In-person session 10

March 21, 2024

PMAP 8521: Program evaluation
Andrew Young School of Policy Studies

Plan for today

Diff-in-diff effect sizes

Miscellaneous R stuff

RDD fun times

Diff-in-diff effect sizes

What the heck is happening at the end of problem set 5?!

Miscellaneous R stuff

Searching past code

Learning with the example pages

Lines across categories

RDD fun times

Is there a rule of thumb to determine which quasi-experimental method we should use?

How do we know which method applies to which circumstance? Does the data tell us?

With RDD we rely on "the rule" to determine treatment and control groups

How do you decide on the rule?
You mentioned that it's arbitrary—
we can choose whatever rule we want?

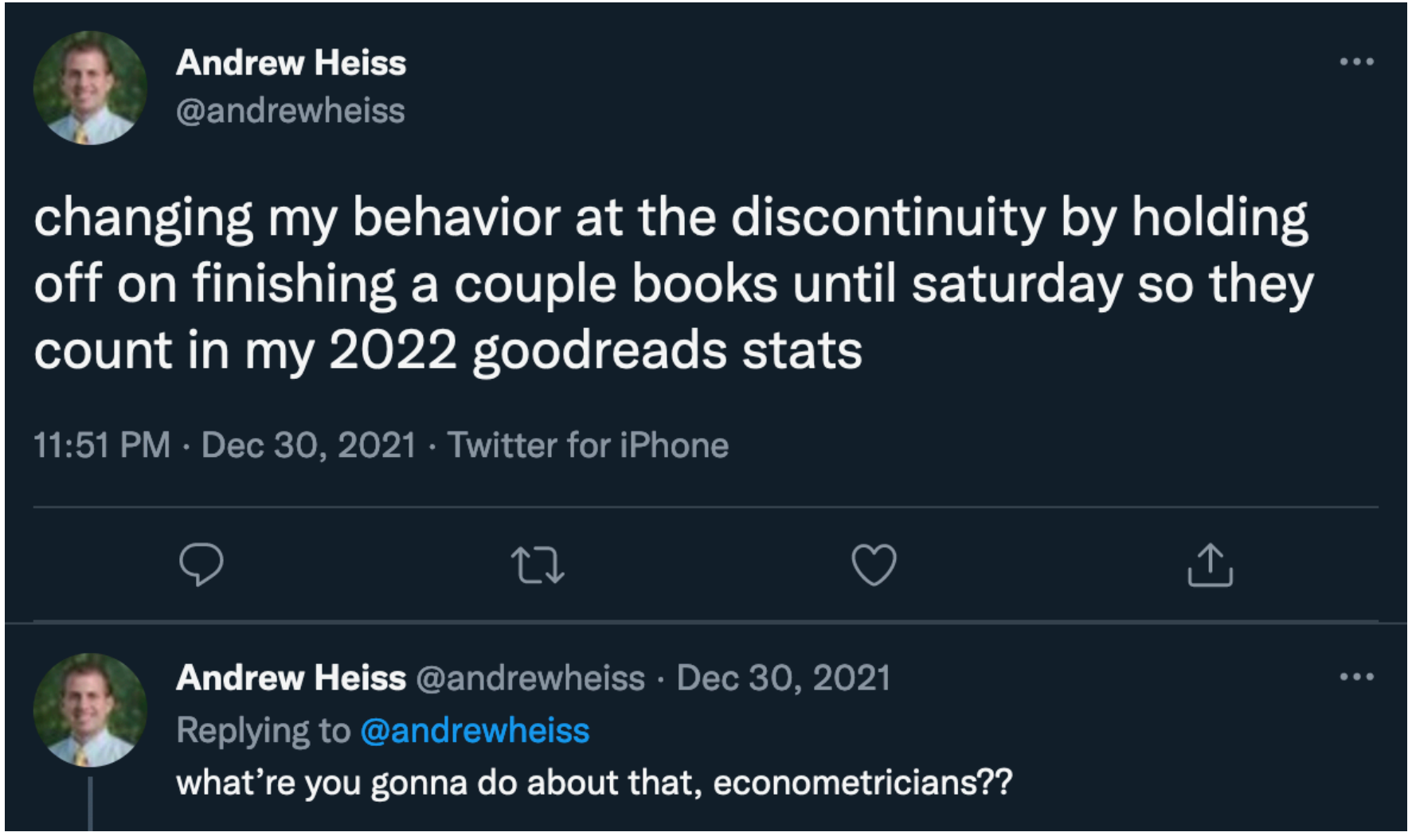
Can we use RDD to evaluate a program that doesn't have a rule for participation?

Can we use a binary running variable?

e.g. someone is eligible for a program if they complete a course

Do we have to limit the data to a bandwidth?

How common are these kinds of rules in the real world?


Andrew Heiss
@andrewheiss
changing my behavior at the discontinuity by holding off on finishing a couple books until saturday so they count in my 2022 goodreads stats
11:51 PM · Dec 30, 2021 · Twitter for iPhone
Andrew Heiss @andrewheiss · Dec 30, 2021
Replying to @andrewheiss
what’re you gonna do about that, econometricians??

Where do these eligibility thresholds come from? Do policy makers research them first and reexamine them later?

Discontinuities everywhere!

Size	Annual	Monthly	138%	150%	200%
1	$12,760	$1,063	$17,609	$19,140	$25,520
2	$17,240	$1,437	$23,791	$25,860	$34,480
3	$21,720	$1,810	$29,974	$32,580	$43,440
4	$26,200	$2,183	$36,156	$39,300	$52,400
5	$30,680	$2,557	$42,338	$46,020	$61,360
6	$35,160	$2,930	$48,521	$52,740	$70,320
7	$39,640	$3,303	$54,703	$59,460	$79,280
8	$44,120	$3,677	$60,886	$66,180	$88,240

Medicaid
138%*

ACA subsidies
138–400%*

CHIP
200%

SNAP/Free lunch
130%

Reduced lunch
130–185%

The US's official poverty measure

Mollie Orshansky

Formula created in 1963

Based solely on food expenses from a survey of household budgets in 1955

The US's official poverty measure

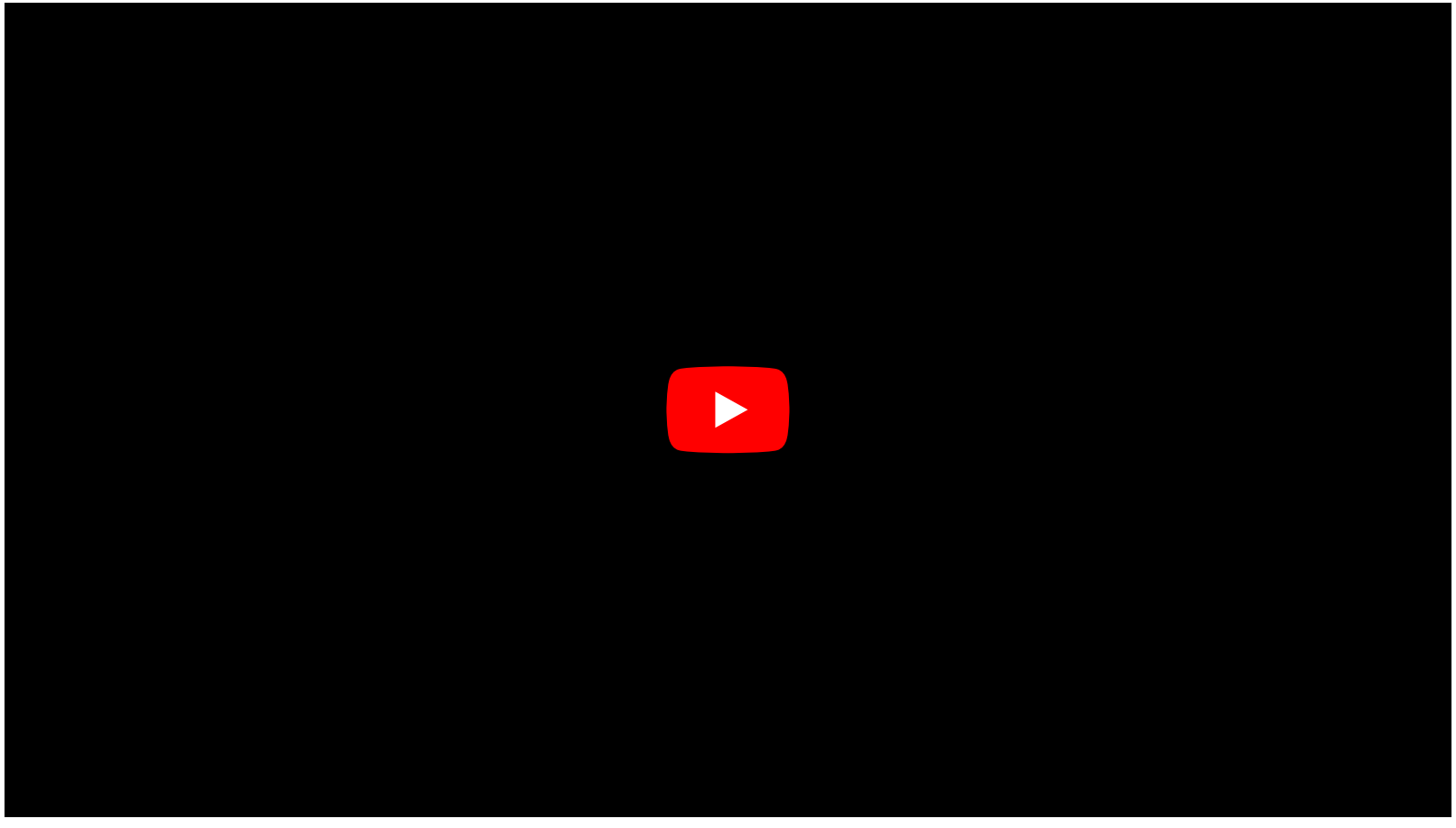
Official formula:

1955 annual food budget × 3

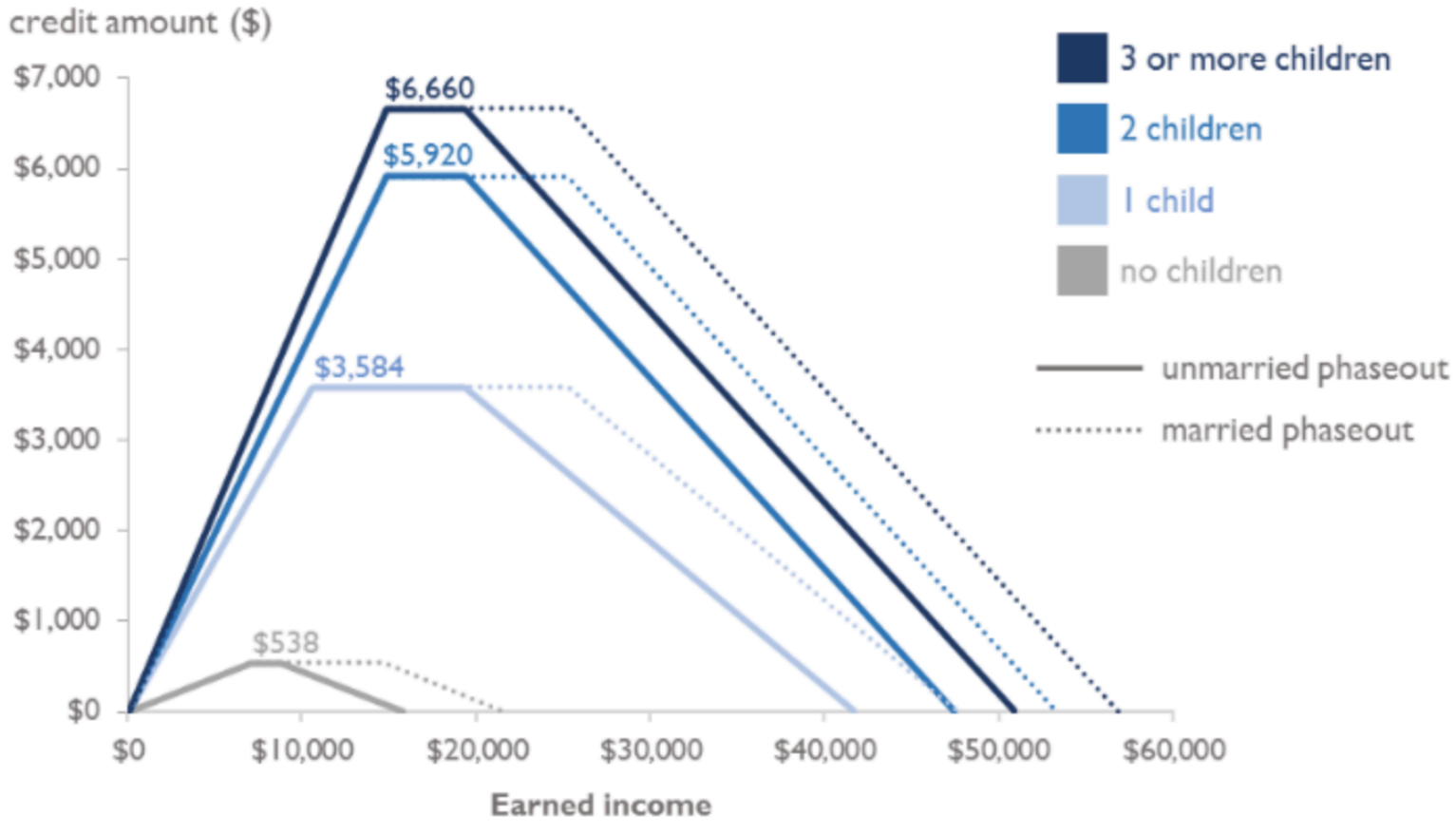
That's all!

In 1963 poverty line was 50% of median income;
in 2005 it was 28%; 18% today

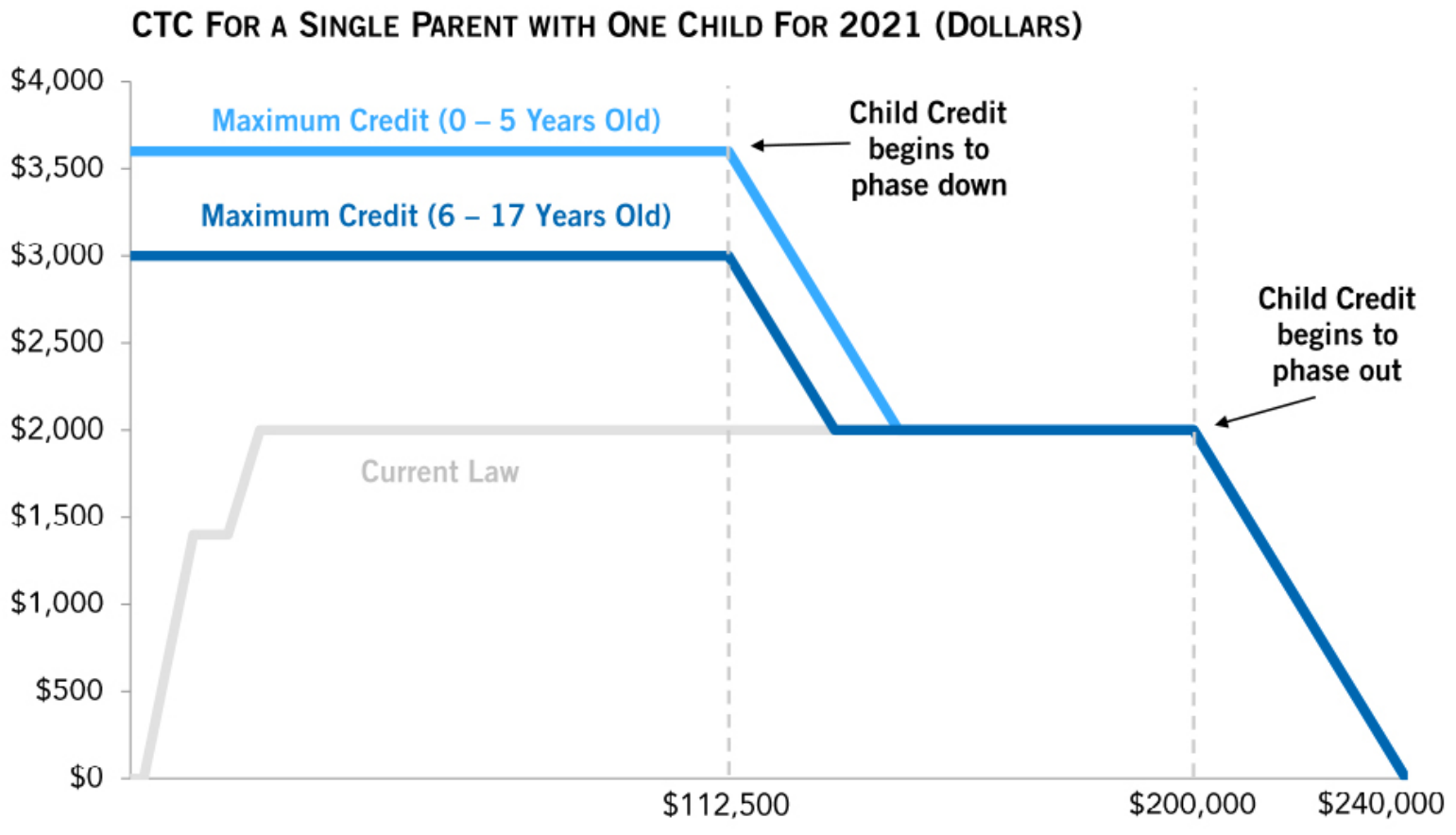
Why don't we change it?



EITC Amount by Number of Qualifying Children, Marital Status, and Income, 2020



The Child Tax Credit is fully refundable under the American Rescue Plan



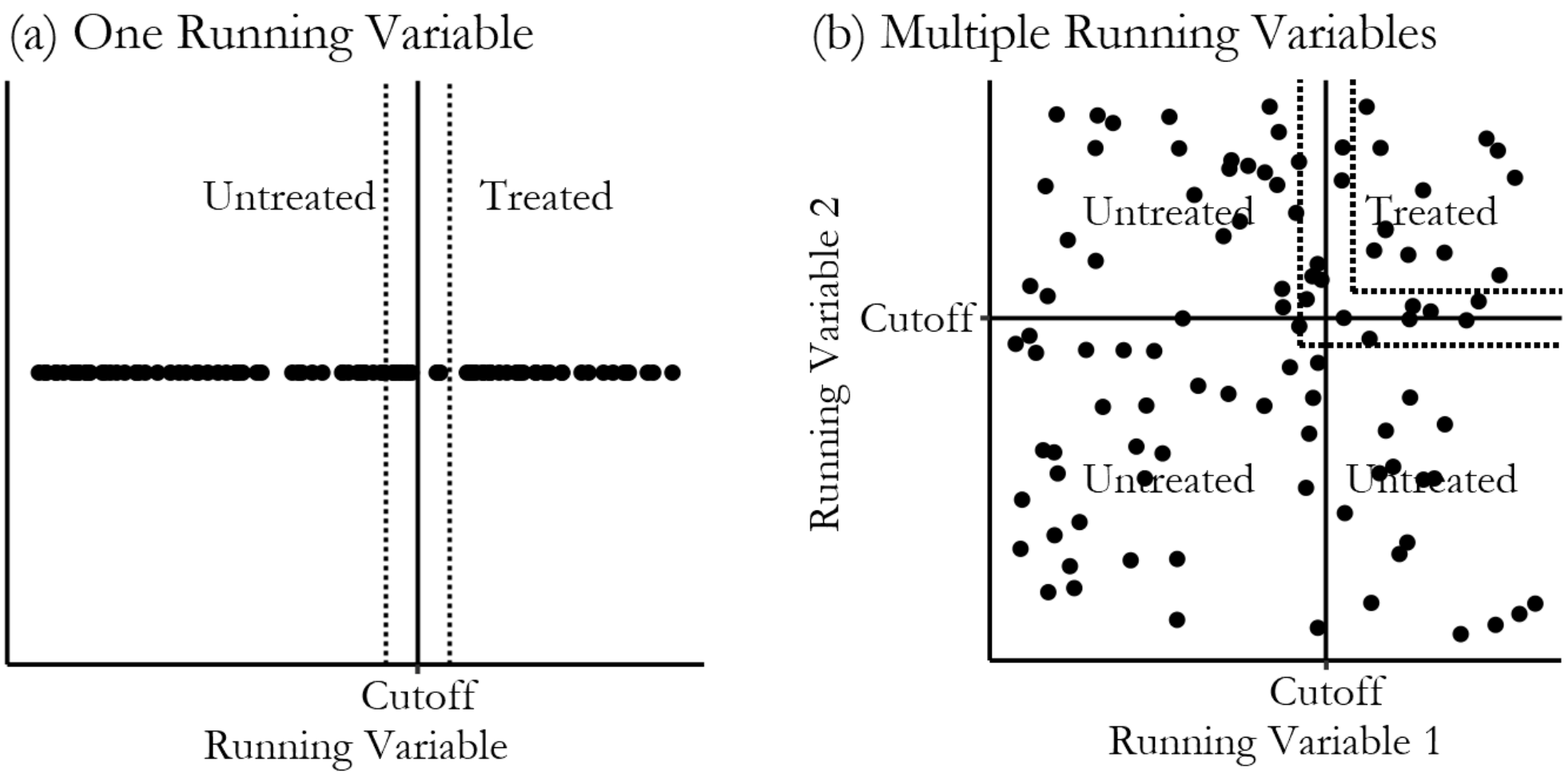
SOURCE: Congressional Research Service, *The Child Tax Credit: Temporary Expansion for 2021 Under the American Rescue Plan Act of 2021*, March 2021.
NOTE: Single parent is assumed to file as head of household.

Why does the cutoff need to be unique to the program of interest?

What if there are multiple cutoffs?

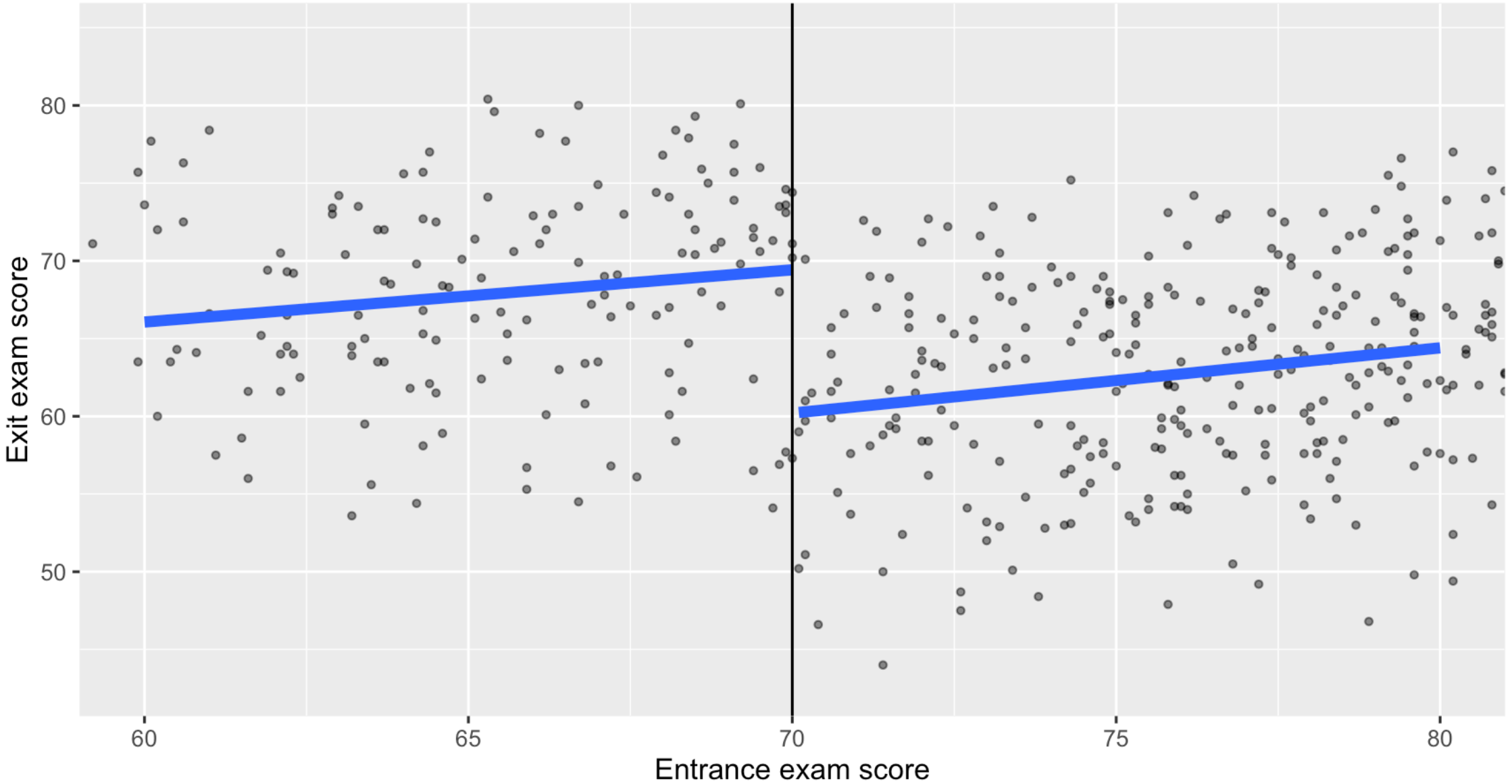
College admission is based on GPA *and* test scores...

WIC/SNAP/Medicaid are based on income *and* family size...


(a) One Running Variable
Untreated
Treated
Cutoff
Running Variable
(b) Multiple Running Variables
Running Variable 2
Cutoff
Untreated
Treated
Untreated
Untreated
Cutoff
Running Variable 1

Why do we center the running variable?

Regression is just fancy averages!


Exit exam score
80
70
60
50
60
65
70
75
80
Entrance exam score

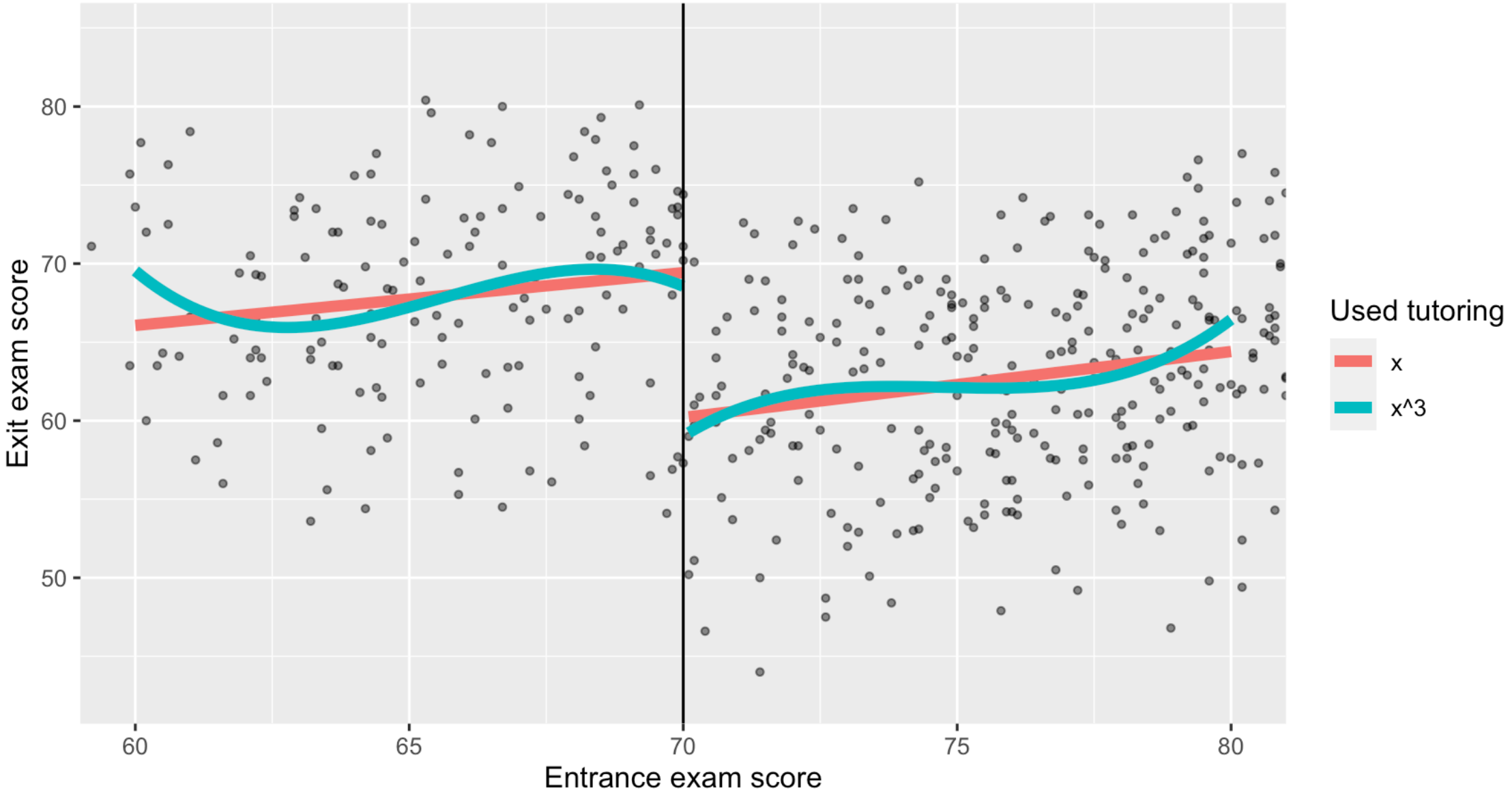
```
lm(exit_exam ~ entrance_exam + tutoring,
   data = filter(tutoring, entrance_exam <= 80,
                 entrance_exam >= 60)) %>%
  tidy()
```

```
## # A tibble: 3 × 5
##   term           estimate std.error statistic  p.value
##   <chr>             <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept)     33.2       8.64       3.84 1.43e- 4
## 2 entrance_exam    0.388     0.114      3.40 7.45e- 4
## 3 tutoringTRUE     9.27      1.31       7.09 6.27e-12
```

```
tutoring_centered <- tutoring %>%
  mutate(entrance_centered = entrance_exam - 70)

lm(exit_exam ~ entrance_centered + tutoring,
   data = filter(tutoring_centered, entrance_exam <= 80,
                 entrance_exam >= 60)) %>%
  tidy()
```

```
## # A tibble: 3 × 5
##   term              estimate std.error statistic   p.value
##   <chr>                <dbl>     <dbl>     <dbl>     <dbl>
## 1 (Intercept)          60.4      0.752     80.3  2.99e-249
## 2 entrance_centered     0.388    0.114      3.40 7.45e-  4
## 3 tutoringTRUE          9.27     1.31       7.09 6.27e- 12
```


Exit exam score
80
70
60
50
60
65
70
75
80
Entrance exam score
Used tutoring
x
x^3

What's the difference between weighting with kernels and inverse probability weighting?

There must be some math behind for the non-parametric lines. Should we care about that or should we just trust in R?

Should we control for confounders?

How do we decide on the right model?

Parametric with $y = x$?

With $y = x^2 + x$?

With $y = x^{\text{whatever}} + x^{\text{whatever}} + x$?

Nonparametric?

`rdrobust()` or just `lm()`?

Controls or no controls?

How do you justify a bandwidth?

Does the bandwidth need to be the same on both sides?

How should we think about the impact of the program on people who score really high or low on the running variable?

If we're throwing most of the data away and only looking at a narrow bandwidth of people, what does this say about generalizability?

What do we do about noncompliance and manipulation?

Fuzzy regression discontinuity!

Why wait for fuzzy regression discontinuity?

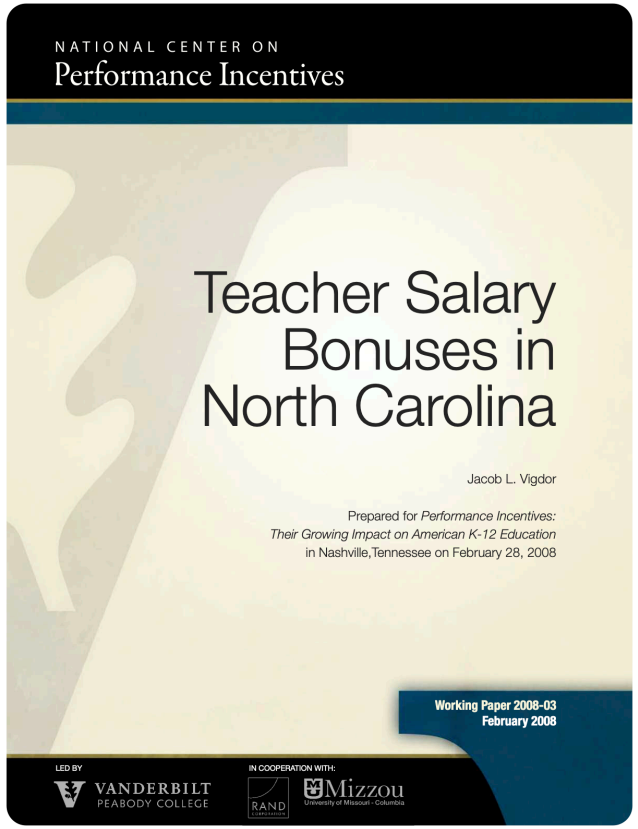
It's RD + instrumental variables

Can other quasi-experimental methods be combined too?

Difference in discontinuity!
Diff-in-diff + RD

https://doi.org/10.1016/j.jebo.2023.12.001

RD play time!



Teachers in North Carolina Public schools earn a bonus of $750 if the students in their school meet a standard called "expected growth." A summary statistic called "average growth" is computed for each school; the expected growth standard is met when this summary measure exceeds zero.

Does getting a bonus in year t cause improved student performance in year $t + 1$?