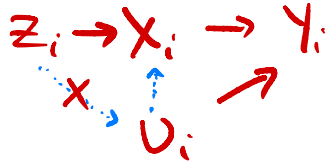
Lecture Note 10: Instrumental Variables

$$Y_i = \beta_0 + \beta_1 X_i + U_i$$

homogeneous effects



For Z_i to be a valid instrument, 2 assumptions:

1. Relevance: $\text{cov}(Z_i, X_i) \neq 0$
2. Exogeneity: $\text{cov}(Z_i, U_i) = 0$
   - ↳ implicitly: a) Z_i only affects Y_i through X_i
     b) Z_i is "as good as random"

$$\text{cov}(Z_i, Y_i) = \text{cov}(Z_i, \beta_0 + \beta_1 X_i + U_i) = \cancel{\text{cov}(Z_i, \beta_0)} + \text{cov}(Z_i, \beta_1 X_i) + \cancel{\text{cov}(Z_i, U_i)}$$

↳

$$\text{cov}(Z_i, Y_i) = \beta_1 \text{cov}(Z_i, X_i)$$

$$\beta_1 = \frac{\text{cov}(Z_i, Y_i)/V[Z_i]}{\text{cov}(Z_i, X_i)/V[Z_i]} = \frac{\text{reduced form}}{1^{st} \text{ stage}}$$

1. Regress Y_i on Z_i (RF)
2. Regress X_i on Z_i (1st)
3. Compute ratio

Two-stage least squares (TSLS)

① First stage: $X_i = \pi_0 + \pi_1 Z_i + V_i$, then form $\hat{X}_i = \hat{\pi}_0 + \hat{\pi}_1 Z_i$

② Second stage: $Y_i = \beta_0^{TSLS} + \beta_1^{TSLS} \hat{X}_i + \varepsilon_i$

$\hat{\beta}_1^{TSLS}$ is same as ratio and consistent estimator of β_1

In R using fixest: y on x, w_1, w_2 and z as instrument

feols(y ~ w1 + w2 | x ~ z, data = df)

Wald Estimator: IV with binary instrument

→ two groups: $Z_i = 0$, $Z_i = 1$

→ $\bar{Y}_0, \bar{Y}_1, \bar{X}_0, \bar{X}_1$

→ $\hat{\beta}_1^{wald} = \dfrac{\bar{Y}_1 - \bar{Y}_0}{\bar{X}_1 - \bar{X}_0}$

↑ in eligibility experiment, $\bar{X}_0 = 0$

Heterogeneous effects

- binary instrument Z_i, binary endogenous variable X_i
- potential outcomes: $Y_i(x, z)$ for treatment level x and instrument level z
- potential treatment states: $X_i(z) < X_i(1), X_i(0)$
- to interpret IV estimates, 3 assumptions:

① Independence: $\{Y_i(x,z), X_i(z)\} \perp Z_i$ ← Z "as good as random"

can estimate 1st-stage and RF regressions

② Exclusion restriction: $Y_i(x,1) = Y_i(x,0)$
$Y_i(x,z) = Y_i(x)$

$Z_i \to X_i \to Y_i$

③ Monotonicity: Either $X_i(1) \geq X_i(0)$ for all i
or $X_i(0) \geq X_i(1)$ for all i

→ people with $X_i(1) \neq X_i(0)$ are "compliers"

		$Z_i = 1$	
		$X_i = 0$	$X_i = 1$
$Z_i = 0$	$X_i = 0$	never takers	compliers
	$X_i = 1$	~~defiers~~ by monotonicity	always takers

LATE: never, always, and compliers

TOT: never and compliers

Under assumption 1-3:

$$\hat{\beta}_1^{TSLS} \xrightarrow{p} \text{LATE} = \text{avg. effect of } X \text{ among compliers}$$

Generalize to non-binary X and Z:

$$X_i = \pi_{0i} + \pi_{1i} Z_i + V_i \qquad 1^{st}$$

$$Y_i = \beta_{0i} + \beta_{1i} X_i + U_i \qquad 2^{nd}$$

Then:

$$\hat{\beta}_1^{TSLS} \xrightarrow{p} E\left[\frac{\overbrace{\pi_{1i}}^{\text{ind. sensitivity to } Z_i}}{\underbrace{E[\pi_{1i}]}_{\text{mean sensitivity to } Z_i}} \underbrace{\beta_{1i}}_{\text{ind. treatment effect}}\right]$$