plm : linear models for panel data

Yves Croissant
Laboratoire d'Economie des Transports
Institut des Sciences de l'Homme
14, avenue Berthelot
F-69363 Lyon cedex 07
yves.croissant@let.ish-lyon.cnrs.fr
33 4.72.72.64.49

`plm` is a package that implements the main estimators and tests used in econometrics for panel data.

Panel datas have an individual and a temporal dimension. `plm` provides specific functions for reading and applying special transformations to panel data sets, and for estimating and testing linear models.

reading data `pdata.frame` takes as main argument a `data.frame`. It returns a `data.frame` with further arguments usefull for panel datas, such as the number of individuals and time observations.

special functions this includes `plag` and `pdiff`, which computes lags and differences of series, `pmean` which computes the mean of a serie, conditionnal on the individual or the time index,

estimation `plm` is a general function which implements the main panel data estimators. The basic usage of `plm` consist of estimating four models :

pooling the ordinary least squares estimator applied to raw observations,

within the ols estimator applied to observations measured as deviations from individual (or time) means,

between the ols estimator applied on individual (or time) means,

random the random effect, a generalized least squares estimator which is a wheighted average of the `within` and the `between` estimator.

`plm` returns by default an object of class `plms`, which is a list of the four models previously described, which are objects of class `plm`. `plms` and `plm` objects have `print` and `summary` methods. These estimators deals with oneway (individual or time) effects or twoways and with unbalanced panel. Different instrumental variable estimators are also available (for example the Hausman and `Taylor` estimator)

tests different tests of model specification are provided :

pFtest a simple test for the presence of individual (or/and time) effects based on the comparison of the `pooling` and the `within` models,

plmtest a set of likelihood ratio tests for the presence of individual (or/and time) effects based on the comparison of the `random` and the `pooling` model,

phausman a Hausman test for the correlation between explanatory variables and individual (or/and time) effects, based on the comparison of the `within` and the `random` models.

Further developments planed for `plm` include :

system estimation seemingly unrelated regression and three stage least squares estimators, using the `systemfit` package,

robust covariance matrix using the `sandwich` package,

autoregressive models Arrelano and Bound general method of moments estimator.

All the functions of `plm` have been tested using the examples provided in the book of B. Baltagi "Econometric analysis of panel data". The data sets used in this book are provided in packages `Ecdat` and `plm`.