```
  ___  ____  ____  ____  ____(R)
 /__    /   ____/   /   ____/
___/   /   /___/   /   /___/
  Statistics/Data analysis
```

Title

rdmse — Mean Squared Error Estimation for Local Polynomial Regression Discontinuity and Regression Kink Estimators.

Syntax

rdmse *depvar* *runvar* [*if*] [*in*] [**,** **c(**#**)** **p(**#**)** **deriv(**#**)** **fuzzy(***fuzzyvar***)** **kernel(***kernelfn***)** **h(**#**)** **b(**#**)** **scalepar(**#**)** **twosided** **pl(**#**)** **pr(**#**)** **hl(**#**)** **hr(**#**)** **bl(**#**)** **br(**#**)**]

rdmse_cct2014 *depvar* *runvar* [*if*] [*in*] [**,** **c(**#**)** **p(**#**)** **deriv(**#**)** **fuzzy(***fuzzyvar***)** **kernel(***kernelfn***)** **h(**#**)** **b(**#**)** **scalepar(**#**)**]

Description

rdmse computes the (asymptotic) mean squared error (MSE) of a local polynomial RD/RK estimator as proposed in Card, Lee, Pei, Weber (2018, 2020). It displays and returns the estimated MSE for the conventional estimator and its bias corrected counterpart as defined in Calonico, Cattaneo, Titiunik (2014a).

rdmse_cct2014 computes the (A)MSE for a conventional RD/RK estimator by gathering the relevant quantities calculated by the 2014 implementation of **rdrobust**, **rdrobust_2014** by Calonico, Cattaneo and Titiunik. It does not estimate the (A)MSE for the bias corrected estimator because some of the quantities required for the calculation are not computed by **rdrobust_2014** (nor **rdrobust**). For the conventional estimator, **rdmse_cct2014** and **rdmse** implement variance estimation slightly differently. Both commands employ a nearest neighbor estimator and set the number of neighbors to three. However, in the event of a tie **rdmse_cct2014** selects all of the closest neighbors following **rdrobust_2014**. In contrast, **rdmse** randomly selects three neighbors and speeds up the computation in doing so.

Options

c(#**)** specifies the RD cutoff in *runvar*. Default is **c(0)**.

p(#**)** specifies the order of the local polynomial. Default is **p(1)** (local linear regression). Consistent with the implementation in **rdrobust**, the maximum value allowed for **p()** is 8. A local polynomial of order (*p*+1) is used to estimate the bias of the estimator.

deriv(#**)** specifies the order of the derivative of the regression functions to be estimated. Default is **deriv(0)** (RD estimator). Use **deriv(1)** for an RK estimator.

fuzzy(*fuzzyvar***)** specifies the treatment variable in a fuzzy RD/RK design. Leave the option unspecified if the underlying design is sharp.

kernel(*kernelfn***)** specifies the kernel function used to construct the local polynomial estimator. Options are **triangular** or **uniform**.

h(#**)** specifies the main bandwidth used to construct the RD/RK estimator. The user has to specify this bandwidth.

b(#**)** specifies the bias bandwidth for estimating the bias of the RD/RK estimator. The user has to specify this bandwidth.

scalepar(#) specifies a scaling factor for the RD/RK parameter of interest. The same option is available in **rdrobust** as per Calonico, Cattaneo, Titiunik (2014b). Default is **scalepar(1)**.

twosided. If specified, the program looks for separate polynomial orders and bandwidths on two sides of the threshold, which need to be specified in **pl()**, **pr()**, **hl()**, **hr()**, **bl()**, and **br()**. The prog rns the estimated mean squared error for the conventional and bias-corrected estimator of the left and right derivatives of order *deriv*, respectively. The two-sided bandwidths can be obtained by specifing the **bwselect(***msetwo*{cmd:) in **rdrobust.** The twosided option can only be used in a sharp RD/RK design (more in Additional Notes below). See Calonico, Cattaneo, Farrell, Titiunik (2017, 2019) for details.

pl(#) and **pr(#)** specify the orders of the local polynomials on the left and right sides of the threshold, respectively. Default is **pl(1)** and **pr(1)** (local linear regressions). Consistent with th entation in **rdrobust**, the maximum value allowed is 8 for both orders. Local polynomials of order (*pl*+1) and (*pr*+1) are used to estimate the biases of the left- and right-side estimators.

hl(#) and **hr(#)** specify the main bandwidths used to construct the estimators of the left and right derivatives of order *deriv*. The user has to supply these bandwidths if the option **twoside** is sp

bl(#) and **br(#)** specify the bias bandwidths used to estimate the biases of the left- and right-side estimators. The user has to supply these bandwidths if the option **twoside** is specified.

Generic Examples:

Let **Y** be the outcome variable and **x** the running variable:

MSE Estimation for local linear sharp RD estimator with uniform kernel and CCT bandwidths (Calonico, Cattaneo, Titiunik 2014a, 2014b)
First estimate the CCT bandwidths using **altrdbwselect** included in the package

```
. altrdbwselect Y x, c(0) deriv(0) p(1) q(2) kernel(uniform)
    bwselect(CCT)
. local bw_h=e(h_CCT)
. local bw_b=e(b_CCT)
```

Then estimate the MSE by passing the CCT bandwidths as arguments

```
. rdmse Y x, deriv(0) c(0) p(1) h(`bw_h') b(`bw_b') kernel(uniform)
```

Estimate the MSE of the sharp local linear RD estimator with manual bandwidths

```
. rdmse Y x, deriv(0) c(0) p(1) h(0.5) b(1.2) kernel(uniform)
```

Estimate the MSE of the sharp local linear RK estimator

```
. rdmse Y x, deriv(0) c(0) p(1) h(0.5) b(1.2) kernel(uniform)
```

Estimate the MSEs of the left- and right- intercept estimators constructed with different polynomial orders and bandwidths on two sides of the threshold

```
. rdmse Y x, c(0) deriv(0) twosided pl(1) pr(2) hl(0.5) hr(0.45)
    bl(1.2) br(1.1) kernel(uniform)
```

Let **T** be the treatment variable.

MSE Estimation for local linear fuzzy RD estimator with uniform kernel and "fuzzy CCT" bandwidths (Card, Lee, Pei, Weber 2015)

First estimate the fuzzy CCT bandwidths using **altfrdbwselect** included in the package

. altfrdbwselect Y x, c(0) fuzzy(T) deriv(0) p(1) q(2) kernel(uniform) bwselect(CCT)

. local fbw_h=e(h_F_CCT)

. local fbw_b=e(b_F_CCT)

Estimate the MSE by passing the "fuzzy CCT" bandwidths as arguments

. rdmse Y x, c(0) fuzzy(T) deriv(0) p(1) h(`fbw_h') b(`fbw_b') kernel(uniform)

Estimate the MSE of the fuzzy local linear RD estimator with manual bandwidths

. rdmse Y x, c(0) fuzzy(T) deriv(0) p(1) h(0.5) b(1.2) kernel(uniform)

Estimate the MSE of the fuzzy local linear RK estimator

. rdmse Y x, c(0) fuzzy(T) deriv(1) p(1) h(0.5) b(1.2) kernel(uniform)

Saved results

If **fuzzy()** and **twosided** are unspecified, **rdmse** saves the scalars:

e(amse_cl)	estimated (asymptotic) MSE of the conventional sharp estimator
e(amse_bc)	estimated (asymptotic) MSE of the bias-corrected sharp estimator

If **twosided** is specified, **rdmse** saves the scalars:

e(amse_l_cl)	estimated (asymptotic) MSE of the conventional left-side estimator
e(amse_l_bc)	estimated (asymptotic) MSE of the bias-corrected left-side estimator
e(amse_r_cl)	estimated (asymptotic) MSE of the conventional right-side estimator
e(amse_r_bc)	estimated (asymptotic) MSE of the bias-corrected right-side estimator

If **fuzzy()** is specified, **rdmse** saves the scalars:

e(amse_F_cl)	estimated (asymptotic) MSE of the conventional fuzzy estimator
e(amse_F_bc)	estimated (asymptotic) MSE of the bias-corrected fuzzy estimator

Since **rdmse_cct2014** only estimates the (asymptotic) MSE of the conventional estimator, it returns **e(amse_cl)** in the sharp case and **e(amse_F_cl)** in the fuzzy case.

Additional Notes

altrdbwselect is an alternative implementation of the CCT bandwidth selector from Calonico, Cattaneo, Titiunik (2014a). As with **rdmse**, it speeds up the computation in Calonico, Cattaneo, Titiunik (2014b) by adopting a random tie breaking scheme in variance estimation. The syntax is the same as **rdbwselect** in Calonico, Cattaneo, Titiunik (2014b).

In the current implementation of **rdrobust**, the two-sided bandwidths in a fuzzy design are optimal for estimating the left and right derivatives of order *deriv* in the the reduced-form relationship between the outcome variable and ning variable. In this spirit, we do not allow **twosided** to be specified in conjunction with **fuzzy()**, and the user should apply the **twosided** option to the reduced-form only by treating it as a sharp design.

References

Calonico, S., M. D. Cattaneo, and R. Titiunik. 2014a. Robust Nonparametric Confidence Intervals for Regression Discontinuity Designs. *Econometrica* 82(6): 2295-2326. https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA11757.

Calonico, S., M. D. Cattaneo, and R. Titiunik. 2014b. Robust Data Driven Inference in the Regression Discontinuity Design. *Stata Journal* 14(4): 909-946. https://journals.sagepub.com/doi/abs/10.1177/1536867X1401400413.

Card, D., D. S. Lee, Z. Pei, and A. Weber. 2015. Inference on Causal Effects in a Generalized Regression Kink Design. *Econometrica* 83(6): 2453-2483. https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA11224.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik. 2017. **rdrobust**: Software for Regression-Discontinuity Designs. *Stata Journal* 17(2): 372-404. https://journals.sagepub.com/doi/abs/10.1177/1536867X1701700208.

Card, D., D. S. Lee, Z. Pei, and A. Weber. 2018. Princeton University Industrial Relations Section Working Paper #622. https://irs.princeton.edu/publications/working-papers/local-polynomial-> order-regression-discontinuity-designs.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik. 2019. Regression Discontinuity Designs Using Covariates. *Review of Economics and Statistics* 101(3): 442-451. https://www.mitpressjournals.org/doi/abs/10.1162/rest_a_00760.

Card, D., D. S. Lee, Z. Pei, and A. Weber. 2020. NBER Working Paper #622. https://www.nber.org/papers/w27424.

Author

Zhuan Pei, Cornell University, Ithaca, NY. zhuan.pei@cornell.edu.