Applied Microeconometrics I

Lecture 7: Instrumental variables (continued)

Stefano Lombardi

Aalto University

September 26, 2023

Lecture Slides

What did we do last time?

- "Chain reaction" $Z \to D \to Y$
- Instrument validity:
  1. *First stage*: Z affects D
  2. *Independence*: Z is as good as randomly assigned
  3. *Exclusion restriction*: Z has an effect on Y only through D
- What does valid Z identify?
  Setting: RCT with incomplete compliance
  - Treatment effect heterogeneity (we also need *Monotonicity*)
  - Assignment to treatment: $Z_i = \{0, 1\}$
  - Counterfactual treatments: D_{0i}, D_{1i}. For each i, conceptually both exist, but only one of the two is observed, D_i:

$$D_i = \begin{cases} D_{1i} & \text{if } Z_i = 1 \\ D_{0i} & \text{if } Z_i = 0 \end{cases}$$

  - Incomplete compliance: $D_{0i} = \{0, 1\}\,; D_{1i} = \{0, 1\}$

What did we do last time?

- Four group types:
  1. Never takers: $D_{0i} = D_{1i} = 0$
  2. Always takers: $D_{0i} = D_{1i} = 1$
  3. Defiers: $D_{0i} = 1, D_{1i} = 0$
  4. Compliers: $D_{0i} = 0, D_{1i} = 1$

  Note: you can **always** think in terms of groups 1.–4.: abstractly, IV *is* an RCT's with imperfect compliance.

- Define $Y(d, z) \equiv Y(D_{zi} = d, Z_i = z)$. Four potential outcome-treatment combinations. For each group type above we calculated $Y_i(D_{1i}, 1) - Y_i(D_{0i}, 0)$.

- Instrumental variable estimates the effect on compliers.

What did we do last time?

- More formally, the assumptions we need are:
  1. Independence: $\{Y_i(D_{1i}, 1), Y_i(D_{0i}, 0), D_{1i}, D_{i0}\} \perp Z_i$
  2. First stage:* $E[D_i|Z_i = 1] - E[D_i|Z_i = 0] \neq 0$
  3. Exclusion: $Y_i(d, 0) = Y_i(d, 1) \equiv Y_{di}$, for $d = 0, 1$
  4. Monotonicity: $D_{1i} \geq D_{0i}, \forall i$ (or vice versa)

  *under 1., the FS assumption becomes $E[D_{1i} - D_{0i}] \neq 0$

- IV estimates LATE (effect the compliers):

  $$E[Y_i(1) - Y_i(0)|D_{1i} > D_{i0}] = \frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[D_i|Z_i=1]-E[D_i|Z_i=0]}$$

- You can always use this framework when doing IV, not just when having an actual RCT with imperfect compliance

Today: Instrumental variables (continued)

- Angrist (2006): When is LATE same as ATET?
- Instrumental variables in the regression framework
  - Instrument validity and weak instruments
  - Two-stage least squares (2SLS)
  - IV and control variables
  - The Wald estimator
  - Multiple instruments
- Examples of instruments – treatments:
  - Angrist and Krueger (1991): quarter of birth – schooling
  - Angrist & Evans (1998): siblings sex mix – third child
  - Lundeburg et al (2016): first IVF treatment – first child

When is LATE same as the effect on the treated?

- There is an important special case when instrumental variables actually give the treatment on the treated
- If there are no always-takers so that $E[D_i|Z_i = 0] = 0$
- Then:
$$E[Y_i(1) - Y_i(0)|D_{1i} > D_{oi}] = \frac{E[Y_i|Z_i=1] - E[Y_i|Z_i=0]}{E[D_i|Z_i=1]}$$
- In these cases IV estimates the effect on the treated population (since all treated are compliers)
- Example: Angrist (2006)

Example: Instruments and criminology

Angrist (2006)

- Angrist (2006) revisits a famous RCT on the treatment of domestic disturbance by the police force in Minneapolis
- RCT tried to address the question whether the officer should arrest the offender or "coddle" (=advise/separate)
- Upon arriving at the scene the officers were supposed to randomize by drawing a card with a coded color for each treatment
- The goal of the RCT was to find out how coddling affects recidivism: *coddling* is the treatment.

Assigned and delivered treatments

Angrist (2006)

Table 1. Assigned and delivered treatments in spousal assault cases.

	Delivered treatment			
		Coddled		
Assigned treatment	*Arrest*	*Advise*	*Separate*	*Total*
Arrest	98.9 (91)	0.0 (0)	1.1 (1)	29.3 (92)
Advise	17.6 (19)	77.8 (84)	4.6 (5)	34.4 (108)
Separate	22.8 (26)	4.4 (5)	72.8 (83)	36.3 (114)
Total	43.4 (136)	28.3 (89)	28.3 (89)	100.0 (314)

The table shows statistics from Sherman and Berk (1984), Table 1.

Assigned and delivered treatments

Angrist (2006)

- We see that when told to coddle 80% ($\frac{(84+5)+(5+83)}{108+114}$) actually coddled
- However, when not told to arrest only 1% ($\frac{1}{92}$) coddled
- Hence, there practically are no always-takers in this experiment

First stage and reduced form

Angrist (2006)

Table 2. First stage and reduced forms for Model 1.

Endogenous variable is coddled				
	First stage		*Reduced form (ITT)*	
	(1)	*(2)**	*(3)*	*(4)**
Coddled-assigned	0.786 (0.043)	0.773 (0.043)	0.114 (0.047)	0.108 (0.041)
Weapon		−0.064 (0.045)		−0.004 (0.042)
Chem. influence		−0.088 (0.040)		0.052 (0.038)
Dep. var. mean	0.567 (Coddled–delivered)		0.178 (Re-arrested)	

The table reports OLS estimates of the first-stage and reduced form for Model 1 in the text.
*Other covariates include year and quarter dummies, and dummies for non-white and mixed race.

First stage and reduced form

Angrist (2006)

- We see that being told to coddle lead to 78.6 percentage point increase in coddling
- We are interested in the effect of coddling on re-arrest rates
- The reduced form effect is 11.4 percentage points

OLS and IV

Angrist (2006)

Table 3. OLS and 2SLS estimates for Model 1.

Endogenous variable is coddled				
	OLS		*IV/2SLS*	
	(1)	*(2)**	*(3)*	*(4)**
Coddled–delivered	0.087 (0.044)	0.070 (0.038)	0.145 (0.060)	0.140 (0.053)
Weapon		0.010 (0.043)		0.005 (0.043)
Chem. influence		0.057 (0.039)		0.064 (0.039)

The Table reports OLS and 2SLS estimates of the structural equation in Model 1.
*Other covariates include year and quarter dummies, and dummies for non-white and mixed race.

Discussion

Angrist (2006)

- We see that if we would only compare coddles and arrests the effect would be 8.7 percentage points
- The reduced form effect is 11.4 percentage points
- LATE estimate is 14.5 percentage points (which is what we get if divide the reduced form with the first stage)
- Why do these estimates differ even though this was an RCT?
  - police officers didn't comply with the coddle assignment if they thought that an arrest was necessary
- Note that non-compliance is one sided:
  - there are **no always-takers**: those encouraged to arrest, always do it (except in one case)
  - therefore all the treated are compliers

Instrumental variables in the regression framework

Schooling example

- Consider instrumental variables in a **regression framework**
- Goal: estimate the causal effect of schooling (treatment, S_i) on wages (outcome, Y_i)
- Lets assume we do not observe everything that affects both selection into schooling and earnings (ability, A_i)
- The relationship between earnings and schooling is:

$$Y_i = \alpha + \rho S_i + u_i$$
$$u_i \equiv \gamma A_i + \varepsilon_i$$

- A_i is assumed to be the only reason why u_i and S_i are correlated, i.e.:

$$E[S_i \varepsilon_i] = 0$$

Instrumental variables in the regression framework

Schooling example

- If we observed A_i, we could simply control for it and consistently estimate ρ with OLS of the (long) regression:

$$Y_i = \alpha + \rho S_i + \gamma A_i + \varepsilon_i$$

- But A_i is not measured (or badly proxied)
- Instrumental variable (IV) allows us to estimate ρ when A_i is unobserved and correlated with both Y_i and S_i.

Instrumental variables in the regression framework

Schooling example

- A valid instrument consistently estimates ρ in the short regression:

$$Y_i = \alpha + \rho S_i + u_i$$

- We can write ρ in terms of the population moments:

$$\text{Cov}(Z_i, Y_i) = \rho \text{Cov}(Z_i, S_i) + \text{Cov}(Z_i, u_i)$$

- By instrument validity, $\text{Cov}(Z_i, u_i) = 0$:

$$\rho = \frac{Cov(Z_i, Y_i)}{Cov(Z_i, S_i)} = \frac{\frac{Cov(Z_i, Y_i)}{Var(Z_i)}}{\frac{Cov(Z_i, S_i)}{Var(Z_i)}}$$

- ρ is the ratio between regression of Y_i on Z_i (the reduced form) and regression of S_i on Z_i (the first stage).

What is a valid instrumental variable?

- Assumptions:

  1. **Relevance (first stage)**: $\text{Cov}(Z_i, S_i) \neq 0$
     The instrument is correlated with the treatment of interest, S_i

  2. **Validity assumption:** $\text{Cov}(Z_i, u_i) = 0$
     The instrument is uncorrelated with any other determinants of Y_i. This requirement can be decomposed in:

     2.1 **Independence (exogeneity)**: Z is not correlated with any unobservable factors that affect Y (as good as randomly assigned)

     2.2 **Exclusion restriction**: Z only affects Y through its effect on S

- A *valid instrumental variable* requires *all* assumptions listed above (relevance + validity assumption).

- Note: we cannot see it in the regression framework, but under heterogeneous treatment effects, we also need *monotonicity*.

Weak instruments

- Without a first stage, ρ is not even defined. What if $\text{Cov}(Z_i, S_i) \neq 0$ but Z_i is only weakly correlated with S_i ?
- Again, write again ρ as:

$$\text{Cov}(Z_i, Y_i) = \rho\text{Cov}(Z_i, S_i) + \text{Cov}(Z_i, u_i)$$

- Suppose $\text{Cov}(Z_i, u_i) \neq 0$ since Z_i is slightly correlated with u_i (in real-world scenarios, this often the case).
- The IV estimator for ρ is now biased:

$$\hat{\rho} = \rho + \frac{Cov(Z,u)}{Cov(S,Z)}$$

- Even when Z_i is only slightly correlated with error term, if the FS is weak the bias will be extremely large.
  - even worse in small samples
  - the bias might be even larger than the OLS one!

Two-stage Least Squares (2SLS)

- We can do IV through a two stage procedure.
- Suppose we have **one instrument** (valid) for one endogenous variable S_i.
- Econometric model of interest:

$$Y_i = X_i'\alpha + \rho S_i + u_i \qquad (1)$$

- First-stage equation:

$$S_i = X_i'\pi_0 + \pi_1 Z_i + v_i \qquad (2)$$

- Note that:
  - X_i is a vector (which includes the intercept).
  - X_i *must* be the same in (1) and (2) (more on this later).
  - S_i appears in (1), so we can substitute it into (1).

Two-stage Least Squares (2SLS)

- **OLS of the FS**. By *exogeneity assumption*, the OLS fitted values of the FS, $\hat{S}_i \equiv X_i'\hat{\pi}_0 + \hat{\pi}_1 Z_i$, are consistent.
- The FS OLS residuals are defined as $\hat{v}_i = S_i - \hat{S}_i$. Substitute the estimated FS in (1):

$$\begin{aligned} Y_i &= X_i'\alpha + \rho[X_i'\hat{\pi}_0 + \hat{\pi}_1 Z_i + \hat{v}_i] + u_i \\ &= X_i'\alpha + \rho\hat{S}_i + \xi_i \end{aligned} \tag{3}$$

  since $S_i = \hat{S}_i + \hat{v}_i$ and $\xi_i = \rho\hat{v}_i + u_i$ (second stage error term)
- **OLS of the second stage**. X_i, Z_i are uncorrelated with $\hat{v}_i$ by construction (OLS residuals of S_i on (X_i, Z_i)) and with u_i (by instrument validity). Hence, ρ is consistently estimated.
- Idea: Conditional on X_i, 2SLS retains only the (exogenous) variation in S_i generated by the instrument Z_i

Instrumental variables and control variables

- Recall that X_i was the same in model (1) and first stage (2).
- Suppose that instead we do not include X_i in the FS:

$$\begin{aligned} Y_i &= X_i'\alpha + \rho S_i + u_i & (1) \\ S_i &= \pi_0 + \pi_1 Z_i + v_i & (2) \end{aligned}$$

- With a valid Z_i, the FS is consistent. The second stage is:

$$Y_i = X_i'\alpha + \rho \hat{S}_i + \underbrace{\rho \hat{v}_i + u_i}_{\xi_i} \qquad (3)$$

- However, now by construction the FS residuals $\hat{v}_i$ in ξ_i are *uncorrelated only with* Z_i (since X_i was not in (2))
- Since the second stage error term is in general correlated with X_i, the OLS estimator of the second stage is inconsistent.

Instrumental variables and control variables

- Another way to think about it: consider the assumption underlying the instrument validity, $Cov(Z_i, u_i) = 0$.
- If X is in the second stage, the assumption holds conditional on X (so if we omit X from the FS, it won't hold anymore).
- Note: randomization might occur only after conditioning on X (*conditional* independence):
  - in this case, we must add the same X to FS and second stage.
  - *Judge leniency*: randomization occurs *within* court-year.
- Once Z is as good as randomly assigned and the exclusion restriction holds, it's unnecessary to condition on X.

2SLS and standard errors

- We showed that with 2SLS we can:
  - do OLS of the FS to get the fitted values $\hat{S}_i$
  - plug $\hat{S}_i$ in (1), do OLS of the second stage to get consistent $\hat{\rho}$.
- In practice, we never run 2SLS "manually" in two steps.
- With the manual two stage procedure, you do not get the correct standard errors
  - The residual that is used to calculate standard errors in second stage includes an extra error since $Y_i - [X_i'\alpha - \rho\hat{S}_i] = \rho\hat{v}_i + u_i$
  - Intuition: even though we treat it as a regular regressor, $\hat{S}$ is a generated regressor and inflates the variance
  - Stata `ivreg`/`ivreg2` fix this: the original endogenous regressor S_i is used to construct residuals: $Y_i - [X_i'\alpha - \rho S_i] = u_i$

The Wald estimator

- Consider the case when we have:
  - Model with one endogenous regressor and no covariates
  - Single binary instrument $z_i \in \{0, 1\}$
- If z_i equals 1 with probability p, then the 2SLS estimator is equivalent to the **Wald estimator**:

$$\rho = \frac{Cov(y_i, z_i)}{Cov(s_i, z_i)} = \frac{E[y_i|z_i=1] - E[y_i|z_i=0]}{E[s_i|z_i=1] - E[s_i|z_i=0]}$$

- This is equal to the IV estimator RF/FS (sometimes called "indirect least squares").

Multiple instruments

- So far we worked with only one valid Z_i. What happens when we have multiple instruments for S_i?
- Instead of using the instruments separately, we might want to combine multiple estimates into one, more precise estimate
- The 2SLS intuition is the same. The FS is now a linear combination of all instruments:

$$S_i = X_i'\pi_0 + \pi_1 Z_{1i} + \pi_2 Z_{2i} + \cdots + \pi_p Z_{pi} + v_i$$

- 2SLS estimates a *weighted average* of the causal effects for instrument-specific compliers:
  - With two Wald estimators, 2SLS is a weighted average of the two Wald estimates obtained by using one instrument at a time
  - The two instruments need not to have the same compliers.

Quarter of birth and an instrument for schooling

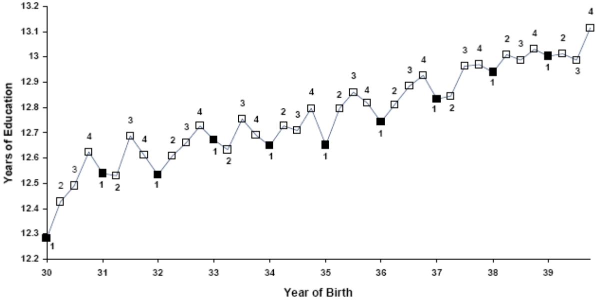
Angrist and Krueger (1991)

- 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?

Angrist and Krueger (1991)

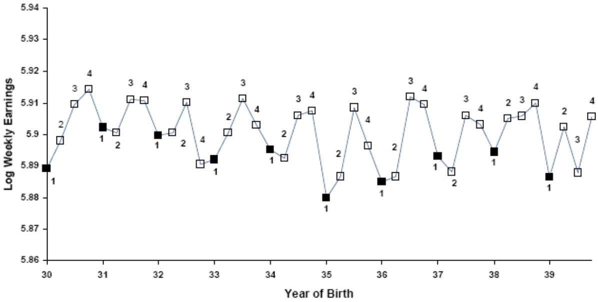
A. Average Education by Quarter of Birth (first stage)



The reduced form for earnings

Angrist and Krueger (1991)

B. Average Weekly Wage by Quarter of Birth (reduced form)



Wald estimates of the returns to schooling

Angrist and Krueger (1991)

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.

Validity of the instrument

Angrist and Krueger (1991)

- Validity:
  1. Power of the instrument?
  2. Exogeneity?
  3. Exclusion restriction?

- Overall, are the OLS estimates mostly larger than the corresponding IV estimates?

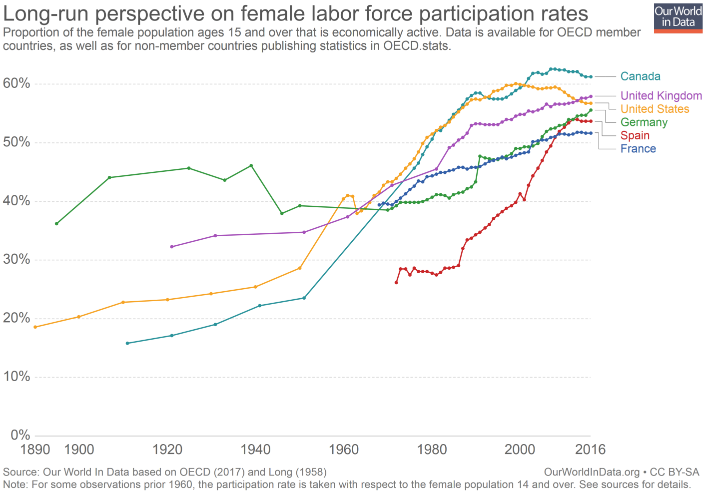
- What does this tell us about the omitted variable bias?

- How are the IV estimates “local” and who are the compliers in this case?

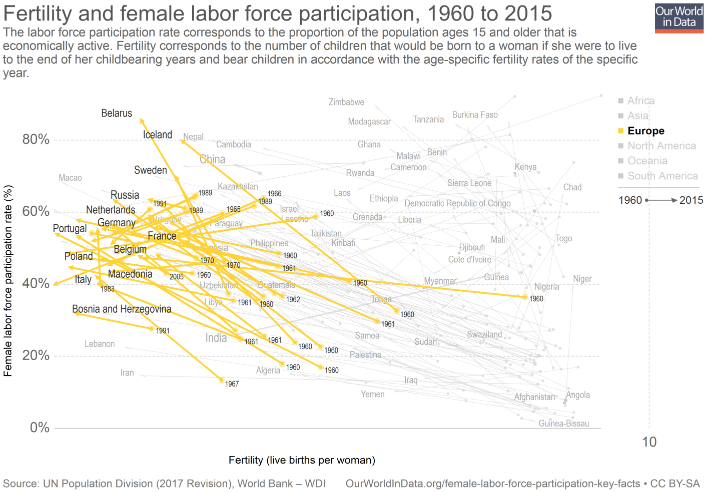
Example: Fertility and female labor supply

- Powerful global trend: Increasing female labor force participation
- Similarly powerful decline in fertility
- Are these trends linked? Does childbearing keep women from developing their careers?
- Huge number of studies that show a negative correlation

Female labor force participation 1890-2016



Fertility and female labor force participation 1960-2015



Fertility and female labor supply

- Why wouldn't this correlation be causal?
- Strong theoretical reasons to believe that fertility and labor supply decisions are jointly determined
- Fertility is not allocated at random
- Unclear whether observed differences in labor market outcomes reflect the causal effects of having children

Fertility and female labor supply

- We must look for variation in the number of children that is as good as randomly assigned
- The population of interest is typically couples who want to have any/more children
- First attempts: twin births instrument
    - Why could this work?
    - What could go wrong here?
- We cover:
    - Angrist & Evans (1998): Sibling sex mix in families with two or more children
    - Lundborg et al (2016): The success of IVF treatments

Siblings sex mix instrument for family size

Angrist & Evans (1998)

- Angrist & Evans (1998): Sibling sex mix of the first two children as an instrument for the decision to have a third child
- Based on two assumptions:
    - Parental preference for mixed sibling-sex composition
    - Sex mix is virtually randomly assigned
- Why does this mean that sex mix can work as an instrument?

First stage

Angrist & Evans (1998)

Sex of first two children in families with two or more children	All women				Married women			
	1980 PUMS (394,835 observations)		1990 PUMS (380,007 observations)		1980 PUMS (254,654 observations)		1990 PUMS (301,588 observations)	
	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child
one boy, one girl	0.494	0.372 (0.001)	0.495	0.344 (0.001)	0.494	0.346 (0.001)	0.497	0.331 (0.001)
two girls	0.242	0.441 (0.002)	0.241	0.412 (0.002)	0.239	0.425 (0.002)	0.239	0.408 (0.002)
two boys	0.264	0.423 (0.002)	0.264	0.401 (0.002)	0.266	0.404 (0.002)	0.264	0.396 (0.002)
(1) one boy, one girl	0.494	0.372 (0.001)	0.495	0.344 (0.001)	0.494	0.346 (0.001)	0.497	0.331 (0.001)
(2) both same sex	0.506	0.432 (0.001)	0.505	0.407 (0.001)	0.506	0.414 (0.001)	0.503	0.401 (0.001)
difference (2) − (1)	—	0.060 (0.002)	—	0.063 (0.002)	—	0.068 (0.002)	—	0.070 (0.002)

Wald estimates

Angrist & Evans (1998)

TABLE 5—WALD ESTIMATES OF LABOR-SUPPLY MODELS

Variable	1980 PUMS: Mean difference by *Same sex*	1980 PUMS: Wald estimate using as covariate: *More than 2 children*	1980 PUMS: Wald estimate using as covariate: *Number of children*	1990 PUMS: Mean difference by *Same sex*	1990 PUMS: Wald estimate using as covariate: *More than 2 children*	1990 PUMS: Wald estimate using as covariate: *Number of children*	1980 PUMS: Mean difference by *Twins-2*	1980 PUMS: Wald estimate using as covariate: *More than 2 children*	1980 PUMS: Wald estimate using as covariate: *Number of children*
More than 2 children	0.0600 (0.0016)	—	—	0.0628 (0.0016)	—	—	0.6031 (0.0084)	—	—
Number of children	0.0765 (0.0026)	—	—	0.0836 (0.0025)	—	—	0.8094 (0.0139)	—	—
Worked for pay	−0.0080 (0.0016)	−0.133 (0.026)	−0.104 (0.021)	−0.0053 (0.0015)	−0.084 (0.024)	−0.063 (0.018)	−0.0459 (0.0086)	−0.076 (0.014)	−0.057 (0.011)
Weeks worked	−0.3826 (0.0709)	−6.38 (1.17)	−5.00 (0.92)	−0.3233 (0.0743)	−5.15 (1.17)	−3.87 (0.88)	−1.982 (0.386)	−3.28 (0.63)	−2.45 (0.47)
Hours/week	−0.3110 (0.0602)	−5.18 (1.00)	−4.07 (0.78)	−0.2363 (0.0620)	−3.76 (0.98)	−2.83 (0.73)	−1.979 (0.327)	−3.28 (0.54)	−2.44 (0.40)
Labor income	−132.5 (34.4)	−2208.8 (569.2)	−1732.4 (446.3)	−119.4 (42.4)	−1901.4 (670.3)	−1428.0 (502.6)	−570.8 (186.9)	−946.4 (308.6)	−705.2 (229.8)
ln(*Family income*)	−0.0018 (0.0041)	−0.029 (0.068)	−0.023 (0.054)	−0.0085 (0.0047)	−0.136 (0.074)	−0.102 (0.056)	−0.0341 (0.0223)	−0.057 (0.037)	−0.042 (0.027)

Comparison of OLS and 2SLS

Angrist & Evans (1998)

TABLE 7—OLS AND 2SLS ESTIMATES OF LABOR-SUPPLY MODELS USING 1980 CENSUS DATA

	All women			Married women			Husbands of married women		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimation method	OLS	2SLS	2SLS	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Instrument for *More than 2 children*	—	*Same sex*	*Two boys, Two girls*	—	*Same sex*	*Two boys, Two girls*	—	*Same sex*	*Two boys, Two girls*
Dependent variable:									
Worked for pay	−0.176 (0.002)	−0.120 (0.025)	−0.113 (0.025) [0.013]	−0.167 (0.002)	−0.120 (0.028)	−0.113 (0.028) [0.013]	−0.008 (0.001)	0.004 (0.009)	0.001 (0.008) [0.013]
Weeks worked	−8.97 (0.07)	−5.66 (1.11)	−5.37 (1.10) [0.017]	−8.05 (0.09)	−5.40 (1.20)	−5.16 (1.20) [0.071]	−0.82 (0.04)	0.59 (0.60)	0.45 (0.59) [0.030]
Hours/week	−6.66 (0.06)	−4.59 (0.95)	−4.37 (0.94) [0.030]	−6.02 (0.08)	−4.83 (1.02)	−4.61 (1.01) [0.049]	0.25 (0.05)	0.56 (0.70)	0.50 (0.69) [0.71]
Labor income	−3768.2 (35.4)	−1960.5 (541.5)	−1870.4 (538.5) [0.126]	−3165.7 (42.0)	−1344.8 (569.2)	−1321.2 (565.9) [0.703]	−1505.5 (103.5)	−1248.1 (1397.8)	−1382.3 (1388.9) (0.549)
ln(*Family income*)	−0.126 (0.004)	−0.038 (0.064)	−0.045 (0.064) [0.319]	−0.132 (0.004)	−0.051 (0.056)	−0.053 (0.056) [0.743]	—	—	—
ln(*Non-wife income*)	—	—	—	−0.053 (0.005)	0.023 (0.066)	0.016 (0.066) [0.297]	—	—	—

Validity of the instrument

Angrist & Evans (1998)

- Validity:
  1. Power of the instrument?
  2. Exogeneity?
  3. Exclusion restriction?
- Overall, are the OLS estimates mostly smaller than the corresponding IV estimates?
- What does this tell us about the omitted variable bias?
- Based on this, what would you say about the effect of childbearing on female labor market outcomes?
- How are the IV estimates local in this case?

IVF treatment as an instrument for having a child

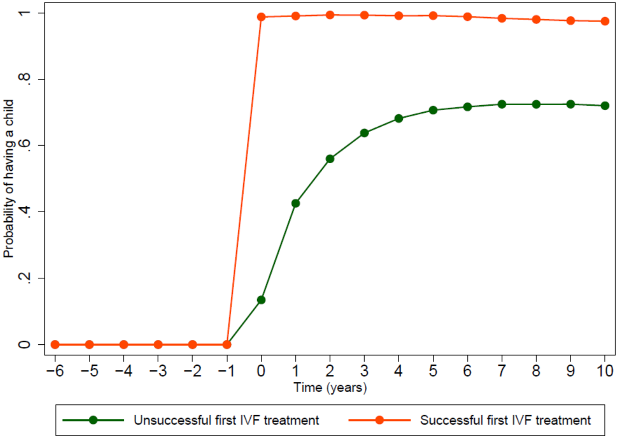
Lundborg et al (2016)

- Difference: effect of childbearing at the *intensive* margin (1 to 2, 2 to 3, ..., children) and *extensive* margin (0 to 1 child)
- Angrist & Evans only studied the effect of going from 2 to 3 children. Is this relevant?
- Lundborg et al. (2016): Success of IVF treatments as an instrument for having a first child
    - Focus on women who are going through IVF treatment
    - Is IVF treatment affecting the probability of having a child?
    - Is the success in the *first* IVF treatment as good as random?
    - Is the success of first IVF treatment affecting labor market participation only through increased fertility?

Fertility and IVF success

Lundborg et al (2016)

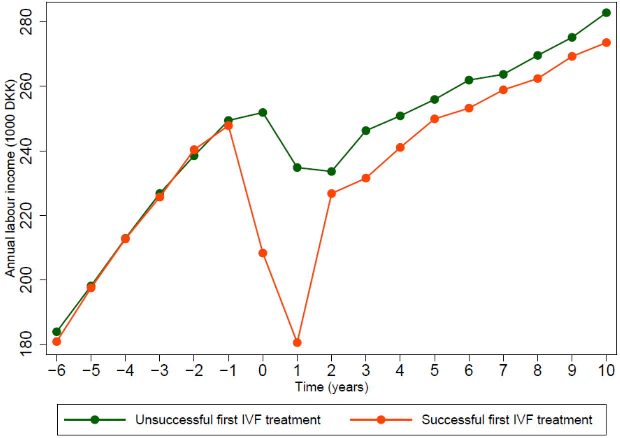
Figure 1: Fertility at the extensive margin before and after the first IVF treatment.



IVF success and annual earnings

Lundborg et al (2016)

Figure 2: Annual earnings before and after the (potential) birth of a child.



First stage, reduced form, and Wald estimates

Lundborg et al (2016)

Table 3: Fertility effects on female labor earnings: Results from first-stage, reduced form, and instrumental variable regressions.

Independent variable	(1) t=0	(2) t=1	(3) t=2	(4) t=3	(5) t=4	(6) t=5	(7) t=6	(8) t=7	(9) t=8	(10) t=9	(11) t=10
Panel A: First stage regressions using any children (0/1) as dependent variable											
IVF success (0/1)	0.84***	0.54***	0.41***	0.33***	0.28***	0.26***	0.24***	0.23***	0.23***	0.23***	0.23***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
N	18538	18494	18435	18381	17404	15599	13779	11983	10173	8342	6620
F-stat.	57424	13303	7844	5493	4055	3199	2482	1874	1520	1176	943
Panel B: Reduced form regressions using annual earnings as dependent variable											
IVF success (0/1)	-43299***	-53082***	-6009***	-14340***	-9050***	-5269**	-8679***	-5992**	-7536**	-5566	-10975***
	(1541)	(1700)	(1830)	(1934)	(2055)	(2244)	(2463)	(2738)	(3038)	(3491)	(3981)
N	18538	18494	18435	18381	17404	15599	13779	11983	10173	8342	6620
Panel C: IV regressions using annual earnings as dependent variable											
Any children (0/1)	-51251***	-97914***	-14781***	-43798***	-32147***	-20493**	-35866***	-26064**	-32814**	-24338	-47079***
	(1826)	(3099)	(4474)	(5872)	(7275)	(8710)	(10168)	(11901)	(13237)	(15262)	(17114)
N	18538	18494	18435	18381	17404	15599	13779	11983	10173	8342	6620

Validity of the instrument

Lundborg et al (2016)

- Validity:
  1. Power of the instrument?
  2. Exogeneity?
  3. Exclusion restriction?
- Compare the estimates with Angrist & Evans
- What does this comparison say about the effect of having the first vs. the third child?
- How are the IV estimates local in this case?

Let us wrap it up

- IV estimates are a powerful tool to identify causal links
- But IV power relies on the quality of the instruments
- Always discuss instrument plausibility. Three dimensions:
  1. Strength of the first stage:
     - Always report the first stage (F-test above 10)
     - Weak instruments have very unpleasant consequences
  2. Exogeneity:
     - Is it sensible to assume that the instrument randomly assigned?
     - Check if the instrument is correlated with predetermined variables
  3. Exclusion restriction:
     - Cannot be tested, but discuss the possible links between z and u
- Specify the group which is affected by the instrument (LATE)

Wald estimator derivation (optional)

- Since z only take values 0 and 1, $E[yz] = E[y|z=1]p$. Also $E[y]E[z] = \{E[y|z=1]p + E[y|z=0](1-p)\}$
- Therefore the numerator of the Wald estimator is:

$$\begin{aligned} Cov(yz) &= E[y-E[y]][z-E[z]] \\ &= E[yz] - E[y]E[z] \\ &= E[y|z=1]p - \{E[y|z=1]p + E[y|z=0](1-p)\}\, p \\ &= p\,\{E[y|z=1] - E[y|z=1]p - E[y|z=0](1-p)\} \\ &= p(1-p)\,\{E[y|z=1] - E[y|z=0]\} \end{aligned}$$

- And similarly for the denominator:

$$Cov(sz) = p(1-p)\,\{E[s|z=1] - E[s|z=0]\}$$