Package 'rdstagger'

May 9, 2026

Type Package

Title Staggered Regression Discontinuity with Network Interference

Version 0.1.0

Description Implements a unified framework combining staggered difference-in-differences with regression discontinuity designs and network interference. Extends Callaway and Sant'Anna (2021) <doi:10.1016/j.jeconom.2020.12.001> to settings where treatment assignment is determined by a running variable crossing a cutoff, adoption timing is heterogeneous across units, and spillover effects operate through a known network structure. Provides group-time average treatment effects (direct and spillover), aggregation schemes, bandwidth selection, and pre-treatment falsification tests.

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.3.3

Depends R (>= 4.0.0)

Imports stats, sandwich, ggplot2, rdrobust,

Suggests testthat (>= 3.0.0), knitr, rmarkdown, dplyr, covr

VignetteBuilder knitr

Config/testthat/edition 3

URL https://github.com/causalfragility-lab/rdstagger

BugReports https://github.com/causalfragility-lab/rdstagger/issues

NeedsCompilation no

Author Subir Hait [aut, cre]

Maintainer Subir Hait <haitsubi@msu.edu>

Repository CRAN

Date/Publication 2026-05-05 18:50:02 UTC

Contents

`rdstagger-package` *rdstagger: Staggered Regression Discontinuity with Network Interference*

Description

Implements a unified framework combining staggered difference-in-differences with regression discontinuity designs and network interference. Extends Callaway and Sant'Anna (2021) to settings where:

1. Treatment assignment is determined by a running variable crossing a cutoff (RD)
2. Treatment adoption timing is heterogeneous across units (staggered DiD)
3. Spillover effects operate through a known network structure (interference)

Main functions

`sim_rdstagger` Simulate a staggered RD panel dataset with interference

`rdstagger_bw` Optimal bandwidth selection per cohort-time cell

`rdstagger_attgt` Estimate ATT(g,t) — direct and spillover effects

`rdstagger_spillover` Estimate spillover effects at network distance d

`rdstagger_agg` Aggregate ATT(g,t) into event-study or overall ATT

`rdstagger_pretest` Pre-treatment parallel trends falsification tests

References

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200-230.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6), 2295-2326.

Manski, C. F. (2013). Identification of treatment response with social interactions. *The Econometrics Journal*, 16(1), S1-S23.

Author(s)

Maintainer: Subir Hait <haitsubi@msu.edu>

See Also

Useful links:

- https://github.com/causalfragility-lab/rdstagger
- Report bugs at https://github.com/causalfragility-lab/rdstagger/issues

---

plot.rdstagger_agg *Plot Aggregated ATT Estimates*

Description

Produces a ggplot2 event-study or aggregation plot from an "rdstagger_agg" object.

Usage

```
## S3 method for class 'rdstagger_agg'
plot(x, ...)
```

Arguments

x	An object of class "rdstagger_agg".
...	Additional arguments (currently unused).

Value

A ggplot2 object.

Examples

```
sim <- sim_rdstagger(n = 300, nperiods = 6, n_cohorts = 2,
                     true_direct = 0.3, seed = 42)
res <- rdstagger_attgt(data = sim$data, yname = "y", xname = "x",
                       gname = "g", tname = "period", idname = "id",
                       bw = 1.5, boot = FALSE)
agg <- rdstagger_agg(res, type = "dynamic")
plot(agg)
```

rdstagger_agg *Aggregate ATT(g,t) Estimates*

Description

Aggregates group-time average treatment effects ATT(g,t) from rdstagger_attgt into summary estimands: event-study (dynamic), cohort-level, calendar-time, or overall ATT.

Usage

```
rdstagger_agg(
  x,
  type = c("dynamic", "group", "calendar", "overall"),
  min_periods = 1L
)
```

Arguments

x	An object of class "rdstagger_attgt" from rdstagger_attgt.
type	Character. Aggregation type: "dynamic" (event-study), "group" (by cohort), "calendar" (by calendar period), or "overall". Default "dynamic".
min_periods	Integer. Minimum number of cohort-time cells required to include an event-time bin. Default 1.

Value

An object of class "rdstagger_agg", a list with:

agg Data frame of aggregated estimates

type Aggregation type used

overall_att Simple overall ATT (post-treatment average)

attgt The original ATT(g,t) data frame

Examples

```
sim <- sim_rdstagger(n = 300, nperiods = 6, n_cohorts = 2,
                     true_direct = 0.3, seed = 42)
res <- rdstagger_attgt(data = sim$data, yname = "y", xname = "x",
                       gname = "g", tname = "period", idname = "id",
                       bw = 1.5, boot = FALSE)

# Event study
agg_dyn <- rdstagger_agg(res, type = "dynamic")
print(agg_dyn)
plot(agg_dyn)

# Overall ATT
```

```
agg_ov <- rdstagger_agg(res, type = "overall")
print(agg_ov)
```

`rdstagger_attgt` *Estimate Group-Time Average Treatment Effects in Staggered RD*

Description

Main estimation function for the staggered RD framework with network interference. Estimates ATT(g, t) — the average treatment effect for cohort g at time t — separately for direct effects on treated units and spillover effects on their network neighbors, within an RD bandwidth around the cutoff.

Usage

```
rdstagger_attgt(
  data,
  yname,
  xname,
  cutoff = 0,
  gname,
  tname,
  idname,
  network = NULL,
  bw = "optimal",
  control_group = c("nevertreated", "notyetreated"),
  xformla = NULL,
  doubly_robust = TRUE,
  boot = TRUE,
  nboot = 999L,
  alpha = 0.05,
  kernel = c("triangular", "epanechnikov", "uniform")
)
```

Arguments

`data`	A `data.frame` in long (panel) format.
`yname`	Character. Outcome variable name.
`xname`	Character. Running variable name.
`cutoff`	Numeric. RD cutoff. Default `0`.
`gname`	Character. Cohort variable name (`Inf`/`NA` for never-treated).
`tname`	Character. Time period variable name.
`idname`	Character. Unit identifier variable name.
`network`	Matrix or `NULL`. $n \times n$ adjacency matrix for spillover estimation. If `NULL`, spillovers are not estimated.

bw	Numeric or "optimal". Bandwidth around the cutoff. If "optimal", calls rdstagger_bw internally. Default "optimal".
control_group	Character. Which units form the control group. "nevertreated" (default) or "notyetreated".
xformla	Formula or NULL. Covariate formula for outcome regression (e.g., ~ age + income). Default NULL.
doubly_robust	Logical. Use doubly-robust estimator. Default TRUE.
boot	Logical. Compute bootstrap standard errors. Default TRUE.
nboot	Integer. Number of bootstrap replications. Default 999.
alpha	Numeric. Significance level for confidence intervals. Default 0.05.
kernel	Character. RD kernel. Default "triangular".

Value

An object of class "rdstagger_attgt", a list with:

attgt Data frame of ATT(g,t) estimates (direct effects)

spillgt Data frame of spillover ATT(g,t) estimates (if network supplied)

args List of call arguments

bandwidth Bandwidth used

Examples

```
sim <- sim_rdstagger(n = 300, nperiods = 6, n_cohorts = 2,
                     true_direct = 0.3, true_spill = 0.1, seed = 42)

res <- rdstagger_attgt(
  data    = sim$data,
  yname   = "y",
  xname   = "x",
  cutoff  = 0,
  gname   = "g",
  tname   = "period",
  idname  = "id",
  network = sim$network,
  bw      = 1.5,
  boot    = FALSE
)
print(res)
```

rdstagger_bw *Bandwidth Selection for Staggered RD*

Description

Computes an optimal bandwidth for each cohort-time cell using the mean-squared-error-optimal bandwidth selector from `rdrobust`. Separate bandwidths are estimated for pre-treatment and post-treatment periods to ensure appropriate comparison groups.

Usage

```
rdstagger_bw(
  data,
  yname,
  xname,
  cutoff = 0,
  gname,
  tname,
  kernel = c("triangular", "epanechnikov", "uniform"),
  bw_common = FALSE
)
```

Arguments

`data`	A `data.frame` containing the panel data.
`yname`	Character. Name of the outcome variable column.
`xname`	Character. Name of the running variable column.
`cutoff`	Numeric. The RD cutoff value. Default 0.
`gname`	Character. Name of the cohort variable column (`Inf` or `NA` for never-treated units).
`tname`	Character. Name of the time period column.
`kernel`	Character. Kernel for RD estimation. One of `"triangular"` (default), `"epanechnikov"`, or `"uniform"`.
`bw_common`	Logical. If `TRUE`, returns a single common bandwidth across all cohort-time cells (the median of cell-specific bandwidths). Default `FALSE`.

Value

A list with elements:

`bw_matrix` A matrix of bandwidths with rows = cohorts, columns = time periods

`bw_common` Single common bandwidth (median across cells)

`bw_summary` A `data.frame` summarising bandwidths by cohort and period

Examples

```
sim <- sim_rdstagger(n = 400, nperiods = 6, n_cohorts = 2, seed = 42)
bw  <- rdstagger_bw(data = sim$data, yname = "y", xname = "x",
                    cutoff = 0, gname = "g", tname = "period")
bw$bw_common
bw$bw_summary
```

`rdstagger_pretest` *Pre-Treatment Falsification Tests for Staggered RD*

Description

Tests the pre-treatment parallel trends assumption within the RD bandwidth. Performs a joint test across all pre-treatment cohort-time cells and individual cell tests, analogous to `pretest` in the `did` package but adapted for the staggered RD setting.

Usage

```
rdstagger_pretest(x, method = c("joint", "individual", "both"))
```

Arguments

`x`	An object of class `"rdstagger_attgt"`.
`method`	Character. Test method: `"joint"` (chi-squared joint test, default), `"individual"` (t-tests per cell), or `"both"`.

Value

A list with elements:

`joint` Joint test statistic, df, and p-value (if requested)

`individual` Data frame of per-cell tests (if requested)

`passes` Logical. `TRUE` if joint test p-value > 0.05

Examples

```
sim <- sim_rdstagger(n = 400, nperiods = 8, n_cohorts = 2,
                     true_direct = 0.3, seed = 42)
res <- rdstagger_attgt(data = sim$data, yname = "y", xname = "x",
                       gname = "g", tname = "period", idname = "id",
                       bw = 1.5, boot = FALSE)
pt  <- rdstagger_pretest(res)
print(pt)
```

`rdstagger_spillover` *Estimate Spillover Effects in Staggered RD*

Description

Estimates the spillover (indirect) treatment effects on network neighbors of treated units within the RD bandwidth. Spillover effects are estimated separately for each cohort-time cell.

Usage

```
rdstagger_spillover(
  data,
  yname,
  xname,
  cutoff = 0,
  gname,
  tname,
  idname,
  network,
  bw,
  kernel = c("triangular", "epanechnikov", "uniform"),
  boot = TRUE,
  nboot = 999L,
  alpha = 0.05
)
```

Arguments

`data`	A `data.frame` in long panel format.
`yname`	Character. Outcome variable name.
`xname`	Character. Running variable name.
`cutoff`	Numeric. RD cutoff. Default `0`.
`gname`	Character. Cohort variable name.
`tname`	Character. Time period variable name.
`idname`	Character. Unit identifier variable name.
`network`	Matrix. $n \times n$ adjacency matrix.
`bw`	Numeric. Bandwidth around the cutoff.
`kernel`	Character. RD kernel. Default `"triangular"`.
`boot`	Logical. Bootstrap standard errors. Default `TRUE`.
`nboot`	Integer. Bootstrap replications. Default `999`.
`alpha`	Numeric. Significance level. Default `0.05`.

Value

A `data.frame` with columns:

`cohort` Treatment cohort

`period` Time period

`spill_att` Spillover ATT estimate

`se` Standard error

`ci_lower`, `ci_upper` Confidence interval

`pval` p-value

`n_exposed` Number of exposed neighbors

Examples

```
sim <- sim_rdstagger(n = 300, nperiods = 6, n_cohorts = 2,
                     true_direct = 0.3, true_spill = 0.15, seed = 42)
sp  <- rdstagger_spillover(
  data    = sim$data,  yname = "y",   xname = "x",
  gname   = "g",       tname = "period", idname = "id",
  network = sim$network, bw = 1.5, boot = FALSE
)
head(sp)
```

---

sim_rdstagger *Simulate a Staggered RD Panel Dataset with Network Interference*

Description

Generates synthetic panel data suitable for testing and demonstrating the `rdstagger` estimators. The data generating process features a running variable with a cutoff-based treatment assignment, staggered adoption across cohorts, and network spillover effects.

Usage

```
sim_rdstagger(
  n = 500,
  nperiods = 8,
  n_cohorts = 3,
  cutoff = 0,
  bw = 1,
  network_density = 0.1,
  true_direct = 0.3,
  true_spill = 0.1,
  outcome_type = c("continuous", "binary", "count"),
  heterogeneous_te = FALSE,
  seed = NULL
)
```

Arguments

n	Integer. Number of units. Default 500.
nperiods	Integer. Number of time periods. Default 8.
n_cohorts	Integer. Number of treatment cohorts. Default 3.
cutoff	Numeric. RD cutoff value on the running variable. Default 0.
bw	Numeric. True bandwidth around the cutoff. Default 1.
network_density	Numeric. Probability of a network tie between any two units (Erdos-Renyi model). Must be in (0, 1). Default 0.1.
true_direct	Numeric. True direct average treatment effect. Default 0.3.
true_spill	Numeric. True spillover effect on network neighbors. Default 0.1.
outcome_type	Character. One of "continuous", "binary", or "count". Default "continuous".
heterogeneous_te	Logical. If TRUE, treatment effects vary by cohort. Default FALSE.
seed	Integer. Random seed for reproducibility. Default NULL.

Value

A list with three elements:

data A data.frame with columns: id, period, y, x (running variable), g (cohort, Inf for never-treated), treated, neighbor_treated, spillover_share

network An $n \times n$ adjacency matrix

true_params A list of the true parameter values used to generate the data

Examples

```
# Basic continuous outcome
sim <- sim_rdstagger(n = 300, nperiods = 6, n_cohorts = 2,
                     true_direct = 0.3, true_spill = 0.1, seed = 42)
head(sim$data)
sim$true_params

# Binary outcome
sim_bin <- sim_rdstagger(n = 500, nperiods = 8, n_cohorts = 3,
                         outcome_type = "binary", seed = 123)
table(sim_bin$data$y)

# Count outcome
sim_cnt <- sim_rdstagger(n = 400, nperiods = 6, n_cohorts = 2,
                         outcome_type = "count", true_direct = 0.5,
                         seed = 999)
```

Index