

Package: pglm (via r-universe)

September 7, 2024

Version 1.0-0

Date 2023-12-02

Title Panel Generalized Linear Models

Depends R (>= 3.5.0), micr

Imports statmod, Formula, plm, Rdpack

Suggests lmtest, car, knitr

Description Estimation of panel models for glm-like models: this includes binomial models (logit and probit), count models (poisson and negbin) and ordered models (logit and probit), as described in: Baltagi (2013) Econometric Analysis of Panel Data <doi:10.1007/978-3-030-53953-5> Hsiao (2014) Analysis of Panel Data <doi:10.1017/CBO9781139839327> and Croissant and Millo (2018), Panel Data Econometrics with R <doi:10.1002/9781119504641>.

Encoding UTF-8

License GPL (>=2)

URL <https://www.r-project.org/package=pglm>

VignetteBuilder knitr

NeedsCompilation yes

RoxygenNote 7.2.3

Roxygen list(markdown = TRUE)

LazyData true

RdMacros Rdpack

Repository <https://ycroissant.r-universe.dev>

RemoteUrl <https://github.com/ycroissant/pglm>

RemoteRef HEAD

RemoteSha 3638c930392d05dd27aa08c4db6ce08fcd6466dc

Contents

fairness	2
health_ins	3
hedonic	4
patents_rd	5
pglm	5
union_wage	7
Index	9

fairness	<i>Perceived Fairness of rules for allocating seats in trains and parking spaces</i>
----------	--

Description

a pseudo-panel of 401 individuals from 2003

Format

a tibble containing:

- id: the individual index
- answer: a factor with levels 0 (very unfair), 1 (essentially unfair), 2 (essentially fair) and 3 (very fair)
- good: one of 'tgv' (French fast train) and 'Parking'
- rule: the allocation rule, a factor with levels 'peak', 'admin', 'lottery', 'addsupply', 'queuing', 'moral' and 'compensation'
- driving: does the individual has the driving license ?
- education: does the individual has a diploma ?
- recurring: does the allocation problem is recurring ?

Source

provided by the authors

References

Raux C, Souche S, Croissant Y (2009). "How Fair Is Pricing Perceived to Be? An Empirical Study." *Public Choice*, **139**(1/2), 227–240. ISSN 00485829, 15737101, <http://www.jstor.org/stable/40270755>.

health_ins	<i>Health Insurance and Doctor Visits</i>
------------	---

Description

a cross-section of 5908 individuals from 1974 to 1982

Format

a tibble containing:

- id: the individual index
- year: the year
- mdu: number of outpatient visits to an MD
- opu: number of outpatient visits to all providers
- coins: coinsurance rate (0, 25, 50 or 100 percent)
- idp: if individual deductible plan: 1, otherwise 0
- lpi: log of the max of 1 and annual participation incentive payment
- fmde: if idp = 1: 0 otherwise ln of the max of 1 and MDE / (0.01 coins)
- income: family income
- size: family size
- age: the age
- sex: a factor with level 'male' and 'female'
- child: a factor with levels 'no' and 'yes'
- race: a factor with levels 'white' and 'black'
- health: self-rated health, a factor with levels poor, fair, good and verygood
- educ: education of the household head in years
- physlim: if the person has a physical limitation: 1
- disease: index of chronic diseases

Source

<http://cameron.econ.ucdavis.edu/musbook/mus.html>

References

Manning WG, Newhouse JP, Duan N, Keeler EB, Leibowitz A (1987). "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment." *The American Economic Review*, **77**(3), 251–277. ISSN 00028282, <http://www.jstor.org/stable/1804094>.

Deb P, Trivedi PK (2002). "The structure of demand for health care: latent class versus two-part models." *Journal of Health Economics*, **21**(4), 601–625. ISSN 0167-6296, [doi:10.1016/S0167-6296\(02\)000085](https://www.sciencedirect.com/science/article/pii/S0167629602000085), <https://www.sciencedirect.com/science/article/pii/S0167629602000085>.

hedonic

Hedonic Prices of Census Tracts in the Boston Area

Description

a cross-section of 506 census tracts

Format

a tibble containing:

- mv: median value of owner-occupied homes
- crim: crime rate
- zn: proportion of 25,000 square feet residential lots
- indus: proportion of no-retail business acres
- chas: is the tract bounds the Charles River?
- nox: annual average nitrogen oxide concentration in parts per hundred million
- rm: average number of rooms
- age: proportion of owner units built prior to 1940
- dis: weighted distances to five employment centers in the Boston area
- rad: index of accessibility to radial highways
- tax: full value property tax rate (\$/\$10,000)
- ptratio: pupil/teacher ratio
- blacks: proportion of blacks in the population
- lstat: proportion of population that is lower status
- townid: town identifier

Source

Online complements to Baltagi (2013): <https://bcs.wiley.com/he-bcs/Books?action=resource&bcsId=4338&itemId=1118672321&resourceId=13452>

References

- Baltagi BH (2001). *Econometric analysis of panel data*. John Wiley and sons.
- Baltagi BH (2013). *Econometric analysis of panel data*. John Wiley and sons.
- Belsley DA, Kuh E, Welsch RE (1980). *Regression diagnostics: identifying influential data and sources of collinearity*. John Wiley.
- Harrison D, Rubinfeld DL (1978). "Hedonic housing prices and the demand for clean air." *Journal of Environmental Economics and Management*, 5(1), 81-102. ISSN 0095-0696, doi:10.1016/0095-0696(78)900062, <https://www.sciencedirect.com/science/article/pii/0095069678900062>.

patents_rd

Dynamic Relation Between Patents and R&D

Description

yearly observations of 346 production units

Format

a tibble containing:

- cusip: Compustat's identifying number for the firm
- year: year
- ardsic: a two-digit code for the applied R
- scisect: is the firm in the scientific sector ?
- capital72: book value of capital in 1972
- sumpat: the sum of patents applied for between 1972-1979
- rd: R and D spending during the year (in 1972 dollars)
- patents: the number of patents applied for during the year that were eventually granted

Source

<http://cameron.econ.ucdavis.edu/racd/racddata.html>, chapter 9.

References

Hall BH, Griliches Z, Hausman JA (1986). "Patents and R and D: Is There a Lag?" *International Economic Review*, **27**(2), 265–283. ISSN 00206598, 14682354, <http://www.jstor.org/stable/2526504>.

pglm

Panel Estimators for Generalized Linear Models

Description

Estimation by maximum likelihood of glm (binomial and Poisson) and 'glm-like' models (Negbin and ordered) on longitudinal data

Usage

```

pglm(
  formula,
  data,
  subset,
  na.action,
  effect = c("individual", "time", "twoways"),
  model = c("random", "pooling", "within", "between"),
  family,
  other = NULL,
  index = NULL,
  start = NULL,
  R = 20,
  method = c("bfgs", "newton"),
  trace = 0,
  ...
)

ordinal(link = c("probit", "logit"))

negbin(link = c("log"), vlink = c("nb1", "nb2"))

```

Arguments

formula	a symbolic description of the model to be estimated,
data	the data: a <code>pdata.frame</code> object or an ordinary <code>data.frame</code> ,
subset	an optional vector specifying a subset of observations,
na.action	a function which indicates what should happen when the data contains NAs,
effect	the effects introduced in the model, one of "individual", "time" or "twoways",
model	one of "pooling", "within", "between", "random",
family	the distribution to be used,
other	for developer's use only,
index	the index,
start	a vector of starting values,
R	the number of function evaluation for the gaussian quadrature method used,
method	the optimization method, one of "bfgs" (the default) and "newton"
trace	an integer
...	further arguments.
link, vlink	arguments of family functions

Value

An object of class "miscr", a list with elements:

Author(s)

Yves Croissant

Examples

```
## a binomial (probit) example
anb <- pglm(union ~ wage + exper + rural, union_wage, family = binomial('probit'),
           model = "pooling", method = "bfgs", trace = 3, R = 5)

## a gaussian example on unbalanced panel data
ra <- pglm(mv ~ crim + zn + indus + nox + age + rm, hedonic, family = gaussian,
          model = "random", trace = 3, method = "newton", index = "townid")

## some count data models
la <- pglm(patents ~ lag(log(rd), 0:5) + scisect + log(capital72) + factor(year), patents_rd,
          family = negbin, model = "within", trace = 3, method = "newton",
          index = c('cusip', 'year'))
la <- pglm(patents ~ lag(log(rd), 0:5) + scisect + log(capital72) + factor(year), patents_rd,
          family = poisson, model = "pooling", index = c("cusip", "year"),
          tracen = 0, method="newton")
```

union_wage

Unionism and wage rate determination

Description

yearly observations of 545 individuals from 1980 to 1987

Format

a tibble containing:

- id: the individual index
- year: the year
- exper: the experience, computed as age - 6 - schooling
- health: does the individual has health disability ?
- hours: the number of hours worked
- married: is the individual married ?
- rural: does the individual lives in a rural area ?
- school: years of schooling
- union: does the wage is set by collective bargaining
- wage: hourly wage in US dollars
- sector: one of agricultural, mining, construction, trade, transportation, finance, businessrepair, personalservice, entertainment, manufacturing, pro.rel.service, pub.admin

- occ: one of proftech, manoffpro, sales, clerical, craftfor, operative, laborfarm, farmlabor, service
- com: one of black, hisp and other
- region: the region, one of NorthEast, NothernCentral, South and other

Source

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

References

Vella F, Verbeek M (1998). "Whose Wages do Unions Raise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men." *Journal of Applied Econometrics*, **13**(2), 163–183. ISSN 08837252, 10991255, <http://www.jstor.org/stable/223257>.

Index

* dataset

- fairness, 2
- health_ins, 3
- hedonic, 4
- patents_rd, 5
- union_wage, 7

fairness, 2

health_ins, 3

hedonic, 4

negbin (pglm), 5

ordinal (pglm), 5

patents_rd, 5

pglm, 5

union_wage, 7