The Two Feasible Seemingly Unrelated Regression Estimator

Ghazal. A. Ghazal, Salwa. A. Hegazy

ABSTRACT: This paper presents the review for the Seemingly Unrelated Regression Equation (SUR) or system model involved several individual relationships that related by their disturbances terms are correlated. This model appears in the beginning of the sixties of the last century, initially introduced by Zellner (1962, 1963), and followed many author .the model has been applied in many fields such as econometric, analysis of economic phenomenon such as, investment, income, consumption, demand functions. We review comprehensively the estimators for (SUR) equation models whereas there are many author write about this model and some properties of the estimators.

Keyword: Seemingly Unrelated Regression Equation, zellner's estimators,(SUR) equations, Two Feasible(SUR) Estimator.

1. INTRODUCTION

Zellner 1962 proposed method of estimating coefficient in encountered sets of regression equations .the Aitken Generalized Least Square (GLS) method to estimate the parameter of sets of system equation regression models, he developed it's estimators for coefficient in linear multivariate system that is popularly known as the (SUR) Equation Model, explored that. The regression coefficient estimators by using Aitken least square procedure method are shown to have smaller variance than the least squares estimator (LSE) The organization of the present paper is the following In section (2) we are going to introduce literature review of the seemingly unrelated regression equation models, in section (3) introduce a details explanation of the specification of model and assumption, in section (4) estimating the coefficient regression of zellner's model , in section (5) we discuss The finite sample moment of the coefficient estimators , with study the some properties of zellner's estimators and present his results by using two equation or more, finally in section (6) present concluding remarks of this chapter

2 LITERATURE REVIEW OF THE

(SUR) EQUATION MODELS

Zellner (1962) In the original paper, proposed a procedure of estimating coefficient in certain generally encountered sets of regression equations.

- *Prof Ghazal. A. Ghazal*
- *Professor of Statistics & Econometrics, Department of Applied Statistics and Econometrics, Institute of Statistical Studies & Research – Cairo University – Giza – Egypt. More than 45 years teaching experience, University of North Carolina and Cairo University Number of published papers specialized Applies Statistics and Econometrics*
- *Salwa. A. Hegazy.*
- *Applied Statistic and Econometrics Department, Institute of Statistical Studies & Research - Cairo University – Giza – Egypt. 14 years teaching experience, of 6 October University, Faculty of Economics and Management – Giza –Egypt* Salwahegazy2@yahoo.com

This procedure is Aitken Generalized Least Square (GLS) method to estimate the parameter of sets of system equation regression models, he developed it's estimators for coefficient in linear multivariate system that is popularly known as the Seemingly Unrelated Regression (SUR) Equation Model, explored that. The regression coefficient estimators by using Aitken least square procedure method are shown to have smaller variance than the least squares estimator (LSE). It is clear that regression coefficient estimator which obtained by applying Aitken least square procedure is more efficient than those obtained by using Ordinary Least Square (OLS) method to estimate coefficient for each equation separately as single equation. The gain occurs when contemporaneous disturbance terms in different regression equations are correlated and when the different sets of independent (exogenous) variables appear in equations of the system, it can be occur when applying Aitken least square procedure to the all set of equations in the system conjunction, and estimate the variance covariance matrix of disturbance terms from single equation least squares residuals. Zellner(1963) In subsequent paper in the following year, consider that system of equation is two equation ,suggested a new approach of estimating the coefficient of a (SURE) model and develops an operational version of (GLS) to estimate sample variance and covariance parameters .where these estimators are based on unrestricted estimate S by replacing their unknown population by the unrestricted sample estimator, this of estimator comes from make regression of endogenous variable (Y) on all exogenous variable (X's).Zellner explored the finite sample properties of these estimators with consider the orthogonal condition of explanatory variables and disturbance term have a normal distribution , provided the finite sample properties of estimators by using two system equation regression models. These estimators are based on the unrestricted estimate S , derived the exact finite sample first and second moment of the two stage Aitken coefficient estimator, this estimation under the orthogonal condition and disturbances are normally distribution, compared the results with the least squared estimator, showed that Aitken (GLS) approach yields minimum variance unbiased linear estimators (MVUL)when the disturbance covariance matrix is known .the estimators are more efficient than single equation least square estimator of any sample size when disturbance are correlated. N.C. Kakwani (1967) studied the zellner's estimators when the case of a multi- regression system equation, that proposed in original paper (1962) under general condition, and explored that the regression coefficient estimators are unbiased even without the orthogonally condition requirement. When the disturbances terms have normally distributed Richard.W.Parks (1967) considered the problem of obtaining efficient estimates for the parameters of a system of regression equations. When the number of equations is greater than two equations .The disturbance terms of this system are assumed to be related by contemporaneous correlation in the different equations under the assumption he developed an estimator that is consistent and has the same normal distribution as the two least square Aitken estimators which assumed the covariance matrix to be known. Jan Kmenta and Roy.F.Gilbert (1968) describe the small sample behavior of these estimators in a number of specific cases by means of a Mont Carlo experiment., discussed the small sample properties of five alternative estimators of a set of linear regression equations when correlated disturbances. two of this procedures are least square and Aitken two stage least square, examined this estimators with the different sample size , studied properties of this estimators and compared the results of five procedures , thus he explored that in many cases zellner's estimators by using Aitken two stage least squares are better than the other estimators when replacing the variance covariance matrix of disturbances terms Σ by its consistent estimate $\left(\hat{\Sigma} = S\right)$ that based on single equation least squares residuals , in zellner's two stage procedure the estimate of the regression coefficient can be used for calculating a new set of residuals leading to obtain a new estimate of variance covariance matrix Σ which can be used for obtaining a new estimates of regression coefficient. Revankar (1974) consider a system of two (SUR) equations model and examined some finite sample properties of coefficient estimators that based on the variance of residuals ,derived the exact variance covariance matrix of zellner's estimator with the normality distribution and additional condition , that where the explanatory variable in the second equation is subset of the explanatory variable of first equation ,Revankar showed that the coefficient estimator based on the unrestricted estimator of variance covariance of disturbances term .that depend upon the residuals from the (OLS) regression by applying regression of dependent variables on all explanatory variables explored that the coefficient estimators are best linear unbiased estimator (BLUE) he derived the exact moments of zellner's estimators and show that the efficiency in increase rapidly as the sample size increases J.S.Mehta and P.A.Swamy (1976) Considered a system of two (SUR) equations. Without imposing any additional restrictions on the matrices of observations of the explanatory variables when deal with the model and derived the exact second moments of Zellner's estimator., Explored that the coefficient estimator is more efficient than the corresponding single equation least-squares in the presence of high contemporaneous correlation coefficient between the random disturbances terms in the two equations, derived the exact second order moments of zellner's estimator and made comparison with those of the single equation lest square estimator to explore efficiency of the estimator that these derived by zellner and revankar Dwivedi & Srivastava (1978) has shown that separate modeling of such equations yield less efficient estimators .than the estimators yield from using SUR approach that considers a joint modeling where disturbance terms are correlated in a system of multi equation regression estimation , The (SUR) procedure was based on the Generalized Least Squares (GLS) approach. , mentioned that Zellner(1962) has shown the gain in efficiency of (SUR) approach was achieved in his Two-Stage Least Squares (2SLS) approach instead of using Ordinary Least Squares(OLS) .The efficiency would be achieve when contemporaneous correlation between the disturbances is high and explanatory variables in different equations are uncorrelated. Aiyi Liu (2002) derived simpler expressions under a certain structure of design matrices for the two-stage Aitken estimates of the regression coefficients of two (SUR) equations model The estimates are shown to have smaller variance than the ordinary least squares estimates under the assumption of independently distributed, studied finite sample efficiency of the two stage estimate, compared the results with (OLS) estimate and compare their dispersion matrices. This simpler version allows studying the performance of such estimates in the model Noted that there are many authors written about seemingly unrelated regression equation models and using it in analysis the different phenomena of econometrics such as Verbon (1980) use (SUR) model to estimate a set of four labor demand equations, Beier-lein et al. (1981) applied (SUR) model to study estimate six equations describing the demand for electricity and natural gas , Denis Conniffe (1982) used model with time series data when study the demand system , Brown et al. (1983) estimate the size of stock returns , in the next year Howrey and Varian (1984) estimated a system of demand equations for electricity, Sickles (1985) used model in the technology and specific factor productivity growth in the US airline industry ,Wan etal. (1992) estimated production functions for some Agricultural crops in regions of China, Egger and Pfaffermayr (2004) used the model in industry when study bilateral outward foreign direct investment in different countries, A. Raknerud et al. (2006) used sure model with time series data when study the demand system,Tiefeng Ma, Rendao Ye (2010)

3 SPECIFICATION THE MODEL AND IT'S ASSUMPTION

In (1962) Zellner suggest a set of equations regression system and proposed the model contain N equations. Consider the following equations model

$$Y_i = X_i B_i = U_i \qquad\qquad i = 1,2,\ldots\ldots N$$

Zellner considered the problem of estimating the regression coefficients efficiently, so proposed applying Aitken's generalized least-squares method to estimate the parameters of a set of regression equations to the whole system of (N) equations in conjunction. Under conditions of disturbance terms are normality distribution and explanatory variable are orthogonal , found that the regression coefficient estimators so obtained are more efficient than those obtained by an using equation-by-equation application of least squares procedure. This gain in efficiency can be quite large if "independent" variables in different equations are not correlated and if disturbance terms in different equations are correlated. In (1963) zellner studied the finite sample properties by using two regression equations model and using two stage least square procedure to estimate the regression coefficient. The basic model

$$\underset{(T.1)}{\boldsymbol{Y_i}} = \underset{(T.k_i)}{\boldsymbol{X_i}}\ \underset{(K_i.1)}{\boldsymbol{B_i}} + \underset{(T.1)}{\boldsymbol{U_i}} \qquad i=1,2 \quad (1)$$

$$K = k_1 + k_2$$

This can be written as

$$\begin{pmatrix} Y_1 \\ \hline Y_2 \end{pmatrix}_{(T.1)} = \left(\begin{array}{c|c} X_1 & 0 \\ \hline 0 & X_2 \end{array}\right)_{(T.K)} \begin{pmatrix} \mathrm{B}_1 \\ \hline \mathrm{B}_2 \end{pmatrix}_{(K.1)} + \begin{pmatrix} U_1 \\ \hline U_2 \end{pmatrix}_{(T.1)} \quad (2)$$

Where

y_i is a $(\mathrm{T}.1)$ vector of observations on the $i-th$ dependent variable X_i is a $(T.\ k_i)$ block diagonal matrix of observation on (K_i) nonstochastic independent variable, each column of which consists of T observation on a regressor (explanatory variable) in the $i-th$ equation of the model, with rank (K_i) .that (x_1) , (x_2) are the matrices of fixed elements , B_i is a $(k_i.1)$ vector of regression coefficients that unknown parameters in the $i-th$ equation of the model., u_i is the corresponding $(\mathrm{T}.1)$ vector of random disturbances term in the $i-th$ regression equation.

Assumption (1)

The disturbance terms have zero mean vectors that is

$$\mathrm{E}(u_1) = \mathrm{E}(u_2) = 0_{(T.1)} \quad (3)$$

And variance covariance matrix

$$E\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\begin{pmatrix} u_1' & u_2' \end{pmatrix} = \begin{pmatrix} \sigma_{11}I & \sigma_{12}I \\ \sigma_{21}I & \sigma_{22}I \end{pmatrix} \quad (4)$$

$$= \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \otimes I = (\Sigma \otimes I)_{(T.T)}$$

Where:

σ_{ii} is scalar and represents the variance of the random disturbance in the $i-th$ equation for each observation in the sample, I_T is $(T.T)$ identity or unit matrix of order T , and $E(.)$ denotes the usual expectation operation , $\Sigma = (\sigma_{ij})$, $i, j = 1,2$ represents the covariance between the disturbances of $i-th$ equations and $j-th$ equations for each observation in the sample, (Σ) is known as variance covariance matrix , $(\otimes)$ denotes the usually kronecker product .

Assumption (2)
The contemporaneous disturbances of the two equation are not linearly dependent, so (Σ) is positive definite matrix and contemporaneous element of (u_i, u_j) , $i, j = 1,2$ have a normal distribution

Assumption (3)
Assume the following orthogonal condition that

$$x_1'x_2 = x_2'x_1 = 0 \quad (5)$$

4 ESTIMATING THE COEFFICIENT OF (SUR) MODEL

In this section we present the zellner's estimators results by conceder the previous assumption, using the Aitken's generalized least-squares (GLS) method to the system of equation in (2) we get the following coefficient estimators.

$$\widehat{\beta}_{GLS} = \begin{pmatrix} b_1^* \\ \hline b_2^* \end{pmatrix} = (X'(\Sigma^{-1} \otimes I)X)^{-1}(X'(\Sigma^{-1} \otimes I)y) \quad (6)$$

Applying the form in (6) ,Then we have

$$\widehat{\beta}_{GLS} = \begin{pmatrix} b_1^* \\ \hline b_2^* \end{pmatrix} = \left[\begin{pmatrix} x_1' & 0 \\ 0 & x_2' \end{pmatrix}\begin{pmatrix} \sigma^{11} & \sigma^{12} \\ \sigma^{21} & \sigma^{22} \end{pmatrix}\begin{pmatrix} x_1 & 0 \\ 0 & x_2 \end{pmatrix}\right]^{-1}\begin{pmatrix} x_1' & 0 \\ 0 & x_2' \end{pmatrix}\begin{pmatrix} \sigma^{11} & \sigma^{12} \\ \sigma^{21} & \sigma^{22} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$= \begin{pmatrix} \sigma^{22}x_1'x_1 & -\sigma^{12}x_1'x_2 \\ -\sigma^{12}x_2'x_1 & \sigma^{11}x_2'x_2 \end{pmatrix}_{(2.2)}^{-1} \begin{pmatrix} \sigma^{22}x_1'y_1 - \sigma^{12}x_1'y_2 \\ -\sigma^{21}x_2'y_1 + \sigma^{11}x_2'y_2 \end{pmatrix}_{(2.1)}$$

Where

$$\sigma^{11} = \sigma_{22}(\sigma_{11}\sigma_{22} - \sigma_{12}^2)^{-1},$$
$$\sigma^{22} = \sigma_{11}(\sigma_{11}\sigma_{22} - \sigma_{12}^2)^{-1}, \quad (7)$$
$$\sigma^{21} = \sigma^{12} = -\sigma_{12}(\sigma_{11}\sigma_{22} - \sigma_{12}^2)^{-1},$$

Zellner used the orthogonal condition that explanatory variables in two equations are independent
Then we have

$$\hat{\beta}_{GLS}=\begin{pmatrix} b_1^* \\ b_2^* \end{pmatrix}=\begin{pmatrix} (x_1'x_1)^{-1}x_1'y_1-\frac{\sigma_{12}}{\sigma_{22}}(x_1'x_1)^{-1}x_1'y_2 \\ (x_2'x_2)^{-1}x_2'y_2-\frac{\sigma_{12}}{\sigma_{11}}(x_1'x_1)^{-1}x_2'y_1 \end{pmatrix} \quad (8)$$

$\hat{\beta}_{GLS}$ estimator is not feasible estimator .where Σ that positive definite matrix is unobservable, it is unknown parameters., Zellner develops an operational version of (GLS) method that Feasible Aitken GLS. Where proposed an estimator of β in the (SUR) equation model by replacing the unknown population parameter Σ by its consistent estimate $(\hat{\Sigma})$ it is an observable matrix (S) ., $\sigma_{12}\neq 0$,when replacement Σ by S in (6), we have Feasible Aitken Generalized Least Square (FGLS) estimator of β Zellner ,Revankar

$$\hat{\beta}_{FGLS}=\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}=[X'(S^{-1}\otimes I_T)X)^{-1}X'(S^{-1}\otimes I_T)Y \quad (9)$$

Where, S_{ij} is the some estimator of σ_{ij} , $i,j=1,2$., $S_{ij}=\frac{1}{T}\hat{u}_i'\hat{u}_j$ ($\hat{u}_{(T.1)}$) is the vector of ordinary least-squares residual (LSR) of i-th equation obtained from applying (OLS) method, Then

$$S=\left(\frac{1}{T}\right)(\hat{u}_1' \,\vdots\, \hat{u}_2')\begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \end{pmatrix}$$

$$=\begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} \quad (10)$$

Where

(S) is the disturbance variance covariance matrix $,(s_{ij})$ is non-singular matrix $,(s_{ij}^{'s})$ are estimators of $\sigma_{ij}^{'s}$.

From (7),(8) ,(10) we have

$$\hat{\beta}_{FGLS}=\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}=\begin{pmatrix} s^{22}x_1'x_1 & -s^{12}x_1'x_2 \\ -s^{12}x_2'x_1 & s^{11}x_2'x_2 \end{pmatrix}^{-1}\begin{pmatrix} s^{22}x_1'y_1-s^{12}x_1'y_2 \\ -s^{21}x_2'y_1+s^{11}x_2'y_2 \end{pmatrix}$$

Where

$$(X'(S^{-1}\otimes I_T)X)=\left[\begin{pmatrix} x_1' & 0 \\ 0 & x_2' \end{pmatrix}\left(\begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}^{-1}\otimes I_T\right)\begin{pmatrix} x_1 & 0 \\ 0 & x_2 \end{pmatrix}\right]$$

$$=\frac{1}{(s_{11}s_{22}-s_{12}^2)}\begin{pmatrix} s_{22}x_1'x_1 & -s_{12}x_1'x_2 \\ -s_{12}x_2'x_1 & s_{11}x_2'x_2 \end{pmatrix}$$

And

$$[X'(S^{-1}\otimes I_T)Y)=\frac{1}{(s_{11}s_{22}-s_{12}^2)}\begin{pmatrix} s_{22}x_1'y_1 & -s_{12}x_1'y_2 \\ -s_{12}x_2'y_1 & s_{11}x_2'y_2 \end{pmatrix}$$

Then by applying the inverse of partition matrix to get the estimators of coefficient we get

$$\hat{\beta}_{FGLS}=\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}=\begin{pmatrix} (x_1'x_1)^{-1}x_1'y_1-\frac{s_{12}}{s_{22}}(x_1'x_1)^{-1}x_1'y_2 \\ (x_2'x_2)^{-1}x_2'y_2-\frac{s_{12}}{s_{11}}(x_1'x_1)^{-1}x_2'y_1 \end{pmatrix} \quad (11)$$

Where

$\hat{\beta}_{FGLS}$ is the Feasible Generalized Least Square (FGLS) estimator of β

5 FINITE SAMPLE MOMENTS OF COEFFICIENT ESTIMATORS

In this section, showed the exact first and second moment of the two-stage estimators, compare the result of the single equation least squares estimators with those derived from the Aitken procedure equation. To determine the moment of two coefficient b_1 *and* b_2 .Let us consider the system of two equations as a follows

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}=\begin{pmatrix} x_1 & x_2 & 0 & 0 \\ 0 & 0 & x_1 & x_2 \end{pmatrix}\begin{pmatrix} B_1 \\ B_{10} \\ B_2 \\ B_{20} \end{pmatrix}+\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad (12)$$

The sets of independent (exogenous) variables appearing in each equation as the same as (12) is a usual regression model which has as maximum likelihood estimators (MLE) for the coefficient vectors., If we put the parameters $B_{10}=0$ *and* $B_{20}=0$ it is the seem as the reduce to the system equation in (2),

Then we get

$$\begin{pmatrix} B_1 \\ B_{10} \\ B_2 \\ B_{20} \end{pmatrix}=\begin{pmatrix} (x_1'x_1)^{-1}x_1'y_1 \\ (x_2'x_2)^{-1}x_2'y_1 \\ (x_1'x_1)^{-1}x_1'y_2 \\ (x_2'x_2)^{-1}x_2'y_2 \end{pmatrix} \quad (13)$$

The coefficients estimators (B_1,B_{10},B_2,B_{20}) in (13) are normal distributed and distributed independently of the variance covariance matrix $T(\hat{\Sigma})=T\{s_{ij}\}$ the elements of estimate the variance covariance matrix (S) have the Wishart distribution $W(\Sigma,N)$,Then

$S_{ij}\sim W(\Sigma,N)$ Where

$N=T-K$, $(K=k_1+k_2)$

The elements of the disturbance variance covariance matrix are

$S_{ij}=\frac{1}{T}\hat{u}_i'\hat{u}_j$,

$(\hat{u}_i)$ is a $(T.1)$ vector of residual

From (11) we get

$$\widehat{\beta}_{FGLS} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} \hat{B}_1 - \left(\frac{s_{12}}{s_{22}}\right)\hat{B}_{20} \\ \hat{B}_2 - \left(\frac{s_{12}}{s_{11}}\right)\hat{B}_{10} \end{pmatrix}$$

5-1 UNBIASEDNESS ESTIMATOR

We can write the estimator b_1 as follows

$$b_1 = \hat{B}_1 - v_1\hat{B}_{20},$$

$$(v_1) = \left(\frac{s_{12}}{s_{22}}\right)$$

take the expectation of both sides. We get $E(b_1) = E(\hat{B}_1) - E(v_1).E(\hat{B}_{20})$,Since (v_1) and $\hat{B}_{20}$ are independently distributed and the expectation of coefficient estimator $\hat{B}_{20}$ equal zero as $E(\hat{B}_{20}) = 0_{(K.1)}$, Then we have $E(b_1) = B_1$,Thus (b_1) is unbiased estimator of (B_1) , Similarly from the second equation

Since $\left(\frac{s_{12}}{s_{11}}\right)$ and $\hat{B}_{10}$ are independently distributed and the expectation of coefficient estimator $\hat{B}_{10}$ equal zero Thus (b_2) is unbiased estimator of (B_2)

5-2 VARIANCE COVARIANCE MATRIX OFCOEFFICIENT ESTIMATORS

To find the variance covariance matrix for $\hat{B}_1$ and $\hat{B}_{20}$ let consider

$$(\hat{B}_1 - B_1) = (x_1'x_1)^{-1}x_1'y_1 - B_1,$$

$$= (x_1'x_1)^{-1}x_1'x_1B_1 + (x_1'x_1)^{-1}x_1'u_1 - B_1$$

$$= (x_1'x_1)^{-1}x_1'u_1$$

$$\operatorname{var}(\hat{B}_1) = E(\hat{B}_1 - B_1)(\hat{B}_1 - B_1)',$$

$$= E((x_1'x_1)^{-1}x_1'u_1)((x_1'x_1)^{-1}x_1'u_1)'$$

$$= (x_1'x_1)^{-1}x'E(u_1u_1')x_1(x_1'x_1)^{-1}$$

$$\operatorname{var}(\hat{B}_1) = \sigma_{11}(x_1'x_1)^{-1} \quad (14)$$

Similarly we can find the expectation of $(\hat{B}_{20}\hat{B}_{20}')$
Then we get

$$E(\hat{B}_{20}\hat{B}_{20}') = \sigma_{22}(x_1'x_1)^{-1} \quad (15)$$

And the expectation of $((\hat{B}_1 - B_1)\hat{B}_{20}')$
Then we get

$$E((\hat{B}_1 - B_1)\hat{B}_{20}') = \sigma_{12}(x_1'x_1)^{-1} \quad (16)$$

Thus, the variance covariance matrix for $\hat{B}_1$ and $\hat{B}_{20}$ is given by

$$E\begin{pmatrix} (\hat{B}_1 - B_1) \\ \hat{B}_{20} \end{pmatrix}\left[(\hat{B}_1 - B_1)' \;\vdots\; \hat{B}_{20}'\right]$$

$$= \begin{pmatrix} \sigma_{11}(x_1'x_1)^{-1} & \sigma_{12}(x_1'x_1)^{-1} \\ \sigma_{12}(x_1'x_1)^{-1} & \sigma_{22}(x_1'x_1)^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \otimes (x_1'x_1)^{-1} = \Sigma \otimes (x_1'x_1)^{-1}$$

Then

$$E(b_1 - B_1)(b_1 - B_1)' = E\left[\hat{B}_1 - B_1 - \frac{s_{12}}{s_{22}}\hat{B}_{20}\right]\left[\hat{B}_1 - B_1 - \frac{s_{12}}{s_{22}}\hat{B}_{20}\right]'$$

$$= E\left[(\hat{B}_1 - B_1)(\hat{B}_1 - B_1)' - 2\frac{s_{12}}{s_{22}}(\hat{B}_1 - B_1)\hat{B}_{20}' + \frac{s_{12}}{s_{22}}\hat{B}_{20}\hat{B}_{20}'\right]$$

And

$$= E(\hat{B}_1 - B_1)(\hat{B}_1 - B_1)' - 2E\left(\frac{s_{12}}{s_{22}}\right)E(\hat{B}_1 - B_1)\hat{B}_{20}' + E\left(\frac{s_{12}}{s_{22}}\right)E(\hat{B}_{20}\hat{B}_{20}')$$

From (14),(15),(16) we get

$$= \sigma_{11}(x_1'x_1)\left(1 - 2E\left(\frac{s_{12}}{s_{22}}\right)\frac{\sigma_{12}}{\sigma_{11}} + E\left(\frac{s_{12}}{s_{22}}\right)^2\frac{\sigma_{22}}{\sigma_{11}}\right) \quad (17)$$

Then we get

$$E(b_1 - B_1)(b_1 - B_1)' = \sigma_{11}(x_1'x_1)\left(1 - 2E(v)\frac{\sigma_{12}}{\sigma_{11}} + E(v^2)\frac{\sigma_{22}}{\sigma_{11}}\right),$$

In order to get the variance .it must be get the expectation of $(v),(v^2)$, then Zellner studied the probability density function of the estimator and show that

$(v) = \left(\frac{s_{12}}{s_{22}}\right)$ has the following the probability density function (Pdf) given by

$$h(v)dv = \left[\frac{\sigma_{22}}{\sigma_{11}(1-\ell^2)}\right]^{\frac{1}{2}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)} \cdot \frac{dv}{\left[1 + \frac{\left(v - \ell\sqrt{\frac{\sigma_{11}}{\sigma_{22}}}\right)^2}{\frac{\sigma_{11}(1-\ell^2)}{\sigma_{22}}}\right]^{\left(\frac{n+1}{2}\right)}}$$

$$-\infty < v < \infty$$

Where

$$\ell = \frac{\sigma_{22}}{\sqrt{\sigma_{11}\sigma_{22}}}, \quad (18)$$

(Γ) denote the Gamma function
See Zellner A [1963, equation (2-16),P,981]
By using Wishart distribution to find the expectation of

(v) *and* (v^2)

obtained the following results

$$E(v)=E\left(\frac{s_{12}}{s_{22}}\right)=\ell\sqrt{\frac{\sigma_{11}}{\sigma_{22}}} \quad (19)$$

Then

$$E(v^2)=E\left(\frac{s_{12}}{s_{22}}\right)^2=\left[\frac{1}{2}\ \frac{\sigma_{11}}{\sigma_{22}}(1-\ell)^2 \frac{\Gamma(\frac{n+1}{2})\ \Gamma(\frac{n-2}{2})}{\Gamma(\frac{1}{2})\left[\Gamma(\frac{n}{2})\right]^2}+\frac{\sigma_{11}}{\sigma_{22}}\ell^2\right] \quad (20)$$

From (17) we are going to find the second terms then we have

$$\left(1-2E\left(\frac{s_{12}}{s_{22}}\right)\frac{\sigma_{12}}{\sigma_{11}}+E\left(\frac{s_{12}}{s_{22}}\right)^2\frac{\sigma_{22}}{\sigma_{11}}\right)$$

Substitute (19), (20) in (16)we get

$$=\left(1-2\ell\frac{\sigma_{11}}{\sigma_{22}}\sqrt{\frac{\sigma_{12}}{\sigma_{11}}}+\left[\frac{1}{2}\ \frac{\sigma_{11}}{\sigma_{22}}(1-\ell)^2 \frac{\Gamma(\frac{n+1}{2})\ \Gamma(\frac{n-2}{2})}{\Gamma(\frac{1}{2})\left[\Gamma(\frac{n}{2})\right]^2}+\frac{\sigma_{11}}{\sigma_{22}}\ell^2\right]\frac{\sigma_{22}}{\sigma_{11}}\right)$$

$$=\left(1-2\ell^2+\left[\frac{1}{2}\ (1-\ell)^2\frac{\Gamma(\frac{n+1}{2})\ \Gamma(\frac{n-2}{2})}{\Gamma(\frac{1}{2})\left[\Gamma(\frac{n}{2})\right]^2}+\ell^2\right]\right)$$

Thus

$$=1-2\ell^2+\left[\frac{(1-\ell^2)\Gamma\left(\frac{n+1}{2}\right)}{2\Gamma(\frac{1}{2})\left(\frac{n}{2}-1\right)\left(\Gamma(\frac{n}{2})\right)}+\ell^2\right]$$

Subtract & take $(1-\ell^2)$ as a common factor

Then

$$=(1-\ell^2)\left[1+\frac{1}{\Gamma(\frac{1}{2})(n-2)}\cdot\frac{\Gamma\left(\frac{n+1}{2}\right)}{\left[\Gamma(\frac{n}{2})\right]}\right] \quad (21)$$

$$\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$$

,

Then we have

$$\left(1-2E\left(\frac{s_{12}}{s_{22}}\right)\frac{\sigma_{12}}{\sigma_{11}}+E\left(\frac{s_{12}}{s_{22}}\right)^2\frac{\sigma_{22}}{\sigma_{11}}\right)$$

$$=(1-\ell^2)\left[1+\frac{1}{(n-2)\sqrt{\pi}}\cdot\frac{\Gamma\left(\frac{n+1}{2}\right)}{\left[\Gamma(\frac{n}{2})\right]}\right] \quad (22)$$

Substitute form (22) in (16)

Thus the exact second moment matrix is given by

$$E(b_1-B_1)(b_1-B_1)'$$

$$=\sigma_{11}(x_1'x_1)^{-1}(1-\ell^2)\left[1+\frac{1}{(n-2)\sqrt{\pi}}\cdot\frac{\Gamma\left(\frac{n+1}{2}\right)}{\left[\Gamma(\frac{n}{2})\right]}\right] \quad (23)$$

That is the results of zellner (1963)

Note that

The estimator of coefficient regression (B) when applying least square estimator procedure is $(\hat{B}_1)=(x'x)^{-1}(x'y)$ and the moment matrix of the single equation least square estimator is $\sigma_{11}(x_1'x_1)^{-1}$ but the moment matrix of the estimator of (SUR) model b_1^* in (8) where using the Aitken GLS procedure to estimate the sets of equation is based on a known variance covariance matrix since replace the unknown Σ by its consistent estimate S then the moment matrix of $\hat{B}_1$ estimator is $\sigma_{11}(x_1'x_1)^{-1}(1-\ell^2)$ as shown in original paper of zellner in (1962), deduce that the Aitken two stage estimator is less variance than least square , thus we can say that , the Aitken two stage estimator is more efficient than single equation least square.

6 Concluding Remarks

In Zellner (1962) the usual optimal properties of zellner's estimators is the Best Linear Unbiased Estimator (BLUE) estimator .with addition assumption of normality distributed, Zellner noted that when the disturbance terms in the different equation in system are correlated .it is denoted as the contemporaneous between the disturbance terms in the different equations not equal zero . These prove to be correct also when applying the Maximum Likelihood (ML) Method to obtain the (ML) estimator which it different from the single equation least squares estimators. When deal with each equation separately, and the Aitken two stage estimators is more efficient than estimator of single equation least square, Kakwani (1967) proved that the (SUR) estimators are unbiased estimators

ACKNOWLEDGMENT

The author thank prof Ghazal for his support and valuable great remarks , and we are welcome for helpful comments

REFRENCES

[1] Aiyi Liu,"efficient estimation of two seemingly unrelated equation", journal of multivariate analysis ,vol 82,p 445-456,(2002).

[2] Abadir.K.M. and Magnus .J.R, "Matrix algebra", Cambridge (2005).

[3] Ghazal ,A. G,"moment of the ratio of two dependent quadratic forms", institute of statistic, Cairo university,(1993).

[4] Greene, W.H., "Econometric Analysis", 5th, Macmillan: New York ,(2002).

[5] Jan Kamenta,Roy F.Gilbert "small sample properties of alternative estimators of seemingly unrelated regressions" , Journal of the American Statistical Association , vol.63,No 324,pp.1180-1200. (1968).

[6] Kakwani, N.C.,"The Unbiasedness of Zellner's Seemingly Unrelated Regression Equations Estimators, "Journal of the American Statistical Association, vol 62, No 317, pp141-142. (1967).

[7] Mehta , J.S. and swamy , P.A.V.B, "further evidence on the relative efficiencies of Zellner's seemingly unrelated regressions estimator ", Journal of the

[8] Revankar , N.S., "some finite sample results in the context of two seemingly unrelated regressions" , Journal of the American Statistical Association , vol . 68,pp.187-190 ,(1974).

[9] Srivastava,V .k. , "seemingly unrelated regression equations models estimation and inference" ,.Marcel Dekker, Inc.New York (1987).

[10] Zellner , A., "an efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias" , Journal of the American Statistical Association , vol.57,pp. 348-368. (1962).

[11] ______"estimators for seemingly unrelated regression equations. some exact finite sample results", Journal of the American Statistical Association ,vol. 58, pp.977-992,(1963).