$y_1 = Y_1\beta_1 + Z_1\gamma_1 + \varepsilon_1 = X_1\delta_1 + \varepsilon_1$

IV estimator: $\hat{\delta}_1 = (W_1' X_1)^{-1} W_1' y_1$

$W_1 = Z A$

2SLS: $\hat{W}_1 = Z\hat{A} = Z(Z'Z)^{-1}Z'X_1 = P_Z X_1$
$= [P_Z Y_1 \;\; P_Z Z_1] = [\hat{Y}_1 \;\; Z_1] = \hat{X}_1$

2SLS estimator: $\hat{\delta}_1 = (X_1' P_Z X_1)^{-1} X_1' P_Z y_1$
$= (\hat{X}_1' X_1)^{-1} \hat{X}_1' y_1$
$= (\hat{X}_1' \hat{X}_1)^{-1} \hat{X}_1' y_1$

Prove: $0 = \text{plim}(\hat{\delta}_1 - \delta_1) = \text{plim}(\hat{X}_1' X_1)^{-1} \hat{X}_1' \varepsilon_1$

(1) $\text{plim}\left(\frac{1}{T}\hat{A}' Z' [Y_1 \;\; Z_1]\right)$

assume $\text{plim}\frac{Z'Z}{T} = M$

$\text{plim}\; \hat{A} = A = (\Pi_1 \;\; I_1)$

(1) $= \text{plim}\frac{1}{T}\hat{A}' Z'(Z\Pi_1 + V_1) = A' M \Pi_1$

(2) $= \text{plim}\frac{1}{T}\hat{X}_1'\varepsilon_1 = \text{plim}\frac{1}{T}\hat{A}' Z'\varepsilon_1 = \text{plim}\,\hat{A}'\; \underbrace{\text{plim}\frac{Z'\varepsilon_1}{T}}_{=0}$

$\Rightarrow 0 = \text{plim}(\hat{\delta}_1 - \delta_1)$

- IV in general is biased, since you need to estimate A.
- In general, these estimators need not have moments
  - 2SLS has the same # of moments as degrees of overidentification.

Normality

$$\sqrt{T}(\hat{\delta}_1 - \delta_1) = \underbrace{\left(\frac{\hat{A}' Z'}{T}[Y_1 \; Z_1]\right)^{-1}}_{\xrightarrow{P} R^{-1}} \underbrace{\frac{\hat{A}' Z' \varepsilon_1}{\sqrt{T}}}_{(1)}$$

$$(1) = \frac{1}{\sqrt{T}} \hat{A}' Z' \varepsilon_1 = \hat{A}' \underbrace{\frac{Z' \varepsilon_1}{\sqrt{T}}}_{(2)}$$ need unid CLT: Lindberg-Feller or Lyapunov $\rightarrow 2+\delta, \delta > 0$ moments must exist

$$(2) \; \frac{Z' \varepsilon_1}{\sqrt{T}} \xrightarrow{d} N\left(0, \sigma_{11} \operatorname{plim} \frac{Z'Z}{T}\right) \stackrel{d}{=} N(0, \sigma_{11} M)$$

$$\Rightarrow \frac{\hat{A}' Z' \varepsilon_1}{\sqrt{T}} \xrightarrow{d} N(0, \sigma_{11} A' M A)$$

$$\Rightarrow \sqrt{T}(\hat{\delta}_1 - \delta_1) \xrightarrow{d} N(0, \sigma_{11} R^{-1} \underbrace{A' M A}_{N} (R^{-1})')$$

Let $\hat{A} = [\hat{\Pi}_1 \; I_1] \Rightarrow \operatorname{plim} \hat{A} = [\Pi_1 \; I_1]$

$$R = (\Pi_1' Z' Z \Pi_1 \quad M_{11})$$

$$N = [\Pi_1' Z' Z \Pi_1 \quad I_1 Z' Z I_1]$$

$$= [\Pi_1' Z' Z \Pi_1 \quad M_{11}] = R$$

$$\Rightarrow \sqrt{T}(\hat{\delta}_1 - \delta_1) \xrightarrow{d} N(0, \sigma_{11} R^{-1} R R^{-1})$$

$$\stackrel{d}{=} N(0, \sigma_{11} R^{-1})$$

Need to estimate $\hat{\sigma}_{11} = \frac{1}{T-(R_1+S_1)}(y_1 - Z\hat{\delta}_1)'(y_1 - Z\hat{\delta}_1)$

- This will be a downward biased estimate of σ_{11} if you actually do the two stages.

Minimum distance or minimum χ^2 interpretation

OLS: $\min_{\delta} (y_1 - X_1\delta_1)' I (y_1 - X_1\delta_1)$ is inconsistent

2SLS: $\min_{\delta} (y_1 - X_1\delta_1)' \underbrace{P_Z}_{\text{eliminates the joint endogeneity}} (y_1 - X_1\delta_1)$ is consistent

- minimum distance interpretation
- transformation to orthogonality

$$\Leftrightarrow \min_{\delta} \frac{1}{\sigma_{11}} (y_1 - X_1\delta_1)' P_Z (y_1 - X_1\delta_1)$$

The solution: $\frac{\hat{\varepsilon}' P_Z \hat{\varepsilon}}{\sigma_{11}} \overset{A}{\sim} \chi^2 \underbrace{(K-K_1)}_{\text{degree of overidentification}}$ $K_1 = r_1 + s_1$

- This is a test of overidentifying restrictions.
  "$E[\chi^2(K-K_1)] = K-K_1$"

How to prove this is χ^2?

$\frac{1}{\sigma_{11}} \varepsilon_1' B \varepsilon_1$ $B = P_Z - P_Z X_1 (X_1' P_Z X_1)^{-1} X_1' P_Z$
$= P_Z - P_{\hat{X}_1}$ $\hat{X}_1 = P_Z X_1 \subset \text{span } Z$

B is symmetric and idempotent

$\Rightarrow$ Cochrane's thm: $\frac{\varepsilon_1' B \varepsilon_1}{\sigma_{11}} \to \chi^2(tr(B))$

$\hat{\varepsilon}_1 = Z\alpha + w$ (*)

$F: \alpha = 0$ $\sim \chi^2(K-K_1)$

$\parallel$

TR^2 where R^2 is from the regression (*)

- Suppose this rejects. What instruments should you not have used? (Hopefully things will still be just identified.)

Prove that 2SLS is optimal among class of IV estimators. (under conditional homoskedasticity.)

$\hat{W}_1$ here uses $\hat{A} = (Z'Z)^{-1}Z'X_1$

$\hat{W}_1 = Z'\hat{A}$

Consider $\hat{\hat{W}}_1 = Z'D$

$$Var(\hat{\delta}_1) = \sigma_{11}\left(\frac{\hat{W}_1'X_1}{T}\right)^{-1}\left(\frac{\hat{W}_1'\hat{W}_1}{T}\right)\left(\frac{\hat{W}_1'X_1}{T}\right)^{-1}$$

- Set $T\sigma_{11} = 1$ (or else change units)

$$Var(\hat{\delta}_1)^{-1} = (\hat{W}_1 X_1')(\hat{W}_1'\hat{W}_1)^{-1}(\hat{W}_1'X_1) = (X_1'ZD)(D'Z'ZD)^{-1}(D'Z'X_1)$$

Choose any $g \neq 0$ $\quad g'[Var(\hat{\delta}_{2SLS})^{-1} - Var(\hat{\delta}_D)^{-1}]g \geq 0$

$$g'(X_1'ZD)(D'Z'ZD)^{-1}(D'Z'X_1)g$$

Let $Z'Z = NN'$, $h = N^{-1}Z'X_1 g$

$$\Rightarrow h'h = g'X_1'Z(N^{-1})'Z'ZX_1 g$$
$$= g'X_1'Z(Z'Z)^{-1}Z'X_1 g$$
$$= g'(V(\hat{\delta}_{2S}))^{-1}g$$

$$g'(X_1'ZD)(D'Z'ZD)^{-1}(D'Z'X_1)g = h'ND(D'N'ND)^{-1}D'N'h$$
$$= h'G(G'G)^{-1}G'h \quad \text{where } G = ND$$

Thus, comparison: $h'\underbrace{[I - G(G'G)^{-1}G']}_{=M_G}h \geq 0$

$h'h \geq h'P_G h$

$$\Rightarrow h'M_G M_G' h = (M_G'h)'(M_G'h) \geq 0$$

Thus, $Var(\hat{\delta}_{2SLS}) \leq Var(\hat{\hat{\delta}}_D)$

- Could nonlinear combinations of our instruments do better?

Third interpretation of 2SLS.

$$\text{plim } \frac{Y_1' \varepsilon_1}{T} \neq 0 \qquad \hat{Y}_1 = Z\hat{\Pi}_1$$

$$\hat{\Pi}_1 = (Z'Z)^{-1} Z' Y_1$$

$$= (Z'Z)^{-1} Z'(Z\Pi + v_1)$$

$$= \Pi + (Z'Z)^{-1} Z' v_1$$

$$\Rightarrow E[\hat{\Pi}_1] = \Pi + E[(Z'Z)^{-1} Z' v_1]$$

We know $\text{plim } \hat{\Pi}_1 = \Pi_1$

$\hat{Y}_1 = Y_1 - \hat{v}_1 \Leftrightarrow Y_1 = \hat{Y}_1 + \hat{v}_1$

$$y_1 = \hat{Y}_1 \beta_1 + Z_1 \gamma_1 + \varepsilon_1 + \hat{v}_1 \beta_1$$

OLS on this is 2SLS.

Need to have $\hat{v}_1$ orthogonal to Z_1

(it will be orthogonal to $\hat{Y}_1$ by construction.)

$$y_1 = \hat{X}_1' \delta_1 + \underbrace{\varepsilon_1 + \hat{v}_1 \beta_1}_{w_1}$$

$$\Rightarrow \hat{\delta}_1 = (\hat{X}_1' \hat{X}_1)^{-1} \hat{X}_1 y_1 = (X_1' P_Z X_1)^{-1} X_1' P_Z y_1$$

$$= \hat{\delta}_{1,IV}$$

Test of over ID:

$$y_1 = \hat{X}_1 \delta_1 + \underbrace{\tilde{Z}_1}_{Z \setminus Z_1} \mu + w$$

$F: \mu = 0$ $K - K_1$ degrees of freedom