J. Basic. Appl. Sci. Res., 6(7)7-14, 2016
ISSN 2090-4304
Journal of Basic and Applied Scientific Research
www.textroad.com

Comparative Analysis of Three-Stage Least Squares and Multivariate Regression Method of Estimating Parameters of Simultaneous Equation Models

Enobong Francis Udoumoh[1*]; Bishop Enebi Olubo-ojo Ohioma[2]; Theresa O. Subeno[3]

[1]Department of Mathematics/Statistics/Computer Science, University of Agriculture, Makurdi, Nigeria.
[2]National Bureau of Statistics, Abuja, Nigeria. E-mail:bishiopohios@yahoo.com
[3] Department of Mathematics and Statistics, Delta State Polytechnic, Ogwashi-Uku, Nigeria.

Received: April 23, 2016
Accepted: June 30, 2016

ABSTRACT

Two sets of simultaneous equation describing the demand and supply of maize using some economic variables were used to compare the Three-Stage Least Squares (3SLS) methods and Multivariate Regression (MVR) of parameter estimation. A sample of empirical data was collected from which other samples were simulated using a normal distribution for $n = 30$, 60 and 100. The findings indicated that the multivariate Regression (MVR) method gives a better estimate of the parameters and has a higher performance when the sample sizes are small (i.e. $n = 12$ and 30), while the Three-Stage Least Squares (3SLS) method gives a better estimate of the parameters and has a better performance when the sample size is large, say $n \geq 60$.
KEYWORDS: Three-Stage Least Squares (3SLS); Multivariate Regression (MVR); Parameter Estimation; simulation.

1 INTRODUCTION

The method of least squares application to a single equation assumes that the explanatory variables are truly exogenous. This means that there is only one way association between the dependent variable Y and the independent variables $X's$. If the association are two ways, that is, if the explanatory variables, $X's$ are also determined by Y, the assumption of ordinary least squares (OLS) which states that the error term, u is independent of the explanatory, X $[E(Xu) = 0]$ will be violated. Hence, the least square method gives biased and also inconsistent estimates, (Koutsoyiannis, 1977). If there is a two-way association in a function, then the function should not be treated in isolation as a single equation model but rather as a wider system of equations which can effectively describe the relationships among all variables. In particular, if $Y = F(X)$ and also $X = F(Y)$, it is not advisable to use a single equation model for the description of the relationship between Y and X. Rather, a multi-equation model which include separate equations which each Y and X appear as endogenous variables, even though they may appear as explanatory variables in other equations of the model. The system describing this joint dependence of variables is called System of Simultaneous Equation. It is therefore of interest and great importance to examine some statistical methods of estimating parameters of the model that contain such variables. Two possible estimation techniques that can be useful in the above context are the three stage least squares (3SLS) and multivariate regression (MVR) methods. The 3SLS is an extension of the two stage least squares (2SLS) method. 2SLS consist of two steps, namely; the estimation of the moment matrix of the reduced form of the simultaneous equations and the estimation of the coefficients of one single structural equation after its reduction. As an extension, the 3SLS uses the 2SLS estimated moment matrix of the structural equation to estimate the coefficients of the entire system simultaneously. When there is more than one dependent variable in a set of multiple regression equations then the result is a multivariate regression (MVR) model. Both 3SLS and MVR models have rich theories and applications in literature. Recently, the multivariate regression methods have been widely applied to predict the quality of red wine based on some chemical and phenolic parameters, (Beaver & Harbertson, 2016; Aleixandre-Tudo et. al. 2015). Elsewhere, MVR has been applied to predict reservoir indicator in oil field management. Kapteyn & Fiebig (1981) derived some necessary and sufficient conditions for the numerical equivalence of the two-stage and three-stage least squares (3SLS) estimators in a linear simultaneous equations model. The efficiency of the 2SLS and 3SLS has been discussed in (Baltagi, 1998).

In this study, an econometric model of two equations shall be built for predicting the quantity of Maize produced and quantity of Maize consumed; using their predictor variables like Price (P_t), price of Maize substitute

***Corresponding Author:** Enobong Francis Udoumoh, Department of Mathematics/Statistics/Computer Science, University of Agriculture, Makurdi, Nigeria. E-mail: uenobong@gmail.com Mobile Phone: +2347032364395

(P_s), lagged price of Maize (M_{t-1}) and investment expenditure on Maize (I_t). We will estimate the parameters of the structural equations of the simultaneous model using multivariate regression (MVR) method and the three-stage least squares (3SLS) regression method . We will also assess the asymptotic properties of the two estimation methods based on empirical evidence. In conclusion, a comparative analysis of MVR and 3SLS methods will be carried out.

2. MODEL EQUATIONS

The demand and supply equations are respectively given as:

$D_t = \beta_0 + \beta_1 S_t + \beta_2 P_t + \beta_3 P_s + U_t$

and

$S_t = \alpha_0 + \alpha_1 P_t + \alpha_2 P_{t-1} + \alpha_3 I_s + V_t$

Where the endogenous variables are:

D_t is the quantity of maize demanded/consumed in thousand metric tones

S_t is the quantity of Maize supplied/produced in thousand metric tones

The exogenous variables are:

P_t is the price of Maize

P_s is the price of Maize substitute (wheat)

P_{t-1} is the lag price of Maize

I_t is the investment expenditure on Maize

Where, $\beta_0, \beta_1, \beta_2, \beta_3, \alpha_0, \alpha_1, \propto_2, \alpha_3$, are structural parameters of the model;

U_t and V_t are the stochastic error terms for the structural equations

The system is a complete simultaneous equation model, and by order condition the model is over-identified; see (Koutsoyiannis, 1977).

2.1: Three-Stage Least Squares

Suppose that we are left with a system of 2-equations in the form:

$$D_t = \beta_0 + \beta_1 S_t + \beta_2 P_t + \beta_3 P_s + U_t \qquad \text{i}$$

$$S_t = \alpha_0 + \alpha_1 P_t + \alpha_2 P_{t-1} + \alpha_3 I_s + V_t \qquad \text{ii}$$

Pre-multiply each equation by the four predetermined variables to obtain a system of 4 x 2 equations, i.e. we have four-forms for each of the two equations.

The set of 4-forms of the first structural equation is:

$$P_t D_t = \beta_0 P_t + \beta_1 P_t S_t + \beta_2 P_t^2 + \beta_3 P_t P_s + P_t U_t$$
$$P_s D_t = \beta_0 P_s + \beta_1 P_s S_t + \beta_2 P_s P_t + \beta_3 P_s^2 + P_s U_t$$
$$P_{t-1} D_t = \beta_0 P_{t-1} + \beta_1 P_{t-1} S_t + \beta_2 P_{t-1}^2 + \beta_3 P_{t-1} P_s + P_{t-1} U_t$$
$$I_t D_t = \beta_0 I_t + \beta_1 I_t S_t + \beta_2 I_t P_t + \beta_3 I_t P_s + I_t U_t$$

The set of 4-forms for the second structural equation is:

$$P_t S_t = \alpha_0 P_t + \alpha_1 P_t^2 + \alpha_2 P_t P_{t-1} + \alpha_3 P_t I_t + P_t V_t$$
$$P_s S_t = \alpha_0 P_s + \alpha_1 P_s + \alpha_2 P_s P_{t-1} + \alpha_3 P_s I_t + P_s V_t$$
$$P_{t-1} S_t = \alpha_0 P_{t-1} + \alpha_1 P_{t-1} P_t + \alpha_2 P_{t-1}^2 + \alpha_3 P_{t-1} I_t + P_{t-1} V_t$$
$$I_t S_t = \alpha_0 I_t + \alpha_1 I_t P_t + \alpha_2 I_t P_{t-1} + \alpha_3 I_t^2 + I_t V_t$$

It can be seen that the disturbances of these equations are heteroscedastic, since the composite random terms $\left[u^{*s} = X_i u_j^{s}, where\ X_i\ are\ exogenous\ variables\right]$ tend to change together with the exogenous variables. Hence, the appropriate method for the estimation of the parameters of the system is generalized least squares. The transformation required involves the variances and the co-variances of the original error terms $u^{,s}$ which however are unknown. We can obtain an estimate of these variance-covariances by first applying the two-Stage Least Squares (2SLS) to each of the structural equation of the original model. Thus we have the following three stages of estimation:

STAGE I: Obtain the reduced form of all the equations of the model

$$D_t = f(P_{t,} P_{t-1,} P_s I_t) \qquad \text{iii}$$

$$S_t = f(P_{t,} P_{t-1,} P_s I_t) \qquad \text{iv}$$

Do OLS on iii and iv and obtain the predicted values; $\widehat{D}_t$ and $\hat{S}_t$

STAGE II: Substitute the value of $\widehat{D}_t$ and $\hat{S}_t$ in the right-hand side of the structural equation i.e. equation (i) where it appears and apply OLS to the transformed equations. We then obtain the two stage least square (2SLS) of $\alpha^{,s}$ and $\beta^{,s}$ which is use for the estimation of the error terms of the two equations (e_{1i} and e_{2i}) each corresponding structural equation i.e. for each equation we have n-values of the error term (n being the sample size). The variance-co-variances of the estimated error terms may easily be computed by the formula:

$$\hat{\sigma}_{e_1}^2 = \frac{\sum_{i=1}^{n} e_{1i}^2}{n},$$

$$\hat{\sigma}_{e_2}^2 = \frac{\sum_{i=1}^{n} e_{2i}^2}{n},$$

$$\hat{\sigma}_{e_1 e_2} = \hat{\sigma}_{e_2 e_1} = \frac{\sum_{i=1}^{n} e_{1i} e_{2i}}{n}$$

The complete set of the variance-covariance of the error terms is as follows:

$$\begin{vmatrix} \hat{\sigma}_{e_1}^2 & \hat{\sigma}_{e_1 e_2} \\ \hat{\sigma}_{e_2 e_1} & \hat{\sigma}_{e_2}^2 \end{vmatrix} = \begin{vmatrix} \frac{\sum_{i=1}^{n} e_{1i}^2}{n} & \frac{\sum_{i=1}^{n} e_{1i} e_{2i}}{n} \\ \frac{\sum_{i=1}^{n} e_{1i} e_{2i}}{n} & \frac{\sum_{i=1}^{n} e_{2i}^2}{n} \end{vmatrix}$$

STAGE III: We use the above variance-covariance of the error terms in order to obtain the transformation of the original variables for the application of the generalized least squares (GLS).

2.2: Multivariate Regression Method

Suppose $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N(\mu, \sum)$ and that

X_1 is p-vector and X_2 is a q-vector

Then the conditional density function of X_1 given that the elements of X_2 are fixed (say X_2) is then defined by:

$g(X_1/X_2 = x_2) = \frac{f(X_1, X_2)}{h(X_2)}$

Where $f(X_1, X_2)$ is the joint density function of X_1 and X_2

$h(X_2)$ is the marginal density function of X_2

$g(X_1/X_2 = x_2) = (2\pi)^{-2P} |\sum_{1.2}|^{-1} \text{Exp-}\frac{1}{2}(X_1 - (\mu_1 + \sum_{12}\sum_{22}^{-1}(X_2 - \mu_2)))' \sum_{1.2}^{-1}(X_1 - (\mu_1 + \sum_{12}\sum_{22}^{-1}(X_2 - \mu_2)))$

Recall that a univariate normal has a density function given as follows:

$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \text{Exp} - \frac{1}{2\sigma^2}(x - \mu)^2$

If X is p-vector of a univariate normal, then

$f(x) = (2\pi)^{-2P} |\sum|^{-1} \text{Exp-}\frac{1}{2}(X - \mu)' \sum^{-1}(x - \mu)$ v

We then obtain $E(X_1/X_2 = x_2) = \mu_1 + \sum_{12}\sum_{22}^{-1}(X_2 - \mu_2)$

We can conclude that for

i. Univariate case, $E(Y/X = x) = \alpha + \beta X + e$, for simple univariate regression so that
$E(Y/X_1 = x_1, X_2 = x_2 \dots X_p = x_p) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + e$
is a multiple regression.

ii. Multivariate case $E(\underline{Y}/\underline{X})$ is given as

$$E[\underline{X_1}/\underline{X_2} = \underline{x_2}] = \mu_1 + \sum_{12} + \sum_{22}^{-1}(X_2 - \mu_2)$$
$$= \mu_1 - \sum_{12}\sum_{22}^{-1}\mu_2 + \sum_{12}\sum_{22}^{-1}X_2 \qquad \text{vi}$$

Meaning that $E[\underline{X_1}/\underline{X_2} = \underline{x_2}] = \alpha + \beta X_2$

$$\hat{Y} = \alpha + \beta X \qquad \text{vii}$$

Comparing vi and vii, we obtained the unbiased estimate of the parameter

$\beta = \sum_{12}\sum_{22}^{-1}$ from an estimator $\hat{\beta} = S_{12}S_{22}^{-1}$ which gives the fitted model for a simultaneous equation as

$$Y_1 = f_1(X_1, X_2 \dots X_P),$$
$$Y_2 = f_2(X_1, X_2 \dots X_P),$$
$$\vdots$$
$$Y_P = f_P(X_1, X_2 \dots X_P)$$

Interested readers may refer to (Hidalgo and Goodman, 2013), (Schervish, 1987) for more information on multivariate regression method and applications.

2.3 Measure of Accuracy of Models

We have adopted two techniques to measure the accuracy of forecast by the 3SLS and MVR, namely; Theil's Inequality Coefficient and χ^2 – Statistic Performance Estimability. Theil's Inequality Coefficient, denoted by U, is a systematic measure of the accuracy of the forecasts obtained from an econometric model, see (Bliemel, 1973), Leuthold, 1975), (Song, et. al. 2013)). The χ^2 – statistic is another measure of evaluating the performance of an estimated model. It measures the discrepancy between the predicted value and the actual value of an estimated model.

3. APPLICATION AND RESULTS

The data for the study was collected from National Bureau of Statistics Abuja. The Bureau of Statistics provides comprehensive, timely, relevant, responsive and user-focused statistical information relating to the social and economic life as well as conditions of the inhabitants of Nigeria. Data on economic activities of maize, such as production, consumption, price of maize, lagged price of Maize, price of maize substitute and investment expenditure on maize was collected for twelve (12) years. From the sample of size 12, simulation was carried out on the basis of a normally distributed disturbance term using MINITAB and data were generated for values of n = 30, 60, and 100. Further analysis was done using Stata 11.

Table 1: Estimates of parameters of MVR and 3SLS with standard error and p-values for n = 12 (the actual data collected from NBS)

	Parameters		Standard Error		Test Statistic		P- Values	
Demand	MVR	3SLS	MVR	3SLS	MVR=t	3SLS=z	MVR	3SLS
S_t	0.7315	-0.6000	1.1545	1.2935	0.63	0.46	0.544	0.643
P_t	-0.6841	-0.7206	0.4089	0.3672	-1.67	-1.96	0.133	0.050
P_S	0.9831	1.0294	0.4258	0.3699	2.31	2.78	0.050	0.005
Const.	-1952.313	-1624.11	2537.124	2665.331	-0.77	-0.61	0.464	0.542
Supply	MVR	3SLS	MVR	3SLS	MVR=t	3SLS=z	MVR	3SLS
P_t	0.0237	0.0294	0.0269	0.0219	0.88	1.35	0.404	0.178
P_{t-1}	-0.0110	-0.0117	0.0041	0.0033	-2.68	-0.50	0.028	0.000
I_t	-0.0029	-0.0109	0.2428	0.0195	-0.12	-0.56	0.906	0.576
Const.	2773.696	2811.046	387.6421	316.1151	7.16	8.89	0.000	0.000

METHOD I (MVR), n = 12

D_{tm} = -0.7315S_t - 0.6841P_t + 0.9831P_s - 1952.313; R^2 = 0.7432

S_{tm} = 0.0237P_t - 0.0110P_{t-1} - 0.0029I_t + 2773.696; R^2 = 0.6814

METHOD II (3SLS), n = 12

D_{tt} = -0.6000$M_{t,}$ – 0.7206P_t + 0.0294P_s – 1624.11; R^2 = 0.7418

S_{tt} = 0.0294P_t – 0.0117P_{t-1} + 0.0110I_t + 2811.046; R^2 = 0.6748

NOTE:

D_{tm} = Quantity of maize consumed (using multivariate regression method of estimating parameters).
S_{tm} = Quantity of maize produced (using multivariate regression method of estimating parameters).
D_{tt} = Quantity of maize consumed (using three stages least square method of estimating parameters).
S_{tt} = Quantity of maize produced (using three stages least square method of estimating parameters).

Table 2: Model performance using the X^2-statistic and the Theil's inequality coefficient, U; for n =12.

Endogenous Variables	X^2-statistic		U- statistic	
	MVR	3SLS	MVR	3SLS
D_t	2071.24	2101.11	0.2175	0.2895
S_t	145.85	147.06	0.0650	0.0651

CRITICAL VALUES

For $\chi^2 \sim \chi^2_{crit} = \frac{1}{2}(Z_\alpha + \sqrt{2n-1})^2 Z_\alpha = \frac{1}{2}(1.96 + \sqrt{23})^2 \text{x} 1.96 = 86.59$

For U, 0<U<1

Table 3: Estimates of parameters of MVR and 3SLS with standard errors and p-values for n =30

	Parameters		Standard Error		Test Statistic		P- Values	
Demand	MVR	3SLS	MVR	3SLS	MVR=t	3SLS=z	MVR	3SLS
S_t	0.2477	0.3358	0.1198	0.1731	2.07	1.94	0.049	0.052
P_t	0.1069	0.0837	0.0264	0.0254	4.05	3.29	0.000	0.001
P_S	0.1364	0.1285	0.0631	0.0815	2.16	1.58	0.040	0.115
Const.	-1478.79	-1361.55	247.0941	331.1034	-5.98	-4.11	0.000	0.000
Supply	MVR	3SLS	MVR	3SLS	MVR=t	3SLS=z	MVR	3SLS
P_t	-0.0343	-0.0211	0.9944	0.0888	-0.34	-0.24	0.733	0.812
P_{t-1}	-0.0940	0.0794	0.1047	0.0930	0.93	0.85	0.377	0.393
I_t	-0.3379	0.3398	0.5066	0.0470	6.67	7.23	0.000	0.000
Const.	-1153.407	-1161.09	193.7796	179.7474	-5.95	-6.46	0.000	0.000

METHOD I (MVR), n = 30

D_{tm} = -0.2479S_t - 0.1069P_t + 0.1364P_s - 1478.79; R^2 = 0.9895

S_{tm} = -0.0343P_t - 0.0940P_{t-1} + 0.3379I_t – 1153.407; R^2 = 0.9784

METHOD II (3SLS), n = 30

D_{tt} = -0.3358S_t – 0.0837P_t + 0.1285P_s – 1361.55; R^2 = 0.9889

S_{tt} = -0.0211P_t – 0.0794P_{t-1} + 0.3398I_t – 1161.09; R^2 = 0.9784

Table 4: Model performance using the X^2-statistic and the Theil's inequality coefficient, U; for n =30.

Endogenous Variables	X^2-statistic		U- statistic	
	MVR	3SLS	MVR	3SLS
D_t	218.09	230.57	0.0379	0.0388
S_t	420.76	422.24	0.0488	0.0489

CRITICAL VALUES

For $\chi^2 \sim \chi^2_{crit} = \frac{1}{2}(Z_\alpha + \sqrt{2n-1})^2 Z_\alpha = \frac{1}{2}(1.96 + \sqrt{59})^2$x1.96 = 91.0927, for U, 0<U<1

Table 5: Estimates of parameters of MVR and 3SLS with standard errors and p-values for n =60

	Parameters		Standard Error		Test Statistic		P- Values	
Demand	MVR	3SLS	MVR	3SLS	MVR=t	3SLS=z	MVR	3SLS
S_t	0.4730	6.0186	0.4873	0.4851	0.97	3.37	0.336	0.001
P_t	0.2943	-0.1627	0.0466	0.0462	6.32	-1.44	0.000	0.150
P_S	-0.0374	-0.0033	0.0434	0.0488	-0.86	-0.07	0.393	0.946
Const.	-2065.70	-8986.29	608.7939	618.606	-3.39	-4.05	0.001	0.000
Supply	MVR	3SLS	MVR	3SLS	MVR=t	3SLS=z	MVR	3SLS
P_t	0.0598	0.0598	0.0051	0.0049	11.75	12.18	0.000	0.0.000
P_{t-1}	0.0002	0.0001	0.0060	0.0003	0.30	0.33	0.763	0.0740
I_t	-0.2090	-0.2100	0.0143	0.0058	3.47	3.62	0.000	0.000
Const.	1210.741	1210.009	20.6442	19.5288	58.65	61.99	0.000	0.000

METHOD I (MVR), n = 60

D_{tm} = -0.4730S_t + 0.2943P_t - 0.0374P_s - 2065.696; R^2 = 0.9920

S_{tm} = 0.0597P_t + 0.0002P_{t-1} + 0.0210I_t + 1210.671; R^2 = 0.9923

METHOD II (3sls), n = 60

D_{tt} = 6.0186S_t – 0.1627P_t + 0.0033P_s – 8986.293; R^2 = 0.9933

S_{tt} = 0.0598P_t + 0.0001P_{t-1} + 0.2100I_t + 1210.009; R^2 = 0.9923

Table 6: Model performance using the X^2-statistic and the Theil's inequality coefficient, U; for n =60.

Endogenous Variables	X^2-statistic		U- statistic	
	MVR	3SLS	MVR	3SLS
D_t	343.73	326.79	0.0471	0.0307
S_t	24.34	24.27	0.0123	0.0122

CRITICAL VALUES

For $\chi^2 \sim \chi^2_{crit} = \frac{1}{2}(Z_\alpha + \sqrt{2n-1})^2 Z_\alpha = \frac{1}{2}(1.96 + \sqrt{119})^2$x1.96 = 162.2917; for U, 0<U<1

Table 7: Estimates of parameters of MVR and 3SLS with standard errors and p-values for n =100

Demand	Parameters		Standard Error		Test Statistic		P- Values	
	MVR	3SLS	MVR	3SLS	MVR=t	3SLS=z	MVR	3SLS
S_t	1.0822	1.9424	0.8040	0.1060	22.66	18.32	0.000	0.000
P_t	0.1557	0.1467	0.0079	0.0065	27.03	20.96	0.000	0.000
P_S	0.0060	0.0058	0.0224	0.0147	1.37	1.27	1.174	0.204
Const.	-3838.705	11616.63	110.5500	143.5545	-34.72	-27.88	0.000	0.000
Supply	MVR	3SLS	MVR	3SLS	MVR=t	3SLS=z	MVR	3SLS
P_t	-0.0583	-0.0850	0.0342	0.0324	-1.71	-2.62	0.091	0.009
P_{t-1}	0.03064	0.0581	0.0334	0.0307	0.92	1.84	0.361	0.065
I_t	0.1072	0.1064	0.0084	0.0083	12.70	12.87	0.000	0.000
Const.	1351.871	1356.156	26.2617	25.6944	51.48	52.78	0.000	0.000

METHOD I (MVR), n = 100
D_{tm} = $1.0822S_t + 0.1557P_t + 0.0060P_s$ - 3838.705; R^2 = 0.8930
S_{tm} = $0.0583P_t + 0.0304P_{t-1} + 0.1072I_t$ + 1351.871; R^2 = 0.9689
METHOD II (3sls), n = 100
D_{tt} = $1.9424R_t - 0.1467P_t + 0.0058P_s$ + 11616.63; R^2 = 0.9693
S_{tt} = $-0.0850P_t + 0.0581P_{t-1} + 0.1064I_t$ + 1356.156; R^2 = 0.9977

Table 8: Model performance using the X^2-statistic and the Theil's inequality coefficient, U; for n =100.

Endogenous Variables	χ^2 statistic		U- statistic	
	MVR	3SLS	MVR	3SLS
D_t	136.05	81.47	0.1143	0.0169
S_t	176.50	171.58	0.0246	0.0244

CRITICAL VALUES

For $\chi^2 \sim \chi^2_{critical} = \frac{1}{2}\left(Z_\alpha + \sqrt{2n-1}\right)^2 Z_\alpha = \frac{1}{2}\left(1.96 + \sqrt{199}\right)^2$x1.96 = 252.9772

For U, 0<U<1

3.1 Discussion of Results

From tables 1 and 2 above, it can be observed that,

- The data fits the demand equation (i.e. equation (i)) better than supply equation (i.e. equation (ii) for both MVR and 3SLS
- The MVR method gives a higher R^2 - values compared to 3SLS, hence, the model estimated by MVR is a better estimation method than 3SLS
- Most of the parameters are significant though the MVR has better estimate of the parameters than the 3SLS since their standard errors are less than those for 3SLS.
- The Theil's inequality Coefficient and the X^2- statistic revealed that the MVR performed better than the 3SLS for the two equations of the model.

From tables 3 and 4 above, it can be observed that,

- The data fits the demand equation (i.e. equation (i)) better than supply equation (i.e. equation (ii) for both MVR and 3SLS
- The MVR method gives a higher R^2 - values compared to 3SLS, hence, the model estimated by MVR is a better estimation method than 3SLS
- Most of the parameters are significant though the MVR has better estimate of the parameters than the 3SLS since their standard errors are less than those for 3SLS.
- The Theil's inequality Coefficient and the X^2- statistic revealed that the MVR performed better than the 3SLS for the two equations of the model.

From tables 5 and 6 above, it can be observed that the 3SLS compete favourably with the MVR as n gets larger such that

- The R^2 - values for 3SLS is higher or equal to that of MVR for equation I and II respectively, hence, models estimated by 3SLS tend to be slightly more adequate compared to those estimated by MVR as n becomes large
- Most of the parameters are significant, though the 3SLS has a slightly better estimate of the parameters than the MVR since their standard errors are less than those for MVR

- The Theil's inequality Coefficient and the X^2- statistic reveals that as n becomes large, the 3SLS performed slightly better than MVR for the two equations of the model.

From tables 7 and 8 above, it can be observed that the 3SLS compete favourably with the MVR as n gets larger such that

- The R^2 - values for 3SLS is higher or equal to that of MVR for equation I and II respectively, hence, models estimated by 3SLS tend to be slightly more adequate compared to those estimated by MVR as n becomes large
- Most of the parameters are significant, though the 3SLS has a slightly better estimate of the parameters than the MVR since their standard errors are less than those for MVR
- The Theil's inequality Coefficient and the X^2- statistic revealed that as n becomes large, the 3SLS performed slightly better than MVR for the two equations of the model.

4: CONCLUSIONS

The conclusion will be based on the two equations representing demand and supply of maize and the behaviour of the two methods of estimation under study at various sample sizes, i.e. n = 12, 30, 60, and 100 as it relates to model adequacy, significance of the parameters, model predictive power and model performance. Adequacy of a model is measured by R^2 –value that is a higher R^2 –value suggests that the model is adequate.

From the result obtained from the analysis (table 9); it is observed that:

- The R^2 –values for MVR is higher than that of 3SLS for n = 12 and 30
- The R^2 –values for3SLS is higher than that of MVR for n = 60 and 100
- The R^2 –values for the two methods increases as the sample size increases from n = 12, 30, 60 and 100.

The models by MVR method gives a more adequate model compared to that estimated by 3SLS method when the sample size is small, i.e. n=12, 30. On the other hand, the model estimated by 3SLS becomes more adequate when the sample size is large, say n≥60.

Significance of the model parameters can be established by the use of the standard error of the parameter estimates such that a high standard error implies that the model parameters are not significant whereas a lower standard error implies that model parameters are significant.

From the analysis, (tables 10) it is observed that:

- The standard error for MVR is smaller than those for 3SLS for small sample sizes, say, n = 12 and 30.
- The standard error for 3SLS is smaller than those for MVR for larger sample, say, n≥60.
- The standard error for the two methods of estimation (3SLS and MVR) decreases as the sample size increased from 12, 30, 60 to 100

The model estimated by MVR tends to have better significant parameters than that of 3SLS for small sample sizes. On the other hand, if the sample size is large, the parameters of the 3SLS become more significant than that for MVR.

The predictive power of model is measured by how close the predictive value is to the actual value. When the deviation between the actual value and the predictive value is zero then we say that the model has a perfect predictive power. Hence a model whose predictive value is not close to the actual value is said to be less powerful. A close observation of the analysis revealed the followings:

- The predicted values of models estimated by MVR is closer to the actual values for n =12 and 30 than those estimate by 3SLS.
- The predicted values of models estimated by 3SLS is closer to the actual values for n = 60 and 100 than those estimated by MVR.

Models estimated by MVR have a higher predictive power when the sample size is small. On the other hand, if the sample is large, then the model estimated by 3SLS becomes more powerful.

The performance of the two equations of the model was measured using the Theil's Inequality Coefficient and X^2- statistic for the various values of the sample sizes and it revealed that:

- The MVR method performed better than the 3SLS for n =12, and 30
- The 3SLS method performed better than the MVR for n = 60 and 100

The 3SLS will perform better if the sample size is considerably large, say n≥60

In general, it is clear that the economic variables in the two equations of the model are well combined since the R^2 –values are high for all values of n = 12, 30, 60 and 100 for the two methods of estimation considered.

Furthermore, multivariate regression method should be used in parameter estimation for simultaneous equation model if the sample size is small since it gives a better estimate of the parameters; but, if the sample size is

considerably large sample say n≥ 60, and the equations of the model are over-identified, the Three-Stage Least Squares Regression will give a better estimate of the parameter hence it should be used in such estimation.
The Theil,s Inequality Coefficient U and the X^2- statistic are in agreement that MVR is better when the sample size is small say 12≤n≤30 while the 3SLS is better when the sample size is large say $n \geq 60$, hence, they are both ideal in assessing the performance of model estimation techniques (MVR and 3SLS) in a simultaneous equation model.
Finally, Simultaneous equation should be used to describe the relationship among economic variables especially when y = f(x) and x = f(y), where y is the dependent variable and x is the independent variable. Hence when estimating the parameter of a simultaneous equation model, the sample size should be considered in selecting the best method of estimation.

REFERENCES

Aleixandre-Tudo, J. L., Alvarez, I., Garcia, M. J., Lizama, V. and Aleixandre, J. L. (2015). Application of Multivariate Regression Methods to Predict Sensory Quality of Red Wine. *Czech J. Food Science*, Vol. 33 (3), 217-227.

Anderson T.W. (1971). An Introduction to multivariate Statistical analysis, John Wiley and sons, New York.

Beaver, C. W. and Harbertson, J. F. (2016). Comparison of Multivariate Regression Methods for the Analysis of Phenolics in Wine Made from Two *Vitis Vinifera* Cultivars. *American Journal of Ecology and Viticulture*. Vol. 67 (1), 56-64.

Baltagi, Badi H. (1998). On the Efficiency of Two-Stage and Three-Stage Least Squares Estimators. *Econometric Reviews*. Vol. 7 (2). 165-169.

Bliemel, F. (1973). Theil's Forecast Accuracy Coefficient: A Clarification. *Journal of Marketing Research*. Vol. 10, No. 4. 444-446.

CBN (2009) Annual Report of Agricultural Product in Nigeria, Produced by Central Bank of Nigeria.

F.M.A.R.D. (2003). Annual Report of the Federal Ministry of Agriculture and Rural Development, Nigeria.

Gujarati, D.N (2003). Basic Econometrics (4th edition.), Tata McGraw- Hill; New Delhi.

Hidalgo, Bertha and Melody Goodman (2013). “Multivariate or Multivariable Regression?” *American Journal of Public Health*. 103, 1. 39-40.

Kapteyn, Arie and Fiebig, Denzil G. (1981). When are two-stage and three-stage least squares estimators identical?. *Economic Letters*. Vol. 8 (1). 53-57.

Kousoyiannis, A. (1977) .Theory of Econometrics (2nd edition.). New York; Palgrave.

Leuthold, R. M. (1975). On the Use of Theil’s Inequality Coefficients. *American Journal of Agricultural Economics*. Vol 57, No.2, 344-346.

Li, Q, Qin, M, Wang, H. and Zang, J. (2014), Application of Multivariate Regression Method in the Prediction of Oilfield Development Indexes. *Advanced Material Research*. Vols.1010-1012, pp 1645-1649.

Maddalla, G.S. (2001). Introduction to Econometrics. John Wiley and Sons, New York.

NBS (2006). Annual Abstract of Statistics, a publication of the National Bureau of Statistics, Nigeria.

Neil, H.T. (1990). Multivariate Analysis with Application in Education and Psychology. (2nd edition); John Wiley and sons, New York.

Norman, C. (2009). Analyzing Multivariate Data. (2nd edition.), Academic Press, New York

Schervish, Mark j. (1987). A Review of Multivariate Analysis. *Statistical Science*, Vol. 2, No.4, 396-433.

Song, Jia, Wei, Li and Ming, Yang (2015). A Method for Simulation Model Validation Based on Theil’s Inequality Coefficient and Principal Component Analysis. *Communications in Computer and Information Science*. Vol. 402, 126-135.

Wonnacott and Wonnaccot, J.R. (1996). Regression; a second course in Statistics, New York: John Wiley and sons