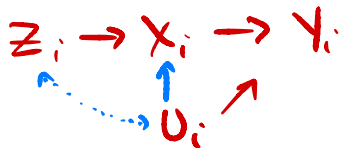


Lecture Note 10: IV

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

↑
homogeneous
effect



Assumptions:

① Relevance: $\text{cov}(Z_i, X_i) \neq 0$

② Exogeneity: $\text{cov}(Z_i, U_i) = 0$

a) Z_i has no direct effect on Y_i

b) Z_i is "as good as random"

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

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

$$\beta_1 = \frac{\text{cov}(Z_i, Y_i) / \sqrt{V(Z_i)}}{\text{cov}(Z_i, X_i) / \sqrt{V(Z_i)}}$$

① Regress Y_i on Z_i (reduced form)

② Regress X_i on Z_i (1st stage)

③ Ratio $\frac{RF}{1st}$

Two-stage least squares (TSLS)

① Regress X_i on Z_i : $X_i = \pi_0 + \pi_1 Z_i + V_i$

Form predicted vals: $\hat{X}_i = \hat{\pi}_0 + \hat{\pi}_1 Z_i$

② Regress Y_i on \hat{X}_i : $Y_i = \beta_0 + \beta_1 \hat{X}_i + \varepsilon_i$

$\hat{\beta}_1^{TSLS}$ is same as ratio $\frac{RF}{2^{st}}$ and consistent estimator for β_1 .

In R using `fixest`, regress Y on X, W_1, W_2 , instrumenting for X using Z

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

Wald estimator: Z_i is binary

$$\rightarrow \bar{y}_0, \bar{y}_1, \bar{x}_0, \bar{x}_1$$

$$\rightarrow \hat{\beta}_1^{\text{wald}} = \frac{\bar{y}_1 - \bar{y}_0}{\bar{x}_1 - \bar{x}_0} = 0 \text{ in eligibility experiment}$$

Heterogeneous effects

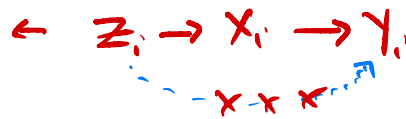
- single binary z_i , single binary X_i
- potential outcome: $Y_i(x, z)$ for treatment level x and instrument level z
- potential treatment status: $X_i(z) \begin{cases} X_i(1) \\ X_i(0) \end{cases}$

• to interpret IV/TSLs, 3 assumptions:

① Independence: $\{Y_i(x, z), X_i(z)\} \perp Z_i$

$\leftarrow Z_i$ is "as good as random"
 \leftarrow allows us to estimate 2nd-stage and RF regressions

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



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

| | | $Z_i = 1$ | |
|-----------|-----------|-----------------------------------|------------------|
| | | $X_i = 0$ | $X_i = 1$ |
| $Z_i = 0$ | $X_i = 0$ | never takers | compliers |
| | $X_i = 1$ | defers monotonicity | always takers |

TOT \rightarrow never,
compliers

LATE \rightarrow never,
always,
compliers

Under assumptions 1-3:

$$\hat{\beta}_i^{\text{TSLS}} \rightarrow \text{LATE} = \text{avg effect of } X \text{ on } Y \text{ among compliers}$$

Generalizing to non-binary Z_i and X_i :

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

$$Y_i = \beta_{0i} + \beta_{1i} Z_i + U_i$$

1st
2nd

Then:

$$\hat{\beta}_i^{\text{TSLS}} \rightarrow E \left[\frac{\hat{\pi}_{1i}}{E[\hat{\pi}_{1i}]} \beta_{1i} \right]$$

ind. sensitivity to Z_i

ind. causal effect of X_i on Y_i

avg sensitivity to Z_i