Package: rewie (via r-universe)

September 10, 2024

Type Package

Title Data Preparation and Diagnostics for Random Effects Within Estimator

Version 0.1.0

Maintainer Scott Duxbury <sduxburycpt@gmail.com>

URL https://github.com/sduxbury/rewie

BugReports https://github.com/sduxbury/rewie/issues

Description Diagnostics and data preparation for random effects within estimator, random effects within-idiosyncratic estimator, between-within-idiosyncratic model, and cross-classified between model. Mundlak, Yair (1978) <doi:10.2307/1913646>. Hausman, Jeffrey (1978) <doi:10.2307/1913827>. Allison, Paul (2009) <doi:10.4135/9781412993869>. Neuhaus, J.M., and J. D. Kalbfleisch (1998) <doi:10.2307/3109770>.

License GPL (>=2)

Encoding UTF-8

LazyData true

Imports rockchalk, lme4, MASS, plm, stats

Repository https://sduxbury.r-universe.dev

RemoteUrl https://github.com/sduxbury/rewie

RemoteRef HEAD

RemoteSha f5ec7ca384f1c433d0aac6e1314ece3ff03fc027

Contents

hausman	*Conducts auxiliary Hausman tests*

Description

Tests correlations between level 1 and level 2 errors in the pooled random effects estimator and REWE. If the pooled test is significant, the pooled estimator is biased and the within estimator should be considered. If the within test is significant, then the within estimator is biased and REWIE, BWI, or two-way fixed effects should be consider.

Usage

```
hausman(formula,timevar,csvar,df)
```

Arguments

formula	is a character string or formula object denoting the models to be tested, i.e., "y ~ x1+x2". NOTE:do not provide a lmer formula object (ie."y~x+(1|x").
timevar	is a character string providing the name of the time indicator variable.
csvar	is a character string providing the name of the cross-section indicator variable.
df	is the dataframe containing the data. NOTE: do not provide preprocessed variables to hausman(). All variable transformations used to conduct the tests are done natively within the function.

Details

Conducts auxiliary Hausman test of homogeneity in panel data. If the pooled test is failed, it means that the between variation adn within variation are correlated and so the pooled estimator is biased. The within test tests for correlation between the idiosyncratic and homogenous-within variation. If the test is filed, time heterogeneity exists and the one-way within estimator is biased. If both tests are failed, BWI, REWIE, and two-way fixed effects are unbiased estimators. If only the pooled test is failed, then the one-way within estimator (one-way fixed effects and between within) are unbiased. If neither test is failed, the pooled model is appropriate.

Value

Returns the results from the two Hausman tests (called from the plm package using phtest) in a list.

pooled_vs_FE	is the result from the pooled Hausman test
REWE_vs_2FE	is the result from the within Hausman test

Author(s)

Scott Duxbury, Assistant Professor of Sociology at University of North Carolina, Chapel Hill

See Also

phtest

Examples

```
library(plm)
library(lme4)


##not run
data("Crime")


hausman("lcrmrte~ldensity+polpc",csvar="county", timevar="year",Crime)
```

ICC   *Computes intraclass correlation coefficient (ICC)*

Description

Calculates the ICC for the idiosyncratic, within, and between variation. Calculates the ICC for total variation and for within variation.

Usage

```
ICC(y,timevar,csvar,df)
```

Arguments

y	is the name of the response vector provided as a character string.
timevar	is a character string providing the name of the time indicator variable.
csvar	is a character string providing the name of the cross-section indicator variable.
df	is a dataframe containing y, timevar, and csvar

Details

Calculates ICC foreach level of variation.

Value

Returns the ICC results in a list.

ICC.mat	is the matrix of ICCs and variance components for each level of variation.
within.ICC.mat	is the matrix of ICCs and variance components only examining the within variation.

Author(s)

Scott Duxbury, Assistant Professor of Sociology at University of North Carolina, Chapel Hill

Examples

```
library(plm)
library(lme4)


##not run
data("Crime")


ICC("lcrmrte",csvar="county", timevar="year",Crime)
```

---

rewie.dat   *Function to transform data for mixed modeling with REWE, REWIE,*
                             BW, and BWI.

Description

Centers data for the specified variables. For REWE and BW, data are centered to provide cross-section means and within variation. For REWIE and BWI, data are centered to provide cross-section means, time means for the within variation, and the idiosyncratic variation centered on both cross-section and time means.

Usage

```
rewie.dat(df,vars.to.center,csvar,timevar=NULL,model="BWI")
```

Arguments

df   df is the data frame containing the variables to transform.

vars.to.center   The variables to be transformed.

timevar   is a character string providing the name of the time indicator variable.

csvar   is a character string providing the name of the cross-section indicator variable.

model   specifies the type of transformation. Can be one of the following: "BWI","REWE","REWIE",
                  or"BW". For REWE and BW, the within variation and cross-section means are
                  returned. For REWIE and BWI, the cross-section means, idiosyncratic varia-
                  tion, and time means of the within variation are returned.

Details

Transforms panel data to conduct panel analysis with random effects within estimators and its extensions. The variables are centered on their time and cross-section means to be passed to lmer() for mixed modeling. The time means are the means of the differences between the cross-section means and the observations. The idiosyncratic variation is the observation twice centered (minus the cross-section mean, then minus the time mean). When passed to a mixed model, the idiosyncratic variation returns the same estimate as two-way fixed effects, the between variation gives the between estimator (_between), and the time mean (_common) gives the period effect of the variable. Other time invariant and cross-section invariant variables can also be included, but these models do not need to be transformed.

Value

Returns a dataframe including the transformed values to be used in mixed modeling.

Author(s)

Scott Duxbury, Assistant Professor of Sociology at University of North Carolina, Chapel Hill

See Also

edgeprob

Examples

```
library(plm)
library(lme4)


##not run
data("Crime")


#####Ceate data for REWE and BW
crime.data<-rewie.dat(Crime,c("lcrmrte","ldensity"),csvar="county",model="BW")

#random effects within estimator. Equivalent to one-way fixed effects with a random intercept
REWE<-lmer(lcrmrte_within~ldensity_within+(1|year),data=crime.data)

#Between within model w/o time intercept
BW<-lmer(lcrmrte~ldensity_within+ldensity_between+(1|county),data=crime.data)

#cross-classified between within model including time intercept
ccBW<-lmer(lcrmrte~ldensity_within+ldensity_between+(1|county)+(1|year),data=crime.data)



###Create data for REWIE and BWI
crime.data<-rewie.dat(Crime,c("lcrmrte","ldensity"),csvar="county",timevar="year",model="BWI")
```

```
#Random effects within-idiosyncratic estimator. Idiosyncratic estimator is the
  #two-way fixed effects estimator.
  #_common variables yield the effect of common time trends
    #(period effects/homogenous within effect) on the outcome
REWIE<-lmer(lcrmrte_within~ldensity_within_idiosyncratic+
            ldensity_within_common+
            (1|year),data=crime.data)

#Between-within-idiosyncratic model. Idiosyncratic estimator and common estimator are
  #the same as REWIE, but also includes between effects
BWI<-lmer(lcrmrte~ldensity_within_idiosyncratic+
          ldensity_within_common+
          ldensity_between+
          (1|year)+(1|county),data=crime.data)
```

---

rewie.rsq   *Computes R-squared for RE panel models*

Description

Calculates R-squared for BW, REWE, REWIE, and BW. Includes idiosyncratic R-squared, between R-squared, homogneous-within R-squared, and within R-Squared.

Usage

```
rewie.rsq(model,timevar,csvar,df)
```

Arguments

model	is an lmerMod object fitted by calling the lmer() function in lme4.
timevar	is a character string providing the name of the time indicator variable.
csvar	is a character string providing the name of the cross-section indicator variable.
df	is a dataframe containing y, timevar, and csvar

Details

Calculates the R-squared for each level of variation. It is calculated by computing the proportion of remaining variance to overall variance the model and then subtracting the quotient from 1.

Value

Returns the results for R-squares

Rsq.total	is the overall R-squared.
Rsq.within	is the within R-squared.
Rsq.time	is the time (homogenous-within) R-squared.
Rsq.idio	is the idiosyncratic R-squared.
Rsq.betw	is the between R-squared.

Author(s)

Scott Duxbury, Assistant Professor of Sociology at University of North Carolina, Chapel Hill

Examples

```
require(plm)
require(lme4)

data("Crime")
output<-lmer(lcrmrte~ldensity+(1|county)+(1|year),data=Crime)
rewie.rsq(output,csvar="county",timevar="year",df=Crime)
```

Index