

SUR MODEL Assumption / Estimation / special cases / testing  
(seemingly unrelated regressions)

in the previous lesson CAPM

$M$  = equations

$N$  = number of obs

$$\begin{cases} y_1 = x_1 \theta_1 + \epsilon_1 \\ y_2 = x_2 \theta_2 + \epsilon_2 \\ \vdots \end{cases}$$

$(N \times 1) \quad (N \times k_1) \quad (k_1 \times 1) \quad (N \times 1)$   
 $(N \times 1) \quad (N \times k_2) \quad (k_2 \times 1) \quad (N \times 1)$   
 $\vdots$

$k_2$  can be equal or not to  $k_1$   
and also  $k_1$  can be different from  $k_2$

$$y = X\theta + \epsilon$$

$$y = y_1 | y_2 | y_3 | \dots | y_M \quad N \times 1$$

$$\epsilon = \epsilon_1 | \dots | \epsilon_M \quad N \times 1$$

$$X = \begin{bmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_M \end{bmatrix}$$

$N \times \sum_{i=1}^M k_i$

on the main diagonal there are  
vectors, not scalars, hence this  
is a block matrix

$$\theta = [\theta_1 | \theta_2 | \dots | \theta_M]$$

$\sum_{i=1}^M k_i \times 1$

$$\theta_{OLS} = (X'X)^{-1} X'y$$

important assumption about the error terms

① Exogeneity  $E[\epsilon | X] = 0$

error term does not depend on any  
of the  $x$ s

$$\textcircled{2} \in [\varepsilon_{i,s} \cdot \varepsilon_{j,t} | x] = \textcircled{\delta_{ij}}$$

from 1 to  $M$ , they  
refer to the number of  
equations

1...N  
they refer to the  
observations

$$\delta_{ij} \neq 0 \text{ if } t=s$$

$$\delta_{ij} = 0 \text{ if } t \neq s$$

for the same individual the  
covariance across equations is  
positive, while for different  
individuals the observations

do not covary

↳ and the  $\delta_{ij}$  is also constant

$$N=3 \quad M=2$$

$$\varepsilon = \begin{bmatrix} \varepsilon_{1,1} \\ \varepsilon_{1,2} \\ \varepsilon_{1,3} \\ \varepsilon_{2,1} \\ \varepsilon_{2,2} \\ \varepsilon_{2,3} \end{bmatrix}$$

error first equation first individual

second equation, first individual

$$\Omega = \begin{matrix} (6 \times 6) \\ (NM \times NM) \end{matrix} \in [\varepsilon \varepsilon' | x] = \varepsilon \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{23} \end{bmatrix} \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} & \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \end{bmatrix} \bigg| x$$

$$= \mathbb{E} \begin{bmatrix} \epsilon_{11}^2 & \epsilon_{11} \epsilon_{12} & \epsilon_{11} \epsilon_{13} & \epsilon_{11} \epsilon_{21} & \epsilon_{11} \epsilon_{22} & \epsilon_{11} \epsilon_{23} \\ \epsilon_{12} \epsilon_{11} & \epsilon_{12}^2 & \epsilon_{12} \epsilon_{13} & \epsilon_{12} \epsilon_{21} & \epsilon_{12} \epsilon_{22} & \epsilon_{12} \epsilon_{23} \\ \epsilon_{13} \epsilon_{11} & \epsilon_{13} \epsilon_{12} & \epsilon_{13}^2 & \epsilon_{13} \epsilon_{21} & \epsilon_{13} \epsilon_{22} & \epsilon_{13} \epsilon_{23} \\ \epsilon_{21} \epsilon_{11} & \epsilon_{21} \epsilon_{12} & \epsilon_{21} \epsilon_{13} & \epsilon_{21}^2 & \epsilon_{21} \epsilon_{22} & \epsilon_{21} \epsilon_{23} \\ \epsilon_{22} \epsilon_{11} & \epsilon_{22} \epsilon_{12} & \epsilon_{22} \epsilon_{13} & \epsilon_{22} \epsilon_{21} & \epsilon_{22}^2 & \epsilon_{22} \epsilon_{23} \\ \epsilon_{23} \epsilon_{11} & \epsilon_{23} \epsilon_{12} & \epsilon_{23} \epsilon_{13} & \epsilon_{23} \epsilon_{21} & \epsilon_{23} \epsilon_{22} & \epsilon_{23}^2 \end{bmatrix}$$

→ Because two individuals use the same equation, we are working with homoskedasticity within each equation  
 → errors associated to the same individual between different equations

$$= \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_1^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_1^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_2^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_2^2 \end{bmatrix}$$

equation by equation  
 no homoskedasticity

Kronecker product

$$\Omega = \sum_{(M \times M)} \otimes I_N$$

Consider

$$A \otimes B = C = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1M}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{NM}B & \dots & \dots & a_{NM}B \end{bmatrix}$$

$M \times M$        $(N \times M)$        $(NM \times NM)$

not commutative

$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$

→ variance w/ the first equation, for N observations

→ variance for the same observations across equations

→ variance for the second equation

$$\Sigma \cdot I_3 = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \sigma_{12} & 0 & 0 \\ 0 & \sigma_1^2 & 0 & 0 & \sigma_{12} & 0 \\ 0 & 0 & \sigma_1^2 & 0 & 0 & \sigma_{12} \\ \hline \sigma_{12} & 0 & 0 & \sigma_2^2 & 0 & 0 \\ 0 & \sigma_{12} & 0 & 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_{12} & 0 & 0 & \sigma_2^2 \end{bmatrix}$$

$$y = X\beta + \varepsilon \quad E[\varepsilon\varepsilon' | X] = \Sigma \otimes I_N$$

OLS should not be a good model, it ignores the correlation

$$\hat{\beta}_{GLS} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} y$$

$$\hat{\beta}_{GLS} = (X' (\Sigma \otimes I_N) X)^{-1} X' (\Sigma \otimes I_N) y$$

(can take the inverse inside because  $\Sigma$  is invertible)

$$\hat{\beta}_{OLS} = (X' \hat{\Omega}^{-1} X)^{-1} X' \hat{\Omega}^{-1} y$$

$$\hat{\Omega} = \hat{\Sigma} \otimes I_N \quad \text{the predictor is exogenous } \Sigma$$

(N x N)  $\rightarrow$

1)  $\hat{\beta}_{OLS} = (X'X)^{-1}X'y$  exogenous equation by equation (limited information approach)

$$\hat{\epsilon} = y - X\hat{\beta}_{OLS}$$

(N x 1)

$$\hat{\epsilon} = \begin{bmatrix} \hat{\epsilon}_1 & \hat{\epsilon}_2 & \dots & \hat{\epsilon}_M \end{bmatrix}$$

(N x M) concatenated horizontally (residuals common)

$$\hat{\epsilon}'\hat{\epsilon} = \begin{bmatrix} \hat{\epsilon}_1' \\ \hat{\epsilon}_2' \\ \vdots \\ \hat{\epsilon}_M' \end{bmatrix} \begin{bmatrix} \hat{\epsilon}_1 & \hat{\epsilon}_2 & \dots & \hat{\epsilon}_M \end{bmatrix}$$

(1 x N) it is a row (N x M)

(M x N)

$$= \begin{bmatrix} \sum_{i=1}^N \hat{\epsilon}_{1i}^2 & \sum_{i=1}^N \hat{\epsilon}_{1i} \hat{\epsilon}_{2i} & \dots & \sum_{i=1}^N \hat{\epsilon}_{1i} \hat{\epsilon}_{Mi} \\ \sum_{i=1}^N \hat{\epsilon}_{2i} \hat{\epsilon}_{1i} & \sum_{i=1}^N \hat{\epsilon}_{2i}^2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^N \hat{\epsilon}_{Mi} \hat{\epsilon}_{1i} & \dots & \dots & \sum_{i=1}^N \hat{\epsilon}_{Mi}^2 \end{bmatrix}$$

$\downarrow$   
 variance errors last equation

$$\hat{\Sigma} = \frac{1}{N} \hat{E}' \hat{E} = \begin{bmatrix} \frac{1}{N} \sum_{i=1}^N \hat{\epsilon}_i & \frac{1}{N} \sum_{i=1}^N \hat{\epsilon}_i \hat{\epsilon}_i & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

plus  $\hat{\Sigma} = \Sigma$  it is an estimator of the variance covariance matrix term

$$\text{Var}(\hat{\beta}_{OLS}) = (X' \Omega^{-1} X)^{-1}$$

Two special cases where  $\hat{\beta}_{OLS} = \hat{\beta}_{GLS}$

KELNER CONDITIONS (do not need to be satisfied jointly)

↳ i can estimate as equation by equation in this case

1)  $\Sigma$  is diagonal  $\Omega = \Sigma \otimes I_N$

if  $\Sigma$  is diagonal the correlation across equations is zero

2) Same regressors in each equation

$$y = X\beta + \epsilon \quad \text{regressor } \beta_0$$

$$X = \begin{bmatrix} \beta_0 & 0 & \dots & 0 \\ \vdots & \beta_0 & \ddots & \vdots \\ 0 & \dots & \beta_0 \end{bmatrix}$$

$(N \times K)$

when they repeat it is useful to use the Kronecker product

$x_0$  is  $(n \times 1)$

$$I_n \otimes x_0 = X \quad \rightarrow \text{equivalent of scalar multiplied by the identity but for matrices}$$

$$\hat{\beta}_{OLS} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} y$$

$$= \left[ (I_n \otimes x_0)' (\Sigma^{-1} \otimes I_n) (I_n \otimes x_0) \right]^{-1} (I_n \otimes x_0)' (\Sigma^{-1} \otimes I_n) y$$

the transpose can go inside

$$(A \otimes B)(C \otimes D) = (AC \otimes BD)$$

$$(\Sigma^{-1} \otimes (x_0' x_0))^{-1} (\Sigma^{-1} \otimes x_0') y$$

$$(\Sigma \otimes (x_0' x_0)^{-1}) (\Sigma^{-1} \otimes x_0') y$$

$$[I_n \otimes (x_0' x_0)^{-1} x_0'] y$$

We are repeating equation by equation the OLS

$$\hat{\beta}_{OLS} = \begin{bmatrix} (x_0' x_0)^{-1} x_0' y_1 \\ \vdots \\ (x_0' x_0)^{-1} x_0' y_n \end{bmatrix}$$