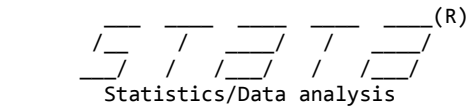

help fairlie

Title

fairlie — Nonlinear decomposition of binary outcome

Syntax

fairlie *depvar indepvars* [*if*] [*in*] [*weight*], **by(***groupvar***)** [*options*]

where the syntax for *indepvars* is

term [*term* ...]

with *term* as a variable name or, alternatively, a set of variables specified as

([*name*:] *varlist***)**

name is any valid Stata name and labels the set. If *name* is omitted, the name of the first variable is used to l

options	Description
by(*groupvar***)**	specify the groups (required); *groupvar* must be 0/1
reps(*#***)**	number of decomposition replications; default is 100
nodots	suppress the replication dots
ro	randomize ordering of variables in the detailed decomposition
reference(*#***)**	specify the reference model; # must be 0 (use group 0 model) or 1 (use group 1 model); defau
pooled[**(***varlist***)**]	use a pooled model as reference; *varlist* is added to the model if specified
probit	use a **probit** model; default is to use a **logit** model
noest	suppress model estimation output
saveest(*name***)**	store model estimation results under *name*
level(*#***)**	set confidence level; default is **level(95)**
nolegend	suppress legend
estopts	options passed through to the internal call of **logit** or **probit**

fweights, **pweight**s, and **iweight** are allowed; see weight.

Description

fairlie computes the nonlinear decomposition of binary outcome differentials proposed by Fairlie (1999, 2003, 20
the two groups defined by *groupvar* and quantifies the contribution of group differences in the *indepvars* to the
contributions of the individual independent variables (or groups of independent variables). **fairlie** also report
covariances are set to zero. Therefore, do *not* use post-estimation commands such as **test** or **lincom** after **fairlie**

The implementation of the decomposition technique closely follows the suggestions provided by Fairlie (2003). If
computation of the detailed decomposition (see below).

The decomposition technique involves one-to-one matching of cases between the two groups. If the groups have dif
results depend on the specific sample, the process is repeated and mean results are reported. Use **reps()** to spec
replicable results; see help **set seed**.

The separate contributions from independent variables or groups of independent variables may be sensitive to the
the **ro** option described below to randomize the ordering of variables, thus approximating results over all possib

Alternative decomposition approaches for binary response variables are provided, e.g., by Gomulka and Stern (199

Algorithm for weighted data: The algorithm by Fairlie for the detailed decomposition is based on matching observ
random sample (without replacement) from the larger group. The goal of the matching is to generate a hypothetica
the first group and some from the second group. In the case of weighted data, the original algorithm cannot be u
variables in the hypothetical sample. However, an appropriate hypothetical sample can be constructed by matching
proportional to the weights. In the present implementation the sizes of the two sub-samples are set to half the
replacement. The choice of the sub-sample size is arbitrary but that does not matter much since the precision o
observations, which is a function of the sub-sample size and the number of decomposition replications as set by
counterbalanced by an increase (a decrease) in the number of replications. The results from the original algorit
but they have the same expectation if the weights are uninformative (i.e. if the weights are equal for all obser

Options

by(*groupvar***)** defines the groups between which the decomposition is to be performed. *groupvar* must be 0/1.

reps(#) specifies the number of decomposition replications to be performed. The default is 100.

nodots suppresses the display of replication dots.

ro causes the ordering of variables to be randomized in the detailed decomposition. The default is to estimate t
groups of independent variables) one after another in the specified order. Note that results are sensitive t
the variables in each replication and, therefore, approximate average results over all possible orderings. T
variables. It may also be useful to increase the number of replications when using this option.

reference(#) specifies the reference estimates to be used with the decomposition. **reference(0)**, the default, ind
reference(1) specifies that the coefficients from the *groupvar*==1 model are used.

pooled[**(***varlist***)**] specifies that the coefficients from the pooled model over all cases be used for the decomposi
restricted to the non-missing cases of *groupvar*. This is reasonable because it is sometimes desirable to in
sample of the pooled model.) Optionally, *varlist* will be added as additional control variables to the pooled
important that the reference group in the pooled model is the group for which *groupvar*==0.

probit specifies that the **probit** command is used for model estimation. The default is to use **logit**.

noest suppresses the display of the model estimates.

saveest(*name***)** stores the model estimation results under *name* using **estimates store**.

level(#); see estimation options.

nolegend suppresses the legend for the variable sets.

estopts are options passed through to the internal call of **logit** or **probit**.

Examples

```
. use http://fmwww.bc.edu/RePEc/bocode/h/homecomp.dta

. fairlie homecomp female age (educ:hsgrad somecol college)
    (marstat:married prevmar) if white==1|black==1, by(black)

. fairlie homecomp female age (educ:hsgrad somecol college)
    (marstat:married prevmar) if white==1|black==1, by(black)
    pooled(black)

. generate black2 = black==1 if white==1|black==1

. fairlie homecomp female age (educ:hsgrad somecol college)
    (marstat:married prevmar), by(black2)
    pooled(black latino asian natamer)

. fairlie homecomp female age (educ:hsgrad somecol college)
    (marstat:married prevmar) if white==1|black==1 [pw=wgt],
    by(black)
```

Saved Results

Scalars

e(N)	number of observations
e(N_0)	number of obs for which *groupvar*==0
e(N_1)	number of obs for which *groupvar*==1
e(N_match)	sample size used for one-to-one matching
e(reps)	number of decomposition replications
e(pr_0)	outcome probability for *groupvar*==0
e(pr_1)	outcome probability for *groupvar*==1
e(diff)	differential **e(pr_0)-e(pr_1)**
e(expl)	total contribution of group differences in regressors

Macros

e(cmd)	**fairlie**
e(depvar)	name of dependent variable
e(by)	name group variable
e(_cmd)	command used for model estimation (**logit** or **probit**)
e(wtype)	weight type
e(wexp)	weight expression
e(reference)	reference estimates (**0**, **1**, or **pooled**)
e(legend)	definitions of regressor sets
e(ro)	**ro**, if the random order option was specified
e(properties)	**b V**

Matrices

e(b)	detailed decomposition results
e(V)	variances for **e(b)** (covariances are set to zero)
e(_b)	reference coefficients
e(_V)	variance-covariance matrix of **e(_b)**

Functions

e(sample)	marks estimation sample

References

Fairlie, Robert W. (1999). The Absence of the African-American Owned Business: An Analysis of the Dynamics of Se
DOI:10.1086/209914
Fairlie, Robert W. (2003). An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models
http://ideas.repec.org/p/egc/wpaper/873.html
Fairlie, Robert W. (2005). An extension of the Blinder-Oaxaca decomposition technique to logit and probit models
DOI:10.3233/JEM-2005-0259
Gomulka, Joanna, and Nicholas Stern (1990). The Employment of Married Women in the United Kingdom 1970-83. Econo
Yun, Myeong-Su (2003). Decomposing Differences in the First Moment. IZA Discussion Paper No. 877. http://ideas.

Author

Ben Jann, University of Bern, ben.jann@unibe.ch

Thanks for citing this software as follows:

> Jann, B. (2006). fairlie: Stata module to generate nonlinear decomposition of binary outcome differentials.

Acknowledgments

I thank Sonia Bhalotra, Rob Fairlie, Julia Horstschräer, George Leckie, and Steven Samuels for their comments ar

Also see

Online: help for **logit**, **probit**, **set seed**