Package: contdid (via r-universe)

October 30, 2026

Title Difference-in-Differences with a Continuous Treatment

Version 0.1.2

Description Provides methods for difference-in-differences with a continuous treatment and staggered treatment adoption. Includes estimation of treatment effects and causal responses as a function of the dose, event studies indexed by length of exposure to the treatment, and aggregation into overall average effects. Uniform inference procedures are included, along with both parametric and nonparametric models for treatment effects. The methods are based on Callaway, Goodman-Bacon, and Sant'Anna (2025) <doi:10.48550/arXiv.2107.02637>.

Depends R (>= 4.1.0),

License GPL-3

Encoding UTF-8

Imports BMisc (>= 1.4.8), ptetools (>= 1.0.1), checkmate, splines2, sandwich, ggplot2, MASS, npiv

URL https://bcallaway11.github.io/contdid/, https://github.com/bcallaway11/contdid

BugReports https://github.com/bcallaway11/contdid/issues

Suggests testthat (>= 3.0.0), tidyr

Config/testthat/edition 3

Config/roxygen2/version 8.1.0

NeedsCompilation no

Author Brantly Callaway [aut, cre], Andrew Goodman-Bacon [aut], Pedro H. C. Sant'Anna [aut]

Maintainer Brantly Callaway <brantly.callaway@uga.edu>

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

Date/Publication 2026-10-30 19:33:27 UTC

RemoteUrl https://github.com/cran/contdid

RemoteRef HEAD

RemoteSha e6094f41f6d0ad28dc7b9f4b1ae2cd48df2d9f3a

Contents

choose_knots_even 2
choose_knots_quantile 2
cont_did 3
cont_did_acrt 8
cont_two_by_two_subset 9
ggcont_did 10
setup_pte_cont 11
simulate_contdid_data 13

Index **15**

`choose_knots_even` *Choose Evenly Spaced Knots*

Description

A function to place equally spaced knots for fitting b-splines

Usage

```
choose_knots_even(x, num_knots)
```

Arguments

`x`	vector of treatment doses
`num_knots`	the number of knots to use

Value

a vector containing the locations of the knots

`choose_knots_quantile` *Choose Knots at Quantiles*

Description

A function to choose knots for fitting b-splines by the quantile of x

Usage

```
choose_knots_quantile(x, num_knots)
```

Arguments

x	vector of treatment doses
num_knots	the number of knots to use

Value

a vector containing the locations of the knots

---

cont_did *Difference-in-differences with a Continuous Treatment*

Description

A function for difference-in-differences with a continuous treatment in a staggered treatment adoption setting.

cont_did currently supports staggered treatment with continuous treatments using B-splines under the hood.

Usage

```
cont_did(
  yname,
  dname,
  gname = NULL,
  tname,
  idname,
  xformula = ~1,
  data,
  target_parameter = c("level", "slope"),
  aggregation = c("dose", "eventstudy", "none"),
  treatment_type = c("continuous", "discrete"),
  dose_est_method = c("parametric", "cck"),
  dvals = NULL,
  degree = 3,
  num_knots = 0,
  allow_unbalanced_panel = FALSE,
  control_group = c("notyettreated", "nevertreated", "eventuallytreated"),
  anticipation = 0,
  weightsname = NULL,
  alp = 0.05,
  bstrap = TRUE,
  cband = FALSE,
  boot_type = "multiplier",
  biters = 1000,
  clustervars = NULL,
  est_method = NULL,
```

```
  base_period = "varying",
  print_details = FALSE,
  cl = 1,
  ...
)
```

Arguments

yname	The name of the outcome variable
dname	The name of the treatment variable in the data. The functionality of cont_did is different from the did package in that the treatment variable is the "amount" of the treatment in a particular period, rather than gname which gives the time period when a unit becomes treated. The dname variable should, for a particular unit, be constant across time periods—even in pre-treatment periods. For units that never participate in the treatment, the amount of the treatment may not be defined in some applications—it is ignored in this function.
gname	The name of the timing-group variable, i.e., when treatment starts for a particular unit. The value of this variable should be set to be 0 for units that do not participate in the treatment in any time period.
tname	The name of the column containing the time periods
idname	The individual (cross-sectional unit) id name
xformula	A formula for additional covariates. This is not currently supported.
data	The name of the data.frame that contains the data
target_parameter	Two options are "level" and "slope". In the first case, the function will report level effects, i.e., ATT's. In the second case, the function will report slope effects, i.e., ACRT's
aggregation	"dose" averages across timing-groups and time periods and provides results as a function of the dose. "eventstudy" averages across timing-groups and doses and reports results as a function of the length of exposure to the treatment. "none" is a stub for reporting fully disaggregated results that can be processed as desired by the user. This is not currently supported though. The combination of the arguments target_parameter and aggregation strongly affects the behavior of the function (and target of the analysis). For example, setting target_parameter="level" and aggregation="eventstudy" is effectively the same thing as binarizing the treatment (i.e., where units are considered treated if they experience any positive amount of the treatment) and reporting an event study.
treatment_type	"continuous" or "discrete" depending on the nature of the treatment. Default is "continuous". "discrete" is not yet supported.
dose_est_method	The method used to estimate the dose-specific effects. The default is "parametric", where the user needs to specify the number of knots and degree for a B-spline which is assumed to be correctly specified. The other option is "cck" which uses the a data-driven nonparametric method to estimate the dose-specific effects based on the npiv package and Chen, Christensen, and Kankanala (ReStud, 2025).

`dvals` The values of the treatment at which to compute dose-specific effects. If it is not specified, the default choice will be use the percentiles of the dose among all ever-treated units.

`degree` The degree of the B-Spline used in estimation. The default is 3, which in combination with the default choice for the `num-knots`, leads to fitting models for the group of treated units that only that is a cubic polynomial in the dose. Setting `degree=1` will lead to a linear model, while setting `degree=2` will lead to a quadratic model.

`num_knots` The number of knots to include for the B-Spline. The default is 0 so that the spline is global (i.e., this will amount to fitting a global polynomial). There is a bias-variance tradeoff for including more or less knots.

`allow_unbalanced_panel`
Whether or not function should "balance" the panel with respect to time and id. The default value is `FALSE` which means that `att_gt()` will drop all units where data is not observed in all periods. The advantage of this is that the computations are faster (sometimes substantially).

`control_group` Which units to use as the control group. The default is "nevertreated" which sets the control group to be the group of units that never participate in the treatment. This group does not change across groups or time periods. The other option is to set `group="notyettreated"`. In this case, the control group is set to the group of units that have not yet participated in the treatment in that time period. This includes all never treated units, but it includes additional units that eventually participate in the treatment, but have not participated yet.

`anticipation` The number of time periods before participating in the treatment where units can anticipate participating in the treatment and therefore it can affect their untreated potential outcomes

`weightsname` The name of the column containing the sampling weights. If not set, all observations have same weight. When weights are time-invariant (constant within each unit across periods), all `fix_weights` options produce identical results and no special handling is needed.

When weights vary across time (e.g., time-varying population sizes), the default behavior differs by panel type:

Balanced panel Each 2x2 DiD comparison uses the weight from the earlier of the two time periods involved. For post-treatment cells, this is the base period (g-1). For pre-treatment cells with `base_period="varying"`, this is the pre-treatment period itself. The panel DRDID estimators are used.

Repeated cross sections and unbalanced panels Both periods' per-observation weights are passed directly to the RC DRDID estimators, so each observation carries its own period-specific weight.

Use the `fix_weights` argument to override the default behavior.

`alp` the significance level, default is 0.05

`bstrap` Boolean for whether or not to compute standard errors using the multiplier bootstrap. Default is `TRUE` (in addition, cband is also by default `TRUE` indicating that uniform confidence bands will be returned). If `bstrap=FALSE`, analytical standard errors are reported; these are cluster-robust when `clustervars` is supplied.

cband — Boolean for whether or not to compute a uniform confidence band that covers all of the group-time average treatment effects with fixed probability 1-alp. In order to compute uniform confidence bands, bstrap must also be set to TRUE. The default is TRUE.

boot_type — should be one of "multiplier" (the default) or "empirical". The multiplier bootstrap is generally much faster, but attgt_fun needs to provide an expression for the influence function (which could be challenging to figure out). If no influence function is provided, then the pte package will use the empirical bootstrap no matter what the value of this parameter. bstrap = FALSE (purely analytical standard errors) takes priority over boot_type: it is only available when attgt_fun returns an influence function, and it overrides boot_type = "empirical".

biters — The number of bootstrap iterations to use. The default is 1000, and this is only applicable if bstrap=TRUE.

clustervars — A vector of variables names to cluster on. At most, there can be two variables (otherwise will throw an error) and one of these must be the same as idname which allows for clustering at the individual level. Clustered standard errors are available with the multiplier bootstrap (bstrap=TRUE) or analytically (bstrap=FALSE).

est_method — the method to compute group-time average treatment effects. The default is "dr" which uses the doubly robust approach in the DRDID package. Other built-in methods include "ipw" for inverse probability weighting and "reg" for first step regression estimators. The user can also pass their own function for estimating group time average treatment effects. The required signature depends on the data structure:

Panel data (panel=TRUE): f(y1, y0, D, covariates, i.weights, inffunc, ...) where y1 is an n x 1 vector of post-treatment outcomes, y0 is an n x 1 vector of pre-treatment outcomes, D is a binary vector indicating treatment group membership, covariates is an n x k matrix, i.weights is a vector of sampling weights, and inffunc is a logical requesting influence-function computation.

Repeated cross sections / unbalanced panel (panel=FALSE): f(y, post, D, covariates, i.weights, inffunc, ...) where y is the outcome vector (length n), post is a binary indicator for the post-treatment period, D is a binary treatment indicator, covariates is an n x k matrix, i.weights is a vector of sampling weights, and inffunc is a logical.

In both cases the function should return a list that includes ATT (the estimated group-time average treatment effect) and att.inf.func (an n x 1 influence function — one entry per observation passed into the estimator). The function can return other things as well, but these are the only two that are required. With no covariates (xformla = NULL), the built-in methods ("dr", "ipw", "reg") all reduce to the unconditional difference-in-differences estimator, so the choice among them is irrelevant; a custom est_method function is still called (with an intercept-only covariates matrix) and determines the estimates.

base_period — Whether to use a "varying" base period or a "universal" base period. Either choice results in the same post-treatment estimates of ATT(g,t)'s. In pre-treatment periods, using a varying base period amounts to computing a pseudo-ATT in each treatment period by comparing the change in outcomes for a particular

group relative to its comparison group in the pre-treatment periods (i.e., in pre-treatment periods this setting computes changes from period t-1 to period t, but repeatedly changes the value of t)

A universal base period fixes the base period to always be (g-anticipation-1). This does not compute pseudo-ATT(g,t)'s in pre-treatment periods, but rather reports average changes in outcomes from period t to (g-anticipation-1) for a particular group relative to its comparison group. This is analogous to what is often reported in event study regressions.

Using a varying base period results in an estimate of ATT(g,t) being reported in the period immediately before treatment. Using a universal base period normalizes the estimate in the period right before treatment (or earlier when the user allows for anticipation) to be equal to 0, but one extra estimate in an earlier period.

`print_details` Whether or not to show details/progress of computations. Default is `FALSE`.

`cl` number of clusters to be used when bootstrapping; default is 1

`...` Additional arguments to be passed to a custom `est_method` function. These are ignored when using built-in estimation methods (`"dr"`, `"ipw"`, `"reg"`).

Value

cont_did_obj

Examples

```
# build small simulated data
set.seed(1234)
df <- simulate_contdid_data(
  n = 1000,
  num_time_periods = 4,
  num_groups = 4,
  dose_linear_effect = 0,
  dose_quadratic_effect = 0
)

# estimate effects of continuous treatment
cd_res <- cont_did(
  yname = "Y",
  tname = "time_period",
  idname = "id",
  dname = "D",
  data = df,
  gname = "G",
  target_parameter = "slope",
  aggregation = "dose",
  treatment_type = "continuous",
  control_group = "notyettreated",
  biters = 50,
  cband = TRUE,
  num_knots = 1,
  degree = 3,
```

```
)

summary(cd_res)
```

`cont_did_acrt` *Compute ACRT's for a Timing Group and Time Period*

Description

This is the main function for computing dose-specific effects of a continuous treatment, given a particular timing group and time period.

Usage

```
cont_did_acrt(gt_data, dvals = NULL, degree = 1, knots = numeric(0), ...)
```

Arguments

`gt_data`	data that is "local" to a particular group-time average treatment effect
`dvals`	The values of the treatment at which to compute dose-specific effects. If it is not specified, the default choice will be use the percentiles of the dose among all ever-treated units.
`degree`	The degree of the B-Spline used in estimation. The default is 3, which in combination with the default choice for the `num-knots`, leads to fitting models for the group of treated units that only that is a cubic polynomial in the dose. Setting `degree=1` will lead to a linear model, while setting `degree=2` will lead to a quadratic model.
`knots`	A vector of placements of knots for b-splines. Since this function is typically called internally, this would typically be set by the calling function.
`...`	additional arguments

Value

ptetools::attgt_if object

cont_two_by_two_subset *Continuous Two-by-Two Subset*

Description

A function for computing a 2x2 subset of original data. This function is adapted from `ptetools::two_by_two_subset` and allows for the treatment to be continuous. This is the subset with post treatment periods separately for the treated group and comparison group and pre-treatment periods in the period immediately before the treated group became treated.

Usage

```
cont_two_by_two_subset(
  data,
  g,
  tp,
  control_group = "notyettreated",
  anticipation = 0,
  base_period = "varying",
  ...
)
```

Arguments

`data`	the full dataset
`g`	the current group
`tp`	the current time period
`control_group`	whether to use "notyettreated" (default) or "nevertreated"
`anticipation`	the number of periods of anticipation (i.e., number of periods before the treatment happens where the treatment can "already" affect the outcome)
`base_period`	The type of base period to use. This only affects the numeric value of results in pre-treatment periods. Results in post-treatment periods are not affected by this choice. The default is "varying", where the base period will "back up" to the immediately preceding period in pre-treatment periods. The other option is "universal" where the base period is fixed in pre-treatment periods to be the period right before the treatment starts. "Universal" is commonly used in difference-in-differences applications, but can be unnatural for other identification strategies.
`...`	extra arguments to get the subset correct

Value

list that contains correct subset of data, `n1` number of observations in this subset, and `disidx` a vector of the correct ids for this subset.

ggcont_did *Plot Results with a Continuous Treatment*

Description

a function to plot results with a continuous treatment

Usage

```
ggcont_did(dose_obj, type = "att")
```

Arguments

dose_obj	a result from running cont_did
type	whether to plot ATT(d) or ACRT(d), defaults to att for plotting ATT(d). For ACRT(d), use "acrt"

Value

A ggplot object

Examples

```
# build small simulated data
set.seed(1234)
df <- simulate_contdid_data(
    n = 5000,
    num_time_periods = 4,
    num_groups = 4,
    dose_linear_effect = 0,
    dose_quadratic_effect = 0
)

# estimate effects of continuous treatment
cd_res <- cont_did(
    yname = "Y",
    tname = "time_period",
    idname = "id",
    dname = "D",
    data = df,
    gname = "G",
    target_parameter = "slope",
    aggregation = "dose",
    treatment_type = "continuous",
    control_group = "notyettreated",
    biters = 50,
    cband = TRUE,
    num_knots = 1,
    degree = 3,
```

```
)

# plot ATT as a function of the dose
ggcont_did(cd_res, type = "att")

# plot ACRT as a function of the dose
ggcont_did(cd_res, type = "acrt")
```

`setup_pte_cont` *Setup for DiD with a Continuous Treatment*

Description

A function that creates a pte_params object, adding several different variables that are needed when there is a continuous treatment.

Usage

```
setup_pte_cont(
  yname,
  gname,
  tname,
  idname,
  data,
  xformula = ~1,
  target_parameter,
  aggregation,
  treatment_type,
  required_pre_periods = 1,
  anticipation = 0,
  base_period = "varying",
  cband = TRUE,
  alp = 0.05,
  boot_type = "multiplier",
  weightsname = NULL,
  gt_type = "att",
  biters = 100,
  cl = 1,
  dname,
  dvals = NULL,
  degree = 1,
  num_knots = 0,
  ...
)
```

Arguments

yname	Name of outcome in data
gname	Name of group in data
tname	Name of time period in data
idname	Name of id in data
data	balanced panel or repeated cross sections data
xformula	A formula for additional covariates. This is not currently supported.
target_parameter	Two options are "level" and "slope". In the first case, the function will report level effects, i.e., ATT's. In the second case, the function will report slope effects, i.e., ACRT's
aggregation	"dose" averages across timing-groups and time periods and provides results as a function of the dose. "eventstudy" averages across timing-groups and doses and reports results as a function of the length of exposure to the treatment. "none" is a stub for reporting fully disaggregated results that can be processed as desired by the user. This is not currently supported though. The combination of the arguments target_parameter and aggregation strongly affects the behavior of the function (and target of the analysis). For example, setting target_parameter="level" and aggregation="eventstudy" is effectively the same thing as binarizing the treatment (i.e., where units are considered treated if they experience any positive amount of the treatment) and reporting an event study.
treatment_type	"continuous" or "discrete" depending on the nature of the treatment. Default is "continuous". "discrete" is not yet supported.
required_pre_periods	The number of required pre-treatment periods to implement the estimation strategy. Default is 1.
anticipation	how many periods before the treatment actually takes place that it can have an effect on outcomes
base_period	The type of base period to use. This only affects the numeric value of results in pre-treatment periods. Results in post-treatment periods are not affected by this choice. The default is "varying", where the base period will "back up" to the immediately preceding period in pre-treatment periods. The other option is "universal" where the base period is fixed in pre-treatment periods to be the period right before the treatment starts. "Universal" is commonly used in difference-in-differences applications, but can be unnatural for other identification strategies.
cband	whether or not to report a uniform (instead of pointwise) confidence band (default is TRUE)
alp	significance level; default is 0.05
boot_type	which type of bootstrap to use
weightsname	The name of the column that contains sampling weights. The default is NULL, in which case no sampling weights are used.
gt_type	which type of group-time effects are computed. The default is "att". Different estimation strategies can implement their own choices for gt_type

biters	number of bootstrap iterations; default is 100
cl	number of clusters to be used when bootstrapping; default is 1
dname	The name of the treatment variable in the data. The functionality of cont_did is different from the did package in that the treatment variable is the "amount" of the treatment in a particular period, rather than gname which gives the time period when a unit becomes treated. The dname variable should, for a particular unit, be constant across time periods—even in pre-treatment periods. For units that never participate in the treatment, the amount of the treatment may not be defined in some applications—it is ignored in this function.
dvals	an optional argument specifying which values of the treatment to evaluate ATT(d) and/or ACRT(d). If no values are supplied, then the default behavior is to set dvals to be the 1st to 99th percentiles of the dose among units that experience any positive dose.
degree	The degree of the B-Spline used in estimation. The default is 3, which in combination with the default choice for the num-knots, leads to fitting models for the group of treated units that only that is a cubic polynomial in the dose. Setting degree=1 will lead to a linear model, while setting degree=2 will lead to a quadratic model.
num_knots	The number of knots to include for the B-Spline. The default is 0 so that the spline is global (i.e., this will amount to fitting a global polynomial). There is a bias-variance tradeoff for including more or less knots.
...	additional arguments

Value

pte_params object

---

simulate_contdid_data *Simulate data for DiD with a Continuous Treatment*

Description

A function that simulates panel data when there is a continuous treatment.

Besides the parameters that can be passed to the function, some values are hard coded. The individual fixed effect is drawn from a normal distribution with mean equal to the group. The time effects are hard coded to be equal to the time period. The dose is drawn from a uniform distribution between 0 and 1.

Usage

```
simulate_contdid_data(
  n = 5000,
  num_time_periods = 4,
  num_groups = num_time_periods,
  pg = rep(1/num_groups, num_groups - 1),
```

```
  pu = 1/(num_groups),
  dose_linear_effect = 0,
  dose_quadratic_effect = 0
)
```

Arguments

`n`	The number of cross-sectional units. Default is 5000.
`num_time_periods`	The number of time periods. Default is 4.
`num_groups`	The number of groups. Default is the number of time periods. In this case, the groups will consist of a never-treated group and groups that become treated in every period starting in the second period.
`pg`	A vector of probabilities that a unit will be in a particular treated group. The default is equal probabilities.
`pu`	The probability that a unit will be in the never-treated group. The default is that it is 1/num_groups.
`dose_linear_effect`	The linear effect of the treatment. Default is 0.
`dose_quadratic_effect`	The quadratic effect of the treatment. Default is 0.

Value

A balanced panel data frame with the following columns:

- id: unit id
- time_period: time period
- Y: outcome
- G: unit's group
- D: amount of the treatment

Index

att_gt(), *5*

choose_knots_even, 2
choose_knots_quantile, 2
cont_did, 3
cont_did_acrt, 8
cont_two_by_two_subset, 9

ggcont_did, 10

setup_pte_cont, 11
simulate_contdid_data, 13