countries
  - `TwinCrises` : Costs of currency and banking crises

- **O30** : Innovation; Research and Development; Technological Change; Intellectual Property Rights: General
  - GiantsShoulders : Impact of institutions on cumulative research
- **O31** : Innovation and Invention: Processes and Incentives
  - Dialysis : Diffusion of haemodialysis technology
- **O32** : Management of Technological Innovation and R&D
  - RDSpillovers : Research and development spillovers data
- **O33** : Technological Change: Choices and Consequences; Diffusion Processes
  - RDSpillovers : Research and development spillovers data
- **O41** : One, Two, and Multisector Growth Models
  - Solow : Growth model
- **O47** : Empirical Studies of Economic Growth; Aggregate Productivity; Cross-Country Output Convergence
  - DemocracyIncome : The relation between democraty and income
  - DemocracyIncome25 : The relation between democraty and income
  - FinanceGrowth : Financial institutions and growth
  - IneqGrowth : Inequality and growth
  - Reelection : Deficits and reelection
  - Solow : Growth model
  - TwinCrises : Costs of currency and banking crises
- **Q11** : Agriculture: Aggregate Supply and Demand Analysis; Prices
  - IncomeMigrationH : Income and Migration, household data
  - IncomeMigrationV : Income and Migration, village data
- **Q12** : Micro Analysis of Farm Firms, Farm Households, and Farm Input Markets
  - IncomeMigrationH : Income and Migration, household data
  - IncomeMigrationV : Income and Migration, village data
- **Q15** : Land Ownership and Tenure; Land Reform; Land Use; Irrigation; Agriculture and Environment
  - LandReform : Politics and land reforms in India
- **R12** : Size and Spatial Distributions of Regional Economic Activity
  - RegIneq : Interregional redistribution and inequalities
- **R23** : Urban, Rural, Regional, Real Estate, and Transportation Economics: Regional Migration; Regional Labor Markets; Population; Neighborhood Characteristics
  - IncomeMigrationH : Income and Migration, household data
  - IncomeMigrationV : Income and Migration, village data
  - RegIneq : Interregional redistribution and inequalities
- **R31** : Housing Supply and Markets
  - HousePricesUS : House Prices data
- **R41** : Transportation: Demand, Supply, and Congestion; Travel Time; Safety and Accidents; Transportation Noise

  - SeatBelt : Seat belt usage and traffic fatalities
- **Z12** : Cultural Economics: Religion
  - Donors : Dynamics of charitable giving
- **Z13** : Economic Sociology; Economic Anthropology; Language; Social and Economic Stratification
  - ScrambleAfrica : The long-run effects of the scramble for Africa

---

IneqGrowth *Inequality and Growth*

Description

5-yearly observations of 266 world from 1961 to 1995

number of observations : 1862

number of time-series : 7

country : country

package : panel

JEL codes: O47, O15, C23, C33, O16

Chapter : 07

Usage

```
data(IneqGrowth)
```

Format

A dataframe containing:

country country name

period the period

growth growth rate

yssw years of secondary schooling among women, lagged

yssm years of secondary schooling among men, lagged

pinv price level of investment, lagged

lgdp log initial gdp per capita

gini gini index

Source

http://www.cgdev.org/content/publications/detail/14256

References

Forbes, Kristin J. (2000) "A Reassessment of the Relationship Between Inequality and Growth", *American Economic Review*, **90(4)**, 869-887, doi: 10.1257/aer.90.4.869 .

Roodman, David (2009) "A Note on the Theme of Two Many Instruments", *Oxford Bulletin of Economics An Statistics*, **71(1)**, 135–158, doi: 10.1111/j.14680084.2008.00542.x .

`LandReform` *Politics and Land Reforms in India*

Description

yearly observations of 89 villages from 1974 to 2003

number of observations : 2670

number of time-series : 30

country : India

package : panellimdep

JEL codes: D72, O13, O17, Q15

Chapter : 08

Usage

```
data(LandReform)
```

Format

A dataframe containing:

mouza village id number

year Year

district District

rplacul ratio of patta land registered to operational land

rpdrhh ratio of pattadar households to total households (hh)

rblacul ratio of barga land registered to operational land

rbgdrrghh ratio of bargadar registered hh to total hh

election election year dummy

preelect preelection year dummy

edwalfco to complete

erlesscu interpolated landless hh, gi

ermgcu interpolated mg hh, gi

ersmcu interpolated sm hh, gi

ermdcu interpolated md hh, gi

ercusmol ratio of land below 5 acres cultivable NOT extrapolated

ercubgol ratio of land above 12.5 acres cultivable

erillnb interpolated ratio of illiterate non big hh

erlow interpolated ratio of low caste hh

ratleft0 Left Front share in GP, == 0 for 1974

dwalfco Assembly average vote difference LF-INC, district

inflat Inflation in last 5 years in CPI for Agricultural Labourers

smfempyv Year variation in Employment in Small Scale Industrial Units registered with Dir

incseats INC seats / Total seats in Lok Sabha

lfseats Ratio of LF seats in parliament

inflflag Interaction between Inflation and ratleft lagged

inclflag Interaction between INC seats and ratleft lagged

lflflag Interaction between LF seats and ratleft lagged

ratleft Left Front share in GP, ==share of assembly seats for 1974

infiw to complete

infumme to complete

infal to complete

gp Gran Panchayat

Source

American Economic Association Data Archive : https://www.aeaweb.org/aer/

References

Bardhan, Pranab and Dilip Mookherjee (2010) "Determinants of Redistributive Politics: An Empirical Analysis of Land Reform in West Bengal, India", *American Economic Review*, **100(4)**, 1572–1600, doi: 10.1257/aer.100.4.1572 .

`LateBudgets` *Late Budgets*

Description

yearly observations of 48 States from 1978 to 2007

number of observations : 1440

number of time-series : 30

country : United States

package : limdeppanel

JEL codes: C78, D72, H61, H72

Chapter : 08

Usage

```
data(LateBudgets)
```

Format

A dataframe containing:

state the state

year the year

late late budget ?

dayslate number of days late for the budget

unempdiff unemployment variation

splitbranch split branch

splitleg split legislature

elecyear election year

endbalance end of year balances in the general fund and stabilization fund

demgov democrat governor ?

lameduck lameduck

govexp number of years since the incumbent governor took office

newgov new governor ?

pop the polulation

kids percentage of population aged 5-17

elderly percentage of population aged 65 or older

nocarry does the state law does not allow a budget deficit to be carried over to the next fiscal year ?

supmaj is a super majority required to pass each budget ?

fulltimeleg full time legislature ?

shutdown shutdown provision ?

black percentage of blacks

graduate percentage of graduates

censusresp census response rate

fiveyear five year dummies, one of '93-97', '98-02', '03-07'

deadline is there a deadline ? one of 'none', 'soft' and 'hard'

Source

American Economic Association Data Archive : https://www.aeaweb.org/aer/

References

Andersen, Asger Lau; Lassen, David Dreyer and Lasse Holboll Westh Nielsen (2012) "Late Budgets", *American Economic Journal, Economic Policy*, **4(4)**, 1-40, doi: 10.1257/pol.4.4.1 .

Examples

```
#### Example 8-4

## ------------------------------------------------------------------------
data("LateBudgets", package = "pder")
library("plm")
LateBudgets$dayslatepos <- pmax(LateBudgets$dayslate, 0)
LateBudgets$divgov <- with(LateBudgets,
                           factor(splitbranch == "yes" |
                                  splitleg == "yes",
                                  labels = c("no", "yes")))
LateBudgets$unemprise <- pmax(LateBudgets$unempdiff, 0)
LateBudgets$unempfall <- - pmin(LateBudgets$unempdiff, 0)
form <- dayslatepos ~ unemprise + unempfall + divgov + elecyear +
    pop + fulltimeleg + shutdown + censusresp + endbalance + kids +
    elderly + demgov + lameduck + newgov + govexp + nocarry +
    supmaj + black + graduate

## ------------------------------------------------------------------------
FEtobit <- pldv(form, LateBudgets)
summary(FEtobit)
```

Mafia *Mafia and Public Spending*

Description

yearly observations of 95 provinces from 1986 to 1999

number of observations : 1330

number of time-series : 14

country : Italy

package : panelivreg

JEL codes: D72, E62, H71, K42

Chapter : 06

Usage

```
data(Mafia)
```

Format

A dataframe containing:

province the province (95)

region the region (19)

year the year

pop the population

y percentage growth of real per-capita value added

g annual variation of the per-capita public investment in infrastructure divided by lagged real per-capita value added

cd number of municipalities placed under the administration of external commissioners

cds1 same as cd, provided that the official deccree is publisehd in the first semester of the year

cds2 same as cd, provided that the average number of days betwen the dismissal of the city concil and the year end is less than 180

u1 change in the log of per-capita employment

u2 change in the log of per-capita hours of wage supplement provided by the unemployment insurance scheme

mafiosi first difference of the number of people reported by the police forces to the judicial authority because of mafia-type association

extortion first difference of the number of people reported by the police forces to the judicial authority because of extorsion

corruption1 first difference of the number of people reported by the police forces to the judicial authority because of corruption

corruption2 first difference of the number of crimes reported by the police forces to the judicial authority because of corruption

murder first difference of the number of people reported by the police forces to the judicial authority because of murder related to mafia activity

Source

American Economic Association Data Archive : https://www.aeaweb.org/aer/

References

Acconcia, Antonio; Corsetti, Giancarlo and Saviero Simonelli (2014) "Mafia and Public Spending: Evidence on the Fiscal Multimplier Form a Quasi-experiment", *American Economic Review*, **104(7)**, 2189-2209, doi: 10.1257/aer.104.7.2185 .

`MagazinePrices` *Magazine Prices*

Description

yearly observations of 38 magazines from 1940 to 1980

number of observations : 1262

number of time-series : 41

country : United States

package : binomialpanel

JEL codes: L82

Chapter : 08

Usage

```
data(MagazinePrices)
```

Format

A dataframe containing:

year the year

magazine the magazine name

price the price of the magazine in january

change has the price changed between january of the current year and january of the following year ?

length number of years since the previous price change

cpi gdp deflator index

cuminf cummulative change in inflation since the previous price change

sales single copy sales of magazines for magazine industry

cumsales cumulative change in magazine industry sales since previous price change

included is the observation included in the econometric analysis ?

id group index numbers used for the conditional logit estimation

Source

Journal of Applied Econometrics Data Archive : http://qed.econ.queensu.ca/jae/

References

Willis, Jonathan L. (2006) "Magazine Prices Revisited", *Journal of Applied Econometrics*, **21(3)**, 337-344, doi: 10.1002/jae.836 .

Cecchetti, Stephen G. (1986) "The Frequency of Price Adjustment, a Study of Newsstand Prices of Magazines", *Journal of Econometrics*, **31**, 255-274, doi: 10.1016/03044076(86)900618 .

Examples

```
#### Example 8-3

## ------------------------------------------------------------------------
data("MagazinePrices", package = "pder")
logitS <- glm(change ~ length + cuminf + cumsales, data = MagazinePrices,
              subset = included == 1, family = binomial(link = 'logit'))
logitD <- glm(change ~ length + cuminf + cumsales + magazine,
              data = MagazinePrices,
              subset = included == 1, family = binomial(link = 'logit'))

if (requireNamespace("survival")){
    library("survival")
    logitC <- clogit(change ~ length + cuminf + cumsales + strata(id),
```

```
                    data = MagazinePrices,
                    subset = included == 1)
    if (requireNamespace("texreg")){
        library("texreg")
        screenreg(list(logit = logitS, "FE logit" = logitD,
                       "cond. logit" = logitC), omit.coef = "magazine")
    }
}
```

RDPerfComp *R and D Performing Companies*

Description

yearly observations of 509 firms from 1982 to 1989

number of observations : 4072

number of time-series : 8

country : United States

package : panel

JEL codes: C51, D24

Chapter : 07

Usage

```
data(RDPerfComp)
```

Format

A dataframe containing:

id firm identifier

year year

y production in logs

n labor in logs

k capital in logs

Source

author's website https://www.nuffield.ox.ac.uk/users/bond/index.html

References

Blundell, Richard and Stephen Bond (2000) "GMM Estimation with Persistent Panel Data: An Application to Production Functions", *Econometric Reviews*, **19(3)**, 321-340, doi: 10.1080/07474930008800475
.

RDSpillovers *Research and Development Spillovers Data*

Description

a cross-section of 119 industries from 1980 to 2005

country : world

package : panel

JEL codes: C51, D24, O32, O33

Chapter : 04, 05, 09

Usage

```
data(RDSpillovers)
```

Format

A dataframe containing:

id country-industry index

year year

country country

sector manufacturing sector as SIC 15-37, excluding SIC 23

lny log output

lnl log of labour input

lnk log of physical capital stock

lnrd log of RD capital stock

Source

author's web site https://sites.google.com/site/medevecon/home

References

Eberhardt, M.; Helmers, C. and H. Strauss (2013) "Do Spillovers Matter in Estimating Private Returns to R and D?", *The Review of Economics and Statistics*, **95(2)**, 436–448, doi: 10.1162/REST_a_00272 .

Examples

```
#### Example 4-10

## ------------------------------------------------------------------------
## Not run:
data("RDSpillovers", package = "pder")
library("plm")
fm.rds <- lny ~ lnl + lnk + lnrd

## ------------------------------------------------------------------------
pcdtest(fm.rds, RDSpillovers)

## ------------------------------------------------------------------------
rds.2fe <- plm(fm.rds, RDSpillovers, model = "within", effect = "twoways")
pcdtest(rds.2fe)

## ------------------------------------------------------------------------
cbind("rho"  = pcdtest(rds.2fe, test = "rho")$statistic,
      "|rho|"= pcdtest(rds.2fe, test = "absrho")$statistic)


#### Example 5-10

## ------------------------------------------------------------------------
data("RDSpillovers", package = "pder")
pehs <- pdata.frame(RDSpillovers, index = c("id", "year"))
ehsfm <- lny ~ lnl + lnk + lnrd
phtest(ehsfm, pehs, method = "aux")

## ------------------------------------------------------------------------
phtest(ehsfm, pehs, method = "aux", vcov = vcovHC)


#### Example 5-15

## ------------------------------------------------------------------------
fm <- lny ~ lnl + lnk + lnrd

## ------------------------------------------------------------------------


if (requireNamespace("lmtest")){
    library("lmtest")
    gglsmodehs <- pggls(fm, RDSpillovers, model = "pooling")
    coeftest(gglsmodehs)

    feglsmodehs <- pggls(fm, RDSpillovers, model = "within")
    coeftest(feglsmodehs)

    phtest(gglsmodehs, feglsmodehs)

    fdglsmodehs <- pggls(fm, RDSpillovers, model = "fd")
```

```
    fee <- resid(feglsmodehs)
    dbfee <- data.frame(fee=fee, id=attr(fee, "index")[[1]])
    coeftest(plm(fee~lag(fee)+lag(fee,2), dbfee, model = "p", index="id"))

    fde <- resid(fdglsmodehs)
    dbfde <- data.frame(fde=fde, id=attr(fde, "index")[[1]])
    coeftest(plm(fde~lag(fde)+lag(fde,2), dbfde, model = "p", index="id"))

    coeftest(fdglsmodehs)
}


#### Example 9-7

## ------------------------------------------------------------------------
ccep.rds <- pcce(fm.rds, RDSpillovers, model="p")
if (requireNamespace("lmtest")){
    library("lmtest")
    ccep.tab <- cbind(coeftest(ccep.rds)[, 1:2],
                      coeftest(ccep.rds, vcov = vcovNW)[, 2],
                      coeftest(ccep.rds, vcov = vcovHC)[, 2])
    dimnames(ccep.tab)[[2]][2:4] <- c("Nonparam.", "vcovNW", "vcovHC")
    round(ccep.tab, 3)
}


## ------------------------------------------------------------------------
autoreg <- function(rho = 0.1, T = 100){
  e <- rnorm(T+1)
  for (t in 2:(T+1)) e[t] <- e[t]+rho*e[t-1]
  e
}
set.seed(20)

f <- data.frame(time = rep(0:40, 2),
                rho = rep(c(0.2, 1), each = 41),
                y = c(autoreg(rho = 0.2, T = 40),
                      autoreg(rho = 1, T = 40)))
if (requireNamespace("ggplot2")){
    library("ggplot2")
    ggplot(f, aes(time, y)) + geom_line() + facet_wrap(~ rho) + xlab("") + ylab("")

    autoreg <- function(rho = 0.1, T = 100){
        e <- rnorm(T)
        for (t in 2:(T)) e[t] <- e[t] + rho *e[t-1]
        e
    }
    tstat <- function(rho = 0.1, T = 100){
        y <- autoreg(rho, T)
        x <- autoreg(rho, T)
        z <- lm(y ~ x)
        coef(z)[2] / sqrt(diag(vcov(z))[2])
```

```
    }
    result <- c()
    R <- 1000
    for (i in 1:R) result <- c(result, tstat(rho = 0.2, T = 40))
    quantile(result, c(0.025, 0.975))
    prop.table(table(abs(result) > 2))


    result <- c()
    R <- 1000
    for (i in 1:R) result <- c(result, tstat(rho = 1, T = 40))
    quantile(result, c(0.025, 0.975))
    prop.table(table(abs(result) > 2))


    R <- 1000
    T <- 100
    result <- c()
    for (i in 1:R){
        y <- autoreg(rho=1, T=100)
        Dy <- y[2:T] - y[1:(T-1)]
        Ly <- y[1:(T-1)]
        z <- lm(Dy ~ Ly)
        result <- c(result, coef(z)[2] / sqrt(diag(vcov(z))[2]))
    }

    ggplot(data.frame(x = result), aes(x = x)) +
        geom_histogram(fill = "white", col = "black",
                       bins = 20, aes(y = ..density..)) +
        stat_function(fun = dnorm) + xlab("") + ylab("")


    prop.table(table(result < -1.64))
}

## End(Not run)
```

Reelection *Deficits and Reelection*

Description

yearly observations of 75 countries from 1960 to 2003

number of observations : 439

number of time-series : 16

country : world

package : panelbinomial

JEL codes: D72, E62, H62, O47

Chapter : 08

Usage

```
data(Reelection)
```

Format

A dataframe containing:

country the country

year the year

narrow TRUE if the observation belongs to the narrow data set

reelect one if the incumbent was reelected and zero otherwise

ddefterm the change in the ratio of the government surplus to gdp in the two years preeceding the election year, relative to the two previous years

ddefey the change in the government surplus ratio to gdpin the election year, compared to the previous year

gdppc the average growth rate of real per capita gdp during the leader's current term

dev one for developped countries, 0 otherwise

nd one for a new democratic country, 0 otherwise

maj one for majoritarian electoral system, 0 otherwise

Source

American Economic Association Data Archive : https://www.aeaweb.org/aer/

References

Adi Brender and Allan Drazen (2008) "How Do Budget Deficits and Economic Growth Affect Reelection Prospects? Evidence From a Large Panel of Countries", *American Economic Review*, **98(5)**, 2203-2220, doi: 10.1257/aer.98.5.2203 .

Examples

```
#### Example 8-1

## ------------------------------------------------------------------------
## Not run:
library("plm")
data("Reelection", package = "pder")

## ------------------------------------------------------------------------
elect.l <- glm(reelect ~ ddefterm + ddefey + gdppc + dev + nd + maj,
          data = Reelection, family = "binomial", subset = narrow)
l2 <- update(elect.l, family = binomial)
l3 <- update(elect.l, family = binomial())
l4 <- update(elect.l, family = binomial(link = 'logit'))
```

```
## ------------------------------------------------------------------------
elect.p <- update(elect.l, family = binomial(link = 'probit'))

## ------------------------------------------------------------------------

if (requireNamespace("pglm")){
    library("pglm")
    elect.pl <- pglm(reelect ~ ddefterm + ddefey + gdppc + dev + nd + maj,
                 Reelection, family = binomial(link = 'logit'),
                subset = narrow)
    elect.pp <- pglm(reelect ~ ddefterm + ddefey + gdppc + dev + nd + maj,
                 Reelection, family = binomial(link = 'probit'),
                subset = narrow)
}

## End(Not run)
```

RegIneq *Interregional Redistribution and Inequalities*

Description

yearly observations of 17 countries from 1982 to 1999

number of observations : 102

number of time-series : 6

country : oecd

package : panel

JEL codes: D72, H23, H71, H73, H77, R12, R23

Chapter : 07

Usage

```
data(RegIneq)
```

Format

A dataframe containing:

country the country

period the period

regineq coefficient of variatio of regional gdp per capita

gdppc real gross domestic product per capita

pop total population

popgini gini coefficient of regional population size

urban share of urban living population

social total government social expenditures as share of gdp

unempl unemployment rate

dec sub-national expenditures as share of total government expenditures

transrev grants received by national and sub-national governments from other levels of government as share of total government revenues

transaut sub-national non autonomous revenues as share of total government revenues

Source

Review of Economic Studies' web site https://academic.oup.com/restud

References

Anke S. Kessler and Nico A. Hansen and Christian Lessmann (2011) "Interregional Redistribution and Mobility in Federations: a Positive Approach", *Review of Economic Studies*, **78(4)**, 1345-1378, doi: 10.1093/restud/rdr003 .

`ScrambleAfrica` *The Long-run Effects of the Scramble for Africa*

Description

a pseudo-panel of 49 countries

number of observations : 1212

number of individual observations : 2-112

country : Africa

package : countpanel

JEL codes: D72, D74, F51, J15, O15, O17, Z13

Chapter : 08

Usage

```
data(ScrambleAfrica)
```

Format

A dataframe containing:

country country code

group ethnic group name

conflicts number of conflicts

split dummy for partitioned ethnic area

spillover spillover index, the fraction of adjacent groups in the same country that are partitioned

region the region

pop population according to the first post-independance census

area land area

lake lakes dummy

river rivers dummy

capital dummy if a capital city falls in the homeland of an ethnic group

borderdist distance of the centroid of the area from the national border

capdist distance of the centroid of the area from the capital

seadist distance of the centroid of the area from the sea coast

coastal dummy for areas that are by the sea coast

meanelev mean elevation

agriculture index of land suitability for agriculture

diamond diamond mine indicator

malaria malaria stability index

petroleum oil field indicator

island island dummy

city1400 dummy for areas with major city in 1400

Source

American Economic Association Data Archive : https://www.aeaweb.org/aer/

References

Michalopoulos, Stelios and Elias Papaioannou (2016) "The Long-run Effects of the Scramble for Africa", *American Economic Review*, **106(7)**, 1802–1848, doi: 10.1257/aer.20131311 .

`SeatBelt` *Seat Belt Usage and Traffic Fatalities*

Description

yearly observations of 51 states from 1983 to 1997

number of observations : 765

number of time-series : 15

country : United States

package : panel

JEL codes: R41, K42

Chapter : 06

Usage

```
data(SeatBelt)
```

Format

A dataframe containing:

state the state code

year the year

farsocc the number of traffic fatalities of drivers and passengers (of any seating position) of a motor vehicule in transport

farsnocc the number of traffic fatalities of pedestrians and bicyclists

usage rate of seat belt usage

percapin median income in current US dollars

unemp unemployment rate

meanage mean age

precentb the percentage of african-americans in the state population

precenth the percentage of people of hispanic origin in the state population

densurb traffic density urban ; registered vehicules per unit length of urban roads in miles

densrur traffic density rural ; registered vehicules per unit length of urban roads in miles

viopcap number of violent crimes (homicide, rape and robbery) per capita

proppcap number of preperty rimes (burglary, larceny and auto theft) per capita

vmtrural vehicule miles traveled on rural roads

vmturban vehicule miles traveled on urban roads

fueltax fuel tax (in curent cents)

lim65 65 miles per hour speed limit (55 mph is the base category)

lim70p 70 miles per hour or above speed limit (55 mph is the base caegory)

mlda21 a dummy variable that is equal to 1 for a minimum for a minimum legal drinking age of 21 years (18 years is the base category)

bac08 a dummy variable that is equal to 1 foe a maximum of 0.08 blood alcohol content (0.1 is the base category)

ds a dummy equal to 1 for the periods in which the state had a secondary-enforcement mandatory seat belt law, or a primary-enforcement law that preceded by a secondary-enforcement law (no seat belt law is the base category)

dp a dummy variable eqal to 1 for the periods in which the state had a primary-enforcement mandatory seat belt law that was not preceded by a secondary-enforcement law (no seat belt is the base category)

dsp a dummy variable equal to 1 for the periods in which the state had a primary-enforcement mandatory seat belt law that was preceded by a secondary enforcement law (no seat belt law is the base category

Source

author's website https://leinav.people.stanford.edu

References

Cohen, Alma and Liran Einav (2003) "The Effects of Mandatory Seat Belt Laws on Driving Behavior and Traffic Fatalities", *The Review of Economics and Statistics*, **85(4)**, 828-843, doi: 10.2139/ssrn.293582 .

Examples

```
#### Example 6-1

## ------------------------------------------------------------------------
## Not run:
library("plm")

## ------------------------------------------------------------------------
y ~ x1 + x2 + x3 | x1 + x3 + z
y ~ x1 + x2 + x3 | . - x2 + z

## ------------------------------------------------------------------------

data("SeatBelt", package = "pder")
SeatBelt$occfat <- with(SeatBelt, log(farsocc / (vmtrural + vmturban)))
ols <- plm(occfat ~ log(usage) + log(percapin) + log(unemp) + log(meanage) +
           log(precentb) + log(precenth)+ log(densrur) +
           log(densurb) + log(viopcap) + log(proppcap) +
           log(vmtrural) + log(vmturban) + log(fueltax) +
           lim65 + lim70p + mlda21 + bac08, SeatBelt,
           effect = "time")
fe <- update(ols, effect = "twoways")
ivfe <- update(fe, . ~ . |  . - log(usage) + ds + dp +dsp)

rbind(ols = coef(summary(ols))[1,],
      fe = coef(summary(fe))[1, ],
      w2sls = coef(summary(ivfe))[1, ])

## ------------------------------------------------------------------------
SeatBelt$noccfat <- with(SeatBelt, log(farsnocc / (vmtrural + vmturban)))
nivfe <- update(ivfe, noccfat ~ . | .)
coef(summary(nivfe))[1, ]

## End(Not run)
```

Seniors *Intergenerationals Experiments*

Description

a pseudo-panel of 159 Individuals

number of observations : 2703

number of individual observations : 17

country : France

package : panellimdep

JEL codes: C90, J14, J26, M12, M51

Chapter : 08

Usage

```
data(Seniors)
```

Format

A dataframe containing:

id individual number of each subject

period from 1 to 17

session from 1 to 12

firm 1 if working subject, 0 otherwise

firmx 1 if the firm is X, 0 if the firm is Y

order 1 if the treatment with no information on the generation of the group is played first in the Public Good game, 0 otherwise

gender 1 if male subject, 0 if female subject

manager 1 if the subject is a manager, 0 otherwise

student 1 if the subject is a student, 0 otherwise

retir 1 if retiree, 0 otherwise

senior 1 if the subject is a senior, 0 otherwise

seniord 1 if the subject reports s/he is a senior, 0 if junior

workingsenior 1 if the subject is a working senior, 0 otherwise

workingjunior 1 if the subject is a working junior, 0 otherwise

information 1 if information is given on the generation composition of the group, 0 otherwise

nbseniors number of seniors in the group, excluding the subject

homogend 1 if the group is homogenous in terms of declared generation, 0 otherwise

homodgenck 1 if the group is homogenous in terms of declared generation and this is common information, 0 otherwise

contribution amount of the contribution to the public good (from 0 to 20)

pot amount of the public good (from 0 to 60)

potlag amount of the public good in the previous period (from 0 to 60)

potimean amount of the public good, excluding the subject's contribution (from 0 to 40)

potimeanlag amount of the public good in the previous period, excluding the subject's contribution (from 0 to 40)

payoffpggame payoff in the public good game

desirnbseniors desired number of seniors co-participants in the Selection treatment (from 0 to 2)

invest amount invested in the risky lotery

payoffriskgame payoff in the investment game

letters 1 if letters are A M F U R I P , 0 if they are OATFNED

idicompet individual number of the co-participant in the Task game

seniordopponent 1 if the co-participant in the Task game reports s/he is a senior, 0 otherwise

seniori 1 if the co-participant in the Task game is a senior

option 1 if the subject has chosen the tournament, 0 otherwise

option0 1 if the co-participant has chosen the tournament, 0 otherwise

twoperstour 1 if both participants have chosen the tournament, 0 otherwise

beliefself number of words the subject believes s/he will create

beliefseniors number of words the subject believes the seniors will create on average

beliefjuniors number of words the subject believes the juniors will create on average

beliefsmatchs number of words the subject believes the seniors will create on average when matched with a senior

beliefjmatchj number of words the subject believes the juniors will create on average when matched with a junior

relatabil 1 if the subject believes s/he can create more words than the generation of his/her co-participant, 0 otherwise

performance number of words actually created

perfi number of words actually created by the co-participant

payoffcompetitiongame payoff in the Task game

expesenck 1 if the subject has been informed that s/he was interacting with seniors in the Public Good game, 0 otherwise

potlagsenior Amount of the pot in the previous period * the subject is a senior

heterogend 1 if the group mixes the two generations, 0 otherwise

Source

American Economic Association Data Archive : https://www.aeaweb.org/aer/

References

Charness, Gary and Marie-Claire Villeval (2009) "Cooperation and Competition in Intergenerational Experiments in the Field and the Laboratory", *American Economic Review*, **99(3)**, 956–978, doi: 10.1257/aer.99.3.956 .

Solow *Growth Model*

Description

yearly observations of 97 countries from 1960 to 1985

number of observations : 576

number of time-series : 6

country : world

package : panel

JEL codes: O47, O41

Chapter : 07

Usage

```
data(Solow)
```

Format

A dataframe containing:

id country id

year year

lgdp log of gdp per capita

lsrate log of the saving rate, approximated by the investement rate

lpopg log of population growth + 0.05 (which is an approximation of the sum of the rate of labor-augmenting technological progress and of the rate of depreciation of physical capital)

Source

author's website https://www.nuffield.ox.ac.uk/users/bond/index.html

References

Caselli, Francesco; Esquivel, Gerardo and Fernando Lefort (1996) "Reopening the Convergence Debate: a New Look at Cross-country Growth Empirics", *Journal of Economic Growth*, **1**, 363-389, doi: 10.1007/BF00141044 .

Bond, Stephen; Hoeffler, Anke and Johnatan Temple (2001) "GMM Estimation of Empirical Growth Model", *CEPR Discussion Paper*, **3048**, 1-33.

TexasElectr *Production of Electricity in Texas*

Description

yearly observations of 10 firms from 1966 to 1983

number of observations : 180

number of time-series : 18

country : Texas

package : productionpanel

JEL codes: D24, C13, C51, C23, J31

Chapter : 02, 03

Usage

```
data(TexasElectr)
```

Format

A dataframe containing:

id the firm identifier

year the year, from 1966 to 1983

output output

pfuel price of fuel

plab price of labor

pcap price of capital

expfuel expense in fuel

explab expense in labor

expcap expense in capital

Source

Journal of Applied Econometrics Data Archive : http://qed.econ.queensu.ca/jae/

References

Kumbhakar SC (1996) "Estimation of Cost Efficiency with Heteroscedasticity: An Application to Electric Utilities", *Journal of the Royal Statistical Society, Series D*, **45**, 319–335.

Horrace and Schmidt (1996) "Confidence Statements for Efficiency Estimates From Stochastic Frontier Models", *Journal of Productity Analysis*, **7**, 257–282, doi: 10.1007/BF00157044 .

Horrace and Schmidt (2012) "Multiple Comparisons with the Best, with Economic Applications", *Journal of Applied Econometrics*, **15(1)**, 1–26, doi: 10.1002/(SICI)10991255(200001/02)15:1<1::AID-JAE551>3.0.CO;2Y .

Examples

```
#### Example 2-6

## ------------------------------------------------------------------------
data("TexasElectr", package = "pder")
library("plm")
TexasElectr$cost <- with(TexasElectr, explab + expfuel + expcap)
TE <- pdata.frame(TexasElectr)
summary(log(TE$output))
ercomp(log(cost) ~ log(output), TE)
models <- c("within", "random", "pooling", "between")
sapply(models, function(x)
       coef(plm(log(cost) ~ log(output), TE, model = x))["log(output)"])

#### Example 3-2

## ------------------------------------------------------------------------
data("TexasElectr", package = "pder")

if (requireNamespace("dplyr")){
    library("dplyr")
    TexasElectr <- mutate(TexasElectr,
                          pf = log(pfuel / mean(pfuel)),
                          pl = log(plab / mean(plab)) - pf,
                          pk = log(pcap / mean(pcap)) - pf)

## ------------------------------------------------------------------------
    TexasElectr <- mutate(TexasElectr, q = log(output / mean(output)))

## ------------------------------------------------------------------------
    TexasElectr <- mutate(TexasElectr,
                          C = expfuel + explab + expcap,
                          sl = explab / C,
                          sk = expcap / C,
                          C = log(C / mean(C)) - pf)

## ------------------------------------------------------------------------
    TexasElectr <- mutate(TexasElectr,
                          pll = 1/2 * pl ^ 2,
                          plk = pl * pk,
                          pkk = 1/2 * pk ^ 2,
                          qq = 1/2 * q ^ 2)

## ------------------------------------------------------------------------
    cost <- C ~ pl + pk + q + pll + plk + pkk + qq
    shlab <- sl ~ pl + pk
    shcap <- sk ~ pl + pk

## ------------------------------------------------------------------------
    R <- matrix(0, nrow = 6, ncol = 14)
    R[1, 2] <- R[2, 3] <- R[3, 5] <- R[4, 6] <- R[5, 6] <- R[6, 7] <- 1
    R[1, 9] <- R[2, 12] <- R[3, 10] <- R[4, 11] <- R[5, 13] <- R[6, 14] <- -1
```

```
## ------------------------------------------------------------------------
    z <- plm(list(cost = C ~ pl + pk + q + pll + plk + pkk + qq,
                  shlab = sl ~ pl + pk,
                  shcap = sk ~ pl + pk),
             TexasElectr, model = "random",
             restrict.matrix = R)
    summary(z)
}
```

`Tileries` *Production of Tileries in Egypt*

Description

weeklyly observations of 25 firms from 1982 to 1983

number of observations : 483

number of time-series : 22

country : Egypt

package : panelproduction

JEL codes: D24, C13, C51, C23, J31

Chapter : 01, 03

Usage

```
data(Tileries)
```

Format

A dataframe containing:

id firm id

week week (3 weeks aggregated)

area one of `"fayoum"` and `"kalyubiya"`

output output

labor labor hours

machine machine hours

Source

Journal of Applied Econometrics Data Archive : http://qed.econ.queensu.ca/jae/

References

Horrace and Schmidt (1996) "Confidence Statements for Efficiency Estimates From Stochastic Frontier Models", *Journal of Productity Analysis*, **7**, 257–282, doi: 10.1007/BF00157044 .

Horrace and Schmidt (2012) "Multiple Comparisons with the Best, with Economic Applications", *Journal of Applied Econometrics*, **15(1)**, 1–26, doi: 10.1002/(SICI)10991255(200001/02)15:1<1::AID-JAE551>3.0.CO;2Y .

Seale J.L. (1990) "Estimating Stochastic Frontier Systems with Unbalanced Panel Data: the Case of Floor Tile Manufactories in Egypt", *Journal of Applied Econometrics*, **5**, 59–79, doi: 10.1002/jae.3950050105 .

Examples

```
#### Example 1-2

## ------------------------------------------------------------------------
data("Tileries", package = "pder")
library("plm")
coef(summary(plm(log(output) ~ log(labor) + machine, data = Tileries,
             subset = area == "fayoum")))

## ------------------------------------------------------------------------
coef(summary(plm(log(output) ~ log(labor) + machine, data = Tileries,
             model = "pooling", subset = area == "fayoum")))


#### Example 1-5

## ------------------------------------------------------------------------
data("Tileries", package = "pder")
til.fm <- log(output) ~ log(labor) + log(machine)
lm.mod <- lm(til.fm, data = Tileries, subset = area == "fayoum")

## ------------------------------------------------------------------------
if (requireNamespace("car")){
    library("car")
    lht(lm.mod, "log(labor) + log(machine) = 1")

## ------------------------------------------------------------------------
    library("car")
    lht(lm.mod, "log(labor) + log(machine) = 1", vcov=vcovHC)
}


#### Example 1-6

## ------------------------------------------------------------------------
plm.mod <- plm(til.fm, data = Tileries, subset = area == "fayoum")

## ------------------------------------------------------------------------
if (requireNamespace("car")){
    library("car")
```

```
    lht(plm.mod, "log(labor) + log(machine) = 1", vcov = vcovHC)
}

#### Example 3-1

## ------------------------------------------------------------------------
library(plm)
data("Tileries", package = "pder")
head(Tileries, 3)
pdim(Tileries)

## ------------------------------------------------------------------------
Tileries <- pdata.frame(Tileries)
plm.within <- plm(log(output) ~ log(labor) + log(machine), Tileries)
y <- log(Tileries$output)
x1 <- log(Tileries$labor)
x2 <- log(Tileries$machine)
lm.within <- lm(I(y - Between(y)) ~ I(x1 - Between(x1)) + I(x2 - Between(x2)) - 1)
lm.lsdv <- lm(log(output) ~ log(labor) + log(machine) + factor(id), Tileries)
coef(lm.lsdv)[2:3]
coef(lm.within)
coef(plm.within)

## ------------------------------------------------------------------------
tile.r <- plm(log(output) ~ log(labor) + log(machine), Tileries, model = "random")
summary(tile.r)

## ------------------------------------------------------------------------
plm.within <- plm(log(output) ~ log(labor) + log(machine),
                  Tileries, effect = "twoways")
lm.lsdv <- lm(log(output) ~ log(labor) + log(machine) +
                  factor(id) + factor(week), Tileries)
y <- log(Tileries$output)
x1 <- log(Tileries$labor)
x2 <- log(Tileries$machine)
y <- y - Between(y, "individual") - Between(y, "time") + mean(y)
x1 <- x1 - Between(x1, "individual") - Between(x1, "time") + mean(x1)
x2 <- x2 - Between(x2, "individual") - Between(x2, "time") + mean(x2)
lm.within <- lm(y ~ x1 + x2 - 1)
coef(plm.within)
coef(lm.within)
coef(lm.lsdv)[2:3]

## ------------------------------------------------------------------------
wh <- plm(log(output) ~ log(labor) + log(machine), Tileries,
          model = "random", random.method = "walhus",
          effect = "twoways")
am <- update(wh, random.method = "amemiya")
sa <- update(wh, random.method = "swar")
ercomp(sa)

## ------------------------------------------------------------------------
re.models <- list(walhus = wh, amemiya = am, swar = sa)
```

```
sapply(re.models, function(x) sqrt(ercomp(x)$sigma2))
sapply(re.models, coef)
```

TobinQ *The Q Theory of Investment*

Description

yearly observations of 188 firms from 1951 to 1985

number of observations : 6580

number of time-series : 35

country : United States

package : panel

Chapter : 02

Usage

```
data(TobinQ)
```

Format

A dataframe containing:

cusip compustat's identifying number

year year

isic sic industry classification

ikb investment divided by capital : broad definition

ikn investment divided by capital : narrow definition

qb Tobin's Q : broad definition

qn Tobin's Q : narrow definition

kstock capital stock

ikicb investment divided by capital with imperfect competition : broad definition

ikicn investment divided by capital with imperfect competition : narrow definition

omphi one minus phi (see the article p. 320)

qicb Tobin's Q with imperfect competition : broad definition

qicn Tobin's Q with imperfect competition : narrow definition

sb S (see equation 10 p. 320) : broad definition

sn S (see equation 10 p. 320) : narrow definition

Source

Journal of Applied Econometrics Data Archive : http://qed.econ.queensu.ca/jae/

References

Schaller, Huntley (1990) "A Re-examination of the Q Theory of Investment Using U.S. Firm Data", *Journal of Applied Econometrics*, **5(4)**, 309–325, doi: 10.1002/jae.3950050402 .

Examples

```
#### Example 2-1

## ------------------------------------------------------------------------
## Not run:
library("plm")
data("TobinQ", package = "pder")

## ------------------------------------------------------------------------
pTobinQ <- pdata.frame(TobinQ)
pTobinQa <- pdata.frame(TobinQ, index = 188)
pTobinQb <- pdata.frame(TobinQ, index = c('cusip'))
pTobinQc <- pdata.frame(TobinQ, index = c('cusip', 'year'))

## ------------------------------------------------------------------------
pdim(pTobinQ)

## ----results = 'hide'----------------------------------------------------
pdim(TobinQ, index = 'cusip')
pdim(TobinQ)

## ------------------------------------------------------------------------
head(index(pTobinQ))

## ------------------------------------------------------------------------
Qeq <- ikn ~ qn
Q.pooling <- plm(Qeq, pTobinQ, model = "pooling")
Q.within <- update(Q.pooling, model = "within")
Q.between <- update(Q.pooling, model = "between")

## ------------------------------------------------------------------------
Q.within
summary(Q.within)

## ------------------------------------------------------------------------
head(fixef(Q.within))
head(fixef(Q.within, type = "dfirst"))
head(fixef(Q.within, type = "dmean"))

## ------------------------------------------------------------------------
head(coef(lm(ikn ~ qn + factor(cusip), pTobinQ)))


#### Example 2-2

## ------------------------------------------------------------------------
Q.swar <- plm(Qeq, pTobinQ, model = "random", random.method = "swar")
```

```
Q.swar2 <- plm(Qeq, pTobinQ, model = "random",
               random.models = c("within", "between"),
               random.dfcor = c(2, 2))
summary(Q.swar)

## ------------------------------------------------------------------------
ercomp(Qeq, pTobinQ)
ercomp(Q.swar)

## ------------------------------------------------------------------------
Q.walhus <- update(Q.swar, random.method = "swar")
Q.amemiya <- update(Q.swar, random.method = "amemiya")
Q.nerlove <- update(Q.swar, random.method = "nerlove")
Q.models <- list(swar = Q.swar, walhus = Q.walhus,
                 amemiya = Q.amemiya, nerlove = Q.nerlove)
sapply(Q.models, function(x) ercomp(x)$theta)
sapply(Q.models, coef)


#### Example 2-3

## ------------------------------------------------------------------------
sapply(list(pooling = Q.pooling, within = Q.within,
            between = Q.between, swar = Q.swar),
       function(x) coef(summary(x))["qn", c("Estimate", "Std. Error")])

## ------------------------------------------------------------------------
summary(pTobinQ$qn)

## ------------------------------------------------------------------------
SxxW <- sum(Within(pTobinQ$qn) ^ 2)
SxxB <- sum((Between(pTobinQ$qn) - mean(pTobinQ$qn)) ^ 2)
SxxTot <- sum( (pTobinQ$qn - mean(pTobinQ$qn)) ^ 2)
pondW <- SxxW / SxxTot
pondW
pondW * coef(Q.within)[["qn"]] +
  (1 - pondW) * coef(Q.between)[["qn"]]

## ------------------------------------------------------------------------
T <- 35
N <- 188
smxt2 <- deviance(Q.between) * T / (N - 2)
sidios2 <- deviance(Q.within) / (N * (T - 1) - 1)
phi <- sqrt(sidios2 / smxt2)

## ------------------------------------------------------------------------
pondW <- SxxW / (SxxW + phi^2 * SxxB)
pondW
pondW * coef(Q.within)[["qn"]] +
  (1 - pondW) * coef(Q.between)[["qn"]]

#### Example 2-8
```

```
## ------------------------------------------------------------------------
Q.models2 <- lapply(Q.models, function(x) update(x, effect = "twoways"))
sapply(Q.models2, function(x) sqrt(ercomp(x)$sigma2))
sapply(Q.models2, function(x) ercomp(x)$theta)

## End(Not run)
```

TradeEU *Trade in the European Union*

Description

yearly observations of 91 pairs of countries from 1960 to 2001

number of observations : 3822

number of time-series : 42

country : Europe

package : gravity

JEL codes: C51, F14

Chapter : 06

Usage

```
data(TradeEU)
```

Format

A dataframe containing:

year the year

pair a pair of countries

trade the sum of logged exports and imports, bilateral trade flow

gdp the sum of the logged real GDPs

sim a measure of similarity between two trading countries;

rlf a measure of relative factor endowments;

rer the logged bilateral real exchange rate;

cee a dummy equal to 1 when both belong to European Community;

emu a dummy equal to 1 when both adopt the common currency;

dist the geographical distance between capital cities;

bor a dummy equal to 1 when the trading partners share a border;

lan a dummy equal to 1 when both speak the same language;

rert the logarithm of real exchange rates between the European currencies and the U.S. dollar;

ftrade the time specific common factors (individual means) of the variables trade

fgdp the time specific common factors (individual means) of the variables gdp

fsim the time specific common factors (individual means) of the variables sim

frlf the time specific common factors (individual means) of the variables rlf

frer the time specific common factors (individual means) of the variables rer

Source

Journal of Applied Econometrics Data Archive : http://qed.econ.queensu.ca/jae/

References

Serlenga, Laura and Yongcheol Shin (2007) "Gravity Models of Intra-eu Trade: Application of the Ccep-ht Estimation in Heterogenous Panels with Unobserved Common Time-specific Factors", *Journal of Applied Econometrics*, **22**, 361–381, doi: 10.1002/jae.944 .

Examples

```
#### Example 6-3

## ------------------------------------------------------------------------
## Not run:
data("TradeEU", package = "pder")
library("plm")

## ------------------------------------------------------------------------
ols <- plm(trade ~ gdp + dist + rer + rlf + sim + cee + emu + bor + lan, TradeEU,
          model = "pooling", index = c("pair", "year"))
fe <- update(ols, model = "within")
fe

## ------------------------------------------------------------------------
re <- update(fe, model = "random")
re

## ------------------------------------------------------------------------
phtest(re, fe)

## ----results='hide'------------------------------------------------------
ht1 <- plm(trade ~ gdp + dist + rer + rlf + sim + cee + emu + bor + lan |
           rer + dist + bor | gdp + rlf + sim + cee + emu + lan ,
           data = TradeEU, model = "random", index = c("pair", "year"),
           inst.method = "baltagi", random.method = "ht")
ht2 <- update(ht1, trade ~ gdp + dist + rer + rlf + sim + cee + emu + bor + lan |
              rer + gdp + rlf + dist + bor| sim + cee + emu + lan)

## ------------------------------------------------------------------------
phtest(ht1, fe)
phtest(ht2, fe)

## ------------------------------------------------------------------------
ht2am <- update(ht2, inst.method = "am")
```

```
## ------------------------------------------------------------------------
phtest(ht2am, fe)

## End(Not run)
```

`TradeFDI` *Trade and Foreign Direct Investment in Germany and the United States*

Description

yearly observations of 490 combinations of countries / industries from 1989 to 1999

number of observations : 3860

number of time-series : 11

country : Germany and United States

package : gravity

JEL codes: F12, F14, F21, F23

Chapter : 06

Usage

```
data(TradeFDI)
```

Format

A dataframe containing:

id id

year time period

country country name

indusid industry code

importid importer code

lrex log real bilateral exports

lrfdi log real bilateral outward stocks of FDI

lgdt log sum of bilateral real GDP

lsimi log (1-[exporter GDP/(exporter+importer GDP)]^2- [exporter GDP/(exporter+importer GDP)]^2)

lrk log (real capital stock of exporter/real capital stock of importer)

lrh log (secondary school enrolment of exporter/secondary school enrolment of importer)

lrl log (labor force of exporter/labor force of importer)

ldist log bilateral distance between exporter and importer

lkldist (lrk-lrl) * ldist

lkgdt abs(lrk)*lgdt

Source

Journal of Applied Econometrics Data Archive : http://qed.econ.queensu.ca/jae/

References

Peter Egger and Michael Pfaffermayr (2004) "Distance, Trade, and Fdi: A Hausman-taylor Sur Approach", *Journal of Applied Econometrics*, **19(2)**, 227–246, doi: 10.1002/jae.721 .

`TurkishBanks` *Turkish Banks*

Description

yearly observations of 53 banks from 1990 to 2000

number of observations : 583

number of time-series : 11

country : Turkey

package : productionpanel

JEL codes: D24, G21, L33

Chapter : 02

Usage

```
data(TurkishBanks)
```

Format

A dataframe containing:

id bank id

year the years

type one of `"conventional"` and `"islamic"`

pl price of labor

pf price of borrowed funds

pk price of physical capital

output output, total loans

cost total cost

empexp employee expenses

nbemp number of employees

faexp assets expenses

fa fixed assets

intexp total interest expenses (interest on deposits and non-deposit funds + other interest expenses),

bfunds borrowed funds (deposits + non-deposit funds)

dep deposits

nondep non-deposits

npl non performing loans

ec equity capital

quality quality index

rindex risk index

ta total assets

ts total securities (only for conventional banks)

Source

Journal of Applied Econometrics Data Archive : http://qed.econ.queensu.ca/jae/

References

Mahmoud A. El-Gamal and Hulusi Inanoglu (2005) "Inefficiency and Heterogeneity in Turkish Banking: 1990-2000", *Journal of Applied Econometrics*, **20(5)**, 641–664, doi: 10.1002/jae.835 .

Examples

```
#### Example 2-5

## ------------------------------------------------------------------------
data("TurkishBanks", package = "pder")
library("plm")
TurkishBanks <- na.omit(TurkishBanks)
TB <- pdata.frame(TurkishBanks)
summary(log(TB$output))
ercomp(log(cost) ~ log(output), TB)
models <- c("within", "random", "pooling", "between")
sapply(models, function(x)
       coef(plm(log(cost) ~ log(output), TB, model = x))["log(output)"])
```

`TwinCrises` *Costs of Currency and Banking Crises*

Description

yearly observations of 22 countries from 1970 to 1997

number of observations : 616

number of time-series : 28

country : world

package : panel

JEL codes: F32, G15, G21, O16, O19, O47

Chapter : 06

Usage

```
data(TwinCrises)
```

Format

A dataframe containing:

country the country name

year the year

gdp real gdp growth

pubsurp change in budget surplus to real gdp ratio

credit credit growth

extgdp external growth rates (weight average)

exr real exchange rate overvaluation

open openess

curcrises currency crises

bkcrises banking crises

twin twin crises

area a factor with levels 'other', 'asia' and 'latam' (for latin America)

Source

Journal of Money, Credit and Banking : https://jmcb.osu.edu/archive

References

Hutchison, Michael M. and Ilan Noy (2005) "How Bad Are Twins ? Output Costs of Currency and Banking Crises", *Journal of Money, Credit and Banking*, **37(4)**, 725–752.

---

usaw *Spatial weights matrix - 49 US states*

Description

Spatial weights matrix of the 48 continental US States plus District of Columbia based on the queen contiguity criterium.

Usage

```
data(usaw49)
data(usaw46)
```

Format

A matrix with elements different from zero if state i and j are neighbors. Weights are row standardized. According to the queen contiguity criterium, Arizona and Colorado are considered neighbours. Two versions are provided, one for 49 States, the other one for 46 States.

Author(s)

Giovanni Millo

Index