SUR MODEL

(SEEMINGLY UNRELATED REGRESSIONS)

Basic Idea:
We have two (or more) linear models. We consider specific examples: like investment (sales, production) of two firms producing similar product (i.e. operating on the same market) explained by the same set of explanatory variables.

: or: demand (supply) for goods which are to some extent substitutes again regressed on the same set of explanatory variables.

Theoretically we can do 2 things.

1) estimate the two equations separately by OLS assuming they have nothing in common

2) estimate the two equations jointly by GLS assuming there is "something" creating interdependen: we model it by assuming errors in the 1st equation correlated with errors in the 2nd.

Potential gain:
- closer to reality
- more PRECISE PARAMETER ESTIMATES (i.e. lower variances of $\hat{\beta}$ by GLS)

As example we consider two firms producing similar goods (services) : General Electric and Westinghouse.

We regress their investment over 20 YEARS on CONSTANT (J) expected profits (firms value) (V) and CAPITAL (K).

THIS IS A TIME SERIES MODEL (observations accross 20 YEARS AS OPPOSED TO ACROSS INDIVIDUALS)

NOTATION:

WE WILL USE INDEX 1 for FIRM 1 (GENERAL EL)
INDEX 2 FIRM 2 (Westin)

WE WRITE 2 EQUATIONS

$$i_{1t} = \beta_{11} + V_{1t}\beta_{12} + K_{1t}\beta_{13} + e_{1t}$$

$$i_{2t} = \beta_{21} + V_{2t}\beta_{22} + K_{2t}\beta_{23} + e_{2t}$$

otherwise:

$$Y_1 = X_1\beta_1 + e_1$$
$$Y_2 = X_2\beta_2 + e_2$$

where

$$X_1 = \begin{bmatrix} j & v_1 & K_1 \end{bmatrix} = \begin{array}{ccc} \begin{bmatrix} x_{11} & x_{12} & x_{13} \end{bmatrix} \\ \text{reg 1 var 1} \quad \text{reg 1 var 2} \quad \text{reg 1 var 3} \end{array}$$

$$\beta_1 = \begin{bmatrix} \beta_{11} \\ \beta_{12} \\ \beta_{13} \end{bmatrix}$$

reg 1 coef 1, 2 and 3

$$Y_1 = i_1$$

$$X_2 = \begin{bmatrix} j & v_2 & K_2 \end{bmatrix} = \begin{bmatrix} x_{21} & x_{22} & x_{23} \end{bmatrix}$$

$$\beta_2 = \begin{bmatrix} \beta_{21} \\ \beta_{22} \\ \beta_{23} \end{bmatrix}$$

$$Y_2 = i_2$$

1) WE CAN ESTIMATE BOTH MODELS SEPARATELY BY OLS. THE RESULT OF THIS IS IDENTICAL TO GLS DESCRIBED BELOW:

GLS WILL BE APPLIED TO EQUATIONS 1 and 2 JOINTLY:

$$y_1 = x_{11}\beta_{11} + x_{12}\beta_{12} + x_{13}\beta_{13} + e_1$$
$$y_2 = x_{21}\beta_{21} + x_{22}\beta_{22} + x_{23}\beta_{23} + e_2$$

yielding:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & 0 & 0 & 0 \\ 0 & 0 & 0 & x_{21} & x_{22} & x_{23} \end{bmatrix} \begin{bmatrix} \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \beta_{21} \\ \beta_{22} \\ \beta_{23} \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

or otherwise:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ \vdots \\ y_{1T} \\ y_{21} \\ y_{22} \\ y_{23} \\ \vdots \\ y_{2T} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & & & \\ x_{21} & x_{22} & x_{23} & & 0 & \\ x_{31} & x_{32} & x_{33} & & & \\ \vdots & \vdots & \vdots & & & \\ x_{T1} & x_{T2} & x_{T3} & & & \\ & & & x_{11} & x_{12} & x_{13} \\ & & & x_{21} & x_{22} & x_{23} \\ & 0 & & x_{31} & x_{32} & x_{33} \\ & & & \vdots & \vdots & \vdots \\ & & & x_{T1} & x_{T2} & x_{T3} \end{bmatrix} \begin{bmatrix} \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \beta_{21} \\ \beta_{22} \\ \beta_{23} \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{1T} \\ e_{21} \\ e_{22} \\ e_{23} \\ \vdots \\ e_{2T} \end{bmatrix}$$

We assume .

→ DIFFERENT VARIANCES σ_1^2 and σ_2^2 FOR EQUATIONS

→ ERRORS IN EQUATION 1 ARE CONTEMPORANEOUSLY UNCORRELATED WITH ERRORS IN 2.

(intuition: no interdependencies between investment decisions of General and Westin)

no contem corr

$$W = E\left[\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \begin{bmatrix} e_1' & e_2' \end{bmatrix}\right] = \begin{bmatrix} E\, e_1 e_1' & E\, e_1 e_2' \\ E\, e_2 e_1' & E\, e_2 e_2' \end{bmatrix} = \begin{bmatrix} \sigma_1^2 I_T & 0 \\ 0 & \sigma_2^2 I_T \end{bmatrix}$$

More explicitely:

$$W = \left[\begin{array}{ccc:ccc} \sigma_1^2 & & 0 & & & \\ & \sigma_1^2 & & & 0 & \\ 0 & & \sigma_1^2 \ddots & & & \\ \hdashline & & & \sigma_2^2 & & 0 \\ & 0 & & & \sigma_2^2 \ddots & \\ & & & 0 & & \sigma_2^2 \\ \hdashline \end{array}\right]$$

reason for GLS: different variances accross the sample (pooled)

We can assume normality to have:

$$e \sim N\left(0,\ \begin{bmatrix} \sigma_1^2 I_T & 0 \\ 0 & \sigma_2^2 I_T \end{bmatrix}\right)$$

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

or

$$= \left(\begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}' \begin{bmatrix} \sigma_1^2 I_T & 0 \\ 0 & \sigma_2^2 I_T \end{bmatrix}^{-1} \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}\right)^{-1} \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}' \begin{bmatrix} \sigma_1^2 I_T & 0 \\ 0 & \sigma_2^2 I_T \end{bmatrix}^{-1} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sigma_1^2} X_1' X_1 & 0 \\ 0 & \frac{1}{\sigma_2^2} X_2' X_2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_1^2} X_1' y \\ \frac{1}{\sigma_2^2} X_2' y \end{bmatrix} = \begin{bmatrix} (X_1' X_1)^{-1} X_1' Y_1 \\ (X_2' X_2)^{-1} X_2' Y_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

where $\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = b$ is the OLS estimate.

THIS PROVES THE EQUIVALENCE OF 2 SEPARATE OLS TO ONE STEP GLS. (theoretically)

note also that in order to apply GLS we need to know the matrix W. INDEED WE HAVE TO ESTIMATE IT! TO DO THAT WE WOULD RUN 2 OLS REGESSIONS TO ESTIMATE σ_1^2 AND σ_2^2, AND AFTER USE $\hat{W}$ IN THE GLS STEP. HENCE: WHAT I CALL "ONE STEP GLS" refers to joint estimation of 2 models, but technically consists of two steps. $\hat{\hat{\beta}}$

WE NOW ALLOW FOR INTERDEPENDENCIES

interdependencies btw investment of 1 and 2 will be accomodated by CORRELATION IN ERROR TERMS

$$W = \begin{bmatrix} E e_1 e_1' & E e_1 e_2' \\ E e_2 e_1' & E e_2 e_2' \end{bmatrix} = \begin{bmatrix} \sigma_1^2 I_T & \sigma_{12} I_T \\ \sigma_{12} I_T & \sigma_2^2 I_T \end{bmatrix}$$

CONTEMPORANEOUS CORRELATION

$$W = \left[\begin{array}{ccc|ccc} \sigma_1^2 & & & \sigma_{12} & & \\ & \sigma_1^2 & & & \sigma_{12} & \\ & & \ddots & & & \ddots \\ & & \sigma_1^2 & & & \sigma_{12} \\ \hline \sigma_{12} & & & \sigma_2^2 & & \\ & \sigma_{12} & & & \sigma_2^2 & \\ & & \ddots & & & \ddots \\ & & \sigma_{12} & & & \sigma_2^2 \end{array}\right]$$

the "empty" fields are filled with zeros.

To see that consider the first row:

$$E(e_{11} e_{11}) \; E(e_{11} e_{21}) \ldots E(e_{11} e_{T1}) \; E(e_{11} e_{12}) \; (E\, e_{11} e_{22}) \ldots$$

$$\sigma_1^2 \qquad E(e_{t1} e_{t2}) = \sigma_{12} \qquad \sigma_{12} \qquad \cdots E(e_{11} e_{T2})$$

$$E(e_{t2} e_{t1}) = \sigma_{12}$$

Where σ_{12} IS THE COVARIANCE BETWEEN ERRORS IN EQUATIONS 1 and 2 HAVING THE SAME INDEX OF TIME t, HENCE BEING CONTEMPORANEOUS

WE NOW APPLY THE GLS ESTIMATOR

(cause variances unequal across sample)

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

being BEST UNBIASED. IT'S VARIANCE IS:

$$var(\hat{\beta}) = (X' W^{-1} X)^{-1}$$

BUT... IN PRACTISE W IS UNKNOWN AND HAS TO BE ESTIMATED:

$$\hat{\hat{\beta}} = (X' \hat{W}^{-1} X') X' \hat{W}^{-1} Y$$

$\hat{\hat{\beta}}$ is UNBIASED. IT IS THE ONLY FINITE SAMPLE PROPERTY WE CAN NOT SAY: IT IS MINIMUM VARIANCE, AS IT DEPENDS on $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$ and $\hat{\sigma}_{12}$ and ANALYTICALLY CAN NOT BE FOUND. IT IS NO LONGER A LINEAR FUNCTION OF Y.

BUT: $\hat{\hat{\beta}}$ is CONSISTENT AND HAS THE SAME ASYMPTOTIC DISTRIBUTION AS $\hat{\beta}$.

HENCE TESTS ARE ASYMPTOTICALLY VALID

We need to estimate

$$\hat{\sigma}_{ij} = (y_i - x_i b_i)'(y_j - x_j b_j)/T$$

where i and $j = 1$ or 2, and b are OLS estimators of individual equations separately.

It is an estimator with T denominator, biased in small samples but CONSISTENT

$(T-K)$ is no longer right as both regressions may have different # of regressors.

We have hence:

$$\hat{\sigma}_{11} = (y_1 - x_1 b_1)'(y_1 - x_1 b_1)/T \qquad \text{(variance in eq. 1)}$$

$$\hat{\sigma}_{22} = (y_2 - x_2 b_2)'(y_2 - x_2 b_2)/T \qquad \text{(variance in eq 2)}$$

$$\hat{\sigma}_{12} = \hat{\sigma}_{21} = (y_1 - x_1 b_1)'(y_2 - x_2 b_2)/T \qquad \text{(contemp. correlation)}$$

and

$$\hat{W} = \begin{bmatrix} \hat{\sigma}_{11} I_T & \hat{\sigma}_{12} I_T \\ \hat{\sigma}_{12} I_T & \hat{\sigma}_{22} I_T \end{bmatrix}$$

! note:

- $\hat{\sigma}_{11} = \dfrac{\sum_{t=1}^{T} \hat{e}_{1t}^2}{T}$
- $\hat{\sigma}_{12} = \dfrac{\sum_{t=1}^{T} \hat{e}_{1t}\hat{e}_{2t}}{T}$
- $\hat{\sigma}_{22} = \dfrac{\sum_{t=1}^{T} \hat{e}_{2t}^2}{T}$

we can use $\hat{W}^{-1}$ to estimate

$$\hat{var}(\hat{\hat{\beta}}) = \left(X' \hat{W}^{-1} X\right)^{-1}$$

$\hat{\beta}$ is asymptotically NORMALLY DISTRIBUTED WITH MEAN β and $var = var(\hat{\beta}) = \left(X'W^{-1}X\right)^{-1}$

THE ASYMPTIOTIC NORMALITY IS USED TO BUILT ASYMPTOTICALLY VALID TESTS.

IN this setup, the t statistics produced by the SAS prog. are in fact WALD tests of $H_0: \beta = 0$ for any coefficient

BEFORE WE CONTINUE ON THIS, A SIMPLE ASYMPTOTIC TEST OF NO CORRELATION BETWEEN REGESSION 1 and 2:

$$r_{12}^2 = \frac{(\hat{\sigma}_{12})^2}{\hat{\sigma}_1^2 \, \hat{\sigma}_2^2}$$

is the squared correlation

note: $\hat{\sigma}_1^2 = \hat{\sigma}_{11}$

$\hat{\sigma}_2^2 = \hat{\sigma}_{22}$ NOTATION!

$H_0: \hat{\sigma}_{12} = 0$

UNDER H_0 $\quad T r_{12}^2 \sim \chi^2_{(1)}$

x_c at 0,05 and 1 deg. of freedom is 3.84.
We compare $T r_{12}^2$ with 3.84 and if the stat $> x_c^2$ we reject Ho: no contemporaneous correlation

RESTRICTED ESTIMATION

claim: both equations have same values of coefficients, i.e. investment of Westin and General depends in exactly the same way on profits and capital.

bivariate SUR Model :

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$Y = X \beta + e$$

Assume non-diagonal variance matrix, i.e. presence of contemporaneous correlation btw 1 and 2.

IMPOSE RESTRICTION : $\beta_1 = \beta_2$, which is:

$$\begin{bmatrix} \beta_{11} \\ \beta_{12} \\ \beta_{13} \end{bmatrix} = \begin{bmatrix} \beta_{21} \\ \beta_{22} \\ \beta_{23} \end{bmatrix}$$

and can be written

$$R\beta = 0$$

$$R\beta = R\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0\\ 0 & 1 & 0 & 0 & -1 & 0\\ 0 & 0 & 1 & 0 & 0 & -1\end{bmatrix}\begin{bmatrix}\beta_{11}\\ \beta_{12}\\ \beta_{13}\\ \beta_{21}\\ \beta_{22}\\ \beta_{23}\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

yielding

$$\begin{aligned}\beta_{11} &= \beta_{21}\\ \beta_{12} &= \beta_{22}\\ \beta_{13} &= \beta_{23}\end{aligned} \quad \Rightarrow \text{ call it } \begin{bmatrix}\gamma_1\\ \gamma_2\\ \gamma_3\end{bmatrix} = \gamma$$

a common coeff. vector

and rewrite the model

$$y = \begin{bmatrix}y_1\\ y_2\end{bmatrix} = \begin{bmatrix}x_1\\ x_2\end{bmatrix}\gamma + \begin{bmatrix}e_1\\ e_2\end{bmatrix} = Z\gamma + e$$

We use the GLS SUR estimator to estimate $\hat{\gamma}$

$\hat{W}$ is estimated as before from OLS residuals.

Next we test the hypothesis : $R\beta = 0$

$$H_0 : R\beta = 0$$

by computing the statistics

$$q = R\hat{\hat{\beta}} \left[R \left(X' \hat{W}^{-1} X \right)^{-1} R' \right]^{-1} \hat{\hat{\beta}} R \Big/ J$$

IT IS AGAIN AN ASYMPTOTIC TEST WHICH RELIES ON THE ASYMPTOTIC NORMALITY OF $\hat{\hat{\beta}}$.

UNDER H_0: $q \sim F_{J,\, 2(T-J)}$, $J = K$, common # of regressors

and we reject H_0 when $q > F_c$.

Besides that, SAS produces t statistic (WALD tests) for $H_0 : \lambda = 0$ for every individual hypo:

$\beta_{11} = \beta_{21}$, $\beta_{12} = \beta_{22}$, $\beta_{13} = \beta_{23}$

SUMMARY:

1. STEPS TO FOLLOW

- estimate both models by OLS and keep $\hat{e}$

- find $\hat{\sigma}_{11}$, $\hat{\sigma}_{22}$, $\hat{\sigma}_{12}$
- test: $\hat{\sigma}_{12} = 0$
- if rejected, estimate SUR by GLS.

2. WHEN DOES OLS AND SUR GLS YIELD SAME PARAMETER ESTIMATES?

2 CASES

1) contemporaneous correlations are zero, i.e. $\sigma_{12} = 0$
2) both regressions have exactly same DATA FOR THE SAME REGRESSORS.

IN THESE CASES RATHE USE OLS because

$$\hat{var}(\hat{\beta}) = (X' \hat{W}^{-1} X')^{-1} \geq \hat{var}\, b = (X'X)^{-1} \sigma^2$$

PANEL DATA MODELS

1) $y_{it} = \alpha + \beta x_{it} + e_{it}$, $e_{it} \sim IIN(0, \sigma^2)$

$i = 1 \dots N = 100$

$t = 1 \dots T$

POOLED REGRESSION

USE OLS FOR STACKED OBS

$$\begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1T} \\ \hline y_{21} \\ y_{22} \\ \vdots \\ y_{2T} \\ \vdots \\ \text{etc} \end{bmatrix} = \alpha + \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1T} \\ \hline x_{21} \\ x_{22} \\ \vdots \\ x_{2T} \\ \vdots \\ \text{etc} \end{bmatrix} \beta + \begin{bmatrix} e_{11} \\ e_{12} \\ \vdots \\ e_{1T} \\ \hline e_{21} \\ e_{22} \\ \vdots \\ e_{2T} \\ \vdots \\ \text{etc} \end{bmatrix}$$

the properties of the <u>OLS</u> in small sample <u>and</u> asymptotically hold.

<u>Unusual</u> asymptotics: $n \to \infty$, T fixed

$T \to \infty$, n fixed

$n \to \infty$, $T \to \infty$

2) $y_{it} = \alpha_i + \beta x_{it} + e_{it}$ $\quad e_{it} \sim IIN(0, \sigma^2)$

FIXED EFFECT

USE OLS, the asymptotic properties of OLS

MAY NOT HOLD because # of parameters

is $n + 1$.,

if n fixed, $T \to \infty \Rightarrow$ OLS convergent

$n \to \infty$, T fixed $\Rightarrow$ OLS not convergent

$n \to \infty$, $T \to \infty$ $\Rightarrow$ OLS convergent.

regress $y_{it} - \bar{y}_i$ on $x_{it} - \bar{x}_i$

3) $y_{it} = \alpha + \beta x_{it} + \boxed{e_{it} + u_i}$

RANDOM EFFECT

$u_i \sim IIN(0, \eta^2)$

$e_{it} \sim IIN(0, \sigma^2)$

USE ~~R~~FGLS (~~[illegible]~~ FGLS