Non-Experimental Data: IV

Rajeev Dehejia

Today's class

- Introduction to Instrumental Variables
  – What are they
  – How do we estimate IV
  – Tests for specification/fit

Use cases/jobs

- Common use cases:
  - Research questions around the effects of voluntary programs (e.g. jobs programs)
  - Long-term impact evaluations (i.e. relies heavily on historical data)
- Fields that tend to use IV:
  - Epidemiology
  - Labor economics

IV: Current Events & Papers

- Using Machine Learning to improve IV estimates: "Machine Labor" December 2019
  Joshua Angrist, Brigham Frandsen

- Social and private returns to education: " Signaling and Employer Learning with Instruments" May 2019
  Gaurab Aryal, Manudeep Bhuller, Fabian Lange

- IV analyses of the effects of prescription opioid use on infant health and maternal behaviors: "Effect of Prescription Opioids and Prescription Opioid Control Policies on Infant Health" February 2020, Engy Ziedan, Robert Kaestner

Recap of the problem

$$Y = \alpha X + \rho T + \varepsilon$$

- There is some part of the error (ε) that we don't observe (maybe behavioral parameters, maybe simultaneously determined component, etc.)
- This component might not be:
  - Fixed within a group *Can't control with group fixed effects*
  - Fixed over time/space *Can't control with time effects*
  - Related to observables *Endogeneity*
- BUT…this component IS correlated with the treatment/variable of interest *But only predicts the outcome insofar as it is correlated with the treatment*

Our treatment effects model

Lifetime earnings — *Attended college* — *Underlying ability, parental pressure, motivation…*

$$Y = \alpha + \rho T + \varepsilon$$

- Think of the example of schooling
- How much more will you earn if you go to college?
- Can't observe true underlying ability which is correlated with college attendance decision *and* future earnings

What's correlated and what's not

- The model we want to estimate:

$$Y_i = \alpha + \rho T_i + \gamma A_i + v_i$$

- We have that:
  - E[Tv] = 0 (by assumption)
  - E[Av] = 0 (by construction)

- The idea: if *A* could be observed, we'd just include it in the regression and be done

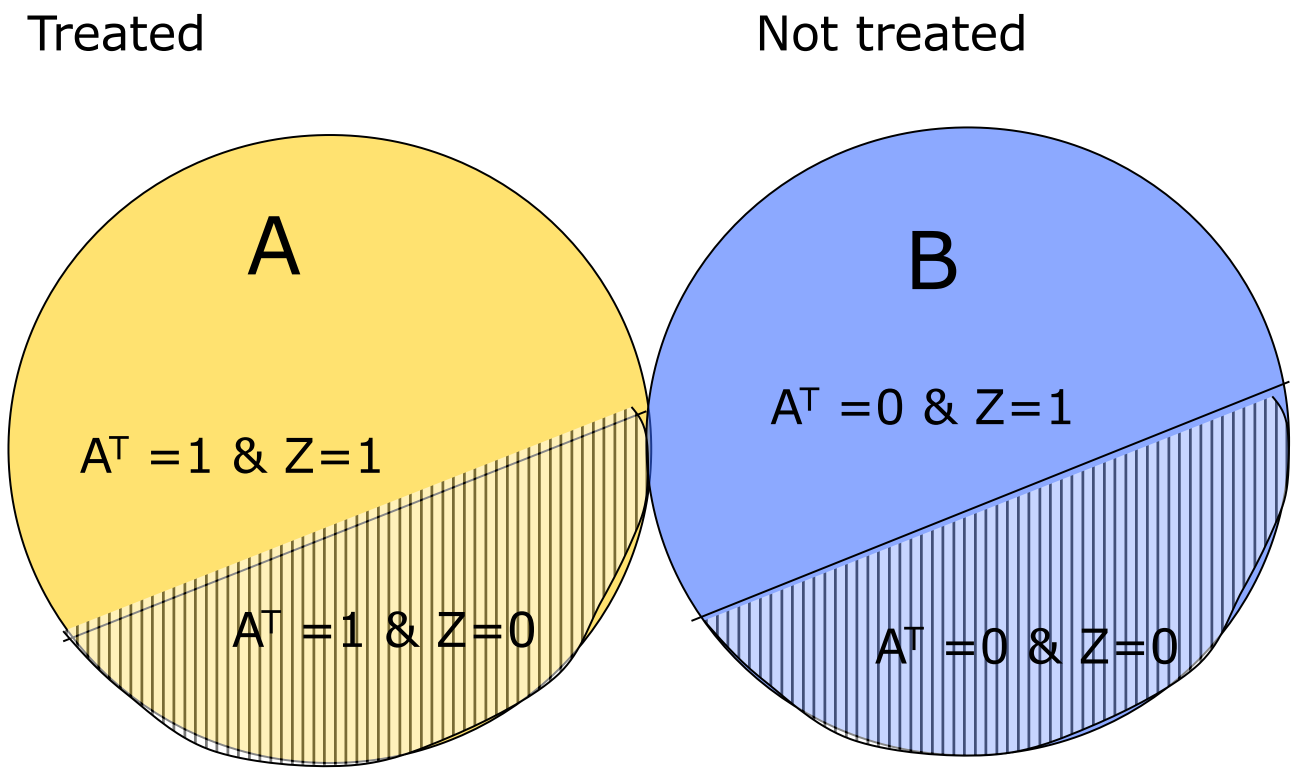
Motivation: partial compliance

- Suppose you run an experiment in which not everyone randomly assigned to the treatment (z=1) actually takes it (T=1), and where not everyone assigned to control (z=0) is excluded from treatment (T=0)?
- Can't run

$$Y = a + bT + e$$

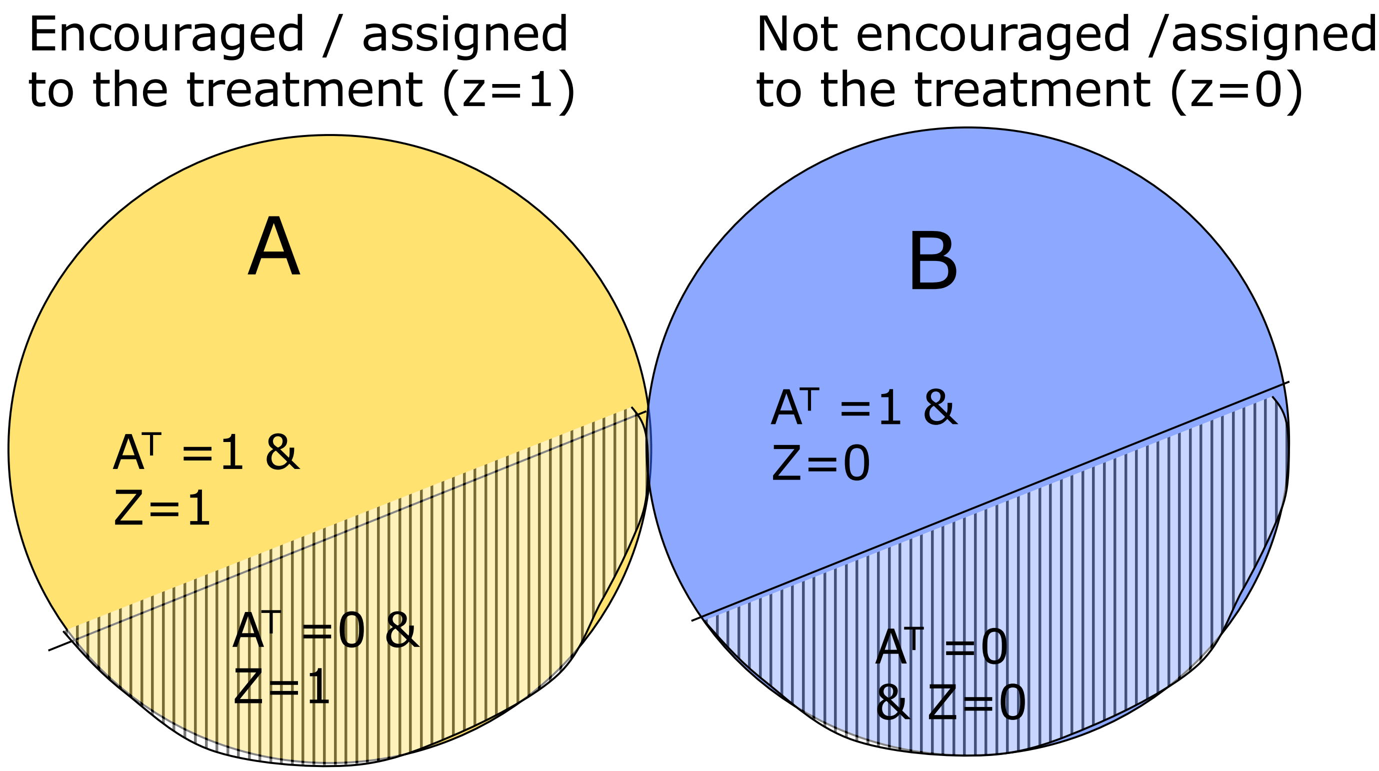
as OLS since T is self-selected. But can run it as IV with Z (assignment) as the instrument.

IV example: non-compliance



A **biased comparison is *all* of A (A^T=1, treated) to *all* of B (A^T=0, untreated)**: this mixes together the compliers (A^T=1 & Z=1; A^T=0 & Z=0) and the non-compliers (A^T=0 & Z=1; A^T=1 & Z=0

IV example: non-compliance



Using the instrument, Z , we can: compare outcome (Y) for z=1 vs z=0 for causal ITT effect (z is randomly assigned so A^H and A^L balance); and can also compare Prob(T=1) for z=1 vs z=0 (for "first stage" or effect of z on treatment take-up). Ratio of the two gives the effect of *T*.

The Wald estimator

- Intent-to-treat (ITT) effect.
  - Is causal (effect of Z on Y).
  - But probably understates treatment effect since contrast between first and second term isn't all vs none treatment but more vs less treated.

$$E\left(y \middle| z=1\right) - E\left(y \middle| z=0\right)$$

- The first stage captures the causal effect of instrument (assignment, or "z") on treatment take up (T)…
  - Or the proportion by which first term in ITT has more treated than second term.

$$E\left(T \middle| z=1\right) - E\left(T \middle| z=0\right)$$

- The ratio scales up ITT by first stage, so is causal effect of *T* on *Y*.

$$\delta = \frac{E\left(y \middle| z=1\right) - E\left(y \middle| z=0\right)}{E\left(T \middle| z=1\right) - E\left(T \middle| z=0\right)}$$

The Wald estimator

1. **First stage:** Difference in treatment take-up rate (T=1) for those that were (z=1) and were not (z=0) selected for treatment

   $$E(T|z=1)-E(T|z=0)$$

   Also known as ϕ, "phi"

2. **Second stage:** Difference in outcome measure (Y) for those that were (z=1) and were not (z=0) selected for treatment

   $$E(y \mid z=1) - E(y \mid z=0)$$

   Also known as "Reduced form"; ρ, "rho"; ITT

3. **Wald Estimator:** Ratio of stage 2: stage 1

   $$\delta = \frac{E(y|z=1)-E(y|z=0)}{E(T|z=1)-E(T|z=0)}$$

   Also known as the Local Average Treatment Effect (LATE); TOT

Introducing instruments

$$Y_i = a + \rho T_i + \gamma A_i + v_i = a + \rho T_i + e_i$$

- The problem: How to estimate ρ *(treatment effect)* when
  - *A* is not observed (or unobservable)
  - *A* is related to *Y*
  - Cov(*AT*) ≠0 (*A* is related to the uptake of treatment)
- The solution: find something (an instrument, z) that is
  - Correlated with *T* ["Relevance"]
  - Uncorrelated with any other determinant of the outcome variable *Y* [Exogenous + "Exclusion Restriction"]

How does IV work

- Our two "must have" instrument characteristics can be written as
  - E[z T] ≠0 "Relevance"
    - There is a relationship between the instrument and treatment uptake
  - E[z e] = 0 "Exogeneity" + "Exclusion restriction"
    - There is no relationship between the instrument and the error term
    - The instrument is not correlated with the outcome Y
- Then from our equations we can write our population estimate of ρ as:

$$\frac{\mathrm{cov}(Y,z)}{\mathrm{cov}(T,z)} = \frac{\mathrm{cov}(a+\rho T+e,z)}{\mathrm{cov}(T,z)} = \frac{\rho\,\mathrm{cov}(Tz)}{\mathrm{cov}(T,z)} = \rho$$

Simplest case for IV

- Homogeneous treatment effects (same ρ for all i)
- Dummy Variable for instrument
  - $z = 1$ with probability q

- For now—don't worry about covariates
  - Simple extension: just include these in both stages
  - Simplify our notation later…

How does IV work – simplest case

$$\rho = \frac{\mathrm{cov}(y,z)}{\mathrm{cov}(T,z)}$$

$$= \frac{E(yz) - E(y)E(z)}{E(Tz) - E(T)E(z)}$$

$$= \frac{E(y\cdot 1 \mid z=1)\Pr(Z=1) + E(y\cdot 0 \mid z=0)\Pr(Z=0) - \left[E(y \mid z=1)\Pr(Z=1) + E(y \mid Z=0)\Pr(Z=0)\right]\Pr(Z=1)}{E(T\cdot 1 \mid z=1)\Pr(Z=1) + E(T\cdot 0 \mid z=0)\Pr(Z=0) - \left[E(T \mid z=1)\Pr(Z=1) + E(T \mid Z=0)\Pr(Z=0)\right]\Pr(Z=1)}$$

$$= \frac{E(y\cdot 1 \mid z=1)\Pr(Z=1) - E(y \mid z=1)\Pr(Z=1)\Pr(Z=1) - E(y \mid Z=0)\Pr(Z=0)\Pr(Z=1)}{E(T\cdot 1 \mid z=1)\Pr(Z=1) - E(T \mid z=1)\Pr(Z=1)\Pr(Z=1) - E(T \mid Z=0)\Pr(Z=0)\Pr(Z=1)}$$

$$= \frac{E(y\cdot 1 \mid z=1)\Pr(Z=1)(1-\Pr(Z=1)) - E(y \mid Z=0)(1-\Pr(Z=1))\Pr(Z=1)}{E(T\cdot 1 \mid z=1)\Pr(Z=1)(1-\Pr(Z=1)) - E(T \mid Z=0)(1-\Pr(Z=1))\Pr(Z=1)}$$

$$= \frac{(E(y\cdot 1 \mid z=1) - E(y \mid Z=0))(1-\Pr(Z=1))\Pr(Z=1)}{(E(T\cdot 1 \mid z=1) - E(T \mid Z=0))(1-\Pr(Z=1))\Pr(Z=1)}$$

$$= \frac{E(y\cdot 1 \mid z=1) - E(y \mid Z=0)}{E(T\cdot 1 \mid z=1) - E(T \mid Z=0)}$$

How IV works – simplest case

First stage

$$x = c + dz + f$$

$$\hat{d} = \text{cov}(x,z) / \text{var}(z)$$

$$\hat{x} = \hat{c} + \hat{d}z$$

Second stage

$$y = a + b\hat{x} + e$$

- *In an RCT, x is usually binary or categorical (received treatment, did not receive treatment*
- *Here, x-hat is the predicted treatment uptake rate given assignment*
- *Thereby introducing (and statistically mitigating) the effects of partial compliance*

$$\hat{b}^{2SLS} = \frac{\text{cov}(y,\hat{x})}{\text{var}(\hat{x})} = \frac{\text{cov}(y,\hat{c}) + \hat{d}\,\text{cov}(y,z)}{\hat{d}^2\,\text{var}(z)} = \frac{\text{cov}(y,z)}{\frac{\text{cov}(x,z)}{\text{var}(z)}\text{var}(z)} = \frac{\text{cov}(y,z)}{\text{cov}(x,z)} = \hat{b}^{IV}$$

Partial compliance

- So far:
  - In controlled RCTs…
    - control group implies no treatment
    - treatment group implies get treatment
- Often things are not as clean as this
  - Treatment is an opportunity/choice
  - Close substitutes available to those in control group
  - Implementation not perfect e.g. pushy parents getting their kid into a charter school even if they were not selected in a voucher lottery

An example: JTPA

- The Job Training Partnership Act (JTPA) included a large randomized trial to evaluate the effect of training on earnings
- The JTPA offered treatment randomly; participation was voluntary
- Roughly 60 percent of those offered training received it
- IV setup
  - T_i indicates those who **received JTPA services**
  - z_i indicates the **random offer of treatment** ('assignment' to treatment)
  - Y_i is **earnings** in the 30 months since random assignment
- The first-stage here is approximately the compliance rate ~
- $E[T_i \mid zi = 1] - E[T_i \mid zi = 0] = P[T_i = 1 \mid z_i = 1] = .60$
  - (.62 of z_i =1 group trained; .02 of zi =0 group also trained)

An example: JTPA

TABLE 4.4.1

Results from the JTPA experiment: OLS and IV estimate

	Treatment effect on treated: ***E[y | T=1] - E[y | T=0]*** **(TOT)**		**Intent to Treat:** **E[y | z=1] - E[y | z=0]** **(ITT)**		**2SLS:** **E[y | z=1] - E[y | z=0]** **E[T | z=1] - E[T | z=0]** **(ITT)**	
	Without Covariates (1)	With Covariates (2)	Without Covariates (3)	With Covariates (4)	Without Covariates (5)	With Covariates (6)
A. Men	3,970	3,754	1,117	970	1,825	1,593
	(555)	(536)	(569)	(546)	(928)	(895)
B. Women	2,133	2,215	1,243	1,139	1,942	1,780
	(345)	(334)	(359)	(341)	(560)	(532)

Notes: Authors' tabulation of JTPA study data. The table reports OLS, ITT, and IV estimates of the effect of subsidized training on earnings in the JTPA experiment. Columns 1 and 2 show differences in earnings by training status; columns 3 and 4 show differences by random-assignment status. Columns 5 and 6 report the result of using random-assignment status as an instrument for training. The covariates used for columns 2, 4, and 6 are high school or GED, black, Hispanic, married, worked less than 13 weeks in past year, AFDC (for women), plus indicators for the JTPA service strategy recommended, age group, and second follow-up survey. Robust standard errors are shown in parentheses. There are 5,102

Another example: Domestic violence

- What's the best police response to domestic violence? The Minneapolis Domestic Violence Experiment (MDVE; Sherman and Berk, 1984) boldly tried to find out. . .
- Police were randomly assigned to advise, separate, or arrest
- Substantial compliance problems as officers made their own judgements in the field

Table 1: Assigned and Delivered Treatments
in Spousal Assault Cases

Assigned Treatment	Delivered Treatment			Total
		Coddled		
	Arrest	Advise	Separate	
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)

Another example: Domestic violence

- Outcome: recidivism

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 (failed)	

Another example: Domestic violence

- Outcome: recidivism

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)

Motivation: units of measurement

- Causal effect measured in units of 'experiment' – not very helpful
- Often want to convert causal effects to more meaningful units e.g. in Project STAR what is effect of reducing class size by one child

Using IV to scale experiments

- For example the Project Star experiment that improved outcomes by reducing class size.

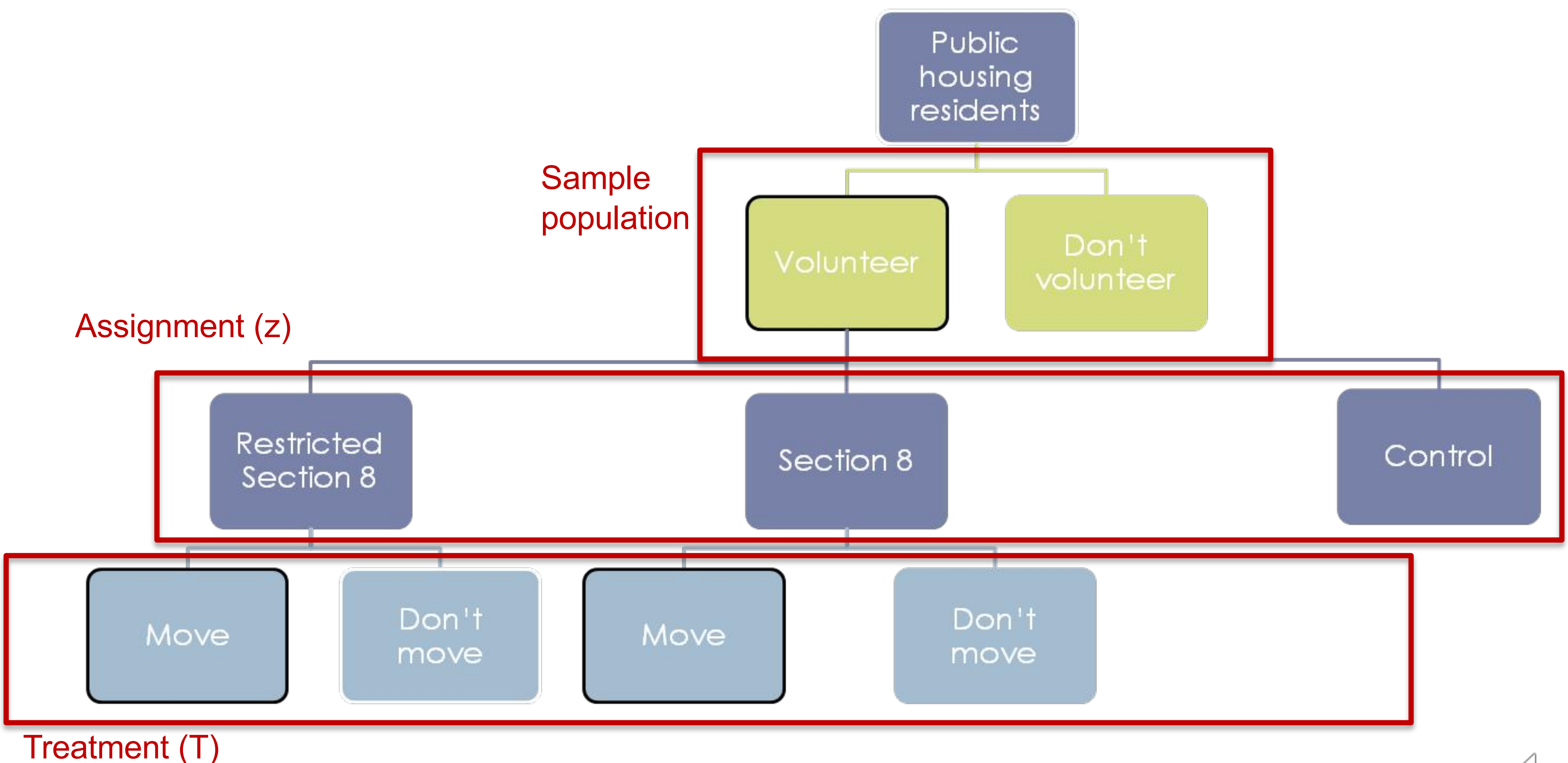
$$\delta = \frac{E\left(y \middle| z=1\right) - E\left(y \middle| z=0\right)}{E\left(T \middle| z=1\right) - E\left(T \middle| z=0\right)}$$

- Where *T* is class size
- Takes the treatment effect on outcome variable and divides by treatment effect on class size
- Wald again!
- Not hard to compute but how to get standard error?

An example: Moving to Opportunity

- Designed to investigate the impact of living in bad neighborhoods on outcomes
- Gave some residents of public housing projects chance to move out
- Two treatments:
  - Voucher for private rental housing (Section 8)
  - Voucher for private rental housing restricted for use in ‘good’ neighborhoods (Section 8 with restrictions)
  - + Control
- No-one forced to move so imperfect compliance – 60% and 40% did use it

Moving to Opportunity



Some terminology

- z denotes whether in control or treatment group – 'intention-to-treat'
- T denotes whether actually get treatment
- With perfect compliance:
  - $\Pr(T=1|z=1)=1$
    - *Those in 'treatment group' housing always move*
  - $\Pr(T=1|z=0)=0$
    - *Among those in the control group, moving cannot be at all predicted by group assignment*
- With imperfect compliance:

  $1>\Pr(T=1|z=1)>\Pr(T=1|z=0)>0$

  Sometimes, those in 'treatment group' housing do not move, and sometimes those in control group housing move

What do we want to estimate?

- 'Intention-to-Treat' :

$$ITT = E(y|Z=1) - E(y|Z=0)$$

- This can be estimated in usual way
- Treatment Effect on Treated

$$TOT = \frac{E(y|Z=1) - E(y|Z=0)}{E(T|Z=1) - E(T|Z=0)}$$

Estimating TOT

- Can't use simple regression of y on Z
- But should recognize TOT as Wald estimator
- Can estimated by regressing y on T using Z as instrument
- Relationship between TOT and ITT:

$$ITT = TOT \bullet \left[\Pr\left(T = 1 \middle| Z = 1\right) - \Pr\left(T = 1 \middle| Z = 0\right) \right]$$

MTO: What might this example look like in Stata?

- ***Y*: Outcome measure** (one of many)
  - Graduated from high school (for example)
  - `grad`
- ***T*: Actually received treatment**
  - Moved to housing for which they got a voucher
  - `moved`
- ***z:* Instrument (treatment group to which they were assigned)**
  - Voucher for private rental housing (Section 8)
  - Voucher for private rental housing for use in ‘good’ neighborhoods
  - Control group (no voucher)
  - `voucher`

```
ivreg grad (moved = voucher)
```

Most important results from MTO

- No effects on adult economic outcomes
- Improvements in adult mental health
- Beneficial outcomes for teenage girls
- Adverse outcomes for teenage boys

Sample results from MTO

	E/S (i)	CM (ii)	ITT (iii)	TOT (iv)
A. Adult outcomes				
Obese, BMI≥30	E-C	.468	-.051 (.022)	-.108 (.048)
Calm and peaceful	E-C	.466	.061 (.022)	.131 (.047)
Psychological distress, K6 z-score	E-C	.050	-.092 (.046)	-.196 (.099)

- TOT approximately twice the size of ITT
- Consistent with 50% use of vouchers

Bombs, Bones and Breakpoints:
The Geography of Economic Activity
Davis and Weinstein, AER, 2002

- Existence of agglomerations (e.g. cities) a puzzle
- Land and labour costs higher so why don’t firms relocate to increase profits
- Must be some compensatory productivity effect
- Different hypotheses about this:
  - Locational fundamentals
  - Increasing returns (Krugman) – path-dependence

Testing these hypotheses

- Consider a temporary shock to city population
- Locational fundamentals theory would predict no permanent effect.
- Increasing returns would suggest permanent effect.
- Would like to do experiment of randomly assigning shocks to city size.
- This is not going to happen.

The Davis-Weinstein idea

- Use US bombing of Japanese cities in WW2.
- This is a ‘natural experiment’ not a true experiment because:
  - WW2 not caused by desire to test theories of economic geography
  - Pattern of US bombing not random
- Sample is 303 Japanese cities, data is:
  - Population before and after bombing
  - Measures of destruction

Basic equation

$$\Delta s_{i,60-47} = \beta_0 + \beta_1 \Delta s_{i,47-40} + \beta_2 x_i + \varepsilon_i$$

- $\Delta s_{i,47\text{-}40}$ is change in population just before and after war
- $\Delta s_{i,60\text{-}47}$ is change in population at later period
- How to test hypotheses:
  - Locational fundamentals predicts β_1=-1
  - Increasing returns predicts β_1=0

The IV approach

- $\Delta s_{i,47\text{-}40}$ might be influenced by both permanent and temporary factors
- Only want part that is transitory shock caused by war damage
- Instrument $\Delta s_{i,47\text{-}40}$ by measures of death and destruction

The first-stage: Correlation of $\Delta s_{i,47\text{-}40}$ with z

TABLE 2—INSTRUMENTAL VARIABLES EQUATION (DEPENDENT VARIABLE = RATE OF GROWTH IN CITY POPULATION BETWEEN 1940 AND 1947)

Independent variable	Coefficient
Constant	0.213
	(0.006)
Deaths per capita	−0.665
	(0.506)
Buildings destroyed per capita	−2.335
	(0.184)
R^2:	0.409
Number of observations:	303

Note: Standard errors are in parentheses.

Why do we need first-stage?

- Establishes instrument relevance – correlation of *T* and *z*
- Gives an idea of how strong this correlation is – ‘weak instrument’ problem (next time)
- In this case reported first-stage not obviously that implicit in what follows
  - That would be bad practice

The IV estimates

TABLE 3—TWO-STAGE LEAST-SQUARES ESTIMATES OF IMPACT OF BOMBING ON CITIES (INSTRUMENTS: DEATHS PER CAPITA AND BUILDINGS DESTROYED PER CAPITA)

	Dependent variable = growth rate of population between 1947 and 1960		Dependent variable = growth rate of population between 1947 and 1965
Independent variable	(i)	(ii)	(iii)
Growth rate of population between 1940 and 1947	−1.048 (0.097)	−0.759 (0.094)	−1.027 (0.163)
Government reconstruction expenses	1.024 (0.387)	0.628 (0.298)	0.392 (0.514)
Growth rate of population between 1925 and 1940		0.444 (0.054)	0.617 (0.092)
R^2:	0.279	0.566	0.386
Number of observations:	303	303	303

Note: Standard errors are in parentheses.

Why are these other variables included?

- Potential criticisms of instrument exogeneity
  - Government post-war reconstruction expenses correlated with destruction and had an effect on population growth
  - US bombing heavier of cities of strategic importance (perhaps they had higher growth rates)
- Inclusion of the extra variables designed to head off these criticisms
- Assumption is that of exogeneity conditional on the inclusion of these variables
- Conclusion favours locational fundamentals view

An additional piece of supporting evidence....

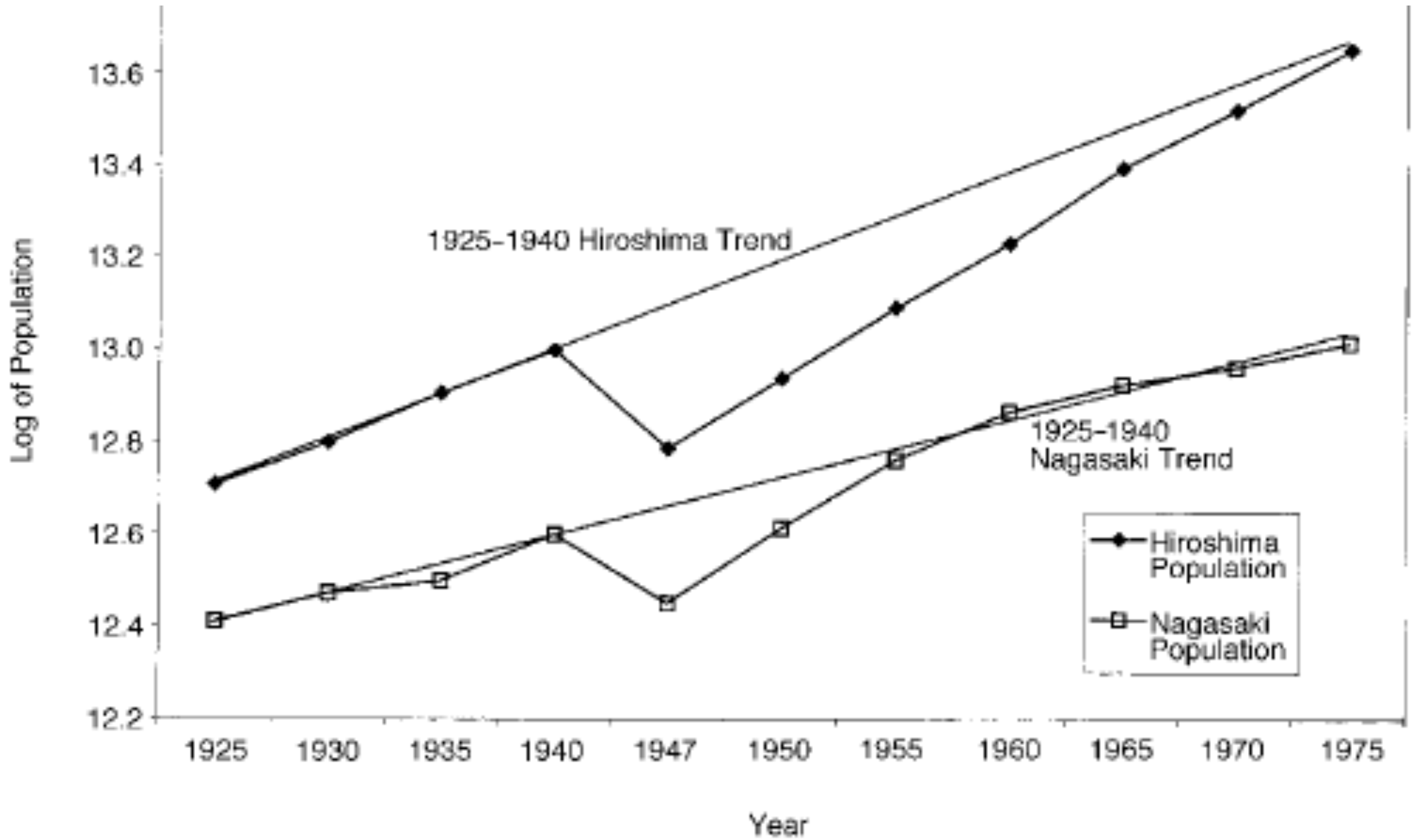


FIGURE 2. POPULATION GROWTH

- Always trying to build a strong evidence base – many potential ways to do this, not just estimating equations

Miguel and Roland, Vietnam

- Use pattern of bombing in Vietnam to examine the effect of war on poverty.
- Is bombing random? No (next slide).
- This is a case where they try to argue for conditional exogeneity.

Miguel and Roland

Table 3: Predicting bombing intensity

	Dependent variable: Total U.S. bombs, missiles, and rockets per km^2			
	(1)	(2)	(3)	(4)
| Latitude – 17°N |	-14.8***	-17.0***	-10.2***	-27.8
	(5.3)	(6.0)	(2.2)	(16.2)
Population density (province), 1960-61	0.0050	-0.0035**	-0.0034**	-.163*
	(0.0043)	(0.0016)	(0.0014)	(0.083)
Former South Vietnam	-138.5*	-134.5	-37.1	-171.3
	(74.9)	(87.2)	(27.7)	(118.8)
Proportion of land area 250-500m	89.5*	-27.6	-26.6*	-104.5*
	(47.1)	(20.5)	(14.2)	(54.9)
Proportion of land area 500-1000m	-49.6	-17.7	-10.5	-52.2
	(65.3)	(18.9)	(16.8)	(31.8)
Proportion of land area over 1000m	156.3*	-6.0	-19.8	-50.6
	(81.4)	(30.4)	(19.1)	(31.2)
Average precipitation (cm)	0.26	0.22	0.15*	0.09
	(0.17)	(0.18)	(0.08)	(0.31)
Average temperature (celsius)	15.2	-0.2	-0.6	7.6
	(0.8)	(4.4)	(3.6)	(5.6)
Latitude (°N)	-8.7	-10.0	-2.3	-15.5
	(6.3)	(7.1)	(2.6)	(12.9)
District soil controls	No	Yes	Yes	Yes
Exclude Quang Tri province	No	No	Yes	No
Central Region sample	No	No	No	Yes
Observations	55	584	576	229
R^2	0.54	0.33	0.25	0.43
Mean (s.d.) dependent variable	30.6 (51.7)	32.3 (68.5)	27.1 (50.6)	56.7 (91.0)

Miguel and Roland

Table 4: Local bombing impacts on estimated 1999 poverty rate

	Dependent variable: Estimated poverty rate, 1999						
	OLS	OLS	OLS	OLS	OLS	OLS	IV-2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Total U.S. bombs, missiles, and rockets per km^2	-0.00087*	-0.00040*	-0.00065***	-0.00079***	-0.00017		0.00026
	(0.00048)	(0.00022)	(0.00012)	(0.00016)	(0.00019)		(0.00042)
Population density (province), 1960-61 (÷100)	-0.0089***	-0.0021**		-0.0023**	-0.013	-0.0021**	-0.0020*
	(0.0016)	(0.0009)		(0.0010)	(0.010)	(0.0010)	(0.0010)
Former South Vietnam	-0.317***	-0.174**		-0.122*	-0.005	-0.139**	-0.104
	(0.087)	(0.071)		(0.071)	(0.047)	(0.058)	(0.082)
Proportion of land area 250-500m	0.341***	0.339***	0.182***	0.325***	0.285***	0.342***	0.349***
	(0.096)	(0.070)	(0.067)	(0.069)	(0.111)	(0.070)	(0.073)
Proportion of land area 500-1000m	0.386**	0.261***	0.157**	0.261***	0.161**	0.253***	0.257***
	(0.172)	(0.052)	(0.062)	(0.053)	(0.064)	(0.054)	(0.055)
Proportion of land area over 1000m	0.571**	-0.048	-0.001	-0.066	-0.187**	-0.044	-0.043
	(0.231)	(0.113)	(0.159)	(0.111)	(0.086)	(0.120)	(0.116)
Average precipitation (cm)	0.00027	0.00111***	0.00060	0.00110***	0.00070*	0.00068*	0.00063
	(0.00044)	(0.00035)	(0.00046)	(0.00033)	(0.00036)	(0.00038)	(0.00044)
Average temperature (celsius)	0.033	-0.012	-0.034	-0.013	-0.0373	-0.0143	-0.0143
	(0.029)	(0.019)	(0.022)	(0.020)	(0.0219)	(0.0196)	(0.0199)
Latitude (°N)	-0.0127	-0.0088	0.038	-0.0044	0.0211**	-0.0051	-0.0025
	(0.0108)	(0.0088)	(0.026)	(0.0088)	(0.0092)	(0.0081)	(0.0100)
| Latitude – 17°N |						-0.0044	
						(0.0069)	
District soil controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	No	No	Yes	No	No	No	No
Exclude Quang Tri province	No	No	No	Yes	No	No	No
Central Region sample	No	No	No	No	Yes	No	No
Observations	55	584	584	576	229	584	584
R^2	0.75	0.61	0.79	0.63	0.72	0.60	-
Mean (s.d.) dependent variable	0.39 (0.16)	0.41 (0.20)	0.41 (0.20)	0.41 (0.20)	0.43 (0.20)	0.41 (0.20)	0.41 (0.20)

IV is proximity to the 17th parallel, and arbitrary international border.

Miguel and Roland

Table 7: Local war impacts on infrastructure and human capital

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)	IV-2SLS (7)
Panel A: Dependent variable: Proportion of households with access to electricity, 1999							
Total U.S. bombs, missiles, and rockets per km^2	0.00168***	0.00036***	0.00025	0.00043**	0.00025*		0.0019**
	(0.00055)	(0.00012)	(0.00016)	(0.00017)	(0.00013)		(0.0009)
| Latitude – 17°N |						-0.033***	
						(0.009)	
District soil controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	No	No	Yes	No	No	No	No
Exclude Quang Tri province	No	No	No	Yes	No	No	No
Central Region sample	No	No	No	No	Yes	No	No
Observations	55	584	584	576	229	584	584
R^2	0.59	0.57	0.75	0.57	0.62	0.58	-
Mean (s.d.) dependent variable	0.72 (0.21)	0.71 (0.27)	0.71 (0.27)	0.71 (0.27)	0.67 (0.26)	0.71 (0.27)	0.71 (0.27)
Panel B: Dependent variable: Proportion of literate respondents, 1999							
Total U.S. bombs, missiles, and rockets per km^2	0.00005	0.00003	0.00009	0.00012**	-0.00003		0.00041
	(0.00012)	(0.00006)	(0.00006)	(0.00006)	(0.00006)		(0.00037)
| Latitude – 17°N |						-0.0070	
						(0.0052)	
District soil controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	No	No	Yes	No	No	No	No
Exclude Quang Tri province	No	No	No	Yes	No	No	No
Central Region sample	No	No	No	No	Yes	No	No
Observations	55	584	584	576	229	584	584
R^2	0.65	0.59	0.75	0.59	0.55	0.59	-
Mean (s.d.) dependent variable	0.89 (0.07)	0.88 (0.11)	0.88 (0.11)	0.88 (0.11)	0.86 (0.11)	0.88 (0.11)	0.88 (0.11)

Angrist and Evans

- Looking at the effect of fertility on labor supply.
- Many reasons not to use an OLS regression of labor supply on fertility (such as?)
- What is a good instrument? Something correlated with fertility but uncorrelated with error term… What does that mean?
- Think of a randomized encouragement for an extra child.

Randomized encouragement

- If your first two children are of the same sex then you are more likely to want a third child. Why? Gender balance.

Gender balance

TABLE 3—FRACTION OF FAMILIES THAT HAD ANOTHER CHILD BY PARITY AND SEX OF CHILDREN

	All women				Married women			
	1980 PUMS (649,887 observations)		1990 PUMS (627,362 observations)		1980 PUMS (410,333 observations)		1990 PUMS (477,798 observations)	
Sex of first child in families with one or more children	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
(1) one girl	0.488	0.694 (0.001)	0.489	0.665 (0.001)	0.485	0.720 (0.001)	0.487	0.698 (0.001)
(2) one boy	0.512	0.694 (0.001)	0.511	0.667 (0.001)	0.515	0.720 (0.001)	0.513	0.699 (0.001)
difference (2) − (1)	—	0.000 (0.001)	—	0.002 (0.001)	—	0.000 (0.001)	—	0.001 (0.001)

	All women				Married women			
	1980 PUMS (394,835 observations)		1990 PUMS (380,007 observations)		1980 PUMS (254,654 observations)		1990 PUMS (301,588 observations)	
Sex of first two children in families with two or more children	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)

Is this a good IV?

- Need to believe that sex of first two kids is random (precludes sex selection of first two).
- Need to believe that having same sex first kids doesn't affect anything other than having a third child. Examples?

The Wald estimator

TABLE 5—WALD Es

Variable	1980 PUMS: Mean difference by *Same sex*	Wald estimate using as covariate: *More than 2 children*	Wald estimate using as covariate: *Number of children*
More than 2 children	0.0600 (0.0016)	—	—
Number of children	0.0765 (0.0026)	—	—
Worked for pay	−0.0080 (0.0016)	−0.133 (0.026)	−0.104 (0.021)
Weeks worked	−0.3826 (0.0709)	−6.38 (1.17)	−5.00 (0.92)
Hours/week	−0.3110 (0.0602)	−5.18 (1.00)	−4.07 (0.78)
Labor income	−132.5 (34.4)	−2208.8 (569.2)	−1732.4 (446.3)
ln(*Family income*)	−0.0018 (0.0041)	−0.029 (0.068)	−0.023 (0.054)

Arsenic and new wells

- One of the greatest successes of public health: getting rural population to drink water from tube wells, bypassing contaminated surface water.
- But in Bangladesh, well water has naturally occurring arsenic – affects 60+ million people.
- What is the effect of arsenic poisoning on economic and cognitive outcomes?

OLS?

- Would you run an OLS regression, e.g., income or IQ test score on arsenic content in the body?
- No.
- Why?

IV

- So what you need is a randomized encouragement to have less/ more arsenic in your body.
- You could imagine using proximity to a well with low/high arsenic. Is this a good IV?

Pitt and Rosenzweig

- Use a biology literature indicating that there is natural genetic variation in ability to break down arsenic.
- Could this be a good IV?
- Yes: exogenous.
- No: not excluded because people sharing common genes also share a common environment and habits. Affects many things.

Solution

- Look at arsenic in the body, removing any systematic part due to where you live (village fixed effects) or what you eat (using information on food habits).
- What is left is genetic variation that doesn't operate through location or habits – presumably then it is *excluded* so a valid IV.

Intuition

- An instrumental variable, ***Z* is uncorrelated with the disturbance *e* but is correlated with *X***(e.g., proximity to college might be correlated with schooling but not with wage residuals)

- With this new variable, the **IV estimator should capture only the shift-in-*X*-induced effects on *Y*,** whereas the OLS estimator captures not only the direct effect of on but also the effect of the included measurement error and/or endogeneity

- **IV is not as efficient as OLS (especially if *Z* only weakly correlated with *X***, i.e. when we have so-called 'weak instruments') and only has large sample properties (consistency)

- IV results in biased coefficients. **The bias can be large in the case of weak instruments**

Conclusion

- Natural experiments useful source of knowledge
- Often requires use of IV
- Instrument exogeneity and relevance need justification
- Weak instruments potentially serious
- Good practice to present first-stage regression
- Finding more robust alternative to IV an active research area