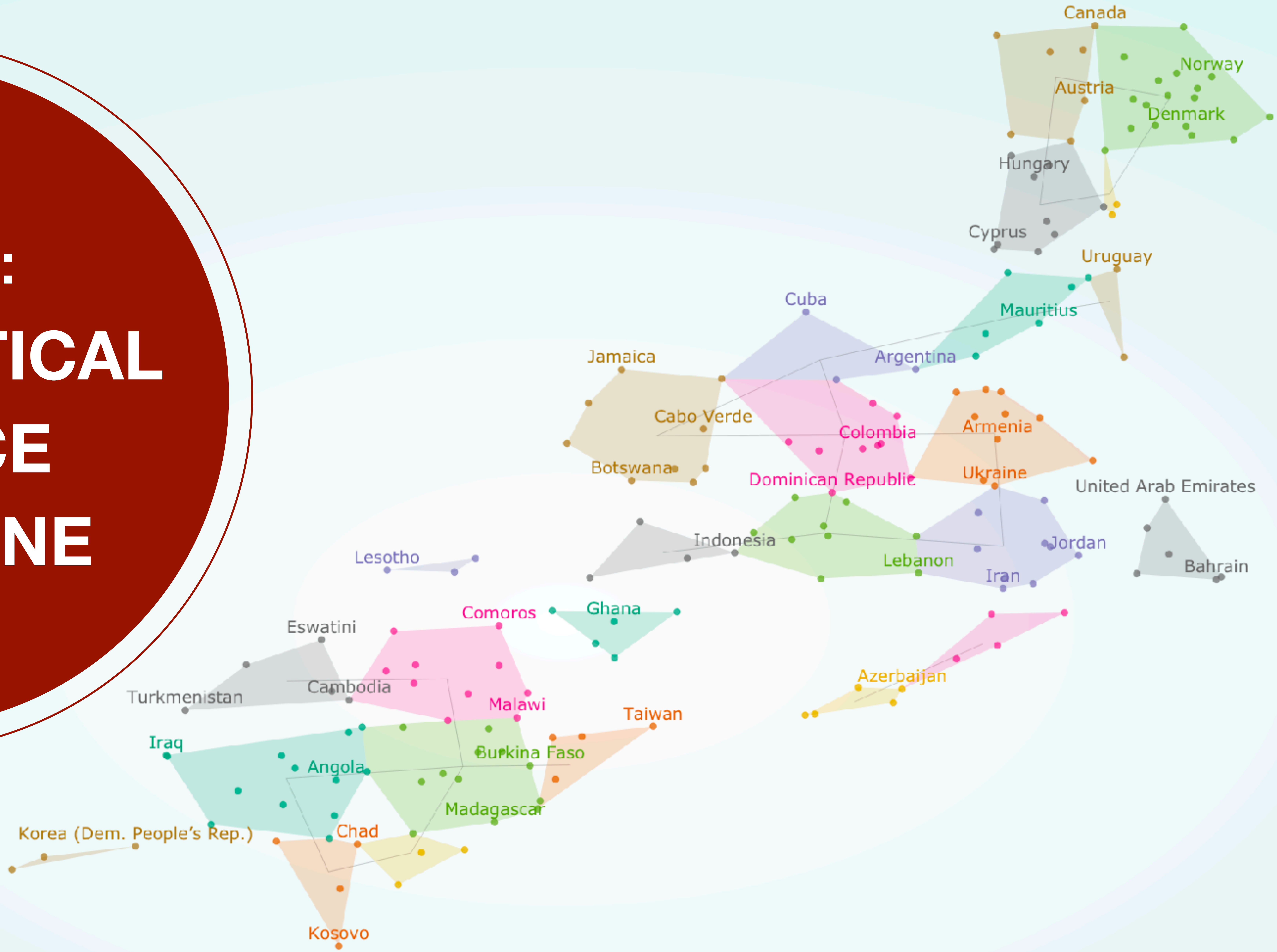
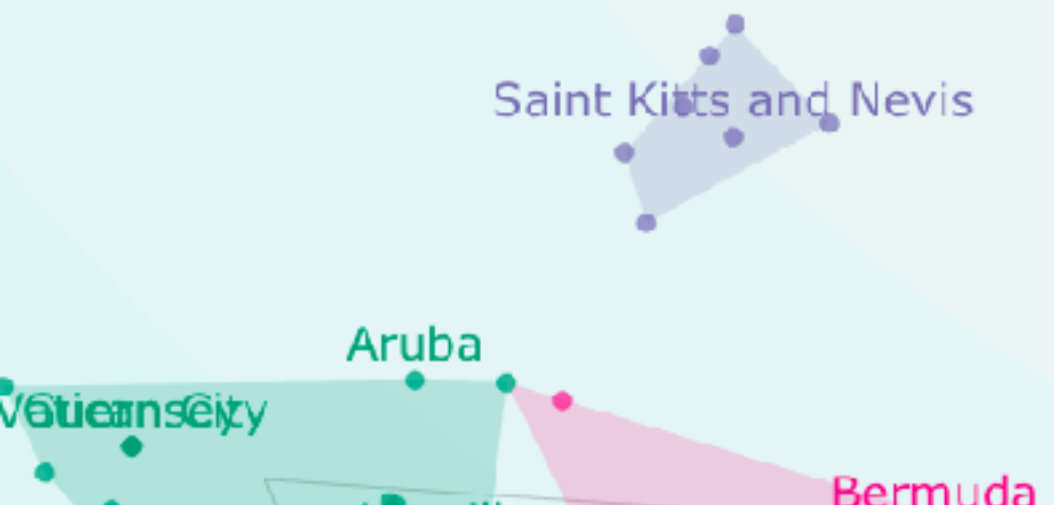
Poli-301: The Political Science Discipline





TODAY'S AGENDA

1. Quick recap of Homework
2. Finish up diff-in-diff
3. Regression discontinuity designs
4. Begin uncertainty

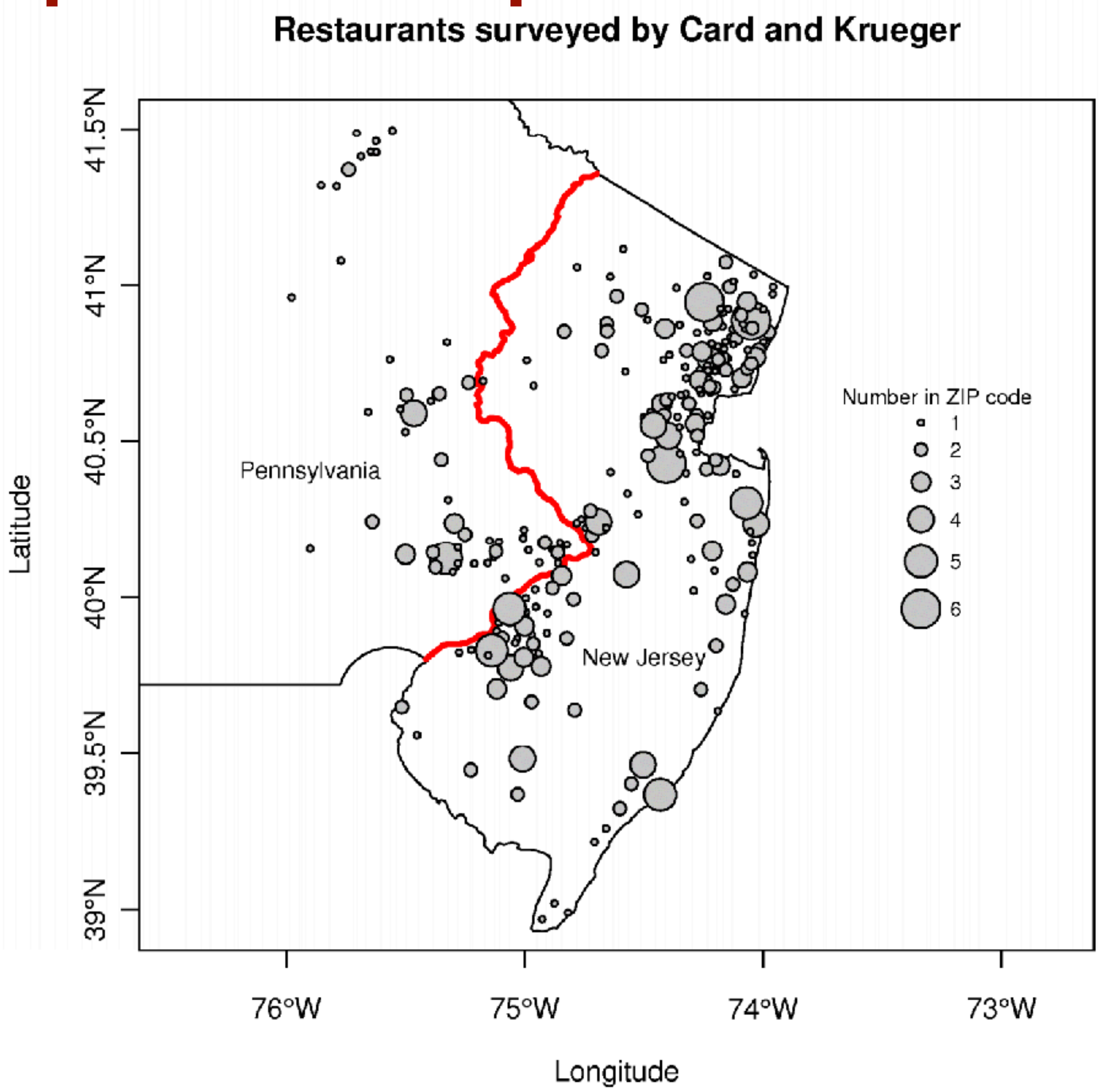
Other examples: maps

NJ implements min. Wage increase; what's the effect on employment?

Naively comparing states with and without laws is bad

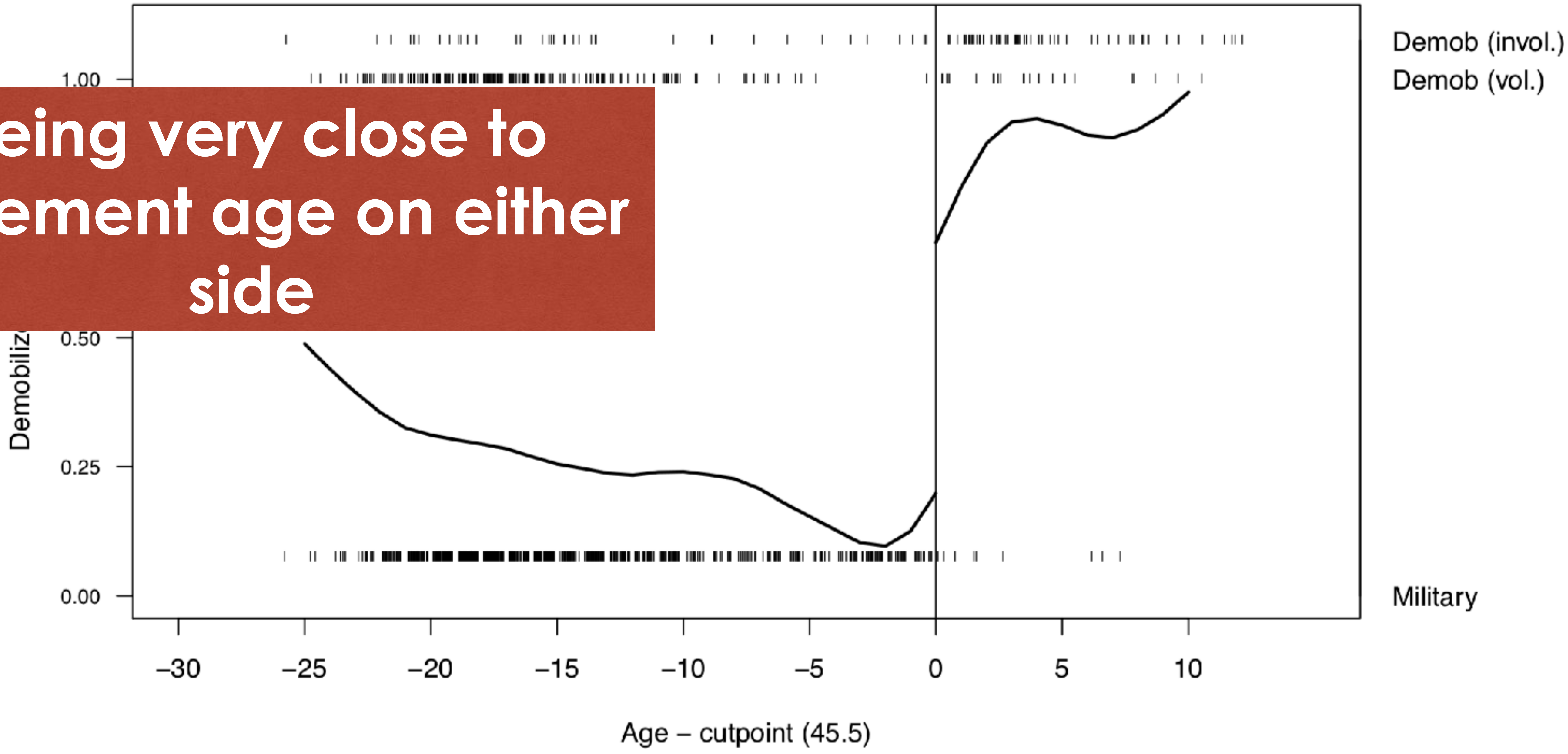
Compare states very close to border

Claim: otherwise identical, except some got law



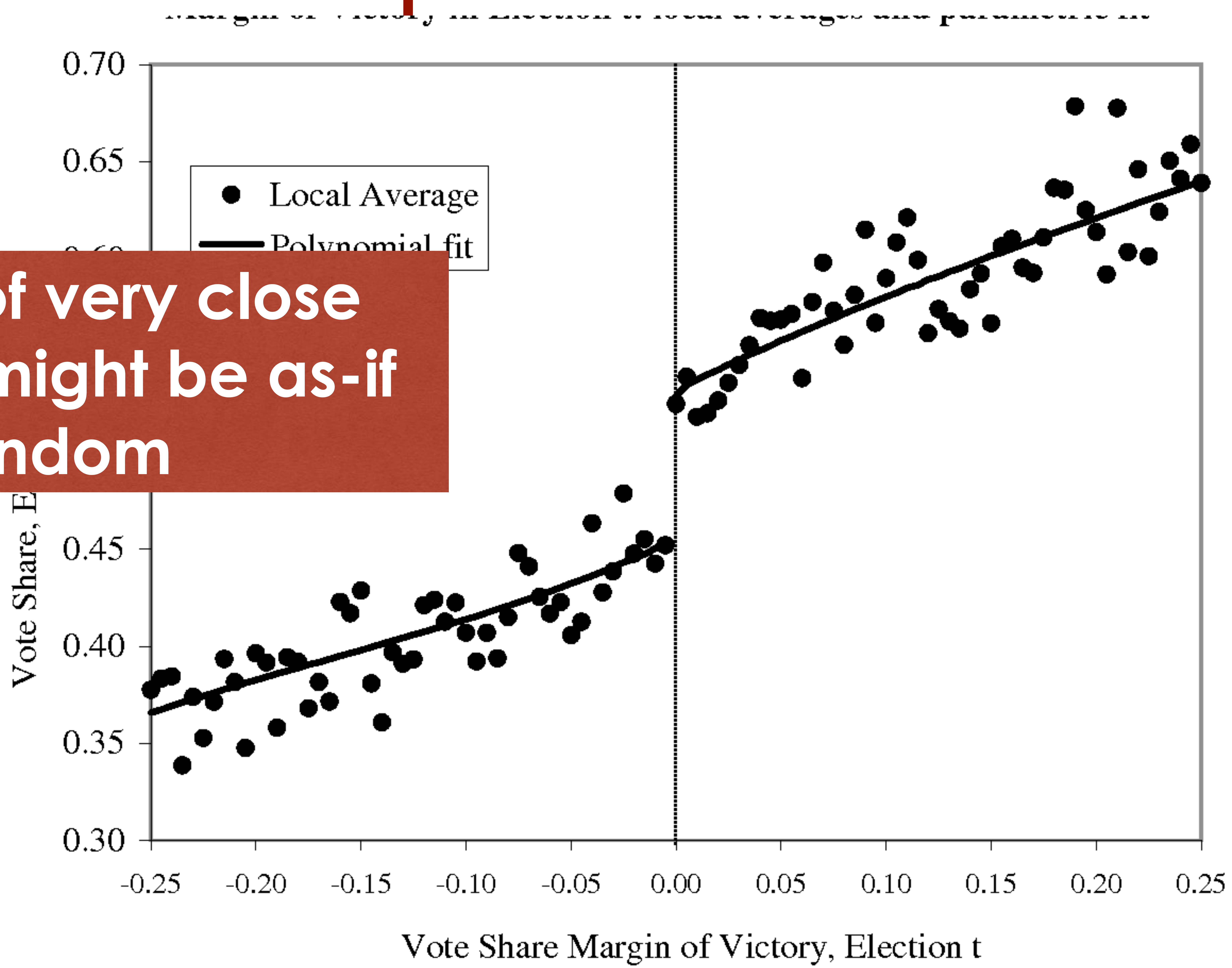
Other examples: age cutoffs

Figure 2: **Demobilization rates by age (centered at eligibility cut-point of 45.5 years old)**



Being very close to retirement age on either side

Other examples: close elections



Winner of very close election might be as-if random

Regression discontinuity

Benefits:
Mimics randomized experiment, close all backdoors

Remember:
In a true experiment there are no backdoors since nothing “causes” treatment assignment

Application:
Discontinuities are everywhere: anytime a law, policy, or rule has to arbitrarily separate people/land into groups

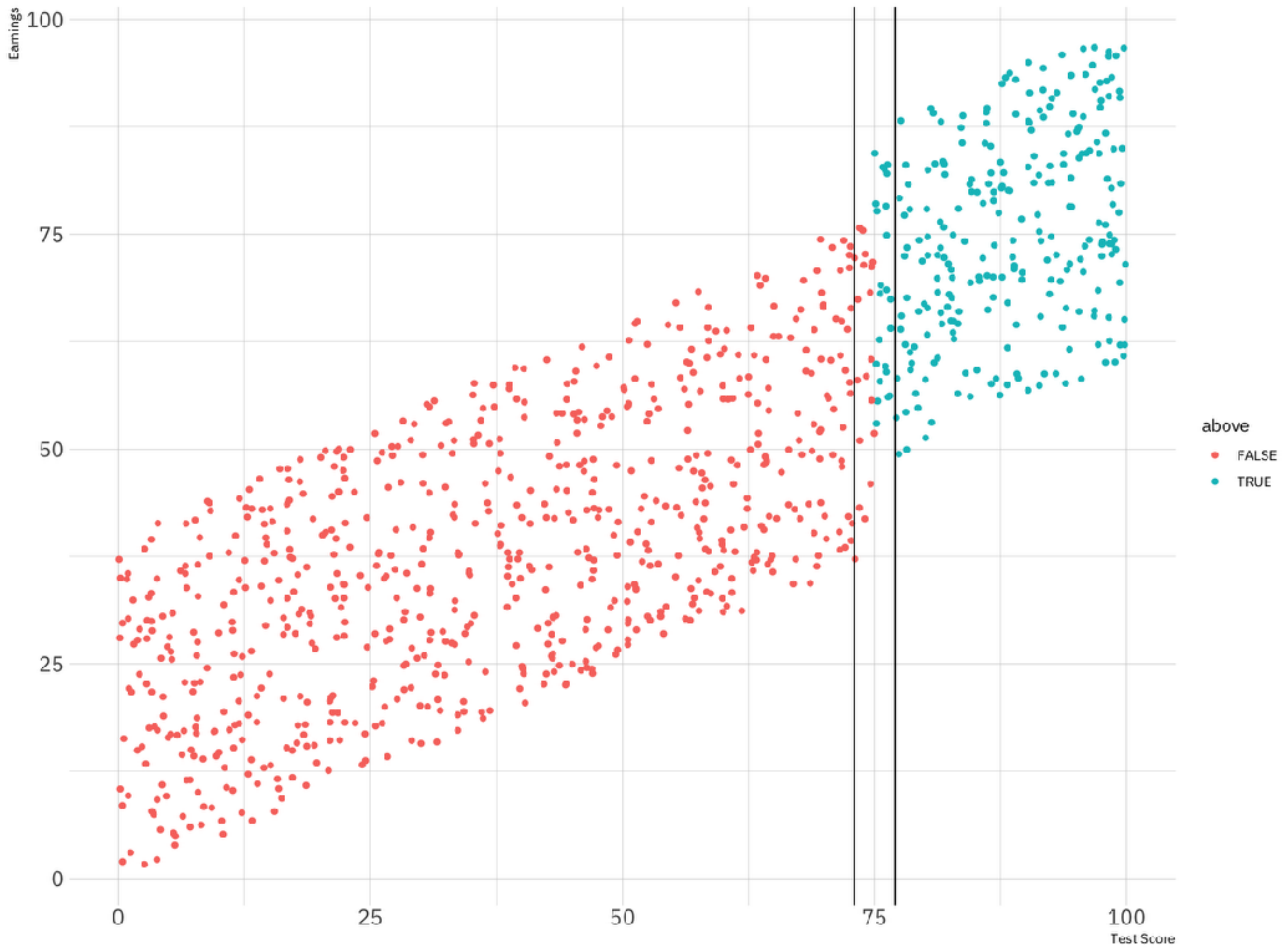
Limitations

You throw away a lot of data

Your estimate is "local"

~~What is the effect of being gifted on earnings?~~

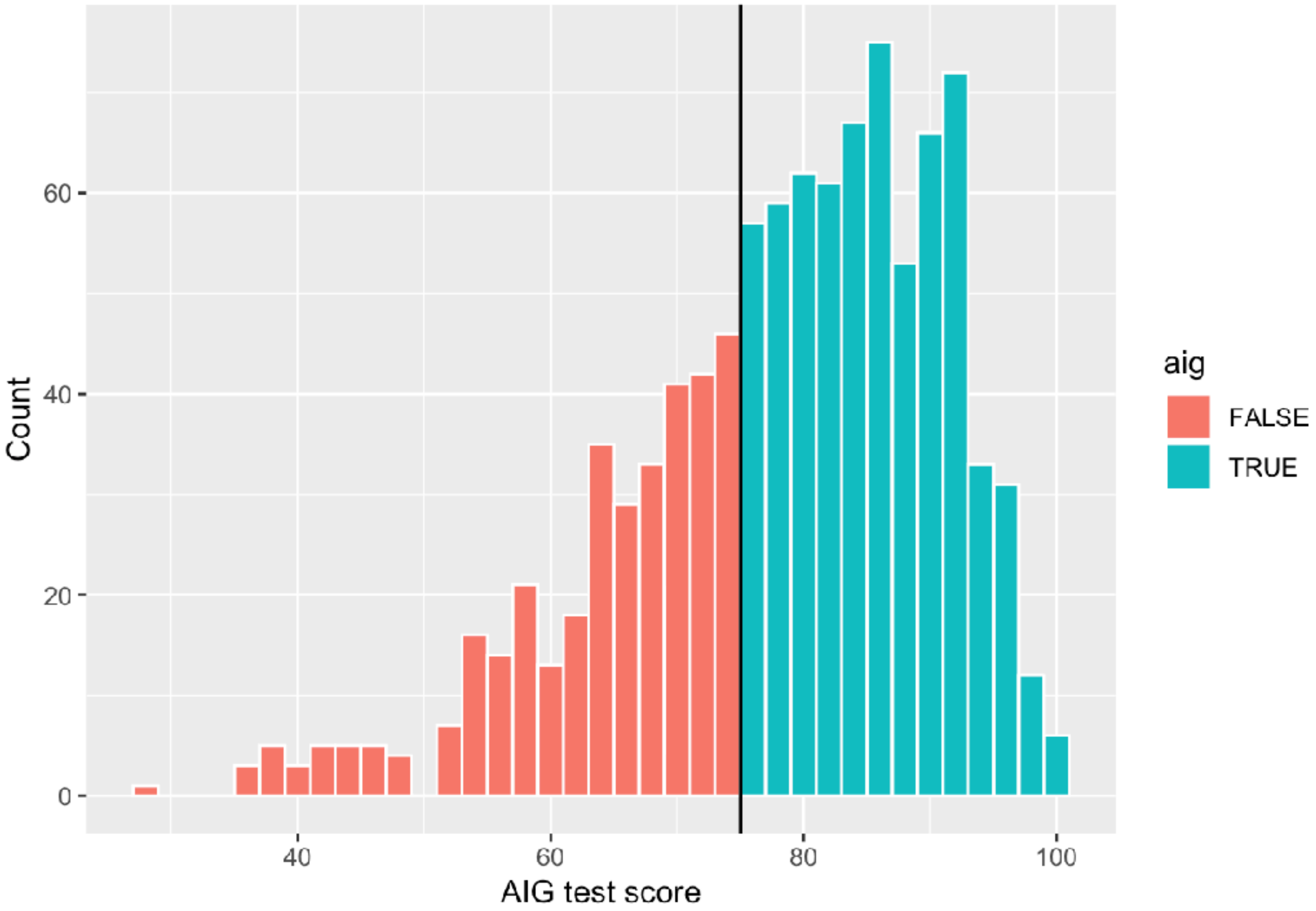
What is the effect of being gifted on earnings among students +/- 2 75 IQ?



Limitations

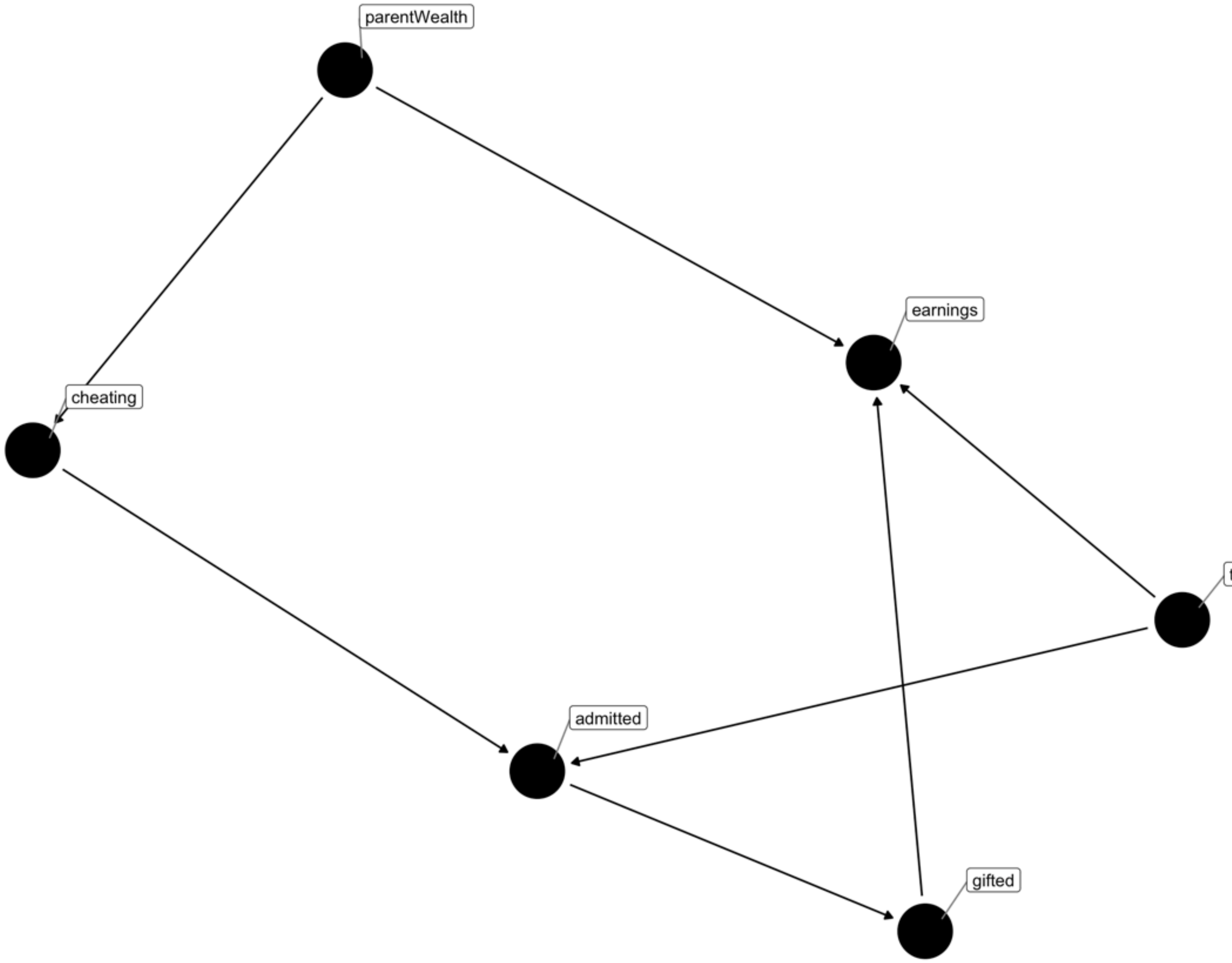
“Sorting around the threshhold”

People know the rule exists and are able to manipulate running variable



Example

Parents know 75 is the cutoff; wealthier/better-connected parents able to whine their way to gifted



In-class example

Go on website

Regression discontinuity:
Design that takes advantage of *natural experiment* created by *arbitrary* nature of rules/laws/policies

Approach:
Compare units right around cutoff; units nearby got "as-if" random assignment to treatment. Smaller bandwidths —> more convincing results

Limitations:
You throw away a lot of data and might run out! Units might sort around discontinuity

We can try to block backdoors	**Controls**
Select similar observations	**Matching**
Use a group at different times as its own control	**Fixed Effects**
When treatment kicks in at particular time, select reasonable control to compare	**Difference-in-differences**
When treatment is assigned by rule, select observations close to cutoff	**Regression Discontinuity**

So far…

How can we get close to true effect of X on Y?

How can we close backdoors, avoid colliders between X and Y?

Now…

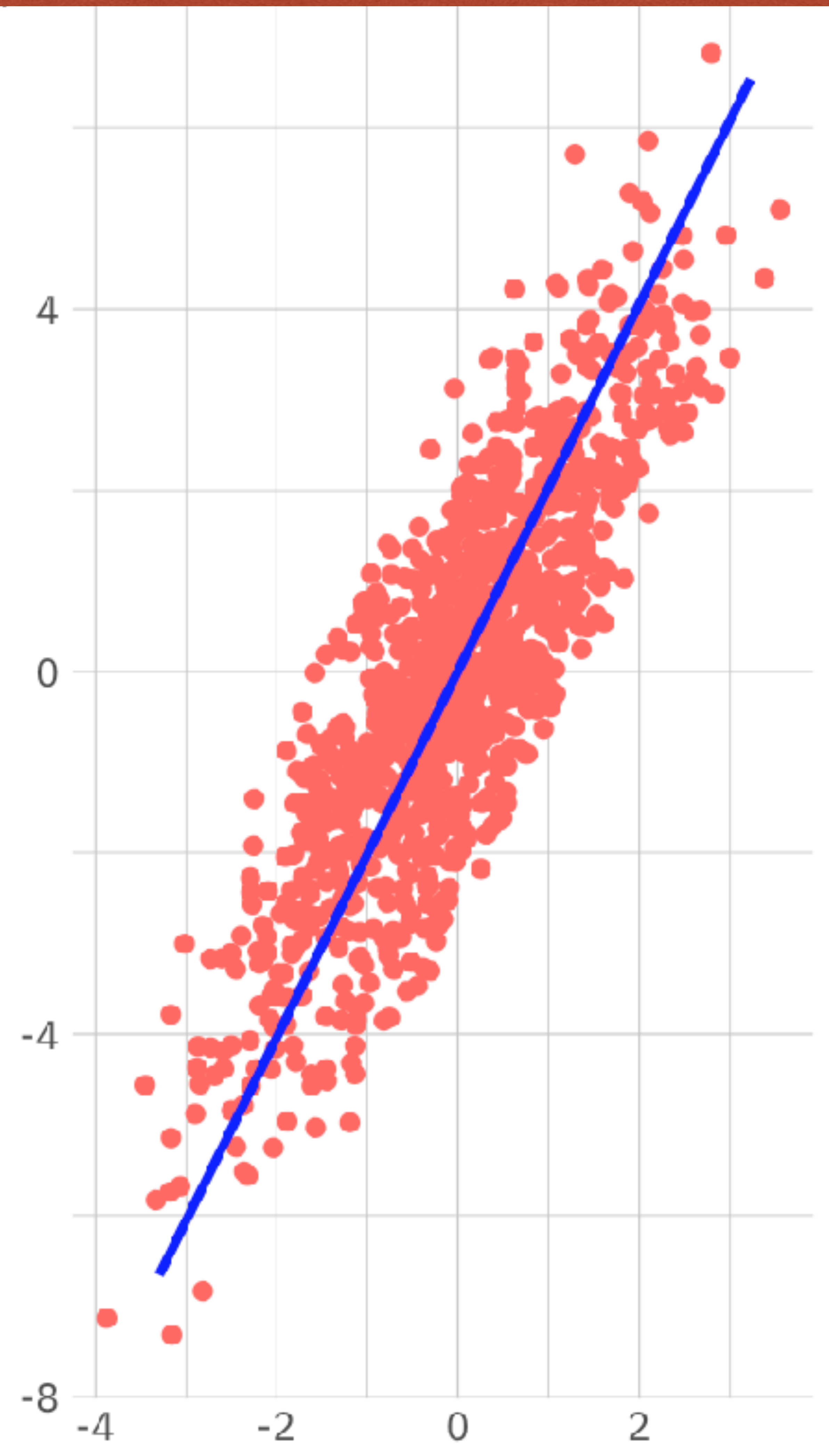
How confident are we in our estimate?

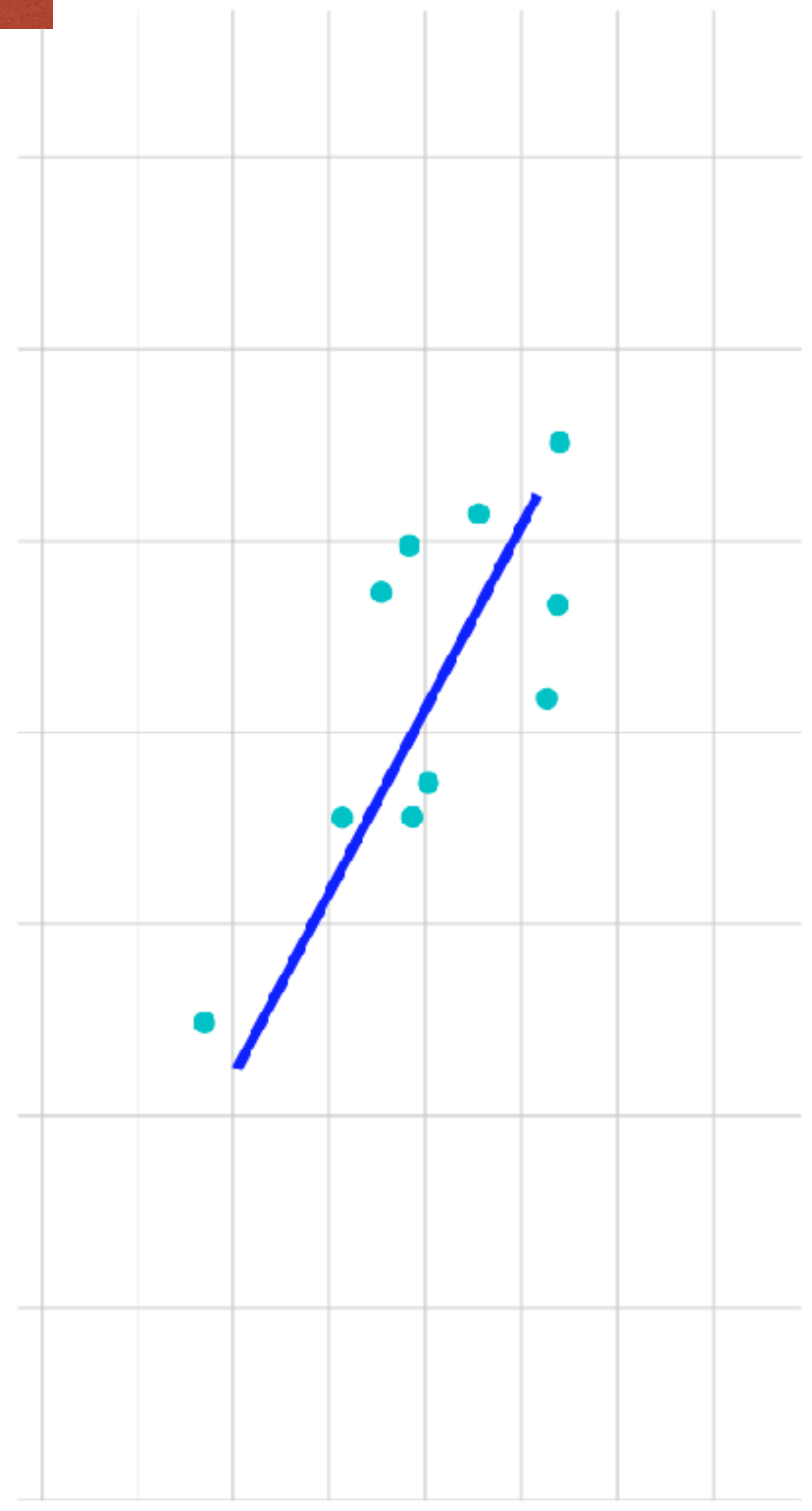
How likely is our estimate to replicate?

```
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.59846    0.02993  19.993   <2e-16 ***
## partners_rc1     0.02260    0.03528   0.640   0.5220
## partners_rc2     0.15869    0.06489   2.445   0.0146 *
## partners_rc3     0.03117    0.09742   0.320   0.7490
## partners_rc4+    0.15710    0.07780   2.019   0.0437 *
## ---
```

What are these other numbers? And the stars?

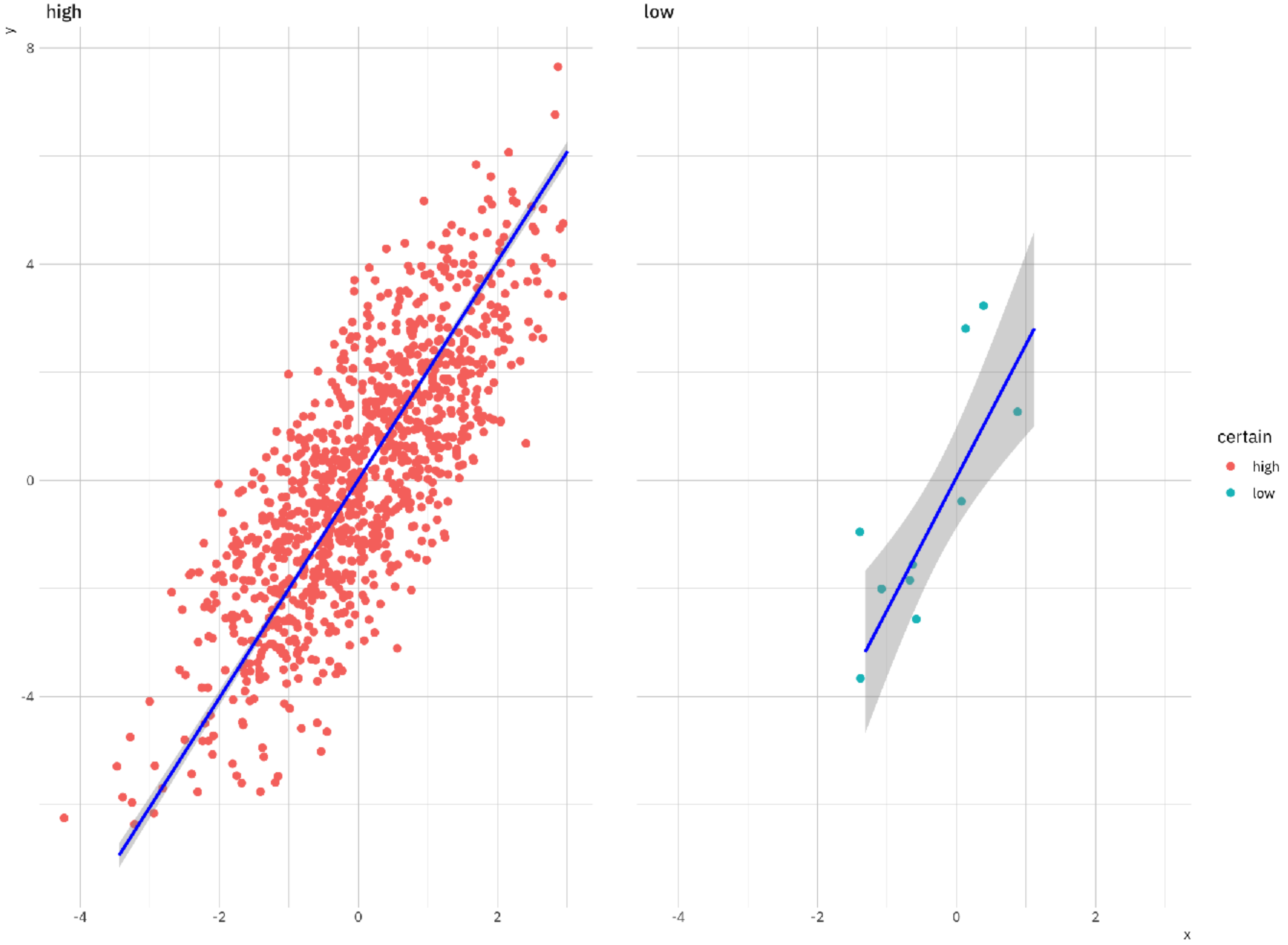
$$\mathbf{y} = 0.01 + 2.02(\mathbf{x}) + \epsilon$$





$$\mathbf{y} = 0.24 + 1.91(\mathbf{x}) + \epsilon$$

Intuitively:
There is something qualitatively different about the two plots

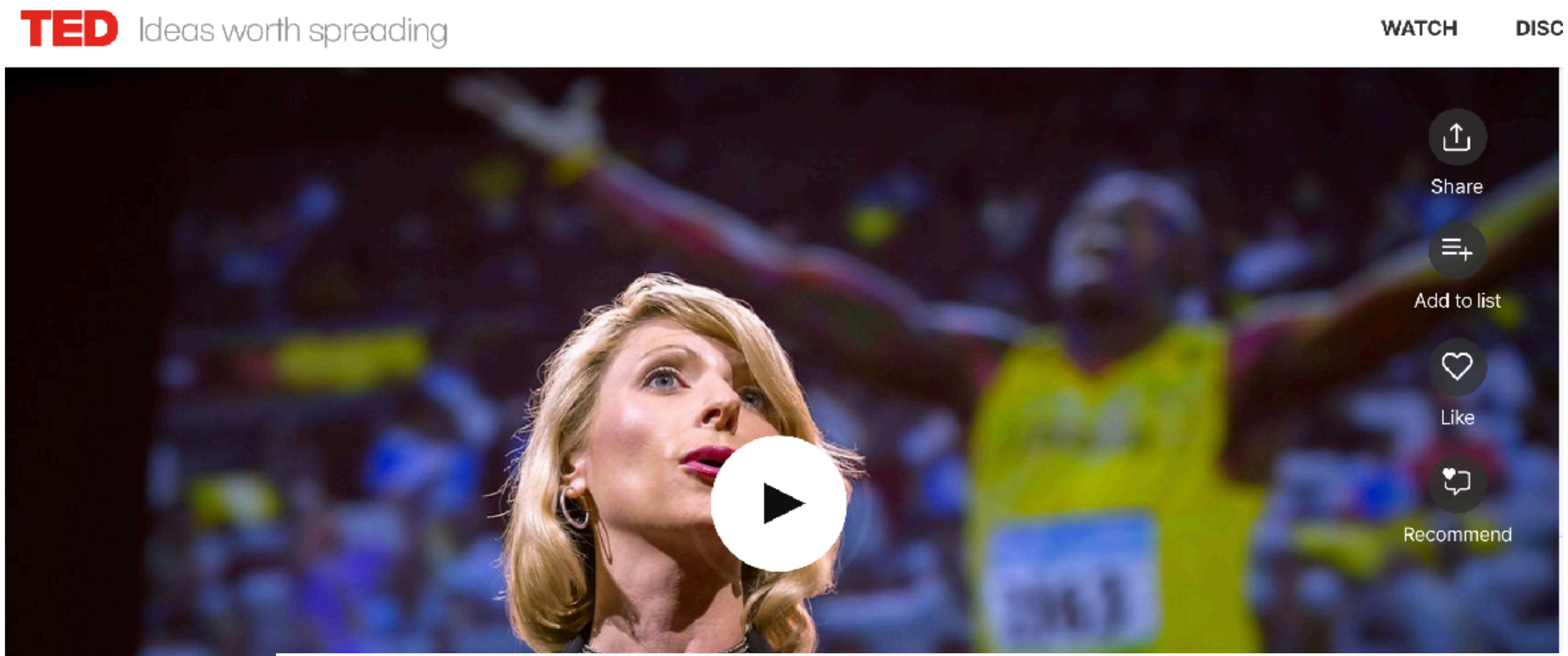

high
low
y
8
4
0
-4
-4
-2
0
2
x
certain
high
low

```
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.59846    0.02993  19.993   <2e-16 ***
## partners_rc1    0.02260    0.03528   0.640   0.5220
## partners_rc2    0.15869    0.06489   2.445   0.0146 *
## partners_rc3    0.03117    0.09742   0.320   0.7490
## partners_rc4+   0.15710    0.07780   2.019   0.0437 *
## ---
```

These numbers describe our certainty about our estimate

Normally regression and uncertainty are taught hand-in-hand

Why did I put this at the end of class?

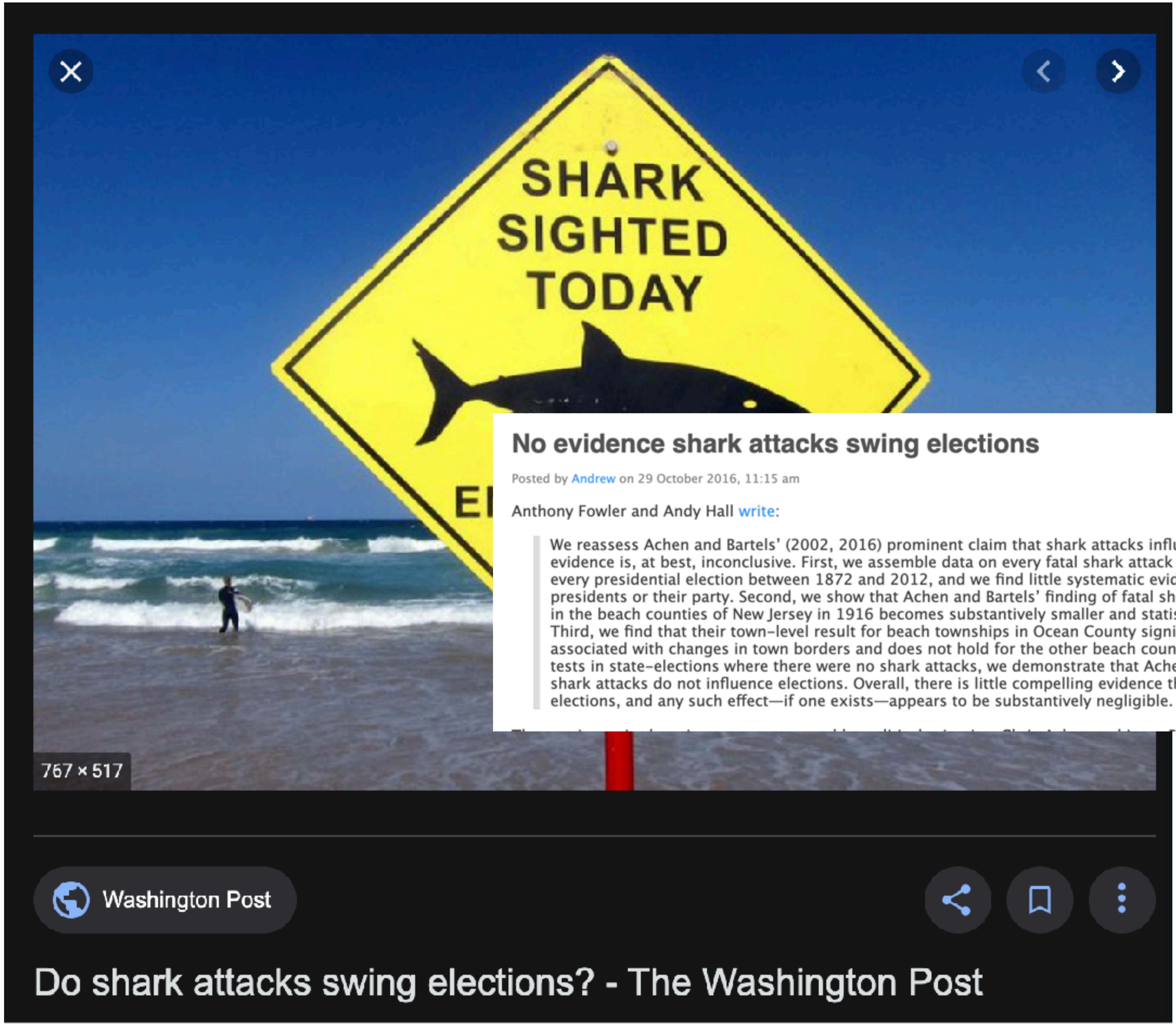


My position on "Power Poses"

Regarding: Carney, Cuddy & Yap (2010).

Reasonable people, whom I respect, may disagree. However since early 2015 the evidence has been mounting suggesting there is unlikely any embodied effect of nonverbal expansiveness (vs. contractiveness)—i.e.., "power poses" -- on internal or psychological outcomes.

As evidence has come in over these past 2+ years, my views have updated to reflect the evidence. As such, **I do not believe that "power pose" effects are real.**



No evidence shark attacks swing elections

Posted by Andrew on 29 October 2016, 11:15 am

Anthony Fowler and Andy Hall write:

> We reassess Achen and Bartels' (2002, 2016) prominent claim that shark attacks influence presidential elections, and we find that the evidence is, at best, inconclusive. First, we assemble data on every fatal shark attack in U.S. history and county-level returns from every presidential election between 1872 and 2012, and we find little systematic evidence that shark attacks hurt incumbent presidents or their party. Second, we show that Achen and Bartels' finding of fatal shark attacks hurting Woodrow Wilson's vote share in the beach counties of New Jersey in 1916 becomes substantively smaller and statistically weaker under alternative specifications. Third, we find that their town-level result for beach townships in Ocean County significantly shrinks when we correct errors associated with changes in town borders and does not hold for the other beach counties in New Jersey. Lastly, implementing placebo tests in state-elections where there were no shark attacks, we demonstrate that Achen and Bartels' result was likely to arise even if shark attacks do not influence elections. Overall, there is little compelling evidence that shark attacks influence presidential elections, and any such effect—if one exists—appears to be substantively negligible.


SCIENCE
Many Psychology Findings Not as Strong as Claimed, Study Says
The New York Times
Science
AAAS
nature
Over half of psychology studies fail reproducibility test
Vox
SCIENCE & HEALTH
Scientists replicated 100 recent psychology experiments. More than half of them failed.

The replication crisis

Wave of studies in social and health sciences fail to replicate

Replication: someone, later, performs a very similar study to confirm effect

We think this has something to do with how researchers think about uncertainty

"Big if true"

Basically: it is much easier to get your study published if you can show an effect is highly certain

~~5,000 people study, find no effect of X on Y~~

5,000 people study, find effect of X on Y

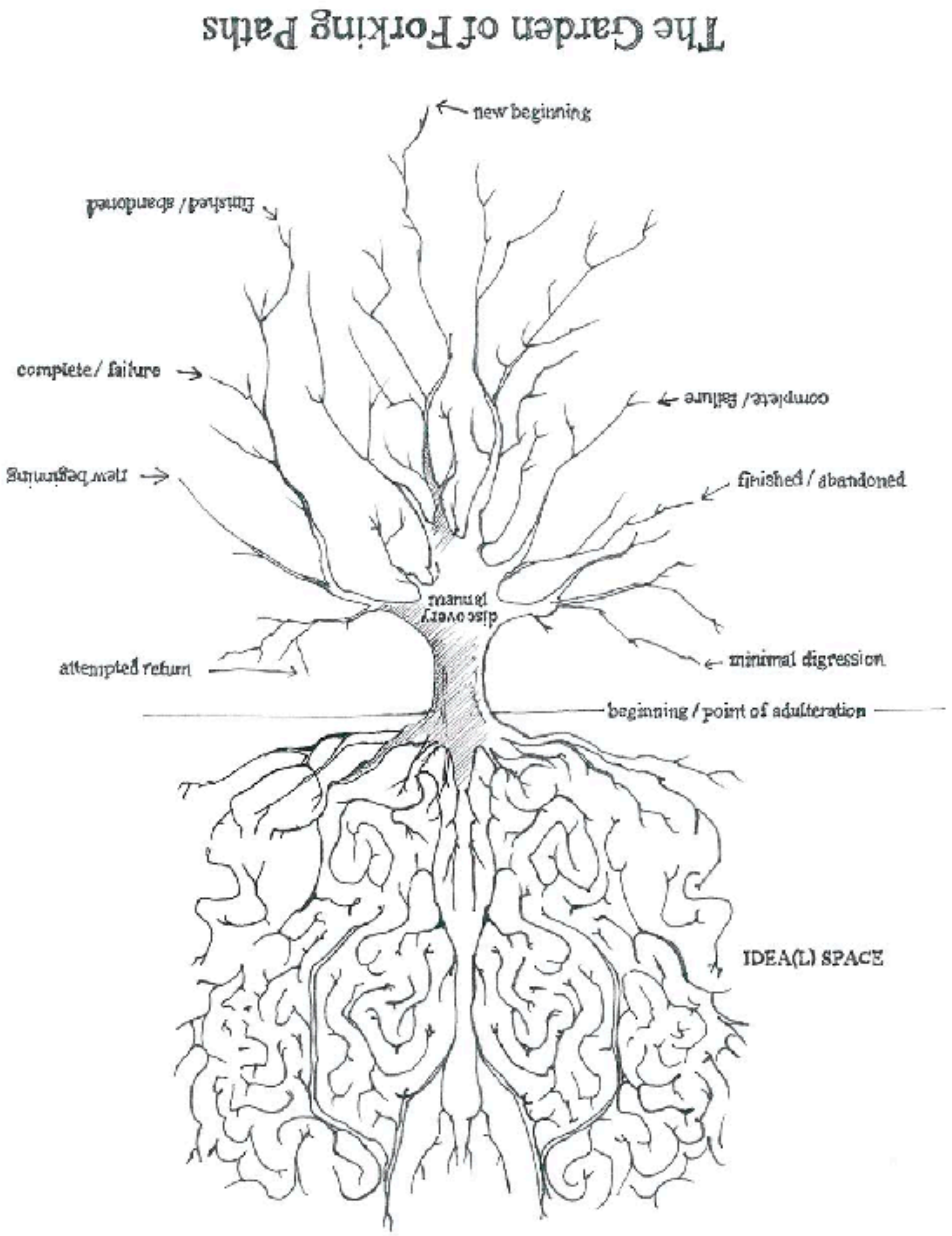
The garden of forking paths

This means researchers often make analysis decisions with this in mind

"No effect… hm. Well, what if the effect is only true among women?"

"Still nothing… hm. Maybe it depends on how sick you were to begin with?"

Might sound nefarious/bad faith, but in most cases it's not!



The Garden of Forking Paths

It’s complicated

A lot of moving pieces in the “replication crisis”

But part of it is that we **fetishize** uncertainty measures

We’re trained to run regressions and hone in on significant results (“**star-gazing**”)

So I separate them in this class; one thing is **how** to get the right answer, another is how **certain** we are of that answer

We rarely have **all the data** on the thing we care about

state	hourslw	lwage	year	after	miami	class	person
93	40	2.37	82	TRUE	FALSE	1	Pariz
93	35	1.95	83	TRUE	FALSE	1	Vimala
93	40	1.11	81	TRUE	FALSE	1	Fardosa
93	40	1.16	84	TRUE	FALSE	1	Virgiline
58	40	1.53	82	TRUE	FALSE	1	Hopie
59	40	2.24	79	FALSE	TRUE	2	Bethann
59	48	1.75	81	TRUE	TRUE	1	Haliey
59	42	2.14	79	FALSE	TRUE	1	Olawale
59			83	TRUE	TRUE	1	Barnetta
59	45		80	FALSE	TRUE	1	Krysteen

We instead have a **sample**

Problem:

Different samples will produce different **estimates**

```
gss_sm %>%
  sample_n(10) %>%
  lm(obama ~ sex, data = .) %>%
  get_regression_table()
```

Run this, tell me what you get

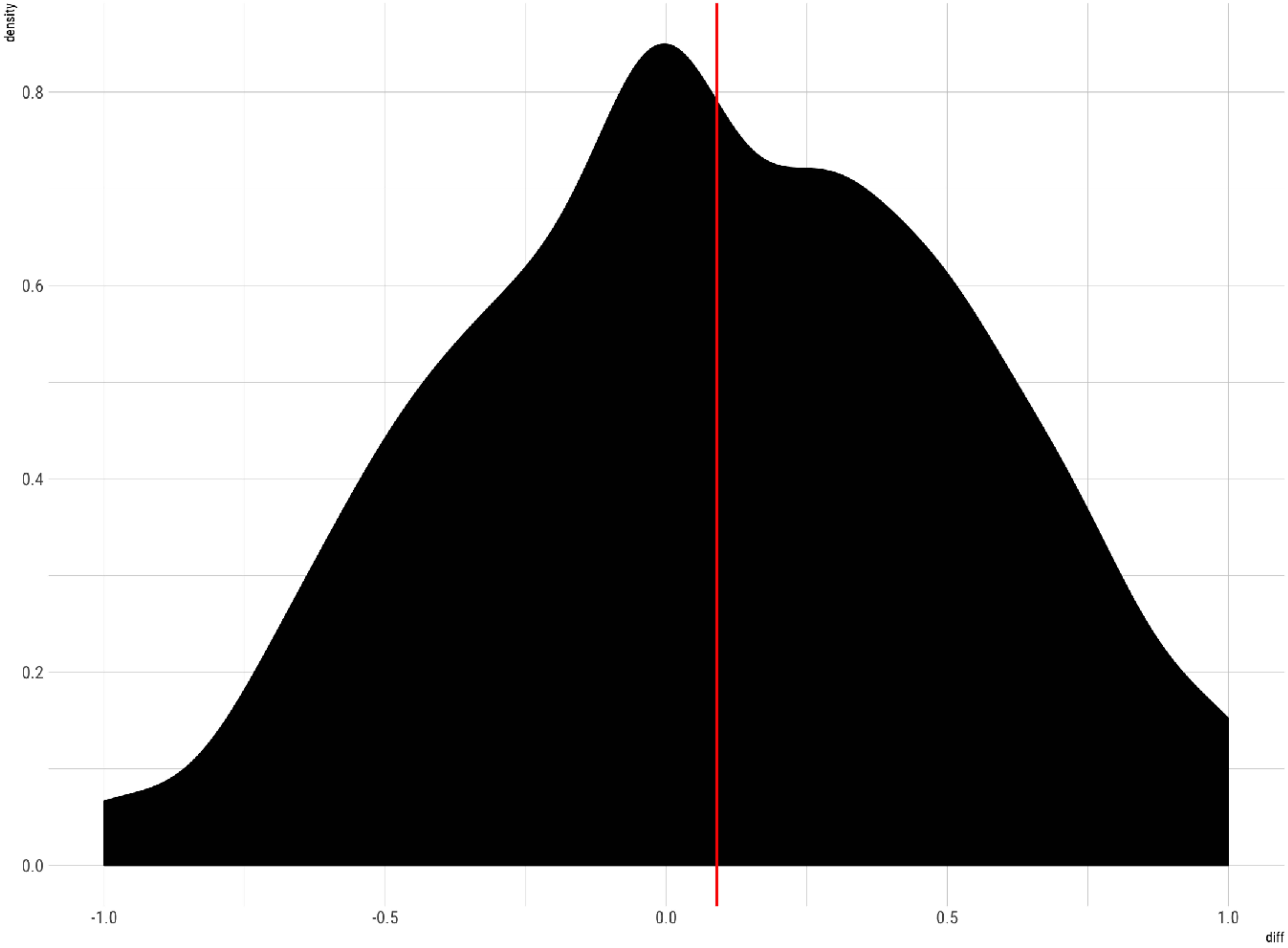
The truth:

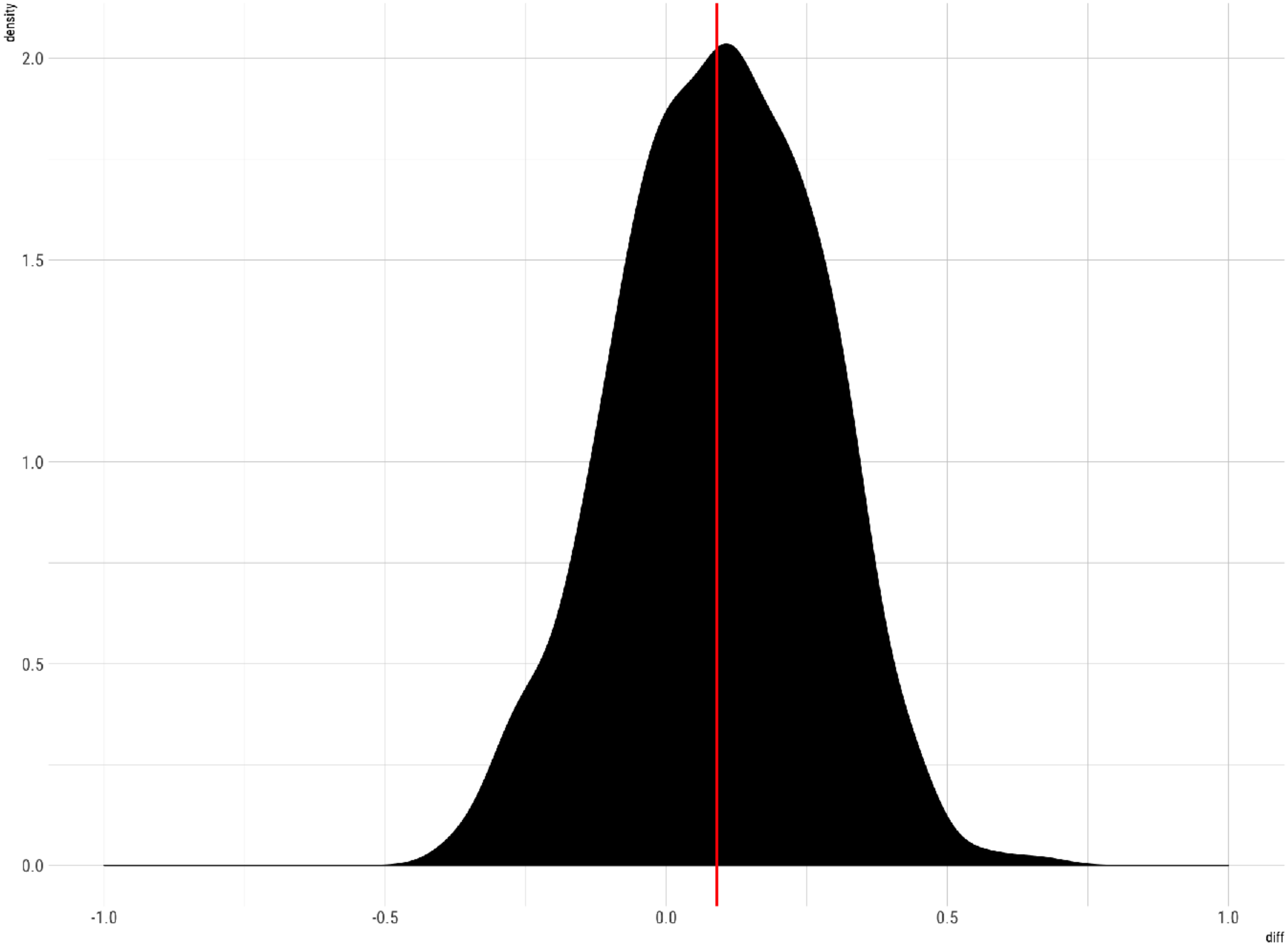
Intercept = .58

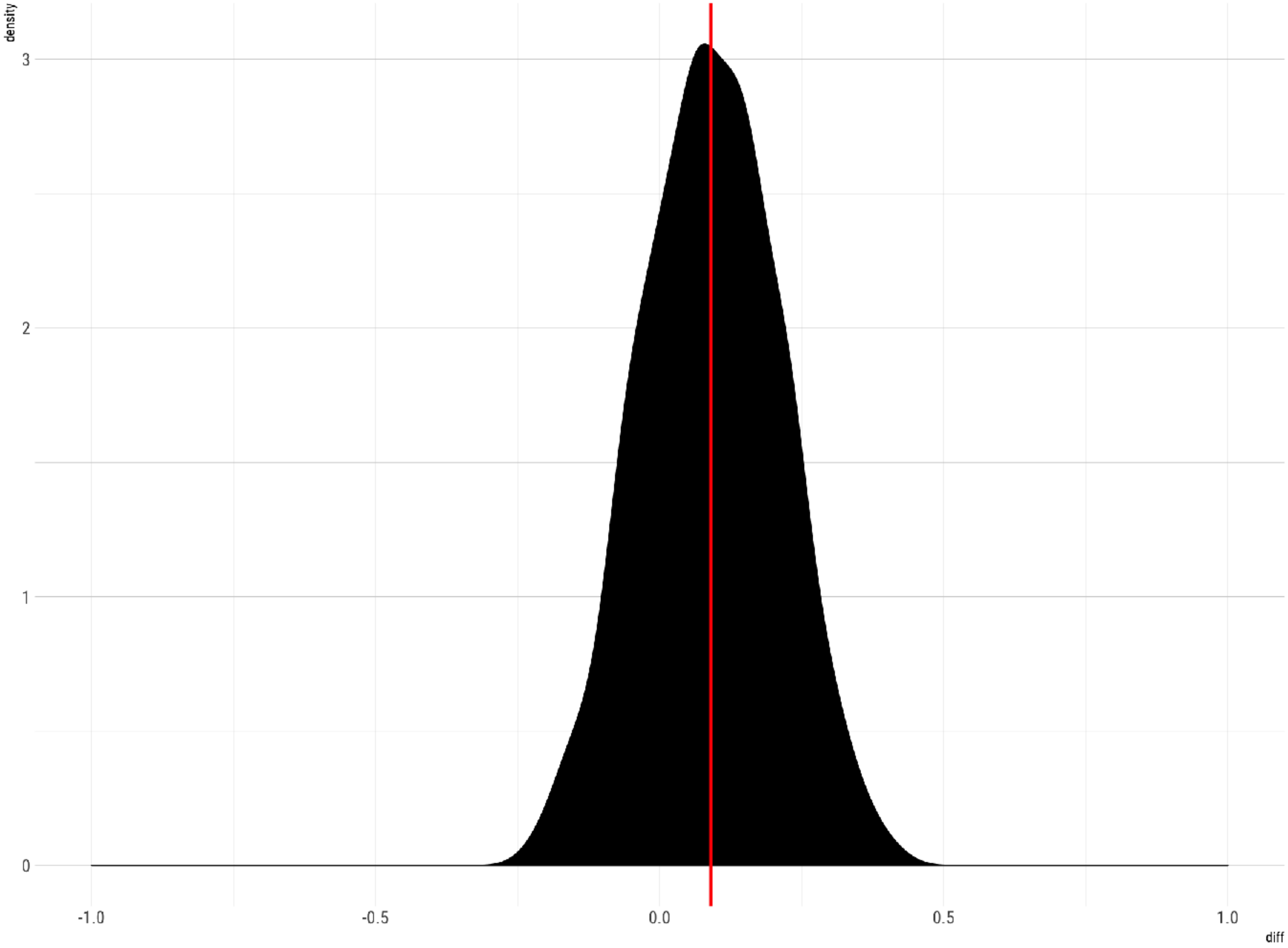
Female = .09

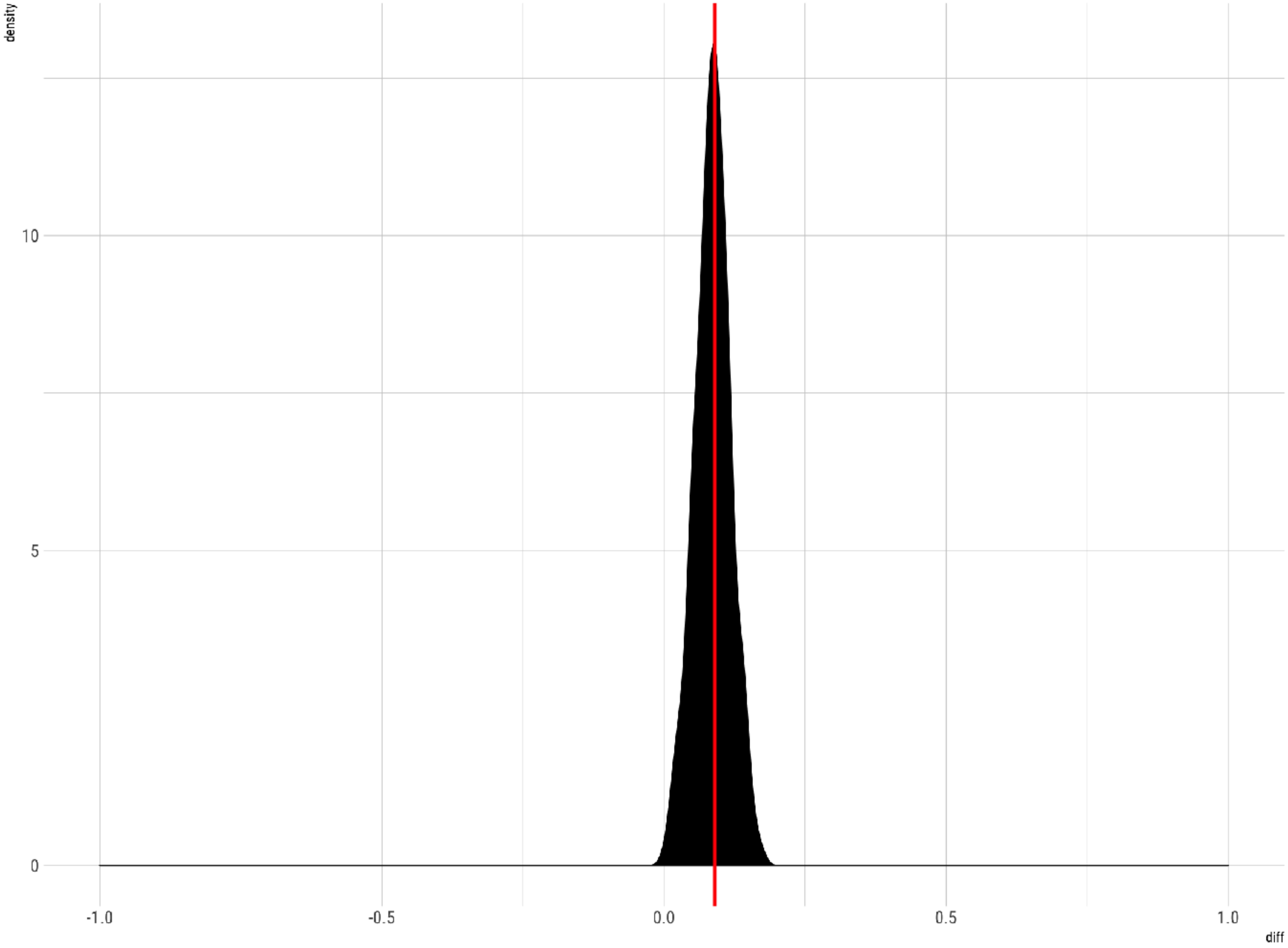
Table 1

replicate	sexFemale
1	-0.167
2	-0.167
3	-0.25
4	0.0833
5	0
6	-0.1
7	-0.4
8	-0.167
9	1
10	0


density
0.8
0.6
0.4
0.2
0.0
-1.0
-0.5
0.0
0.5
1.0
diff

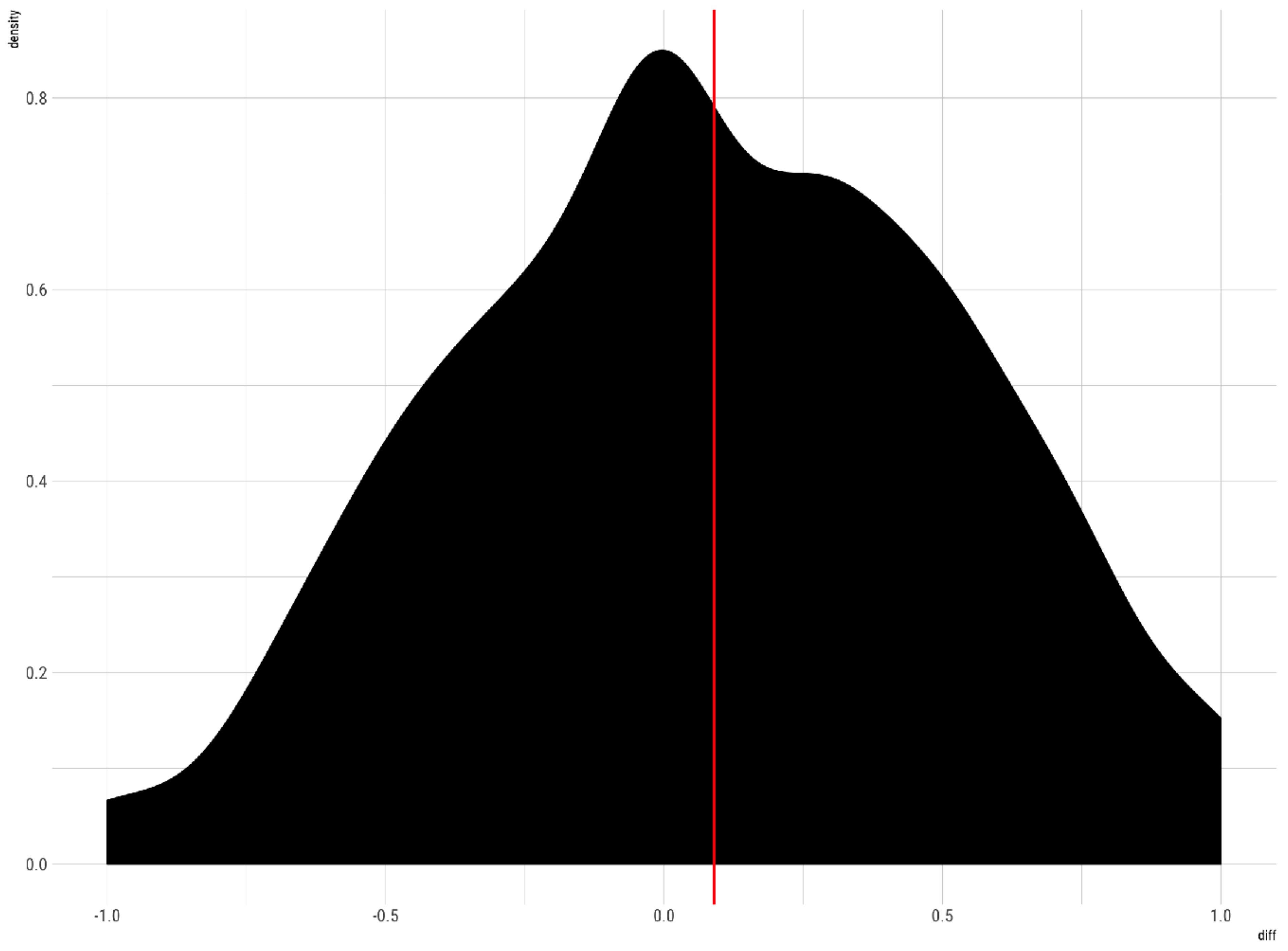

density
2.0
1.5
1.0
0.5
0.0
-1.0
-0.5
0.0
0.5
1.0
diff

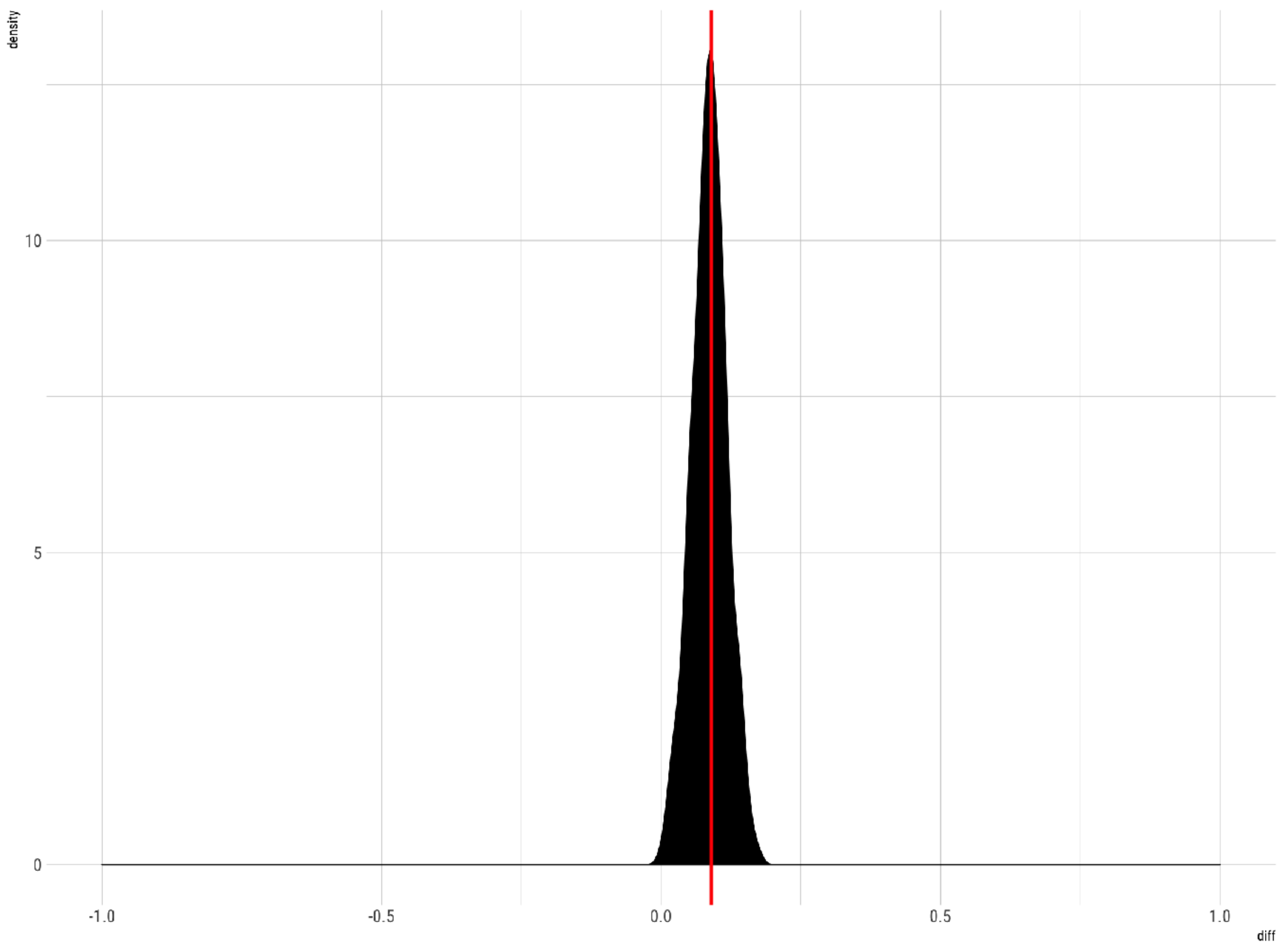

density
3
2
1
0
-1.0
-0.5
0.0
0.5
1.0
diff


density
10
5
0
-1.0
-0.5
0.0
0.5
1.0
diff

Detour on variation

Lot more variation on left than right; how to measure this?





Standard deviation

Measures the spread of the data around its mean

$$Standard\,deviation = \sqrt{\frac{(x_1 - Mean)^2 + (x_2 - Mean)^2 + \cdots + (x_n - Mean)^2}{n-1}}$$

The bigger the standard deviation the greater the spread

```r
# variation, standard deviation
county_data %>%
  group_by(state) %>%
  select(per_gop_2012) %>%
  summarise(mean = mean(per_gop_2012, na.rm = TRUE),
            sd = sd(per_gop_2012, na.rm = TRUE))
```

Uncertainty

The more data we have the closer we get to the truth and the less uncertain we are

Even if we're not 100% right, we'll be less wrong with more data!

Vocabulary

Population

All of the instances of the thing we care about

Population paramater

The true difference in the Obama's vote share btwn men/women

Sample

A subset of the population

Sample Statistic

The difference in each of those 10 random draws

Good samples

There are good and bad samples in the world

A good sample is **representative** of the population and **unbiased**

What does this mean?

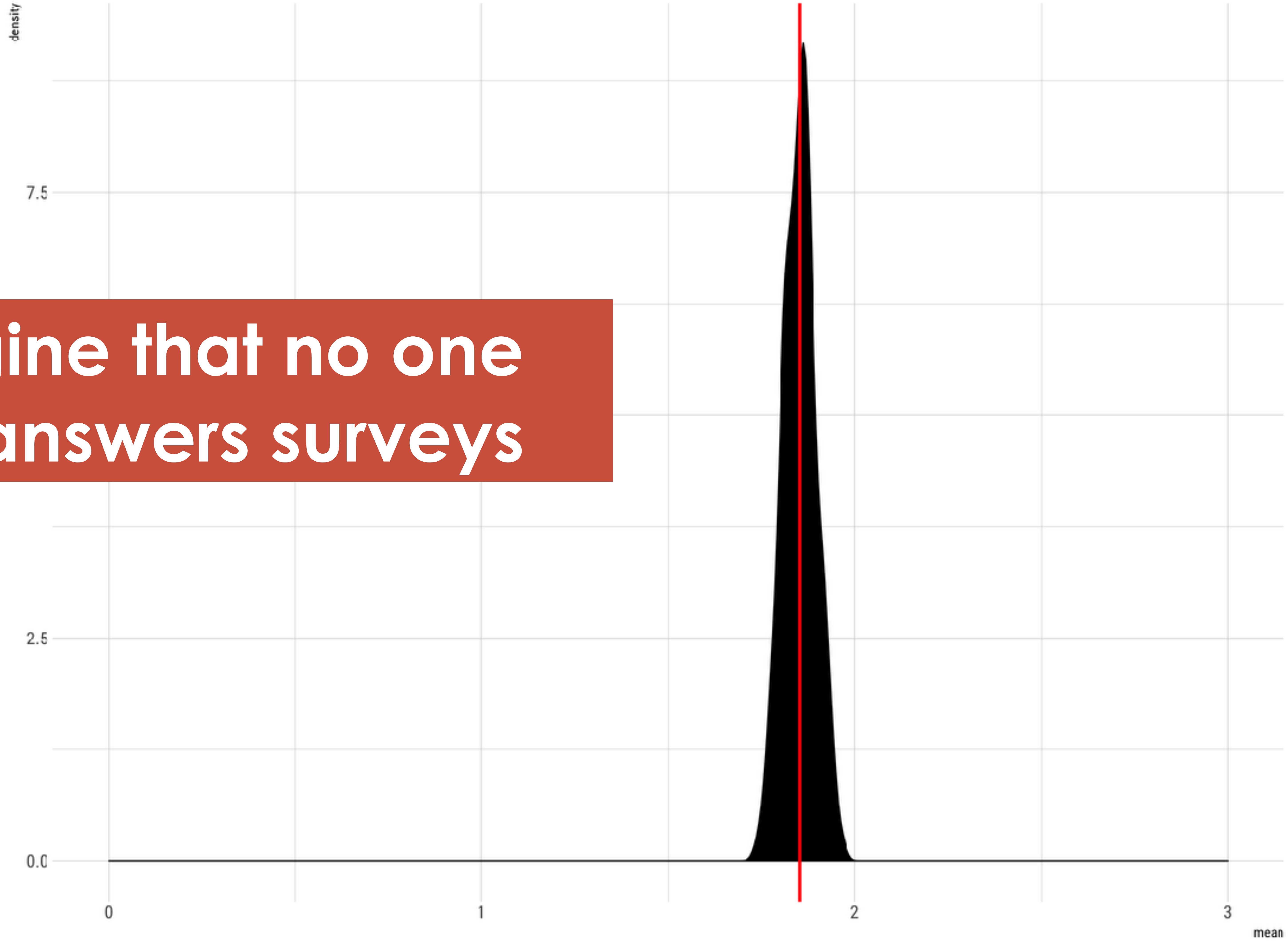
Average number of kids

Find average number of kids in survey

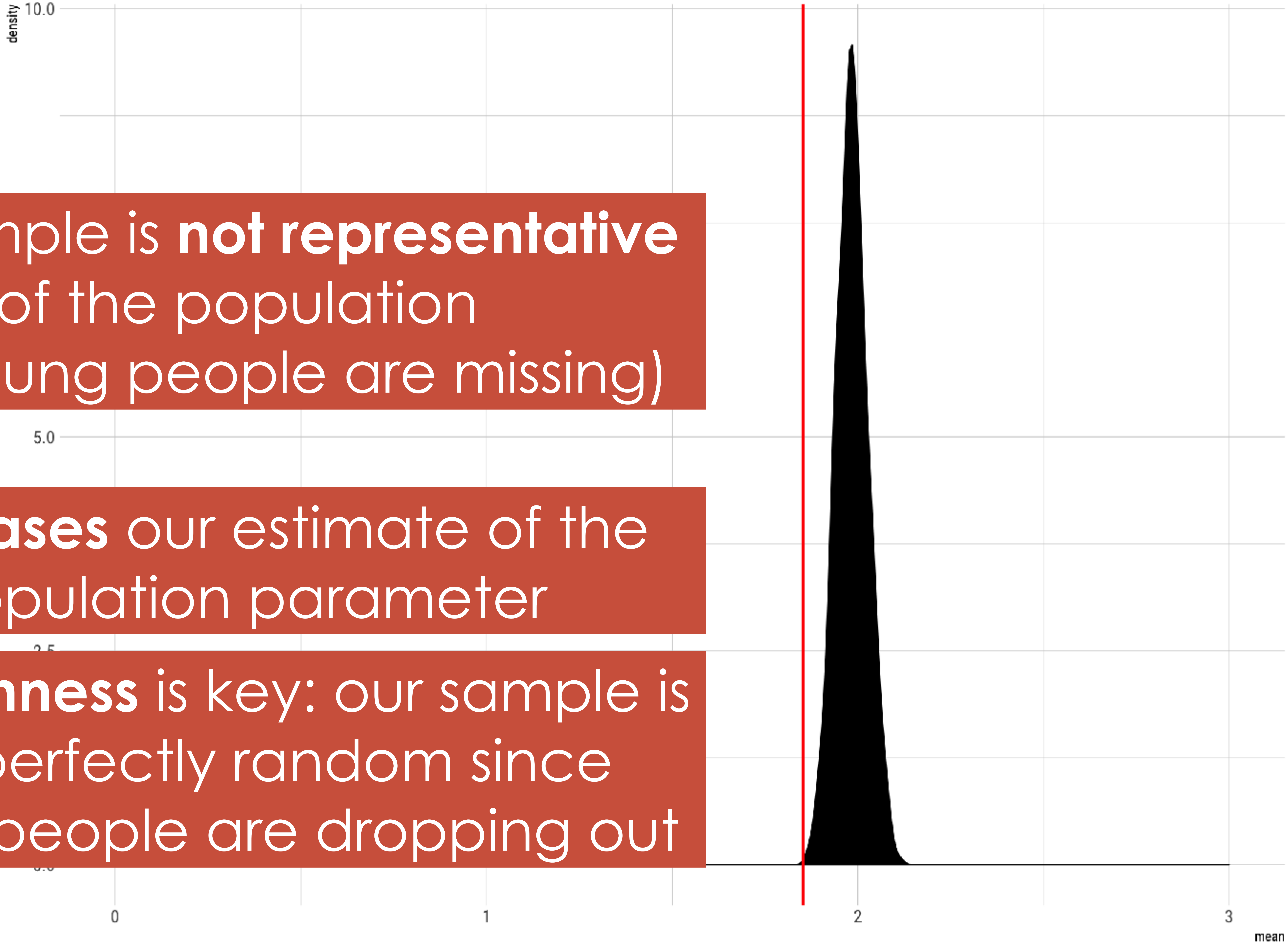
Take samples with a big size and big n, and draw histogram

Now, make it so that younger people are less sampled and redraw histogram

Average number of kids



Now imagine that no one under 25 answers surveys



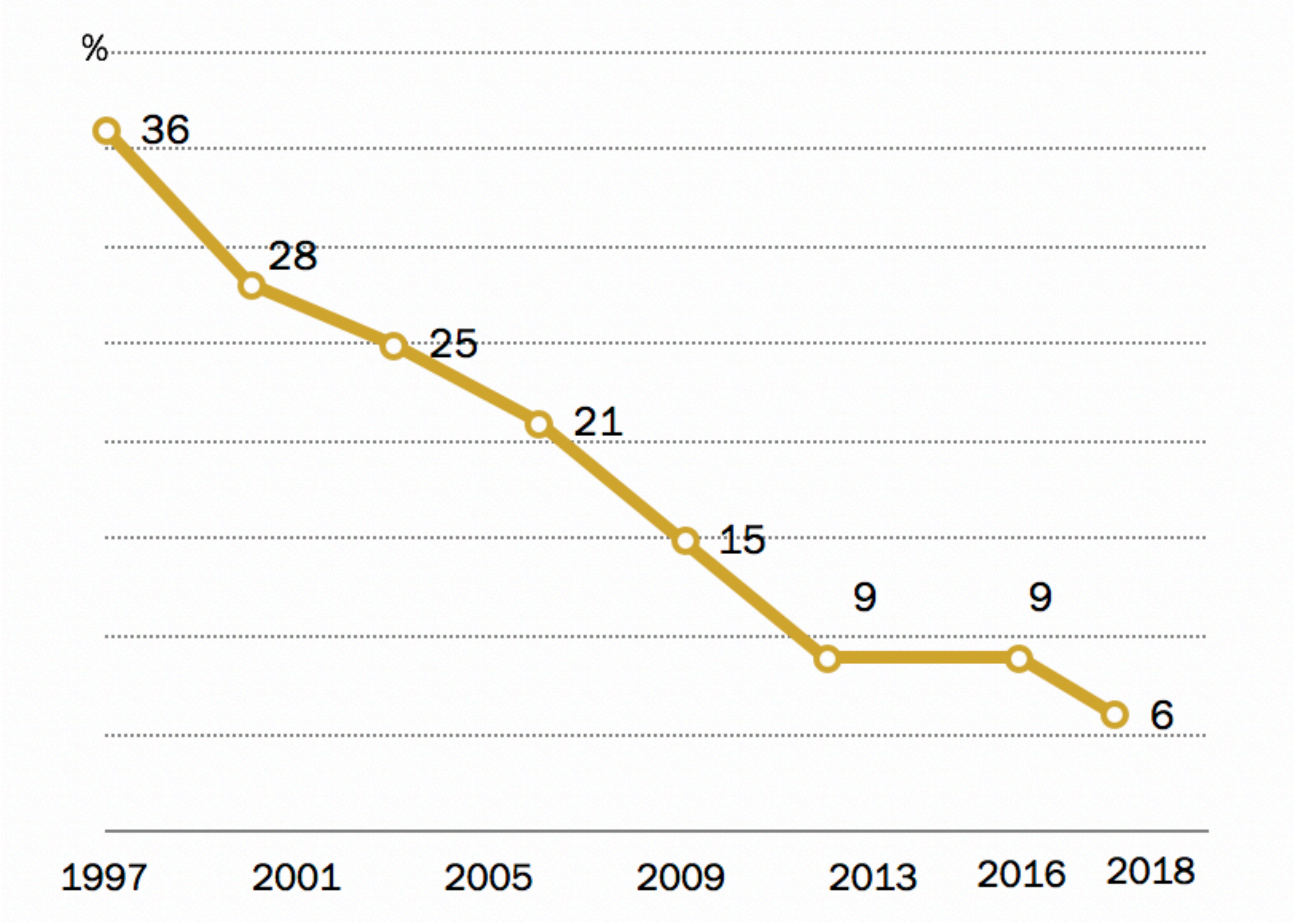
The sample is **not representative** of the population (the young people are missing)

This **biases** our estimate of the population parameter

Randomness is key: our sample is not perfectly random since young people are dropping out

After brief plateau, telephone survey response rates have fallen again

Response rate by year (%)



Note: Response rate is AAPOR RR3. Only landlines sampled 1997-2006. Rates are typical for surveys conducted in each year.
Source: Pew Research Center telephone surveys conducted 1997-2018.

PEW RESEARCH CENTER

What about when we have all the data?

Table 1

name		state		pop	female	white	black
Dickinson	County	MI		25957	50.1	96.9	0.5
Hart Count	y	KY		18597	50.6	93.4	4.7
Erie Count	y	NY		922835	51.6	80.7	13.9
Preble Cou	nty	OH		41586	50.4	97.5	0.6
Palm Beach	County	FL	1	397710	51.6	76.5	18.5

Electoral results from every county in the US: shouldn't we be perfectly certain?