Package: didimputation (via r-universe)

October 28, 2026

Type Package

Title Imputation Estimator from Borusyak, Jaravel, and Spiess (2021)

Version 0.6.0

Description Estimates Two-way Fixed Effects
difference-in-differences/event-study models using the
imputation-based approach proposed by Borusyak, Jaravel, and
Spiess (2021).

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

Depends R (>= 3.5.0),

Imports Matrix, fixest (>= 0.13.2), data.table (>= 1.10.0)

Suggests haven, testthat (>= 3.0.0)

Config/testthat/edition 3

URL https://github.com/kylebutts/didimputation

License MIT + file LICENSE

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

Date/Publication 2026-09-28 00:22:11 UTC

RemoteUrl https://github.com/kylebutts/didimputation

RemoteRef HEAD

RemoteSha 5645e6b14bfb3fe3faa88a1c14a2339bbb03fbc3

Contents

df_het 2
df_hom 2
did_imputation 3

Index **6**

df_het *Simulated data with two treatment groups and heterogenous effects*

Description

Generated using the following call: `did2s::gen_data(panel = c(1990, 2020), g1 = 2000, g2 = 2010, g3 = 0, te1 = 2, te2 = 1, te3 = 0, te_m1 = 0.05, te_m2 = 0.15, te_m3 = 0)`

Usage

```
df_het
```

Format

A data frame with 31000 rows and 15 variables:

unit individual in panel data

year time in panel data

g the year that treatment starts

dep_var outcome variable

treat T/F variable for when treatment is on

rel_year year relative to treatment start. Inf = never treated.

rel_year_binned year relative to treatment start, but <=-6 and >=6 are binned.

unit_fe Unit FE

year_fe Year FE

error Random error component

te Static treatment effect = te

te_dynamic Dynamic treatmet effect = te_m

state State that unit is in

group String name for group

df_hom *Simulated data with two treatment groups and homogenous effects*

Description

Generated using the following call: `did2s::gen_data(panel = c(1990, 2020), g1 = 2000, g2 = 2010, g3 = 0, te1 = 2, te2 = 2, te3 = 0, te_m1 = 0, te_m2 = 0, te_m3 = 0)`

Usage

```
df_hom
```

Format

A data frame with 31000 rows and 15 variables:

unit individual in panel data

year time in panel data

g the year that treatment starts

dep_var outcome variable

treat T/F variable for when treatment is on

rel_year year relative to treatment start. Inf = never treated.

rel_year_binned year relative to treatment start, but <=-6 and >=6 are binned.

unit_fe Unit FE

year_fe Year FE

error Random error component

te Static treatment effect = te

te_dynamic Dynamic treatmet effect = te_m

group String name for group

state State that unit is in

weight Weight from runif()

`did_imputation` *Borusyak, Jaravel, and Spiess (2021) Estimator*

Description

Treatment effect estimation and pre-trend testing in staggered adoption diff-in-diff designs with an imputation approach of Borusyak, Jaravel, and Spiess (2021)

Usage

```
did_imputation(
  data,
  yname,
  gname,
  tname,
  idname,
  first_stage = NULL,
  wname = NULL,
  wtr = NULL,
  horizon = NULL,
  pretrends = NULL,
  cluster_var = NULL
)
```

Arguments

data	A data.frame
yname	String. Variable name for outcome. Use fixest c() syntax for multiple lhs, e.g. "c(y1, y2)".
gname	String. Variable name for unit-specific date of treatment (never-treated should be zero or NA).
tname	String. Variable name for calendar period.
idname	String. Variable name for unique unit id.
first_stage	Formula for Y(0). Formula following fixest::feols. Fixed effects specified after "|". If not specified, then just unit and time fixed effects will be used.
wname	String. Variable name for estimation weights of observations. This is used in estimating Y(0) and also augments treatment effect weights.
wtr	Character vector of treatment weight names (see horizon for standard static and event-study weights)
horizon	Integer vector of event_time or TRUE. This only applies if wtr is left as NULL. if specified, weighted averages/sums of treatment effects will be reported for each of these horizons separately (i.e. tau0 for the treatment period, tau1 for one period after treatment, etc.). If TRUE, all horizons are used. If wtr and horizon are null, then the static treatment effect is calculated.
pretrends	Integer vector or TRUE. Which pretrends to estimate. If TRUE, all pretrends are used. pretrends is used exclusively to perform a pre-trends test and does not affect the main estimation. It runs a separate OLS regression, as suggested by Borusyak, Jaravel, and Spiess (2024, Section 4.4, Test 1). This regression estimates deviations for each event time specified in pretrends relative to a reference period and the never-treated group. The reference period consists of all pre-treatment event times omitted from pretrends. To maintain full control over the reference period, avoid specifying all pre-treatment periods in pretrends. The pre-treatment estimates in the output are from this regression.
cluster_var	String. Variable name for clustering groups. If not supplied, then idname is used as default.

Details

The imputation-based estimator is a method of calculating treatment effects in a difference-in-differences framework. The method estimates a model for Y(0) using untreated/not-yet-treated observations and predicts Y(0) for the treated observations hat(Y_it(0)). The difference between treated and predicted untreated outcomes Y_it(1) - hat(Y_it(0)) serves as an estimate for the treatment effect for unit i in period t. These are then averaged to form average treatment effects for groups of (i, t).

Value

A data.frame containing treatment effect term, estimate, standard error and confidence interval. This is in tidy format.

Examples

Load example dataset which has two treatment groups and homogeneous treatment effects

```
data("df_hom", package="didimputation")
```

Static TWFE:

You can run a static TWFE fixed effect model for a simple treatment indicator

```
did_imputation(data = df_hom, yname = "dep_var", gname = "g",
               tname = "year", idname = "unit")
#>      term estimate  std.error conf.low conf.high
#>    <char>    <num>      <num>    <num>     <num>
#> 1:  treat 2.024639 0.03243596 1.961065  2.088214
```

Event Study:

Or you can use relative-treatment indicators to estimate an event study estimate

```
did_imputation(data = df_hom, yname = "dep_var", gname = "g",
               tname = "year", idname = "unit", horizon=TRUE)
#>       term estimate  std.error conf.low conf.high
#>     <char>    <num>      <num>    <num>     <num>
#>  1:      0 2.117232 0.07368419 1.972811  2.261653
#>  2:      1 1.856536 0.07672104 1.706163  2.006909
#>  3:      2 1.986357 0.07137180 1.846468  2.126246
#>  4:      3 2.004843 0.07653409 1.854836  2.154850
#>  5:      4 1.950228 0.07543636 1.802372  2.098083
#>  6:      5 2.038302 0.07580288 1.889728  2.186875
#>  7:      6 2.031571 0.07223098 1.889999  2.173144
#>  8:      7 2.025286 0.07541719 1.877468  2.173104
#>  9:      8 1.976081 0.07493409 1.829210  2.122951
#> 10:      9 2.121434 0.07268404 1.978974  2.263895
#> 11:     10 2.087984 0.08271442 1.925864  2.250105
#> 12:     11 1.942825 0.11421421 1.718965  2.166685
#> 13:     12 1.940532 0.11200348 1.721005  2.160059
#> 14:     13 1.964569 0.11361969 1.741875  2.187264
#> 15:     14 2.023456 0.11753255 1.793092  2.253820
#> 16:     15 2.235051 0.12110086 1.997693  2.472409
#> 17:     16 2.178438 0.11552325 1.952013  2.404864
#> 18:     17 1.935576 0.11278311 1.714521  2.156631
#> 19:     18 2.134953 0.10993120 1.919488  2.350418
#> 20:     19 2.111984 0.11146282 1.893517  2.330451
#> 21:     20 1.925168 0.11214206 1.705370  2.144967
#>       term estimate  std.error conf.low conf.high
```

Index

∗ **datasets**
    df_het, 2
    df_hom, 2

df_het, 2
df_hom, 2
did_imputation, 3

fixest::feols, *4*