Applied Microeconometrics I

Lecture 12: Short recap

Tuomas Pekkarinen

Aalto University

October 13, 2022

`Lecture Slides`

Short recap of the course

Example: Returns to schooling

- We will now shortly go over the methods that have been covered in this course
- We do this with a help of an example: Returns to schooling
- Whether education really increases earnings is one of the classic questions in economics
- Subject on intensive study since Jacob Mincer's work in the 1960's
- Methods used: DD, IV, RDD

Short recap of the course

Example: Returns to schooling

- Early work on returns to schooling relied on identification based on observables
- Typical regression would look like this

  $$\log Y_i = \alpha + \rho S_i + \beta_1 X_i + \beta_2 X_i^2 + \epsilon_i$$

  where $\log Y_i$ is the logarithm of annual earnings, S_i is years of education, and X is potential work experience
- How credible is the CIA assumption here?

- Which factors are we omitting, when trying to estimate returns to schooling relying on identification based on observables?
- Denote "ability" with A_i
- Suppose that real regression of Y_i on S_i looks like this (we ignore experience X for convenience):

$$\log Y_i = \alpha + \rho S_i + \gamma A_i + \epsilon_i$$

- If A_i is not observable and is omitted from the regression, our estimates are biased:

$$\hat{\rho} = \rho + \gamma \frac{Cov(S, A)}{Var(S)}$$

- $\gamma \frac{Cov(S,A)}{Var(S)}$ is the "ability bias"
- What is the likely sign of this bias?

Short recap of the course

Example: Returns to schooling

- Many early studies tried to control for ability with proxies such as measures of IQ
- Think of this strategy in terms of potential outcomes
- Suppose that $D_i = 1$ if individual has graduated from university
- Then the observed earnings difference between university graduates and non-graduates conditional on IQ can be written as:

$$E[Y_i|D_i = 1, \text{IQ}] - E[Y_i|D_i = 0, \text{IQ}] =$$
$$E[Y_{1i} - Y_{0i}|D_i = 1, \text{IQ}] - \{E[Y_{0i}|D_i = 1, \text{IQ}] - E[Y_{0i}|D_i = 0, \text{IQ}]\}$$

- Two serious problems with this strategy:
  1. IQ may not capture all relevant abilities
  2. IQ may be a bad control

- Introducing controls can give rise to more problems than they actually solve
- Bad controls: Control variables that are themselves outcomes caused by our causal variable of interest
- For example think of controlling for white collar status
- For simplicity assume that going to college is randomly assigned

- Presumably going to college has a positive effect on the probability of working in a white collar occupation
- We can distinguish three groups of workers based on the effect of college on their white collar status
  - AB: workers who are always blue collar workers
  - AW: workers who are always white collar workers
  - BW: workers who are white collar only if they went to college
- Our goal is to estimate returns to schooling, controlling for white collar status

Bad control example from Mastering 'Metrics

TABLE 6.1
How bad control creates selection bias

	Potential occupation		Potential earnings		Average earnings by occupation	
Type of worker	Without college (1)	With college (2)	Without college (3)	With college (4)	Without college (5)	With college (6)
Always Blue (AB)	Blue	Blue	1,000	1,500	Blue 1,500	Blue 1,500
Blue White (BW)	Blue	White	2,000	2,500		White 3,000
Always White (AW)	White	White	3,000	3,500	White 3,000	

Short recap of the course

Example: Returns to schooling

- Limiting the college/non-college comparisons to those who have white collar jobs leads us to conclude that returns to college are zero
- However, the average effect of going to college is 500
- Conditioning on bad controls changes the composition of the treatment and control group

Short recap of the course

Example: Returns to schooling

- In lecture 9, we saw how we can use differences-in-differences to estimate the effect of schooling on earnings when we have access to data on twins

$$\text{Twin 1: } Y_{1f} = \alpha + \rho S_{1f} + \gamma A_f + \epsilon_{1f}$$
$$\text{Twin 2: } Y_{2f} = \alpha + \rho S_{2f} + \gamma A_f + \epsilon_{2f}$$

- If A_f is common to the pair of twins, then differencing yields:

$$Y_{1f} - Y_{2f} = \rho(S_{1f} - S_{2f}) + (\epsilon_{1f} - \epsilon_{2f})$$

- Under these assumptions estimating ρ with the differenced equation gives us the causal effect of schooling on earnings

OLS estimates in the population and in the twin sample

TABLE II
OLS ESTIMATES OF THE (MEAN) RETURN TO SCHOOLING USING THE CPS AND TWINS DATA

	CPS[a]	Identical twins
	OLS (1)	OLS (2)
Own education	0.085	0.110
	(0.0003)	(0.009)
Age	0.071	0.104
	(0.0004)	(0.010)
Age^2 (÷100)	−0.074	−0.106
	(0.0005)	(0.013)
Female	−0.253	−0.318
	(0.001)	(0.040)
White	0.087	−0.100
	(0.002)	(0.072)
Sample size	476,851	680
R^2	0.332	0.339

Standard errors are in parentheses. All regressions include a constant.

a. The Current Population Survey (CPS) sample is drawn from the 1991–1993 Outgoing Rotation Group files. The sample includes workers age 18–65 with an hourly wage greater than $1 per hour in 1993 dollars; the regression is weighted using the earnings weight. We converted the 1992 and 1993 education categories into a continuous measure according to the categorization suggested by Park [1994].

First difference estimates

	GLS (1)	GLS (2)	3SLS (3)	First-difference (4)	First-difference by IV (5)
Own education	0.102	0.066	0.091	0.070	0.088
	(0.010)	(0.018)	(0.024)	(0.019)	(0.025)
Avg. education		0.051	0.033		
$[(S_1 + S_2)/2]$		(0.022)	(0.028)		
Age	0.104	0.103	0.103		
	(0.013)	(0.013)	(0.013)		
Age^2 (÷100)	−0.107	−0.104	−0.104		
	(0.015)	(0.015)	(0.015)		
Female	−0.315	−0.309	−0.306		
	(0.049)	(0.049)	(0.049)		
White	−0.106	−0.105	−0.101		
	(0.090)	(0.091)	(0.091)		
Covered by a union					
Married					
Tenure (years)					
Sample size	680	680	680	340	340
R^2	0.262	0.264	0.267	0.039	

- This strategy is very sensitive to measurement error
- Ashenfelter and Rouse solution: Assume that twins report each other's schooling with independent measurement errors
- Then we can use one's twins reporting of one's own schooling as an instrument for one's own reporting
- Intuition: Both my recollection and my twin sibling's recollection are mismeasured assessments of my real level of schooling. Instrumenting my own recollection with my sibling's recollection will clean away the measurement error

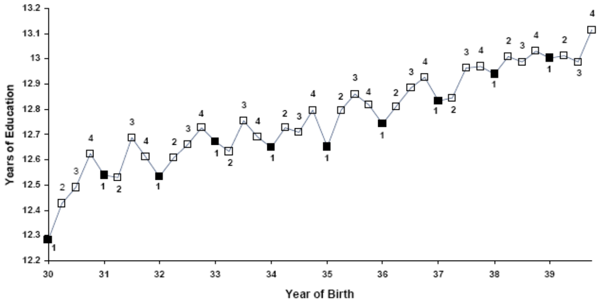
Short recap of the course

Example: Returns to schooling

- In Lecture 7, we saw how one could use instrumental variables to estimate the returns to schooling
- Angrist and Krueger: Quarter of birth as an instrument for schooling
- Students enter schooling in the September of the calendar year in which they turn 6
- And compulsory school law requires them to remain in school until they become 16
- Hence people born late in the year are more likely to stay at school longer

Is the first stage right?

A. Average Education by Quarter of Birth (first stage)



The reduced form for earnings

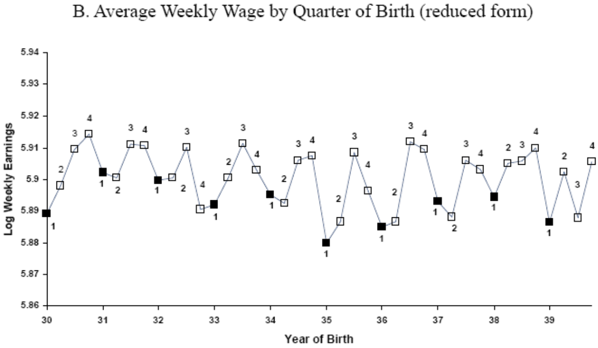


Table 4.1.2: Wald estimates of the returns to schooling using quarter of birth instruments

	(1) Born in the 1st or 2nd quarter of year	(2) Born in the 3rd or 4th quarter of year	(3) Difference (std. error) (1)-(2)
ln (weekly wage)	5.8916	5.9051	-0.01349 (0.00337)
Years of education	12.6881	12.8394	-0.1514 (0.0162)
Wald estimate of return to education			0.0891 (0.0210)
OLS estimate of return to education			0.0703 (0.0005)

Notes: Adapted from a re-analysis of Angrist and Krueger (1991) by Angrist and Imbens (1995). The sample includes native-born men with positive earnings from the 1930-39 birth cohorts in the 1980 Census 5 percent file. The sample size is 329,509.

- IV estimates the local average treatment effect (LATE) which is often different from the average treatment effect on the treated
- How is the effect local in the Angrist and Krueger case?
- When is LATE the same as ATT?

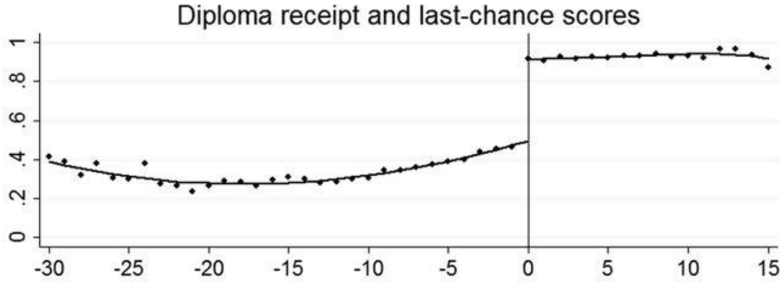
- Finally, we go over an RDD example on the effects of schooling
- Particular question, what is the effect of the high school diploma as such?
- Sheepskin effect: The effect of diploma as a piece of paper, ceteris paribus

Short recap of the course

Example: Do degrees matter? Clark and Martorelli, 2014

- In Texas, getting a high school diploma is conditional on passing an exit exam
- Clark and Martorelli exploit the fact that the probability of getting the diploma jumps discontinuously at the passing of exit exam threshold, to identify the effect of diplomas on earnings
- Identifying assumption: Getting a diploma is randomly assigned near the passing threshold
- Results:
  - The probability of getting the diploma increases by 50 percentage points at the passing threshold
  - Yet, the earnings don't change discontinuously
  - No evidence of sheepskin effects

Clark and Martorelli: The first stage



Clark and Martorelli: Reduced form