No class Oct. 9

1] If all equations are just identified, then

$2SLS \equiv 3SLS \equiv ILS \equiv FIML \equiv ILS$

2] If equation 1 is overid and all other equations are just id, then

- $\hat{\delta}_{1,2SLS} \equiv \hat{\delta}_{1,3SLS}$

$$\left.\begin{array}{l} y_1 = y_2\beta_1 + Z_1 + Z_1\gamma_1 + u_1 \\ y_2 = Z\Pi_1 + v_1 \end{array}\right\}$$ Nothing to be gained by doing 3SLS here for 1st equation.

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & I\otimes Z \end{bmatrix}\begin{bmatrix} \delta_1 \\ \Pi_2 \end{bmatrix} + \begin{bmatrix} u_1 \\ v_2 \end{bmatrix} \qquad Var\begin{pmatrix} u_1 \\ v_2 \end{pmatrix} = \begin{bmatrix} \sigma_{11} & \Psi_{12} \\ & \underbrace{\Omega}_{(M-1)\times(M-1)} \end{bmatrix} \otimes I_T$$

$$X_1 = [y_2 \quad Z_1] \qquad \delta_1 = \begin{bmatrix} \beta_1 \\ \gamma_1 \end{bmatrix}$$

3SLS:

$$\begin{bmatrix} \hat{\delta}_1 \\ \hat{\Pi}_2 \end{bmatrix} = \begin{bmatrix} X_1'(\Psi^{11}\otimes P_Z)X_1 & X_1'(\Psi^{12}\otimes Z) \\ & \Psi^{22}\otimes Z'Z \end{bmatrix}^{-1} \cdot$$

$$\begin{bmatrix} X_1(\Psi^{11}\otimes P_Z)y_1 + X_1'(\Psi^{12}\otimes P_Z)y_2 \\ (\Psi^{21}\otimes Z')y_1 + (\Psi^{22}\otimes Z')y_2 \end{bmatrix}$$

Let $G_1 = \sigma^{11}\otimes(Z'Z)^{-1}$

$H = X_1'Z$

$Q = G - GH'[HGH']^{-1}HG'$

Inverse matrix:

$$\begin{bmatrix} (X_1'(\sigma^{11}\otimes P_Z)X_1)^{-1} & (X_1'(\sigma^{11}\otimes P_Z X_1)^{-1}X_1'(\sigma^{11}\Psi_{12}\otimes Z(Z'Z)^{-1}) \\ & \Omega_{22}\otimes(Z'Z)^{-1}-(\Psi_{21}\otimes I)\,Q\,(\Psi_{12}\otimes I) \end{bmatrix}$$

$$\sigma^{11} = \Psi^{11} + \sigma^{11}\Psi_{12}\Psi^{21}$$

$$\Psi^{11}\Psi_{12}\Omega_{22}^{-1} + \Psi^{12} = 0$$

$$\Rightarrow \hat{\delta}_{1,3SLS} = (X_1'(\sigma^{11}\otimes P_Z X)^{-1} X_1'(\sigma^{11}\otimes P_Z)y_1$$

$$= (X_1'(I\otimes P_Z)X_1)^{-1}X_1'(I\otimes P_Z)y_1$$

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

$$= \hat{\delta}_{2SLS}$$

$$YB + Z\Gamma = U \qquad U \sim N(0, \Sigma\otimes I)$$

$$\mathcal{L} = c + \frac{T}{2}\log[\det(\Sigma)]^{-1} - \frac{T}{2}\operatorname{tr}[\tfrac{1}{T}\Sigma^{-1}U'U]$$

log likelihood of multivariate normal

$$= c + \frac{T}{2}\log[\det\Sigma]^{-1} + T\log|\det B|$$
$$- \frac{T}{2}\operatorname{tr}[\tfrac{1}{T}\Sigma^{-1}(YB+Z\Gamma)'(YB+Z\Gamma)]$$

$$B = \frac{\partial U}{\partial Y}$$

$B = I$ if no joint endogeneity

(*) What happens if we have a triangular system?

$(\underset{\text{unknown}}{B^u}, \Gamma^u, \Sigma) \leftarrow$ maximize $\mathcal{L}$ wrt this

$$\frac{\partial \log \det(A)}{\partial A} = (A')^{-1}$$

wrt unknown elements

$$\text{(i)} \quad \frac{\partial h}{\partial B^u} : \Big[\overbrace{T(B')^{-1}}^{\text{Jacobian Term}} - Y'\underbrace{(YB + Z\Gamma)}_{\equiv U}\Sigma^{-1} \Big]^{u} = 0$$

$$\text{(ii)} \quad \frac{\partial h}{\partial \Gamma^u} : \Big[\qquad - Z'(YB + Z\Gamma)\Sigma^{-1} \Big]^{u} = 0$$

$$\text{(iii)} \quad \frac{\partial h}{\partial (\Sigma^{-1})} : \frac{1}{2}\Big[T\Sigma - (YB + Z\Gamma)'(YB + Z\Gamma) \Big]^{u} = 0$$

$$\Rightarrow \quad \Sigma = \frac{(YB + Z\Gamma)'(YB + Z\Gamma)}{T}$$

(i) becomes: $\Big[(B')^{-1}\Sigma\Sigma^{-1} - \frac{1}{T}Y'U\Sigma^{-1}\Big]^{u} = 0$

$$\Leftrightarrow \quad \Big[\Big\{(B')^{-1}\Sigma - \frac{1}{T}Y'U\Big\}\Sigma^{-1}\Big]^{u} = 0$$

It turns out $\operatorname{plim} \frac{Y'U}{T} = (B')^{-1}\Sigma$

$$\Rightarrow \quad T(B')^{-1}\Sigma = (B')^{-1}\underbrace{U'}_{YB+Z\Gamma}U = Y'U + \underbrace{(B')^{-1}\Gamma'Z'U}_{-\Pi'Z'U}$$

FOC: $\Big[(B')^{-1}\Gamma'Z'(YB + Z\Gamma)\Sigma^{-1}\Big]^{u} = 0$

Thus, we have:

(i) $\Big[(B')^{-1}\Gamma'Z'(YB + Z\Gamma)\Sigma^{-1}\Big]^{u} = 0$

(ii) $\Big[-Z'(YB + Z\Gamma)\Sigma^{-1}\Big]^{u} = 0$

Thus $\left[\begin{bmatrix} (B')^{-1}\Gamma' Z' \\ -Z' \end{bmatrix}(YB + Z\Gamma)\Sigma^{-1}\right]^u = 0$

$\beta_{ii} = 1$ gives us (using the stacked form $y = \tilde{X}\delta + u$)

$$z_j[(\Sigma^{-1})_i \otimes I_T]y = z_j'[(\Sigma^{-1})_i \otimes I_T]x_i\delta$$

and $\hat{y}_j[(\Sigma^{-1})_i \otimes I_T]y = \hat{y}_j[(\Sigma^{-1})_i \otimes I_T]x_i\delta$

where $\hat{y}_j = (-Z\hat{\Gamma}\hat{B}^{-1})_j$

Thus, the stacked FOCs for FIML become

$$\hat{\delta}_{FIML} = (\hat{W}'\tilde{X})^{-1}\hat{W}'y$$

where $\hat{W} = \hat{\tilde{X}}'(\hat{\Sigma} \otimes I)^{-1}$

$$\hat{\tilde{X}}_F = \begin{bmatrix} \hat{X}_1 & & 0 \\ & \ddots & \\ 0 & & \hat{X}_m \end{bmatrix}$$

and $\hat{X}_i = [-Z(\hat{\Gamma}\hat{B})^{-1}_i \quad Z_i]$

FIML is "internally consistent" by the invariance principle.

- Need to iterate on FIML

$$\sqrt{T}(\hat{\delta}_{FIML} - \delta) \xrightarrow{d} N(0, \operatorname{plim}[(\tfrac{1}{T}X'(\Sigma\otimes I)^{-1}X)^{-1}])$$

$$\text{plim}\sqrt{T}(\hat{\delta}_{FIML}-\hat{\delta}_{3SLS}) \equiv \text{plim}\sqrt{T}(\hat{\delta}_F-\hat{\delta}_3)$$
$$=\text{plim}\left(\frac{1}{T}\hat{\tilde{X}}_F'(\hat{\Sigma}_F\otimes I)^{-1}\tilde{X}\right)^{-1}\frac{1}{\sqrt{T}}\left(\hat{\tilde{X}}_F'(\hat{\Sigma}_F\otimes I)^{-1}u\right)$$
$$-\text{plim}\left(\frac{1}{T}\hat{\tilde{X}}_3'(\hat{\Sigma}_3\otimes I)^{-1}\tilde{X}\right)^{-1}\frac{1}{\sqrt{T}}\left(\hat{\tilde{X}}_3'(\hat{\Sigma}_3\otimes I)^{-1}u\right)$$

use the facts that

1] $\text{plim} -\hat{\Gamma}_F\hat{B}_F^{-1} = \text{plim}\,\hat{\Pi}_3 = \Pi$

2] $\text{plim}\ \hat{\Sigma}_F = \text{plim}\ \hat{\Sigma}_3 = \Sigma$

Thus,

$$\text{plim}\sqrt{T}(\hat{\delta}_F-\hat{\delta}_3) = \text{plim}\left(\frac{1}{T}\hat{\tilde{X}}_3(\hat{\Sigma}_3\otimes I)^{-1}\tilde{X}\right)^{-1}\cdot$$
$$\frac{1}{\sqrt{T}}\left[\hat{\tilde{X}}_F'(\hat{\Sigma}_F^{-1}\otimes I)-\hat{\tilde{X}}_3'(\hat{\Sigma}_3^{-1}\otimes I)\right]u$$

but $\text{plim}\dfrac{Z'u}{T}=0$

and $\sqrt{T}(\hat{\Pi}_3-\hat{\Pi}_F)$, $\sqrt{T}(\hat{\Sigma}_3-\hat{\Sigma}_F)$ both have limiting distributions and are thus bounded in probability.

$\Rightarrow \sqrt{T}(\hat{\delta}_F-\hat{\delta}_3) = o_p(1)\cdot O_p(1) = o_p(1)$

and $\sqrt{T}(\hat{\delta}_F-\hat{\delta}_3)\xrightarrow{p} 0$ and thus $\sqrt{T}(\hat{\delta}_F-\hat{\delta}_3)\xrightarrow{d} 0$,

so $\hat{\delta}_F$ and $\hat{\delta}_3$ are asymptotically equivalent.

Test of overid restrictions:

compare $\hat{\Pi}_R$ and $\underbrace{\hat{\Pi}_{UR}}_{}$,

least squares eqn by eqn

$$\hat{\Pi}_R = -\underbrace{\hat{\Gamma}\hat{B}^{-1}}_{\text{restricted estimates}}$$

$$LR = 2(\mathcal{L}_u - \mathcal{L}_R) \xrightarrow{p} \chi^2(\#\text{ overid})$$

alternatively, can test using

$$\hat{u}_3 = Z\alpha + W$$

check if $\alpha \neq 0$.

$$\text{Var}(W) = \Sigma \otimes I$$

quasi-maximum likelihood

- can do away with initial normality assumption

$$y_1 = X_1\delta_1 + u_1$$

$$\hat{y}_1 = Y_1 - \hat{v}_1$$

$$\hat{\delta}_1 = \begin{bmatrix} Y_1'Y_1 - \hat{v}_1'\hat{v}_1 & Y_1 Z_1 \\ & Z_1'Z_1 \end{bmatrix}^{-1} \begin{bmatrix} (Y - \hat{v}_1)y_1 \\ Z_1' y_1 \end{bmatrix}$$

Let $\hat{Y}_1 = Y_1 - K\hat{v}_1$

2SLS: $K=0$

OLS: $K=1$

$$W_1 = [\hat{Y}_1 \; Z_1]$$

Consistency $\Rightarrow K=1$

Need plim $k=1$ if k is stochastic

What if we want this to be as efficient as 2SLS?

Need $\sqrt{T}(k-1) \xrightarrow{d} 0$

Nagar: 2SLS is biased but consistent. Can do 2nd order Taylor expansion to get an unbiased estimator up to second order.

- $k = 1 + \frac{K - K_1}{T}$, K = # instruments, K_1 = # rhs variables

- $\text{plim}\sqrt{T}(1-k) = -\text{plim}\left(\frac{K-K_1}{\sqrt{T}}\right) = 0 \Rightarrow$ optimality

suppose $\frac{K}{T} \to \alpha$ (ie have many instruments)

bias doesn't go away Bekker EMA 1994

- many instruments problem.