Quant II

Lab 5: Placebos and intro to RDD

Giacomo Lemoli

March 7, 2022

Today's plan

- Placebos
- RDD with `rdrobust`

Housekeeping

- Midterm this Wednesday
- Class time: 10-11.50 in 217
- Papers will be provided

Placebos

This section uses the exposition in Eggers, Tuñón and Dafoe (2021)

- In medical studies a placebo is a "fake treatment" delivered to patients in the control group
- In observational studies, a placebo analysis means testing a relationship that our causal theory suggests to be 0
- Idea: if our theory/assumptions dictate an effect must not be there, finding it means the theory/assumptions are incorrect

Typologies of placebo

From Eggers, Tuñón and Dafoe (2021): placebo analyses change elements of the research design while maintaining the same design

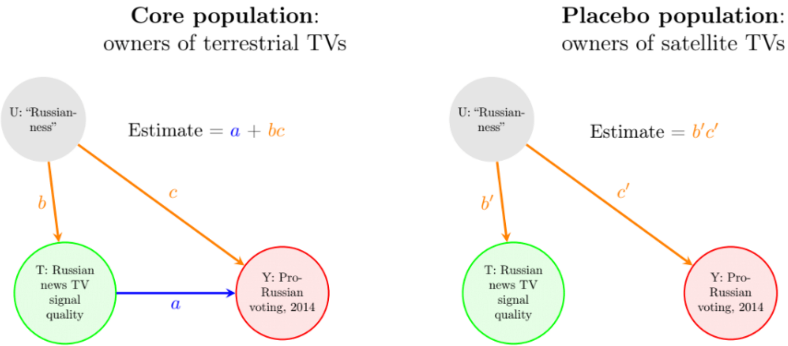
- Placebo population: same design on a different (sub-)population
- Placebo outcome: change the outcome variable
- Placebo treatment: change the treatment variable

Placebos: general concepts

- In a placebo test, only the feature of the design that is a "placebo" has to change
- All the other features of the design must be the same
- Otherwise not a real placebo: we just pick up different things so we can't compare the two results

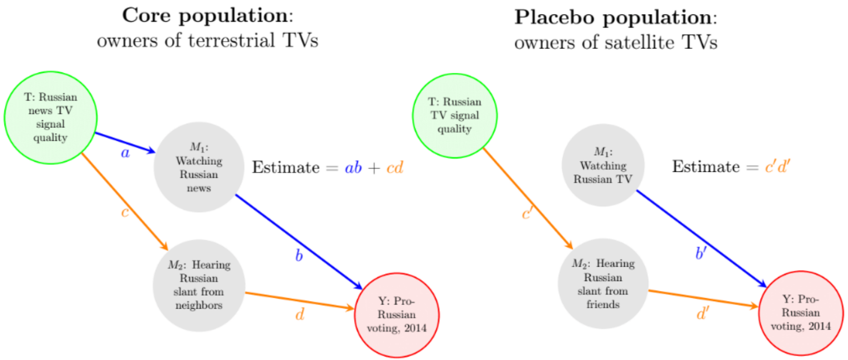
Placebo population: bias

Figure 4: The logic of placebo population tests for confounding bias



Placebo population: mechanisms/exclusion restriction violations

Figure 5: The logic of placebo population tests for alternative mechanisms



Placebo population: other examples

TABLE 7—REDUCED FORM RELATIONSHIP BETWEEN THE DISTANCE FROM THE COAST AND TRUST WITHIN AFRICA AND ASIA

	Trust of local government council			
	Afrobarometer sample		Asiabarometer sample	
	(1)	(2)	(3)	(4)
Distance from the coast	0.00039***	0.00031***	−0.00001	0.00001
	(0.00009)	(0.00008)	(0.00010)	(0.00009)
Country fixed effects	Yes	Yes	Yes	Yes
Individual controls	No	Yes	No	Yes
Number of observations	19,913	19,913	5,409	5,409
Number of clusters	185	185	62	62
R^2	0.16	0.18	0.19	0.22

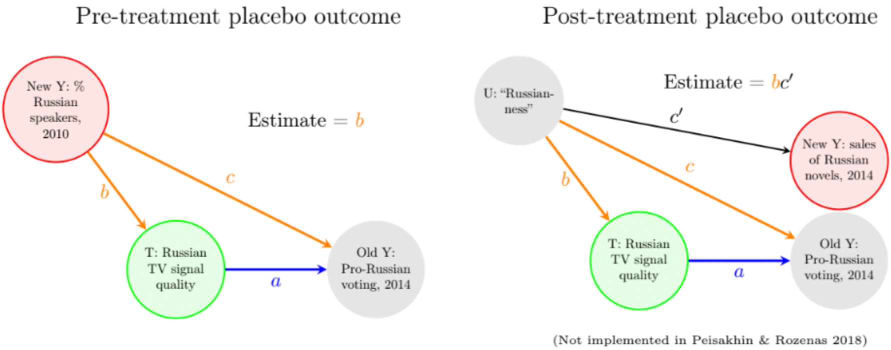
Notes: The table reports OLS estimates. The unit of observation is an individual. The dependent variable in the Asiabarometer sample is the respondent's answer to the question: "How much do you trust your local government?" The categories for the answers are the same in the Asiabarometer as in the Afrobarometer. Standard errors are clustered at the ethnicity level in the Afrobarometer regressions and at the location (city) level in the Asiabarometer and the WVS samples. The individual controls are for age, age squared, a gender indicator, education fixed effects, and religion fixed effects.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

Placebo outcomes

Figure 6: The logic of placebo outcome tests



Placebo outcomes

Allow for different types of tests:

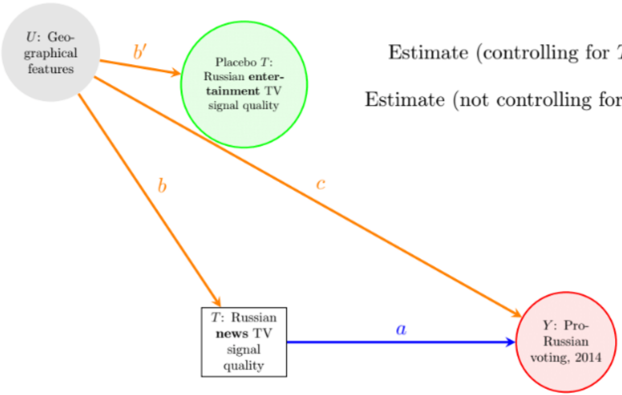
- Assessing *balance* across treated and controls: e.g. Cinelli and Hazlett (2020)
- Also assess different mechanisms

Placebo treatments

- We just use a treatment that we know has 0 effect
- Main question is: do we include the actual treatment in the test or not?
  - Tradeoff: if we don't include it, we may fail the placebo test because we pick up the "actual" treatment effects. If we do, we lose power (because we take away variation in the placebo variable)

Placebo treatments

Figure 7: Logic of a placebo treatment test



$$\text{Estimate (controlling for } T) = \frac{b'c(1-b^2)}{1-(b'b)^2}$$

$$\text{Estimate (not controlling for } T) = b'(c + ba)$$

Placebo treatments: other examples

TABLE 6. Regression: Placebos and Robustness

	(1)	(2)	(3)	(4)	(5)	(6)
A) Coefficient on Overseas Port in 1907 (OLS)						
# H-M Riots, 1850–1950	0.326	0.085	0.201	0.306	0.591	0.690
	[0.549]	[0.616]	[0.573]	[0.919]	[0.807]	[0.770]
R-squared	0.26	0.36	0.48	0.43	0.46	0.57
Any H-M Riot, 1850–1950	0.084	0.034	0.057	0.148	0.082	0.095
	[0.117]	[0.119]	[0.116]	[0.142]	[0.120]	[0.101]
R-squared	0.26	0.30	0.39	0.41	0.43	0.57
B) Coefficient on Medieval Port Silted by 1901 (OLS)						
# H-M Riots, 1850–1950	−1.308***	−1.245**	−1.298**	−1.187*	−1.439**	−1.375*
	[0.417]	[0.533]	[0.564]	[0.635]	[0.671]	[0.716]
R-squared	0.26	0.39	0.51	0.44	0.49	0.59
Any H-M Riot, 1850–1950	−0.272**	−0.233*	−0.198**	−0.203	−0.201	−0.096
	[0.111]	[0.111]	[0.086]	[0.126]	[0.144]	[0.150]
R-squared	0.27	0.31	0.40	0.41	0.44	0.57
C) Coefficient on Medieval Port (2SLS)						
# H-M Riots, 1850–1950	−3.938	−3.550*	−2.056	−3.363*	−2.374**	−2.118**
	[2.531]	[2.005]	[1.421]	[1.979]	[1.034]	[0.966]
Any H-M Riot, 1850–1950	−0.253	−0.657	−0.359	−0.240	−0.637	−0.648*
	[0.543]	[0.526]	[0.298]	[0.370]	[0.415]	[0.333]
Sample	Full	Coastal, <200 km	Coastal, <100 km	Full	Coastal, <200 km	Coastal, <100 km
Controls	Medieval	Medieval	Medieval	Medieval	Medieval	Medieval
Province/NS × Annex FE	No	No	No	Yes	Yes	Yes
Observations	248	110	89	248	110	89

Notes: Each cell represents a regression. All regressions include quadratic polynomials in Longitude and Latitude and Log. Distances from the Modern Coast, Navigable Rivers and the Ganges, Coastal Town and Natural Disasters, Medieval Town, Mughal Mint, Other Patronage Center, Inland Trade Route, Skilled Crafts in Town, Major Shi'a State, Centuries Muslim Rule. Robust standard errors (clustered at Native State × Annexation level): *significant at 10%; **5%; ***1%.

Placebos: summary

- Placebos give credibility to identification assumptions and allow to detect problems
- We should always think about them at the **design stage** (to collect appropriate data)
- Remember that a placebo is valid if the design remains the same except for the thing that you change

RDD

In this section: hands-on approach to RDD based on Cattaneo, Idrobo, and Titiunik (2019)

We will use the state of the art: the packages `rdrobust`, `rdlocrand`, `rddensity` (available in both Stata and R format).

Next week: Topics in RDD (estimation issues, fuzzy RDD, local randomization approach)

Working example: Meyersson (2014)

Econometrica, Vol. 82, No. 1 (January, 2014), 229–269

ISLAMIC RULE AND THE EMPOWERMENT OF THE POOR AND PIOUS

BY ERIK MEYERSSON[1]

Does Islamic political control affect women's empowerment? Several countries have recently experienced Islamic parties coming to power through democratic elections. Due to strong support among religious conservatives, constituencies with Islamic rule often tend to exhibit poor women's rights. Whether this reflects a causal relationship or a spurious one has so far gone unexplored. I provide the first piece of evidence using a new and unique data set of Turkish municipalities. In 1994, an Islamic party won multiple municipal mayor seats across the country. Using a regression discontinuity (RD) design, I compare municipalities where this Islamic party barely won or lost elections. Despite negative raw correlations, the RD results reveal that, over a period of six years, Islamic rule increased female secular high school education. Corresponding effects for men are systematically smaller and less precise. In the longer run, the effect on female education remained persistent up to 17 years after, and also reduced adolescent marriages. An analysis of long-run political effects of Islamic rule shows increased female political participation and an overall decrease in Islamic political preferences. The results are consistent with an explanation that emphasizes the Islamic party's effectiveness in overcoming barriers to female entry for the poor and pious.

KEYWORDS: Political Islam, regression discontinuity, education.

Meyersson (2014)

Causal effect of interest: Victory of Islamic candidate on educational attainment of women

Elements:

- **Outcome (Y)**: percentage of women aged 15-20 in 2000 who had completed high school by 2000
- **Running variable (X)**: vote percentage of the Islamic party minus vote percentage of the strongest secular opponent
- **Treatment (T)**: 1 if Islamic party won in 1994, 0 otherwise

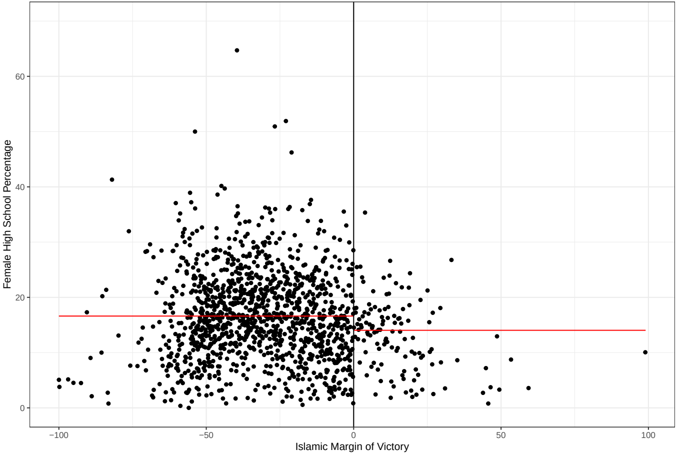
Let's begin

```
library(rdrobust); library(rddensity); library(haven)

# Import data and define variables
data <- read_dta("CIT_2019_Cambridge_polecon.dta")

Y <- data$Y
X <- data$X
T <- data$T
T_X <- T*X
```

Visualization

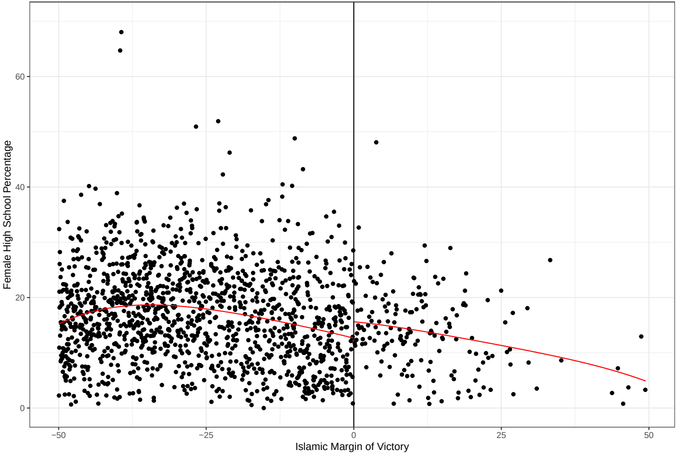
Visualization

```
# Before we begin:
# ?rdplot

# Raw means comparison
rdplot(Y, X, nbins = c(2500, 500), p = 0, col.lines = "red", col.dots = "black", title = "",
       x.label = "Islamic Margin of Victory", y.label = "Female High School Percentage", y.lim = c(0,70))
```



Visualization

```
# Local means comparison
rdplot(Y[abs(X) <= 50], X[abs(X) <= 50], nbins = c(2500, 500), p = 4, col.lines = "red", col.dots = "black",
       title = "", x.label = "Islamic Margin of Victory", y.label = "Female High School Percentage",
       y.lim = c(0,70))
```



Binning

It may be hard to spot visually discontinuities in raw data. A common approach is to “smooth” the data by binning. We have two things to do:

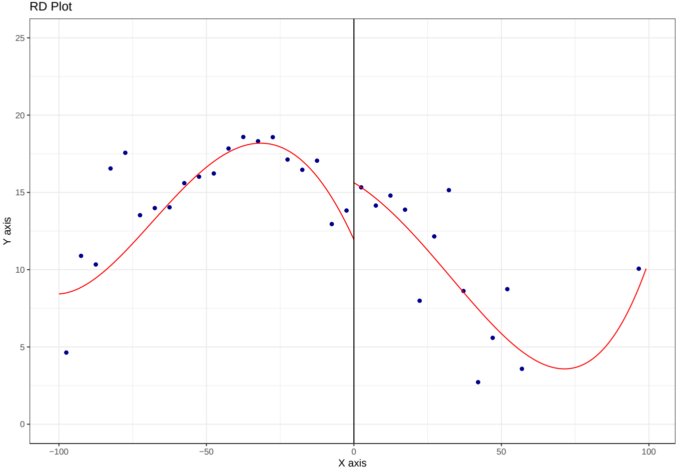
- Split the raw data into segments (bins) of the running variable, compute the mean outcome in each bin, plot the mean outcome against the mid point of the bin
- Overlay global polynomial fit of the outcome on the running variable, estimated **separately** on each side of the cutoff and using the **raw data**

We can do this automatically with the `rdplot` command

Binning

Set bin size manually: evenly-spaced.

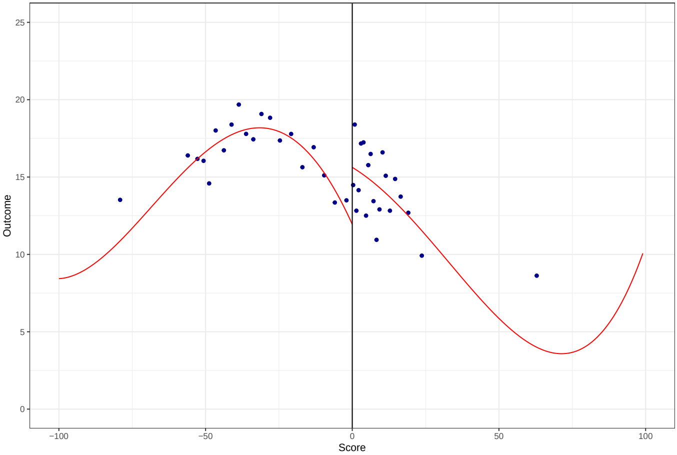
```
# Default is 4th polynomial degree on each side
rdplot(Y, X, nbins = c(20,20), binselect = "es", y.lim = c(0,25))
```



Principled binning

Some procedures retain information about the actual distribution of the data and reduce discretion. For instance, quantile-spaced bins.

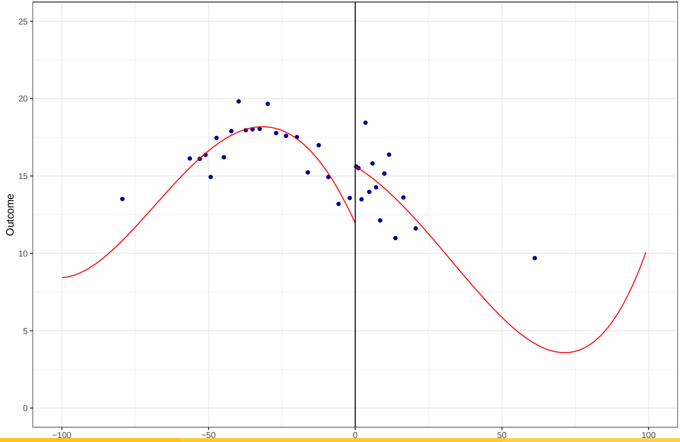
```
rdplot(Y, X, nbins = c(20,20), binselect = 'qs', x.label = 'Score',
      y.label = 'Outcome', title = '', x.lim = c(-100,100), y.lim = c(0,25))
```



Principled binning

We can also have a data-driven approach to the number of bins: default is minimize the IMSE of the local means estimator (optimizing along bias-variance)

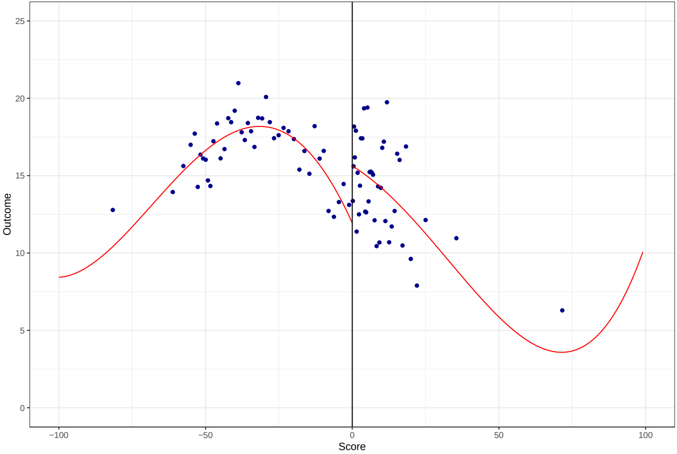
```
rdplot(Y, X, binselect = 'qs', x.label = 'Score',
       y.label = 'Outcome', title = '', x.lim = c(-100,100), y.lim = c(0,25))
```



Principled binning

Mimicking variance: choose number of bins so that the variability of means "mimicks" that of the raw data.

```
rdplot(Y, X, binselect = 'qsmv', x.label = 'Score',
      y.label = 'Outcome', title = '', x.lim = c(-100,100), y.lim = c(0,25))
```



Binning in sum

"Which method of implementation is most appropriate depends on the researcher's particular goal, for example, illustrating/testing for the overall functional form versus showing the variability of the data. We recommend to start with MV bins to better illustrate the variability of the outcome as a function of the score, ideally comparing ES and QS bins to highlight the distributional features of the score. Then, if needed, the researcher can select the number of bins to be IMSE-optimal in order to explore the global features of the regression function." (Cattaneo, Idrobo, and Titiunik 2019)

Estimation of causal effects

Continuity-based framework

If the CEF functions are continuous at the cutoff, RDD identifies a causal effect at the cutoff.

But in practice there are never observations with exactly the cutoff value. So we need to approximate the CEF on both sides of the cutoff.

Global approximations are good for plots (descriptions), but are not suitable for estimation of the treatment effect: see Gelman and Imbens (2019)

The best practice now is to use local polynomial functions with low order near the cutoff.

Estimation of causal effects

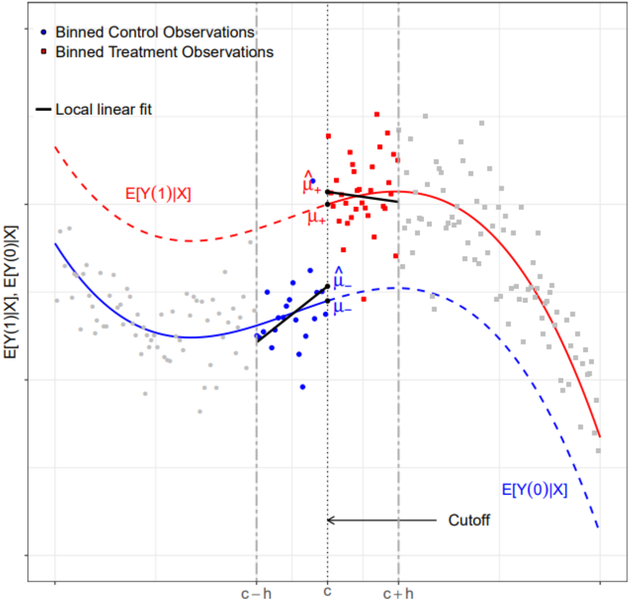
In a sharp RD:

- Choose polynomial of order p and a kernel function $K(.)$
- Choose a bandwidth h around the cutoff c
- Estimate on each side of the cutoff a WLS regression with weights $K(\frac{X_i - c}{h})$:

$$\hat{Y}_i = \hat{\mu}_+ + \hat{\mu}_{+,1}(X_i - c) + \hat{\mu}_{+,2}(X_i - c)^2 + ... + \hat{\mu}_{+,p}(X_i - c)^p$$
$$\hat{Y}_i = \hat{\mu}_- + \hat{\mu}_{-,1}(X_i - c) + \hat{\mu}_{-,2}(X_i - c)^2 + ... + \hat{\mu}_{-,p}(X_i - c)^p$$

- Calculate the sharp RD estimate: $\hat{\tau}_{SRD} = \hat{\mu}_+ - \hat{\mu}_-$, the difference of the two functions when $X_i = c$

Estimation of causal effects



Parameters

The relevant parameters are: bandwidth, kernel function, polynomial order

- **Kernel**: triangular one is recommended (weight $= 0$ outside h and $\uparrow$ as we get closer to c) and default in `rdrobust`. Uniform and Epanechnikov are also available
- **Bandwidth**: the most important thing. Usually chosen by a data-driven approach to minimize the MSE of the local polynomial point estimator
- **Polynomial order**: low to avoid overfitting, generally local linear is the default choice

Estimation

```r
# By default c = 0
out <- rdrobust(Y, X, kernel = "uniform", p = 1, h = 20)
summary(out)
```

```
## Call: rdrobust
##
## Number of Obs.                 2629
## BW type                      Manual
## Kernel                      Uniform
## VCE method                       NN
##
## Number of Obs.                 2314         315
## Eff. Number of Obs.             608         280
## Order est. (p)                    1           1
## Order bias  (q)                   2           2
## BW est. (h)                  20.000      20.000
## BW bias (b)                  20.000      20.000
## rho (h/b)                     1.000       1.000
## Unique Obs.                    2311         315
##
## ============================================================================
##         Method     Coef. Std. Err.         z     P>|z|      [ 95% C.I. ]
## ============================================================================
##   Conventional     2.927     1.235     2.371     0.018     [0.507 , 5.347]
##         Robust         -         -     1.636     0.102    [-0.582 , 6.471]
## ============================================================================
```

Estimation

`rdrobust` has a function `rdbwselect` which can select a variety of optimal bandwidths. It is a stand-alone function, but can be called from inside `rdrobust` using the option `bwselect`

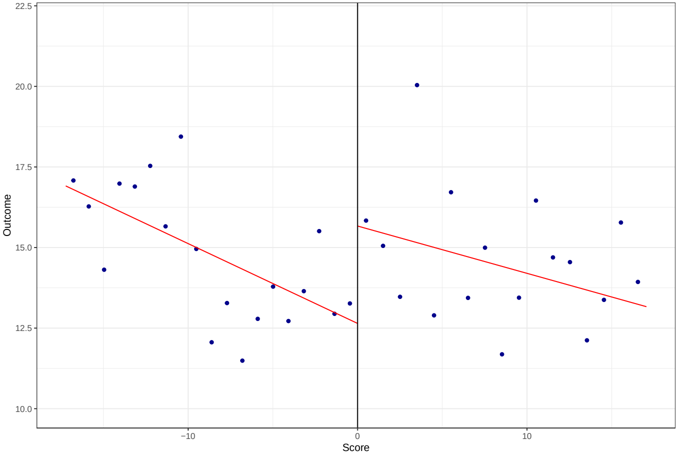
Estimation

```
# Choose h to minimize MSE
out <- rdrobust(Y, X, kernel = "triangular", p = 1, bwselect = "mserd")
summary(out)
```

```
## Call: rdrobust
##
## Number of Obs.                 2629
## BW type                       mserd
## Kernel                   Triangular
## VCE method                       NN
##
## Number of Obs.                 2314          315
## Eff. Number of Obs.             529          266
## Order est. (p)                    1            1
## Order bias  (q)                   2            2
## BW est. (h)                  17.240       17.240
## BW bias (b)                  28.576       28.576
## rho (h/b)                     0.603        0.603
## Unique Obs.                    2311          315
##
## =============================================================================
##         Method     Coef. Std. Err.         z     P>|z|      [ 95% C.I. ]
## =============================================================================
##   Conventional     3.020     1.427     2.116     0.034     [0.223 , 5.816]
##         Robust         -         -     1.776     0.076    [-0.309 , 6.276]
## =============================================================================
```

Illustrate the main effect

```
bw <- out$bws[1,1]
rdplot(Y[abs(X)<=bw], X[abs(X)<=bw], p = 1, kernel = 'triangular',
       x.label = 'Score', y.label = 'Outcome', title = '', y.lim = c(10,22))
```



Inference

The local polynomial function is just an approximation, as such may be biased. In order to make inference, we need to take into account the bias.

Routine approaches differ in whether

- They remove the estimated bias from the derivation of the confidence intervals
- They take incorporate extra variability from bias removal in the standard error estimate

Inference

Table 3: Local Polynomial Confidence Intervals

	Centered at	Standard Error
Conventional: $\mathtt{CI}_{\mathtt{us}}$	$\hat{\tau}_{\mathtt{SRD}}$	$\sqrt{\hat{\mathscr{V}}}$
Bias-Corrected: $\mathtt{CI}_{\mathtt{bc}}$	$\hat{\tau}_{\mathtt{SRD}} - \hat{\mathscr{B}}$	$\sqrt{\hat{\mathscr{V}}}$
Robust bias-corrected: $\mathtt{CI}_{\mathtt{rbc}}$	$\hat{\tau}_{\mathtt{SRD}} - \hat{\mathscr{B}}$	$\sqrt{\hat{\mathscr{V}}_{\mathtt{bc}}}$

Inference

Robust bias correction enables to do valid inference using the same bandwidth used for the point estimate.

Another approach is to use a different bandwidth for the calculation of the standard error.

Inference

```r
out <- rdrobust(Y, X, kernel = 'triangular',  p = 1, bwselect = 'mserd', all = TRUE)
summary(out)
```

```
## Call: rdrobust
##
## Number of Obs.                 2629
## BW type                       mserd
## Kernel                   Triangular
## VCE method                       NN
##
## Number of Obs.                 2314         315
## Eff. Number of Obs.             529         266
## Order est. (p)                    1           1
## Order bias  (q)                   2           2
## BW est. (h)                  17.240      17.240
## BW bias (b)                  28.576      28.576
## rho (h/b)                     0.603       0.603
## Unique Obs.                    2311         315
##
## =============================================================================
##         Method     Coef. Std. Err.         z     P>|z|      [ 95% C.I. ]
## =============================================================================
##   Conventional     3.020     1.427     2.116     0.034     [0.223 , 5.816]
## Bias-Corrected     2.983     1.427     2.090     0.037     [0.186 , 5.780]
##         Robust     2.983     1.680     1.776     0.076    [-0.309 , 6.276]
## =============================================================================
```

Conclusion

- Look at the documentation of the package
- See CIT for additional examples and code snippets in Stata as well