2SLS (IV)

X - stochastic :

$\hat{\beta}_{OLS}$	consistent	$cov(X, \varepsilon) = 0$
	unbiased	$X \perp \varepsilon$

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

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

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

deterministic

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

stochastic

bias = 0 if $X \perp \varepsilon$

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

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

$E(\varepsilon \mid X) = 0 \Rightarrow E(\varepsilon) = 0$

$\Rightarrow cov(X,\varepsilon) = 0$

2SLS

X - endogenous reg.

z - instrumental variable → exogenous $cov(\varepsilon, z) = 0$

→ relevance $cov(x, z) \neq 0$

$$y = \beta_1 + \beta_2 \cdot X_i + \varepsilon_i$$

(1) $\hat{X}_i = \hat{\theta}_1 + \hat{\theta}_2 \cdot z_i$ $\qquad \hat{\theta}_2 = \dfrac{\widehat{cov}(x, z)}{\widehat{var}(z)}$

(2) $y_i = \hat{\beta}_1 + \hat{\beta}_2 \cdot \hat{X}_i + \hat{\varepsilon}_i$

$\updownarrow$

x

Problem 1.

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

$$= \frac{\cancel{\hat{\theta}_2}\, \widehat{cov}(z, y)}{(\hat{\theta}_2)^{\cancel{2}}\, \widehat{var}(z)} = \frac{\widehat{cov}(z, y)}{\frac{\widehat{cov}(z, x)}{\widehat{var}(z)} \cdot \widehat{var}(z)} = \frac{\widehat{cov}(z, y)}{\widehat{cov}(z, x)}$$

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

$$\operatorname*{plim}_{n \to \infty} \hat{\beta}_2^N = \operatorname*{plim}_{n \to \infty} \frac{\widehat{cov}(z, y)}{\widehat{cov}(z, x)} = \frac{cov(z_i, y_i)}{cov(z_i, x_i)} =$$

$$= \frac{cov(z_i, \beta_1 + \beta_2 X_i + \varepsilon_i)}{cov(z_i, X_i)} = \beta_2 + \frac{\overbrace{cov(z_i, \varepsilon_i)}^{= 0 \text{ (exogeneity)}}}{\underbrace{cov(z_i, X_i)}_{\neq 0 \text{ (relevance)}}}$$

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

$$se(\hat{\beta}_2^{IV}) = \sqrt{\underbrace{\frac{s^2}{\Sigma (x_i - \bar{x})^2}}_{var(\hat{\beta}_2^{OLS})} \cdot \underbrace{\frac{1}{\hat{corr}^2(x, z)}}_{R^2 \text{ from 1st step}}}$$

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

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

2SLS (multinomial reg)

$x_1, \dots, x_p$ – endogenous vars

$w_1, \dots, w_2$ – exogenous var

$z_1, \dots, z_m$ – instruments

1 $x_1 \mid z_1, \dots, z_m, w_1, \dots, w_2 \Rightarrow \hat{x}_1$

$\vdots$

$x_p \mid z_1, \dots, z_m, w_1, \dots, w_2 \Rightarrow \hat{x}_p$

2 $y \mid \hat{x}_1, \dots, \hat{x}_p, w_1, \dots, w_2$

$m < p$ under identified $\hat{x}_1 = \hat{\alpha}_1 + \hat{\alpha}_2 \cdot z_i$

$\hat{x}_2 = \hat{\beta}_1 + \hat{\beta}_2 \cdot z_i$

$\Rightarrow$ on 2 step perf.mu

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

$m > p$ over identified (2SLS)

Test : 1) relevance

2) exogeneity (Sargan's test)

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

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

H_0: z - weak instruments $F < 10$

H_a: z - strong instruments $F \geq 10$

2 Sargan test $(m > p)$

H_0 : z - exogenous

H_a : z - endogenous

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

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

3 Hausman test $\hat{\theta}, \forall \tilde{\theta} \in C_{LUE}$

$Var(\hat{\theta}) \leq Var(\tilde{\theta})$

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

H_a: $\hat{\beta}_{OLS}$ - inconsistent $\Rightarrow$ $\hat{\beta}_{IV}$ - consistent

k - # regr. in 2step

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