Regression Discontinuity Lab

Discussion 10

Wednesday, Oct 25

Reminders & Announcements

- Peer reviews for PS4 are due tomorrow at 5pm

- PS5 released tomorrow

  ‣ Due Thursday, Nov 2nd by 5pm

  ‣ Will cover instrumental variables and regression discontinuity

- Final project write-up due Tuesday, November 21 @ 5pm

  ‣ Connect with your group if you haven't yet

  ‣ Let us know if you have questions/concerns

Group Activity
Regression Discontinuity Review

- Work with the people around you

- Favorite Fall or Halloween tradition? Any fun costumes planned this year?

- Review terms: What do we mean by running variable, cutoff, and bandwidth in regression discontinuity?

  ‣ Look for definitions in https://theeffectbook.net/ch-RegressionDiscontinuity.html#how-does-it-work-4

- What is an example scenario where regression discontinuity might be useful in estimating some causal effect?

Regression Discontinuity
Big Picture

- When treatment is assigned discontinuously according to some cutoff

  ‣ Under the cutoff, you get no treatment

  ‣ Above the cutoff, you get treatment

- Around the cutoff, we expect people to be similar (the only difference is whether or not they receive treatment and we think of it as kinda random)

- Estimate a local average treatment effect (LATE): the effect of treatment on individuals near the cutoff

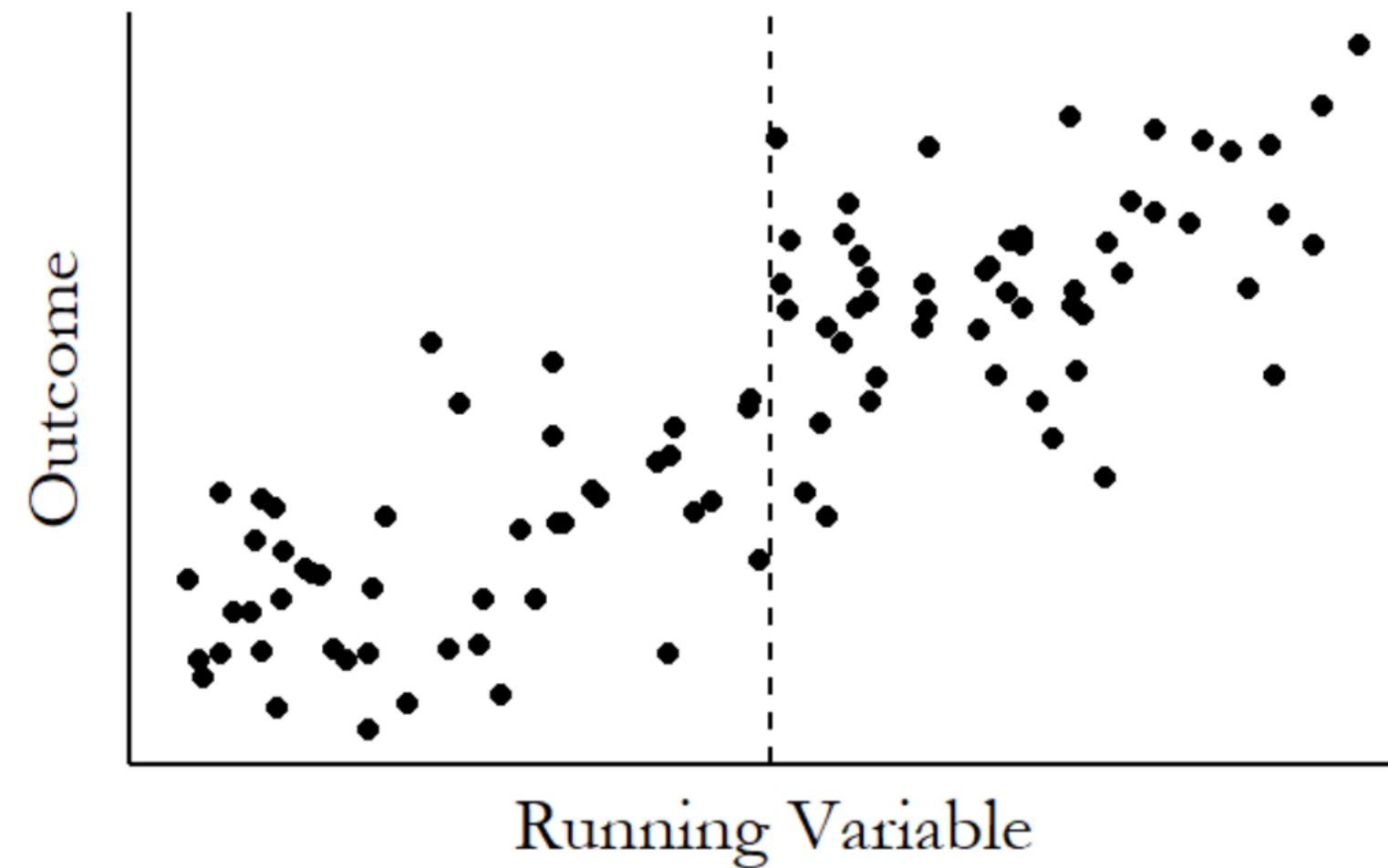# Regression Discontinuity Big Picture

Big Picture

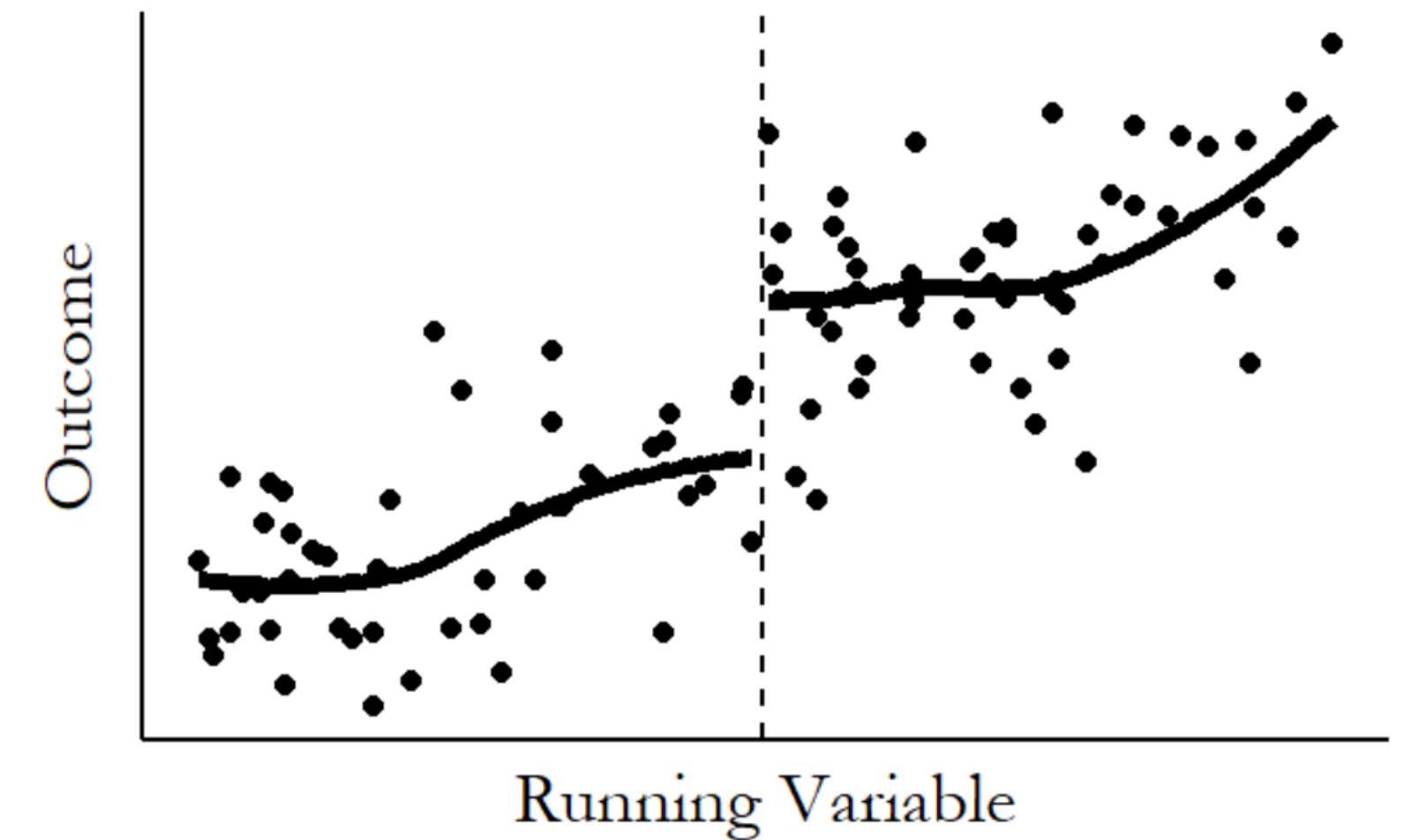(b) Predict values near the cutoff using regression models

(c) Determine how far away from the cutoff you're willing to look
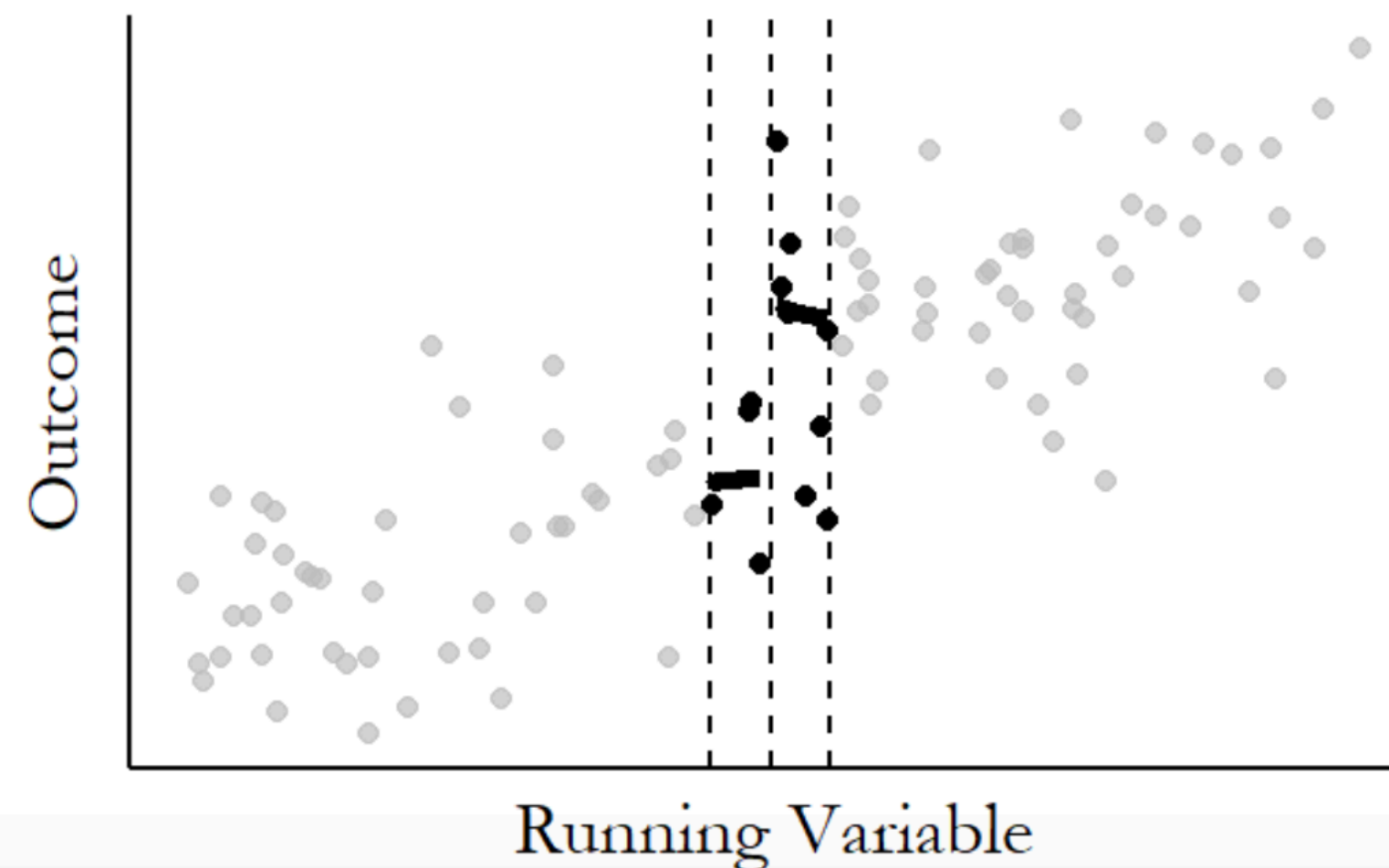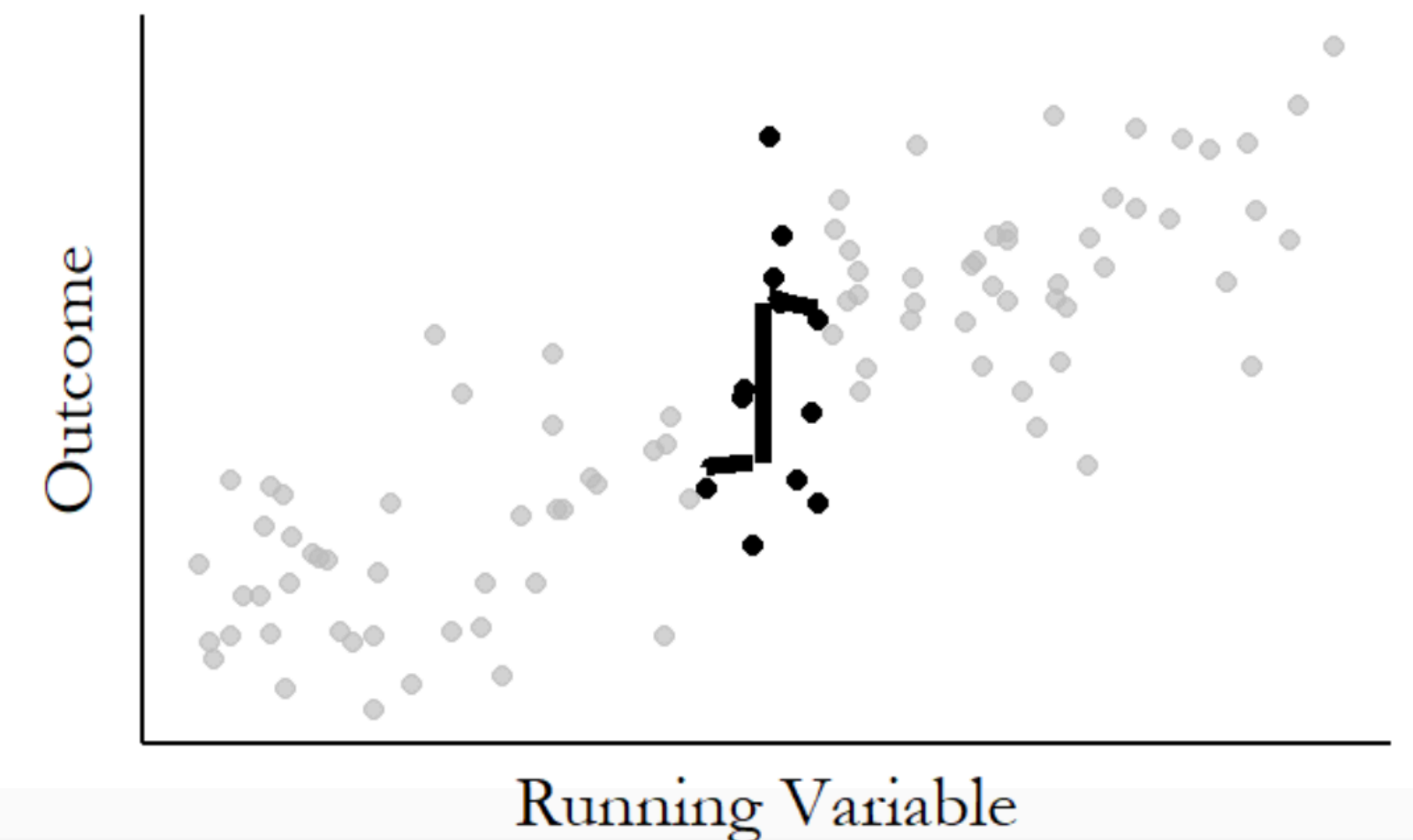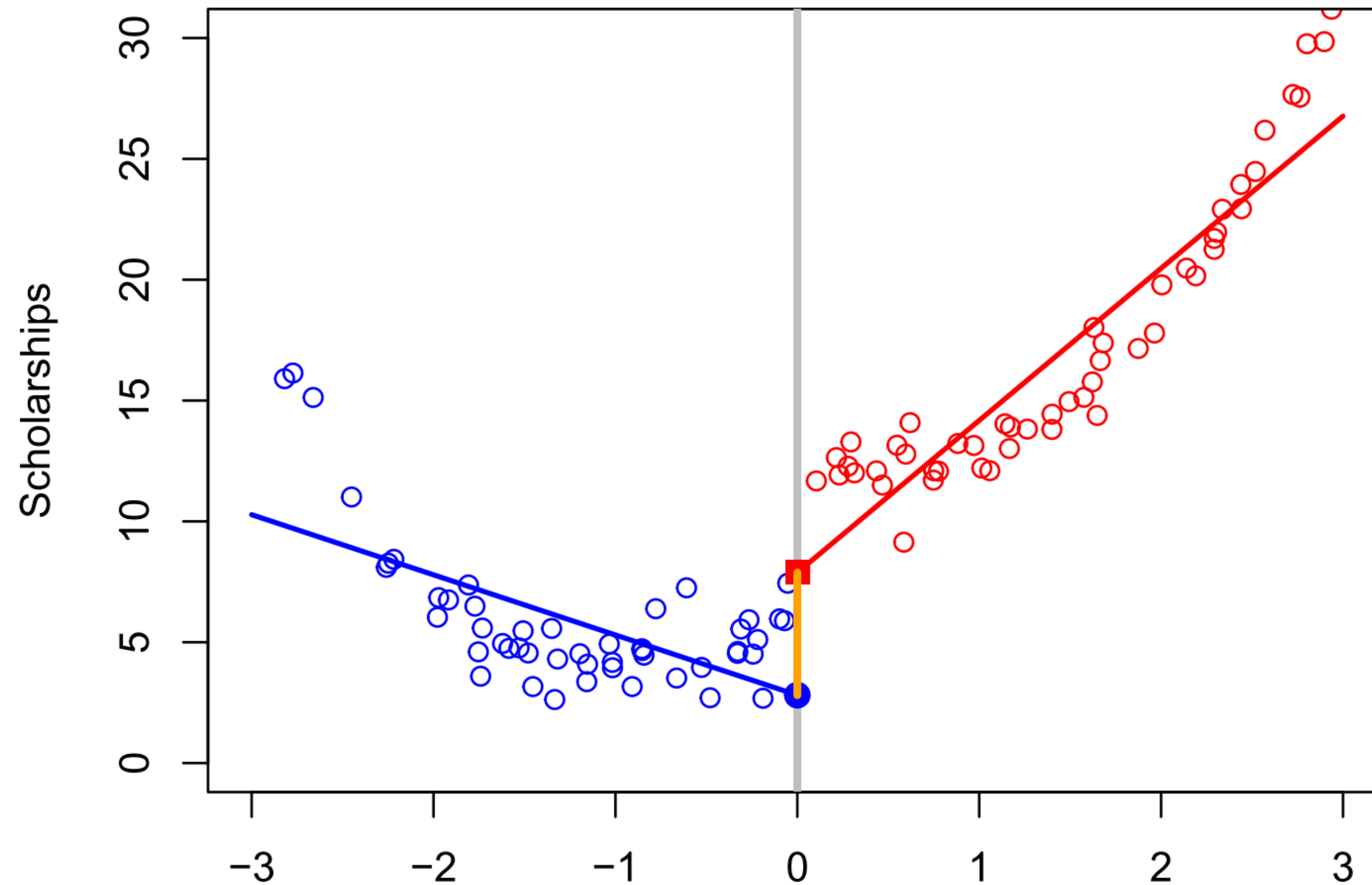
(d) Measure how far the jump is at the cutoff
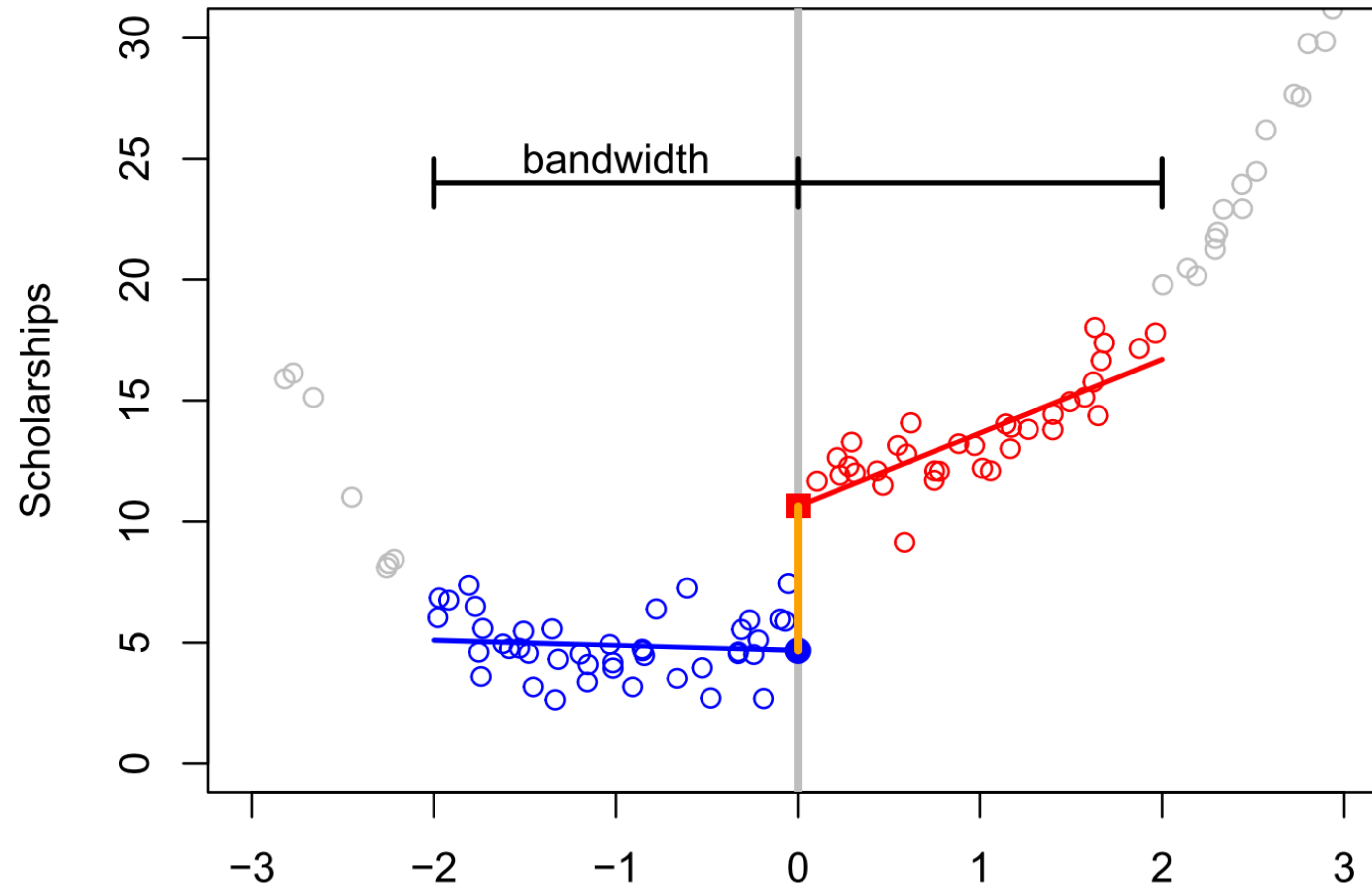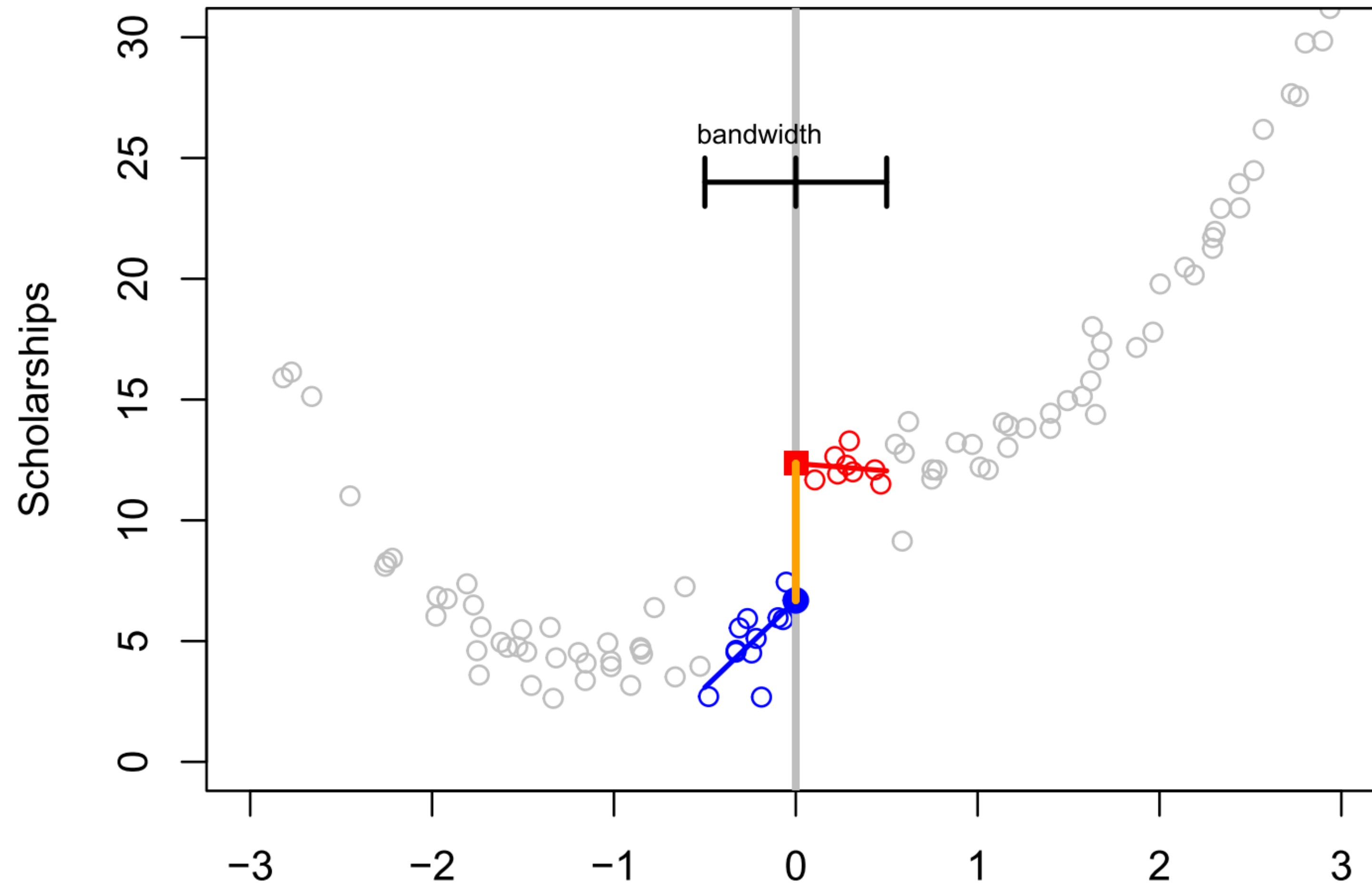
Choosing a Bandwidth

Non-linear setting

What if $E[Y^{a=1} \mid X]$ is non-linear?

Choosing a Bandwidth

Non-linear setting

What if $E[Y^{a=1} \,|\, X]$ is non-linear?

Choosing a Bandwidth

Non-linear setting

What if $E[Y^{a=1} \mid X]$ is non-linear?
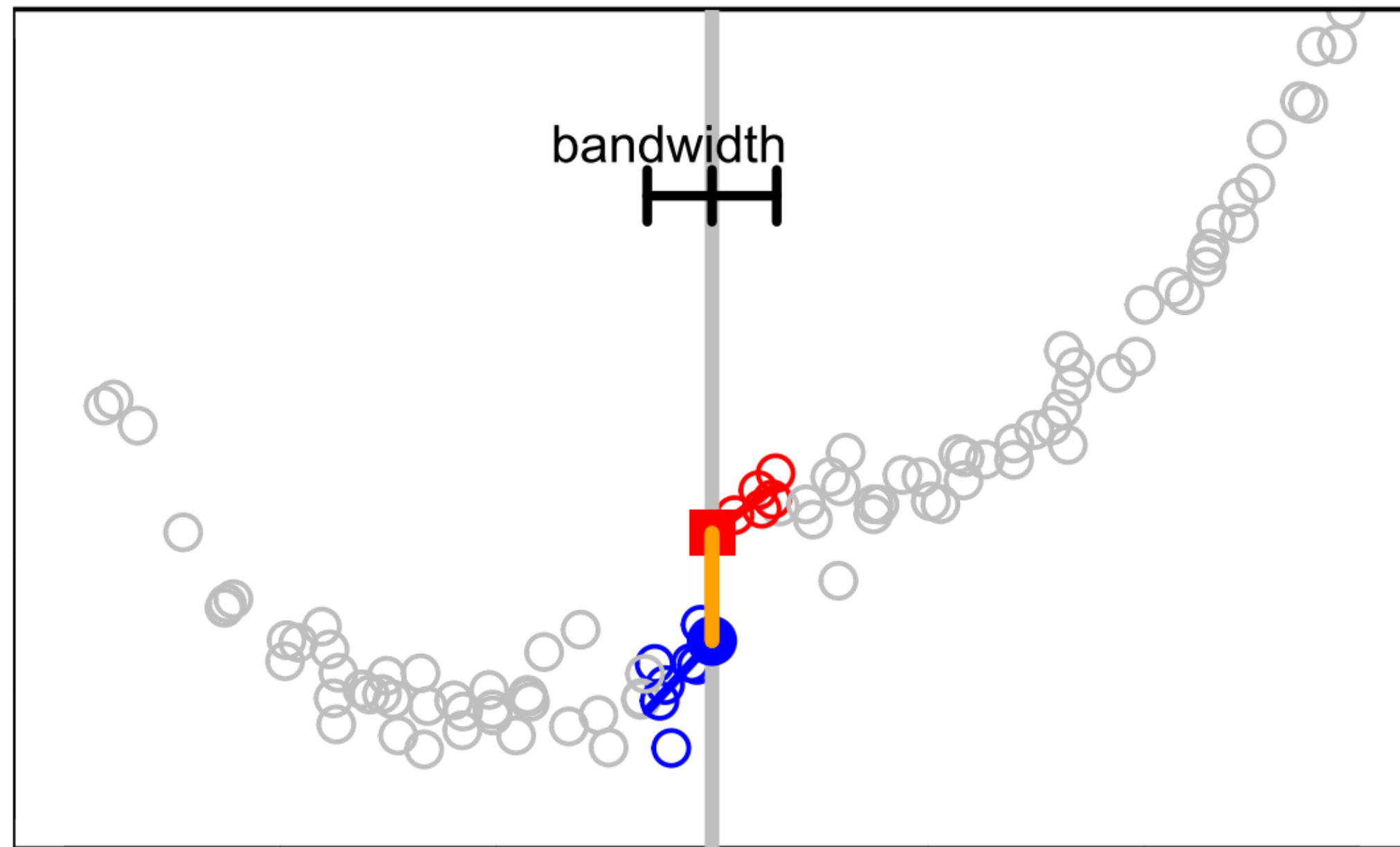
Choosing a Bandwidth

How do we choose a bandwidth?
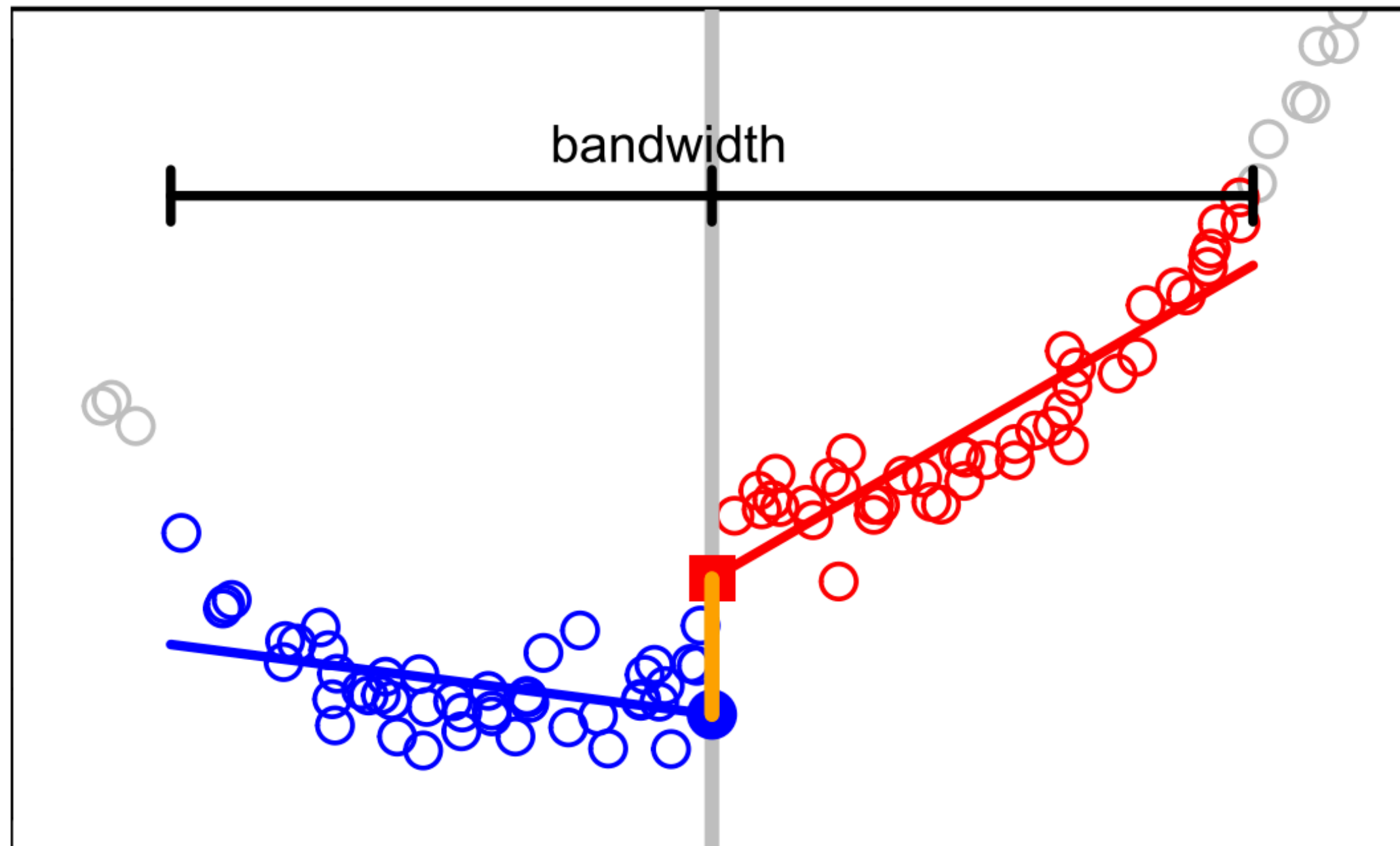
- **Bias:** How far from the truth are we with infinite data?

Choosing a Bandwidth

How do we choose a bandwidth?

- **Bias:** How far from the truth are we with infinite data?

- **Variance:** How much would my estimate change in a new sample?

- Typically, bandwidth should decrease as sample size increases

Weighting
Big Picture

Triangular Kernel
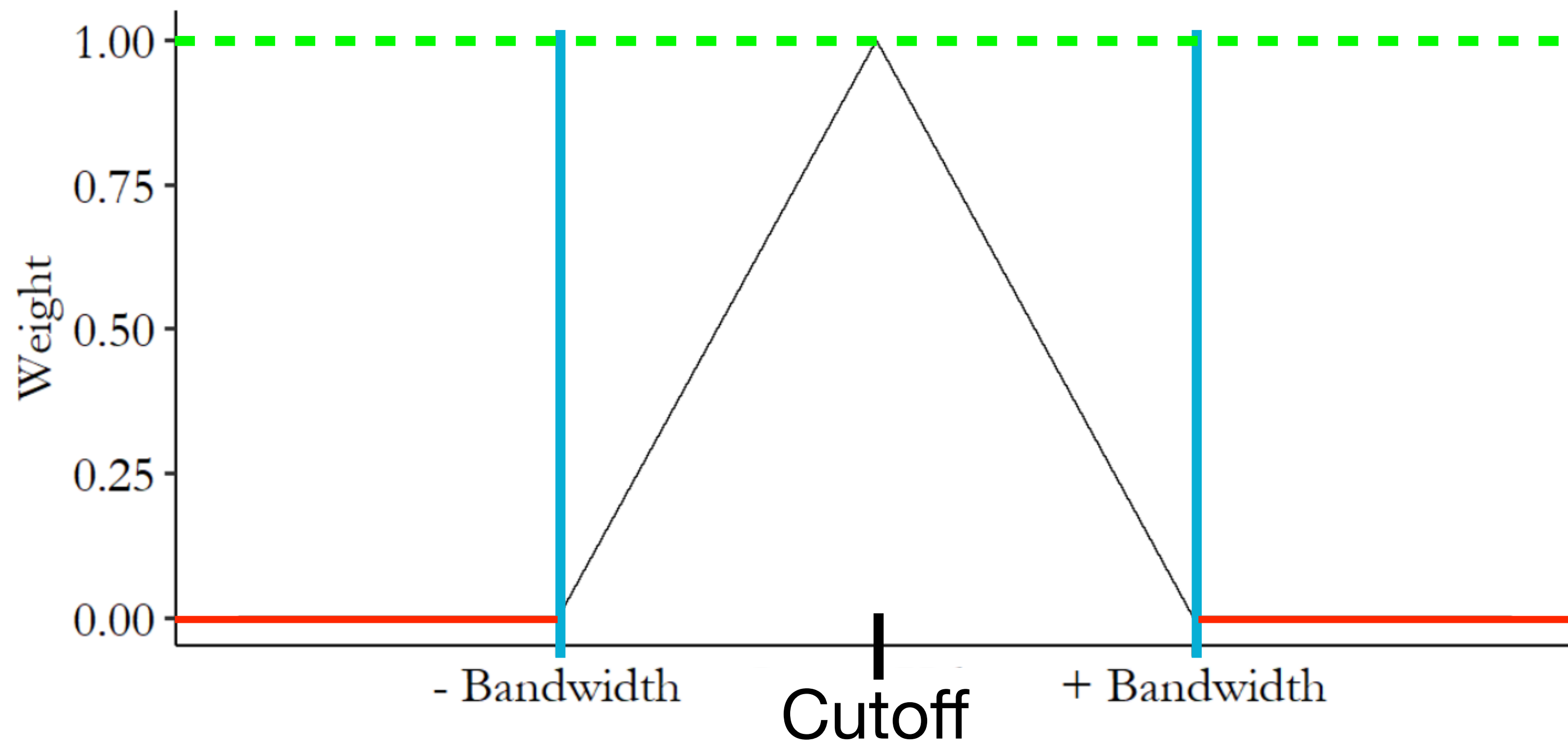
- Observations closest to the cutoff are the "most accurate"

- We just discussed tradeoffs between bigger versus smaller bandwidths

- Even within a bandwidth, we may want to weigh observations closer to the cutoff more heavily than observations farther from it

Regression Discontinuity in Code
The `rdrobust` **package in R**

- The `rdrobust` package in R basically takes care of everything for us!

  ‣ Chooses bandwidth, estimates causal effects, gives standard error

  ‣ `results = rdrobust(y, x, kernel, p, h)`

    ‣ y = dependent variable, x = running variable, `kernel` = optional weighting, p = degree of polynomial for regression, h = pre-selected bandwidth

```
# uniform kernel with bandwidth 10
out <- rdrobust(dem_vote_t2, dem_margin_t0, kernel = 'uniform',  p = 1, h = 10)
summary(out)
```

Regression Discontinuity in Code
The rdrobust package in R

```
## Sharp RD estimates using local polynomial regression.
##
## Number of Obs.                1297
## BW type                     Manual
## Kernel                     Uniform
## VCE method                      NN
##
## Number of Obs.                 595          702
## Eff. Number of Obs.            245          206
## Order est. (p)                   1            1
## Order bias  (q)                  2            2
## BW est. (h)                 10.000       10.000
## BW bias (b)                 10.000       10.000
## rho (h/b)                    1.000        1.000
## Unique Obs.                    595          702
##
## =================================================================
##           Method     Coef. Std. Err.        z      P>|z|     [ 95% C.I. ]
## =================================================================
##     Conventional     6.899     1.722    4.007      0.000    [3.525 , 10.273]
##           Robust         -     3.891    0.000      [5.156 , 15.624]
## =================================================================
```