Package 'rdrobust'

October 26, 2014

Type Package

Title Robust data-driven statistical inference in Regression-Discontinuity designs

Version 0.70

Date 2014-10-14

Author Sebastian Calonico <scalonico@bus.miami.edu>, Matias D. Cattaneo
<cattaneo@umich.edu>, Rocio Titiunik <titiunik@umich.edu>

Maintainer Sebastian Calonico <scalonico@bus.miami.edu>

Description Regression-discontinuity (RD) designs are quasi-experimental
research designs popular in social, behavioral and natural sciences.
The RD design is usually employed to study the (local) causal effect
of a treatment, intervention or policy. This package provides tools
for data-driven graphical and analytical statistical inference in RD
designs: rdrobust to construct local-polynomial point estimators
and robust confidence intervals for average treatment effects at the
cutoff in Sharp, Fuzzy and Kink RD settings, rdbwselect to perform
bandwidth selection for the different procedures implemented, and
rdplot to conduct exploratory data analysis (RD plots).

Depends R (>= 3.0.2)

License GPL-2

NeedsCompilation no

Repository CRAN

Date/Publication 2014-10-26 07:23:08

R topics documented:

---

  rdrobust-package   *Robust data-driven statistical inference in RD designs*

Description

Regression-discontinuity (RD) designs are quasi-experimental research designs popular in social, behavioral and natural sciences. The RD design is usually employed to study the (local) causal effect of a treatment, intervention or policy. This package provides tools for data-driven graphical and analytical statistical inference in RD designs: rdrobust to construct local-polynomial point estimators and robust confidence intervals for average treatment effects at the cutoff in Sharp, Fuzzy and Kink RD settings, rdbwselect to perform bandwidth selection for the different procedures implemented, and rdplot to conduct exploratory data analysis (RD plots).

Details

Package:	rdrobust
Type:	Package
Version:	0.70
Date:	2014-10-14
License:	GPL-2

Function for statistical inference: rdrobust
Function for bandwidths selection: rdbwselect
Function for exploratory data analysis (RD plots): rdplot

Author(s)

Sebastian Calonico, University of Miami, Coral Gables, FL. <scalonico@bus.miami.edu>.

Matias D. Cattaneo, University of Michigan, Ann Arbor, MI. <cattaneo@umich.edu>.

Rocio Titiunik, University of Michigan, Ann Arbor, MI. <titiunik@umich.edu>.

---

  rdbinselect   *Data-driven optimal length selector of evenly-spaced and quantile-based bins employed to approximate the underlying regression functions in RD estimation*

Description

NOTE: this command is no longer mantained, and is included only for back compatibility. Please use the new command rdplot.

rdbinselect employs the results in Calonico, Cattaneo and Titiunik (2014d) to implement several novel data-driven optimal choices of evenly-spaced and quantile-based bins used to approximate the underlying regression functions by local sample means, which is then used to construct the RD plots usually found in empirical applications.

Calonico, Cattaneo and Titiunik (2014b) provides an introduction to this command. For a review on RD methods see Imbens and Lemieux (2008), Lee and Lemieux (2010), Dinardo and Lee (2011), Calonico, Cattaneo and Titiunik (2014a), and references therein.

A companion Stata package is described in Calonico, Cattaneo and Titiunik (2014c).

Usage

```
rdbinselect(y, x, data, subset = NULL, c = 0, p = 4,
          numbinl = NULL, numbinr = NULL, binselect = "es",
          lowerend = NULL, upperend = NULL, scale = 1, hide = FALSE,
          par=NULL, title = NULL, x.label = NULL, y.label = NULL,
          x.lim = NULL, y.lim = NULL, model = FALSE, frame = FALSE)
```

Arguments

y	is the dependent variable.
x	is the running variable (a.k.a. score or forcing variable).
data	an optional data frame.
subset	an optional vector specifying a subset of observations to be used.
c	specifies the RD cutoff in x; default is c = 0.
p	specifies the order of the global-polynomial used to approximate the population conditional mean functions for control and treated units; default is p = 4.
numbinl	specifies the number of bins used to the left of the cutoff, denoted J_-. If not specified, J_- is estimated using the method and options chosen below.
numbinr	specifies the number of bins used to the right of the cutoff, denoted J_+. If not specified, J_+ is estimated using the method and options chosen below.
binselect	specifies the partition length selection procedure to be used. This option is available only if J_- and J_+ are not set manually. Options are:
	es: evenly-spaced method using spacings estimators; default option.
	espr: evenly-spaced method with polynomial regression.
	esdw: density-weighted evenly-spaced method.
	qs: quantile-spaced method with spacings estimators.
	qspr: quantile-spaced method using polynomial regression.
	qsdw: density-weighted quantile-spaced method.
lowerend	specifies the lower bound for x to the left of the cutoff. Default is the minimum value in sample.

upperend	specifies the upper bound for x to the right of the cutoff. Default is the maximum value in sample.
scale	specifies a multiplicative factor to be used with the optimal numbers of bins selected. Specifically, the number of bins used for the treatment and control groups will be $\text{scale} \times \hat{J}_+$ and $\text{scale} \times \hat{J}_-$, where $\hat{J}$. denotes the estimated optimal numbers of bins originally computed for each group; default is scale = 1.
hide	logical. If TRUE, it omits the RD plot; default is hide = FALSE.
par	optional graphical parameters.
title	optional title for the RD plot.
x.label	optional label for the x-axis of the RD plot.
y.label	optional label for the y-axis of the RD plot.
x.lim	optional setting for the range of the x-axis in the RD plot.
y.lim	optional setting for the range of the x-axis in the RD plot.
frame	logical. If TRUE, the data frame used in model fitting will be returned.
model	logical. If TRUE, the model object will be returned.

Value

method	method used to compute the optimal number of bins.
results	matrix containing the main estimation results (sample size, polynomial order, number of bins, scales and bin lengths).
coef	matrix containing the coefficients of the pth order global polynomial estimated both sides of the cutoff.

Author(s)

Sebastian Calonico, University of Miami, Coral Gables, FL. <scalonico@bus.miami.edu>.

Matias D. Cattaneo, University of Michigan, Ann Arbor, MI. <cattaneo@umich.edu>.

Rocio Titiunik, University of Michigan, Ann Arbor, MI. <titiunik@umich.edu>.

References

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014a. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica*, forthcoming. http://www-personal.umich.edu/~cattaneo/papers/Calonico-Cattaneo-Titiunik_2014_ECMA.pdf.

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014b. "rdrobust: An R Package for Robust Inference in Regression-Discontinuity Designs." University of Michigan, Department of Economics. http://www-personal.umich.edu/~cattaneo/papers/Calonico-Cattaneo-Titiunik_2014_Rpkg.pdf.

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014c. "Robust Data-Driven Inference in the Regression-Discontinuity Design." University of Michigan, Department of Economics. http://www-personal.umich.edu/~cattaneo/papers/Calonico-Cattaneo-Titiunik_2014_Stata.pdf.

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014d. "Optimal Data-Driven Regression Discontinuity Plots." University of Michigan, Department of Economics. http://www-personal.umich.edu/~cattaneo/papers/RD-rdplot.pdf.

Dinardo, J., and D. S. Lee. 2011. "Program Evaluation and Research Designs." In *Handbook of Labor Economics*, ed. O. Ashenfelter and D. Card, vol. 4A, 463-536. Elsevier Science B.V.

Imbens, G., and T. Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142(2): 615-635.

Lee, D. S., and T. Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48(2): 281-355.

See Also

[rdbwselect](), [rdrobust]()

Examples

```
x<-runif(1000,-1,1)
y<-5+3*x+2*(x>=0)+rnorm(1000)
rdbinselect(y,x)
```

---

rdbwselect   *Bandwidth selection procedures for local-polynomial RD estimators*

Description

[rdbwselect]() implements bandwidth selection procedures for local-polynomial RD estimators proposed in Calonico, Cattaneo and Titiunik (2014a). It also computes bandwith selection procedures proposed in Ludwig and Miller (2007) and Imbens and Kalyanaraman (2012).

Calonico, Cattaneo and Titiunik (2014b) provides an introduction to this command. Additional details for conventional approaches to conduct inference in the RD design can be found in Imbens and Lemieux (2008), Lee and Lemieux (2010), Dinardo and Lee (2011), and references therein.

For data-driven RD plots use the command [rdplot](), which implements the results in Calonico, Cattaneo and Titiunik (2014d).

A companion Stata package is described in Calonico, Cattaneo and Titiunik (2014c).

Usage

```
rdbwselect(y, x, data, subset = NULL, c = 0, p = 1, q = 2, deriv = 0,
              rho = NULL, kernel = "tri", bwselect = "CCT", scaleregul = 1,
              delta = 0.5, cvgrid_min = NULL, cvgrid_max = NULL,
              cvgrid_length = NULL, cvplot = FALSE, vce = "nn", matches = 3,
              all = FALSE, precalc = TRUE, model = FALSE, frame = FALSE)
```

Arguments

y	is the dependent variable.
x	is the running variable (a.k.a. score or forcing variable).
data	an optional data frame.

subset	an optional vector specifying a subset of observations to be used.
c	specifies the RD cutoff in x; default is c = 0.
p	specifies the order of the local-polynomial used to construct the point-estimator; default is p = 1 (local linear regression).
q	specifies the order of the local-polynomial used to construct the bias-correction; default is q = 2 (local quadratic regression).
deriv	specifies the order of the derivative of the regression function to be estimated; default is deriv = 0 (Sharp RD, or Fuzzy RD if fuzzy is also specified). Setting it equal to 1 results in estimation of a Kink RD design (or Fuzzy Kink RD if fuzzy is also specified).
rho	if specified, sets the pilot bandwidth b equal to h/rho, where h is computed using the method and options chosen below.
kernel	is the kernel function used to construct the local-polynomial estimator(s). Options are triangular (default option), epanechnikov and uniform.
bwselect	selects the bandwidth selection procedure to be used. By default it computes both h and b, unless rho is specified, in which case it only computes h and sets b = h/rho. Options are:
	CCT for bandwidth selector proposed by Calonico, Cattaneo and Titiunik (2014a) (default option).
	IK for bandwidth selector proposed by Imbens and Kalyanaraman (2012) (only available for Sharp RD design).
	CV for cross-validation method proposded by Ludwig and Miller (2007) (only available for Sharp RD design).
scaleregul	specifies scaling factor for the regularization terms of CCT and IK bandwidth selectors. Setting scaleregul = 0 removes the regularization term from the bandwidth selectors; default is scaleregul = 1.
delta	sets the quantile that defines the sample used in the cross-validation procedure. This option is used only if bwselect(CV) is specified; default is delta = 0.5, that is, the median of the control and treated samples.
cvgrid_min	sets the minimum value of the bandwidth grid used in the cross-validation procedure. This option is used only if bwselect = "CV" is specified.
cvgrid_max	sets the maximum value of the bandwidth grid used in the cross-validation procedure. This option is used only if bwselect = "CV" is specified.
cvgrid_length	sets the bin length of the (evenly-spaced) bandwidth grid used in the cross-validation procedure. This option is used only if bwselect = "CV" is specified.
cvplot	generates a graph of the CV objective function. This option is used only if bwselect = "CV" is specified.
vce	specifies the procedure used to compute the variance-covariance matrix estimator. This option is used only if CCT or IK bankdwith procedures are employed. Options are:
	nn for nearest-neighbor matches residuals using matches number of matches. This is the default option (with matches = 3, see below).
	resid for estimated plug-in residuals using h bandwidth.

matches	specifies the number of matches in the nearest-neighbor based variance-covariance matrix estimator. This options is used only when nearest-neighbor matches residuals are employed; default is `matches = 3`.
all	if specified, [rdbwselect](#) reports three different procedures:
	CCT for bandwidth selector proposed by Calonico, Cattaneo and Titiunik (2014a).
	IK for bandwidth selector proposed by Imbens and Kalyanaraman (2012).
	CV for cross-validation method proposed by Ludwig and Miller (2007).
model	logical. If `TRUE`, the model object will be returned.
frame	logical. If `TRUE`, the data frame used in model fitting will be returned.
precalc	internal option.

Value

bws	matrix containing the estimated bandwidths for each selected procedure.
bwselect	bandwidth selection procedure employed.
kernel	kernel function used to construct the local-polynomial estimator(s).
p	order of the local-polynomial used to construct the point-estimator.
q	order of the local-polynomial used to construct the bias-correction estimator.

Author(s)

Sebastian Calonico, University of Miami, Coral Gables, FL. <scalonico@bus.miami.edu>.

Matias D. Cattaneo, University of Michigan, Ann Arbor, MI. <cattaneo@umich.edu>.

Rocio Titiunik, University of Michigan, Ann Arbor, MI. <titiunik@umich.edu>.

References

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014a. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica*, forthcoming. http://www-personal.umich.edu/~cattaneo/papers/Calonico-Cattaneo-Titiunik_2014_ECMA.pdf.

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014b. "rdrobust: An R Package for Robust Inference in Regression-Discontinuity Designs." University of Michigan, Department of Economics. http://www-personal.umich.edu/~cattaneo/papers/Calonico-Cattaneo-Titiunik_2014_Rpkg.pdf.

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014c. "Robust Data-Driven Inference in the Regression-Discontinuity Design." University of Michigan, Department of Economics. http://www-personal.umich.edu/~cattaneo/papers/Calonico-Cattaneo-Titiunik_2014_Stata.pdf.

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014d. "Optimal Data-Driven Regression Discontinuity Plots." University of Michigan, Department of Economics. http://www-personal.umich.edu/~cattaneo/papers/RD-rdplot.pdf.

Dinardo, J., and D. S. Lee. 2011. "Program Evaluation and Research Designs." In *Handbook of Labor Economics*, ed. O. Ashenfelter and D. Card, vol. 4A, 463-536. Elsevier Science B.V.

Imbens, G., and T. Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142(2): 615-635.

Imbens, G. W., and K. Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79(3): 933-959.

Lee, D. S., and T. Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48(2): 281-355.

Ludwig, J., and D. L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *Quarterly Journal of Economics* 122(1): 159-208.

See Also

[rdrobust](), [rdplot]()

Examples

```
x<-runif(1000,-1,1)
y<-5+3*x+2*(x>=0)+rnorm(1000)
rdbwselect(y,x)
```

---

rdplot   *Data-driven Regression Discontinuity Plots*

Description

[rdplot]() implements several data-driven RD plots, using either evenly-spaced or quantile-spaced partitioning. Two type of RD plots are constructed: (i) RD plots with binned sample means tracing out the underlying regression function (IMSE-optimal selectors), and (ii) RD plots with binned sample means mimicking the underlying variability of the data. For all technical and methodological details see Calonico, Cattaneo and Titiunik (2014d).

Calonico, Cattaneo and Titiunik (2014b) provides an introduction to this command. For a review on RD methods see Imbens and Lemieux (2008), Lee and Lemieux (2010), Dinardo and Lee (2011), Calonico, Cattaneo and Titiunik (2014a), and references therein.

For local polynomial inference methods see the commands [rdrobust]() and [rdbwselect]().

A companion Stata package is described in Calonico, Cattaneo and Titiunik (2014c).

Usage

```
rdplot(y, x, data, subset = NULL, c = 0, p = 4,
          numbinl = NULL, numbinr = NULL, binselect = "esmv",
          lowerend = NULL, upperend = NULL, scale = 1, scalel = 1, scaler = 1,
          hide = FALSE, par=NULL, title = NULL, x.label = NULL, y.label = NULL,
       x.lim = NULL, y.lim = NULL, col.dots = NULL, col.lines = NULL, type.dots = NULL,
          model = FALSE, frame = FALSE)
```

Arguments

y	is the dependent variable.
x	is the running variable (a.k.a. score or forcing variable).
data	an optional data frame.
subset	an optional vector specifying a subset of observations to be used.
c	specifies the RD cutoff in x; default is c = 0.
p	specifies the order of the global-polynomial used to approximate the population conditional mean functions for control and treated units; default is p = 4.
numbinl	specifies the number of bins used to the left of the cutoff, denoted J_-. If not specified, J_- is estimated using the method and options chosen below.
numbinr	specifies the number of bins used to the right of the cutoff, denoted J_+. If not specified, J_+ is estimated using the method and options chosen below.
binselect	specifies the procedure to select the number of bins. This option is available only if J_- and J_+ are not set manually. Options are:
	es: IMSE-optimal evenly-spaced method using spacings estimators.
	espr: IMSE-optimal evenly-spaced method using polynomial regression.
	esmv: mimicking variance evenly-spaced method using spacings estimators. This is the default option.
	esmvpr: mimicking variance evenly-spaced method using polynomial regression.
	qs: IMSE-optimal quantile-spaced method using spacings estimators.
	qspr: IMSE-optimal quantile-spaced method using polynomial regression.
	qsmv: mimicking variance quantile-spaced method using spacings estimators.
	qsmvpr: mimicking variance quantile-spaced method using polynomial regression.
lowerend	specifies the lower bound for x to the left of the cutoff. Default is the minimum value in sample.
upperend	specifies the upper bound for x to the right of the cutoff. Default is the maximum value in sample.
scale	specifies a multiplicative factor to be used with the optimal numbers of bins selected. Specifically, the number of bins used for the treatment and control groups will be scale$\times \hat{J}_+$ and scale$\times \hat{J}_-$, where $\hat{J}.$ denotes the estimated optimal numbers of bins originally computed for each group; default is scale = 1.
scalel	same as scale but only for the data to the left of the cutoff.
scaler	same as scale but only for the data to the right of the cutoff.
hide	logical. If TRUE, it omits the RD plot; default is hide = FALSE.
par	optional graphical parameters.
title	optional title for the RD plot.
x.label	optional label for the x-axis of the RD plot.
y.label	optional label for the y-axis of the RD plot.
x.lim	optional setting for the range of the x-axis in the RD plot.

`y.lim`	optional setting for the range of the y-axis in the RD plot.
`col.dots`	optional setting for the color of the dots in the RD plot.
`col.lines`	optional setting for the color of the lines in the RD plot.
`type.dots`	optional setting for the type of lines in the RD plot.
`frame`	logical. If TRUE, the data frame used in model fitting will be returned.
`model`	logical. If TRUE, the model object will be returned.

Value

`method`	method used to compute the optimal number of bins.
`results`	matrix containing the main estimation results (sample size, polynomial order, number of bins, scales and bin lengths).
`coef`	matrix containing the coefficients of the pth order global polynomial estimated both sides of the cutoff.

Author(s)

Sebastian Calonico, University of Miami, Coral Gables, FL. <scalonico@bus.miami.edu>.

Matias D. Cattaneo, University of Michigan, Ann Arbor, MI. <cattaneo@umich.edu>.

Rocio Titiunik, University of Michigan, Ann Arbor, MI. <titiunik@umich.edu>.

References

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014a. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica*, forthcoming. http://www-personal.umich.edu/~cattaneo/papers/Calonico-Cattaneo-Titiunik_2014_ECMA.pdf.

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014b. "rdrobust: An R Package for Robust Inference in Regression-Discontinuity Designs." University of Michigan, Department of Economics. http://www-personal.umich.edu/~cattaneo/papers/Calonico-Cattaneo-Titiunik_2014_Rpkg.pdf.

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014c. "Robust Data-Driven Inference in the Regression-Discontinuity Design." University of Michigan, Department of Economics. http://www-personal.umich.edu/~cattaneo/papers/Calonico-Cattaneo-Titiunik_2014_Stata.pdf.

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014d. "Optimal Data-Driven Regression Discontinuity Plots." University of Michigan, Department of Economics. http://www-personal.umich.edu/~cattaneo/papers/RD-rdplot.pdf.

Dinardo, J., and D. S. Lee. 2011. "Program Evaluation and Research Designs." In *Handbook of Labor Economics*, ed. O. Ashenfelter and D. Card, vol. 4A, 463-536. Elsevier Science B.V.

Imbens, G., and T. Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142(2): 615-635.

Lee, D. S., and T. Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48(2): 281-355.

See Also

[rdbwselect](), [rdrobust]()

Examples

```
x<-runif(1000,-1,1)
y<-5+3*x+2*(x>=0)+rnorm(1000)
rdplot(y,x)
```

| rdrobust | *Local-polynomial RD estimation with robust confidence intervals* |

Description

[rdrobust]() implements a local-polynomial RD estimator with robust confidence intervals as proposed in Calonico, Cattaneo and Titiunik (2014a). It also computes alternative procedures available in the literature.

Calonico, Cattaneo and Titiunik (2014b) provides an introduction to this command. Additional details for conventional approaches to conduct inference in the RD design can be found in Imbens and Lemieux (2008), Lee and Lemieux (2010), Dinardo and Lee (2011), and references therein.

For data-driven RD plots use the command [rdplot](), which implements the results in Calonico, Cattaneo and Titiunik (2014d).

A companion `Stata` package is described in Calonico, Cattaneo and Titiunik (2014c).

Usage

```
rdrobust(y, x, data, subset = NULL, c = 0,  p = 1, q = 2, deriv = 0,
            fuzzy = NULL, h = NULL, b = NULL, rho = NULL, scalepar = 1,
            kernel = "tri", bwselect = "CCT",  scaleregul = 1, delta = 0.5,
            cvgrid_min = NULL, cvgrid_max = NULL, cvgrid_length = NULL,
            cvplot = FALSE, vce = "nn", matches = 3, level = 95,
            all = FALSE, model = FALSE, frame = FALSE)
```

Arguments

y	is the dependent variable.
x	is the running variable (a.k.a. score or forcing variable).
data	an optional data frame.
subset	an optional vector specifying a subset of observations to be used.
c	specifies the RD cutoff in x; default is c = 0.
p	specifies the order of the local-polynomial used to construct the point-estimator; default is p = 1 (local linear regression).
q	specifies the order of the local-polynomial used to construct the bias-correction; default is q = 2 (local quadratic regression).

deriv	specifies the order of the derivative of the regression function to be estimated; default is deriv = 0 (Sharp RD, or Fuzzy RD if fuzzy is also specified). Setting it equal to 1 results in estimation of a Kink RD design (or Fuzzy Kink RD if fuzzy is also specified).
fuzzy	specifies the treatment status variable used to implement Fuzzy RD estimation (or Fuzzy Kink RD if deriv = 1 is also specified). Default is Sharp RD design and hence this option is not used. *Note*: For fuzzy RD designs, bandwidths are estimated using sharp RD bandwidth selectors for the reduced form outcome equation.
h	directly sets the main bandwidth. If not specified, it is computed by the companion command [rdbwselect]().
b	directly sets the pilot bandwidth. If not specified, it is computed by the companion command [rdbwselect]().
rho	sets the value of rho, so that the pilot bandwidth b equals h/rho. Default is rho = 1 if h is specified but b is not.
scalepar	specifies scaling factor for RD parameter of interest. This option is useful when the population parameter of interest involves a known multiplicative factor (e.g., sharp kink RD). Default is scalepar = 1 (no scaling).
kernel	is the kernel function used to construct the local-polynomial estimator(s). Options are triangular (default option), epanechnikov and uniform.
bwselect	selects the bandwidth selection procedure to be used. By default it computes both h and b, unless rho is specified, in which case it only computes h and sets b = h/rho. Options are:
	CCT for bandwidth selector proposed by Calonico, Cattaneo and Titiunik (2014a) (default option).
	IK for bandwidth selector proposed by Imbens and Kalyanaraman (2012) (only available for Sharp RD design).
	CV for cross-validation method proposded by Ludwig and Miller (2007) (only available for Sharp RD design).
scaleregul	specifies scaling factor for the regularization terms of CCT and IK bandwidth selectors. Setting scaleregul = 0 removes the regularization term from the bandwidth selectors; default is scaleregul = 1.
delta	sets the quantile that defines the sample used in the cross-validation procedure. This option is used only if bwselect(CV) is specified. See companion command [rdbwselect]() for more details; default is delta = 0.5, that is, the median of the control and treated samples.
cvgrid_min	sets the minimum value of the bandwidth grid used in the cross-validation procedure. This option is used only if bwselect = "CV" is specified. See companion command [rdbwselect]() for more details.
cvgrid_max	sets the maximum value of the bandwidth grid used in the cross-validation procedure. This option is used only if bwselect = "CV" is specified. See companion command [rdbwselect]() for more details.
cvgrid_length	sets the bin length of the (evenly-spaced) bandwidth grid used in the cross-validation procedure. This option is used only if bwselect = "CV" is specified. See companion command [rdbwselect]() for more details.

cvplot	generates a graph of the CV objective function. This option is used only if bwselect = "CV" is specified. See companion command [rdbwselect](#) for more details.
vce	specifies the procedure used to compute the variance-covariance matrix estimator. This option is used only if CCT or IK bankdwith procedures are employed. Options are:
	nn for nearest-neighbor matches residuals using matches number of matches. This is the default option (with matches = 3, see below).
	resid for estimated plug-in residuals using h bandwidth.
matches	specifies the number of matches in the nearest-neighbor based variance-covariance matrix estimator. This options is used only when nearest-neighbor matches residuals are employed; default is matches = 3.
level	sets the confidence level for confidence intervals; default is level = 95.
all	if specified, [rdrobust](#) reports three different procedures:
	(i) conventional RD estimates with conventional standard errors.
	(ii) bias-corrected estimates with conventional standard errors.
	(iii) bias-corrected estimates with robust standard errors.
model	logical. If TRUE, the model object will be returned.
frame	logical. If TRUE, the data frame used in model fitting will be returned.

Value

c	cutoff value.
N_l	sample size used to the left of the cutoff.
N_r	sample size used to the right of the cutoff.
N	overall sample size.
p	order of the polynomial used for estimation of the regression function.
q	order of the polynomial used for estimation of the bias of the regression function.
h	bandwidth used for estimation of the regression function.
b	bandwidth used for estimation of the bias of the regression function estimator.
coef	vector containing conventional and bias-corrected local-polynomial RD estimates.
se	vector containing conventional and robust standard errors of the local-polynomial RD estimates.
pv	vector containing the p-values associated with conventional, bias-corrected and robust local-polynomial RD estimates.
ci	matrix containing the confidence intervals associated with conventional, bias-corrected and robust local-polynomial RD estimates.

Author(s)

Sebastian Calonico, University of Miami, Coral Gables, FL. <scalonico@bus.miami.edu>.

Matias D. Cattaneo, University of Michigan, Ann Arbor, MI. <cattaneo@umich.edu>.

Rocio Titiunik, University of Michigan, Ann Arbor, MI. <titiunik@umich.edu>.

References

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014a. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica*, forthcoming. http://www-personal. umich.edu/~cattaneo/papers/Calonico-Cattaneo-Titiunik_2014_ECMA.pdf.

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014b. "rdrobust: An R Package for Robust Inference in Regression-Discontinuity Designs." University of Michigan, Department of Economics. http://www-personal.umich.edu/~cattaneo/papers/Calonico-Cattaneo-Titiunik_2014_ Rpkg.pdf.

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014c. "Robust Data-Driven Inference in the Regression-Discontinuity Design." University of Michigan, Department of Economics. http:// www-personal.umich.edu/~cattaneo/papers/Calonico-Cattaneo-Titiunik_2014_Stata.pdf.

Calonico, S., Cattaneo, M. D., and R. Titiunik. 2014d. "Optimal Data-Driven Regression Discontinuity Plots." University of Michigan, Department of Economics. http://www-personal.umich. edu/~cattaneo/papers/RD-rdplot.pdf.

Dinardo, J., and D. S. Lee. 2011. "Program Evaluation and Research Designs." In *Handbook of Labor Economics*, ed. O. Ashenfelter and D. Card, vol. 4A, 463-536. Elsevier Science B.V.

Imbens, G., and T. Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142(2): 615-635.

Imbens, G. W., and K. Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79(3): 933-959.

Lee, D. S., and T. Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48(2): 281-355.

Ludwig, J., and D. L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *Quarterly Journal of Economics* 122(1): 159-208.

See Also

rdbwselect, rdplot

Examples

```
x<-runif(1000,-1,1)
y<-5+3*x+2*(x>=0)+rnorm(1000)
rdrobust(y,x)
```

---

rdrobust-internal   *Package internal functions*

Description

Internal functions which are not part of the package application programming interface.

Author(s)

Sebastian Calonico, University of Miami, Coral Gables, FL. <scalonico@bus.miami.edu>.

Matias D. Cattaneo, University of Michigan, Ann Arbor, MI. <cattaneo@umich.edu>.

Rocio Titiunik, University of Michigan, Ann Arbor, MI. <titiunik@umich.edu>.

---

rdrobust_RDsenate   *RD Senate Data*

Description

Extract of the dataset constructed by Cattaneo, Frandsen, and Titiunik (2014), which include measures of incumbency advantage in the U.S. Senate for the period 1914-2010.

Usage

```
data(rdrobust_RDsenate)
```

Format

A data frame with 1390 observations on the following 2 variables.

margin a numeric vector.

vote a numeric vector.

Source

Cattaneo, M. D., Frandsen, B., and R. Titiunik. 2014. "Randomization Inference in the Regression Discontinuity Design: An Application to the Study of Party Advantages in the U.S. Senate." *Journal of Causal Inference*, forthcoming. http://www-personal.umich.edu/~cattaneo/papers/Cattaneo-Frandsen-Titiunik_2014_JCI.pdf.

References

Cattaneo, M. D., Frandsen, B., and R. Titiunik. 2014. "Randomization Inference in the Regression Discontinuity Design: An Application to the Study of Party Advantages in the U.S. Senate." *Journal of Causal Inference*, forthcoming. http://www-personal.umich.edu/~cattaneo/papers/Cattaneo-Frandsen-Titiunik_2014_JCI.pdf.

Index