IV (2SLS)

X - stochastic

$\hat{\beta}_{OLS}$ consistency $\quad cov(x, \varepsilon) = 0$

unbiased $\quad x \perp \varepsilon$

$$\hat{\beta} = (X'X)^{-1}X'y = (X'X)^{-1}X'(X\beta + \varepsilon) =$$

$$\beta + (X'X)^{-1}X'\varepsilon$$

$$E(\hat{\beta}) = \beta + E((X'X)^{-1}X'\varepsilon)$$

deterministic

$$(X'X)^{-1}X'E(\varepsilon) = 0$$

bias $= 0$ if $X \perp \varepsilon$

($cov(X, \varepsilon) = 0$ is not enough)

$$E(\hat{\beta}|X) = \beta + E(\underbrace{(X'X)^{-1}X'}\varepsilon|X)$$

$E(\varepsilon|X) = 0 \rightarrow E(\varepsilon) = 0$

$E(\varepsilon|X) = 0 \rightarrow Cov(\varepsilon, x) = 0$

IV

X - endogeneous var

z - instrument

exogenous $cov(z, \varepsilon) = 0$

relevance $cov(z, x) \neq 0$

$$\rightarrow \quad y_i = \beta_1 + \beta_2 X_i + \varepsilon_i$$

$$d(z)$$

$$\textcircled{1} \quad \hat{X}_i = \hat{\theta}_1 + \hat{\theta}_2 \cdot z_i \qquad \hat{\theta}_2 = \frac{\hat{cov}(x, z)}{\hat{Var}(z)}$$

$$\textcircled{2} \quad y_i = \hat{\beta}_1 + \hat{\beta}_2 \cdot \hat{X}_i + \hat{\varepsilon}_i \quad \Rightarrow \quad \hat{\beta}_2^{IV}$$

Problem 1.

$$\hat{\beta}_2 = \frac{\hat{cov}(\hat{x}, y)}{\hat{Var}(\hat{x})} = \frac{\hat{cov}(\hat{\theta}_1 + \hat{\theta}_2 \cdot z, y)}{\hat{Var}(\hat{\theta}_1 + \hat{\theta}_2 \cdot z)} =$$

$$= \frac{\cancel{\hat{\theta}_1 \hat{cov}(1, y)}^{\to 0} + \cancel{\hat{\theta}_2} \hat{Cov}(z, y)}{(\hat{\theta}_2)^{\cancel{2}} \cdot \hat{Var}(z)} =$$

$$= \frac{\hat{cov}(z, y)}{\frac{\hat{cov}(z, x)}{\hat{Var}(z)} \cdot \hat{Var}(z)} = \frac{\hat{cov}(z, y)}{\hat{cov}(z, x)}$$

$$\frac{\hat{cov}(\overset{z}{x}, y)}{\hat{cov}(\underset{z}{x}, x)}$$

$$\underset{n\to\infty}{\text{plim}}\ \hat{\beta}_2^{IV} = \underset{n\to\infty}{\text{plim}} \frac{\widehat{\text{Cov}}(z,y)}{\widehat{\text{Cov}}(z,X)} = \frac{\text{Cov}(z_i, y_i)}{\text{Cov}(z_i, X_i)} =$$

$$= \frac{\text{Cov}(z_i, \beta_1 + \beta_2 \cdot X_i + \varepsilon_i)}{\text{Cov}(z_i, X_i)} = \beta_2 + \frac{\overbrace{\text{Cov}(z_i, \varepsilon_i)}^{=0 \text{ exogenous}}}{\underbrace{\text{Cov}(z_i, X_i)}_{\neq 0 \text{ relevant}}}$$

$$se(\hat{\beta}_2^{IV}) = \sqrt{\underbrace{\frac{s^2}{\Sigma(X_i - \bar{X})^2}}_{Var(\hat{\beta}_2^{OLS})} \cdot \frac{1}{\widehat{\text{corr}}^2(X, z)}}$$

$$Var(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

$$s^2 = \frac{\Sigma e_i^2}{n-2}$$

$$e_i = y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i$$

2SLS

$X_1, \dots, X_p$ - endogenous vars

$W_1, \dots, W_r$ - exogenous vars

$Z_1, \dots, Z_m$ - instrumental vars

1. $\hat{X}_1 \mid Z_1, \dots, Z_m, W_1, \dots, W_r$

$\vdots$

$\hat{X}_p \mid Z_1, \dots, Z_m, W_1, \dots, W_r$

2. $y \mid \hat{X}_1, \dots, \hat{X}_p, W_1, \dots, W_r$

$m < p$

under identified

$\hat{X}_{1i} = \hat{\alpha}_1 + \hat{\alpha}_2 \cdot z_i$

$\hat{X}_{2i} = \hat{\beta}_1 + \hat{\beta}_2 \cdot z_i$

2 step! $y_i = \hat{\theta}_1 + \hat{\theta}_2 \hat{X}_{1i} + \hat{\theta}_3 \hat{X}_{2i} + \hat{\varepsilon}_i$

$\rightarrow$ perfect multicollinearity

$m = p$ exact identified $\quad \hat{\beta}_{IV} = (Z'X)^{-1} Z'y$

$m > p$ over identified $\quad \hat{\beta}_{2SLS}$

Test: ① Relevance (Weak instruments)

② Exogeneity (Sargan test)

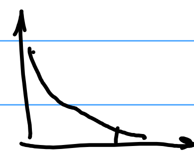
③ $\hat{\beta}_{OLS}$ vs $\hat{\beta}_{IV}$ (Hausmann test)

① $x \mid z_1, \dots, z_m, w_1, \dots, w_r \Rightarrow F$

$F \geq 10$ strong instruments z

$F < 10$ weak instruments z

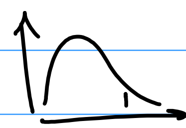
② Sargan's test $(m > p)$



$\frac{\sum_{i=1}^{k} \xi_i^2}{k}$

H_0: z - exogeneous

H_a: z - endogenous



$\hat{\varepsilon}_i \mid z_1, \dots, z_m, w_1, \dots, w_r \Rightarrow F$

$J = m \cdot F \sim \chi^2_{m-p}$

③ Wu-Hausmann

H_0: $\hat{\beta}_{OLS}$ - consistent $\Rightarrow$ $\hat{\beta}_{OLS}$ - efficient

H_a: $\hat{\beta}_{OLS}$ - inconsistent $\Rightarrow$ $\hat{\beta}_{2SLS}$ - consistent k - # reg in 2step

$$(\hat{\beta}_{2SLS} - \hat{\beta}_{OLS})' \left(Var(\hat{\beta}_{2SLS}) - Var(\hat{\beta}_{OLS}) \right)^{-1} (\hat{\beta}_{2SLS} - \hat{\beta}_{OLS}) \overset{H_0}{\sim} \chi^2_k$$