*i*Journals: International Journal of Social Relevance & Concern
ISSN-2347-9698
**Volume 6 Issue11 November 2018**

# Estimation of Seemingly Unrelated Regression Equations with Non-Normal Disturbances

**E. I. Olamide**
Department of Statistics, Federal University of Technology, Akure
*E-mail: eiolamide@futa.edu.ng*

***DOI< 10.26821/IJSRC.6.11.2018.61101>***

## ABSTRACT

One of the basic assumptions of Seemingly Unrelated Regression Equations (SURE) is the normality of error terms in the regressions model. This paper thus considers estimating SUR model when the normality assumption of the error term is violated.

Small and asymptotic properties are examined using Ordinary Least Squares (OLS) and Feasible Generalized Least Squares (FGLS) estimators.

The results revealed that the SUR estimator (FGLS) was diminishingly efficient as sample size increases with their standard errors converging at large sample size of 1000.

**Keywords**: Feasible Generalized Least Squares, Ordinary Least Squares, Monte Carlo, Seemingly Unrelated Regressions.

## 1. INTRODUCTION

Seemingly Unrelated Regressions (SUR) is a set of regression equations that are related by their error terms. Each equation is a valid linear regression on its own thus satisfying the assumptions of the Classical Linear Regression Model (including the normality of the error terms).

Each equation can be estimated individually by the standard Ordinary Least Squares (OLS), Iterative Ordinary Least Squares, Generalized Least Squares (using true error covariances), Iterative Generalized Least Squares, Feasible Generalized Least Squares (using estimated error covariances), e.t.c.

Many researchers have carried out studies on the SUR model in diverse forms. Zellner (1962), Jan Kmenta and Roy F. Gilbert (1970) have shown that joint estimation of a system of linear regression equations with mutually correlated disturbances leads, in general, to more efficient estimates than estimation of each equation separately. A method of joint estimation developed by Arnold Zellner on the assumption that the disturbances are non-auto correlated was based on the Generalized Least Squares (GLS) approach. In their work, Jan Kmenta and Roy F. Gilbert developed several alternative estimators and compare their small-sample efficiency.

Arnold Zellner (1962) reported a method of estimating the parameters of a set of regression equations which involves application of Aitken's generalized least squares to the whole system of regression equations and established that the regression coefficient estimators so obtained are at least asymptotically more efficient than those obtained by an equation-by-equation application of ordinary least squares. He stressed the fact that the gain in efficiency can be quite large if "explanatory" variables in different equations are not highly correlated and if disturbance terms in different equations are highly correlated.

Binkley and Nelson (1988) presented a note on the efficiency of seemingly unrelated regression. They emphasized that the variance-covariance matrix for the seemingly unrelated regression estimator is expressed as an ordinary-least-squares variance-covariance matrix. This consents new insight into the efficiency of seemingly unrelated regression, especially the role of correlations among variables.

K. Viraswami (1998) delivered a working paper on some efficiency results on SURE model. In his study, he examined a two equation SUR model in which the equations have some common independent variables and obtained the asymptotic efficiency of the OLS estimator of a parameter of interest relative to its FGLS estimator. He also provided the small-sample relative efficiency of the ordinary least squares estimator and the seemingly unrelated residuals estimator.

W. B. Yahya, S. B. Adebayo, E. T. Jolayemi, B. A. Oyejola and O. O. M. Sanni (2008) examined the relative gain/loss in efficiency of Seemingly Unrelated Regression (SUR) estimators when one or more pair of the predictors in the system of equations is correlated (non-orthogonal).

Their paper addressed the challenges of multicollinearity which often affects the efficiency of SUR estimators by determining the 'Tolerable Non-

ISSN-2347-9698
Volume 6 Issue11 November 2018

orthogonal Correlation Points' (TNCP) among the predictors at which the efficiency of SUR estimators will still be preserved.

This paper thus studies the gain in efficiency of the FGLS over the OLS estimator when the error terms are not normally but uniformly distributed.

The rest of the paper is organized as follows. In section 2, the methodology is presented; section 3 gives the parametric SUR framework while the simulation study is discussed in section 4. Results and discussions are presented in section 5. Concluding remarks are given in section 6.

## 2. METHODOLOGY

The Monte Carlo approach is a method of inferring conclusion from simulated data. Essentially, it is used to infer the desired solution from the behaviour of generated random numbers that have been chosen in such a way that they directly simulate the physical random process of the original problem. Whether theoretical or analytical approach, conclusion are deduced from postulates.

In Econometrics, attention is often focused on evaluating the behaviour of estimators in small sample and asymptotic theory fails most of the time in providing enough useful information about these estimators. This inexorably creates a problem in trying to choose from among competing estimates. One way of studying the small sample properties of estimators is to utilize the Monte Carlo method.

The Monte Carlo method has been applied in past researches not only to the choice of alternative estimators but also in determining the impact of sample size, serial correlation, multicollinearity and other factors on the different possible estimators. This approach creates a "laboratory environment" where controlled experiments on estimators are performed.

Small sample properties of various econometric techniques are studied from simulated data in Monte Carlo studies and not with direct application of the techniques to actual observations. This approach is due to the fact that actual observations on economic variables are usually infected by multicollinearity, autocorrelation, errors of measurement and other econometric "defects" and in some cases simultaneously. Studies on small sample properties of estimators are usually based on the assumption of the simultaneous occurrence of all these problems.

By Monte Carlo approach we can generate data sets and stochastic terms that are free from these problems, thereby generating data resembling those obtained from controlled experiments.

This study examines the extent to which the correlated errors are uniformly distributed and Σ is estimated.

## 3. PARAMETRIC SUR FRAMEWORK

Consider a system of regression equations with T response variables each containing S observations with associated distinct vectors of explanatory variables as:

$$\begin{aligned} y_1 &= X_1\beta_1 + \varepsilon_1, \\ y_2 &= X_2\beta_2 + \varepsilon_2, \\ &\vdots \\ y_T &= X_T\beta_T + \varepsilon_T \end{aligned} \qquad (1)$$

The Seemingly Unrelated Regressions (SUR) model above is:

$$y_i = X_i\beta_i + \varepsilon_i\,, \quad i = 1, \ldots, T, \qquad (2)$$

where $y_i$ is a T ($S$ x 1) vector of observations on the dependent variable y,

$X_i$ is a $S$ x $k_i$ matrix of non-stochastic regressors, $\beta_i$ is a $k_i$ x 1 vector of unknown regression coefficients and $\varepsilon_i$ is a $S$ x 1 vector of unobservable disturbances.

We can then further compress the notion by stacking the $T$ equations in the compact form

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \qquad (3)$$

where y is an $TS$ x 1 vector, X is an $TS$ x $k$ matrix, $\beta$ is a $k$ x 1 and $\varepsilon$ is an $TS$ x 1 vector of disturbances.

i.e,

$$y = \begin{bmatrix} y_1 \\ y_{2a} \\ \vdots \\ y_T \end{bmatrix}, \qquad X = \begin{bmatrix} X_1 & 0 & \ldots & 0 \\ 0 & X_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & X_T \end{bmatrix},$$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_T \end{bmatrix} \quad \text{and } \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{bmatrix}$$

and with

$$E(\varepsilon \,/\, X_1, X_2, \ldots, X_T) = 0,$$

$$E(\varepsilon\,\varepsilon' \,/\, X_1, X_2, \ldots, X_T) = \Omega.$$

From the stacked model in (3), the Ordinary Least Squares (OLS) estimator of the parameter $\beta$ is given by

$$\hat{\beta}_{OLS} = (X'X)^{-1}\, X'Y \qquad (4)$$

We assume that a total of $S$ observations are used in estimating the parameters of the $T$ equations. Each equation involves $K_t$ regressors, for a total of
$K = \Sigma_{i=1}^{T} K_i$ .
We will require $S > K_i$. The data are assumed to be well behaved. We also assume that disturbances are uncorrelated across observations. Therefore,
$E\ [\varepsilon_{it}\varepsilon_{js}\ | X_1, X_2, \ldots, X_T] = \sigma_{ij}$, if $t = s$ and 0 otherwise.
The disturbance formulation is therefore,

$$E[\varepsilon_i \varepsilon_j' | X_1, X_2, \ldots, X_T] = \sigma_{ij} I_S$$

or

$$E(\varepsilon\ \varepsilon' / X_1, X_2, \ldots, X_T) = \Omega$$

$$= \begin{bmatrix} \sigma_{11} I & \sigma_{12} I & \ldots & \sigma_{1T} I \\ \sigma_{21} I & \sigma_{22} I & \ldots & \sigma_{2T} I \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{T1} I & \sigma_{T2} I & \ldots & \sigma_{TS} I \end{bmatrix}$$

(5)

Where,

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1T} \\ \sigma_{21} & \sigma_{22} & \ldots & \sigma_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{T1} & \sigma_{T2} & \ldots & \sigma_{TS} \end{bmatrix}$$

This implies that in (5),

$$\Omega = \Sigma \otimes I,$$

and

$$\Omega^{-1} = \Sigma^{-1} \otimes I.$$

If $\Omega$ is known and denoting the ij[th] element of $\Sigma^{-1}$by $\sigma^{ij}$, the generalized least squares estimator for the coefficients in this model is:

$$\hat{\beta} = (X'\Omega^{-1}X)^{-1}\ X'\Omega^{-1}y \qquad (6)$$

i.e,

$$\hat{\boldsymbol{\beta}} = [X'(\boldsymbol{\Sigma}^{-1} \otimes I)\ \mathbf{X}]^{-1} X'(\boldsymbol{\Sigma}^{-1} \otimes I)\ \mathbf{y}.$$

If otherwise, $\Omega$ is not known, it is estimated using

$$\Sigma_{T\times T}^{s} = \frac{1}{s} \Sigma_{i=1}^{S} \begin{bmatrix} \hat{u}_{1i} \\ \vdots \\ \hat{u}_{Ti} \end{bmatrix} [\hat{u}_{1i} \quad \cdots \quad \hat{u}_{Ti}]$$

and this gives rise to the Feasible Generalised Least Squares which is denoted by

$$\hat{\boldsymbol{\beta}}_{FSUR} = \hat{\boldsymbol{\beta}}_{FGLS} = (\mathbf{X}'\ (\hat{\boldsymbol{\Sigma}} \otimes \mathbf{I})^{-1}\ \mathbf{X})^{-1}\ \mathbf{X}'\ (\hat{\boldsymbol{\Sigma}} \otimes \mathbf{I})^{-1}\ \mathbf{Y}$$

This paper assumes the disturbance terms are uniformly distributed and are generated from (0, 1).

## 4. SIMULATION STUDIES

Properties of various econometric methods are studied from simulated data in Monte Carlo studies and not with direct application of the methods to real observations. This is due to the abnormalities usually present in real observations such as multicollinearity, errors of measurement and others. By Monte Carlo approach we can generate data sets and stochastic terms that are free from these problems, thereby generating data resembling those obtained from controlled experiments.
We consider a system of regression equations with the structural form;

$$y_1 = 0.5 + 0.7\ x_{11} + 0.3\ x_{12} + \varepsilon_1$$
$$y_2 = 0.4 + 0.6\ x_{21} + \varepsilon_2$$
$$y_3 = 0.8 + 0.2\ x_{31} + \varepsilon_3$$

Where the explanatory variables $(x_{11}, x_{12}, x_{21}, x_{31})$ are generated from Poisson distribution with $\lambda = 1$ for the various sample sizes of 20, 100 and 1000.
Then, $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3)'$ are series of random uniform deviates of required lengths 20, 100 and 1000 that are generated.

## 5. RESULTS AND DISCUSSIONS

The summary of the results when the model is estimated with uniformly distributed error terms is presented in the table below.

**Table 1: COMPARATIVE STUDY OF THE ESTIMATORS ACROSS DIFFERENT SAMPLE SIZES**

| | T =20 | | | | T = 100 | | | | T = 1000 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | OLS | | FGLS | | OLS | | FGLS | | OLS | | FGLS | |
| Regressions | Estimate | SE | Estimate | SE | Estimate | SE | Estimate | SE | Estimate | SE | Estimate | SE |
| $y_1$ | | | | | | | | | | | | |
| $\beta_{10} = 0.5$<br>$\beta_{11} = 0.7$<br>$\beta_{12} = 0.3$ | 1.0766<br>0.5921<br>0.3243 | 0.1208<br>0.0585<br>0.0577 | 1.0663<br>0.5926<br>0.3332 | 0.1190<br>0.0574<br>0.0566 | 1.0584<br>0.6807<br>0.2788 | 0.0499<br>0.0316<br>0.0259 | 1.0542<br>0.6826<br>0.2810 | 0.0496<br>0.0314<br>0.0258 | 1.0122<br>0.6872<br>0.2913 | 0.0162<br>0.0091<br>0.0092 | 1.0123<br>0.6869<br>0.2914 | 0.0162<br>0.0091<br>0.0092 |
| $y_2$ | | | | | | | | | | | | |
| $\beta_{20} = 0.4$<br>$\beta_{21} = 0.6$ | 0.8932<br>0.5059 | 0.0985<br>0.0864 | 0.9154<br>0.4812 | 0.0974<br>0.0848 | 0.9186<br>0.5625 | 0.0374<br>0.0283 | 0.9204<br>0.5605 | 0.0373<br>0.0281 | 0.9123<br>0.5889 | 0.0125<br>0.0085 | 0.9130<br>0.5884 | 0.0125<br>0.0085 |
| $y_3$ | | | | | | | | | | | | |
| $\beta_{30} = 0.8$<br>$\beta_{31} = 0.2$ | 1.4557<br>0.1484 | 0.0800<br>0.0471 | 1.4620<br>0.1433 | 0.0788<br>0.0460 | 1.2550<br>0.2228 | 0.0425<br>0.0272 | 1.2524<br>0.2250 | 0.0422<br>0.0269 | 1.2948<br>0.2056 | 0.0128<br>0.0094 | 1.2948<br>0.2056 | 0.0128<br>0.0094 |

**TABLE 2: COMPARATIVE STUDY OF THE ESTIMATORS USING ADJUSTED $R^2$**

| Regressions | T = 20 | | T =100 | | T = 1000 | |
|---|---|---|---|---|---|---|
| | OLS | FGLS | OLS | FGLS | OLS | FGLS |
| 1 | 0.8533 | 0.8531 | 0.8518 | 0.8518 | 0.8692 | 0.8692 |
| 2 | 0.6364 | 0.6348 | 0.7995 | 0.7995 | 0.8273 | 0.8273 |
| 3 | 0.3165 | 0.3161 | 0.4003 | 0.4003 | 0.3236 | 0.3236 |

## 5.1. DISCUSSIONS

Table 1 gives the estimates and their standard errors across the different sample sizes considered. It was found that the estimators were consistent and efficient especially for the slopes in all the three regression equations examined. As the sample sizes increased, the standard errors of the estimates converged for the two estimators.

Table 2 gives the adjusted $R^2$ for the two estimators. It was found that there is no appreciable difference between OLS and FGLS estimators.

## 6. CONCLUSION

The results obtained revealed that the FGLS estimator was efficient in small sample but not in large sample. In other words, as the sample size increased, there is no gain in efficiency accruing to the SUR estimator when the error terms are uniformly distributed.

## 7. REFERENCES

[1]. Bartels, R., & Fiebig, D. G., "A simple characterization of seemingly unrelated regressions models in which OLS is BLUE", The American Statistician, vol. 45(2), 1991, pp. 137-140.

[2]. Binkley, J. K., & Nelson, C. H,, "A note on the efficiency of seemingly unrelated regression", The American Statistician, vol. 42(2), 1988, pp. 137-139.

[3]. Dwivedi, T. D. and Srivastava, V. K., "Optimality of least squares in the seemingly unrelated regression equation model, Journal of Econometrics, vol. 7, 1978, pp. 391-395.

[4]. Fomby, T. B., Hill, R. C. and Johnson, S.R., "Advanced econometric methods", New York : Springer-Verlag, 1984, 155-169.

[5]. Gallant, A. R., "Seemingly unrelated nonlinear regressions", Journal of Econometrics, vol. 3, 1975, pp. 35-50.

[6]. Judge, G. G., Griffiths, W. E., Hill, R. C., & Lütkepohl, H., "The Theory and Practice of Econometrics", New York: Wiley, 2nd ed., 1985, pp. 465-514.

[7]. Judge, G. G., Hill, R. C., Griffiths, W. E. & Lütkepohl, H., "Introduction to the theory and practice of econometrics", New York: Wiley, 2nd ed., 1988, pp. 444-468.

[8]. Kmenta, J., & Gilbert, R. F., "Small sample properties of alternative estimators of Seemingly Unrelated Regression", Journal of American Statistical Association, vol. 63, 1968, pp. 1180-1200.

[9]. Parks, R. W., "Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated", Journal of the American Statistical Association, vol. 62(318), 1967, pp. 500-509.

[10]. Percy, D. F., "Prediction for seemingly unrelated regressions", Journal of the Royal Statistical Society Series, vol. 54(1), 1992, pp. 243-252.

[11]. Powell, A. A., "Aitken estimators as a tool in allocating predetermined aggregates", Journal of the American Statistical Association, vol. 64, 1965, pp. 913 – 922.

[12]. Rao, P. and Griliches, Z, "Small-sample properties of several two-stage regression methods in the context of auto-correlated errors", Journal of the American Statistical Association, vol. 64(325), 1969, pp. 253-272.

[13]. Revankar, N. S., "Some Finite Sample Results in the Context of Two Seemingly Unrelated Regression Equations", Journal of American Statistical Association, vol. 69, 1974, pp. 187-190.

[14]. Rocke, D. M., "Bootstrap Bartlett adjustment in seemingly unrelated regression", Journal of the American Statistical Association, vol. 84(406), 1989, pp. 598-601.

[15]. Telser, L. G., "Iterative estimation of a set of linear regression equations", Journal of American Statistical Association, vol. 59, 1964, 845-862.

[16]. Zellner, A., "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias", Journal of the American Statistical Association , vol. 57(298), 1962, pp. 348-368.