

# Regression Discontinuity and Differences in Differences

Fernando Hoces la Guardia

08/03/2022

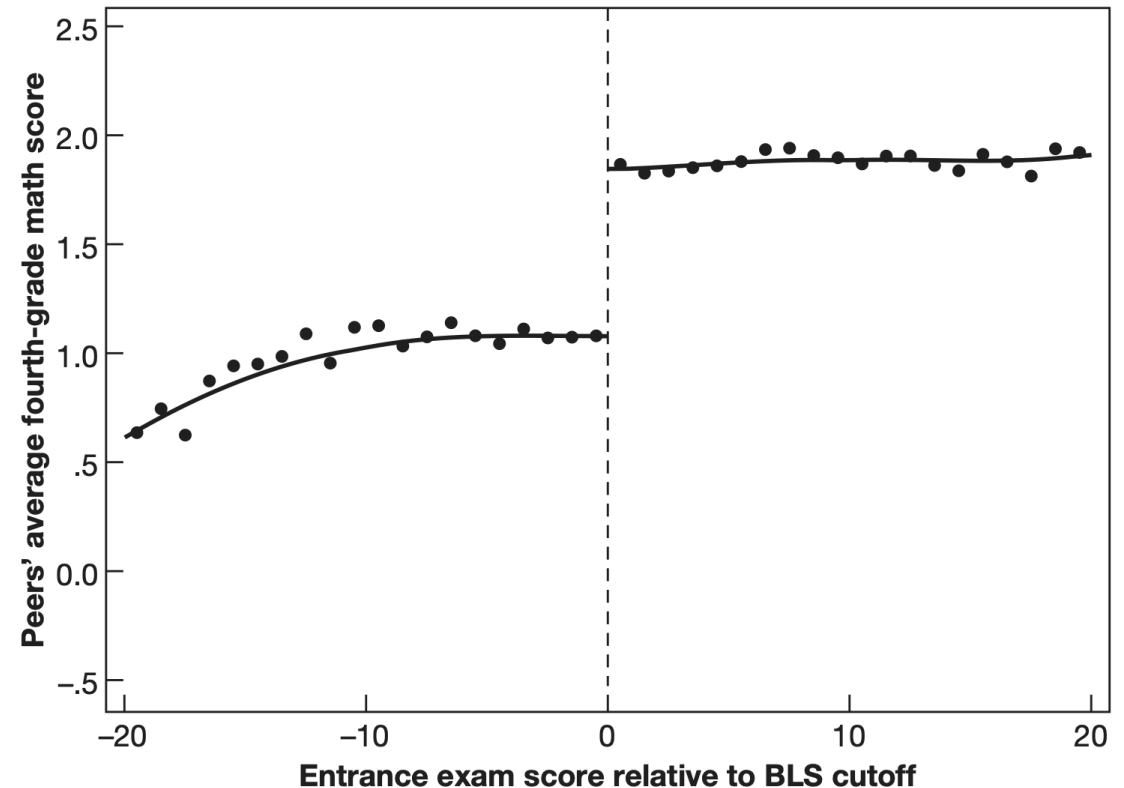
# Today's Lecture

- Finish Fuzzy RDD
- Start Differences in Differences

# Fuzzy RDD is IV in Peer Effect Example 2/3

- Beware of confusions: in sharp RDD this variable represents the treatment, in fuzzy represents the instrument (akin to the offers in KIPP and OHP examples).
- To add to the confusion the instrument here is labeled as  $D_i$  (instead of  $Z_i$ )
- If this is the instrument, what is the first stage?

FIGURE 4.8  
Peer quality around the BLS cutoff



Notes: This figure plots average seventh-grade peer quality for applicants to Boston Latin School (BLS), conditional on admissions test scores, for BLS applicants scoring near the admissions cutoff. Peer quality is measured by seventh-grade schoolmates' fourth-grade math scores. Solid lines show fitted values from a local linear regression, estimated separately on either side of the cutoff (indicated by the vertical dashed line).

# Fuzzy RDD is IV in Peer Effect Example 3/3

- First Stage:

$$\overline{X}_{(i)} = \alpha_1 + \phi D_i + \beta_1 R_i + e_{1i}$$

- Reduced Form:

# Fuzzy RDD is IV in Peer Effect Example 3/3

- First Stage:

$$\overline{X}_{(i)} = \alpha_1 + \phi D_i + \beta_1 R_i + e_{1i}$$

- Reduced Form:

$$Y_i = \alpha_0 + \rho D_i + \beta_0 R_i + e_{0i}$$

- Second Stage (for 2SLS):

# Fuzzy RDD is IV in Peer Effect Example 3/3

- First Stage:

$$\overline{X}_{(i)} = \alpha_1 + \phi D_i + \beta_1 R_i + e_{1i}$$

- Reduced Form:

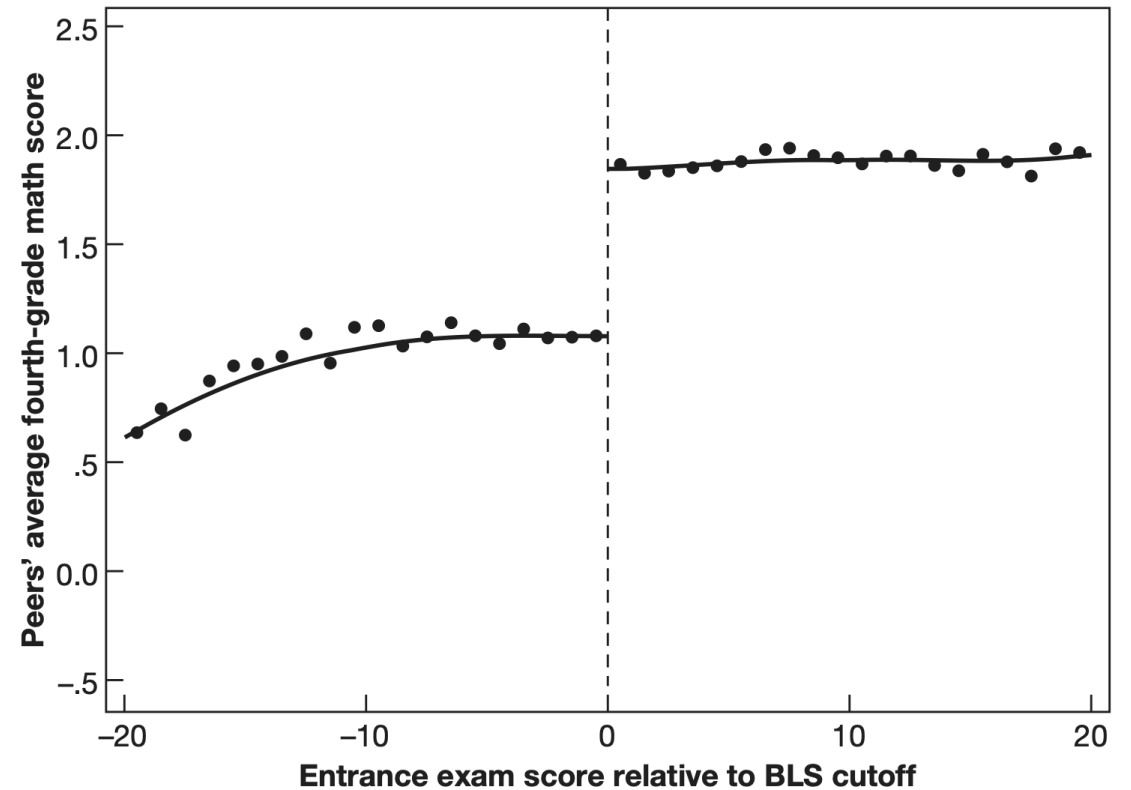
$$Y_i = \alpha_0 + \rho D_i + \beta_0 R_i + e_{0i}$$

- Second Stage (for 2SLS):

$$Y_i = \alpha_2 + \lambda \widehat{\overline{X}_{(i)}} + \beta_2 R_i + e_{2i}$$

# IV Assumptions

FIGURE 4.8  
Peer quality around the BLS cutoff

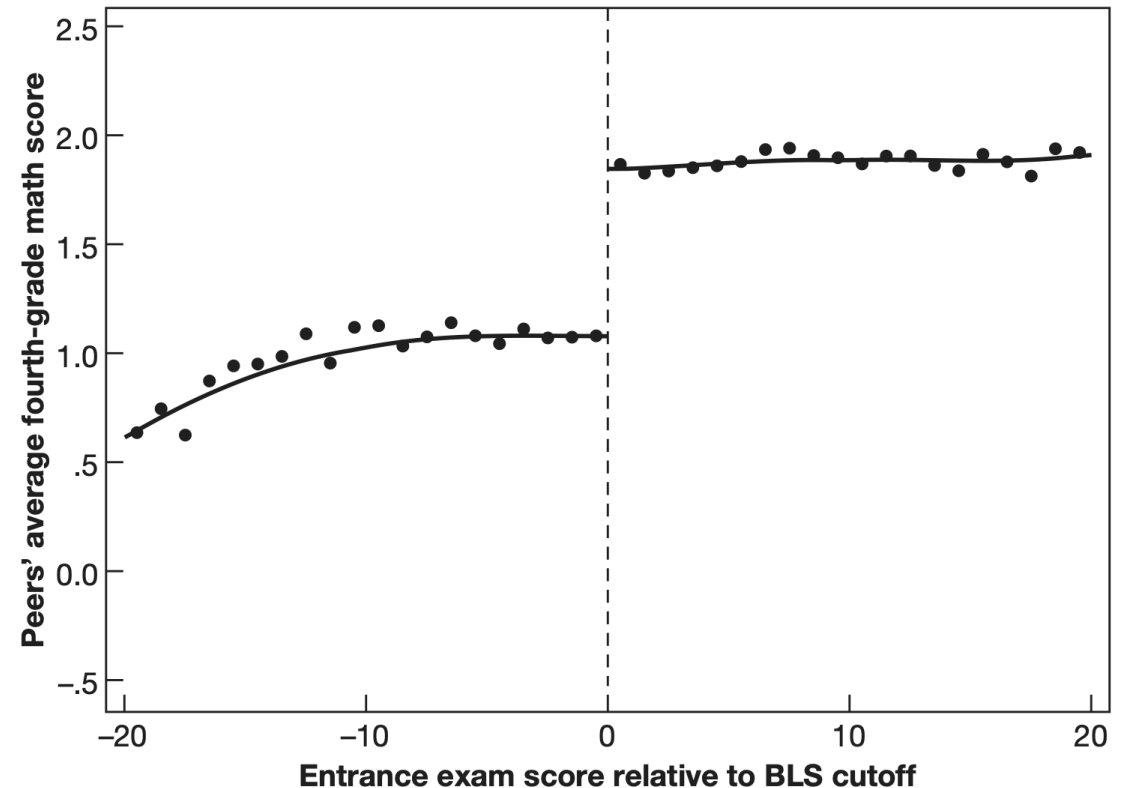


*Notes:* This figure plots average seventh-grade peer quality for applicants to Boston Latin School (BLS), conditional on admissions test scores, for BLS applicants scoring near the admissions cutoff. Peer quality is measured by seventh-grade schoolmates' fourth-grade math scores. Solid lines show fitted values from a local linear regression, estimated separately on either side of the cutoff (indicated by the vertical dashed line).

# IV Assumptions

- **Relevancy:** See figure 4.8. Effect of instrument on treatment is an increase in  $0.8\sigma$  (very big)

FIGURE 4.8  
Peer quality around the BLS cutoff



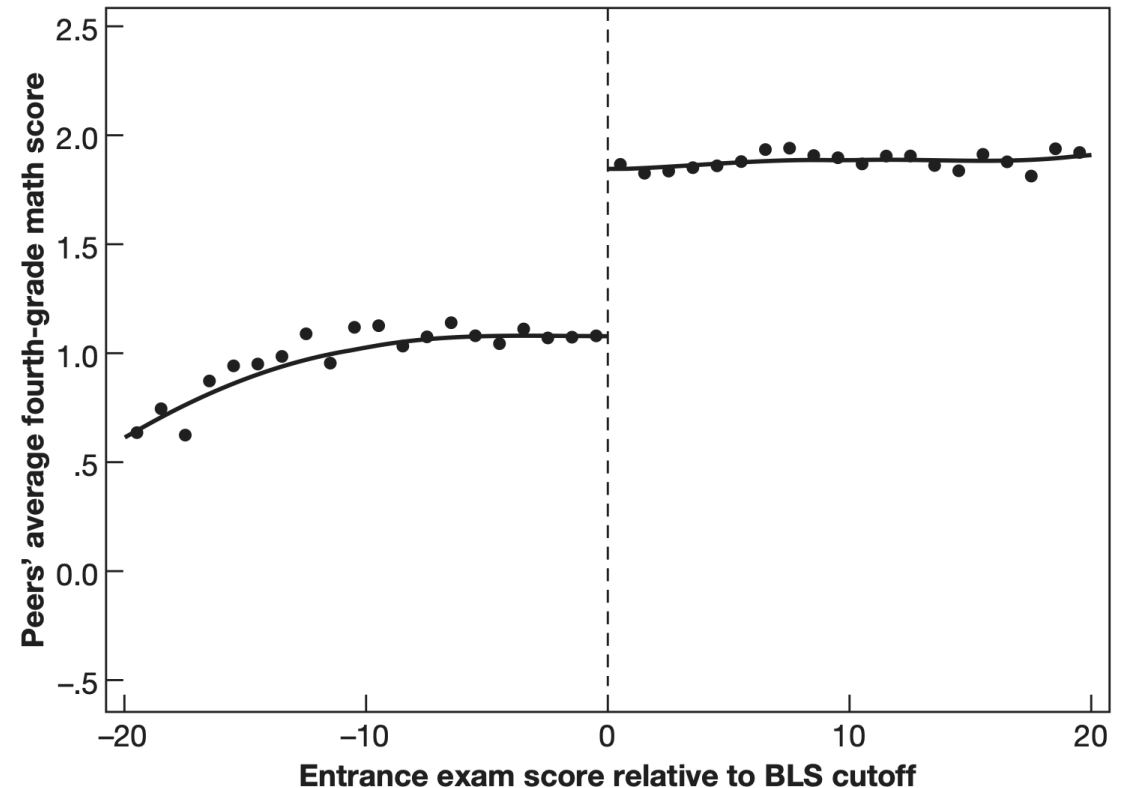
*Notes:* This figure plots average seventh-grade peer quality for applicants to Boston Latin School (BLS), conditional on admissions test scores, for BLS applicants scoring near the admissions cutoff. Peer quality is measured by seventh-grade schoolmates' fourth-grade math scores. Solid lines show fitted values from a local linear regression, estimated separately on either side of the cutoff (indicated by the vertical dashed line).



# IV Assumptions

- **Relevancy:** See figure 4.8. Effect of instrument on treatment is an increase in  $0.8\sigma$  (very big)
- **Independence:** Yes for the same reason that Sharp RDD does not have OVB: Instrument is a deterministic function of a running variable.

FIGURE 4.8  
Peer quality around the BLS cutoff

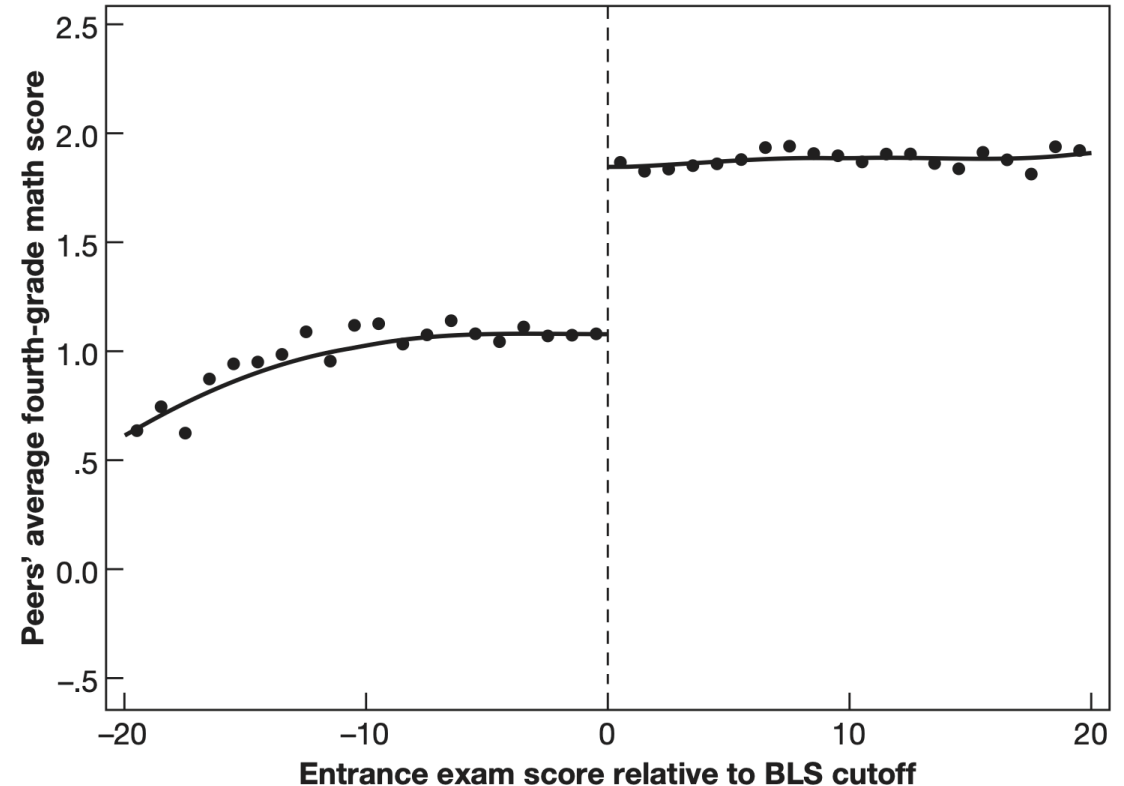


*Notes:* This figure plots average seventh-grade peer quality for applicants to Boston Latin School (BLS), conditional on admissions test scores, for BLS applicants scoring near the admissions cutoff. Peer quality is measured by seventh-grade schoolmates' fourth-grade math scores. Solid lines show fitted values from a local linear regression, estimated separately on either side of the cutoff (indicated by the vertical dashed line).

# IV Assumptions

- **Relevancy:** See figure 4.8. Effect of instrument on treatment is an increase in  $0.8\sigma$  (very big)
- **Independence:** Yes for the same reason that Sharp RDD does not have OVB: Instrument is a deterministic function of a running variable.
- **Exclusion (Restriction):** the cut-off variable (instrument) is influencing the math scores (outcome) only through peer quality (treatment). There are probably other channels so this assumption probably doesn't hold.

FIGURE 4.8  
Peer quality around the BLS cutoff

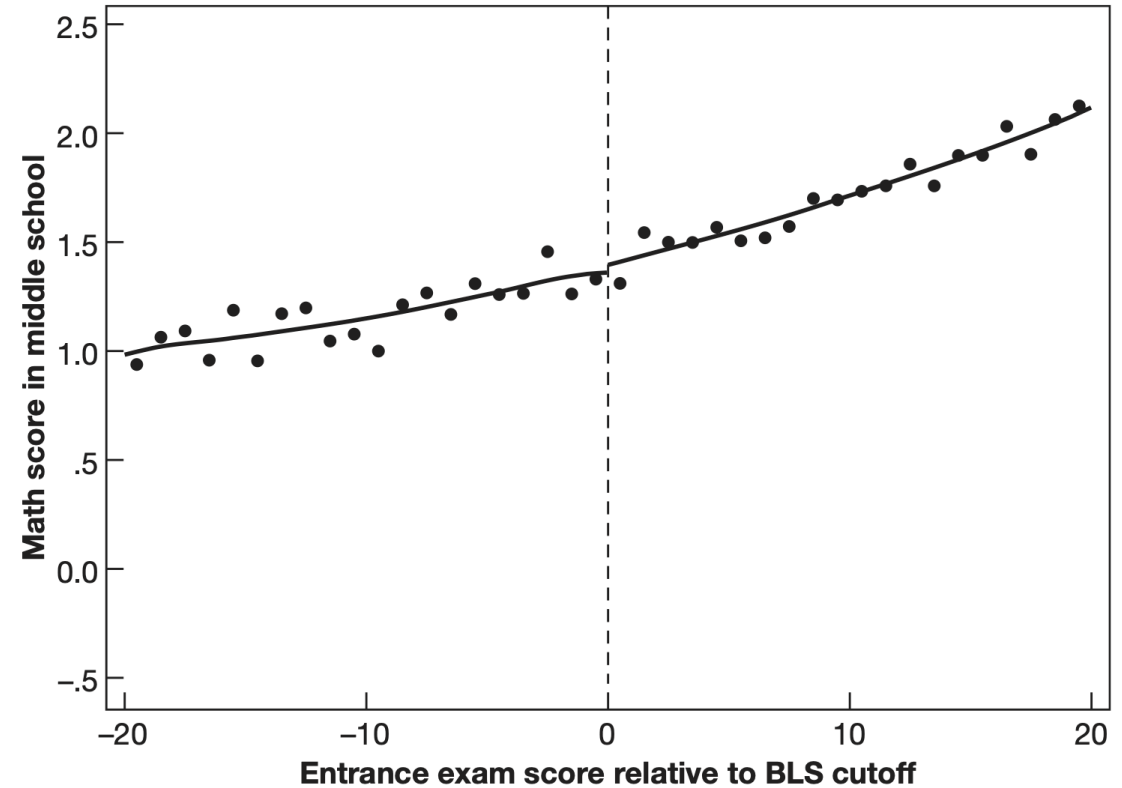


Notes: This figure plots average seventh-grade peer quality for applicants to Boston Latin School (BLS), conditional on admissions test scores, for BLS applicants scoring near the admissions cutoff. Peer quality is measured by seventh-grade schoolmates' fourth-grade math scores. Solid lines show fitted values from a local linear regression, estimated separately on either side of the cutoff (indicated by the vertical dashed line).

# Results

- First Stage:  $\phi = 0.8$   
(no SE reported). Strong first stage.

FIGURE 4.9  
Math scores around the BLS cutoff

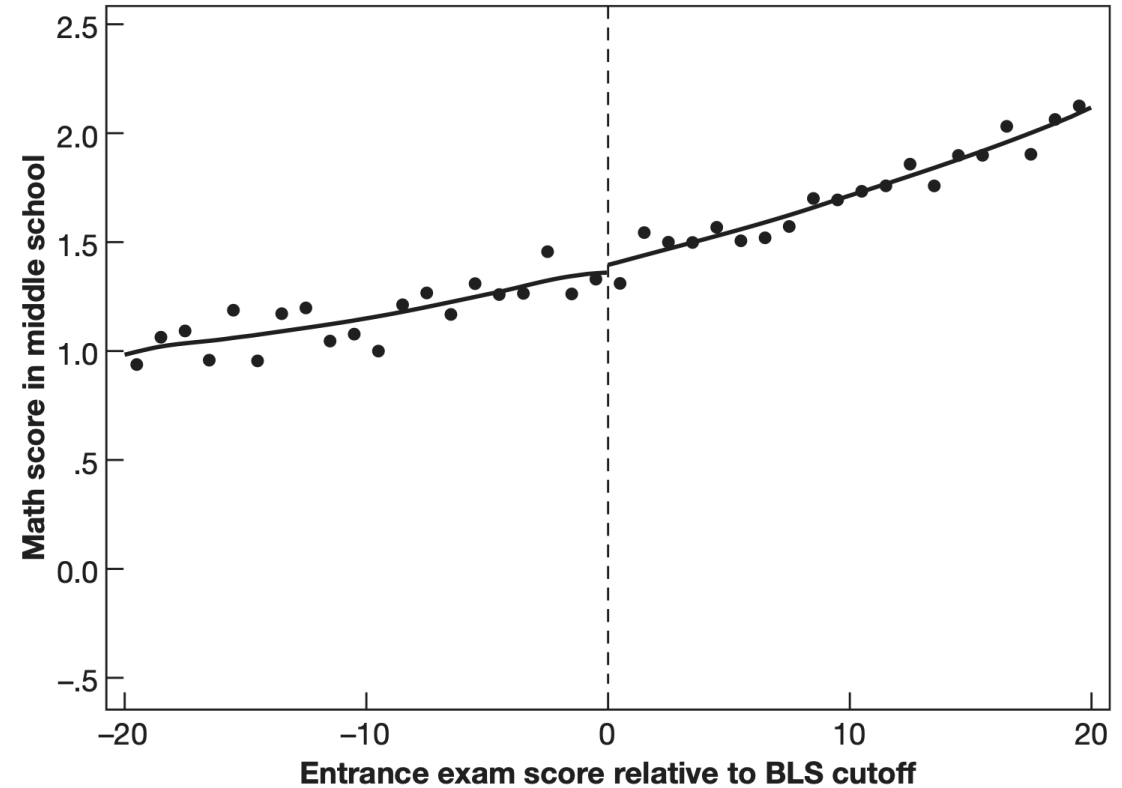


*Notes:* This figure plots seventh- and eighth-grade math scores for applicants to the Boston Latin School (BLS), conditional on admissions test scores, for BLS applicants scoring near the admissions cutoff. Solid lines show fitted values from a local linear regression, estimated separately on either side of the cutoff (indicated by the vertical dashed line).

# Results

- First Stage:  $\phi = 0.8$   
(no SE reported). Strong first stage.
- Reduced Form:  $\rho = -0.02$   
(SE = 0.1). Statistical zero.

FIGURE 4.9  
Math scores around the BLS cutoff

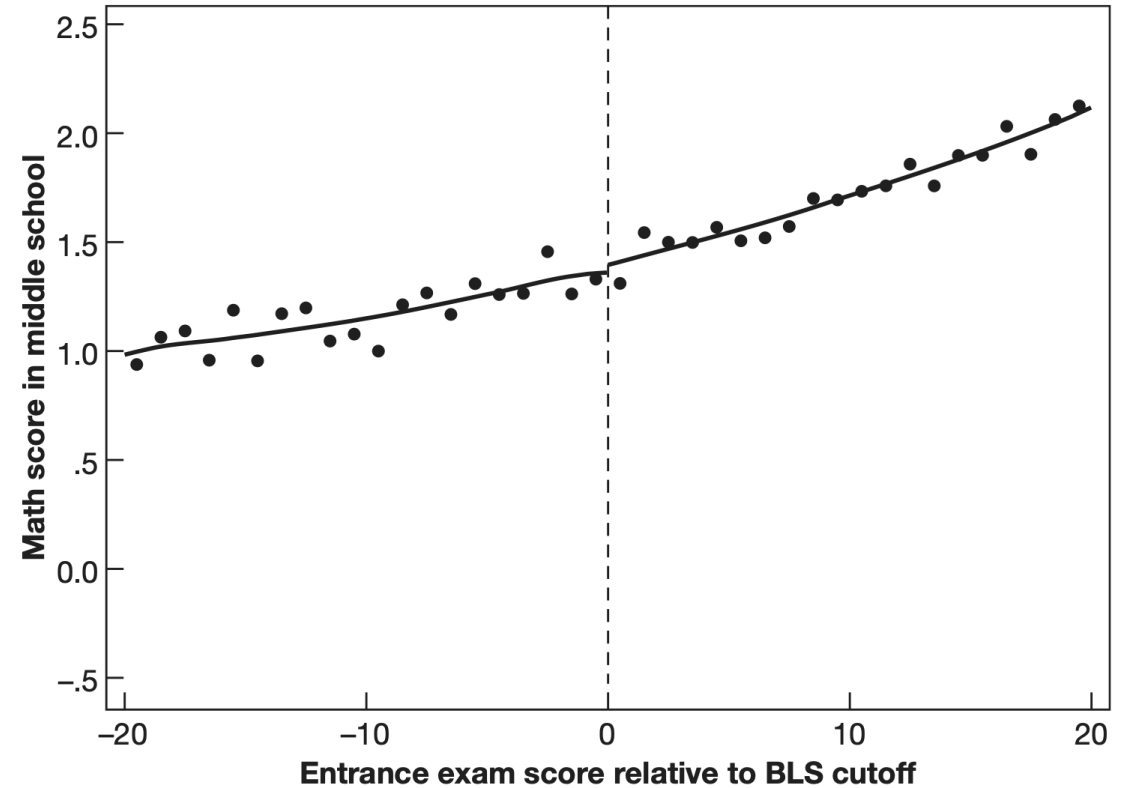


*Notes:* This figure plots seventh- and eighth-grade math scores for applicants to the Boston Latin School (BLS), conditional on admissions test scores, for BLS applicants scoring near the admissions cutoff. Solid lines show fitted values from a local linear regression, estimated separately on either side of the cutoff (indicated by the vertical dashed line).

# Results

- First Stage:  $\phi = 0.8$   
(no SE reported). Strong first stage.
- Reduced Form:  $\rho = -0.02$   
(SE = 0.1). Statistical zero.
- 2SLS LATE:  $\lambda = -0.023$   
(SE = 0.132). Zero again.
- OLS:  $\theta_1 = 0.25$   
(no SE reported). Strong positive "effect"

FIGURE 4.9  
Math scores around the BLS cutoff



*Notes:* This figure plots seventh- and eighth-grade math scores for applicants to the Boston Latin School (BLS), conditional on admissions test scores, for BLS applicants scoring near the admissions cutoff. Solid lines show fitted values from a local linear regression, estimated separately on either side of the cutoff (indicated by the vertical dashed line).

# Back to the Exclusion Restriction

- We saw that the exclusion assumption probably doesn't hold, so why bother with the estimation?
- The key is that the reduce form has zero effect.
- Whatever other channels (of the same instrument) will be captured in the reduce form.
- So no effect in the reduce form for this instrument, means no effect for any treatment/channel this instrument is instrumenting.
- Additionally, an OVB analysis of the OLS estimates shows us that most (all?) potentially omitted variables produce  $OVB > 0$  (practice question for the exam!).

# RDD: Final Considerations 1/2

- Visual inspection of RDD estimates are important but remember to keep an eye on the range of the y-axis
- Notice here that we cannot interpret the result of regression as a matched group, because we do not have individuals in the same cell (say age 20) with both treatment and control. The validity of RDD depends on our willingness to extrapolate across the running variable, at least around a narrow neighborhood around the cut-off.
- This extrapolation limits the policy questions that can be answered with RDD evidence. RDD can answer questions about changes in the margin (from 21 to 22 or 19) but not complete rearrangements of a policy (prohibiting or eliminating restrictions completely).

# RDD: Final Considerations 2/2

- There is one important assumption for RDD that MM does not discuss, and it is pretty important (but I will not test you on it): RDD works as long as the threshold cannot be manipulated. This means that individuals cannot place themselves on either side of the threshold at will. This probably can be connected to the exclusion restriction, but requires a deeper dive into Fuzzy RDD. For those interested in more RDD I suggest following up this class from [Andrew Heiss](#).



# Differences-in-Differences

# Differences-in-Differences (DD)

- Our fifth and last research design tool!
- Aka DD, Diff-in-Diff, Diff-Diff, etc.
- Based on the assumption that sometimes even though treatment and control might differ in unobservables, these differences will be constant over time.

# Policy Example: Effect of Monetary Policy in Times of Crisis 1/3

- Context: Great depression (1930s) in the US.
- Bank runs where a widespread problem during this time
  - A bank run occurs when there is a sudden drop in trust towards the bank's capability to pay back its deposits. No bank holds all its deposits so in the case of a bank run any bank run can go bankrupt.
  - Nowadays there is a clear role for the central banks as lenders of last resort. Back in the 1930s the decision was more discretionary.

# Policy Example: Effect of Monetary Policy in Times of Crisis 2/3

- To avoid bank runs central banks can provide credit to banks at a very low cost.
- The problem with this are two:
  - (i) it prevents the bankruptcy of underperforming (insolvent) banks, at the costs of government funds.
  - (ii) it encourages excessively risky behavior of banks in the future (moral hazard).
- The US Federal Reserve System (Monetary authority in the US) has 12 separate districts, each run by a Federal Reserve Bank.
- In the 1930 each of these banks had significant autonomy in deciding its monetary policy.

# Policy Example: Effect of Monetary Policy in Times of Crisis 3/3

- In December 1930, there was a major bank run in Mississippi (US State).
- It so happens that Mississippi's monetary jurisdiction is split between two of the 12 Federal Reserve Authorities: the 6th and the 8th district.
- It also happens that these authorities reacted very differently to the bank run:
  - 6th District (treatment): made available cheap credit to banks. Expanded bank lending by 40%.
  - 8th District (control): restricted the bank lending by 10%.

# Difference in Difference Estimator

- The DD estimator is defined as the change in outcomes (1st difference) of one group (treatment) over one dimension (typically time) compared to (2nd difference) the same change in another group (control).

$$\delta_{DD} = (\bar{Y}_{T,t+1} - \bar{Y}_{T,t}) - (\bar{Y}_{C,t+1} - \bar{Y}_{C,t})$$

- In this example:
  - Outcomes: Number of banks (later will add number of firms and sales volumes).
  - Groups: 6th and 8th districts.
  - Dimension of change: time.

# Difference in Difference Estimator: Example

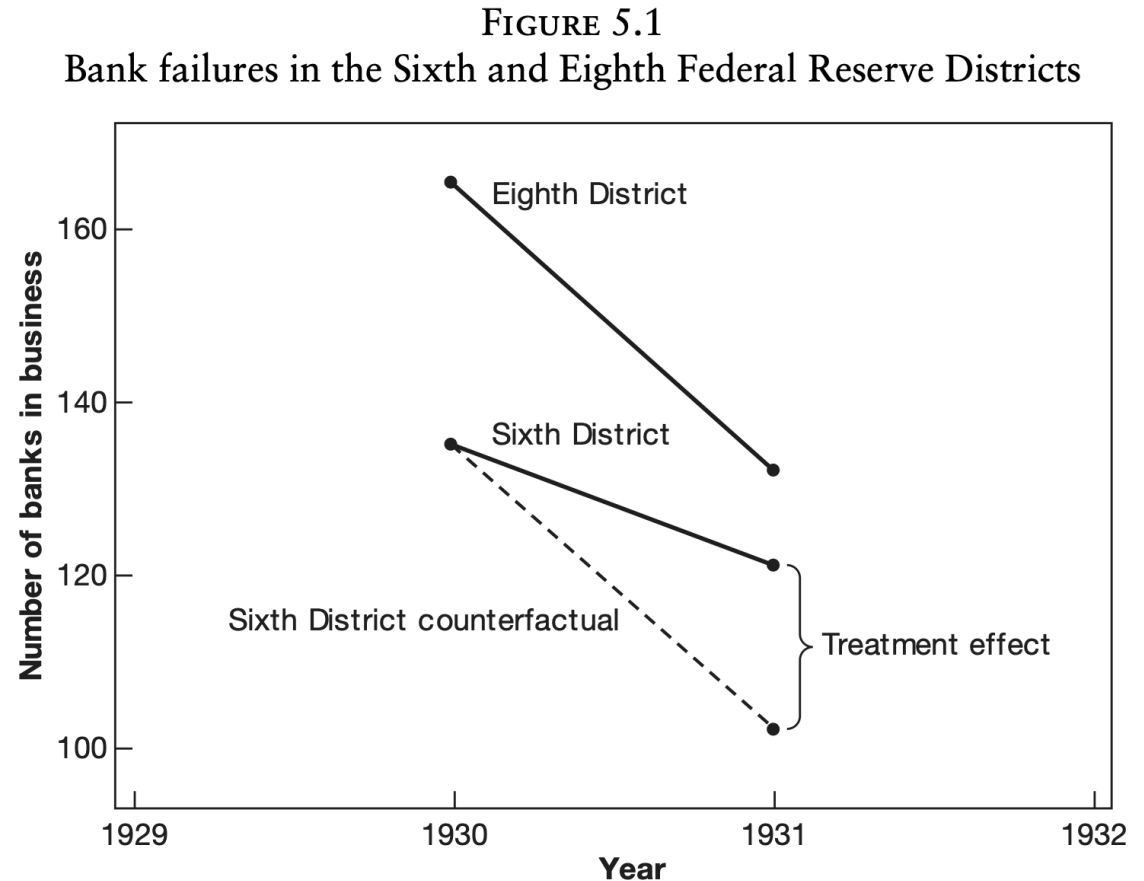
- Change in outcomes of one group over one dimension:  $\bar{Y}_{6th,1931} - \bar{Y}_{6th,1930}$
- Same change in another group:  $\bar{Y}_{8th,1931} - \bar{Y}_{8th,1930}$
- Comparing those two:

$$\begin{aligned}\delta_{DD} &= (\bar{Y}_{6th,1931} - \bar{Y}_{6th,1930}) - (\bar{Y}_{8th,1931} - \bar{Y}_{8th,1930}) \\ &= (121 - 135) - (132 - 165) \\ &= -14 - (-33) \\ &= 19\end{aligned}$$

- Equivalently this can be expressed as the change in comparisons between treatment and control, over time.
- Compare this with a simple difference (in groups, but here each group just has one observation):  $Y_{6th,1931} - \bar{Y}_{8th,1931} = -11$

# Graphically

- “The DD tool amounts to a comparison of trends over time” (MM)
- Think of what is the counterfactual of the treatment group (  $Y_0$  in the terminology of potential outcomes)



*Notes:* This figure shows the number of banks in operation in Mississippi in the Sixth and Eighth Federal Reserve Districts in 1930 and 1931. The dashed line depicts the counterfactual evolution of the number of banks in the Sixth District if the same number of banks had failed in that district in this period as did in the Eighth.

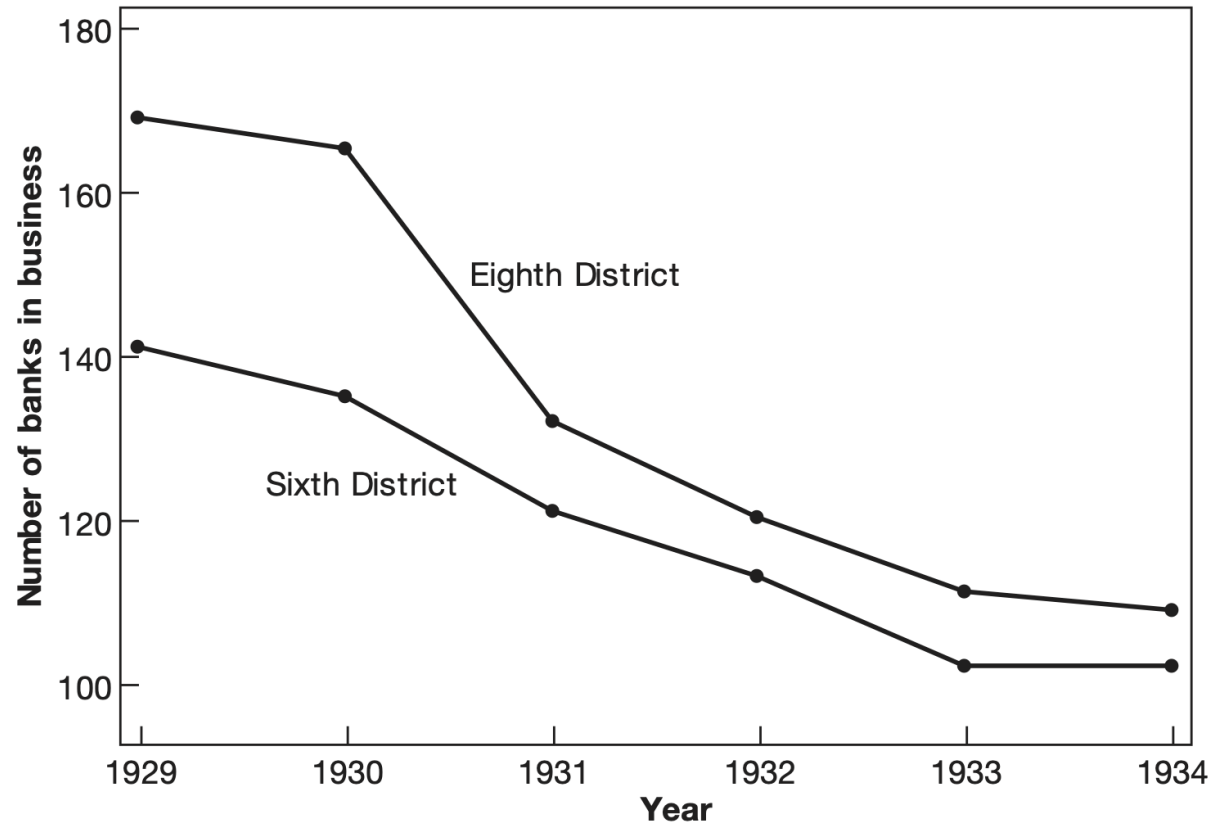


# Key Assumption: Common Trends

- Also known as Parallel Trends Assumption.
- In the absence of an intervention, the treatment and control group would have had the same trend over time.
- In the example: absence the more aggressive lending, the trends in the 6th district would have been the trends of the 8th.
- This is a strong assumption, but can be tested in the data.
  - To test it, we look for trends where the treatments must not have an effect: before the intervention, or after the control group reverse its policy to imitate the treatment (1931)

# Common Trends Graphically

FIGURE 5.2  
Trends in bank failures in the Sixth and Eighth Federal Reserve Districts



*Note:* This figure shows the number of banks in operation in Mississippi in the Sixth and Eighth Federal Reserve Districts between 1929 and 1934.

# DD and Regression 1/2

- Benefits of using regression:
  - Allows to fit any number of observations (not only 4 points!)
  - Allows to implement DD with more than two entities (districts in this example)
  - Facilitates statistical inference.
- Components:
  - (i) A binary variable  $TREAT_d$  that identifies the treated districts *regardless* if the treatment was assign already or not (i.e.  $TREAT_t = 1$  for all  $t$ ).
  - (ii) A binary variable  $POST_t$  that identifies the time period is post treatment or pre-treatment *regardless* of treatment assignment (i.e.  $POST_t = 1$  for controls in the post period too).
  - (iii) The interaction between these two binaries  $TREAT_d \times POST_t$ ; the coefficient on this variable is the DD causal effect.

# DD and Regression 2/2

- Regression equation:

$$Y_{dt} = \alpha + \beta TREAT_d + \gamma POST_t + \delta_{DD}(TREAT_d \times POST_t) + e_{dt}$$

# DD and Regression 2/2

- Regression equation:

$$Y_{dt} = \alpha + \beta TREAT_d + \gamma POST_t + \delta_{DD}(TREAT_d \times POST_t) + e_{dt}$$

- Regression estimates:

$$Y_{dt} = 167 - 29TREAT_d - 49POST_t + 20.5(TREAT_d \times POST_t) + e_{dt}$$

(8.8)                      (7.6)                      (10.7)

- Standard errors of a OLS regression will be too small (overestimate precision) as they assume independent observations.
- Within a unit (district) observations will not be independent, making it less information than with 12 fully independent observations.

# DD Estimates Using Real Outputs

- Beyond number of banks what matters most is a measure of economic activity
- Here there is more limited data (back to the world of 4 points) so we inspect the results without regression.
- DD estimate on number of wholesale firms: 181
- DD estimate on net wholesale sales (\$ millions): 81

TABLE 5.1  
Wholesale firm failures and sales in 1929 and 1933

	1929	1933	Difference (1933–1929)
Panel A. Number of wholesale firms			
Sixth Federal Reserve District (Atlanta)	783	641	–142
Eighth Federal Reserve District (St. Louis)	930	607	–323
Difference (Sixth–Eighth)	–147	34	181
Panel B. Net wholesale sales (\$ million)			
Sixth District Federal Reserve (Atlanta)	141	60	–81
Eighth District Federal Reserve (St. Louis)	245	83	–162
Difference (Sixth–Eighth)	–104	–23	81

*Notes:* This table presents a DD analysis of Federal Reserve liquidity effects on the number of wholesale firms and the dollar value of their sales, paralleling the DD analysis of liquidity effects on bank activity in Figure 5.1.

# Acknowledgments

- MM