8) Sharp Regression Discontinuity Design

Vitor Kamada

July 2018

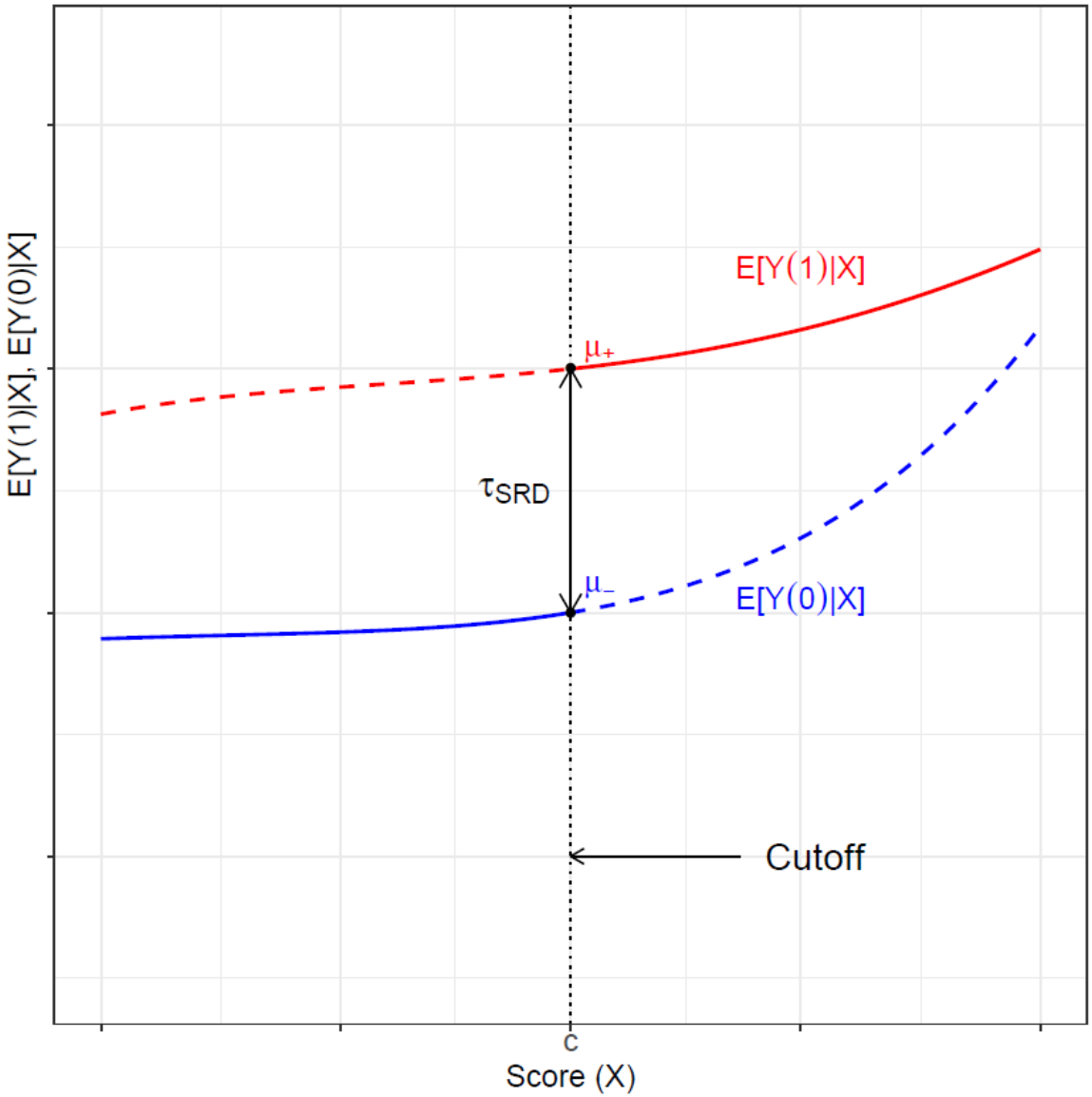
Reference

Tables, Graphics, and Figures from

A Practical Introduction to Regression Discontinuity Designs: Volume I

Cattaneo et al. (2018)

Treatment Effect in Sharp RD Design

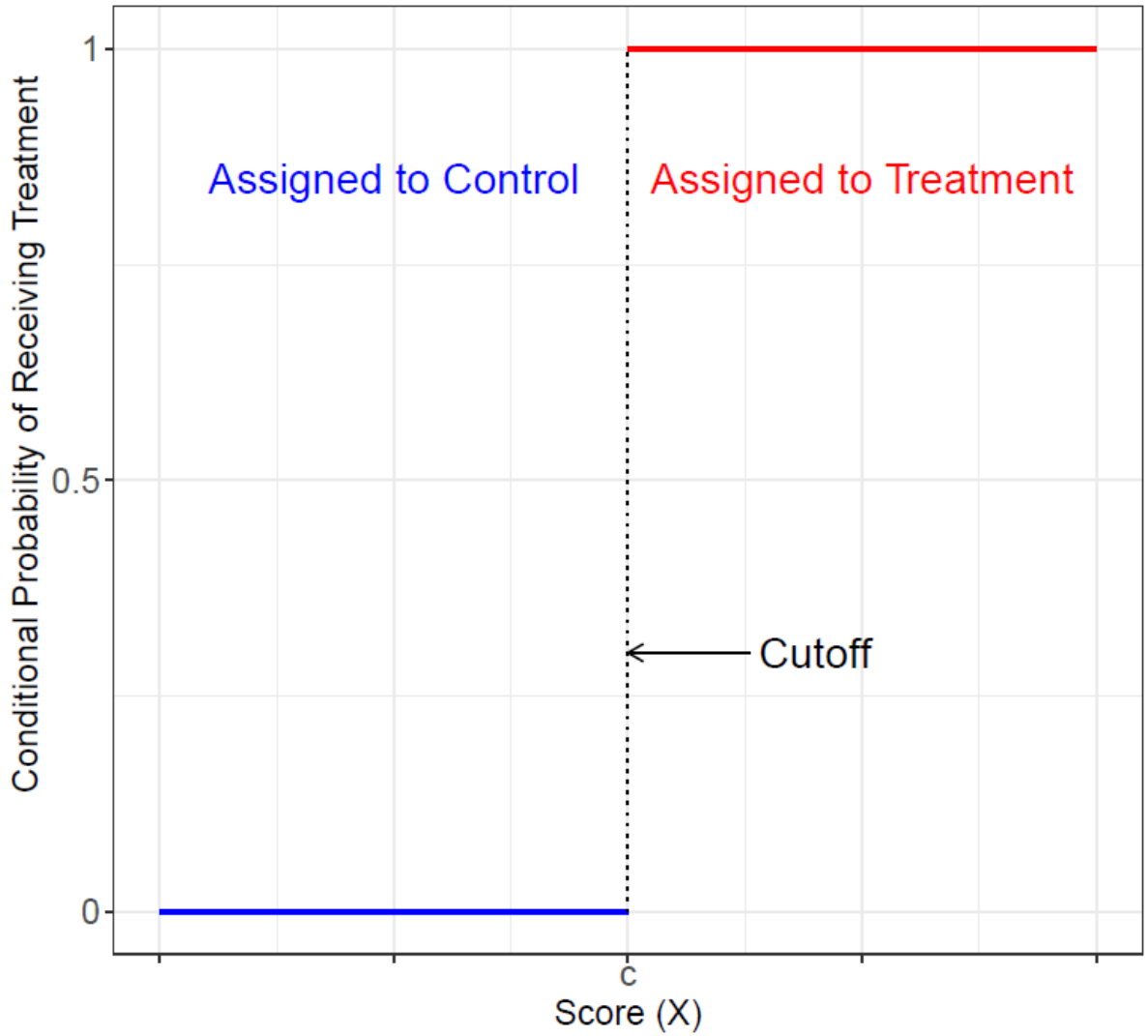


Sharp RD

$$D_i = \begin{cases} 1 \ \ if \ \ x_i \geq x_0 \\ 0 \ \ if \ \ x_i < x_0 \end{cases}$$

$$y_i = \alpha + \rho D_i + \beta x_i + \delta D_i x_i + \eta_i$$

Conditional Probability of Receiving Treatment



Generalization of RD

$$y_i = f(x_i) + \rho D_i + \eta_i$$

$$y_i = \alpha + \beta_1 x_i + \beta_2 x_i^2 + ... + \beta_p x_i^p + \rho D_i + \eta_i$$

$$y_i = \alpha + \beta_{01}\tilde{x}_i + \beta_{02}\tilde{x}_i^2 + ... + \beta_{0p}\tilde{x}_i^p + \rho D_i + \beta_1^* D_i \tilde{x}_i + \beta_2^* D_i \tilde{x}_i^2 + ... + \beta_p^* D_i \tilde{x}_i^p + \eta_i$$

$$\tilde{x}_i = x_i - x_0$$

Nonparametric Estimation Strategy

$$E[y_i|x_0 - \Delta \leq x_i < x_0] \cong E[y_{0i}|x_i = x_0]$$

$$E[y_i|x_0 \leq x_i < x_0 + \Delta] \cong E[y_{1i}|x_i = x_0]$$

$$\lim_{\Delta \to 0} E[y_i|x_0 \leq x_i < x_0+\Delta] - E[y_i|x_0-\Delta \leq x_i < x_0]$$

$$= E[y_{1i} - y_{0i}|x_i = x_0]$$

$$= \rho \textbf{ (Average Causal Effect)}$$

Islamic Rule and the Empowerment of the Poor and Pious

Y: % of women aged 15 to 20 completed high school in 2000

X: vote margin of the Islamic candidate for mayor in the 1994 Turkish elections

T: electoral victory of the Islamic candidate in 1994

Meyersson (2014) in Econometrica

Descriptive Statistics for Meyersson (2014)

Variable	Mean	Median	Std. Dev.	Min.	Max.
Y	16.306	15.523	9.584	0.000	68.038
X	-28.141	-31.426	22.115	-100.000	99.051
T	0.120	0.000	0.325	0.000	1.000
Percentage of men aged 15-20 with high school education	19.238	18.724	7.737	0.000	68.307
Islamic percentage of votes in 1994	13.872	7.029	15.385	0.000	99.526
Number of parties receiving votes 1994	5.541	5.000	2.192	1.000	14.000
Log population in 1994	7.840	7.479	1.188	5.493	15.338
Percentage of population below 19 in 2000	40.511	39.721	8.297	6.544	68.764
Percentage of population above 60 in 2000	9.222	8.461	3.960	1.665	27.225
Gender ratio in 2000	107.325	103.209	25.293	74.987	1033.636
Household size in 2000	5.835	5.274	2.360	2.823	33.634
District center	0.345	0.000	0.475	0.000	1.000
Province center	0.023	0.000	0.149	0.000	1.000
Sub-metro center	0.022	0.000	0.146	0.000	1.000

Note: the number of observations for all variables is 2,629

RD Libraries

library(lpdensity)

library(rddensity)

library(rdrobust)

library(rdlocrand)

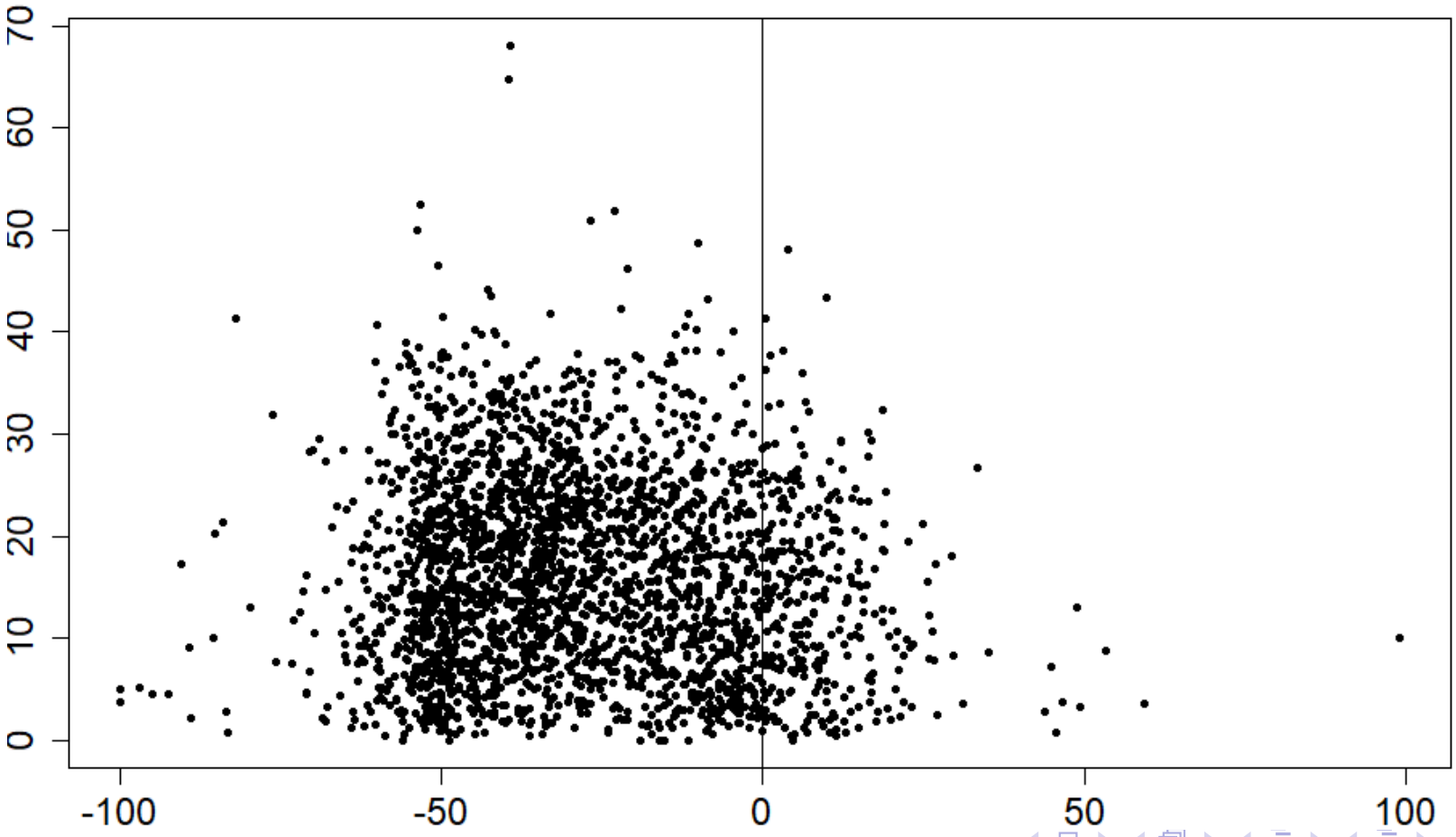
library(foreign); library(ggplot2)

library(TeachingDemos)

https://sites.google.com/site/rdpackages/

plot(X, Y, xlab = "Score", ylab = "Outcome", col = 1, pch = 20, cex.axis = 1.5, cex.lab = 1.5)

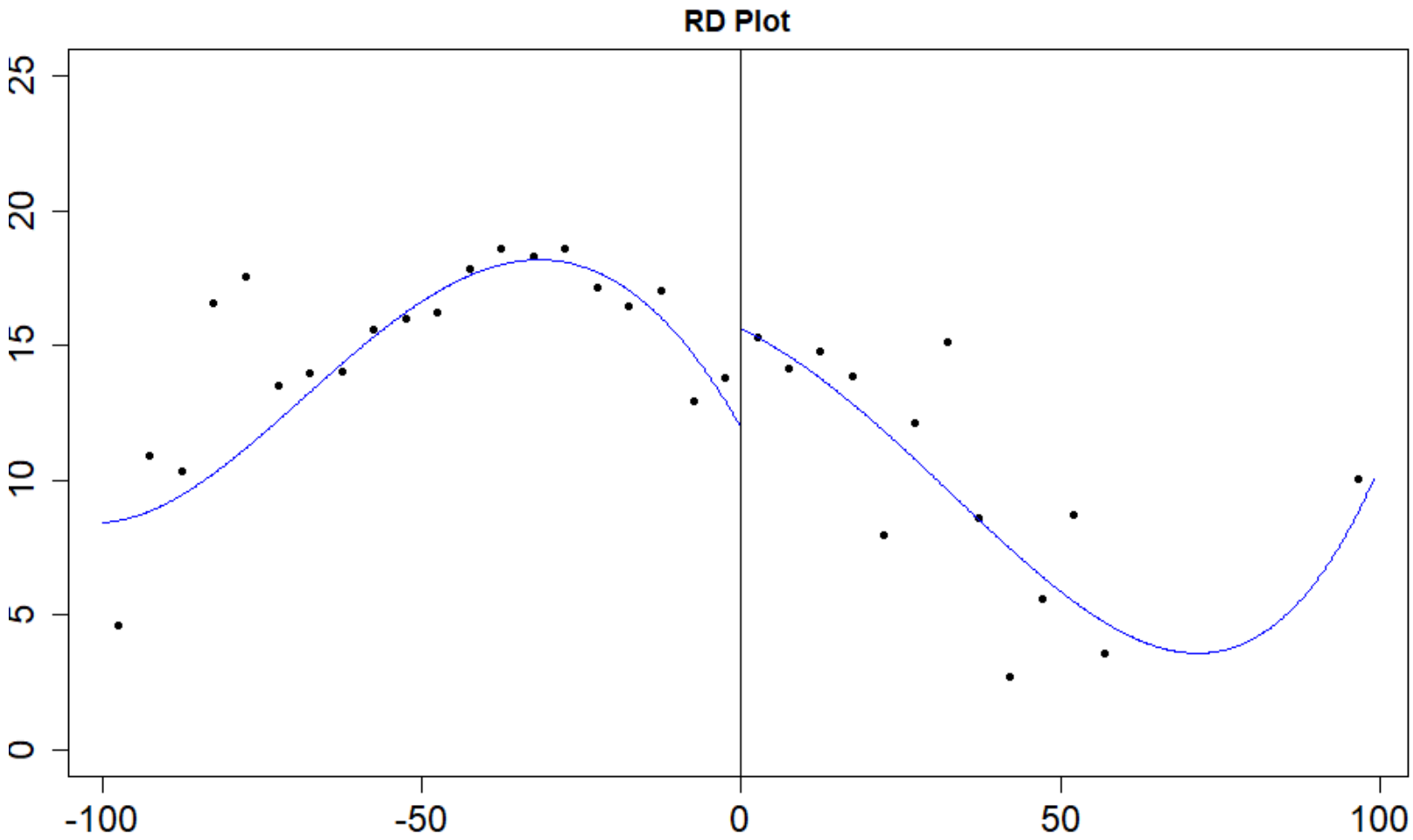
abline(v=0)



Partition of Islamic Margin of Victory into 40 Bins of Equal Length

Bin	Average Outcome in Bin	Number of Observations	Group Assignment
$\mathcal{B}_{-,1} = [-100, -95)$	$\bar{Y}_{-,1} = 4.6366$	4	Control
$\mathcal{B}_{-,2} = [-95, -90)$	$\bar{Y}_{-,2} = 10.8942$	2	Control
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathcal{B}_{-,19} = [-10, -5)$	$\bar{Y}_{-,19} = 12.9518$	149	Control
$\mathcal{B}_{-,20} = [-5, 0)$	$\bar{Y}_{-,20} = 13.8267$	148	Control
$\mathcal{B}_{+,1} = [0, 5)$	$\bar{Y}_{+,1} = 15.3678$	109	Treatment
$\mathcal{B}_{+,2} = [5, 10)$	$\bar{Y}_{+,2} = 13.9640$	83	Treatment
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathcal{B}_{+,19} = [90, 95)$	$\bar{Y}_{+,19} = \text{NA}$	0	Treatment
$\mathcal{B}_{+,20} = [95, 100]$	$\bar{Y}_{+,20} = 10.0629$	1	Treatment

out = rdplot(Y, X, nbins = c(20,20), binselect = 'esmv', y.lim = c(0,25), cex.axis = 1.5, cex.lab = 1.5)



Sharp RD by OLS

```
out = lm(Y[X >= -20 & X <= 20] ~ X[X >= -20 & X <= 20] + T[X >= -20 & X <= 20] + T_X[X >= -20 & X <= 20])

summary(out)
```

```
Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               12.62254    0.77459  16.296  < 2e-16 ***
X[X >= -20 & X <= 20]     -0.24807    0.06723  -3.690 0.000238 ***
T[X >= -20 & X <= 20]      2.92708    1.23529   2.370 0.018024 *
T_X[X >= -20 & X <= 20]    0.12612    0.12459   1.012 0.311667
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

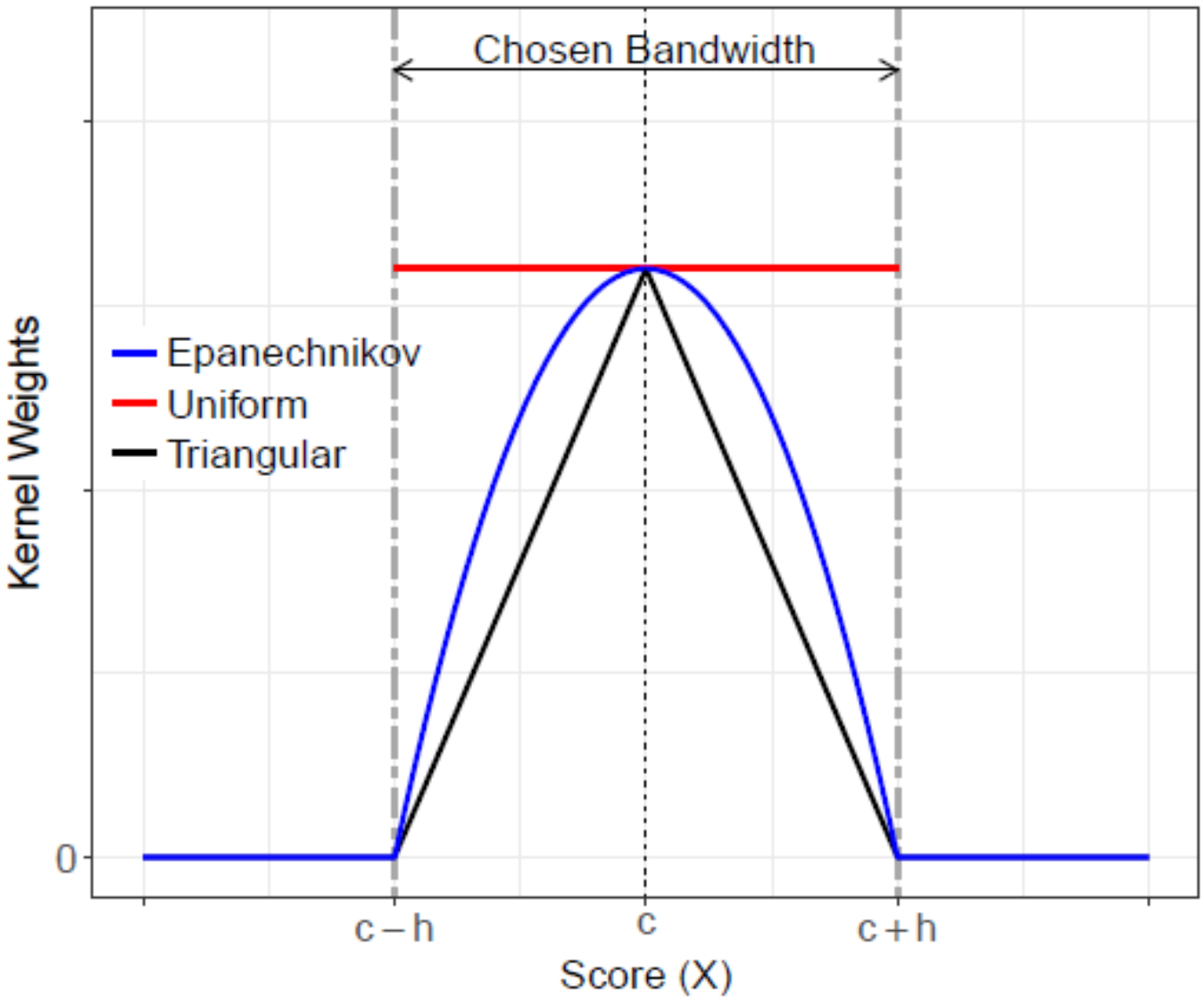
Kernel: Local Linear Estimator

$$\min_{\alpha,\beta} \sum_{i=1}^{N} K\left(\frac{x_i - x}{h}\right)[y_i - \alpha - \beta(x_i - x)]^2$$

$$\begin{pmatrix} \hat{\alpha}(x) \\ \hat{\beta}(x) \end{pmatrix} = (Z'KZ)^{-1}Z'Ky$$

$$h \to \infty,\ \hat{m}(x) \to \hat{\alpha} + \hat{\beta}x$$

Kernel Choice



out = rdrobust(Y, X, kernel = "triangular", p = 1, bwselect = "mserd")

```
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  (p)                    2           2
BW est. (h)                   17.239      17.239
BW bias (b)                   28.575      28.575
rho (h/b)                      0.603       0.603

=============================================================================
        Method     Coef. Std. Err.         z     P>|z|      [ 95% C.I. ]
=============================================================================
  Conventional     3.020     1.427     2.116     0.034     [0.223 , 5.817]
        Robust         -         -     1.776     0.076     [-0.309 , 6.276]
```

Z = cbind(data$vshr_islam1994, data$partycount, data$lpop1994, data$merkezi, data$merkezp, data$subbuyuk, data$buyuk)

out = rdrobust(Y, X, covs = Z, kernel = 'triangular', scaleregul = 1, p = 1, bwselect = 'mserd')

Number of Obs.	2314	315
Eff. Number of Obs.	448	241
Order est. (p)	1	1
Order bias (p)	2	2
BW est. (h)	14.409	14.409
BW bias (b)	23.731	23.731
rho (h/b)	0.607	0.607

Method	Coef.	Std. Err.	z	P>|z|	[95% C.I.]
Conventional	3.108	1.284	2.421	0.015	[0.592 , 5.624]
Robust	-	-	2.088	0.037	[0.194 , 6.132]