

# Package: fairlieR (via r-universe)

July 23, 2026

**Title** Fairlie Decomposition for Nonlinear Models (Logit Models)

**Version** 0.1.1

**Description** Performs the Fairlie decomposition for nonlinear models (logit models). This function breaks down the difference in average outcomes between two groups into the part explained by group differences in observable characteristics and the unexplained part (Fairlie 1999 <[doi:10.1086/209914](https://doi.org/10.1086/209914)>; Fairlie 2005 <[doi:10.3233/JEM-2005-0259](https://doi.org/10.3233/JEM-2005-0259)>).

**License** MIT + file LICENSE

**Encoding** UTF-8

**RoxygenNote** 7.3.2

**Depends** R (>= 3.5)

**LazyData** true

**Suggests** ggplot2

**URL** <https://github.com/CathyBu9/fairlieR>

**BugReports** <https://github.com/CathyBu9/fairlieR/issues>

**NeedsCompilation** no

**Author** Xinyan Bu [aut, cre]

**Maintainer** Xinyan Bu <[xinyan.bu@outlook.com](mailto:xinyan.bu@outlook.com)>

**Repository** <https://cran.r-universe.dev>

**Date/Publication** 2026-07-23 13:10:02 UTC

**RemoteUrl** <https://github.com/cran/fairlieR>

**RemoteRef** HEAD

**RemoteSha** c8dd3323c2ff25dc33f1ce1df7368ce0d3790ed7

## Contents

|                    |          |
|--------------------|----------|
| fairlieR . . . . . | 2        |
| homecomp . . . . . | 4        |
| <b>Index</b>       | <b>6</b> |

**Description**

Performs the Fairlie decomposition for nonlinear models (logit models). This function breaks down the difference in average outcomes between two groups into the part explained by group differences in observable characteristics and the unexplained part (Fairlie 1999 [doi:10.1086/209914](#); Fairlie 2005 [doi:10.3233/JEM-2005-0259](#)). Supports multiple reference coefficient weighting schemes.

**Usage**

```
fairlieR(
  formula,
  data,
  group,
  base_group = 0,
  reference = 0,
  reps = 100,
  ro = FALSE,
  seed = NULL,
  plot = FALSE,
  plot_color = c("#1F77B4", "#BDC3C7"),
  plot_show_title = TRUE,
  plot_show_subtitle = TRUE,
  verbose = TRUE
)
```

**Arguments**

|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| formula    | An object of class "formula" (e.g., $Y \sim X1 + X2$ ).                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| data       | A data frame containing the variables in the model.                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| group      | A character string specifying the grouping variable.                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| base_group | The value of the grouping variable used as the baseline for the difference. Default is 0.                                                                                                                                                                                                                                                                                                                                                                                                       |
| reference  | Determines the coefficients used for decomposition. Options: <ul style="list-style-type: none"> <li>• 0 the value of base_group (<math>Y=0</math>),</li> <li>• 1 the value of the <math>Y=1</math> group,</li> <li>• 0.5 (Reimers 1983: equally weighted),</li> <li>• "cotton" (Cotton 1988: sample size weighted),</li> <li>• -1 (Neumark 1988: pooled regression excluding group indicator),</li> <li>• -2 (Jann 2008: pooled regression including group indicator). Default is 0.</li> </ul> |
| reps       | An integer specifying the number of random permutations. Default is 100.                                                                                                                                                                                                                                                                                                                                                                                                                        |

|                    |                                                                                                                                          |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| ro                 | A logical value. If TRUE, randomizes the ordering of variables. Default is FALSE.                                                        |
| seed               | An optional integer to set the random seed for reproducibility.                                                                          |
| plot               | A logical value. If TRUE, generates an academic forest plot. Default is FALSE.                                                           |
| plot_color         | A character vector of length 2 specifying fill colors for significant and non-significant variables. Default is c("#1F77B4", "#BDC3C7"). |
| plot_show_title    | A logical value. If TRUE, displays the main title on the plot. Default is TRUE.                                                          |
| plot_show_subtitle | A logical value. If TRUE, displays the subtitle (reference method) on the plot. Default is TRUE.                                         |
| verbose            | A logical value. If TRUE, prints the decomposition summary to the console. Default is TRUE.                                              |

### Value

A data frame containing the decomposition results.

### Author(s)

Xinyan Bu

### References

- 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. DOI:10.1086/209914.
- Fairlie, Robert W. (2003). An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models. *Economic Growth Center, Yale University Discussion Paper No. 873*. <https://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. *Journal of Economic and Social Measurement* 30: 305-316. DOI:10.3233/JEM-2005-0259.
- OAXACA R. Male-Female Wage Differentials in Urban Labor Markets (J/OL). *International Economic Review*, 1973, 14(3): 693. DOI:10.2307/2525981.
- BLINDER A S. Wage Discrimination: Reduced Form and Structural Estimates on JSTOR (EB/OL). (1973) (Accessed: 2025-11-04). <https://www.jstor.org/stable/144855>.
- REIMERS CW. EconPapers: Labor Market Discrimination against Hispanic and Black Men (EB/OL). (1983) (Accessed: 2025-11-05). [https://econpapers.repec.org/article/tpreestat/v\\_3a65\\_3ay\\_3a1983\\_3ai\\_3a479.htm](https://econpapers.repec.org/article/tpreestat/v_3a65_3ay_3a1983_3ai_3a479.htm).
- COTTON J. On the Decomposition of Wage Differentials (J). *The Review of Economics and Statistics*, 1988, 70(2): 236-243.
- NEUMARK. EconPapers: Employers' Discriminatory Behavior and the Estimation of Wage Discrimination (EB/OL). (1988) (Accessed: 2025-11-05). [https://econpapers.repec.org/article/uwpjhriss/v\\_3a23\\_3ay\\_3a295.htm](https://econpapers.repec.org/article/uwpjhriss/v_3a23_3ay_3a295.htm).

- Hlavac, Marek (2022). *oaxaca*: Blinder-Oaxaca Decomposition in R.
- Jann, B. (2006). *fairlie*: Stata module to generate nonlinear decomposition of binary outcome differentials.

## Examples

```
# Load the embedded Stata example dataset
data("homecomp")

# Filter the data exactly as in the Stata example: if white==1|black==1
df_subset <- subset(homecomp, white == 1 | black == 1)

# Align factor levels and establish baselines
df_subset$educ_factor <- as.factor(
  ifelse(df_subset$college == 1, "College",
  ifelse(df_subset$somecol == 1, "SomeCol",
  ifelse(df_subset$hsgrad == 1, "HSGrad", "DropOut")))
)
df_subset$educ_factor <- relevel(df_subset$educ_factor, ref = "DropOut")

df_subset$marstat_factor <- as.factor(
  ifelse(df_subset$married == 1, "Married",
  ifelse(df_subset$prevmar == 1, "PrevMar", "Single"))
)
df_subset$marstat_factor <- relevel(df_subset$marstat_factor, ref = "Single")

# Run the native R Fairlie decomposition with visualization enabled
result <- fairlieR(homecomp ~ female + age + educ_factor + marstat_factor,
  data = df_subset,
  group = "black",
  base_group = 0,
  reference = -2,
  reps = 100,
  plot = TRUE)
```

---

homecomp

*Home Ownership Data from Stata*

---

## Description

An example dataset from Stata's *fairlie* module used to demonstrate the decomposition.

## Usage

homecomp

## Format

A data frame containing variables for home ownership, gender, age, education, and marital status.

**Source**

<http://fmwww.bc.edu/RePEc/bocode/h/homecomp.dta>

# Index

## \* **datasets**

homecomp, [4](#)

fairlieR, [2](#)

homecomp, [4](#)