Difference-in-Differences

Xiaoyang Ye

xiaoyang.ye@princeton.edu

Example: Dynarski (2003)

Does Aid Matter?
Measuring the Effect of Student Aid on College Attendance and Completion

By SUSAN M. DYNARSKI*

The United States spends billions of dollars each year on financial aid for college students, but there is little evidence that these subsidies serve their goal of increasing college attendance and completion. Determining whether aid affects schooling decisions is an empirical challenge. The traditional approach has been to regress a person's educational attainment against covariates and the aid for which he is eligible and interpret the coefficient on aid as its casual effect. However, this is problematic, as aid eligibility is correlated with many observed and unobserved characteristics that affect schooling decisions. In order to identify the effect of aid, we need a source of variation in aid that is plausibly exogenous to unobservable attributes that influence college attendance. A shift in aid policy that affects some students but not others is one such source of exogenous variation.

cent of full-time college students aged 18 to 21 were receiving Social Security student benefits.[1]

In 1981, Congress voted to eliminate the program. Enrollment sank rapidly (see Figure 1): by the 1984–1985 academic year, program spending had dropped by $3 billion. Except for the introduction of the Pell Grant program in the early 1970's, and the various G.I. bills, this is the largest and sharpest change in grant aid for college students that has ever occurred in the United States. The program's demise provides an opportunity to measure the incentive effects of financial aid. Using difference-in-differences methodology, and proxying for benefit eligibility with the death of a parent during an individual's childhood, I find that the elimination of the Social Security student benefit program reduced college attendance probabilities by more than a third. These estimates suggest that an offer of

Impact of financial aid on college enrollment

- Dynarski (2003) analyzes the elimination in 1982 of a large benefit to college-student children of Social Security recipients who died.
- Treatment group
  - ▷ = individuals with a deceased father
- Post group
  - ▷ = there is one observation per individual, but these individuals graduated high school in different years, which affected whether they were eligible for the benefit

The Model: Difference-in-differences

$$Y = \beta_0 + \beta_1 \textit{Before} + \beta_2 \textit{Father Deceased} + \beta_3 \textit{Father Deceased} * \textit{Before} + \varepsilon$$

	Father not deceased	Father deceased	Difference	DID
Before 1982				
After 1982				
Difference				
DID				

The Model: Difference-in-differences

$$Y = \beta_0 + \beta_1 \textit{Before} + \beta_2 \textit{Father Deceased} + \beta_3 \textit{Father Deceased} * \textit{Before} + \varepsilon$$

	Father not deceased	Father deceased	Difference	DID
Before 1982	$\beta_0 + \beta_1$	$\beta_0 + \beta_1 + \beta_2 + \beta_3$		
After 1982	β_0	$\beta_0 + \beta_2$		
Difference				
DID				

The Model: Difference-in-differences

$$Y = \beta_0 + \beta_1 Before + \beta_2 Father\ Deceased + \beta_3 Father\ Deceased * Before + \varepsilon$$

	Father not deceased	Father deceased	Difference	DID
Before 1982	$\beta_0 + \beta_1$	$\beta_0 + \beta_1 + \beta_2 + \beta_3$	$\beta_2 + \beta_3$	
After 1982	β_0	$\beta_0 + \beta_2$	β_2	
Difference				
DID				

The Model: Difference-in-differences

$$Y = \beta_0 + \beta_1 \textit{Before} + \beta_2 \textit{Father Deceased} + \beta_3 \textit{Father Deceased} * \textit{Before} + \varepsilon$$

	Father not deceased	Father deceased	Difference	DID
Before 1982	$\beta_0 + \beta_1$	$\beta_0 + \beta_1 + \beta_2 + \beta_3$	$\beta_2 + \beta_3$	
After 1982	β_0	$\beta_0 + \beta_2$	β_2	
Difference	β_1	$\beta_1 + \beta_3$		
DID				

The Model: Difference-in-differences

$$Y = \beta_0 + \beta_1 \textit{Before} + \beta_2 \textit{Father Deceased} + \beta_3 \textit{Father Deceased} * \textit{Before} + \varepsilon$$

	Father not deceased	Father deceased	Difference	DID
Before 1982	$\beta_0 + \beta_1$	$\beta_0 + \beta_1 + \beta_2 + \beta_3$	$\beta_2 + \beta_3$	
After 1982	β_0	$\beta_0 + \beta_2$	β_2	
Difference	β_1	$\beta_1 + \beta_3$		β_3
DID			β_3	

Results

TABLE 2—OLS, EFFECT OF ELIGIBILITY FOR STUDENT BENEFITS ON PROBABILITY OF ATTENDING COLLEGE BY AGE 23

	(1) Difference-in-differences	(2) Add covariates
Deceased father × before	0.182	0.219
	(0.096)	(0.102)
Deceased father	−0.123	Y
	(0.083)	
Before	0.026	Y
	(0.021)	
Senior-year family income/ 10,000 ($2,000)		Y
AFQT score		Y
Black		Y
Hispanic		Y
Father attended college		Y
Mother attended college		Y
Single-parent household		Y
Family size		Y
Female		Y
Age in 1988		Y
State dummies		Y
All covariates × before		Y
All covariates × deceased father		Y
R^2	0.002	0.339
Number of observations	3,986	3,986

Notes: Regressions weighted by 1988 sample weights. Standard errors adjusted for heteroskedasticity and multiple observations within households.

Introduction

Potential outcomes

- *Factual* vs. *Counterfactual*

$$Y_i = T_i \cdot Y_i(1) + (1 - T_i) \cdot Y_i(0)$$

  - T_i: a dummy variable indicating whether individual i receives treatment ($T_i = 1$) or not ($T_i = 0$)
  - $Y_i(1)$: the outcome of individual i if she receives treatment
  - $Y_i(0)$: the outcome of individual i if she does not receive treatment

- A valid causality question must involve well-defined causes (treatments, manipulations), and the counterfactuals should be unambiguously defined.

Fundamental problem of causal inference

- *Individual treatment effect*

$$\tau_i = Y_i(1) - Y_i(0)$$

- Causality is defined by potential outcomes, not by realized (observed) outcomes
- We can only observe one of the two potential outcomes
  - ▷ Missing data problem: Any statistical method dealing with treatment effects necessarily imputes the counterfactual part of the data.

Selection bias in observed outcomes

- Holland (1986):

$$\begin{aligned}&\mathbf{E}[Y_i(1)|T_i=1]-\mathbf{E}[Y_i(0)|T_i=0]\\&=\underbrace{\mathbf{E}[Y_i(1)|T_i=1]-\mathbf{E}[Y_i(0)|T_i=1]}_{\tau_{\text{ATT}}}+\underbrace{\mathbf{E}[Y_i(0)|T_i=1]-\mathbf{E}[Y_i(0)|T_i=0]}_{\text{selection bias}}\end{aligned}$$

- Roy model:

$$\begin{aligned}\text{Potential Outcomes:}\quad & Y_i(0)=\mathbf{X}_i\beta(0)+u_i(0)\\ & Y_i(1)=\mathbf{X}_i\beta(1)+u_i(1)\\ \text{Selection/Assignment Mechanism:}\quad & \mathbf{1}_{\{T_i=1\}}=F(\mathbf{X}_i\gamma)+\epsilon_i\end{aligned}$$

  - ▷ The identification is:

$$\mathbf{X}_i\perp(u_i(0),u_i(1),\epsilon_i)$$

Causal inference designs

1. By knowledge of **Assignment Mechanism**
   - Random assignment (RCT)
   - Regression discontinuity (RD)
2. By **Self-Selection**
   - Difference-in-differences (DID)
     - Influence of “other factors” fixed
   - Selection on unobservables and instrumental variables (IV)
     - Conditional on covariates, instrument “as good as randomly assigned” (uncorrelated with potential outcomes)
     - Another structural approach: Heckman selection model
   - Selection on observables and matching (Matching)
     - Conditional on covariates, treatment “as good as randomly assigned”

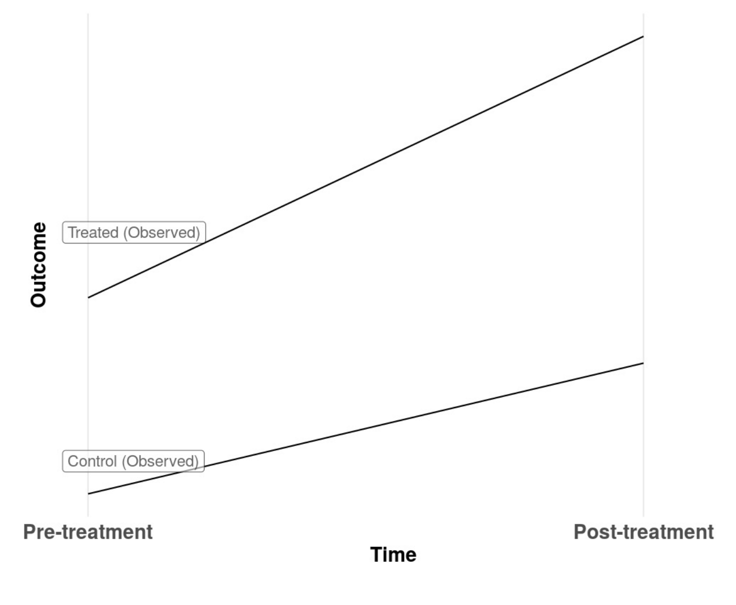
DID

- Use data from the control group to impute untreated outcomes in the treated group
- Arrow of time:

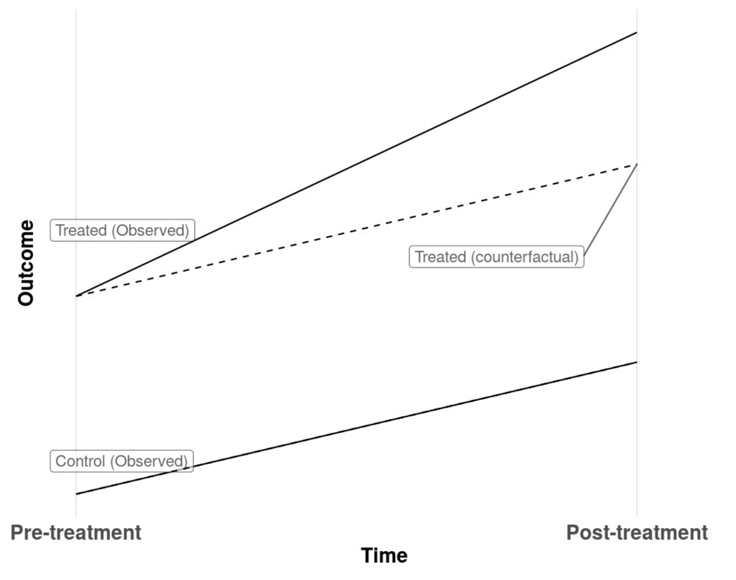
$$Y(t) = Y^0(t) = Y^1(t), for\ t \leq T_0$$

		Potential outcomes			
		Pre-intervention		Post-intervention	
Individual	Group	Untreated	Treated	Untreated	Treated
1	Treated	✓		?	✓
2	Treated	✓		?	✓
3	Treated	✓		?	✓
...					
N-2	Control	✓		✓	?
N-1	Control	✓		✓	?
N	Control	✓		✓	?

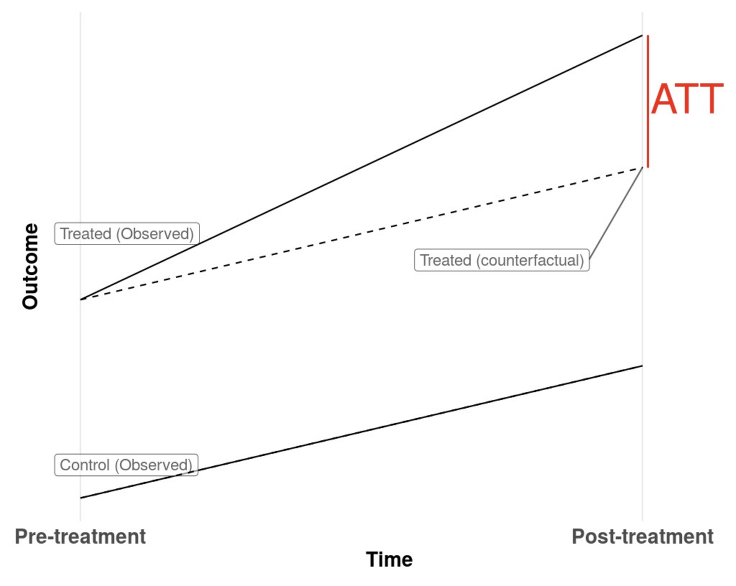
Counterfactual in DID



Counterfactual in DID



Counterfactual in DID



Identification

- Holland (1986):

$$\mathbf{E}[Y_i(1)|T_i = 1] - \mathbf{E}[Y_i(0)|T_i = 0]$$
$$= \underbrace{\mathbf{E}[Y_i(1)|T_i = 1] - \mathbf{E}[Y_i(0)|T_i = 1]}_{\tau_{\text{ATT}}} + \underbrace{\mathbf{E}[Y_i(0)|T_i = 1] - \mathbf{E}[Y_i(0)|T_i = 0]}_{\text{selection bias}}$$

- DID with time machine:

$$Y = \beta_0 + \beta_1 Post + \beta_2 Treatment + \beta_3 Treatment * Post + \varepsilon$$

$$DID\ estimate = \underbrace{\mathbf{E}[Y_i^{post}(1)|T_i = 1] - \mathbf{E}[Y_i^{pre}(0)|T_i = 1]}_{\beta_1+\beta_3} -$$
$$\underbrace{\mathbf{E}[Y_i^{post}(1)|T_i = 0] - \mathbf{E}[Y_i^{pre}(0)|T_i = 0]}_{\beta_1}$$
$$= \beta_3$$

An empirical roadmap

Outline

1. Make assumptions about how the data were generated
2. Connect the untreated outcomes to the observed outcomes
3. Estimate the DID parameter
4. Extensions: Two-way fixed effects and event study
5. Check robustness and sensitivity
6. Related methods

1. Setup: Data generating process

- An exogenous event/treatment
  - Natural experiment
  - Transparent exogenous source of variation that determine treatment assignment (e.g., policy changes, government randomization)
  - Changes should be concentrated around the treatment
- Comparability of the treatment and control groups
  - Recall the counterfactual assumption
- Collect data: {pre, post}* {treatment, control}

2. Parallel trends

1. A picture is worth a thousand words
2. DiD will generally be more plausible if the treatment and control groups are similar in LEVELS to begin with, not just in TRENDS.
    - Any paper should address why the original levels of the experimental and control groups differ, and why we should not think this same mechanism would not impact trends
    - Always show a graph showing the levels of the two series you are comparing over time, not just their difference
        - Alternative: A difference graph + a level comparison table
    - DID on a matched sample for robustness checks

2. Parallel trends

3. Not rejecting the null hypothesis is not equivalent to confirming it
    - Pre-testing is not a substitute for logical reasoning
    - Have an explicit discussion in the paper of why it is reasonable to think the parallel trends assumption is justified, whether there were other policies or sectoral trends going on that might be a threat, etc.
4. Thinking carefully about what sort of violations of parallel trends are plausible, and examining robustness to these
    - Rambachan & Roth, 2019
    - Bilinski & Hatfield, 2019

3. Estimation and inference

- Most commonly used estimator: Regression
- Semiparametric and nonparametric approaches (Athey & Imbens, 2006)
- Matching
- Standard errors
  - Bertrand, Duflo, & Mulainathan (2004), Petersen (2007), Donald & Lang (2007)
  - robust cluster s.e. (to heteroskedasticity and dependence)

4. Extensions

- Fixed effects
  - Multiple time periods or panel data on units
  - Fixed effects eliminate time-specific or unit-specific unobservables
- Event study
  - Estimate the "treatment effect" for each time unit pre and post the event
- Heterogeneous treatment effects
  - DDD
  - Difference in dosage
  - Quantile regressions

5. Robustness and sensitivity tests

- Placebo test: Time
  - ▷ Imagine we artficially move the treatment time to one of those earlier time points (i.e., prior to the time that the treatment was actually received)
  - ▷ In an ideal world, the treatment effect would be null.
- Placebo test: Unit
  - ▷ Estimate impact of policy on a “non-equivalent dependent variable”, i.e., an outcome that should *not* be influenced by the policy but might be influenced by some omitted variable
- Model sensitivity (Candelaria & Shores, 2019)
  - ▷ Secular time trends
  - ▷ Correlated random trends (with different unit levels or different functional forms)
  - ▷ Cross-sectional dependence (different number of common factors)

6. Recent advancements/Related methods

1. Interactive fixed effects
   - (Bai, 2009; Candelaria & Shores, 2019)
2. Synthetic control
   - Original idea by Abadie & Gardeazabal, 2003; Abadie et al., 2010
   - Synthetic DID (Arkhangelsky et al., 2019)
   - Augmented synthetic control (Ben-Michael et al., 2018)
   - Generalized synthetic control (Xu, 2017)
   - similar to matching + DID
3. Comparative interrupted time series (CITS)
   - does not require parallel pre-trends
   - does require a linear model (intercept & slope) to capture the pre-post change

6. Recent advancements/Related methods

4. Instrumental variable for diverging trends due to unobserved confounders
   - (Freyaldenhoven et al., 2019)
   - Use an observed covariate as an instrument for the unobserved confounder (unrelated to treatment)

5. Variation in treatment timing
   - (Goodman-Bacon, 2018)

Back to Dynarski (2003)

Stata practice in 10 minutes

1. Replicate Table 2 Column 1
2. Test parallel pre-trends

Stata practice in 10 minutes

1. Replicate Table 2 Column 1
2. Test parallel pre-trends
3. Plot event study graph
4. Estimate heterogeneous effects by gender, race, or ability (AFQT percentile)

Dynarski (2003) Table 2 Column 1

```
.         * Diff-in-Diff using Regression (table 2, column 1)
.         regress ftby23 f_dead before f_dead_before [aw=wt88], cluster(hhid)
(sum of wgt is 1,302,933,368)

Linear regression                               Number of obs     =      3,986
                                                F(3, 3122)        =       2.19
                                                Prob > F          =     0.0875
                                                R-squared         =     0.0020
                                                Root MSE          =     .49973

                                (Std. Err. adjusted for 3,123 clusters in hhid)
```

ftby23	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
f_dead	-.1234757	.0834565	-1.48	0.139	-.2871109	.0401595
before	.0260081	.0212723	1.22	0.222	-.0157011	.0677173
f_dead_before	.1822297	.0958771	1.90	0.057	-.0057589	.3702183
_cons	.4756935	.018872	25.21	0.000	.4386907	.5126964

Pre-trends

```
.         regress ftby23 f_dead yr1-yr4 f_dead_79 f_dead_80 f_dead_81 f_dead_83  [aw=wt88], cluster
> (hhid)
(sum of wgt is 1,302,933,368)

Linear regression                               Number of obs     =      3,986
                                                F(9, 3122)        =       4.67
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0117
                                                Root MSE          =     .49766

                               (Std. Err. adjusted for 3,123 clusters in hhid)
```

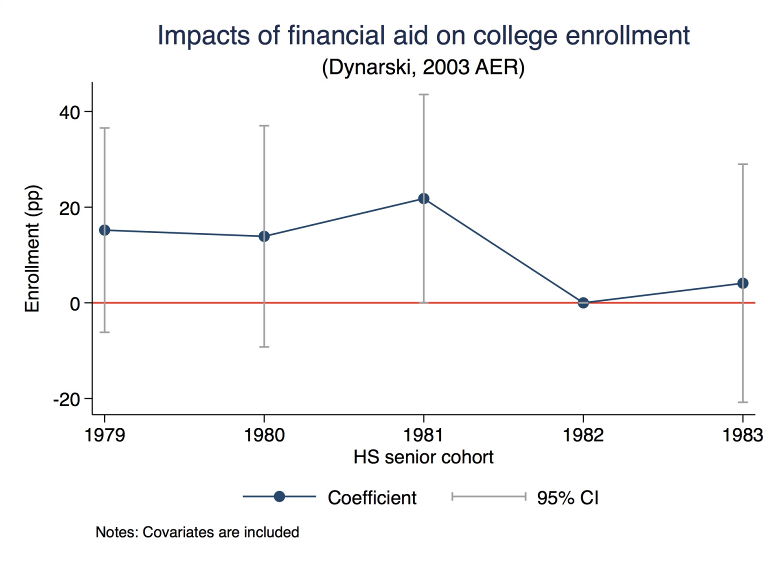
ftby23	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
f_dead	-.1270444	.0975836	-1.30	0.193	-.318379	.0642902
yr1	.1853525	.0442816	4.19	0.000	.0985284	.2721765
yr2	.2264575	.0436311	5.19	0.000	.140909	.3120059
yr3	.2058935	.0442794	4.65	0.000	.1190738	.2927131
yr4	.220878	.0445998	4.95	0.000	.1334302	.3083258
f_dead_79	.2977559	.1261417	2.36	0.018	.0504268	.545085
f_dead_80	.0903417	.1346757	0.67	0.502	-.1737202	.3544035
f_dead_81	.1689355	.1332711	1.27	0.205	-.0923723	.4302434
f_dead_83	.0316845	.1557026	0.20	0.839	-.2736054	.3369745
_cons	.2953085	.0396912	7.44	0.000	.2174851	.3731319

```
.         test f_dead_79+f_dead_80+f_dead_81=0

 ( 1)  f_dead_79 + f_dead_80 + f_dead_81 = 0

       F(  1,  3122) =    2.89
            Prob > F =    0.0890
```

Event study figure



Heterogeneous effects by gender (DDD)

```
.         regress ftby23 f_dead before f_dead_before female f_f_dead f_before f_f_dead_before [aw=w
> t88], cluster(hhid)
(sum of wgt is 1,302,933,368)

Linear regression                               Number of obs     =      3,986
                                                F(7, 3122)        =       2.17
                                                Prob > F          =     0.0342
                                                R-squared         =     0.0044
                                                Root MSE          =     .49936

                                   (Std. Err. adjusted for 3,123 clusters in hhid)
```

ftby23	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
f_dead	-.1752836	.0940599	-1.86	0.062	-.359709	.0091419
before	.0158696	.0296522	0.54	0.593	-.0422702	.0740094
f_dead_before	.20721	.1153185	1.80	0.072	-.0188977	.4333178
female	.0299008	.0372427	0.80	0.422	-.0431218	.1029234
f_f_dead	.1060793	.1618273	0.66	0.512	-.2112194	.423378
f_before	.0202129	.0443298	0.46	0.648	-.0667056	.1071313
f_f_dead_before	-.0516355	.1877993	-0.27	0.783	-.4198582	.3165872
_cons	.4616106	.0253928	18.18	0.000	.4118224	.5113988

Example: School finance reform in China

Economics of Education Review 77 (2020) 101985

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Intergovernmental transfer under heterogeneous accountabilities: The effects of the 2006 Chinese Education Finance Reform☆

Yanqing Ding[a], Fengming Lu[b,c], Xiaoyang Ye*[,d]

[a] Institute of Economics of Education, Graduate School of Education, Peking University
[b] Paul and Marcia Wythes Center on Contemporary China and Department of Politics, Princeton University
[c] Department of Political and Social Change, Coral Bell School of Asia Pacific Affairs, Australian National University
[d] Woodrow Wilson School of Public and International Affairs, Princeton University. Robertson Hall, Princeton, NJ 08544-1013 United States

ARTICLE INFO

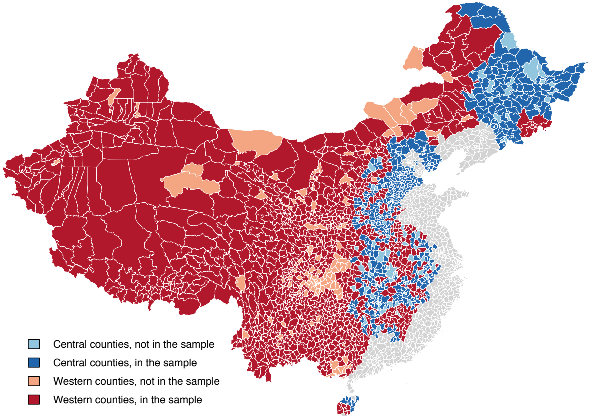
Keywords:
School finance
Intergovernmental transfer
Accountability
Local government incentives

JEL classification:
H52
H75
I22
I28

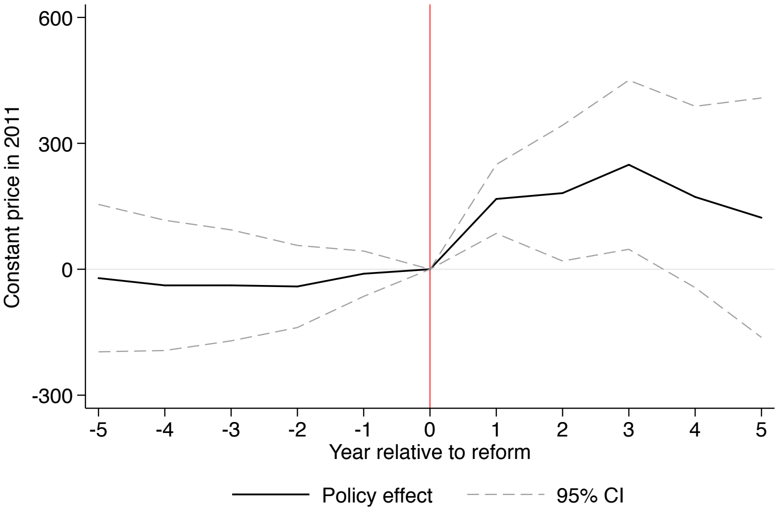
ABSTRACT

While intergovernmental transfers are widely used in improving local education, how local governments in non-democracies allocate fiscal transfers, given they are not electorally accountable, remains unclear. We study the impacts of the 2006 Chinese Education Finance Reform, one of the world's largest education transfer grants, on public school spending. By comparing 1600 Chinese counties that were treated differently in timing and matching ratios, we show natural experimental evidence on how heterogeneous top-down and bottom-up accountabilities affect the allocation of transfer grants. On average, intergovernmental transfers did not increase the total spending levels of local public schools. The causal mechanism is that the transfers crowded out pre-existing local public education investments in extra-budgetary accounts that were not scrutinized and audited by upper-level governments. Heterogeneity analyses further demonstrate that the policy only improved public school spending in counties where public employees had greater means of holding local governments accountable.

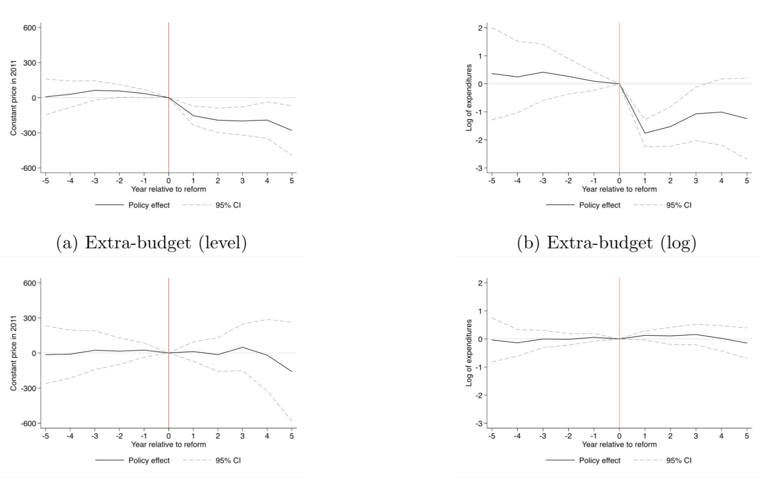
The reform in 2006



Event study graph



Crowding-out effects



(a) Extra-budget (level)

(b) Extra-budget (log)

(c) Total (level)

(d) Total (log)

Fig. 3. Effects of on per-pupil operational spending in rural primary schools. *Notes:* This figure shows estimates from an event study regression with different outcome measures. All the model details are the same as in Fig. 2.

Paper outline

"3. Effects of the transfers on education spending"

- 3.1 Effects on operational expenditures
- 3.2 Other schools as the control group
    - urban schools in east
    - urban schools in west/central regions

Paper outline

"3. Effects of the transfers on education spending"

- 3.1 Effects on operational expenditures
- 3.2 Other schools as the control group
    - urban schools in east
    - urban schools in west/central regions
- 3.3 Compliers as the control group

Paper outline

“3. Effects of the transfers on education spending”

- 3.1 Effects on operational expenditures
- 3.2 Other schools as the control group
  - ▷ urban schools in east
  - ▷ urban schools in west/central regions
- 3.3 Compliers as the control group
- 3.4 Robustness and falsification tests
  - ▷ Choices of samples, measures, weighting

Robustness checks

Table 3

Robustness Checks of the Effects of Additional Central Grants on Budgetary, Total, and Extra-budgetary School Operational Spending (per-pupil) at the County Level, 2002-06 (if not otherwise specified).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Different Winsorizations and Samples								
Sample:	Balanced				Unbalanced	2004-06	2000-06	Drop 4 ARs
Winsorizing:	All	5%	10%	drop 1%	1%	1%	1%	1%
Policy Effect	139.68	140.57	139.10	141.19	139.79	156.67	153.21	153.21
(*Budgetary*)	(46.64)	(44.03)	(41.02)	(45.21)	(45.95)	(36.33)	(44.17)	(52.44)
Policy Effect	35.39	40.15	40.60	44.45	38.78	50.15	30.31	35.39
(*Total*)	(31.76)	(30.59)	(26.78)	(31.71)	(33.84)	(33.64)	(32.10)	(35.33)
Policy Effect	-104.29	-98.93	-97.25	-99.01	-100.65	-105.95	-122.80	-115.23
(*Extra-budgetary*)	(32.14)	(31.09)	(28.72)	(32.93)	(33.13)	(39.28)	(29.47)	(34.35)
Observations	7810	7810	7810	7525	7964	4755	10,542	6345
Clusters (Counties)	1562	1562	1562	1505	1612	1585	1506	1269
B. Different Measures								
	Current prices	Log of outcome	Lagged outcome	# of students in *t*				
Policy Effect	105.88	0.96	136.94	170.51				
(*Budgetary*)	(34.63)	(0.37)	(41.74)	(49.89)				
Policy Effect	28.43	0.15	25.19	79.34				
(*Total*)	(23.54)	(0.10)	(31.06)	(42.57)				
Policy Effect	-77.45	-1.50	-111.27	-101.52				
(*Extra-budgetary*)	(23.70)	(0.31)	(37.32)	(35.52)				
Observations	7810	7810	6248	7810				
Clusters (Counties)	1562	1562	1562	1562				
C. Different Weightings								
	P score weighting	P score × # of students	DFL weighting	Not weighted				
Policy Effect	131.56	158.69	161.58	112.01				
(*Budgetary*)	(44.52)	(43.02)	(52.14)	(47.61)				
Policy Effect	41.05	48.57	21.14	12.56				
(*Total*)	(33.50)	(31.08)	(41.58)	(37.90)				
Policy Effect	-99.37	-110.44	-139.66	-103.05				
(*Extra-budgetary*)	(34.36)	(34.74)	(51.52)	(28.67)				
Observations	7775	7775	7145	7810				
Clusters (Counties)	1555	1555	1551	1562				

Paper outline

“3. Effects of the transfers on education spending”

- 3.4 Robustness and falsification tests
  - ▷ Choices of samples, measures, weighting
  - ▷ High school spending

Paper outline

“3. Effects of the transfers on education spending”

- 3.4 Robustness and falsification tests
  - ▷ Choices of samples, measures, weighting
  - ▷ High school spending
- 3.5 Effects on other public education outcomes
  - ▷ Decrease in teacher salary (constant price)

Paper outline

"3. Effects of the transfers on education spending"

- 3.4 Robustness and falsification tests
  - ▷ Choices of samples, measures, weighting
  - ▷ High school spending
- 3.5 Effects on other public education outcomes
  - ▷ Decrease in teacher salary (constant price)
  - ▷ Small change in Gini coefficient

Paper outline

"3. Effects of the transfers on education spending"

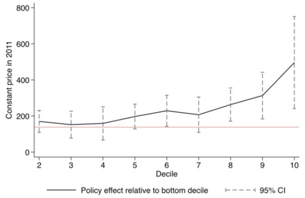
- 3.4 Robustness and falsification tests
  - ▷ Choices of samples, measures, weighting
  - ▷ High school spending
- 3.5 Effects on other public education outcomes
  - ▷ Decrease in teacher salary (constant price)
  - ▷ Small change in Gini coefficient
- 3.6 Null effects on other social spending

Paper outline

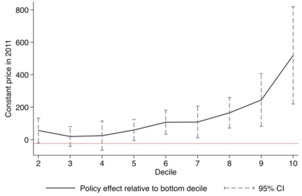
“3. Effects of the transfers on education spending”

- 3.4 Robustness and falsification tests
  - Choices of samples, measures, weighting
  - High school spending
- 3.5 Effects on other public education outcomes
  - Decrease in teacher salary (constant price)
  - Small change in Gini coefficient
- 3.6 Null effects on other social spending
- 3.7 Where do “missing” funds end up

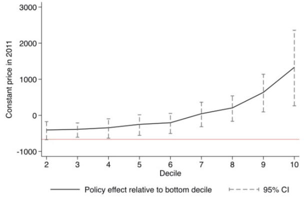
Heterogeneity analysis



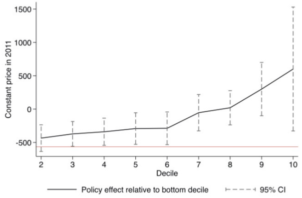
(a) Budgetary operational spending



(b) Total operational spending



(c) Total spending

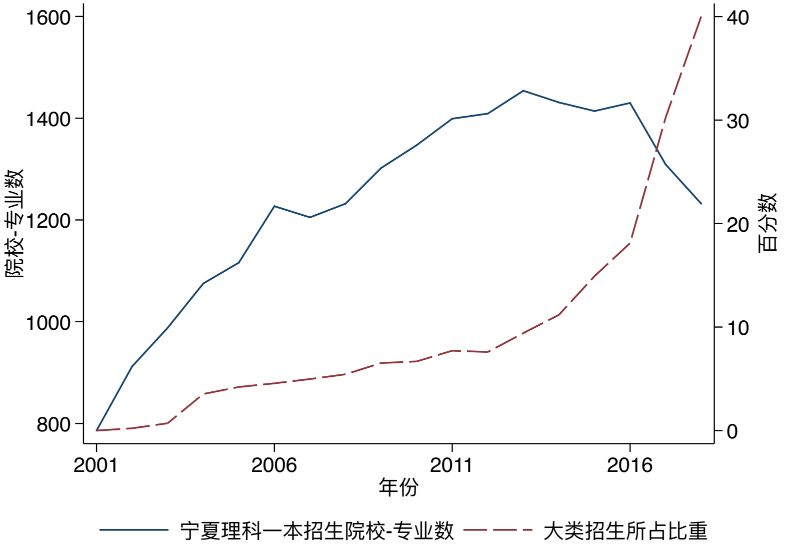


(d) Full-time teachers' salaries

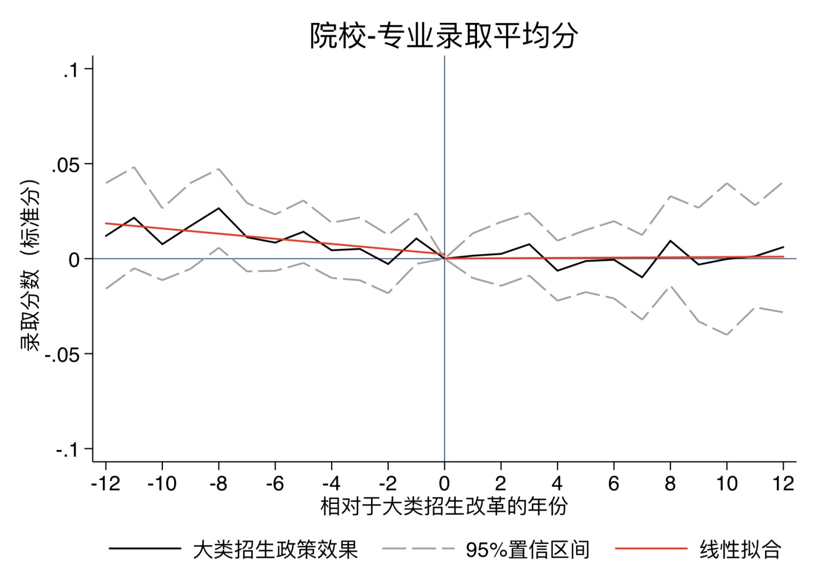
Fig. 8. Effects of intergovernmental transfers by public employment size decile. *Notes:* This figure shows estimates from the DDD regressions by decile of relative sizes of public employment, corresponding the estimates in Table 7. The estimates are the differences between each decile and the bottom decile (interaction effects). All the spending outcomes are measured as per-pupil.

Example: College admissions reform in China

College-major to college-field admissions



Null effects and why?



Summary

Guiding questions of a DID study

1. Ideal experiment
   - How could you use an RCT to answer this causal question?
2. Identification strategy
   - How does the study use observational data to approximate an ideal experiment?

Guiding questions of a DID study

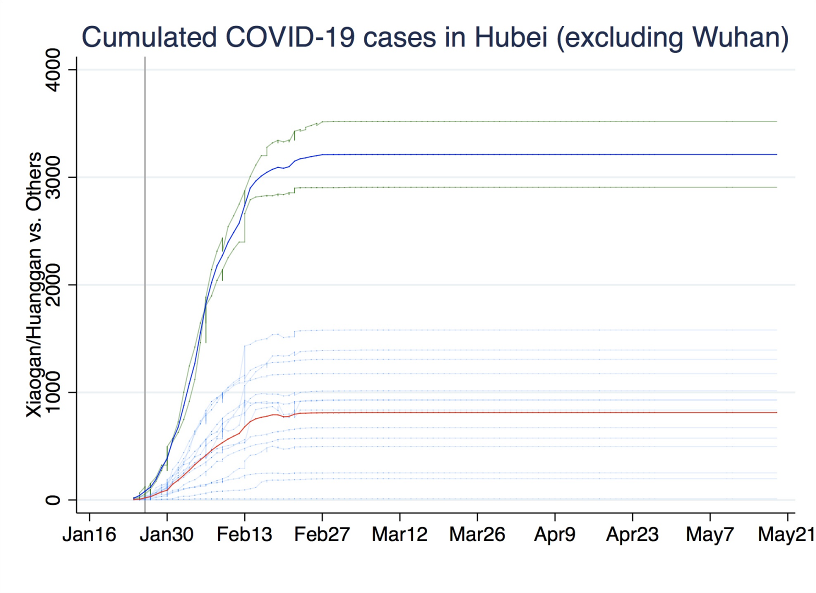
1. Ideal experiment
   - How could you use an RCT to answer this causal question?
2. Identification strategy
   - How does the study use observational data to approximate an ideal experiment?
3. Internal validity
   - First think about (and discuss) why treatment status varies in general. Do people choose treatment? Is it chosen for them? By what process? Is the comparison group plausible?
   - Then describe the identifying variation in this study. This is the variation left after we control for the other variables.
   - What are the key threats to the internal validity of the study?

Guiding questions of a DID study

4. External validity
   - To what populations, programs and places can the results be safely extrapolated?
5. Implication
   - Do the conclusions and/or recommendations follow logically from the empirical evidence presented?

Final discussion

Using DID to estimate the COVID-19 effect



Final discussion

Using DID to estimate the COVID-19 effect (Goodman-Bacon & Marcus, 2020)

- Different policies across place and time
- DID paper racing
  - At least five recent papers use DD methods to show that non-pharmaceutical interventions reduce interactions, infections, or deaths (Dave et al., 2020; Fang et al., 2020; Friedson et al., 2020; Gupta et al., 2020; Hsiang et al., 2020).

Using DID to estimate the COVID-19 effect (Goodman-Bacon & Marcus, 2020)

Challenges

- Packaged policies
- Reverse causality
- Voluntary precautions
- Anticipation
- Spillovers
- Variation in policy timing
- Measurement and scaling of the dependent variable

Using DID to estimate the COVID-19 effect (Goodman-Bacon & Marcus, 2020)

Recommendations

- Estimate dynamics (event study)
- Choose the control group wisely
- Be careful of regression DID
- Sign the bias
- Be clear about what is knowable

Thanks!

Acknowledgments

I have borrowed from various sources to prepare this lecture, including
Brian Jacob's PP 639,
Angrist and Pischke's MHE,
David Mckenzie's blog,
https://www.scunning.com/causalinference_norap.pdf, and
https://diff.healthpolicydatascience.org/.