Lecture 9: Returns to Education and Differences in Differences

Hammad Shaikh and Natalie Bau

Duflo (2001) Introduction

- Duflo uses a natural experiment to measure the returns to schooling

Experimental environment in observational data

- During a short period of time (around 1974), Indonesia built 61,000 schools
  - Targeted at students aged 5 to 14 years old

↳ Primary schools

- Goal is to measure causal impact of this program on educational attainment and earnings

↑schools ⇒ ↑Educ ⇒ ↑wages

returns to education

Duflo (2001) Key Ideas

- Potential problem: schools were built all over Indonesia

Don't have standard control - Treatment

- Some districts got a lot of schools (per student), while others got fewer

Variation in intensity of treatment

- Young children in 1974 are much more affected by the program than older students

- Solution: use variation in **exposure** to school construction program

High Exposure: High intesity + young

Low Exposure: Low intensity + old

Duflo (2001) Endogenity Problems

- Compare individuals in areas that got more schools before and after the school construction has time trend issue

↑Educ over time

- Comparing individuals who were young enough to benefit in places that got more or less schools has selection problem

More schools built in poorer areas

- Solution: Examine how educational attainment changed from the older to the younger cohorts in the places that received more schools relative to places that received fewer schools

Young-old and Low intensity - High intensity environment

Duflo (2001) Differences - Differences Setup

Y = Yrs of Education

- Children born in districts that got more schools, but were born too early to benefit from those schools

Old - High Intensity

- Children born in districts that got more schools and were born late enough to benefit

young - High Intensity (Treatment)

Treat + trend

- Children born in districts that got less schools and were born too early to benefit anyway

Old - low Intensity

- Children born in districts that got less schools and were born later

young - low Intensity

Trend

Duflo (2001) Simplified Model

- $y_{ijk} = \beta_0 + \beta_1 C_j + \beta_2 T_i + \beta_3 P_j + \beta_4 P_j \times T_i + \epsilon_{ijk}$

Wages or Education diff-diff.

- i indexes individual, j indexes districts, and k indexes cohort

- C_j is district level characteristics

  ↳ water & sanitation quality

- $T_i = 1$ if young cohort (2 to 6) and 0 if old cohort (12 to 17)

- $P_j = 1$ if high intensity and 0 if low intensity

Duflo (2001) Diff-Diff Results

→ Program ↑ Yrs Educ. by 0.12 years

→ Program ↑ wages by 2.6%.

TABLE 3—MEANS OF EDUCATION AND LOG(WAGE) BY COHORT AND LEVEL OF PROGRAM CELLS

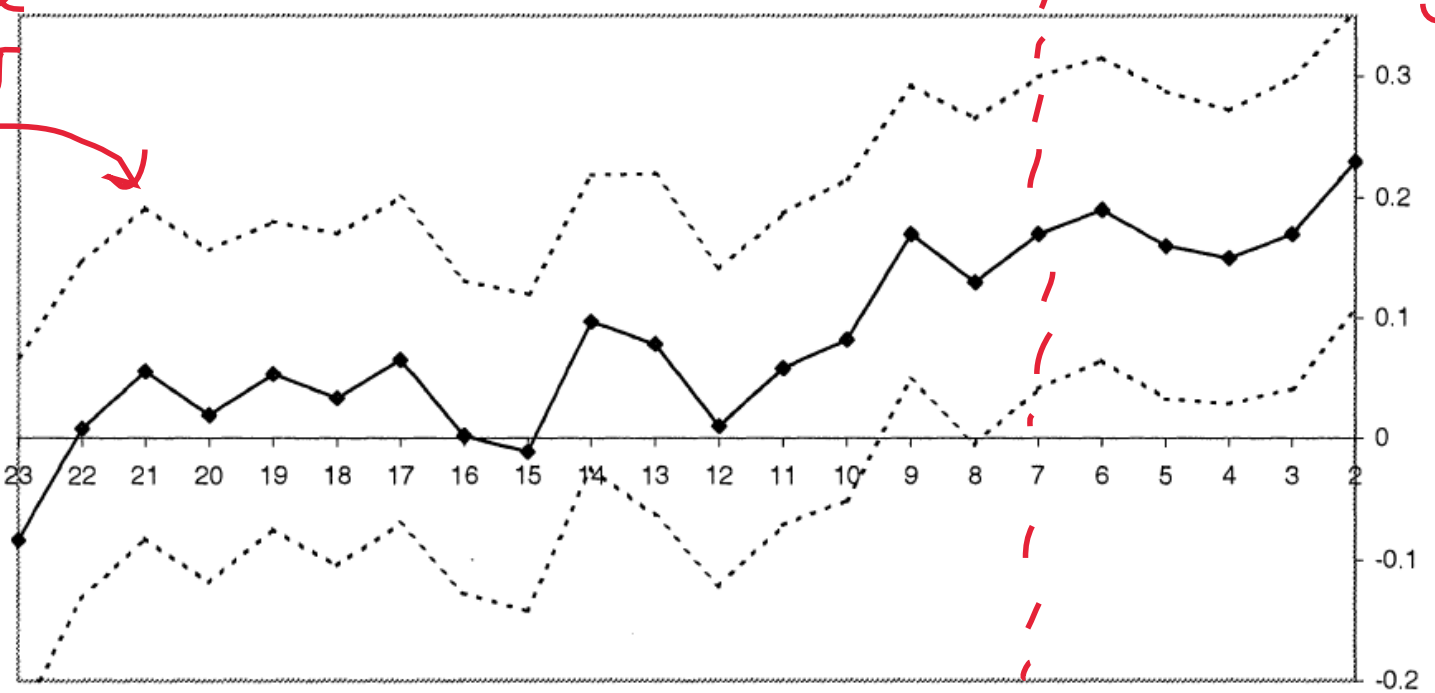
	Years of education			Log(wages)		
	Level of program in region of birth			Level of program in region of birth		
	High (1)	Low (2)	Difference (3)	High (4)	Low (5)	Difference (6)
Panel A: Experiment of Interest						
Aged 2 to 6 in 1974	8.49	9.76	−1.27	6.61	6.73	−0.12
	(0.043)	(0.037)	(0.057)	(0.0078)	(0.0064)	(0.010)
Aged 12 to 17 in 1974	8.02	9.40	−1.39	6.87	7.02	−0.15
	(0.053)	(0.042)	(0.067)	(0.0085)	(0.0069)	(0.011)
Difference	0.47	0.36	0.12	−0.26	−0.29	0.026
	(0.070)	(0.038)	(0.089)	(0.011)	(0.0096)	(0.015)

- Assumption: Time trends are parallel (in the absence of the schools, educational attainment would grow at the same rate in both high/low intensity places over time)

Duflo (2001) Visual Results

= 1 if age 2 in 1974, 0 if age 24

- Plots coefficients of (Age in 1974) $\times$ (Program Intensity)



School construction is not as beneficial

Benefit from being in high intensity

Duflo (2001) Returns to Education

- $log(wage_{ijk}) = \beta_0 + \beta_1 DistrictCharac_j + \beta_2 YrsEduc_{ijk} + \epsilon_{ijk}$

$Cov(YrsEduc, \epsilon) \neq 0$ is problem

- School construction program creates exogenous variation in educational attainment

- (Age in 1974) $\times$ (Program Intensity) IV for $YrsEduc_{ijk}$

IV assump.

1) Program has no direct effect on wages ✓
2) IV related to YrsEduc ✓
3) IV randomly assigned (Maybe)

Duflo (2001) Returns to Education Results

Method	Instrument	(1)
Panel A: Sample of Wage Earners		
Panel A1: Dependent variable: log(hourly wage)		
OLS		0.0776
		(0.000620)
2SLS	Year of birth dummies*program intensity in region of birth	0.0675
		(0.0280)
		[0.96]
2SLS	(Aged 2–6 in 1974)*program intensity in region of birth	0.0752
		(0.0338)
		(0.0338)

- Extra year of education increases wages by 7% on average

Summary of Duflo (2001)

- Each primary school per 1000 children increased educational attainment by at least 0.12 years on average
- The increase in educational attainment due to the program increased wages by at least 1.5 percent on average
- Extra year of education causes at least 7 percent increase in wages on average
- Potential short comings:
  - Very old program in 1974 (results may differ now)
  - Not ideal control and treatment group

→ Only data on men