Package 'rdhte'

July 18, 2025

Title Heterogeneous Treatment Effects in Regression Discontinuity Designs

Version 0.1.0

Description Understanding heterogeneous causal effects based on pretreatment covariates is a crucial step in modern empirical work in data science. Building on the recent developments in Calonico et al (2025) <https://rdpackages.github.io/references/Calonico-Cattaneo-Farrell-Palomba-Titiunik_2025_HTERD.pdf>, this package provides tools for estimation and inference of heterogeneous treatment effects in Regression Discontinuity (RD) Designs. The package includes two main commands: 'rdhte' to conduct estimation and robust bias-corrected inference for conditional RD treatment effects (given choice of bandwidth parameter); 'rdbwhte', which implements automatic bandwidth selection methods; and 'rdhte_lincom' to test linear combinations of parameters.

License GPL-3

Imports rdrobust, sandwich, multcomp

Encoding UTF-8

RoxygenNote 7.3.2

Suggests knitr, rmarkdown

VignetteBuilder knitr

Depends R (>= 3.5)

NeedsCompilation no

Author Sebastian Calonico [aut, cre],
Matias D. Cattaneo [aut],
Max H. Farrell [aut],
Filippo Palomba [aut],
Rocio Titiunik [aut]

Maintainer Sebastian Calonico <scalonico@ucdavis.edu>

Repository CRAN

Date/Publication 2025-07-18 20:10:02 UTC

Contents

`rdhte-package` *rdhte: RD Heterogeneous Treatment Effects Estimation and Inference*

Description

Building on the recent developments in Calonico, Cattaneo, Farrell, Palomba, and Titiunik (2025), this package implements estimation and inference of heterogeneous treatment effects in RD designs. The package includes two main commands: `rdhte` conduct estimation and robust bias-corrected inference for conditional RD treatment effects, for a given choice of bandwidth parameter; and `rdbwhte` implements automatic bandwidth selection methods. We illustrate the methods implemented in the package `rdhte` using a canonical empirical application. We also demonstrate how the package `rdhte` complements, and in very specific cases recovers, the methods available in the packages rdrobust (Calonico, Cattaneo, Farrell, Titiunik (2017) and rdmulti, Cattaneo, Titiunik, VazquezBare (2020).

Commands: `rdhte` for estimation and inference. `rdbwhte` for data-driven bandwidth selection.

Related Stata and R packages useful for inference in regression discontinuity (RD) designs are described in the website: `https://rdpackages.github.io/`.

Author(s)

Sebastian Calonico, University of California, Davis <scalonico@ucdavis.edu>.

Matias D. Cattaneo, Princeton University <cattaneo@princeton.edu>.

Max H. Farrell, University of California, Santa Barbara <maxhfarrell@ucsb.edu>.

Filippo Palomba, Princeton University <fpalomba@princeton.edu>.

Rocio Titiunik, Princeton University <titiunik@princeton.edu>.

References

Calonico, Cattaneo, Farrell, Palomba and Titiunik (2025): rdhte: Learning Conditional Average Treatment Effects in RD Designs. *Working paper*.

Calonico, Cattaneo, Farrell, Palomba and Titiunik (2025): Treatment Effect Heterogeneity in Regression Discontinuity Designs. *Working paper*

rdbwhte *Data-Driven Optimal Bandwidth Selection for RD Heterogeneous Treatment Effects Estimation*

Description

rdbwhte computes MSE- and CER-optimal bandwidths for estimating RD heterogeneous treatment effects based on covariates.

Companion commands: rdhte for RD HTE estimation and inference, and rdhte_lincom for testing linear restrictions of parameters.

Related Stata and R packages useful for inference in RD designs are described in the website: https://rdpackages.github.io/.

Usage

```
rdbwhte(
  y,
  x,
  c = 0,
  covs.hte = NULL,
  covs.eff = NULL,
  p = 1,
  q = 2,
  kernel = "tri",
  vce = "hc3",
  cluster = NULL,
  bwselect = "mserd",
  bw.joint = FALSE
)
```

Arguments

y	Outcome variable.
x	Running variable.
c	RD cutoff in x; default is c = 0.
covs.hte	covariates for heterogeneous treatment effects. Factor variables can be used to distinguish between continuous and categorical variables, select reference categories, specify interactions between variables, and include polynomials of continuous variables.
covs.eff	additional covariates to be used for efficiency improvements.
p	order of the local polynomial used to construct the point estimator (default = 1).
q	order of the local polynomial used to construct the bias correction (default = 2).
kernel	kernel function used to construct the RD estimators. Options are triangular (default option), epanechnikov and uniform.

vce	character string specifying the variance-covariance matrix estimator type (hc0–hc3) (default = "hc3").
cluster	variable indicating the clustering of observations.
bwselect	bandwidth selection procedure to be used. Options are: mserd one common MSE-optimal bandwidth selector for the RD treatment effect estimator. msetwo two different MSE-optimal bandwidth selectors (below and above the cutoff) for the RD treatment effect estimator. msesum one common MSE-optimal bandwidth selector for the sum of regression estimates (as opposed to difference thereof). msecomb1 for min(mserd,msesum). msecomb2 for median(msetwo,mserd,msesum), for each side of the cutoff separately. cerrd one common CER-optimal bandwidth selector for the RD treatment effect estimator. certwo two different CER-optimal bandwidth selectors (below and above the cutoff) for the RD treatment effect estimator. cersum one common CER-optimal bandwidth selector for the sum of regression estimates (as opposed to difference thereof). cercomb1 for min(cerrd,cersum). cercomb2 for median(certwo,cerrd,cersum), for each side of the cutoff separately. Note: MSE = Mean Square Error; CER = Coverage Error Rate. Default is bwselect=mserd.
bw.joint	logical. If TRUE, forces all bandwidths to be the same across groups (default is bw.joint = FALSE).

Value

A list with selected bandwidths and model information.

W.lev	vector of group level identifiers.
kernel	kernel type used.
vce	variance estimator used.
c	cutoff value.
h	vector containing the bandwidths used.
p	order of the polynomial used for estimation.
q	order of the polynomial used for inference.
N	vector with the original number of observations for each group.
Nh	vector with the effective number of observations for each group.
covs.cont	internal value.
rdmodel	rd model.

Author(s)

Sebastian Calonico, University of California, Davis <scalonico@ucdavis.edu>.

Matias D. Cattaneo, Princeton University <cattaneo@princeton.edu>.

Max H. Farrell, University of California, Santa Barbara <maxhfarrell@ucsb.edu>.

Filippo Palomba, Princeton University <fpalomba@princeton.edu>.

Rocio Titiunik, Princeton University <titiunik@princeton.edu>.

References

Calonico, Cattaneo, Farrell, Palomba and Titiunik (2025): rdhte: Learning Conditional Average Treatment Effects in RD Designs. *Working paper*.

Calonico, Cattaneo, Farrell, Palomba and Titiunik (2025): Treatment Effect Heterogeneity in Regression Discontinuity Designs. *Working paper*

See Also

rdhte, rdhte_lincom

Examples

```
set.seed(123)
n <- 5000
X <- runif(n, -1, 1)
W <- rbinom(n, 1, 0.5)
Y <- 3 + 2*X + 1.5*X^2 + 0.5*X^3 + sin(2*X) + 3*W*(X>=0) + rnorm(n)
rdbwhte.1 = rdbwhte(y=Y, x=X, covs.hte=factor(W))
summary(rdbwhte.1)
```

---

rdhte *RD Heterogeneous Treatment Effects Estimation and Inference*

Description

rdhte provides estimation and inference for heterogeneous treatment effects in RD designs using local polynomial regressions, allowing for interactions with pretreatment covariates (Calonico, Cattaneo, Farrell, Palomba and Titiunik, 2025a). Inference is implemented using robust bias-correction methods (Calonico, Cattaneo, and Titiunik, 2014)

Companion commands: rdbwhte for data-driven bandwidth selection and rdhte_lincom for testing linear restrictions of parameters.

Related software packages for analysis and interpretation of RD designs and related methods are available in: https://rdpackages.github.io/.

For background methodology, see Calonico, Cattaneo, Farrell, and Titiunik (2019), Calonico, Cattaneo and Farrell (2020), Cattaneo and Titiunik (2022).

Usage

```
rdhte(
  y,
  x,
  c = 0,
  covs.hte = NULL,
  covs.eff = NULL,
  p = 1,
  q = 2,
```

```
  kernel = "tri",
  weights = NULL,
  h = NULL,
  h.l = NULL,
  h.r = NULL,
  vce = "hc3",
  cluster = NULL,
  level = 95,
  bwselect = NULL,
  bw.joint = FALSE,
  subset = NULL
)
```

Arguments

y	Outcome variable.
x	Running variable.
c	RD cutoff in x; default is c = 0.
covs.hte	covariates for heterogeneous treatment effects. Factor variables can be used to distinguish between continuous and categorical variables, select reference categories, specify interactions between variables, and include polynomials of continuous variables. If not specified, the RD Average Treatment Effect is computed.
covs.eff	additional covariates to be used for efficiency improvements.
p	order of the local polynomial used to construct the point estimator (default = 1).
q	order of the local polynomial used to construct the bias correction (default = 2).
kernel	kernel function used to construct the RD estimators. Options are triangular (default option), epanechnikov and uniform.
weights	variable used for optional weighting of the estimation procedure. The unit-specific weights multiply the kernel function.
h	main bandwidth used to construct the RD estimator. If not specified, bandwidth h is computed by the companion command rdbwhte. More than one bandwidth can be specified for categorical covariates.
h.l	same as h, but only used for observations left of the cutoff c.
h.r	same as h, but only used for observations right of the cutoff c.
vce	character string specifying the variance-covariance matrix estimator type (hc0–hc3) (default = "hc3"). It is based on the R function vcovCL.
cluster	variable indicating the clustering of observations.
level	confidence level for confidence intervals; default is level = 95.
bwselect	bandwidth selection procedure to be used. Options are: mserd one common MSE-optimal bandwidth selector for the RD treatment effect estimator. msetwo two different MSE-optimal bandwidth selectors (below and above the cutoff) for the RD treatment effect estimator. msesum one common MSE-optimal bandwidth selector for the sum of regression estimates (as opposed to difference

thereof). msecomb1 for min(mserd,msesum). msecomb2 for median(msetwo,mserd,msesum), for each side of the cutoff separately. cerrd one common CER-optimal bandwidth selector for the RD treatment effect estimator. certwo two different CER-optimal bandwidth selectors (below and above the cutoff) for the RD treatment effect estimator. cersum one common CER-optimal bandwidth selector for the sum of regression estimates (as opposed to difference thereof). cercomb1 for min(cerrd,cersum). cercomb2 for median(certwo,cerrd,cersum), for each side of the cutoff separately. Note: MSE = Mean Square Error; CER = Coverage Error Rate. Default is bwselect=mserd.

bw.joint — logical. If TRUE, forces all bandwidths to be the same across groups (default is bw.joint = FALSE).

subset — optional vector specifying a subset of observations to be used.

Value

A list with selected RD HTE effects and model information.

Estimate — vector of conventional local-polynomial RD estimates.

Estimate.bc — vector of bias-corrected local-polynomial RD estimates.

se.rb — vector containing robust bias corrected standard errors of the local-polynomial RD estimates.

ci.rb — matrix containing robust bias corrected confidence intervals.

t.rb — vector containing the t-statistics associated with robust local-polynomial RD estimates.

pv.rb — vector containing the p-values associated with robust local-polynomial RD estimates.

coefs — vector containing the coefficients for the jointly estimated p-th order local polynomial model.

vcov — estimated variance-covariance matrix.

W.lev — vector of group level identifiers.

kernel — kernel type used.

vce — variance estimator used.

c — cutoff value.

h — vector containing the bandwidths used.

p — order of the polynomial used for estimation of the regression function.

q — order of the polynomial used for inference on the regression function.

N — vector with the original number of observations for each group.

Nh — vector with the effective number of observations for each group.

covs.cont — internal value.

level — confidence level used.

rdmodel — rd model.

Author(s)

Sebastian Calonico, University of California, Davis <scalonico@ucdavis.edu>.

Matias D. Cattaneo, Princeton University <cattaneo@princeton.edu>.

Max H. Farrell, University of California, Santa Barbara <maxhfarrell@ucsb.edu>.

Filippo Palomba, Princeton University <fpalomba@princeton.edu>.

Rocio Titiunik, Princeton University <titiunik@princeton.edu>.

References

Calonico, Cattaneo, Farrell, Palomba and Titiunik (2025): rdhte: Learning Conditional Average Treatment Effects in RD Designs. *Working paper*.

Calonico, Cattaneo, Farrell, Palomba and Titiunik (2025): Treatment Effect Heterogeneity in Regression Discontinuity Designs. *Working paper*.

Cattaneo, Farrell, and Titiunik. 2022. Regression Discontinuity Designs. *Annual Review of Economics*, 14: 821-851.

Calonico, Cattaneo, and Farrell. 2020. Optimal Bandwidth Choice for Robust Bias Corrected Inference in Regression Discontinuity Designs. *Econometrics Journal*, 23(2): 192-210.

Calonico, Cattaneo, Farrell, and Titiunik. 2019. Regression Discontinuity Designs using Covariates. *Review of Economics and Statistics*, 101(3): 442-451.

Calonico, Cattaneo, and Titiunik. 2014a. Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica* 82(6): 2295-2326.

See Also

rdbwhte, rdhte_lincom

Examples

```
set.seed(123)
n <- 1000
X <- runif(n, -1, 1)
W <- rbinom(n, 1, 0.5)
Y <- 3 + 2*X + 1.5*X^2 + 0.5*X^3 + sin(2*X) + 3*W*(X>=0) + rnorm(n)
m1 = rdhte(y = Y, x = X, covs.hte = factor(W))
summary(m1)
```

---

rdhte_lincom *RD Heterogeneous Treatment Effects. Linear combinations of parameters*

Description

rdhte_lincom computes point estimates, p-values, and robust bias-corrected confidence intervals for linear combinations of parameters after any estimation using rdhte (Calonico, Cattaneo, Farrell, Palomba and Titiunik, 2025a). Inference is implemented using robust bias-correction methods (Calonico, Cattaneo, and Titiunik, 2014). It is based on the R function glht.

Companion commands: rdhte for estimation and inference of RD-HTE and rdbwhte for data-driven bandwidth selection.

Related software packages for analysis and interpretation of RD designs and related methods are available in: https://rdpackages.github.io/.

For background methodology, see Calonico, Cattaneo, Farrell, and Titiunik (2019), Calonico, Cattaneo and Farrell (2020), Cattaneo and Titiunik (2022).

Usage

```
rdhte_lincom(model, linfct, level = 95, digits = 3)
```

Arguments

model	a fitted model returned by rdhte.
linfct	a specification of the linear hypotheses to be tested. Linear functions can be specified by either the matrix of coefficients or by symbolic descriptions of one or more linear hypotheses.
level	Confidence level for intervals; default is level = 95.
digits	Number of decimal places to format numeric outputs (default 3).

Value

A list with two data frames: 'individual' and 'joint', with rounded values.

Author(s)

Sebastian Calonico, University of California, Davis <scalonico@ucdavis.edu>.

Matias D. Cattaneo, Princeton University <cattaneo@princeton.edu>.

Max H. Farrell, University of California, Santa Barbara <maxhfarrell@ucsb.edu>.

Filippo Palomba, Princeton University <fpalomba@princeton.edu>.

Rocio Titiunik, Princeton University <titiunik@princeton.edu>.

References

Calonico, Cattaneo, Farrell, Palomba and Titiunik (2025): rdhte: Learning Conditional Average Treatment Effects in RD Designs. *Working paper*.

Calonico, Cattaneo, Farrell, Palomba and Titiunik (2025): Treatment Effect Heterogeneity in Regression Discontinuity Designs. *Working paper*

See Also

rdhte, rdbwhte

Examples

```
set.seed(123)
n <- 1000
X <- runif(n, -1, 1)
W <- rbinom(n, 1, 0.5)
Y <- 3 + 2*X + 1.5*X^2 + 0.5*X^3 + sin(2*X) + 3*W*(X>=0) + rnorm(n)
m1 <- rdhte(y = Y, x = X, covs.hte = factor(W))
linfct <- c("`factor(W)0` - `factor(W)1` = 0")
rdhte_lincom(model = m1, linfct = linfct)
```

Index