Nonlinear Decomposition Code

I created a decomposition technique for logits, probits and other nonlinear models based on the linear Blinder-Oaxaca decomposition technique. The decomposition technique is useful for explaining the causes of racial and gender inequality in outcomes.

Original Application: Fairlie, Robert W. 1999. "The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment," *Journal of Labor Economics*, 17(1): 80-108. Journal of Labor Economics: The Absence of the African-American Owned Business PDF

Revised to randomly match black/white distributions in: Fairlie, Robert W., and Alicia M. Robb. 2007. "Why are Black-Owned Businesses Less Successful than White-Owned Businesses: The Role of Families, Inheritances, and Business Human Capital," *Journal of Labor Economics*, 25(2): 289-323. Journal of Labor Economics: Why Are Black-Owned Businesses Less Successful than White-Owned Businesses PDF

Revised to randomize variable ordering and incorporate sample weights if needed as discussed in:

Fairlie, Robert W. 2017. "Addressing Path Dependence and Incorporating Sample Weights in the Nonlinear Blinder-Oaxaca Decomposition Technique for Logit, Probit and Other Nonlinear Models," Stanford University (SIEPR) WP

Example Decomposition Programs for SAS

New Versions that Address Path Dependence and Incorporate Sample Weights
decompexample_v7.sas – Original Method of Specifying Order of Variables
decompexamplerandom_v7.sas – Randomized Ordering of Variables to Address Path Dependence

Earlier Versions

decompexample_v6.sas – Full Version with Standard Errors
decompexamplenose_v6.sas – Simplified Version without Standard Errors

Dataset for Example Programs – SAS
Dataset for Example Programs – Stata
Dataset for Example Programs – CSV

Stata Program for Logit or Probit
code written by Ben Jann, ETH Zurich (Swiss Federal Institute of Technology)

In Stata, the program can be installed by typing the following in the command line:
```
ssc install fairlie
```
If the program already exists and you want to update it then type:
```
ssc install fairlie, replace
```
For help and examples on how to use the program type:
```
ssc help fairlie
```

Examples for Using Stata Procedure

1. White-Black Decomposition using Coefficients from Pooled Sample of All Races

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

```
fairlie homecomp female age college (region:midwest south west), by(black2) pooled (black latino asian natamer)
```

Notes: (1) A pooled regression including all racial groups is used to estimate the parameters (which reflects the full market instead of the parameters for only a specific racial group). The full set of race dummies needs to be listed in the command. (2) The black2 dummy is created to define the two comparison groups (black2=0 for whites and black2=1 for blacks). (3) The independent contributions from each region dummy cannot be estimated and thus must be estimated as a group (which is defined in the code).

2. White-Black Decomposition using Coefficients from White Sample

```
fairlie homecomp female age college (region:midwest south west) if white==1|black==1, by(black)
```

Notes: (1) Only white observations (i.e. black=0) are used to estimate the parameters. (2) The black dummy and selecting the sample to only include whites and blacks defines the two comparison groups (black=0 for whites, and black=1 for blacks). (3) The independent contributions from each region dummy cannot be estimated and thus must be estimated as a group (which is defined in the code above).

3. Male-Female Decomposition using Coefficients from Pooled Sample of Men and Women

```
fairlie homecomp black latino asian natamer age college (region:midwest south west), by(female) pooled (female)
```

Notes: (1) A pooled regression including both men and women is used to estimate the parameters (which reflects the full market instead of the parameters for only one gender). The female dummy needs to be listed in the command. (2) The female dummy defines the two comparison groups (female=0 for men and female=1 for women). (3) The independent contributions from each region dummy cannot be estimated and thus must be estimated as a group (which is defined in the code).

4. Male-Female Decomposition using Coefficients from Pooled Sample of Men and Women with Random Ordering of Variables and More Replications

```
fairlie homecomp black latino Asian natamer age college (region:midwest south west), by(female) pooled (female) ro
reps(1000)
```

Notes: (1) A pooled regression including both men and women is used to estimate the parameters (which reflects the full market instead of the parameters for only one gender). The female dummy needs to be listed in the command. (2) The female dummy defines the two comparison groups (female=0 for men and female=1 for women). (3) The independent contributions from each region dummy cannot be estimated and thus must be estimated as a group (which is defined in the code). (4) The variables are ordered randomly in each replication so that contribution estimates are not sensitive to ordering of variables in regression statement. (5) The number of replications is 1000 instead of the default number of replications of 100.