Seemingly Unrelated Regression (SUR): Inference and Testing

Indranil, Himanshu, Sunando, Binamrata

April 19, 2013

Introduction

A seemingly unrelated regression (SUR) system comprises several individual relationships that are linked by the fact that their disturbances are correlated. There are two main motivations for use of SUR. The first one is to gain efficiency in estimation by combining information on different equations. The second motivation is to impose and/or test restrictions that involve parameters in different equations.

The Model

Suppose there are N individuals and T observations on each individual.

The model for the ith individual can be written as

$$Y_{it} = \beta_{i1} + \beta_{i2}X_{i2t} + \beta_{i3}X_{i3t} + ... + \beta_{ik_i}X_{ik_it} + \epsilon_{it}$$

$$t = 1, 2, 3, ..., T$$

X_{ik_it} denotes the tth observation of the ith individual for the $k_i^{th}\beta$

If we stack all the T observations for the ith individual we get,

$$\begin{bmatrix} Y_{i1} \\ Y_{i2} \\ \vdots \\ Y_{iT} \end{bmatrix}_{T \times 1} = \begin{bmatrix} 1 & X_{i21} & X_{i31} \cdots X_{ik_i1} \\ 1 & & \\ \vdots & & \\ 1 & X_{i2T} & X_{i3T} \cdots X_{ik_iT} \end{bmatrix}_{T \times k_i} \begin{bmatrix} \beta_{i1} \\ \beta_{i2} \\ \vdots \\ \beta_{ik_i} \end{bmatrix}_{k_i \times 1} + \begin{bmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \vdots \\ \epsilon_{iT} \end{bmatrix}_{T \times 1}$$

$$Y_i = X_i\beta_i + \epsilon_i$$

$$i=1,2,3,....,N$$

If we look at a typical individual then it appears to be a usual regression problem; since each individual has his own $\beta_i s$.

Here we can have cross sectional as well as time series data.

The N regressions in this problem appear unrelated. However they may be related through the error terms ϵ_i and ϵ_j; where i and j are 2 individuals and ϵ_i and ϵ_j are vectors of order $T \times 1$.

Now suppose we stack on N individuals, we get,

$$
\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix}_{TN \times 1} = \begin{bmatrix} X_1 & & & \\ & X_2 & & \\ \vdots & & & \\ & & \cdots & X_N \end{bmatrix}_{TN \times \Sigma k_i} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{bmatrix}_{\Sigma k_i \times 1} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix}_{TN \times 1}
$$

The full model is given by

$$ Y = X\beta + \epsilon $$

Y has N subvectors each with T observations.

i.e.,

$$
\begin{aligned}
Y_1 &= X_1\beta_1 + \epsilon_1 \\
Y_2 &= X_2\beta_2 + \epsilon_2 \\
&\vdots \\
Y_N &= X_N\beta_N + \epsilon_N
\end{aligned}
$$

The model we consider can be written as

$$ I = \beta_1 + \beta_2 mcap + \beta_3 nfa + \beta_4 a $$

where,

I : Gross Investment at time 't'

$mcap$: Market capital at time 't-1'

nfa: Net Fixed Assets at time 't-1'

a: Current assets at time 't-1'

We collected data for three companies in the heavy automotive industry; Ashok Leyland, Mahindra & Mahindra and Tata Motors. The data has observations for each of these three companies from 1996 to 2012 (17 observations).

$$
\begin{bmatrix} I_{i1} \\ I_{i2} \\ \vdots \\ I_{i17} \end{bmatrix}_{17 \times 1} = \begin{bmatrix} 1 & X_{i12} & X_{i13} & X_{i14} \\ 1 & & & \\ \vdots & & & \\ 1 & X_{i172} & X_{i173} & X_{i174} \end{bmatrix}_{17 \times 4} \begin{bmatrix} \beta_{i1} \\ \beta_{i2} \\ \beta_{i3} \\ \beta_{i4} \end{bmatrix}_{4 \times 1} + \begin{bmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \vdots \\ \epsilon_{i17} \end{bmatrix}_{17 \times 1}
$$

i=Ashok Leyland(al), Mahindra & Mahindra(mm) and Tata Motors(tata)

Now stacking all 3 companies we get

$$
\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix}_{51\times1} = \begin{bmatrix} X_1 & & & \\ & X_2 & & \\ & & X_3 & \\ & & & X_4 \end{bmatrix}_{51\times4} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix}_{4\times1} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix}_{51\times1}
$$

Assumptions

I Regarding explanatory variables

$$
r\left(X_i\right) = k_i < T
$$
$$
r\left(X\right) = k
$$

Order of the X matrix is $TN \times \Sigma k_i$
$X'X$ is a positive definite matrix.
$(X'X)^{-1}$ exists.

II Regarding errors

1) Within individual i

$$
E(\epsilon_{it}) = 0 \ \forall t
$$
$$
E(\epsilon_{it}\epsilon_{is}) = \sigma_{ii} \text{ if t=s (same period)}
$$
$$
E(\epsilon_{it}\epsilon_{is}) = 0 \text{ if t} \neq \text{s (different period)}
$$

Homoskedasticity within an individual.
No direct autocorrelation.

$$
E(\epsilon_i) = 0
$$
$$
E(\epsilon_i\epsilon_i') = \sigma_{ii}I_T
$$

2) Across individuals

$$
E(\epsilon_{it}\epsilon_{js}) = \sigma_{ij} \text{ if t=s}
$$
$$
E(\epsilon_{it}\epsilon_{js}) = 0 \text{ if t} \neq \text{s}
$$

Heteroskedasticity across individuals.
There exists contemporaneous correlation. But there is no cross
autocorrelation.

$$
E(\epsilon_i\epsilon_j') = \sigma_{ij}I_T
$$

Now let us consider our full model,

$$Y = X\beta + \epsilon$$

$$E(\epsilon\epsilon') = E \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix}_{TN \times 1} \begin{bmatrix} \epsilon_1' & \epsilon_2' & \cdots & \epsilon_N' \end{bmatrix}_{1 \times TN}$$

$$E(\epsilon\epsilon') = E \begin{bmatrix} \epsilon_1\epsilon_1' & \epsilon_1\epsilon_2' \cdots & \epsilon_1\epsilon_N' \\ \vdots \\ \epsilon_N\epsilon_1' & \epsilon_N\epsilon_2' \cdots & \epsilon_N\epsilon_N' \end{bmatrix}_{TN \times TN}$$

$$E(\epsilon\epsilon') = \begin{bmatrix} \sigma_{11}I_T & \sigma_{12}I_T \cdots & \sigma_{1N}I_T \\ \sigma_{21}I_T & \sigma_{22}I_T \cdots & \sigma_{2N}I_T \\ \vdots \\ \sigma_{N1}I_T & \sigma_{N2}I_T \cdots & \sigma_{NN}I_T \end{bmatrix}$$

$$E(\epsilon\epsilon') = \begin{bmatrix} \sigma_{11} & \sigma_{12} \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} \cdots & \sigma_{2N} \\ \vdots \\ \sigma_{N1} & \sigma_{N2} \cdots & \sigma_{NN} \end{bmatrix}_{N \times N} \otimes I_T$$

$$E(\epsilon\epsilon') = \Sigma \otimes I_T$$

III Regarding relationship between X and ϵ.
X and ϵ are assumed to be independent.

ESTIMATION
$$Y = X\beta + \epsilon$$

$$E[\epsilon] = 0, E[\epsilon\epsilon'] = \Sigma \bigotimes I_T = V$$

Two cases
1. Simple case : $\sigma_{ij} = 0, \sigma_{ii} = \sigma^2$

$$V = \begin{bmatrix} \sigma^2 & 0 \cdots & 0 \\ 0 & \sigma^2 \cdots & 0 \\ 0 & 0 \cdots & \sigma^2 \end{bmatrix}_{N \times N} \otimes I_T$$

$$V = \sigma^2[I_N] \otimes [I_T] = \sigma^2 I_{NT} \implies CLRM$$

OLS is efficient

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'Y \Rightarrow \hat{\beta}_{i_{OLS}} = (X_i'X_i)^{-1}X_i'Y_i$$

So OLS on the full model boils down to OLS on individual models (which we know is BLUE) $\Rightarrow$ Pooling does not give us any extra efficiency.

2. Pure GLS Estimator

Σ is free implying presence of contemporeneous correlation.

$$\hat{\beta}_{GLS} = [X'\Sigma^{-1} \bigotimes I_T^{-1}X]^{-1}X'\Sigma^{-1} \bigotimes I_T^{-1}Y$$

Here all the coefficients of all the individuals are estimated together $\nRightarrow$ model boils down to individual estimator

Inference

We consider two cases:

1. Simple case $(\sigma_{ij} = 0)$:

Since, for this case, the seemingly unrelated regression structure can be estimated by applying the OLS method of estimation on each equation separately , we can either use the PROC REG statement or the PROC SYSLIN data=sasuser.filename SDIAG SUR statement for estimation. SDIAG imposes the restriction $\sigma_{ij} = 0$ on the error variance-covariance matrix of the SUR model.

We estimate the equation as :

$AshokLeyland : \widehat{i_a} = $ -2648.66 + 0.07$mcap\_$a + 0.14$nfa\_$a + 0.11$a\_$a

$Mahindra\&Mahindra : \widehat{i_m} = $ -15385 + 0.12$mcap\_$m + 0.97$nfa\_$m + 0.88$a\_$m

$TataMotors : \widehat{i_t} = $ -55189 - 0.18$mcap\_$t + 2.25$nfa\_$t + 1.13$a\_$t

For Mahindra & Mahindra, keeping the other explanatory variables constant, a 1 unit increase in $mcap\_$m at time 't-1' results in an average increase of 0.1156 units in $i\_$m at 't'.

Similarly, a 1 unit increase in $nfa\_$m at 't-1' results in an average increase of 0.9741 units in $i\_$m and a 1 unit increase in $a\_$m at 't-1' results in an average increase of 0.8826 units in $i\_$m at 't'.

From P-values, we can see that at 10% level of significance, the estimate of the $mcap\_$m and $a\_$m coefficients are significant.

2. General case (σ_{ij} is free):

In this case, we need to use the GLS method of estimation since the error variance-covariance matrix of the SUR model is not equal to $\sigma^2 I$. We use the the PROC SYSLIN data=sasuser.filename SUR statement for this estimation.

Ashok Leyland: $\widehat{i\_a}$ = -1630.7 + 0.10mcap_a + 0.21nfa_a â 0.065a_a

$Mahindra\&Mahindra : \widehat{i\_m}$ = -14236.2 + 0.126mcap_m + 1.16nfa_m + 0.67a_m

$TataMotors : \widehat{i\_t}$ = -50187.1 - 0.13mcap_t + 2.1nfa_t + 0.96a_t

Now, for Mahindra & Mahindra, keeping the other explanatory variables constant, a 1 unit increase in $mcap\_$m at 't-1' results in an average increase of 0.0283 units in $i\_$m at 't'.

Similarly, a 1 unit increase in $nfa\_$m at 't-1' results in an average increase of 0.6056 units in $i\_$m and a 1 unit increase in $a\_$m at 't-1' results in an average increase of 0.4313 units in $i\_$m at 't'.Note that the results here are different from the equation obtained by OLS method.

From P-values, we can see that at 10% level of significance, the estimate of the $mcap\_$m and $nfa\_$m coefficients are significant.

Testing in SUR Models

A natural question to ask in this framework is equality of behaviour.

$\qquad$ i.e., for $Y_i = X_i\beta + \epsilon_i$ $\qquad\qquad\qquad\qquad\qquad\qquad$ $i = 1, \ldots, N$

Null Hypothesis - $H_0 : \beta_1 = \beta_2 \ldots = \beta_N (k_i = k \ \ \forall i)$

Case 1: Simple Case

$\sigma_{ii} = \sigma^2 \qquad\qquad \forall i,$
$\sigma_{ij} = 0 \qquad\qquad \forall i,$

$consider, \ \ N = 2 \qquad\qquad (Two \ \ individuals)$

$$Y_1 = X_1\beta + \epsilon_1 \qquad (T_1 > K)$$
$$Y_2 = X_2\beta + \epsilon_2 \qquad (T_2 > K)$$
$$H_0 : \beta_1 = \beta_2$$

NOTE: This test is applicable when 1,2 refer to the same individual at two different points in time. It gives natural justification for $\sigma_{i,j} = 0$.

How to test ?
Use the concept of constrained and unconstrained models.

Unconstrained Model:

$$Y_1 = X_1\beta_1 + \epsilon_1$$
$$Y_2 = X_2\beta_2 + \epsilon_2 \qquad \text{estimate by OLS,}$$
$$SS_{un} = SS_1 + SS_2 \qquad \text{Degree of freedom is } (T_1 + T_2 - 2k)$$

Constrained Model:

$$Y_1 = X_1\beta + \epsilon_1 \qquad where, \beta_1 = \beta_2 = \beta$$
$$Y_2 = X_2\beta + \epsilon_2$$
we get SS_c by OLS and DOF is $(T_1 + T_2 - 2k)$

Then we use F statistic,

$$F_{K,T_1+T_2-2K} = \frac{SS_c - SS_{un}/K}{SS_{un}/T_1 + T_2 - 2K}$$

Our Ananlysis

In our case e consider the data for Ashok Leyland and Mahindra Mahindra, each with observations from 1996 to 2012.

The number of explanatory variables must be equal in the two equations. However, the number of observations may vary across the two firms. In this test, we use 3 explanatory variables and $T_1 = T_2 = 17$ observations for each firm.

For testing equality of behavior, we can test the following null hypothesis, $H_0 : \beta_1 = \beta_2$ versus the alternate hypothesis, $H_1 : \beta_1 \neq \beta_2$.

Unconstrained Model:

$$
\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix}
$$

Or,

$I = X\beta + \epsilon$

This structure is estimated by OLS method.

$SS_{al} = 33246407$; $SS_{mm} = 566598063.4$

$SS_{uc} = SS_{al} + SS_{mm} = 599844470$

$DOF_{uc} = T_1 + T_2 - K - K = 26$

Constrained model :

$$
\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \begin{bmatrix} X_1 & \\ & X_2 \end{bmatrix} \begin{bmatrix} \beta \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix}
$$

Or,
$I = X\beta + \epsilon$

This structure is estimated by OLS method in one shot.

$SS_c = 1934598017$

$DOF_c = T_1 + T_2 - K = 30$

Number of Restrictions $= DOF_c - DOF_{uc} = 4$

$F_{cal} = [(SS_c - SS_{uc})/\text{number of restrictions}]/ \ [SS_{uc}/DOF_{uc}] \frown F_{4,26} = 14.4636$

The F_{tab} value at 5% LOS is 2.74

Decision Criteria : We reject H_0 when $F_{cal} \succ F_{tab}$

Therefore, we reject H_0 at 5% LOS.

Not all the coefficients in the two coefficient matrices are equal.

Case 2: Simple Case but N Individuals

$H_0 : \beta_1 = \beta_2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots = \beta_N$

Unconstrained Model:

$SS_{un} = SS_1 + \ldots + SS_N$
$DOF = \sum_{i=1}^{N} T_i - NK$

Constrained Model:

$Y_1 = X_1\beta + \epsilon_1$
$Y_2 = X_2\beta + \epsilon_2$
$\vdots$
$Y_N = X_N\beta + \epsilon_N$

$$F_{(N-1)K, \sum_{i=1}^{N} T_i - NK} = \frac{SS_c - SS_{un}/(N-1)K}{SS_{un}/\sum_{i=1}^{N} T_i - NK}$$

Our Analysis

In our regression Model

Unconstrained model:

$SS_{al} = 33246407$; $SS_{mm} = 566598063.4$; $SS_{tata} = 13445921889$

$SS_{uc} = SS_{al} + SS_{mm} + SS_{tata} = 14045766359$

$DOF_{uc} = T_1 + T_2 + T_3 - K - K - K = 39$

Constrained Model:

$SS_c = 35161183397$

$DOF_c = T_1 + T_2 + T_3 - K = 47$

Number of restrictions $= DOF_c - DOF_{uc} = 8$

$F_{cal} = [(\text{SS}_c \text{ - } SS_{uc})/\text{number of restrictions}]/ \ [SS_{uc}/DOF_{uc}] \frown F_{8,39} = 7.329$

The F_{tab} value at 5% LOS is 2.18

Decision Criteria : We reject H_0 when $F_{cal} \succ F_{tab}$

Therefore, we reject H_0 at 5% LOS.

Not all the coefficients in the two coefficient matrices are equal.

Case 3: Two Individuals but Insufficient Degree of Freedom in the Second Period

Unconstrained Model:

$$Y_1 = X_1\beta_1 + \epsilon_1 \qquad (T_1 > K)$$
$$Y_2 = X_2\beta_2 + \epsilon_2 \qquad (T_2 < k)$$

$$X = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix}$$

$$r(X_1) = K, r(X_2) < K$$

therefore $r(X) < 2K, (X^T X)^{-1}$ cant be computed.

So we can estimate model-1 but cant estimate model-2.

Constrained Model:

$$Y_1 = X_1\beta + \epsilon_1$$
$$Y_2 = X_2\beta + \epsilon_2$$

we get SS_c by OLS and $dof = T_1 + T_2 - K$.

also we get F statistic by,

$$F_{T_2, T_1 - K} = \frac{SS_c - SS_{un}/T_2}{SS_{un}/(T_1 - K)}$$

Our analysis

We consider the data for Ashok Leyland and Mahindra & Mahindra, Ashok Leyland having observations from 1996 to 2012 and Mahindra & Mahindra having observations from 1996 to 1997.

The number of explanatory variables must be equal in the two equations. However, the number of observations are varying across the two firms. In this test, we use 3 explanatory variables, $T_1 = 17$ and $T_2 = 2$.

For testing equality of behavior, we can test the following null hypothesis, $H_0 : \beta_1 = \beta_2$ versus the alternate hypothesis, $H_1 : \beta_1 \neq \beta_2$.

Unconstrained Model:

Or,

I = Xβ + ϵ

This structure is estimated by OLS method.

$SS_{al} = 33246407$

$SS_{uc} = SS_{al} = 33246407$

$DOF_{uc} =$ T$_1 - K = 13$

Constrained model:

Or,
I = Xβ + ϵ

This structure is estimated by OLS method in one shot.
$SS_c = 56881219.77$

$DOF_c = T_1 + T_2 - K = 15$

Number of restrictions $= DOF_c - DOF_{uc} = T_2 = 2$

$F_{cal} = [(SS_c - SS_{uc})/\text{number of restrictions}]/ [SS_{uc}/DOF_{uc}] \backsim F_{2,13} = 4.6208$

The F_{tab} value at 5% LOS is 3.81

Decision Criteria : We reject H_0 when $F_{cal} \succ F_{tab}$

Therefore, we reject H_0 at 5% LOS.

Not all the coefficients in the two coefficient matrices are equal.

Case 4: Partial Test

$Y_1 = X_{11}\beta_{11} + X_{12}\beta_{12} + \epsilon_1$......(1)
$Y_2 = X_{21}\beta_{21} + X_{22}\beta_{22} + \epsilon_2$......(2)

Order of matrices:
$X_{11} = [T_1 \times (K_1 - S)]$
$X_{12} = (T_1 \times S)$

$X_{21} = [T_2 \times (K_2 - S)]$
$X_{22} = (T_2 \times S)$
$\beta_{11} = [(K_1 - S) \times 1]$
$\beta_{12} = (S \times 1)$
$\beta_{21} = [(K_2 - S) \times 1]$
$\beta_{22} = (S \times 1)$
$H_0 : \beta_{12} = \beta_{22}$

Unconstrained Model:

$SS_{un} = SS_1 + SS_2$
$DOF = T_1 + T_2 - K_1 - K_2$

Constrained Model:

$\beta_{12} = \beta_{22} = \gamma (say)$

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} X_{11} & 0 & X_{12} \\ 0 & X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} \beta_{11} \\ \beta_{21} \\ \gamma \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix}$$

Degrees of Freedom$= (T_1 + T_2) - (K_1 - S + K_2 - S + S)$

$F = [(SS_C - SS_{un})/S]/[SS_{un}/T_1 + T_2 - K_1 - K_2]$

Case 5: The General Case

$K_1 = K_2 = ... = K_N = K$
$T_1 = T_2 = ... = T_N = T$

Unconstrained Model:

It is obtained by applying FGLS on the full model.
DOF=TN-NK=N(T-K)

Constrained Model:

DOF=TN-K

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} \beta + \epsilon$$

$$F_{(N-1)K,(T-K)N} = [(SS_C - SS_{un})/(N-1)K]/[SS_{un}/(T-K)N]$$

$$\tilde{SS} \neq \hat{\epsilon}'\hat{\epsilon}$$

$$\tilde{SS} = \hat{\epsilon}'V^{-1}\hat{\epsilon}$$
$$= \hat{\epsilon}'\left[\hat{\Sigma} \otimes I_T\right]^{-1}\hat{\epsilon}$$