

Package ‘fdid’

May 8, 2026

Type Package

Title Factorial Difference-in-Differences

Version 1.0.2

Date 2026-03-18

Description Implements the factorial difference-in-differences (FDID) framework for panel data settings where all units are exposed to a universal event but vary in a baseline factor G . Provides support for various estimators; supports robust, bootstrap, and jackknife variance; returns dynamic, pre/event/post aggregates and raw means; and includes helpers for data preparation and plotting. Methodology follows Xu, Zhao and Ding (2026) <[doi:10.1080/01621459.2026.2628343](https://doi.org/10.1080/01621459.2026.2628343)>.

URL <https://yiqingxu.org/packages/fdid/>,
<https://github.com/xuyiqing/fdid>

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Imports estimatr, dplyr, tidyr, rlang, tidyselect, RColorBrewer,
foreach, doFuture, future, ebal, grf, car, sandwich

Suggests testthat (>= 3.0.0), knitr, rmarkdown

VignetteBuilder knitr

Depends R (>= 3.6.0)

RoxygenNote 7.3.3

NeedsCompilation no

Author Yiqing Xu [aut, cre],
Rivka Lipkowitz [aut],
Enhan Liu [aut]

Maintainer Yiqing Xu <yiqingxu@stanford.edu>

Repository CRAN

Date/Publication 2026-03-23 09:20:07 UTC

Contents

fdid	2
fdid_list	4
fdid_prepare	5
mortality	6
plot.fdid	7
plot.fdid_list	8
print.fdid	9
summary.fdid	10

Index	11
--------------	-----------

fdid	<i>Factorial Difference-in-Differences Estimation</i>
------	-------------------------------------------------------

Description

Performs factorial difference-in-differences (FDID) estimation using various methods and variance estimation techniques.

Usage

```
fdid(
  s,
  tr_period,
  ref_period,
  entire_period = NULL,
  method = "ols1",
  vartype = "robust",
  missing_data = c("listwise", "available"),
  nsims = 1000,
  parallel = FALSE,
  cores = 2,
  target.pop = c("all", "1", "0")
)
```

Arguments

s	A data frame prepared using <code>fdid_prepare</code> .
tr_period	A numeric vector specifying the treatment periods.
ref_period	A numeric scalar specifying the reference period.
entire_period	A numeric vector specifying the total range of time periods. If <code>NULL</code> , estimation is performed on all available time periods. Example: <code>c(1958, 1959, 1960, 1961)</code> .
method	A string specifying the estimation method. Options: <code>"ols1"</code> , <code>"ols2"</code> , <code>"did"</code> , <code>"ebal"</code> , <code>"ipw"</code> , <code>"aipw"</code> . Default is <code>"ols1"</code> .

vartype	A string specifying the variance estimation type. Options: "robust", "bootstrap", "jackknife". Default is "robust".
missing_data	How to handle missing data. Two options: <ul style="list-style-type: none"> "listwise": Drop any row missing any relevant column (including outcomes in the periods used). "available": Drop rows only if they are missing in group/covariates/cluster columns, but allow partial usage of outcomes. Default is "listwise".
nsims	Number of simulations for bootstrap variance estimation. Default is 1000.
parallel	Logical; whether to perform parallel computations. Default is FALSE.
cores	Number of cores for parallel computations. Default is 2.
target.pop	Character; the target population for averaging: "all", "1", or "0". "all" corresponds to the full sample. "1" targets the G=1 population. "0" targets the G=0 population. Default is "all".

Value

A list with the following components:

est	A list with three elements: \$pre, \$event, and \$post containing aggregated pre-treatment, overall event, and post-treatment FDID estimates, respectively.
dynamic	Dynamic FDID estimates for each time in entire_period.
raw_means	Raw mean outcomes by group for each time in entire_period.
tr_period	Treatment periods used.
ref_period	Reference period used.
entire_period	All time periods for dynamic estimation.
method	Method used.
vartype	Variance type used.
times	All numeric time columns found.
G	Group indicator (0/1).
ps	Propensity scores (if ipw or aipw method used).
call	The matched call.
target.pop	Character indicating the target population used.

Author(s)

Rivka Lipkowitz, Enhao Liu

Examples

```
data(fdid)
mortality$uniqueid <- paste(mortality$provid, mortality$countyid, sep = "-")
mortality$G <- ifelse(mortality$pczupu >= median(mortality$pczupu, na.rm = TRUE), 1, 0)
s <- fdid_prepare(
  data = mortality, Y_label = "mortality",
  X_labels = c("avggrain", "lnpop"),
  G_label = "G", unit_label = "uniqueid", time_label = "year"
)
result <- fdid(s, tr_period = 1958:1961, ref_period = 1957)
summary(result)
```

fdid_list

Create an 'fdid_list' Object

Description

Bundles multiple 'fdid' objects into a single list with class "'fdid_list'" for convenient collective handling.

Usage

```
fdid_list(..., validate = TRUE)
```

Arguments

... One or more objects of class "'fdid'", or a single list of them.

validate Logical; if 'TRUE' (default) verify each element inherits from "'fdid'".

Value

A list with classes 'c("fdid_list", "list")'.

Author(s)

Rivka Lipkowitz

Description

Prepares a dataset for factorial difference-in-differences (FDID) analysis by reshaping the data into a wide format, averaging time-varying covariates, and renaming columns for consistency in subsequent analysis.

Usage

```
fdid_prepare(  
  data,  
  Y_label,  
  X_labels = NULL,  
  G_label,  
  unit_label,  
  time_label,  
  cluster_label = NULL  
)
```

Arguments

<code>data</code>	A data frame containing the dataset to be processed.
<code>Y_label</code>	A string specifying the column name of the outcome variable.
<code>X_labels</code>	A character vector specifying the column names of the time-varying covariates.
<code>G_label</code>	A string specifying the column name of the group variable (e.g., treatment vs. control).
<code>unit_label</code>	A string specifying the column name of the unit identifier (e.g., individual or entity).
<code>time_label</code>	A string specifying the column name of the time variable.
<code>cluster_label</code>	An optional string specifying the column name of the clustering variable. Default is 'NULL'.

Value

A data frame in wide format with the following: - Outcome variable pivoted to wide format with time columns. - Time-varying covariates averaged across time. - Columns renamed: - Unit identifier -> 'unit' - Covariates -> 'x1', 'x2', ... - Group variable -> 'G' - Clustering variable (if provided) -> 'c'

Author(s)

Rivka Lipkowitz

Examples

```

data <- data.frame(
  id = rep(1:3, each = 4),
  time = rep(1:4, times = 3),
  outcome = rnorm(12),
  covar1 = runif(12),
  covar2 = runif(12),
  group = c(0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1)
)
fdid_data <- fdid_prepare(
  data = data,
  Y_label = "outcome",
  X_labels = c("covar1", "covar2"),
  G_label = "group",
  unit_label = "id",
  time_label = "time"
)
head(fdid_data)

```

mortality

FDID example dataset

Description

A long-format panel dataset for demonstrating the fdid package.

Usage

```
mortality
```

Format

A data frame with 11973 rows and 17 columns:

provid Province ID

countyid County ID

zupu Genealogy book count

pczupu Genealogy book density (per capita); 45% of counties have zero

lnpczupu Log-transformed genealogy density: $\log(\text{pczupu} + 1)$; used as a continuous treatment in Xu, Zhao, and Ding (2026)

anyzupu Indicator: any genealogy book present

avggrain Average grain output

nograin Indicator: no grain data

urban Urban population share

dis_bj Distance to Beijing

dis_pc Distance to provincial capital
rice Rice cultivation indicator
minority Minority population share
edu Education level
lnpop Log population
year Year (1954–1966)
mortality Mortality rate

plot.fdid

Plot Results from FDID Analysis

Description

Provides visualisations for FDID results, including raw means, dynamic effects, and propensity-score overlap. The comparison plot of multiple methods has been removed; use `plot.fdid_list()` for that.

Usage

```
## S3 method for class 'fdid'
plot(
  x,
  type = c("raw", "dynamic", "overlap"),
  connected = FALSE,
  ci = TRUE,
  shade_periods = x$str_period,
  alpha_shade = 0.2,
  palette = "Set2",
  group_labels = c("Group 0", "Group 1"),
  xlab = NULL,
  ylab = NULL,
  main = NULL,
  ylim = NULL,
  ...
)
```

Arguments

<code>x</code>	An <code>fdid</code> object.
<code>type</code>	One of "raw", "dynamic", or "overlap".
<code>connected</code>	Logical; if TRUE, connects points with lines in the "raw" and "dynamic" plots. Default is FALSE.
<code>ci</code>	Logical; if TRUE, draw 95% CIs when available. Default is TRUE.

shade_periods Shaded intervals on the time axis. Default uses x\$str_period, i.e. event periods. Set to NULL to remove shaded area.

alpha_shade Transparency for shading the treatment period.

palette A palette name from **RColorBrewer**. Default "Set2".

group_labels Labels for the two groups.

xlab, ylab, main Axis labels and main title.

ylim Y-axis limits. Default NULL (computed automatically).

... Additional graphics parameters.

Value

Produces a plot; invisibly returns NULL.

Author(s)

Rivka Lipkowitz, Enhao Liu

Examples

```
data(fdid)
mortality$uniqueid <- paste(mortality$provid, mortality$countyid, sep = "-")
mortality$G <- ifelse(mortality$pczupu >= median(mortality$pczupu, na.rm = TRUE), 1, 0)
s <- fdid_prepare(
  data = mortality, Y_label = "mortality",
  X_labels = c("avggrain", "lnpop"),
  G_label = "G", unit_label = "uniqueid", time_label = "year"
)
result <- fdid(s, tr_period = 1958:1961, ref_period = 1957)
plot(result, type = "raw")
plot(result, type = "dynamic")
```

plot.fdid_list

Plot Multiple FDID Estimates

Description

Creates a comparison plot of point estimates and confidence intervals for every element of an 'fdid_list'.

Usage

```
## S3 method for class 'fdid_list'
plot(
  x,
  xlab = NULL,
  ylab = NULL,
```

```

    main = NULL,
    ylim = NULL,
    vertical = TRUE,
    show_vartype = TRUE,
    ...
)

```

Arguments

x	An object of class <code>"fdid_list"</code> .
xlab, ylab, main	Axis labels and title. If <code>'NULL'</code> , sensible defaults are used.
ylim	Optional numeric vector of length two giving the <i>estimate-axis</i> limits. (Backward compatible: for horizontal plots this is the x-limit; for vertical plots this is the y-limit.)
vertical	Logical; default is TRUE.
show_vartype	Logical; include vartype in labels. Default is TRUE.
...	Additional graphics parameters passed to <code>plot()</code> .

Value

Invisibly returns `'x'`; called for its side-effect of drawing a plot.

Author(s)

Rivka Lipkovitz, Enhao Liu

print.fdid *Print Method for FDID Objects*

Description

Print Method for FDID Objects

Usage

```

## S3 method for class 'fdid'
print(x, ...)

```

Arguments

x	An object of class <code>'fdid'</code> .
...	Additional arguments (not used).

Value

Prints a brief overview of the `'fdid'` object

Author(s)

Rivka Lipkovitz.

summary.fdid

Summary Method for FDID Objects

Description

Summary Method for FDID Objects

Usage

```
## S3 method for class 'fdid'  
summary(object, ...)
```

Arguments

object An object of class fdid.
... Additional arguments (not used).

Value

Prints a summary of the fdid object.

Author(s)

Rivka Lipkovitz, Enhao Liu

Examples

```
data(fdid)  
mortality$uniqueid <- paste(mortality$provid, mortality$countyid, sep = "-")  
mortality$G <- ifelse(mortality$pczupu >= median(mortality$pczupu, na.rm = TRUE), 1, 0)  
s <- fdid_prepare(  
  data = mortality, Y_label = "mortality",  
  X_labels = c("avggrain", "lnpop"),  
  G_label = "G", unit_label = "uniqueid", time_label = "year"  
)  
result <- fdid(s, tr_period = 1958:1961, ref_period = 1957)  
summary(result)
```

Index

* datasets

mortality, 6

fdid, 2

fdid_list, 4

fdid_prepare, 5

mortality, 6

plot.fdid, 7

plot.fdid_list, 8

print.fdid, 9

summary.fdid, 10