Lecture 9 - Instrumental Variables
Part II

1. Review of instrumental variables:

$$y_i = \alpha + \rho s_i + \varepsilon_i \qquad \text{(structural eqn)}$$

s_i is endogenous so instrument:

(a) $$s_i = \pi_{10} + \pi_{11} z_i + \eta_i \qquad \text{(first stage)}$$

$$y_i = \pi_{20} + \pi_{21} z_i + u_i \qquad \text{(reduced form)}$$

Plugging in first-stage into structural:

$$y_i = \alpha + \rho(\pi_{10} + \pi_{11} z_i + \eta_i) + \varepsilon_i$$
$$y_i = (\alpha + \rho\pi_{10}) + \rho\pi_{11} z_i + (\rho\eta_i + \varepsilon_i)$$
$$y_i = \pi_{20} + \pi_{21} z_i + u_i$$

Taking $$\frac{\pi_{21}}{\pi_{11}} = \frac{\text{reduced-form coeff } z_i}{\text{first-stage coeff } z_i} = \frac{\rho\cancel{\pi_{11}}}{\cancel{\pi_{11}}} = \rho$$

$\rho \Rightarrow$ $\boxed{\text{causal estimate?}}$

(b) In order for 2SLS estimate to be causal:

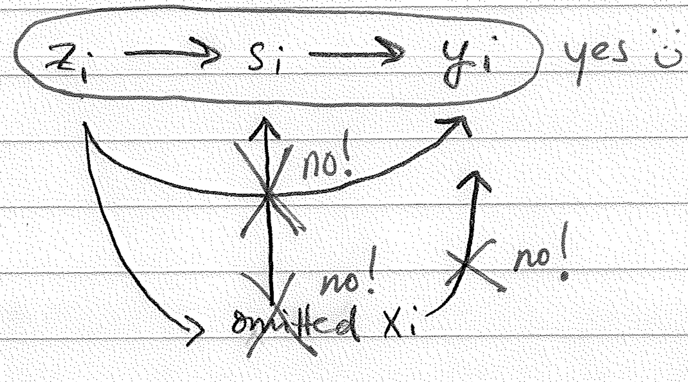
i) strong instrument:

$$s_i = \pi_{10} + \pi_{11} z_i + \eta_i$$

↳ π_{11} is $\neq 0$, high R^2 (or F)

ii) z_i is a "valid" instrument. ⇒

z_i only affects y_i through its effect on s_i.



(c) when you run 2SLS with multiple instruments:

i) F-stat ≥ 10. The larger, the better.

ii) Hansen-Sargan p-value $\geq .10$. The larger, the better (i.e., the closer to 1).

H_0: instruments are <u>not</u> invalid

H_a: instruments are invalid
$E[z_i \hat{\varepsilon}_i] \neq 0$. In other words, instruments are correlated with error term

(If we "accept" H_0, then we think instruments are valid)

"Accepting" H_0 means we cannot reject H_0, which means p-value $\geq .10$.

So we really DO NOT want to reject H_0 in favor of H_a. Note that "accepting" H_0 does not mean instruments are valid. It means we think they are not invalid.

How to run Hansen-Sargan Test:

- estimate by 2SLS:

$$y_i = \alpha + \rho S_i + X_i + \varepsilon_i \text{ where}$$

$$s_i = \pi_{10} + \pi_{11} z_{i1} + \pi_{12} z_{i2} + X_i + \eta_i$$

- obtain $\hat{\varepsilon}_i$
- estimate by OLS

$$\hat{\varepsilon}_i = \gamma_{10} + \gamma_{11} z_i + \gamma_{12} z_{i2} + X_i + \xi_i$$

- obtain R^2

Test $\gamma_{10} = 0$ and $\gamma_{12} = 0$

$R^2 \approx 0$ if z uncorrelated with $\hat{\varepsilon}_i \Rightarrow$ Instrument not invalid.

$NR^2 \sim$ chi-square(# z_i's − # s_i's)

↳ obtain p-value.

2. Wald Estimator : special case of IV
- when have 1 endog variable and 1 instrument

AND - when z_i is binary: $z_i = \begin{cases} 1 \\ 0 \end{cases}$

$$Y_i = \alpha + \rho s_i + \varepsilon_i$$
$$s_i = \pi_{10} + \pi_{11} z_i + \eta_i$$

$$\rho = \frac{E[Y_i | z_i = 1] - E[Y_i | z_i = 0]}{E[s_i | z_i = 1] - E[s_i | z_i = 0]} = \frac{\text{reduced form est.}}{\text{1st stage est.}}$$

$$\Rightarrow E[Y_i | z_i] = \alpha + \rho E[s_i | z_i] + \underbrace{E[\varepsilon_i | z_i]}_{=0 \text{ if } \mathrm{cov}(\varepsilon_i, z_i) = 0}$$

Example: Wald Estimate of Return to Schooling Using Quarter of Birth:

TABLE 4.1.2
Wald estimates of the returns to schooling using quarter-of-birth instruments

	(1) Born in 1st Quarter of Year	(2) Born in 4th Quarter of Year	(3) Difference (Std. Error) (1) – (2)
ln (weekly wage)	5.892	5.905	−.0135 (.0034)
Years of education	12.688	12.839	−.151 (.016)
Wald estimate of return to education			.089 (.021)
OLS estimate of return to education			.070 (.0005)

Notes: From 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 162,515.

$$Y_i = \pi_{20} + \pi_{21} z_i + u_i \quad \text{(reduced form)}$$

$$z_i = \begin{cases} 1 & \text{if born in 1st quarter,} \\ 0 & \text{otherwise} \end{cases}$$

ln wages:

$$5.905 = \pi_{20} + \pi_{21}(0) \quad \text{if born in } 4^{th}\text{ Q}$$

$$\Rightarrow \boxed{\pi_{20} = 5.905}$$

$$5.892 = \pi_{20} + \pi_{21}(1) \quad \text{if born in } 1^{st}\text{ Q}$$

$$= 5.905 + \pi_{21}$$

$$\Rightarrow \boxed{\pi_{21} = -.0135} \Rightarrow$$

Being born in 1st Q is associated with -.0135 lower wages

Years educ:

$$12.839 = \pi_{10} + \pi_{11}(0) \quad \text{if born in } 4^{th}\text{ Q}$$

$$\Rightarrow \boxed{\pi_{10} = 12.839}$$

$$12.688 = \pi_{10} + \pi_{11}(1) \quad \text{if born in 1st Q}$$

$$= 12.839 + \pi_{11}$$

$$\Rightarrow \boxed{\pi_{11} = -.151} \Rightarrow$$

Being born in 1st Q is associated with -.151 less years of schooling.

$$\boxed{\frac{\frac{\Delta \text{ wages}}{\Delta Q \text{ birth}}}{\frac{\Delta \text{ schooling}}{\Delta Q \text{ birth}}} = \frac{\frac{-.0135}{1}}{\frac{-.151}{1}} = \frac{-.0135}{-.151} = .089 = \frac{\Delta \text{ wages}}{\Delta \text{ schooling}}}$$

What's really nice about Wald Estimator?

- nice, intuitive, simple, can compute by hand! (Like we just did).

* - All 2SLS is just fancy versions of Wald.

i) if 1 endog and 2 dummy instrumental vars:

a) $Y_i = \alpha_1 + \rho_1 \hat{S}_{i1} + \varepsilon_i$

where $\hat{S}_{i1} = \hat{\Pi}_{10} + \hat{\Pi}_{11} Z_{i1}$

b) $Y_i = \alpha_2 + \rho_2 \hat{S}_{i2} + \varepsilon_i$

where $\hat{S}_{i2} = \hat{\Pi}_{20} + \hat{\Pi}_{21} Z_{i2}$

c) $Y_i = \alpha + \rho \hat{S}_i + \varepsilon_i$

where $\hat{S}_i = \hat{\Pi}_{30} + \hat{\Pi}_{31} Z_{i1} + \hat{\Pi}_{32} Z_{i2}$

$$\hat{\rho} \approx \frac{\hat{\rho}_1 + \hat{\rho}_2}{2} \approx (0.5)\hat{\rho}_1 + (0.5)\hat{\rho}_2$$

more generally:

$$\hat{\rho} = w\hat{\rho}_1 + (1-w)\hat{\rho}_2$$

$$w = \frac{\frac{n_a}{\sigma^2_{\varepsilon a}}}{\frac{n_a}{\sigma^2_{\varepsilon a}} + \frac{n_b}{\sigma^2_{\varepsilon b}}}$$

$w \in (0,1)$ with w based on variance of $\hat{\varepsilon}_i$ from (a) and $\hat{\varepsilon}_i$ from (b).

ii) Even if z_i is not a dummy $\{0,1\}$ variable, it can often be made into a dummy.

e.g. z_i is $\{1,2,3,4\}$ quarter of birth

gen four quarter-of-birth dummies:

$\Rightarrow$

	z_1	z_2	z_3	z_4
Q1	1	0	0	0
Q2	0	1	0	0
Q3	0	0	1	0
Q4	0	0	0	1

iii) If we're thinking of monitoring and evaluation of "randomised" experiments, $z_i = \{0,1\}$ is <u>assignment</u> to treatment.

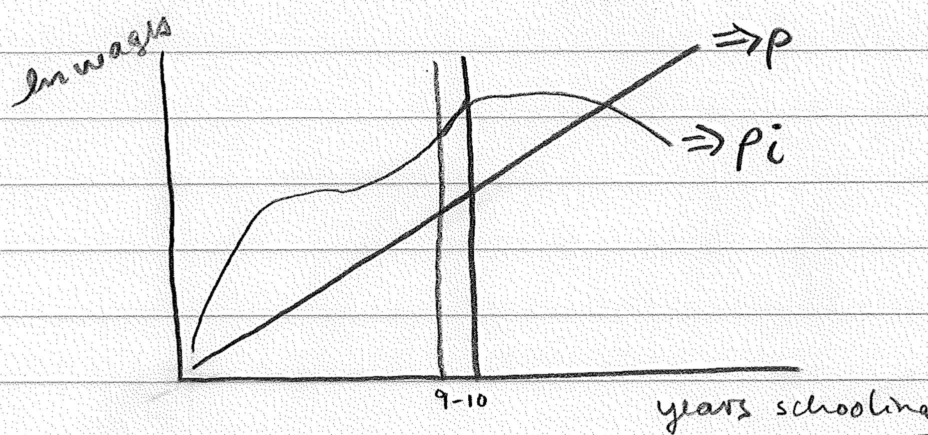
3. Interpreting IV Estimates.

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

or

$$Y_i = \alpha + \underline{\underline{\rho_i}} S_i + \varepsilon_i$$

ρ or ρ_i? Homogenous or heterogenous effects of schooling on wages.



Imagine a policy that compels individuals to remain in school until they're 16. Those that turn 16 earlier drop out at grade 10; those that turn 16 later drop out at grade 11.

From this instrument, $\hat{\rho} = .089 \Rightarrow$
1 extra year = 9% higher earnings.

- Generalizeable?
  - If we had induced one extra year between 8^{th} - 9^{th} grade, would $\rho = .089$?

- Average Population?
  - What kinds of people are induced to get one extra year of schooling but would have dropped out otherwise?

Who does instrument target?

never takers : would have dropped out much earlier anyways (even though illegal)

always takers : would have gotten much more school anyways.

compliers : induced into getting more school because of treatment.

LATE (Local Average Treatment Effect)

↳ Effect of 1 extra year of schooling between 10^{th} - 11^{th} grade for those who are induced into remaining in school because of minimum age laws is $\rho = .089$.