IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 14, Issue 4 Ver. III (Jul - Aug 2018), PP 01-13
www.iosrjournals.org

On The Efficiency of Some Estimators for Modeling Seemingly Unrelated Regression with Heteroscedastic Disturbances

R. B. Afolayan[1]and B. L. Adeleke[2]
Department of Statistics, Faculty of Physical Sciences, University of Ilorin, Ilorin, Nigeria.
Corresponding Author: R. B. Afolayan

Abstract: *This Study investigates the performances of seemingly unrelated regression (SUR) estimators when the homoscedastic assumption of disturbances is violated in each of the regression equation. The finite properties and relative performance of the estimators to Ordinary Least Square (OLS) were examined under four forms of heteroscedasticity of the error terms and levels of Contemporaneous Correlation (Cc). The efficiency of three estimation techniques for SUR model was examined using Root Mean Square Error (RMSE) criterion to determine the best estimator(s) under different conditions at various sample sizes. The simulation results revealed that OLS estimator performed very well in small sample in the absence or low contemporaneous correlation at different forms of heteroscedasticity except when the form of heteroscedasticity is quadratic where Huber estimator was the best. Feasible Generalized Least Square (FGLS) is the most efficient estimator when the contemporaneous correlation (ρ) is moderate* $(\rho = 0.5 \;\&\; \rho = 0.7)$ *or high (ρ=0.95) irrespective of the sample sizes or forms of heteroscedasticity. However, in large samples; OLS, Huber and BISQ showed equivalent performances. Above all, the relative gain of using FGLS increases tremendously when the Cc is high. The study concludes that no single estimator is generally efficient than the other under all different conditions considered.*
Key Words*: Seemingly unrelated Regression, Heteroscedasticity, Homoscedasticity, Contemporaneous correlation and Feasible Generalized Least Square*

Date of Submission: 27-07-2018 Date of acceptance: 11-08-2018

I. Introduction

Seemingly Unrelated Regression (SUR) Equations is a system of M multiple regression equations in which each equation has a single dependent and K (k≥1), independent or exogenous variables as in standard regression model. The m equations have no link or relationship with one another except that their disturbances are said to be correlated.

Zellner (1962) introduced the SUR estimation procedure for systems of regression equation as against the general estimation method of ordinary least square (OLS) estimator. Prior to the introduction of SUR estimation procedure, OLS has been the common method of estimating the coefficient of the parameters of regression model. The OLS is a single-equation estimation method that does not account for the interactions that may exist among the different regression equations. The OLS is unbiased, consistent and remain the most efficient estimator when all the assumptions of classical linear regression model are satisfied. The single by single equation estimation of SUR model without using the information from error terms across equations by OLS is still unbiased and consistent but seizes to be the best linear unbiased estimator (BLUE).

In view of this, the inferences about the parameters of the model using the statistic from OLS estimator become invalid. Zellner (1962) viewed that the disturbance terms of these equations are likely to be contemporaneously correlated because of unobservable factors that influence the disturbance term in one equation may affect the disturbance terms in other equations. Estimating these equations separately without accounting for the non-zero covariance structure of the errors leads to inefficient parameter estimates.

However, the joint estimation procedure of SUR using the information from the correlation among the errors of different equations is more efficient than the separate equation estimation procedure of the ordinary least square (OLS) and the gain in efficiency is achieved if contemporaneous correlation between the disturbances across equation is very high and other assumptions of classical regression model are satisfied, see (Judge et al 1988; Zellner, 1962&1963; Zellner&Theil, 1962).

1.1. Background of the Study

After the Proposed Aitken's Generalized Least Squares (AGLS) estimator often reffered to as Feasible Generalized Least Square (FGLS) by Zellner (1962), much extensive theoretical and empirical applications of the work have been recorded in econometrics, statistics and other areas in recent time. For instance, finite sample properties of SUR estimator (SURE) have been studied in the literature (e.gZellner, 1962 &1963; Kakawani, 1967, Zellner& Huang, 1962; Zellner&Theil, 1962) by asymptotic expansions (e.g. Srivastava and Maekawa, 1995), or by simulation (e.g. Kmenta& Gilbert, 1968). Several other researchers also work on SUR model under the violation of basic model assumptions in different data generating processes, such as time series, Panel, Cross-sectional etc. For example, Kmenta and Gilbert (1970), considered the problem of estimating a system of regression equations in which the disturbances are both serially and contemporaneously correlated. They developed an estimator that is consistent and efficient than SURE in situation in which each of the disturbance follows a first order auto regressive scheme. Other works within the SUR frame-work with autocorrelated disturbances and error component can be found in (Guilkey and Schmidt (1973); Avery (1977); Walter Kramer (1980); Baltagi (1980); Messemer and parks (2004);and Alaba(2010)).

Takada et al (1995), presented methods of resolving the problem of non-singularity in the covariance matrix of the errors in the SUR model and proposed an efficient procedure of estimation. The empirical study of the estimator was investigated by studying the diffusion processes of video cassette recorders across different geographic regions in the US, which exhibits a singular covariance matrix. The empirical results show that the procedure is efficient in tackling the problem and provide plausible estimation results.

The efficiency of SUR model when exogenous variables across-equations are correlated was investigated by Yahya et al (2008). In their work, it was established that at large sample size ($n \geq 50$), the SURE would still be efficient if correlation exists among the exogenous variable in SUR model.

However, under small and moderate sample sizes, they recommended a Tolerable non-orthogonal correlation point (TNCP) of ± 0.2 under which SUR estimator would still be efficient.

Situations in which the assumption of independency of error terms is violated was investigated by Olamide and Adepoju (2013), the performances of FGLS, Iterative Ordinary Least Squares (IOLS) and Ordinary Least Squares (OLS) were compared in SUR model with first order autoregressive error terms.

Olanrewaju (2013) noted that one of the problems of time series data is autocorrelation of the error terms and extended the work of Yahay et al(2008) to situations where disturbances in different equations follow a first order autocorrelation. The effect of multicollinearity, autocorrelation and correlation between the errors terms on some methods of estimation of system of simultaneous equation were investigated through Monte Carlo experiment. He implored seemingly unrelated regression model and assumed first order autocorrelation of the error term. The performances of Ordinary Least Squares (OLS), Three Stage Least Squares (3SLS), Feasible Generalized Least Squares (FGLS), Maximum Likelihood (ML), Full Information Maximum Likelihood (FML) and Multivariate Regression (MR) were investigated extensively at different levels of multicollinearity, autocorrelation and correlation between the error terms. The results of the study shows that ML estimator is preferred when there is presence of autocorrelation and multicollinearity in the model. However, when there is correlation between the error term SUR and 3SLS should be preferred.

SUR model also gained an appreciable application in econometrics and applied sciences, for instance ,Sparks (2004) developed a SUR procedure that is applicable to environmental situations especially when missing and censored data are inevitable.

Singh and Ullah (1974) extended Zellner's (1962) SUR model to credibility regression model with random coefficient and proposed estimators that are asymptotically more efficient than Zellner's estimator. In share equation systems with random coefficient, Mandy & Martins-Filho (1993) proposed a consistent and asymptotically efficient estimator for SUR systems that have additive heteroscedastic contemporaneous correlation. They followed Amemiya (1977) by using Generalized Least Squares (GLS) to estimate the parameter of the covariance matrix.

1.2. Justification For The Study

The violation of each assumption has attracted the attention of many researchers especially in single regression model but the violation of homoscedasticity assumption in system of regression model has not been adequately exploited. When errors are heteroscedastic, the consequences ofusing FGLS estimatorbased statistic in hypothesis testing could be disastrous than using single equation estimation for SUR model such as OLS. In recent time, fewer research works have been done to investigate the performances of SUR estimators under violation of basic model assumptions in SUR model. More so, since many real life data do not satisfy the Constant variance assumption of the error terms there is need for further investigation on the performances of FGLS relative to some alternative estimator of SUR model with different error variance structures. The present study therefore investigate the performances of some estimators of SUR model under different forms of heteroscedasticity at various levels ofcontemporaneous correlation under different samples sizes using Monte Carlo simulation procedures.

II. Sur Framework And Estimation Procedures

The SUR model is given by

$$Y_i = X_i\beta_i + \varepsilon_i \qquad i = 1, \dots\dots, M \qquad (2.0)$$

Where Y_i is *n x 1* vector denoting observations for the i equation and X_i *is* a *n x k* matrix of non-stochastic regressors, $\beta_i$is a *k x 1* vector of parameter and *n x 1* vector of disturbances ε_i.

The equation (2.0) could be written compactly as;

$$Y = X\beta + \epsilon \qquad (2.1)$$

And in matrix form as:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} X_1 & 0 & \cdots & \cdots & 0 \\ 0 & X_2 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & X_m \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \beta_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \vdots \\ \varepsilon_m \end{bmatrix} (2.2)$$

$mn \times 1 \qquad mn \times \sum_{i=1}^{m} K_i \qquad \sum_{i=1}^{m} k_i \times 1 \;\; mn \times 1$

The assumption regarding the disturbances in equation (2.0) is that the disturbances are contemporaneously correlated with mean, $E(\varepsilon) = 0$ and variance-covariance of the disturbances given as Ω is $E(\varepsilon\varepsilon') = Var(\varepsilon) = \Omega = \sum \otimes I$where $\otimes$ is the Kronecker product operator, I is an identity matrix of order $N \times N$ and $\sum = ((\sigma_{ij}))$ is an $M \times M$ symmetric and positive definite matrix such that

$$E(\varepsilon_i \varepsilon'_j) = \begin{cases} \sigma_{ii} I & for\ i = j, \qquad j = 1, \dots\dots\dots\dots m \\ \sigma_{ij} I & for\ i \neq j, \qquad i = 1, \dots\dots\dots\dots, m \end{cases} \qquad (2.3)$$

where $\Omega$is an *m x m* matrix of the form:

$$E(\varepsilon_i \varepsilon'_j) = \text{var.}(\varepsilon_i) = \Omega = \begin{pmatrix} \sigma_{11}I & \sigma_{12}I\dots\dots & \sigma_{1m}I \\ \sigma_{21}I & \sigma_{22}I\dots\dots & \sigma_{2m}I \\ \vdots & \vdots & \vdots \\ \sigma_{m1}I & \sigma_{m2}\dots\dots & \sigma_{mm}I \end{pmatrix} \qquad (2.4)$$

Estimating k $\times$ l vector of parameters in the equation (2.1) is equivalent to applying OLS to each equation since the covariance matrix is no longer scalar of the form σ^2I. The ordinary least-squares (OLS) estimator of β for separate equation in equation (2.1) is given as

$$\hat{\beta} = (X'X)^{-1}X'y \qquad (2.5)$$

The OLS estimator of equation (2.1) is unbiased and consistent but less efficient

And its variance-covariance matrix is

$$V(\hat{\beta}) = (X'X)^{-1}X'\Omega X(X'X)^{-1} \qquad (2.6)$$

Aitken (1962) suggested an infeasible estimator called generalized least square (GLS) estimator for a known m by m positive definite variance-covariance matrixΩ using a Weighted Least Square (WLS) approach.

The idea is to transform the model in (2.1) such that the disturbances covariance matrix in (2.6) becomes$\Omega = \sigma^2$I.

That is, transform the joint model in (2.1) by a square and invertible but non- diagonal weighting matrix A, such that:

$$A\Omega A' = I_n \Leftrightarrow \Omega = (A'A)^{-1} \mathbf{(2.7)}$$

Multiplying the compact SUR model in (2.3) by A, we obtain

$$AY = AX\beta + A\epsilon \qquad (2.8)$$

With $Y^* = AY$, $X^* = AX$, $\varepsilon^* = A\varepsilon$, *then* model (2.1) becomes

$$Y^* = X^*\beta + \varepsilon^* \qquad (2.9)$$

The $E(\varepsilon^* \varepsilon^{*'}) = E(A\varepsilon\varepsilon' A') = AE(\varepsilon\varepsilon')A' = A\Omega A' = I_m$, thus the transformed model now has a diagonal variance- covariance matrix with all its off-diagonal elements zero. That is $E(\varepsilon^*)=0$ and $Var(\varepsilon^*) = \sigma^2 I_{nm}$.

The generalized least square (GLS) estimator of β in (2.9) is given by

$$\widehat{\beta_{GLS}} = (X^*X^*)^{-1}X^{*'}Y^* = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y \qquad (2.10)$$

And its variance- covariance matrix is

$$V(\beta_{GLS}) = \sigma^2(X'\Omega^{-1}X)^{-1} \qquad (2.11)$$

However, in practice the value of the covariance matrix Ω is generally unknown. Zellner (1962) suggested replacing Ω by its consistent estimator $\widehat{\Omega}$ and then estimate the parameters in SUR model in two steps called Feasible GLS (FGLS) estimator given as

$$\widehat{\beta_{FGLS}} = (X'\widehat{\Omega}^{-1}X)^{-1}X'\widehat{\Omega}^{-1}Y \qquad (2.12)$$

Where $\widehat{\Omega} = S \otimes I$ is an *m x m* matrix based on single equation of OLS disturbances and S= (s_{ij}) , where $s_{ij} = \varepsilon_i'\varepsilon_j/n - k$, see (*Yahya et.al, 2008).*

III. Robust M- Regression Estimator

M-estimators for regression which was introduced by Huber (1973), is a generalization of the OLS estimation procedure by minimizing the sum of a less rapidly increasing function (objective function) of the residuals instead of minimizing the sum of squared residuals as;

$$\hat{\beta} = min \sum_{i=1}^{n} \rho(y_{i-}X_i{}' \beta) \qquad (3.1)$$

Equivalent to M-Estimator of location, the robustness of the estimator is determined by the choice of weight function. The solution is not scale equivariance, and thus the residuals must be standardized by a robust scale estimator $\hat{\sigma}_e$ usually the median absolute deviation (MAD). Differentiating the objective function in (3.1) and setting the partial derivative to zero gives the score function:

$$\hat{\beta} = min \sum_{i=1}^{n} \psi((y_{i-}X_i{}' \beta)/\hat{\sigma}_e)\chi_i{}' = 0 \qquad (3.2)$$

Where ψ is the derivative of the objective function ρ.The equation (3.2) is a P system of normal equations. Then ψ is replaced by appropriate weights that decrease as the size of the residual increases, defined by $w(\varepsilon) = \frac{\psi(\varepsilon)}{\varepsilon}$ $and\ w_{i=}w(\varepsilon)$. Hence (3.2) becomes

$$\hat{\beta} = min \sum_{i=1}^{n} \omega_i((y_{i-}X_i{}' \beta)/\hat{\sigma}_e)\chi_i{}' = 0 \qquad (3.3)$$

In order to solve the equation in (3.3) with respect to β , an iterative procedure called Iterative Re-Weighted Least Squares (IRWLS) is employed as follows;

1. Set the iteration counter t=0 and select the initial estimates $\beta^{(0)}$ from initial OLS estimates.
2. Calculate the residual from OLS in (1) and set as $\epsilon^{(0)}$
3. Select any weight function of choice and applied to the initial OLS residuals to create an initial weights, w($\epsilon^{(0)}$)
4. The first iteration, t=1, uses weighted least squares(WLS) to minimize $\sum \omega_i{}^1 \varepsilon_i{}^2$ and obtain $\beta^{(1)}$.In matrix form, W represent an *n x n* diagonal matrix of individual weights, the solution is

$$\widehat{\beta^{(1)}} = (X'WX)^{-1}X'WY$$

5. At each iteration t, calculate the residual $\epsilon^{(t-1)}$ and associated weights w($\epsilon^{(t-1)}$) from previous iteration
6. Solve for new WLS estimates $\widehat{\beta^{(t)}} = (X'W^{(t-1)}X)^{-1}X'W^{(t-1)}Y$, where$W^{(t-1)} = diag(W_i{}^{(t-1)})$, the current weight matrix and X is the model matrix, with $\chi_i{}'$ as its ith row.
7. Steps 4-6 are repeated until the estimated coefficients converge.

M-estimates based on Huber's ψ function were used in this study with Objective and bi-weight function defined respectively as:

$$\rho(\varepsilon)_{huber} = \begin{cases} \frac{1}{2}\varepsilon^2 & for\ |\varepsilon| \le k \\ k|\varepsilon| - \frac{1}{2}K^2 & for\ |\varepsilon| > k \end{cases} \qquad (3.4$$

$$w_{huber} = \begin{cases} 1 & for\ |\varepsilon| \le k \\ \frac{k}{|\varepsilon|} & for\ |\varepsilon| > k \end{cases} \qquad (3.5)$$

Huber's ψ functions have computational advantage but sensitive to leverage points,(See Maronne *et al* (2006) for more details.The choice of c = 1.345 recommended by Huber (1981) produced a relative efficiency of approximately 95% when the error density is normal.

IV. Robust Mm- Regression Estimator

MM-estimator is robust estimator which combines the high breakdown point of S-estimator and high efficiency of M estimator. Breakdown point is a measure of the proportion of outliers that can be addressed before these observations affect the model. The S-estimator proposed by Rousseeuw andYohai (1984), minimizes the dispersion of the residuals, expressed as:

$$min \frac{1}{n} \sum_i^n \rho(\frac{\varepsilon_i}{\hat{\sigma}_e}) \qquad (4.1)$$

Where $\hat{\sigma}_e$ is the estimate of residual scale and ρ is the weight function.

The MM-estimator requires three stages. The first stage finds the regression parameter using S estimator which is consistent and has high breakdown point of 50% but not necessarily efficient. In the second stages, M estimator of the residuals scale is calculated using the residuals obtained from initial S estimator. The M estimator of the regression parameter as described in section 1.34 is then used in the third stage to estimate the regression parameter that is consistent, robust to outlier with high efficiency and asymptotically normally distributed (Maronna*et al* (2006)). The MM uses the Turkey's objective function:

$$w_{bisquare} = \begin{cases} \frac{c^2}{6}\{1-[1-(\varepsilon/c]^3\} & for\ |\varepsilon| \le k \\ \frac{c^2}{6} & for\ |\varepsilon| > k \end{cases} \qquad (4.2)$$

And Turkeybi-weight functions as:

$$w_{bisquare} = \begin{cases} \{[1-(\varepsilon/c)^2]^2\} & for\ |\varepsilon| \le c \\ (0) & for\ |\varepsilon| > c \end{cases} \qquad (4.3)$$

Where the constant, c= 4.685 produced 95% relative efficiency (Susanti et al., 2014).

V. Methodology

The simulation study considers a system of SUR equations containing two distinct linear regression equations to examine the performances of single equation estimators relative to SUR estimation procedure under different forms of heteroscedasticity in SUR model. In view of this,the study considered four heteroscedastic error structures coined from Harvey (1976) additive and multiplicative heteroscedastic model but in our model we assumed that the variance of the error varies as the mean of the responses. These two General forms are:
(i) $Var(\varepsilon^1\varepsilon)= \sigma^2 E(y_i)^\theta$ and (ii) $Var(\varepsilon^1\varepsilon)= \sigma^2 e^{E(yi)}$; $\theta \ge 0$,

Emanating from the two above, four heteroscedastic structures were formulated, namely:

(a) Exponential form (EXP): $h_{i1} = \sigma^2 \exp(\alpha_1 + \gamma_{20}X_{i1} + \gamma_{21}X_{i2})$

(b) Linear forms (LN): $h_{i2} = \sigma^2 (\alpha_1 + \gamma_{20}X_{i1} + \gamma_{21}X_{i2})$

(c) Square Root (SQR): $h_{i3} = \sigma^2 (\alpha_1 + \gamma_{20}X_{i1} + \gamma_{21}X_{i2})^{0.5}$

(d) Quadratic (QUAD): $h_{i4} = \sigma^2 (\alpha_1 + \gamma_{20}X_{i1} + \gamma_{21}X_{i2})^2$ i= 1,2.

Where α_1, γ_{20} and, γ_{21} are arbitrary constant fixed for each sample size as ; $\alpha_1 =$ 0.6, γ_{20}=0.8 , $\gamma_{21} =$ 0.4, and σ^2=1

In order to assess the asymptotic and small sample properties of the various estimators under the violation of normality assumption of the responses, homoscedasticity assumption of the error terms and, the level of correlation among errors across equations, the entire simulation experiments was performed for various sample sizes (n): n=500, 250,100,50,30,20 with 100 replicates in each case. However, the model specification, form of heteroscedasticity and distributions used in this work are in consonant with what is in the literature with little modification, See (Takada et.al, 1995; Carroll, 1982: Kmenta& Gilbert, 1968).

The simulationstudy was generated given the SUR model in equation(1.0) as:

$Y_1 = X_1\beta_1 + \varepsilon_1$

$Y_2 = X_2\beta_2 + \varepsilon_2$

With exogenous variables, error terms and parameters defined as:

X_1= (1, X_{11} , X_{12}); $\beta_1 = (\beta_{10}, \beta_{11}, \beta_{12})$

X_2= (1, X_{21} , X_{22}); $\beta_1 = (\beta_{20}, \beta_{21}, \beta_{22})$ and $\varepsilon = (\varepsilon_1, \varepsilon_2)$

The values of the model parameters were set as follows;

β_{10}= -20, $\beta_{11} =$ 15, $\beta_{12=}$13, β_{20}= 11.5 , β_{21} = 10.0 , and β_{22} =6.5

The contemporaneous correlations between the errors from the two equations were specified as: $\rho =$ 0.0 , 0.2 ,0.5 ,0.7 *and* 0,95.

The formulated SUR model becomes:

$y_{i1} = -20 + 15X_{i1} + 13X_{i2} + \varepsilon_{i1} h_{ij}$; j=1, 2,3,4. (5.1)

$y_{i2} = 11.5 + 10X_{1i} + 6.5X_{2i} + \varepsilon_{i2} h_{ij}$; for i= 1,…, n (5.2)

VI. Data Generation Procedure

The simulations of observations for the two regression equation model of (5.1& 5.2) were given as follows:

1. The exogenous variables $[X_1, X_2]$ were generated from uniform distribution with parameters a=0.5 and b= -0.5, U(0.5, -0.5) for various sample sizes
2. The error terms $[\varepsilon_{i1}, \varepsilon_{i2}]$ in equation (5.1& 5.2) were drawn from normal distribution with mean=0 and variance equals 1 ,i.e. N(0,1)

The heteroscedastic structure for error variance, h_{ij} as given in section 5were generated using the specified values: $\alpha_1 = 0.6$, γ_{20}=0.8, $\gamma_{21} = 0.4$, σ^2=1 based on generated exogenous variables $[X_1, X_2]$
1000 replications were used for different sample sizes 20, 30, 50, 100, 250 and 500. For each replication, the value of the parameters $\beta_1 = (\beta_{10}, \beta_{11}, \beta_{12})$ and $\beta_1 = (\beta_{20}, \beta_{21}, \beta_{22})$ were estimated for the following estimators:

I. Feasible Generalized Least Squares (FGLS)
II. Ordinary Least Squares (OLS)
III. Huber M-Estimator (Huber)
IV. Turkey Bi-Squares MM- Estimator (BISQ)

The performances of the estimators at different heteroscedastic structures, contemporaneous correlations and sample sizes were evaluated using RMSE and their relative efficiency to OLS. The relative efficiency of an estimator is the measure of the degree to which the estimator performs similar to common method (OLS). The relative efficiency of two unbiased estimators, θ_1 and θ_2, of the parameter θ, is defined as:
RE= $\frac{\text{RMSE }(\theta 1)}{\text{RMSE }(\theta 2)}$, Where, RMSE is the Root mean square error, θ_1 is OLS estimator and θ_2 is any other estimator. If the relative efficiency is 1, then it means that the estimator is as efficient as the OLS if the error distribution of the data is normal. A relativeefficiency of 1.2 for instance, implies the estimator is 20% more efficient than OLS estimator.
The RMSE of the regression parameters estimator of β=($\beta_{10}, \beta_{11}, \beta_{12}, \beta_{20}, \beta_{21}, \beta_{22}$) is calculated as:
RMSE=$\frac{1}{K}\sqrt{\sum_{i=1}^{2}\sum_{j=0}^{2}(\beta_{ijk} - \hat{\beta}_{ijk})^2}$, with K = 1000 the number of replications.

VII. Discussion Of Results

Table 1, present the empirical results of parameter estimates under RMSE criterion with Relative Efficiency (RE) when error terms is normally distributed with constant variance at various levels of contemporaneous correlation and different sample sizes

When there is low or absence of contemporaneous correlation ($\rho = 0.0$ or 0.2) between the error terms in different equations with constanterror terms is the best estimator is OLS at small sample sizes (n $\leq$ 20) using RMSE criterion. The OLS estimator performs better than FGLS, Huber estimator and BISQ estimator in small sample size under low contemporaneous correlation.

When the contemporaneous correlation is moderate or high ($\rho \geq 0.5$) at various sample sizes n≤500 with homoscedastic error term using RMSE criterion, the performance of FGLS improves as the contemporaneous correlation increases. In small sample, the relative efficiency of FGLS to OLS ranges between 5% to 51% and even higher in moderate (n= 50 &100) to large sample(n= 250 &500). However, OLS is consistently performed better than BISQ and Huber estimators.

Table 1: Empirical Result of Estimators for Homoscedastic Normal Error Model

		CORRELATION **ρ=0.0**		CORRELATION **ρ=0.2**		CORRELATION **ρ=0.5**		CORRELATION **ρ=0.7**		CORRELATION **ρ=0.95**	
		RMSE	**RE**	**RMSE**	**RE**	**RMSE**	**RE**	**RMSE**	**RE**	**RMSE**	**RE**
20	OLS	4.4535		4.7095		4.8015		4.8411		4.5769	
	FGLS	4.5317	0.98	4.7931	0.98	4.5699	**1.05**	4.2472	**1.14**	1.9805	**2.31**
	HUBER	4.5850	0.97	4.9131	0.96	4.9843	0.96	5.0528	0.96	4.7537	0.96
	BI-SQ	4.6445	0.96	5.0206	0.94	5.1022	0.94	5.2444	0.92	4.9033	0.93
30	OLS	3.4592		3.7493		3.6893		3.6521		3.7431	
	FGLS	3.5417	0.98	3.7559	**1.00**	3.5398	**1.04**	3.3138	**1.10**	1.4901	**2.51**
	HUBER	3.6008	0.96	3.8617	0.97	3.8215	0.97	3.7446	0.98	3.8935	0.96
	BI-SQ	3.6403	0.95	3.9103	0.96	3.8548	0.96	3.8002	0.96	3.8980	0.96
50	OLS	2.7268		2.8150		2.7688		2.8695		2.6579	
	FGLS	2.7414	0.99	2.7787	**1.01**	2.5514	**1.09**	2.5789	**1.11**	1.0673	**2.49**
	HUBER	2.7573	0.99	2.8281	**1.00**	2.7740	1.00	2.9087	0.99	2.7160	0.98
	BI-SQ	2.7806	0.98	2.8354	0.99	2.7831	0.99	2.9227	0.98	2.7230	0.98
100	OLS	1.9524		1.9006		1.9079		2.0032		1.7647	
	FGLS	1.9490	**1.00**	1.8660	**1.02**	1.7360	**1.10**	1.7549	**1.14**	0.6933	**2.55**
	HUBER	1.9931	0.98	1.9162	0.99	1.9354	0.99	2.0321	0.99	1.8079	0.98
	BI-SQ	1.9919	0.98	1.9240	0.99	1.9391	0.98	2.0343	0.98	1.8046	0.98
250	OLS	1.1915		1.2089		1.2068		1.2374		1.1745	
	FGLS	1.1885	**1.00**	1.1876	**1.02**	1.1084	**1.09**	1.1126	**1.11**	0.4579	**2.56**
	HUBER	1.2220	0.98	1.2521	0.97	1.2558	0.96	1.2833	0.96	1.2055	0.97
	BI-SQ	1.2198	0.98	1.2486	0.97	1.2504	0.97	1.2876	0.96	1.2031	0.98
500	OLS	0.8266		0.8239		0.8304		0.8466		0.8475	
	FGLS	0.8267	**1.00**	0.8067	**1.02**	0.7634	**1.09**	0.7585	**1.12**	0.2995	**2.83**

	HUBER	0.8331	0.99	0.8540	0.96	0.8657	0.96	0.8801	0.96	0.8738	0.97
	BI-SQ	0.8330	0.99	0.8522	0.97	0.8655	0.96	0.8787	0.96	0.8719	0.97

From Table 2, the results showed that with sample sizes n=20 &30 at low correlation (ρ =0.2), OLS has the minimum RMSE in estimating the model parameters compared to other estimators. The relative efficiency of all other estimator (compared with OLS) is less than 1. The OLS is the most preferred estimator while FGLS and Huber estimators show equivalent behavior.

At moderate sample size n=50& 100 with low Cc all the estimators are consistent and their RMSE is almost the same. However, FGLS remain the most efficient.

Also, when the correlation is moderate or high (ρ≥0.5), the FGLS has the minimum RMSE with efficiency gain of 4% (RE=1.04) over OLS estimator. Meanwhile the performances of Huber and BISQ estimator are equivalent with that of OLS most especially when the sample size is large (n≥50).

However, the RMSE of FGLS increases tremendously as the Contemporaneous correlation increases. In particular, when Cc is high (ρ=0.95), the relative efficiency of FGLS is 2 to 2.6 times that of OLS estimator.

Table 2: Empirical Result of Estimators for Exponential Heteroscedastic for Normal Error Model

		CORRELATION **ρ=0.0**		CORRELATION **ρ=0.2**		CORRELATION **ρ=0.5**		CORRELATION **ρ=0.7**		CORRELATION **ρ=0.95**	
		RMSE	**RE**	**RMSE**	**RE**	**RMSE**	**RE**	**RMSE**	**RE**	**RMSE**	**RE**
20	OLS	10.0200		10.4146		10.6142	-	10.7729	-	10.1354	-
	FGLS	10.1960	0.98	10.6635	0.98	10.2235	**1.04**	9.4879	**1.14**	4.7483	**2.13**
	HUBER	10.2541	0.98	10.7502	0.97	10.8785	0.98	11.0847	0.97	10.5437	0.96
	BI-SQ	10.3843	0.96	10.9863	0.95	11.1206	0.95	11.5202	0.94	10.7944	0.94
30	OLS	7.7712		8.3260	-	8.2051	-	8.1789	-	8.2941	-
	FGLS	7.9676	0.98	8.3994	0.99	7.9191	**1.04**	7.4135	**1.10**	3.4593	**2.40**
	HUBER	7.9757	0.97	8.4810	0.98	8.3834	0.98	8.2812	0.99	8.5076	0.97
	BI-SQ	8.0965	0.96	8.5565	0.97	8.4880	0.97	8.4239	0.97	8.5379	0.97
50	OLS	6.1907		6.2732	-	6.1810	-	6.4304	-	5.9509	-
	FGLS	6.2386	0.99	6.1816	**1.01**	5.6953	**1.09**	5.7721	**1.11**	2.5192	**2.36**
	HUBER	6.1332	**1.01**	7.2331	0.87	6.1274	1.01	6.4262	1.00	6.0101	0.99
	BI-SQ	6.1511	**1.01**	6.2347	**1.01**	6.1068	1.01	6.4551	1.00	6.0090	0.99
100	OLS	4.3631		4.2235	-	4.2589	-	4.4930	-	3.9530	-
	FGLS	4.3636	**1.00**	4.1370	**1.02**	3.8753	**1.10**	3.9459	**1.14**	1.6292	**2.43**
	HUBER	4.3797	**1.00**	4.2187	1.00	4.2603	1.00	4.4716	1.00	3.9966	0.99
	BI-SQ	4.3902	0.99	4.2593	0.99	4.2996	0.99	4.5010	1.00	4.0142	0.98
250	OLS	2.6025		2.6958	-	2.7006	-	2.7543	-	2.6092	-
	FGLS	2.5993	**1.00**	2.7190	0.99	2.4966	**1.08**	2.4816	**1.11**	1.0975	**2.38**
	HUBER	2.6276	0.99	2.7456	0.98	2.7623	0.98	2.8152	0.98	2.6573	0.98
	BI-SQ	2.6550	0.98	2.7467	0.98	2.7554	0.98	2.8312	0.97	2.6625	0.98
500	OLS	1.8587		1.8388	-	1.8490	-	1.8984	-	1.8820	-
	FGLS	1.8591	**1.00**	1.8000	**1.02**	1.6988	**1.09**	1.7045	**1.11**	0.7199	**2.61**
	HUBER	1.8552	**1.00**	1.8763	0.98	1.8972	0.97	1.9371	0.98	1.9199	0.98
	BI-SQ	1.8556	**1.00**	1.8785	0.98	1.9068	0.97	1.9400	0.98	1.9322	0.97

Table 3 shows the simulation results for estimating the model parameters when error is linearly heteroscedasticity at different levels of contemporaneous correlation. With low correlation (ρ=0.2), OLS estimator maintained minimum RMSE at sample sizes 20 & 30. At sample sizes n≥50, all the estimators performed equally best. FGLS has the least RMSE with relative efficiency of 102% compared to OLS estimator. However, the efficiency of Huber and BISQ relative to OLS is 101%, 100% and 101%, 101% respectively for sample sizes 50 & 100.

When the Cc is high (ρ≥0.5) and error is linearly heteroscedastic the RMSE for FGLS increases drastically with efficiency gain of 2 to 2.4 times that of OLS.The Huber and BISQ compete favorably well with OLS at small sample and equivalent at sample sizes n≥50.

Table 3: Empirical Result of Estimators for Linear Heteroscedastic Normal Error Model

		CORRELATION **ρ=0.0**		CORRELATION **ρ=0.2**		CORRELATION **ρ=0.5**		CORRELATION **ρ=0.7**		CORRELATION **ρ=0.95**	
		RMSE	**RE**	**RMSE**	**RE**	**RMSE**	**RE**	**RMSE**	**RE**	**RMSE**	**RE**
20	OLS	3.5736		3.6955		3.7673		3.8308		3.5980	
	FGLS	3.6355	0.98	3.7899	0.98	3.6421	**1.03**	3.3841	**1.13**	1.7500	**2.06**
	HUBER	3.6337	0.98	3.7909	0.97	3.8325	0.98	3.9084	0.98	3.7329	0.96
	BI-SQ	3.6977	0.97	3.8902	0.95	3.9294	0.96	4.0763	0.94	3.8183	0.94
30	OLS	2.7706		2.9567		2.9150		2.9098		2.9428	

	FGLS	2.8416	0.98	2.9871	0.99	2.8179	**1.03**	2.6414	**1.10**	1.2693	**2.32**
	HUBER	2.8232	0.98	2.9877	0.99	2.9582	0.99	2.9325	0.99	3.0002	0.98
	BI-SQ	2.8699	0.97	3.0227	0.98	3.0024	0.97	2.9852	0.97	3.0121	0.98
50	OLS	2.2077		2.2312		2.2001		2.2899		2.1174	
	FGLS	2.2266	0.99	2.1981	**1.02**	2.0288	**1.08**	2.0581	**1.11**	0.9315	**2.27**
	HUBER	2.1713	**1.02**	2.1999	1.01	2.1640	1.02	2.2728	1.01	2.1222	1.00
	BI-SQ	2.1728	**1.02**	2.2054	1.01	2.1598	1.02	2.2856	1.00	2.1277	1.00
100	OLS	1.5531		1.4992		1.5133		1.5978		1.4065	
	FGLS	1.5542	**1.00**	1.4681	**1.02**	1.3786	**1.10**	1.4062	**1.14**	0.6049	**2.33**
	HUBER	1.5476	**1.00**	1.4876	1.01	1.5023	1.01	1.5766	1.01	1.4107	1.00
	BI-SQ	1.5516	**1.00**	1.5046	1.00	1.5198	1.00	1.5887	1.01	1.4222	0.99
250	OLS	0.9210		0.9574		0.9604		0.9779		0.9273	
	FGLS	0.9202	**1.00**	0.9432	**1.02**	0.8899	**1.08**	0.8833	**1.11**	0.4092	**2.27**
	HUBER	0.9219	**1.00**	0.9675	0.99	0.9737	0.99	0.9912	0.99	0.9393	0.99
	BI-SQ	0.9350	0.99	0.9702	0.99	0.9736	0.99	0.9994	0.98	0.9435	0.98
500	OLS	0.6627		0.6532		0.6564		0.6743		0.6674	
	FGLS	0.6628	**1.00**	0.6398	**1.02**	0.6042	**1.09**	0.6070	**1.11**	0.2688	**2.48**
	HUBER	0.6574	**1.01**	0.6620	0.99	0.6688	0.98	0.6835	0.99	0.6768	0.99
	BI-SQ	0.6582	**1.01**	0.6637	0.98	0.6736	0.97	0.6854	0.98	0.6828	0.98

Table 4 shows similar picture as Table 1.2a does under low contemporaneous correlation (0.2) when error is linearly Heteroscedastic. OLS is the most efficient estimator. The relative efficiency of other estimator is less than 100% indicating OLS minimum RMSE.

However, as the sample size increases (n≥50), all the estimators performed almost equally with FGLS being the most efficient estimator showing relatively minimum RMSE. While the RE of other estimator reduces, FGLS maintained relative efficiency of 102%.

When contemporaneous correlation is moderate (ρ=0.5) under square root heteroscedasticity, OLS sill retained its efficiency at small sample size but OLS performance decreases with increase in contemporaneous correlation. Meanwhile the efficiency of FGLS increases with an increase in contemporaneous correlation.

For sample sizes n≥50 and ρ=0.5, Huber and BISQ estimator behave equally well but below the OLS. However, FGLS remain the best and most efficient estimator.

Table 4: Empirical Result of Estimators for Square-Root Heteroscedastic Normal Error Model

		CORRELATION **ρ=0.0**		CORRELATION **ρ=0.2**		CORRELATION **ρ=0.5**		CORRELATION **ρ=0.7**		CORRELATION **ρ=0.95**	
		RMSE	**RE**	**RMSE**	**RE**	**RMSE**	**RE**	**RMSE**	**RE**	**RMSE**	**RE**
20	OLS	3.9519		4.1325	-	4.2131	-	4.2655	-	4.0172	-
	FGLS	4.0213	0.98	4.2239	0.98	4.0395	**1.04**	3.7457	**1.14**	1.8043	**2.23**
	HUBER	4.0668	0.97	4.2897	0.96	4.3487	0.97	4.4163	0.97	4.1909	0.96
	BI-SQ	4.1029	0.96	4.3686	0.95	4.4205	0.95	4.5822	0.93	4.2761	0.94
30	OLS	3.0663		3.2968	-	3.2467	-	3.2265	-	3.2888	-
	FGLS	3.1420	0.98	3.3177	0.99	3.1259	**1.04**	2.9233	**1.10**	1.3223	**2.49**
	HUBER	2.5519	**1.20**	3.3775	0.98	3.3412	0.97	3.2892	0.98	3.3935	0.97
	BI-SQ	3.2055	0.96	3.4095	0.97	3.3697	0.96	3.3396	0.97	3.4025	0.97
50	OLS	2.4282		2.4832	-	2.4455	-	2.5397	-	2.3502	-
	FGLS	2.4456	0.99	2.4481	**1.01**	2.2513	**1.09**	2.2784	**1.11**	0.9551	**2.46**
	HUBER	2.4293	**1.00**	2.4795	1.00	2.4395	1.00	2.5545	0.99	2.3915	0.98
	BI-SQ	2.4382	**1.00**	2.4812	1.00	2.4331	1.01	2.5656	0.99	2.3875	0.98
100	OLS	1.7240		1.6711	-	1.6821	-	1.7717	-	1.5594	-
	FGLS	1.7232	**1.00**	1.6386	**1.02**	1.5300	**1.10**	1.5519	**1.14**	0.6177	**2.52**
	HUBER	1.7422	0.99	1.6806	0.99	1.6963	0.99	1.7813	0.99	1.5910	0.98
	BI-SQ	1.7459	0.99	1.6890	0.99	1.7035	0.99	1.7853	0.99	1.5885	0.98
250	OLS	1.0368		1.0650	-	1.0652	-	1.0889	-	0.9915	-
	FGLS	1.0350	**1.00**	1.0476	**1.02**	0.9817	**1.09**	0.9788	**1.11**	0.4110	**2.41**
	HUBER	1.0547	0.98	1.0937	0.97	1.0991	0.97	1.1222	0.97	1.0550	0.94
	BI-SQ	1.0591	0.98	1.0916	0.98	1.0944	0.97	1.1260	0.97	1.0558	0.94
500	OLS	0.7327		0.7263	-	0.7309	-	0.7482	-	0.7446	-
	FGLS	0.7328	**1.00**	0.7112	**1.02**	0.6715	**1.09**	0.6708	**1.12**	0.2677	**2.78**
	HUBER	0.7353	**1.00**	0.7472	0.97	0.7561	0.97	0.7707	0.97	0.7647	0.97
	BI-SQ	0.7350	**1.00**	0.7470	0.97	0.7580	0.96	0.7707	0.97	0.7665	0.97

Table 5 presents the RMSE and RE for the model parameters when error is normally distributed and the variance structure for the error is quadratic. At ρ=0.2 under quadratic form of heteroscedasticity, the results reveal that neither OLS nor FGLS could maintain the small sample property as it does under other forms of heteroscedasticity. However, Huber estimator outperforms all other estimator showing relative minimum RMSE at sample sizes n≤30. At sample sizes n≥50, Huber, BISQ and FGLS perform almost in similar manner with efficiency gain of 5%, 4% and 2%, for n=50, 5%, 3% and 2%, for n=100, 4%, 1% and 1% for n=250 and 4%, 2% and 2%, for n=500 respectively, relative to OLS. Therefore at ρ=0.2, quadratic heteroscedastic error terms and n≥50, Huber is the most efficient estimator.

When Cc is moderate (ρ=0.5), the FGLS and Huber estimator perform equally well with both estimators showing relatively minimum RMSE. Most importantly at sample size n≤30, the efficiency of FGLS and Huber relative to OLS are the same (102% and 103% for sample sizes 20 & 30 respectively. Meanwhile OLS performs better than BISQ. At sample sizes (n≥50), the performance of BISQ estimator improves with increase in sample size. The performances of FGLS, Huber and Bi-SQ are almost equivalent but FGLS is the most efficient with efficiency gain of 6 to 9% relative to OLS.

With high Cc (ρ ≥0.7), the performances of Huber and BISQ is very encouraging under Quadratic heteroscedasticity than under Linear, Exponential and Square Root Heteroscedasticity. The two estimators performed better than OLS while FGLS is the most efficient estimator. However, the gain in efficiency of FGLS relatively to OLS increases at relative slow rate in Quadratic heteroscedasticity at ρ=0.95 than in other forms of heteroscedasticity.

VIII. Conclusion

In conclusion, the study revealed that level of contemporaneous correlation and sample sizes affects the performances of the estimators considered. Therefore, we recommend OLS estimator for estimating the parameters of SUR model with heteroscedastic error terms when there is low or absence of Cc with small sample size except when there is quadratic heteroscedastic error terms where Huber estimator is recommended in small sample. However, with moderate to high Cc regardless of sample sizes FGLS is the efficient estimator of SUR model with heteroscedsatic error terms.

Table 5: Empirical Result of Estimators for Quadratic Heteroscedastic Normal Error Model

		CORRELATION **ρ=0.0**		CORRELATION **ρ=0.2**		CORRELATION **ρ=0.5**		CORRELATION **ρ=0.7**		CORRELATION **ρ=0.95**	
		RMSE	**RE**	**RMSE**	**RE**	**RMSE**	**RE**	**RMSE**	**RE**	**RMSE**	**RE**
20	OLS	3.0700		3.1123		3.1733		3.2547		3.0511	
	FGLS	3.1223	0.98	3.2049	0.97	3.1195	1.02	2.9447	**1.11**	1.8609	**1.64**
	HUBER	2.9966	**1.02**	3.0809	**1.01**	3.1246	1.02	3.2184	1.01	3.0790	0.99
	BI-SQ	3.1210	0.98	3.2400	0.96	3.2532	0.98	3.3741	0.96	3.1854	0.96
30	OLS	2.3816		2.5113		2.4837		2.5031		2.4812	
	FGLS	2.4452	0.97	2.5483	0.99	2.4188	**1.03**	2.2998	**1.09**	1.3165	**1.88**
	HUBER	2.3382	**1.02**	2.4356	**1.03**	2.4142	**1.03**	2.4118	1.04	2.4406	1.02
	BI-SQ	2.4040	0.99	2.5083	1.00	2.4967	0.99	2.4830	1.01	2.4746	1.00
50	OLS	1.9256		1.8945		1.8723		1.9568		1.8089	
	FGLS	1.9449	0.99	1.8659	1.02	1.7426	**1.07**	1.7798	**1.10**	0.9945	**1.82**
	HUBER	1.8053	**1.07**	1.8003	**1.05**	1.7581	1.06	1.8744	1.04	1.7300	1.05
	BI-SQ	1.8103	1.06	1.8175	1.04	1.7794	1.05	1.8929	1.03	1.7763	1.02
100	OLS	1.3294		1.2747		1.2933		1.3695		1.2080	
	FGLS	1.3324	1.00	1.2458	1.02	1.1877	**1.09**	1.2281	**1.12**	0.6533	**1.85**
	HUBER	1.2658	**1.05**	1.2087	**1.05**	1.2213	1.06	1.2754	1.07	1.1546	1.05
	BI-SQ	1.2781	1.04	1.2410	1.03	1.2562	1.03	1.3018	1.05	1.1919	1.01
250	OLS	0.7702		0.8176		0.8239		0.8342		0.7963	
	FGLS	0.7709	1.00	0.8084	1.01	0.7768	**1.06**	0.7685	**1.09**	0.4571	**1.74**
	HUBER	0.7348	**1.05**	0.7871	**1.04**	0.7930	1.04	0.8036	1.04	0.7724	1.03
	BI-SQ	0.7625	1.01	0.8062	1.01	0.8100	1.02	0.8217	1.02	0.7872	1.01
500	OLS	0.5718		0.5574		0.5594		0.5774		0.5662	
	FGLS	0.5721	1.00	0.5468	1.02	0.5193	**1.08**	0.5273	**1.10**	0.3041	**1.86**
	HUBER	0.5406	**1.06**	0.5376	**1.04**	0.5458	1.02	0.5573	1.04	0.5492	1.03
	BI-SQ	0.5485	1.04	0.5452	1.02	0.5549	1.01	0.5658	1.02	0.5612	1.01

References

[1]. Aitken, A. C., 1934-35, On Least-Squares and Linear Combination of Observations. *Proceedings of the Royal Society of Edinburgh. 55,* 42 - 48.
[2]. Amemiya, T. (1977). A Note on a Heteroscedastic Model. *Journal of Econometrics, 6,* 365 – 370.
[3]. Avery, R. B. (1977). Error Components and Seemingly unrelated regressions. *Econometrica45* , 199-209.
[4]. Baltagi, B. H. (1980). On seemingly unrelated regressions with error Components. *Econometrica, 48*(6), 1547-1551.
[5]. Carroll, R. J. (1982). Adapting for Heteroscedasticity in Linear models. *Annals of Statistics, 10*(4), 1224-1233.
[6]. Guilkey, D. K. and P. Schmidt. (1973) Estimation of Seemingly Unrelated Regressions with Vector Autoregressive Errors, Journal of the American Statistical Association. 68, 642-647.
[7]. Harvey, A. C. (1976). Estimating Regression Model with Multiplicative heteroscedasticity, *Econometrica, 44(3)* , 461-465.
[8]. Huber, P. J. (1981). Robust Statistics,Wiley,NewYork.

[9]. Huber, P.J. (1973). Robust Regression:Asymptotics, Conjectures and Monte Carlo, *Annals of Statitics, 1(5);* 799-821.
[10]. Judge, G. G., Hill, R.C., Griffiths, W. E. and Lutkepohl, H. (1988). Introduction to theory and practice of Econometrics, *2nd ed. New York ,Wiley*, 444-468.
[11]. Kakwani, N. C. (1967). The unbiasedness of Zellner's Seemingly Unrelated Regression Equations Estimators. *Journal of the American Statistical Association,* 62: 141 – 142.
[12]. Kmenta, J. and Gilbert, R. F., (1970). Estimation of seemingly unrelated regressions with autoregressive disturbances. *Journal of the American Statistical Association*, *65*, 186-197.
[13]. Kmenta, J. and Gilbert, R. F., (1968). Small Sample Properties of Alternative Estimators of Seemingly Unrelated Regression. *Journal of American Statistical Association, 63*, 1180-1200.
[14]. Kramer, W., (1980).Finite Sample Efficiency of Ordinary Least Squares in the Linear Regression Model with Autocorrelated Errors.*Journal of the American Statistical Association, 75(372)*: 1005-1009.
[15]. Mandy, D. M. and Martins-Filho, C., (1993). Seemingly Unrelated Regressions Under Additive Heteroscedasticity: Theory and share equation applications. *Journal of Econometrics, 58*, 315 – 346.
[16]. Maronna, R. A., Martin, R. D. and Yohai, V. J. (2006). Robust Statistics: Theory and Methods. John Wiley and Sons Ltd, England.
[17]. Messemer, C. and Parks, R. W., (2004). Bootstrap Methods for Inference in a SUR model with Autocorrelated Disturbances.University of Washington Economics Working Paper No. UWEC-2004-24
[18]. Olanrewaju, S.O.(2013). Effect of Multicollinearity and Autocorrelation on some Estimators of System of Regression equations. Unpublished PhD thesis Submitted to Department Of Statistics, University of Ilorin.
[19]. R Development Core Team (2008). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
[20]. Rousseeuw, P. J. and Yohai, V. J. (1984). Robust Regression by Mean of S Estimators, Robust and Nonlinear Time Series Analysis, New York, 256-274, doi: 10.1007/978-1-4615-7821-5-15.
[21]. Singh, B. and Ullah, A. (1974). Estimation of seemingly unrelated regressions with random coefficients. *Journal of the American Statistical Association,* 69 , 191-195.
[22]. Sparks, R. ,(2004). SUR Models Applied to an Environmental situation with missing Data and Censored Values. *Journal of Applied mathematics and Decision Sciences. 8*, 15-32.
[23]. Srivastava,V. K. and Maekawa, K.,(1995). Efficiency properties of feasible generalized least squares estimators in SURE models under non-normal.*Journal of Econometrics, 66*, 99-121. https://doi.org/10.1016/0304-4076(94)01609-4
[24]. Takada, H., Ullah, A. and Chen, Yu-Min (1995).Estimation of seemingly unrelated regression model when the error covariance matrix is singular. *Journal of the Applied Statistics, 22(4), 517-530*
[25]. Yahya W. B., Adebayo, S. B., Jolayemi, E.T., Oyejola, B. A. and Sanni, O.O.M.(2008). Effects of non-orthogonally on the efficiency of seemingly unrelated regression (SUR) models. *InterStat Journal, http://interstat.statjournals.net/*.29.
[26]. Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association, 57*, 348-368.
[27]. Zellner, A. (1963). Estimators for seemingly unrelated regression equations: some exact finite sample results. *Journal of the American Statistical Association 58*: 977-992.
[28]. Zellner, A. and Theil, H. (1962). Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations. *Econometrical,30*, 54 – 78.
[29]. Zellner, A. and Huang, D.S. (1962). Further properties of efficient estimators for seemingly unrelated regression equations. *International Economic Review* 3(3): 300-313.

APPENDIX I

GRAPHICAL DESCRIPTION OF THE PERFORMANCES OF ESTIMATORS CONSIDERED IN THE STUDY

Fig.1.1a: Bar chart showing the Performance of Estimators using RMSE Criterion at different Sample size when Error is Normal,Uncorrelated across equations and the form of Heteroscedasticity is Exponential

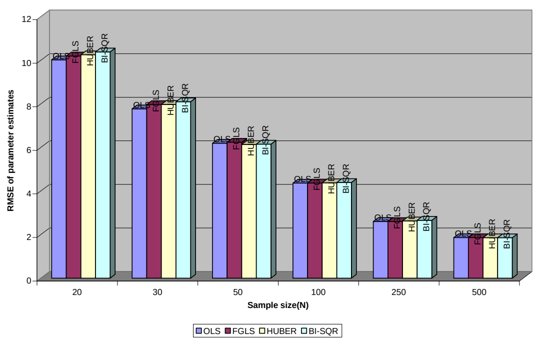


Fig.2.1a: Bar chart showing the Performance of Estimators using RMSE Criterion at different Sample size when Error is Normal,correlated error(R=0.2) and the form of Heteroscedasticity is Exponential

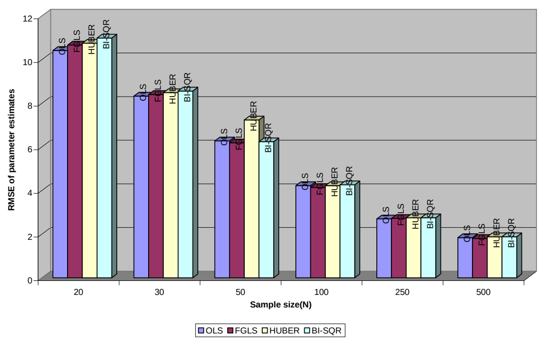


Fig.3.1a: Bar chart showing the Performance of Estimators using RMSE Criterion at different Sample size when Error is Normal,correlated error(R=0.5) and the form of Heteroscedasticity is Exponential

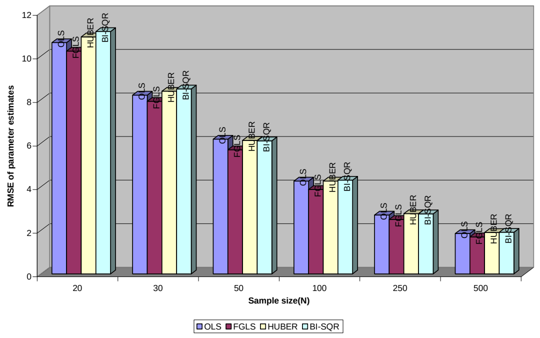


Fig.4.1a: Bar chart showing the Performance of Estimators using RMSE Criterion at different Sample size when Error is Normal,correlated error(R=0.7) and the form of Heteroscedasticity is Exponential

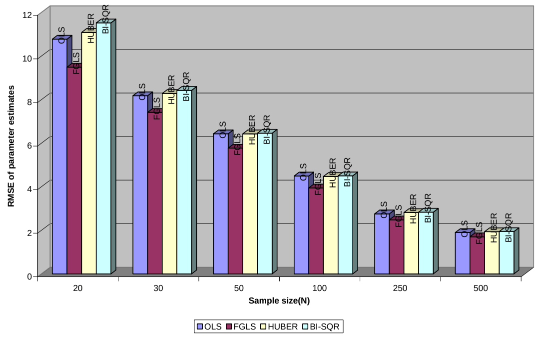


Fig.5.1a: Bar chart showing the Performance of Estimators using RMSE Criterion at different Sample size when Error is Normal,correlated error(R=0.95) and the form of Heteroscedasticity is Exponential

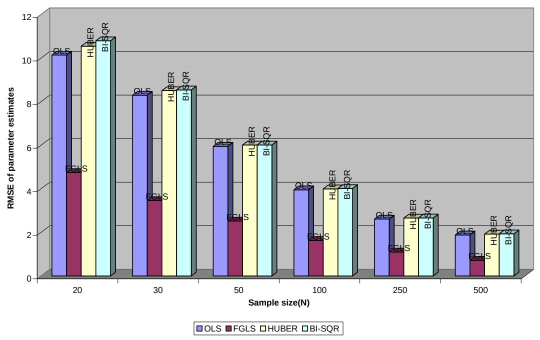


Fig.1.1b:Bar chart showing performance of estimators using RMSE criterion when error is Normal, uncorrelated across the equations and Linearly heteroscedastic(Q=1)

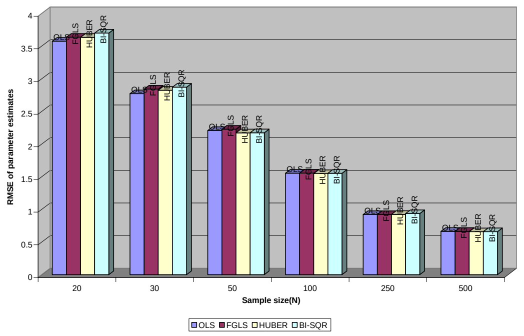


Fig.2.1b:Bar chart showing performance of estimators using RMSE criterion when error is Normal, correlated error (R=0.2) and Linearly heterosedastic(Q=1)

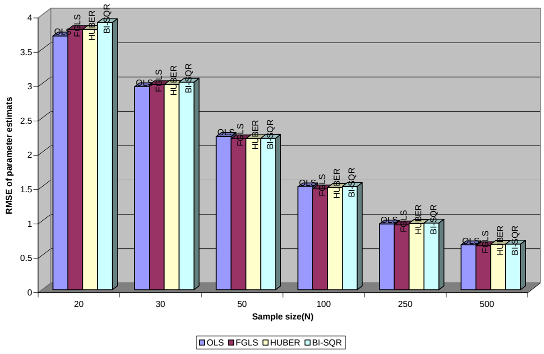


Fig.3.1b:Bar chart showing performance of estimators using RMSE criterion when error is Normal, correlated error (R=0.5) and Linearly heterosedastic(Q=1)

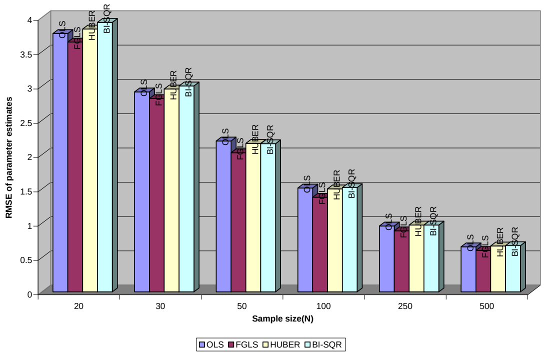


Fig.4.1b:Bar chart showing performance of estimators using RMSE criterion when error is Normal, correlated error (R=0.7) and Linearly heterosedastic(Q=1)

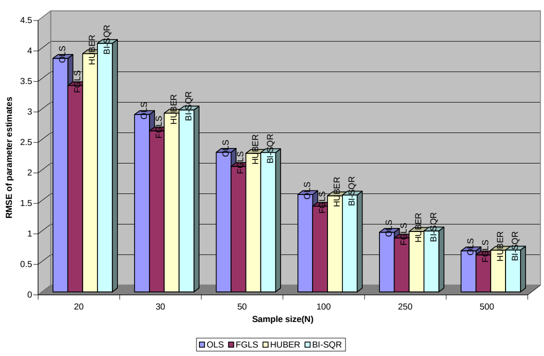


Fig.5.1b:Bar chart showing performance of estimators using RMSE criterion when error is Normal, correlated error (R=0.95) and Linearly heterosedastic(Q=1)

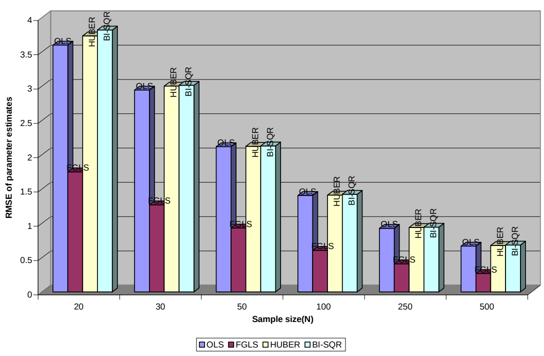


Fig.1.1c: Bar chart showing performance of estimators using RMSE criterion when error is Normal, uncorrelated across the equations and heteroscedastic(Q=0.5)

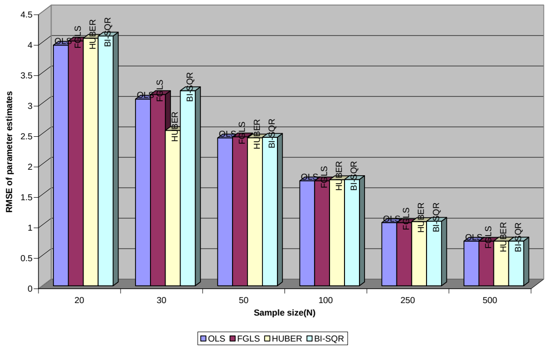


Fig.2.1c:Bar chart showing performance of estimators using RMSE criterion when error is Normal, correlated error(R=0.2) and heteroscedastic(Q=0.5)

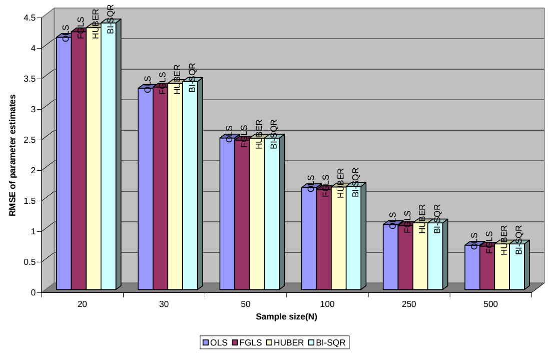


Fig.3.1c:Bar chart showing performance of estimators using RMSE criterion when error is Normal, correlated error(R=0.5) and heteroscedastic(Q=0.5)

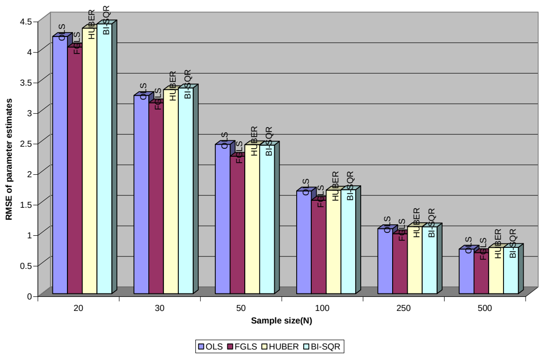


Fig.4.1c:Bar chart showing performance of estimators using RMSE criterion when error is Normal, correlated error(R=0.7) and heteroscedastic(Q=0.5)

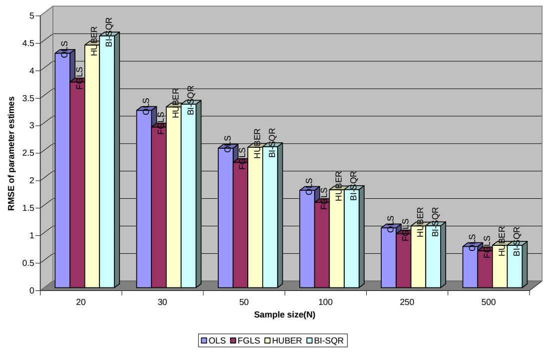


Fig.5.1c:Bar chart showing performance of estimators using RMSE criterion when error is Normal, correlated error(R=0.95) and heteroscedastic(Q=0.5)

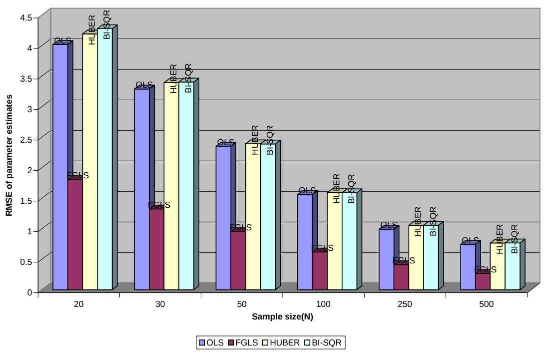


Fig.1.1g:Bar chart showing performance of estimators at diferent level of heteroscedasticity using RMSE criterion when error is Normal, uncorrelated across the equations

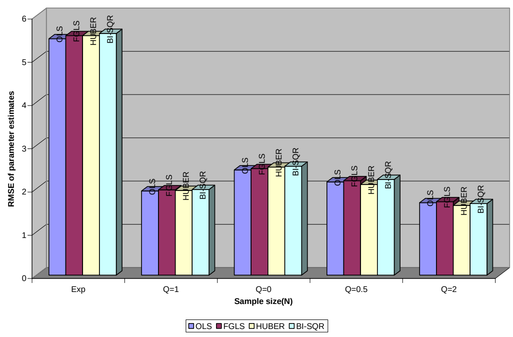


Fig.2.1g:Bar chart showing performance of estimators at diferent level of heteroscedasticity using RMSE criterion when error is Normal, correlated across the equations(R=0.2)

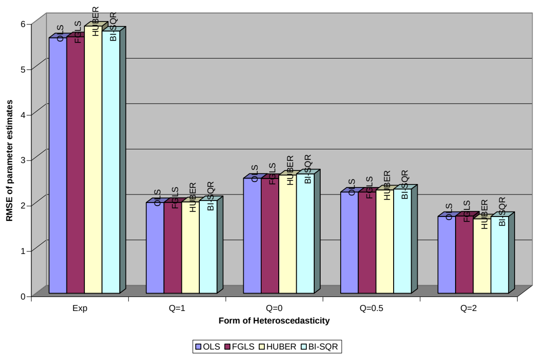


Fig.3.1g:Bar chart showing performance of estimators at diferent level of heteroscedasticity using RMSE criterion when error is Normal, correlated across the equations(R=0.5)

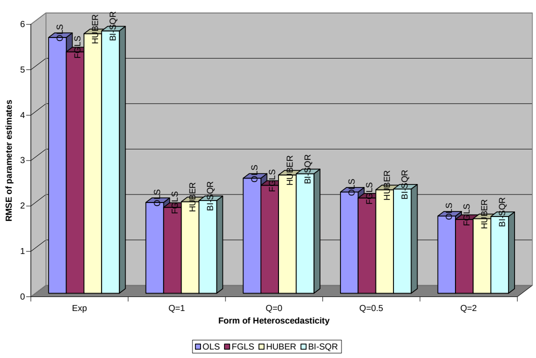


Fig.4.1g:Bar chart showing performance of estimators at diferent level of heteroscedasticity using RMSE criterion when error is Normal, correlated across the equations(R=0.7)

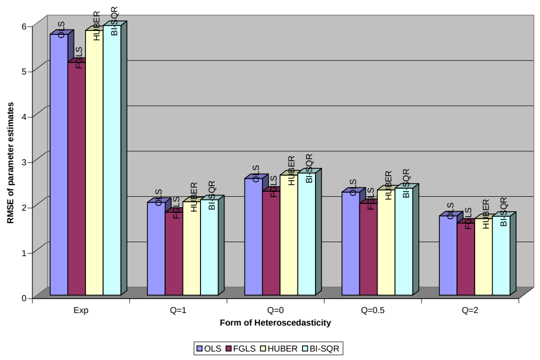
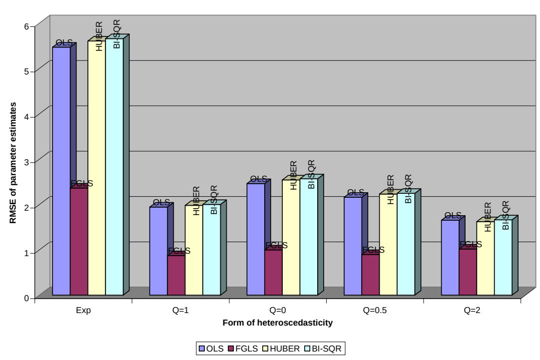
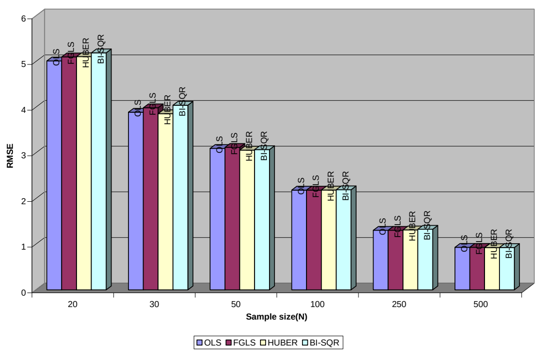


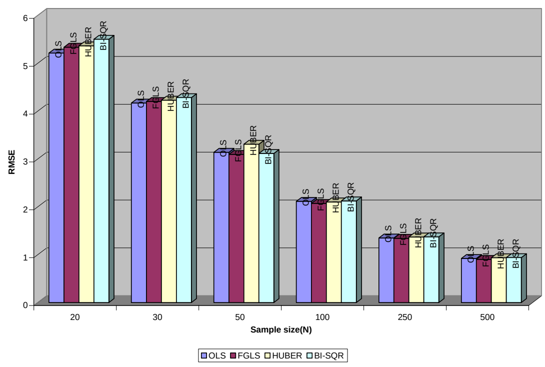
Fig.5.1g:Bar chart showing performance of estimators at diferent level of heteroscedasticity using RMSE criterion when error is Normal, correlated across the equations(R=0.95)



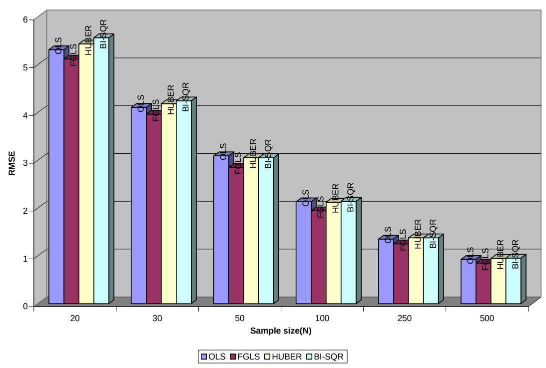
Performance of estimators at different Sample size when Contemporaneous correlation(R=0.0)



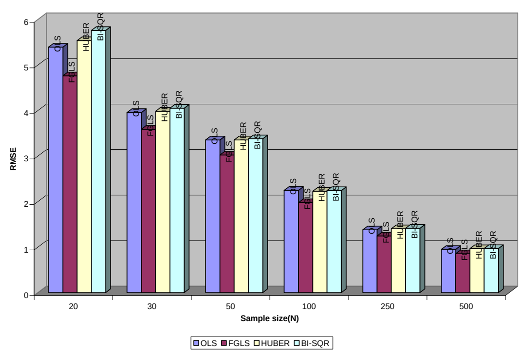
R=0.2



R=0.5



R=0.7



Performances of estimators at different levels of contemporaneous Correlation