



Title

rdwinselect — Window selection procedure for RD designs under local randomization.

Syntax

```
rdwinselect runvar [covariates] [if] [in] [, cutoff(#) obsmin(#) wmin(# #) wobs(#)
wstep(#) wsymmetric wmasspoints nwindows(#) dropmissing statistic(stat_name)
p(#) evalat(point) kernel(kerneltype) approximate level(#) reps(#) seed(#)
plot graph_options(graphopts) genvars obsstep(#) ]
```

Description

rdwinselect implements window selection procedure based balance tests for regression discontinuity (RD) designs under local randomization. Specifically, it constructs a sequence of nested windows around the RD cutoff and reports binomial tests for running variable *runvar* and covariate balance tests for covariates *covariates* (if specified). The recommended window is the largest window around the cutoff such that the minimum p-values of the balance tests is larger than a pre-specified level for all nested (smaller) windows. By default, the p-values are calculated employing randomization inference methods. See [Cattaneo, Frandsen and Titiunik \(2015\)](#) and [Cattaneo, Titiunik and Vazquez-Bare \(2017\)](#) for an introduction to this methodology.

A detailed introduction to this command is given in [Cattaneo, Titiunik and Vazquez-Bare \(2016\)](#).

Companion R functions are also available [here](#).

Companion functions are [rdrandinf](#), [rdsensitivity](#) and [rdrbounds](#).

Related Stata and R packages useful for inference in RD designs are described in the following website:

<https://rdpackages.github.io/>

Options

cutoff(#) specifies the RD cutoff for the running variable *runvar*. Default is **cutoff(0)**.

_____ Window selection _____

obsmin(#) specifies the minimum number of observations above and below the cutoff in the smallest window. Default is **obsmin(10)**.

wmin(# #) specifies the initial window to be used (if **obsmin**(#) is not specified). Can be a single number to specify the length of the (symmetric) initial window, or two numbers to specify the left and right limits of the initial window. Specifying both **wmin**(#) and **obsmin**(#) returns an error.

wobs(#) specifies the number of observations to be added at each side of the cutoff at each step. Default is **wobs(5)**.

wstep(#) specifies the increment in window length. Specifying both **wobs**(#) and **wstep**(#) returns an error.

wsymmetric requires that windows be symmetrized around the cutoff (when **wobs**(#) is specified).

wmasspoints specifies that the running variable is discrete and each masspoint should be used as a window.

nwindows(#) specifies the number of windows to be used. Default is **nwindows(10)**.

dropmissing drop rows with missing values in covariates when calculating windows.

Statistic

statistic(*stat_name*) specifies the statistic to be used. Options are:
diffmeans for difference in means statistic. This is the default option.
ksmirnov for Kolmogorov-Smirnov statistic.
ranksum for Wilcoxon-Mann-Whitney studentized statistic.
hotelling for Hotelling's T-squared statistic.
The option **ttest** is equivalent to **diffmeans** and included for backward compatibility.

p(#) specifies the order of the polynomial for outcome adjustment model. Default is **p(0)**.

evalat(*point*) specifies the point at which the adjusted variable is evaluated. Allowed options are **cutoff** and **means**. Default is **evalat(cutoff)**.

kernel(*kerneltype*) specifies the type of kernel to use as weighting scheme. Allowed kernel types are **uniform** (uniform kernel), **triangular** (triangular kernel) and **epan** (Epanechnikov kernel). Default is **kernel(uniform)**.

Inference

approximate specifies that covariate balance tests should use a large-sample approximation instead of finite-sample exact randomization inference methods.

level(#) specifies the minimum accepted value of the p-value from the covariate balance tests to be used. Default is **level(.15)**.

reps(#) specifies the number of replications to be used. Default is **reps(1000)**.

seed(#) sets the seed for the randomization test. With this option, the user can manually set the desired seed, or can enter the value -1 to use the system seed. Default is **seed(666)**.

Generate plots and variables

plot draws a scatter plot of the minimum p-value from the covariate balance test against window length implemented by the command.

graph_options(*graphopts*) graph options for plot generated by the command.

genvars generates a variable indicating the window number corresponding to each observation and a variable indicating the corresponding window length.

Backward compatibility

obsstep(#) specifies the minimum number of observations to be added on each side of the cutoff. This option is deprecated and only included for backward compatibility. We recommend the use of **wstep** or **wobs** instead.

Example: Cattaneo, Frandsen and Titiunik (2015) Incumbency Data

```
Setup
. use rdlocrand_senate.dta
```

```
Window selection with three covariates and default options
. rdwinselect demmv dopen population demvoteslag1
```

```
Window selection using Kolmogorov-Smirnov statistic
. rdwinselect demmv dopen population demvoteslag1, stat(ksmirnov)
```

Window selection with smallest window including at least 10 observations in each group and adding 3 observations in each step

```
. rdwinselect demmv dopen population demvoteshlag1, obsmin(10) wobs(3)
```

Window selection setting smallest window at .5 and with .125 length increments

```
. rdwinselect demmv dopen population demvoteshlag1, wmin(.5) wstep(.125)
```

Window selection with asymptotic p-values using 40 windows with scatter plot

```
. rdwinselect demmv dopen population demvoteshlag1, nwindows(40) approximate plot
```

Modify graph options: add title and x-axis label

```
. rdwinselect demmv dopen population demvoteshlag1, nwindows(40) approx plot graph_options(title(Main title) xtitle(x-axis title))
```

Saved results

rdwinselect saves the following in **r()**:

Scalars

r(minp)	minimum p-value from covariate test
r(N)	sample size in recommended window
r(N_left)	sample size in recommended window to the left of the cutoff
r(N_right)	sample size in recommended window to the right of the cutoff
r(w_left)	left end of recommended window
r(w_right)	right end of recommended window
r(wobs)	when specified, increment (in observations) in each window
r(wmin)	initial window
r(wstep)	when specified, increment (in window length) in each window
r(nwindows)	total number of windows evaluated

Locals

r(seed)	seed used in permutations
----------------	---------------------------

Matrices

r(wlist)	matrix with window lengths
r(results)	stores the minimum p-value from covariate balance test, p-value from binomial test, sample sizes and window length in each window

References

- Cattaneo, M. D., Frandsen, B., and R. Titiunik. 2015. [Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate.](#) *Journal of Causal Inference* 3(1): 1-24.
- Cattaneo, M.D., Titiunik, R. and G. Vazquez-Bare. 2016. [Inference in Regression Discontinuity Designs under Local Randomization.](#) *Stata Journal* 16(2): 331-367.
- Cattaneo, M. D., Titiunik, R. and G. Vazquez-Bare. 2017. [Comparing Inference Approaches for RD Designs: A Reexamination of the Effect of Head Start on Child Mortality.](#) *Journal of Policy Analysis and Management* 36(3): 643-681.

Authors

Matias D. Cattaneo, Princeton University, Princeton, NJ. cattaneo@princeton.edu.

Rocio Titiunik, Princeton University, Princeton, NJ. titiunik@princeton.edu.

Gonzalo Vazquez-Bare, UC Santa Barbara, Santa Barbara, CA. gvazquez@econ.ucsb.edu.