Review Session 9

Sophie Hill API202 04/01/2022

Agenda

Practice for Quiz #3!

Topics to review:

- **Causal estimands** (Review Session 6)
- **Interpreting regression coefficients** (Review Session 7)
- **Omitted variable bias** (Review Session 8)
- **Fixed effects, Difference-in-differences** (Review Session 9)

**Disclaimer: this is not a comprehensive list!*

Key terms*

Causal inference:

- Individual Treatment Effect
- Average Treatment Effect (on the Treated)
- Observed outcome
- Potential outcome
- Selection bias
- Heterogeneous effects bias

Linear regression:

- Bivariate regression
- Multiple regression
- Intercept
- Slope
- Error
- Residual
- t-statistic
- Omitted variable bias
- Linear probability model
- Reference category
- Interaction term

Causal inference with regression:

- Observational study
- Randomized study
- Fixed effects
- Diff-in-diff
- Parallel trends assumption

Quiz time!

If $\hat{\beta}_1 = 1.4$ with a standard error of 0.2, what is the t-stat? Is it statistically significant?

- $t = 7$, statistically significant at $\alpha = 0.05$
- $t = 0.7$, not statistically significant at $\alpha = 0.05$
- $t = 7$, not statistically significant at $\alpha = 0.05$
- $t = 70$, statistically significant at $\alpha = 0.05$

If $\hat{\beta}_1 = 1.4$ with a standard error of 0.2, what is the t-stat? Is it statistically significant?

- $t = 7$, statistically significant at $\alpha = 0.05$
- $t = 0.7$, not statistically significant at $\alpha = 0.05$
- $t = 7$, not statistically significant at $\alpha = 0.05$
- $t = 70$, statistically significant at $\alpha = 0.05$

What is u_i in the population regression function

$$Y_i = \beta_0 + \beta_1 X_{1i} + u_i \text{ ?}$$

- Individual error term
- Difference between Y_i and mean Y among people with the same X_1
- Everything correlated with Y that is not captured by X_1

What is u_i in the population regression function

$$Y_i = \beta_0 + \beta_1 X_{1i} + u_i \text{ ?}$$

Individual error term

Difference between Y_i and mean Y among people with the same X_1

Everything correlated with Y that is not captured by X_1

$$Income_i = 30000 + 500Age_i + u_i$$

What is the predicted income for a 30-year old?

- $55,000
- $150,000
- $45,000
- $145,000

$$Income_i = 30000 + 500Age_i + u_i$$

What is the predicted income for a 30-year old?

- $55,000
- $150,000
- $45,000
- $145,000

$$Income_i = 30000 + 500Age_i + u_i$$

What is the predicted income difference between a 30-year-old and a 31-year-old?

- $15,500
- $30,500
- $500
- $45,500

$$Income_i = 30000 + 500Age_i + u_i$$

What is the predicted income difference between a 30-year-old and a 31-year-old?

- $15,500
- $30,500
- $500
- $45,500

$$Income_i = 30000 + 500Age_i + u_i$$

What is the predicted income difference between a 50-year-old and a 51-year-old?

- $15,500
- $30,500
- $500
- $45,500

$$Income_i = 30000 + 500Age_i + u_i$$

What is the predicted income difference between a 50-year-old and a 51-year-old?

- $15,500
- $30,500
- $500
- $45,500

$$Income_i = 30000 + 20Age_i + 10Age_i^2 + u_i$$

What is the predicted income for a 30-year old?

- $45,000
- $39,600
- $30,900
- $30,600

$$Income_i = 30000 + 20Age_i + 10Age_i^2 + u_i$$

What is the predicted income for a 30-year old?

- $45,000
- $39,600
- $30,900
- $30,600

$$Income_i = 30000 + 20Age_i + 10Age_i^2 + u_i$$

What is the predicted income difference for a 30-year old vs a 31-year-old?

- $20
- $30
- $620
- $630

$$Income_i = 30000 + 20Age_i + 10Age_i^2 + u_i$$

What is the predicted income difference for a 30-year old vs a 31-year-old?

$20

$30

$620

$630

$$Income_i = 30000 + 20Age_i + 10Age_i^2 + u_i$$

What is the predicted income difference for a 50-year old vs a 51-year-old?

- $56,000
- $57,030
- $1,030
- $1,010

$$Income_i = 30000 + 20Age_i + 10Age_i^2 + u_i$$

What is the predicted income difference for a 50-year old vs a 51-year-old?

- $56,000
- $57,030
- $1,030
- $1,010

If we think the relationship between X and Y might be non-linear we should...

- Not use linear regression
- Include X^2 instead of X
- Include X^2 and X
- Convert X from continuous to categorical

If we think the relationship between X and Y might be non-linear we should...

Not use linear regression

Include X^2 instead of X

Include X^2 and X

Convert X from continuous to categorical

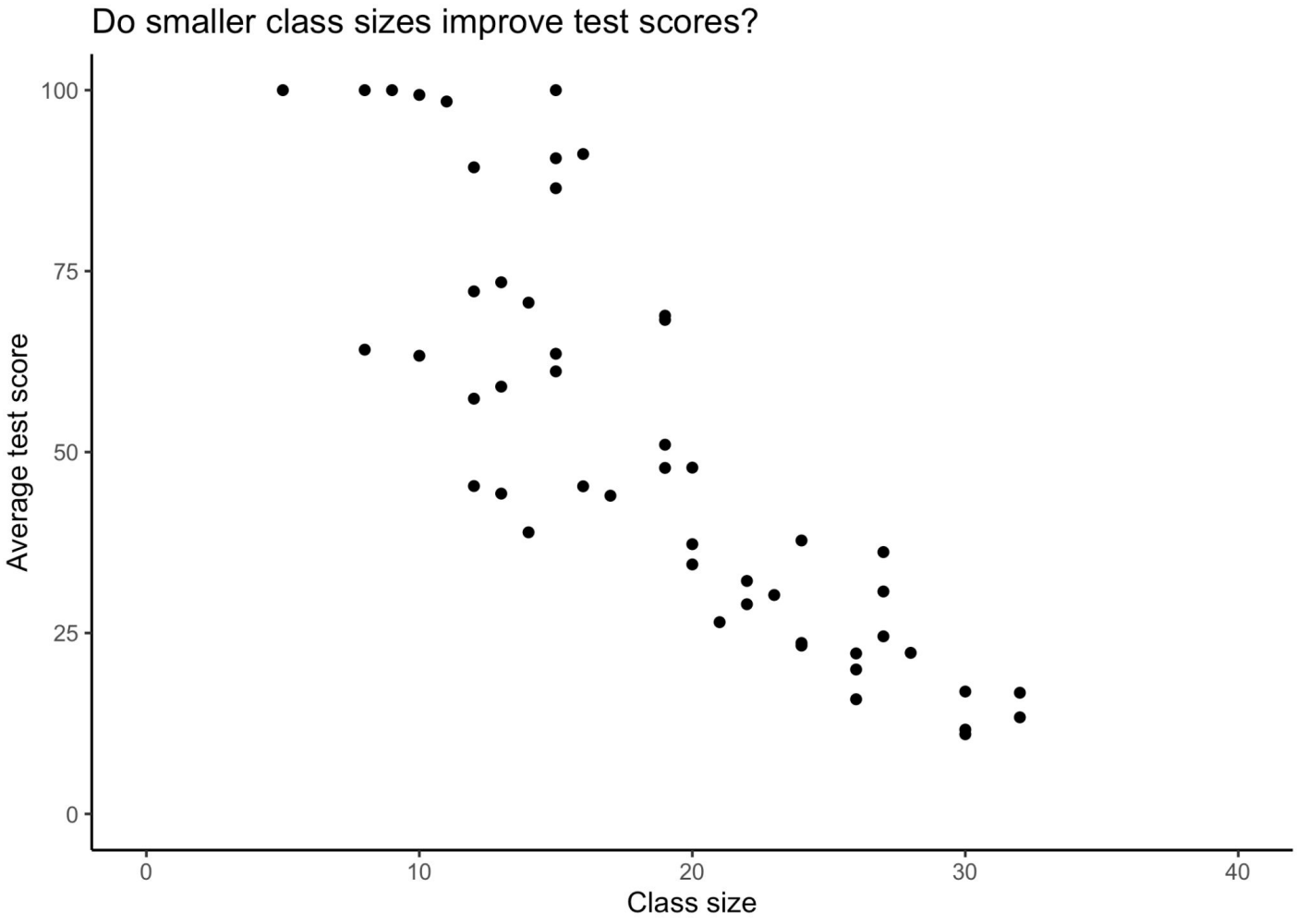
You are interested in the effect of attending Harvard on future earnings. If you include parental wealth as another predictor, the coefficient on Harvard will...

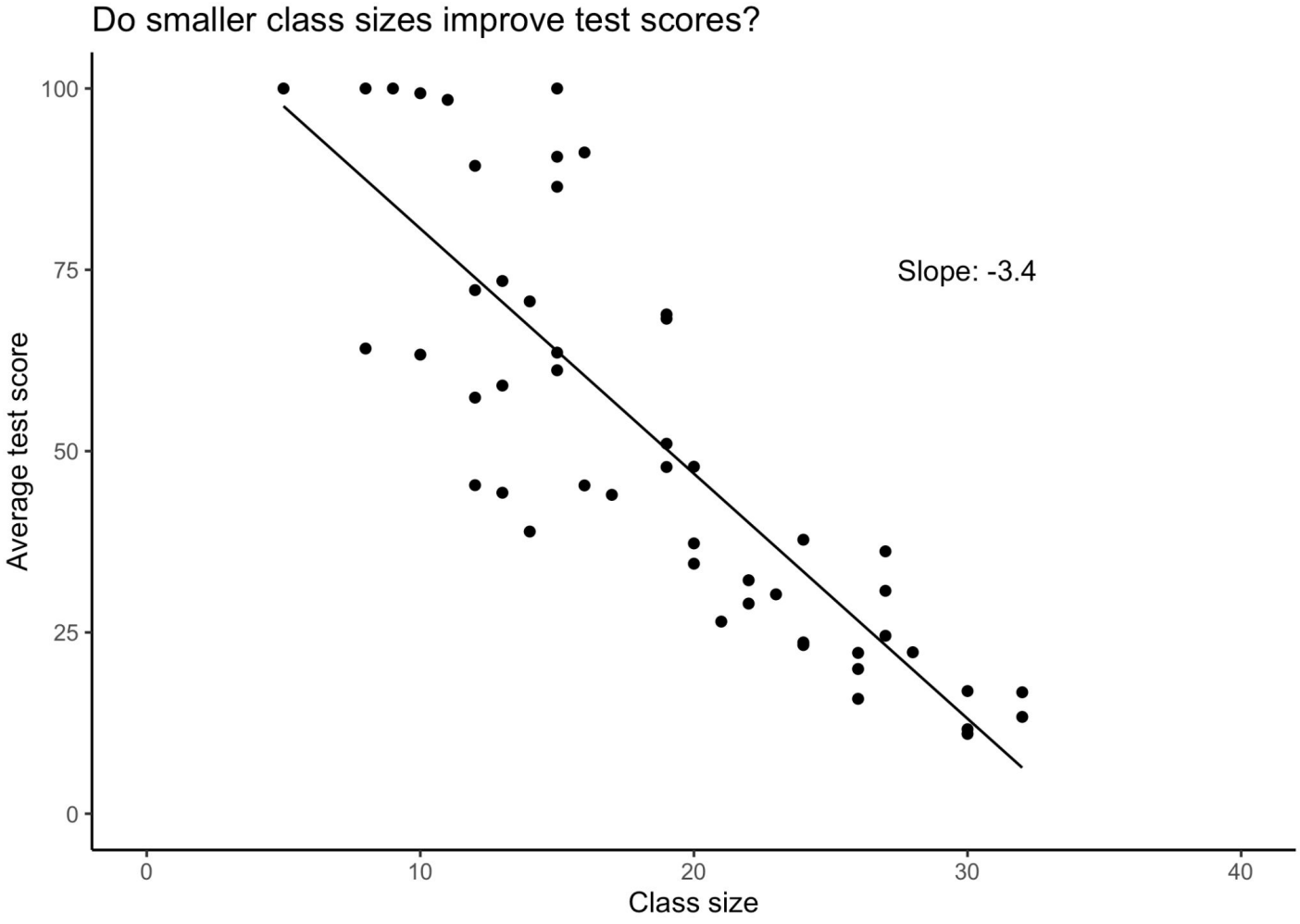
- Stay the same
- Get larger
- Get smaller
- Not enough information

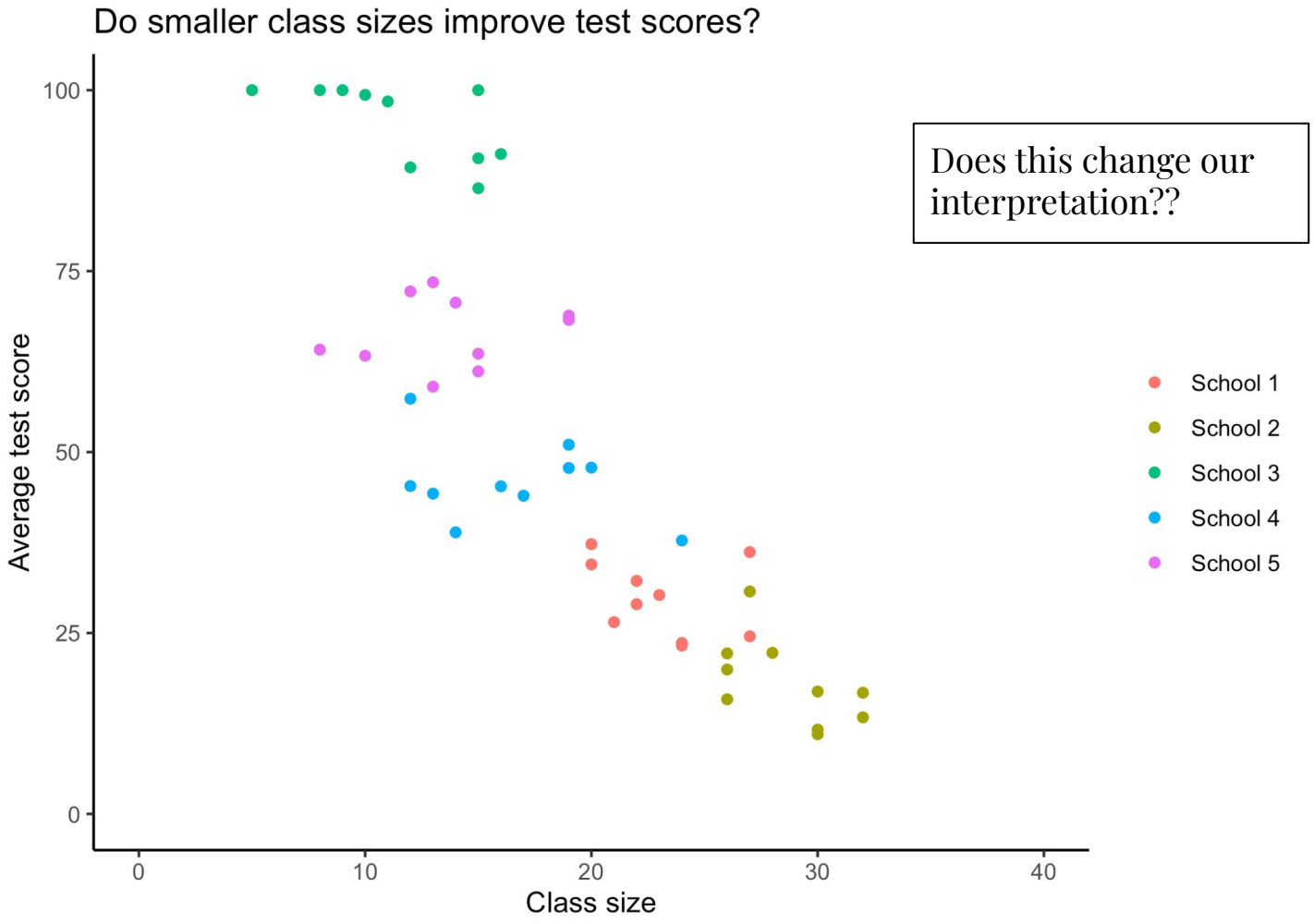
You are interested in the effect of attending Harvard on future earnings. If you include parental wealth as another predictor, the coefficient on Harvard will...

- Stay the same
- Get larger
- Get smaller
- Not enough information

Fixed Effects


Do smaller class sizes improve test scores?
Average test score
100
75
50
25
0
0
10
20
30
40
Class size

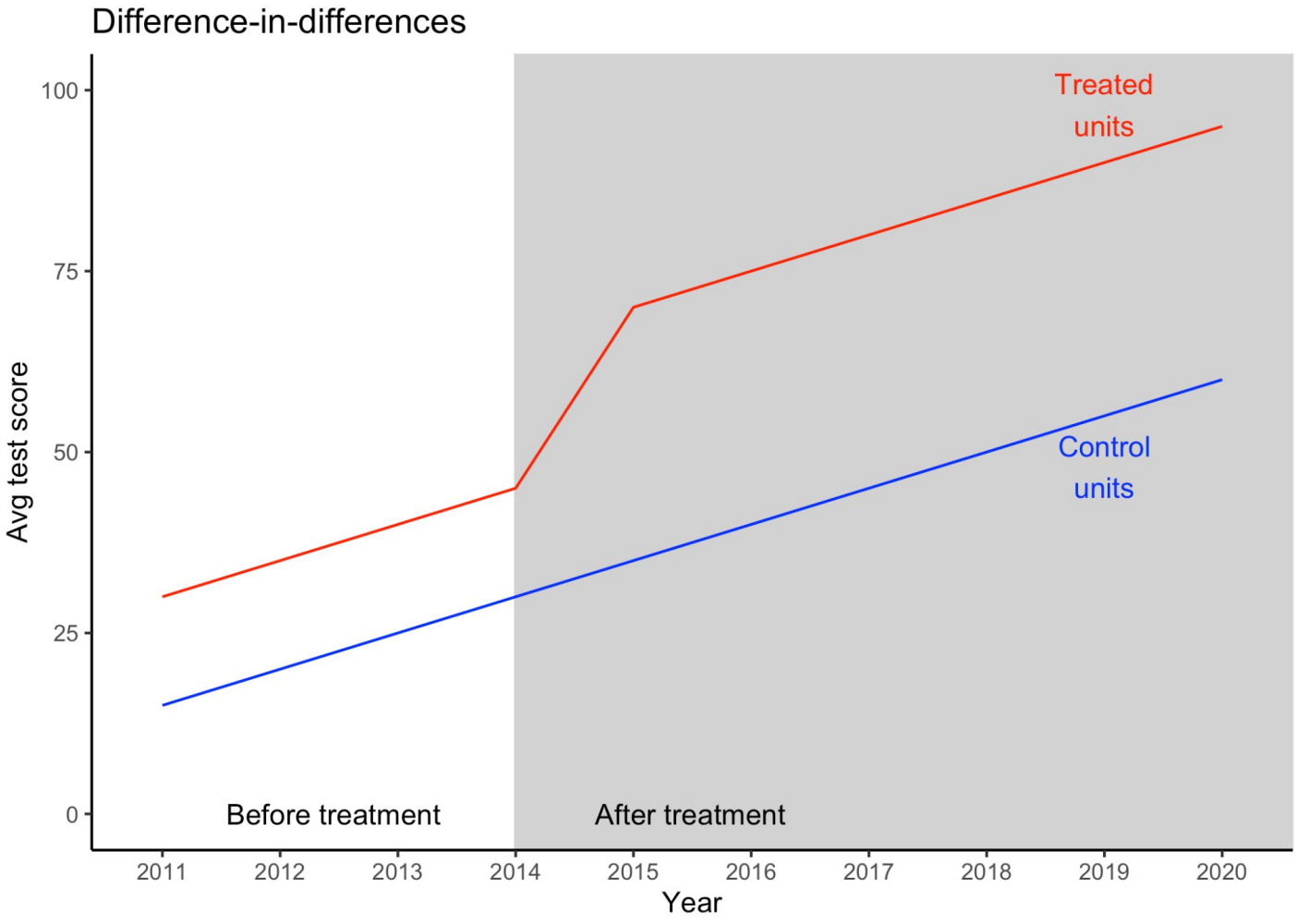

Do smaller class sizes improve test scores?
Average test score
100
75
50
25
0
0
10
20
30
40
Class size
Slope: -3.4

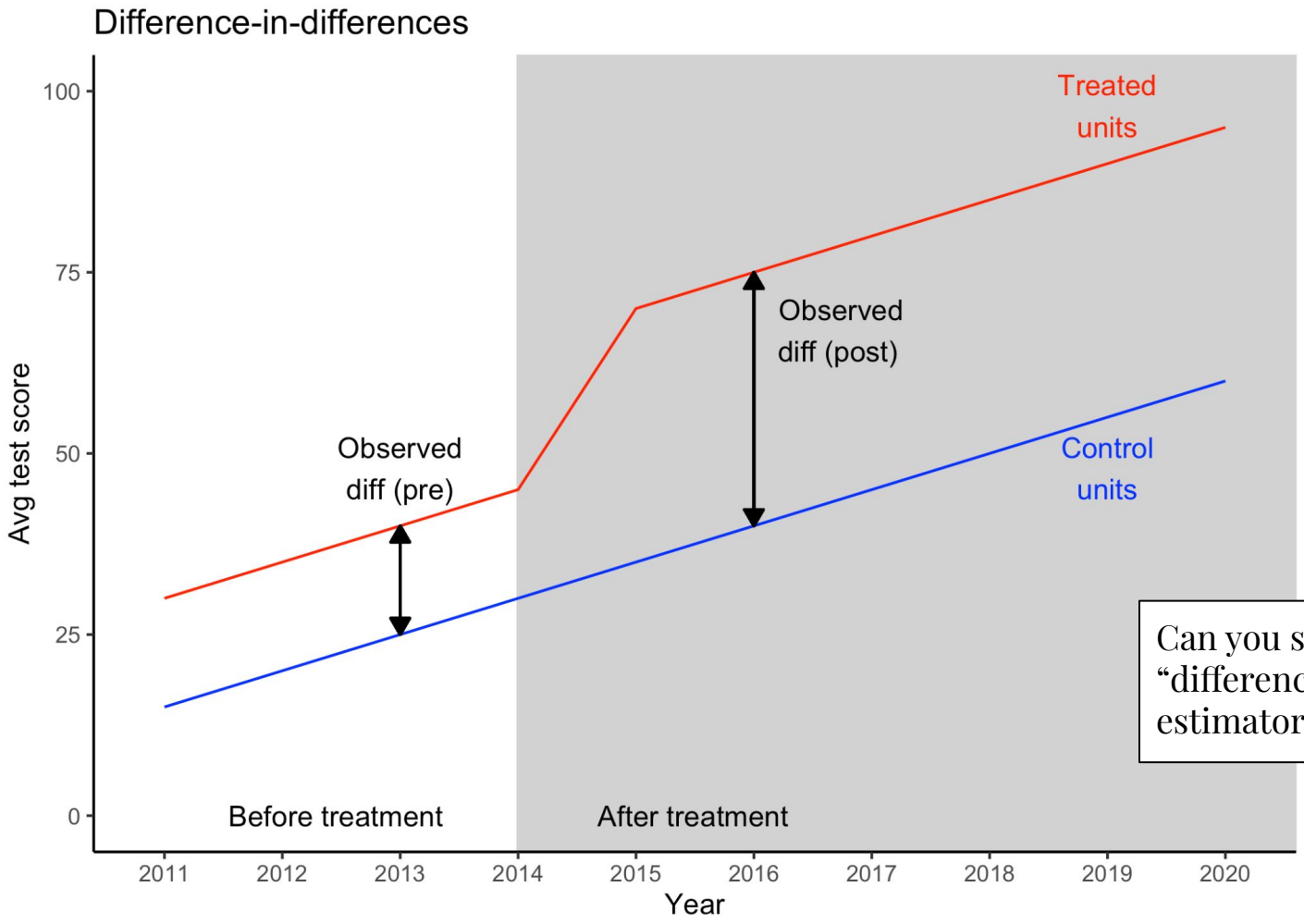


Does this change our interpretation??

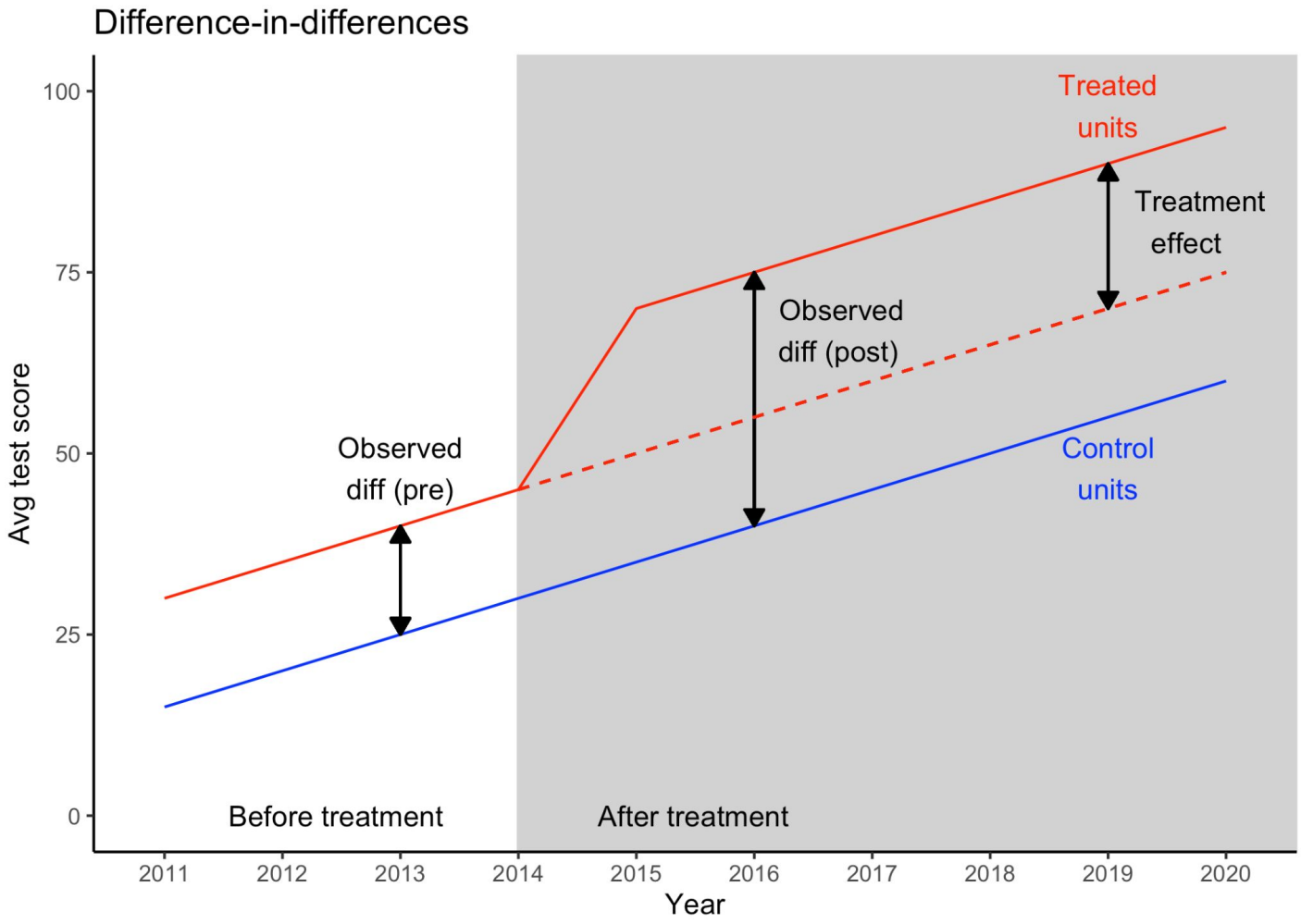
Adding school fixed effects means that the slope of the regression line is the same for all schools, but the intercept can vary.

Difference-in-differences


Difference-in-differences
Treated
units
Control
units
Before treatment
After treatment
Avg test score
Year
100
75
50
25
0
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020

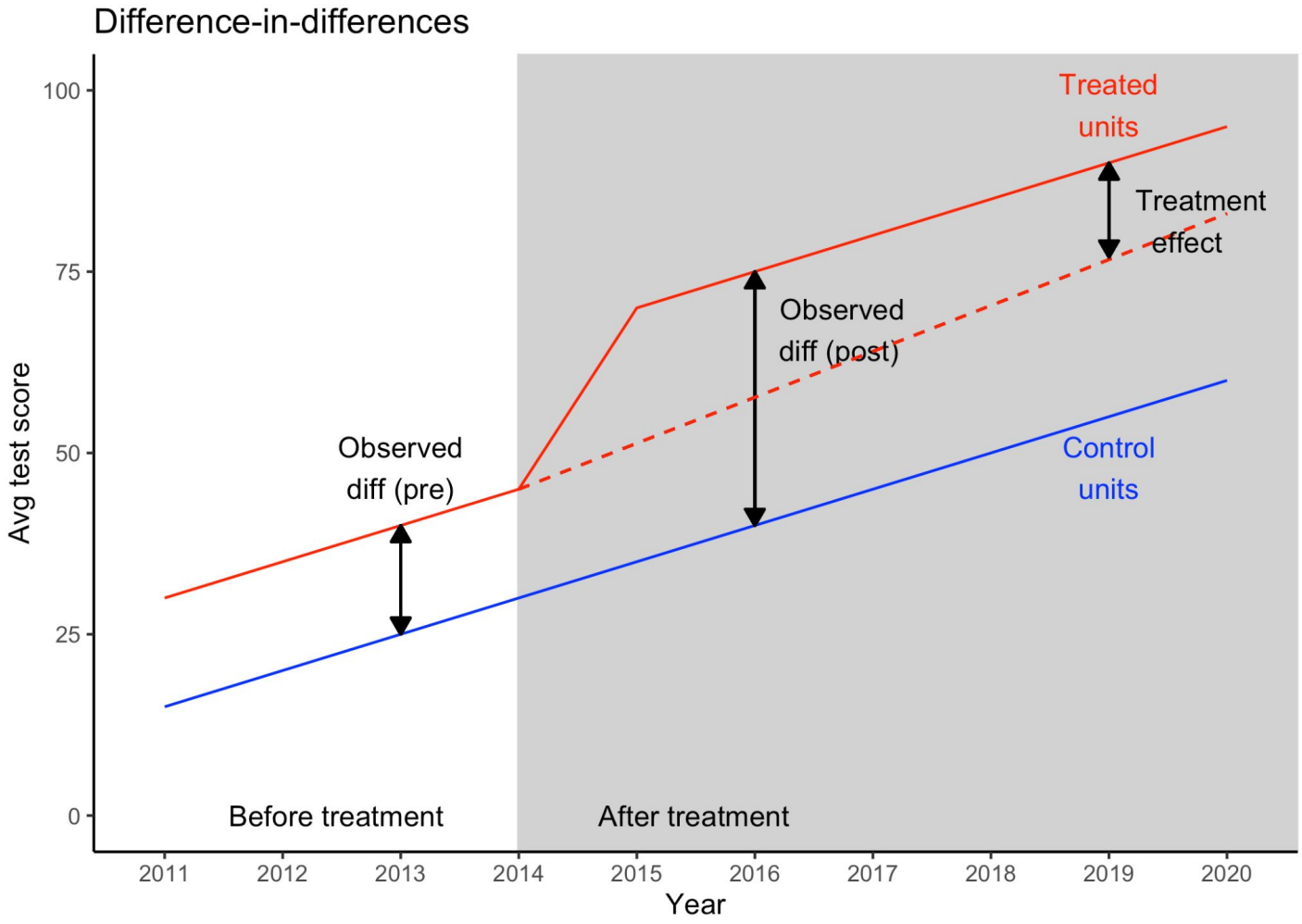


Can you see why we call this the "difference-in-differences" estimator?



The dashed line is the *counterfactual* trend (treated units if they did not get treated) – we cannot observe it!

Relies on the parallel trends assumption.



This would be a violation of the parallel trends assumption!

Here, our estimate of the treatment effect using the DiD estimator would be an **overestimate**.

Parallel trends assumption

- Parallel trends assumption: **in the absence of treatment**, the difference in outcomes between the treated and control units would be **constant over time**.

- Since this assumption relates to an **unobserved** counterfactual, it cannot be empirically tested!

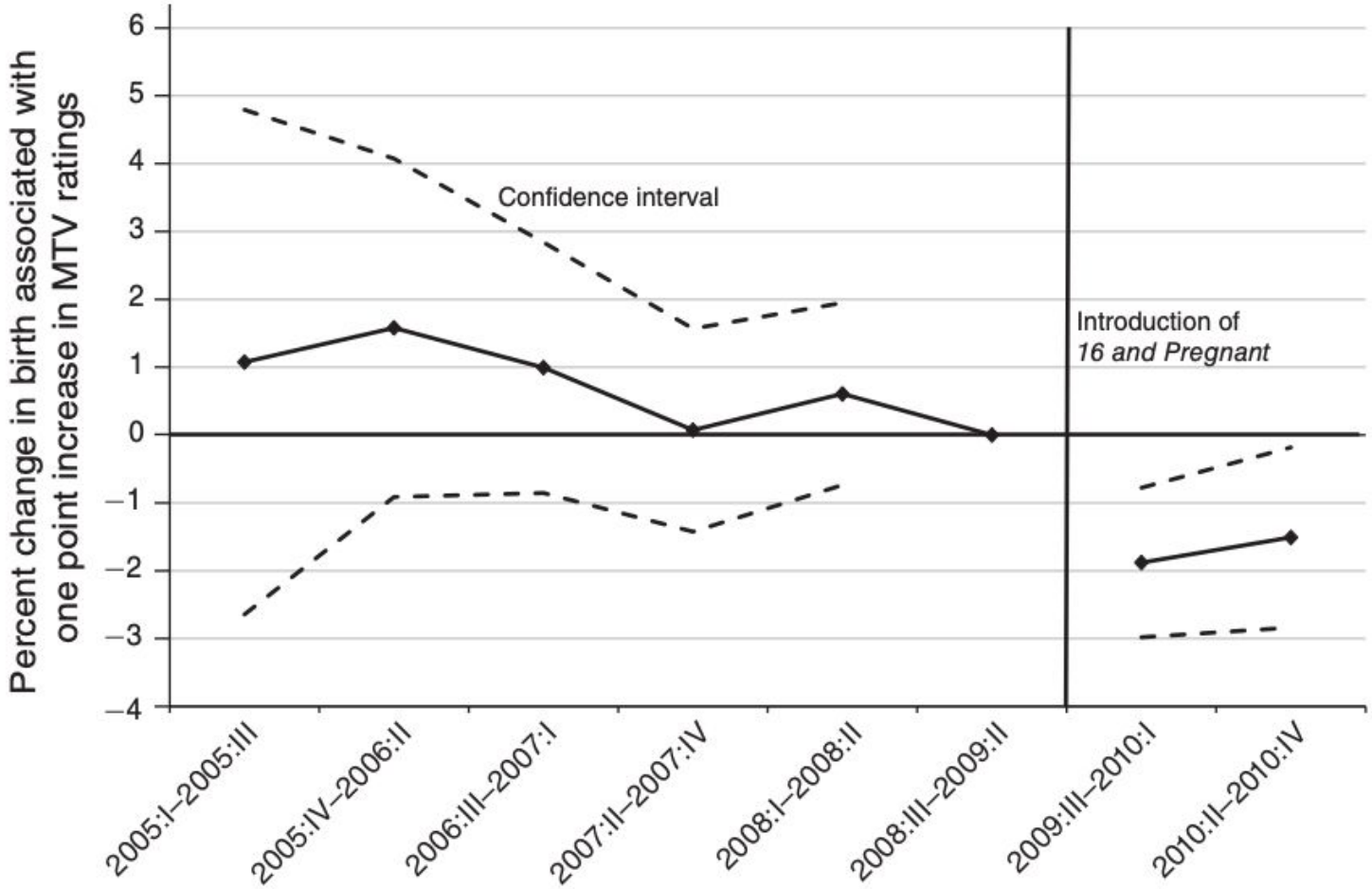
- We can empirically test whether treated and control units have parallel trends in the pre-period… but this does not guarantee that the assumption will hold!

MTV's "16 and pregnant" reduces teen pregnancies?

Kearney & Levine (2015) used a diff-in-diff to argue that the introduction of the MTV reality show "16 and pregnant" caused a 4.3 percent drop in teen birth rates between July 2009 and December 2010.

Treated units = places where lots of people watched MTV before "16 and pregnant was introduced"

Control units = places where few people watched MTV before "16 and pregnant was introduced"


Percent change in birth associated with one point increase in MTV ratings
6
5
4
3
2
1
0
−1
−2
−3
−4
Confidence interval
Introduction of
16 and Pregnant
2005:I–2005:III
2005:IV–2006:II
2006:III–2007:I
2007:II–2007:IV
2008:I–2008:II
2008:III–2009:II
2009:III–2010:I
2010:II–2010:IV

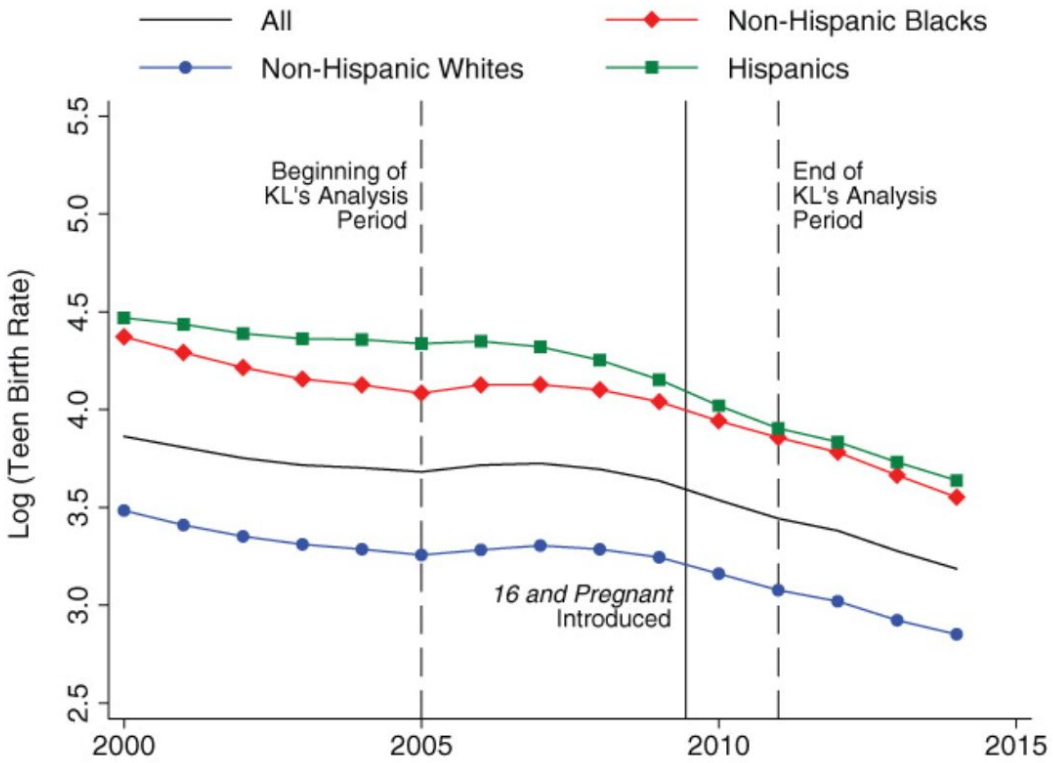
MTV's "16 and pregnant" reduces teen pregnancies? ... or does it??

Jaeger et al. (2020) argued that the paper's conclusions were flawed because the parallel trends assumption required for a diff-in-diff did not hold.

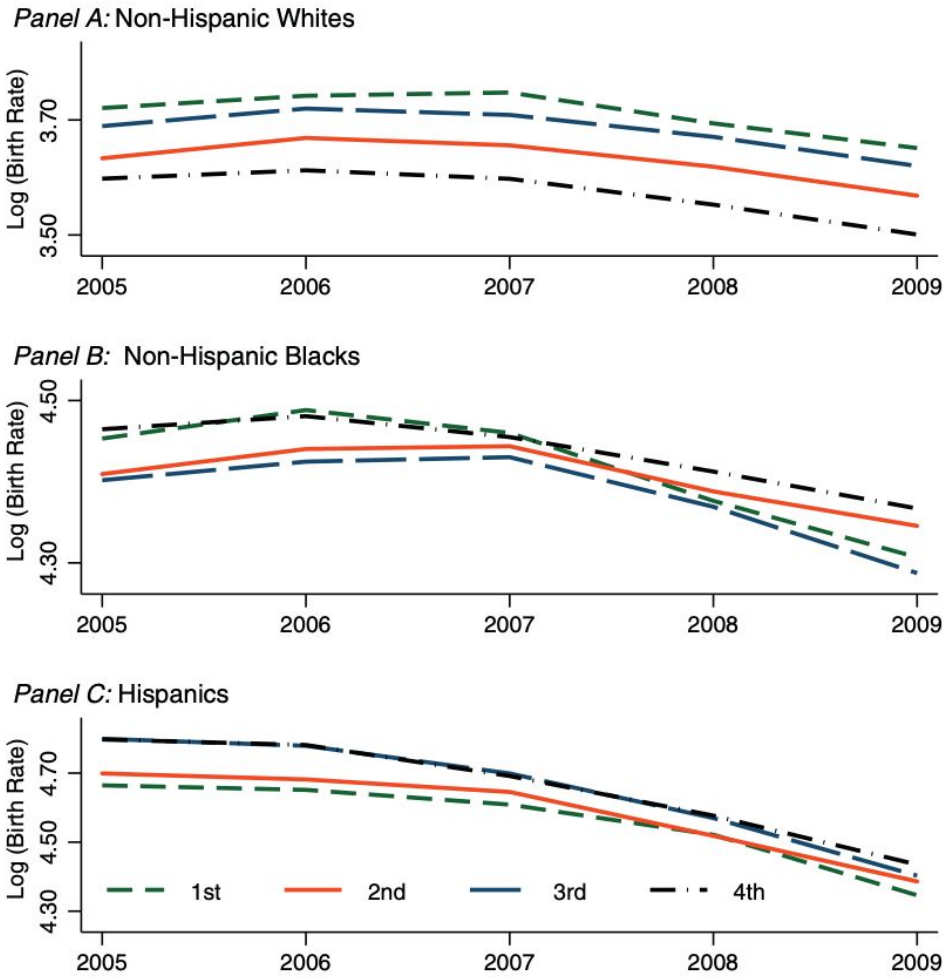
Since MTV viewership is correlated with racial composition, and since racial groups were differentially affected by the Great Recession, we have strong reason to believe the trends would **not** be parallel.

When they include varying time trends for different racial groups in the model, the DiD estimate for "16 and pregnant" is insignificant.

When the parallel trends assumption is violated



When the parallel trends assumption is violated



Takeaways

- To use a diff-in-diff design, the researcher must provide a good justification for why the parallel trends assumption holds.
- Testing for parallel trends in the pre-period is often illuminating but does not tell us whether the assumption actually holds!
- Ultimately, the researcher needs to provide a strong **theoretical** argument for why the assumption holds in their context.
- The original Kearney & Levine paper was published in a prestigious journal 7 years ago… it would probably not get published there today!