Lectures 08-09:
Paper overviews

PPHA 34600
Prof. Fiona Burlig

Harris School of Public Policy
University of Chicago

Lecture 08 recap

TL;DR:

1. Instrumental variables are very powerful
2. ...but they require extremely strong assumptions!
3. Hashtag no free lunch

An example: Impacts of parental incarceration on kids

Policy issue:

- Millions of Americans are in jail annually
- This likely impacts not only them, but also their families
- What is the impact of family member incarceration?

Approach:

- (We're not actually evaluating a program here)
- We need a shock to incarceration

$\rightarrow$ We don't have randomization, so we use IV

- Instrument of choice: random assignment to strict/harsh judges

$\rightarrow$ Do we believe this? Hold that thought...

Treatment effects of parental incarceration on kids

How does parental incarceration affect kids (simplified)?

First stage:

$$\text{Parental incarceration}_i = \alpha + \gamma \text{Judge leniency}_{ij} + \beta X_{ij} + \eta_{ij}$$

where

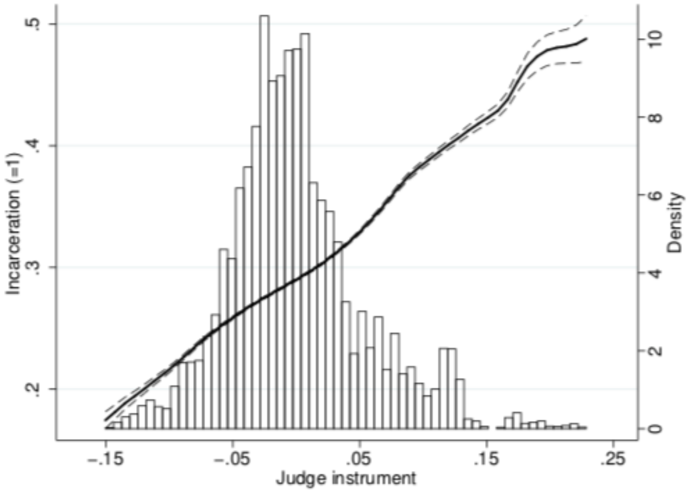
Parental incarceration$_i$ is equal to 1 if i's parent is jailed and 0 otherwise

Judge leniency$_{ij}$ is the strictness of judge j

X_{jt} are controls

η_{it} is an error term

First stage (graphically)



Treatment effects of parental incarceration on kids

How does parental incarceration affect kids (simplified)?

First stage:

$$\text{Parental incarceration}_i = \alpha + \gamma \text{Judge leniency}_{ij} + \beta X_{ij} + \eta_{ij}$$

where
Parental incarceration$_i$ is equal to 1 if i's parent is jailed and 0 otherwise
Judge leniency$_{ij}$ is the strictness of judge j
X_{jt} are controls
η_{it} is an error term

Second stage:

$$Y_i = \alpha + \delta \text{parental } \widehat{\text{incarceration}}_i + \tau X_{ij} + \eta_i$$

where
Y_i are outcomes for child i parental $\widehat{\text{incarceration}}_i$ is the fitted values from the first stage

Second stage (OLS)

Table A4: Effect of parental incarceration on child criminal activity, OLS comparison

	Extensive margin (=1)			Intensive margin (IHS)		
	Charged	Convicted	Incarcerated	Charged	Convicted	Incarcerated
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Criminal activity before age 25 (OLS with no controls)						
Parent incarcerated (=1)	0.024***	0.024***	0.015***	0.054***	0.042***	0.030***
	(0.005)	(0.005)	(0.004)	(0.011)	(0.009)	(0.007)
Index p-value			0.000			0.000
Dependent mean	0.325	0.247	0.124	0.568	0.375	0.205
Observations	83,532	83,532	83,532	83,532	83,532	83,532
Panel B: Criminal activity before age 25 (OLS with controls)						
Parent incarcerated (=1)	-0.004	-0.001	-0.001	-0.009	-0.004	0.000
	(0.005)	(0.005)	(0.003)	(0.010)	(0.008)	(0.006)
Index p-value			0.645			0.645
Dependent mean	0.325	0.247	0.124	0.568	0.375	0.205
Observations	83,532	83,532	83,532	83,532	83,532	83,532

Second stage (IV)

	Extensive margin (=1)			Intensive margin (IHS)		
	Charged	Convicted	Incarcerated	Charged	Convicted	Incarcerated
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Criminal activity before age 25						
Parent incarcerated (=1)	-0.066**	-0.055**	-0.049**	-0.156**	-0.097**	-0.076**
	(0.030)	(0.027)	(0.020)	(0.061)	(0.045)	(0.035)
Index p-value			0.011			0.013
Dependent mean	0.325	0.247	0.124	0.568	0.375	0.205
Observations	83,532	83,532	83,532	83,532	83,532	83,532
Panel B: Juvenile criminal activity						
Parent incarcerated (=1)	-0.064***		-0.033***	-0.113***		-0.030**
	(0.023)		(0.011)	(0.039)		(0.013)
Index p-value			0.001			0.003
Dependent mean	0.202		0.050	0.306		0.052
Observations	64,781		64,781	64,781		64,781
Panel C: Adult criminal activity						
Parent incarcerated (=1)	-0.045	-0.055**	-0.033*	-0.106*	-0.097**	-0.055
	(0.029)	(0.027)	(0.019)	(0.055)	(0.045)	(0.033)
Index p-value			0.044			0.039

Estimating the reduced form

How does judge leniency affect child outcomes (simplified)?

Reduced form:

$$\mathrm{Y}_i = \alpha + \theta \text{Judge leniency}_{ij} + \pi X_{ij} + \eta_i$$

Reduced form

The exclusion restriction is the key to any IV

You should always ask:
What is the exclusion restriction in this analysis saying?

The exclusion restriction is the key to any IV

You should always ask:
What is the exclusion restriction in this analysis saying?

Do we believe this? Why or why not?

Second stage (IV)

	Extensive margin (=1)			Intensive margin (IHS)		
	Charged	Convicted	Incarcerated	Charged	Convicted	Incarcerated
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Criminal activity before age 25						
Parent incarcerated (=1)	-0.066**	-0.055**	-0.049**	-0.156**	-0.097**	-0.076**
	(0.030)	(0.027)	(0.020)	(0.061)	(0.045)	(0.035)
Index p-value			0.011			0.013
Dependent mean	0.325	0.247	0.124	0.568	0.375	0.205
Observations	83,532	83,532	83,532	83,532	83,532	83,532
Panel B: Juvenile criminal activity						
Parent incarcerated (=1)	-0.064***		-0.033***	-0.113***		-0.030**
	(0.023)		(0.011)	(0.039)		(0.013)
Index p-value			0.001			0.003
Dependent mean	0.202		0.050	0.306		0.052
Observations	64,781		64,781	64,781		64,781
Panel C: Adult criminal activity						
Parent incarcerated (=1)	-0.045	-0.055**	-0.033*	-0.106*	-0.097**	-0.055
	(0.029)	(0.027)	(0.019)	(0.055)	(0.045)	(0.033)
Index p-value			0.044			0.039

Lecture 09 recap

TL;DR:

1. Instrumental variables are very powerful
2. With the right assumptions...
3. ...we can handle OVB and ME (and simultaneity)

An example: Returns to schooling

Policy issue:

- How much is a year of school worth?
- This is really important for deciding how much school to invest in
- What role does measurement error play in our estimates?

Approach:

- (We're not actually evaluating a program here)
- We want to estimate the effect of schooling on wages
- **Measurement of years of schooling is poor**
- Instrument of choice: sibling $j \neq i$-reported schooling

Estimating the effects of schooling on wages

The authors will run a (simplified) version of:

$$Wages_i = \tau Schooling_i + \varepsilon_i$$

Where:

$Wages_i$ is a measure of wages

$Schooling_i$ is years of schooling for child i

ε_i is an error term

Estimating the effects of schooling on wages

The authors will run a (simplified) version of:

$$Wages_i = \tau Schooling_i + \varepsilon_i$$

Where:

$Wages_i$ is a measure of wages

$Schooling_i$ is years of schooling for child i

ε_i is an error term

A big concern

- $Schooling_i$ is measured with error
- We are likely to understate the true effect
- **Solution:** $Z_i =$ Schooling as reported by twin$_i$!

First stage estimates (sort of)

				A. *Identical Twins*						
Variable	Y_1	Y_2	S_1^1	S_1^2	S_2^2	S_2^1	E_F^1	E_F^2	E_M^1	E_M^2
Y_1	1.000									
Y_2	0.563	1.000								
S_1^1	0.382	0.168	1.000							
S_1^2	0.375	0.140	0.920	1.000						
S_2^2	0.267	0.272	0.658	0.697	1.000					
S_2^1	0.248	0.247	0.700	0.643	0.877	1.000				
Father's education (E_F^1)	0.155	0.088	0.345	0.266	0.361	0.416	1.000			
Father's education (E_F^2)	0.159	0.091	0.357	0.278	0.320	0.389	0.857	1.000		
Mother's education (E_M^1)	0.102	0.088	0.348	0.343	0.392	0.410	0.614	0.644	1.000	
Mother's education (E_M^2)	0.126	0.087	0.316	0.321	0.322	0.337	0.503	0.579	0.837	1.000

2SLS estimates

Variable	OLS (i)	GLS (ii)	GLS (iii)	IV[a] (iv)	First difference (v)	First difference by IV (vi)
Own education	0.084 (0.014)	0.087 (0.015)	0.088 (0.015)	0.116 (0.030)	0.092 (0.024)	0.167 (0.043)
Sibling's education	—	—	−0.007 (0.015)	−0.037 (0.029)	—	—
Age	0.088 (0.019)	0.090 (0.023)	0.090 (0.023)	0.088 (0.019)	—	—
Age squared (÷100)	−0.087 (0.023)	−0.089 (0.028)	−0.090 (0.029)	−0.087 (0.024)	—	—
Male	0.204 (0.063)	0.204 (0.077)	0.206 (0.077)	0.206 (0.064)	—	—
White	−0.410 (0.127)	−0.417 (0.143)	−0.424 (0.144)	−0.428 (0.128)	—	—
Sample size:	298	298	298	298	149	149
R^2:	0.260	0.219	0.219	—	0.092	—