Journal of Statistical and Econometric Methods, vol.8, no.3, 2019, 65-76
ISSN: 2241-0384 (print), 2241-0376 (online)

Assessment on some estimators of a system of simultaneous equation model on the influence of measurement errors in variables of the model

Umeh Edith U[1]. and Ogu Ezenwa H.[1]

Abstract

This paper discusses a Monte Carlo experiment carried out to assess and examine the performance on the finite sample properties of ordinary least squares, indirect least squares, two stage least squares, and three stage least squares estimates of the parameters of the simultaneous equation model on the influence of measurement errors. A system of three just identified equations is set up and application of the four least square techniques is carried out at four categories; when the variables in the models are free from errors, when only the exogenous variables in the model are contaminated with errors, when only the endogenous variables in the model are contaminated with errors, and when both variables are contaminated. From the analysis, it was observed that the estimates of the parameters in the model at each technique vary at different models. It was also observed that 2SLS is the best estimator when all variables in the model are free from errors. 3SLS took the advantage and became best when only the exogenous variables are contaminated. ILS is best when only the endogenous variables are contaminated. And when both variables are contaminated, ILS and 2SLS became similar and best.

Mathematics Subject classification: 62J99
Keywords: Simultaneous equation, exactly identified equation, Monte-Carlo, Measurement Errors, estimators, endogenous and exogenous variables

1. Introduction

It is fact that actual observations on economic variables usually involve multicollinearity, autocorrelation, errors of measurement and most economic problems (koutsoyiannis, 2003). All along in both single and simultaneous (multiple) equation models, it assume implicitly that the dependent and explanatory variables (single equation case), and endogenous and exogenous

[1] Nnamdi Azikiwe University, Awka, Anambra State, Nigeria

Article Info: *Received*: February 28, 2019. *Revised:* March 25, 2019
Published online: 15 May, 2019

variables (simultaneous equation case) are "error free". Thus, measurement errors (MEs) have widely spread problem, variables measured with error can appear in almost all fields of application, especially fields dealing with humans. Many economic variables are measured with substantial error; this occurs both because economists posit theoretical variables such as permanent income which have no direct measures in the real world, and because the process of measurement is itself inaccurate.

Many econometric models assume that the data on their variables are accurate; they are not guess estimates, extrapolated, interpolated or rounded off in any systematic manner, such as to the nearest hundredth naira, and so on. Unfortunately, this ideal is not met in practice for a variety of reasons, such as nonresponse errors, reporting errors, and computing errors.

Griliches(1974), as noted by Ahmed et al(2014) stated that errors maybe introduced by the wording of the survey questionnaires. These errors arise because; (1) in economics, the data producers and data analyzers are separate, (2) there is fuzziness about what it is we would like to observe, and (3) the phenomena we are trying to measure are complex.

Kmenta(1991) as noted by Ahmed et al(2014), introduced the three main classes of such variables that may enter econometric model; (1) variables for which exact measurements are not available and which are represented by error contaminated substitutes, (2) unobserved variables that can be represented only through restitutes related substitutes called "proxy", and (3) variables that are intrinsically not measurable and frequently not even properly defined such as "permanent income" or "intelligence", but that are related to a number of measurable(manifest) variables such as age, educational attainment, e.t.c.

Unlike the single equation model, where the parameter estimates are affected only when the predetermined variable(s) are contaminated with measurement error, the simultaneous equation model parameters are affected when any variable of the model (endogenous, exogenous or both) are contaminated with errors. This is because endogenous variable in one equation of the model may serve as a predetermined variable in another equation(s) of the same model.

However various works done like; (Agunbiade and Iyaniwura 2010), (Ayinde et al 2011), (Agunbiade 2011), (Agunbiade 2012), (Olubosoye 2001), (Oifohi 2012), (Agunbiade 2007), (Johnson 2009), (Alabi and Oyejola 2015) on studying the finite properties of estimators in estimating the parameters on simultaneous equation system, focus on assessing the effects of some of these economics problems; multicollinearity, correlation among error term e.t.c. There is need to investigate or assess the effect of measurement error in variables of system of simultaneous equation in estimating its parameters. These brought us to the questions; what happens when exogenous or endogenous or both variables are contaminated? Secondly, does it affect the parameters' estimates when only the endogenous variables are contaminated with measurement errors, since in simple linear model as noted and practiced by (Gujarati 2005), the presence of measurement error in dependent variables does not affect its parameters' estimates.

Measurement errors pose a serious problem when they are present in the endogenous variable(s), or exogenous variable(s) or both. Therefore, this study examine and compare the performances of four "Least Squares" methods of estimating the parameters of simultaneous equation model, on the influence of measurement errors in variable, using the Monte Carlo approach.

2. Methodology and Application

Consider a three-equation model of the form:

$$\begin{aligned} y_{1t} &= \beta_{12}y_{2t} + \gamma_{11}x_{1t} + \gamma_{12}x_{2t} + u_{1t} \\ y_{2t} &= \beta_{23}y_{3t} + \gamma_{21}x_{1t} + \gamma_{23}x_{3t} + u_{2t} \\ y_{3t} &= \beta_{31}y_{1t} + \gamma_{32}x_{2t} + \gamma_{33}x_{2t} + u_{3t} \end{aligned} \quad (1)$$

where $\beta_{12}, \beta_{23}, \beta_{31}$ are the respective coefficients of the endogenous variables y_{2t}, y_{3t}, y_{1t}. γ_{11}, γ_{12}, The respective coefficients of the exogenous variables x_{1t}, x_{2t} in the first equation of the model. γ_{21}, γ_{23}, the respective coefficients of the exogenous variables x_{1t}, x_{3t} in the second equation of the model and γ_{22}, γ_{23}; the respective coefficients of the exogenous variables x_{2t}, x_{3t} in the third equation of the model. u_{1t}, u_{2t}, u_{3t} denote stochastic disturbance term which are assumed to be independently and identically normally distributed with zero mean and constant variance-covariance matrix.

Rearranging model (1) yields:

$$\begin{aligned} &y_{1t} - \beta_{12}y_{2t} + 0y_{3t} - \gamma_{11}x_{1t} - \gamma_{12}x_{2t} + 0x_{3t} = u_{1t} \\ &0y_{1t} + y_{2t} - \beta_{23}y_{3t} - \gamma_{21}x_{1t} + 0x_{2t} - \gamma_{23}x_{3t} = u_{2t} \\ &-\beta_{31}y_{1t} + 0y_{2t} - y_{3t} + 0x_{1t} - \gamma_{32}x_{2t} - \gamma_{33}x_{3t} = u_{3t} \end{aligned} \quad (2)$$

Equation (2) can be written in matrix form as below;

$$\begin{pmatrix} 1 & -\beta_{12} & 0 \\ 0 & 1 & -\beta_{23} \\ -\beta_{31} & 0 & 1 \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix} + \begin{pmatrix} -\gamma_{11} & -\gamma_{12} & 0 \\ -\gamma_{21} & 0 & -\gamma_{23} \\ 0 & -\gamma_{32} & -\gamma_{33} \end{pmatrix} \begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{pmatrix} = \begin{pmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{pmatrix}$$

This in standard form becomes;

$$\beta Y + \Gamma X = U$$

Making Y the subject of the formula, we have;

$$Y = \Pi X + V \quad (3)$$

where $\Pi = -\beta^{-1}\Gamma$ and $V = \beta^{-1}U$

This equation (3) is written in matrix form as,

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix} = \begin{pmatrix} \Pi_{11} & \Pi_{12} & \Pi_{13} \\ \Pi_{21} & \Pi_{22} & \Pi_{23} \\ \Pi_{31} & \Pi_{32} & \Pi_{33} \end{pmatrix} \begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{pmatrix} + \begin{pmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{pmatrix} \quad (4)$$

Thus $\Pi = -\beta^{-1}\Gamma$ (5)

and $V = \beta^{-1}U$ (6)

With equation (5) and (6) the reduced form of (1) becomes;

$$Y_{1t} = \left(\frac{\gamma_{11}+\beta_{12}\gamma_{21}}{1-\beta_{12}\beta_{23}\beta_{31}}\right)X_{1t} + \left(\frac{\gamma_{12}+\beta_{12}\beta_{23}\gamma_{32}}{1-\beta_{12}\beta_{23}\beta_{31}}\right)X_{2t} + \left(\frac{\beta_{12}\gamma_{23}+\beta_{12}\beta_{23}\gamma_{33}}{1-\beta_{12}\beta_{23}\beta_{31}}\right)X_{3t} + \left(\frac{u_{1t}+\beta_{12}u_{2t}+\beta_{12}\beta_{23}u_{3t}}{1-\beta_{12}\beta_{23}\beta_{31}}\right)$$

$$Y_{2t} = \left(\frac{\beta_{23}\beta_{31}\gamma_{11}+\gamma_{21}}{1-\beta_{12}\beta_{23}\beta_{31}}\right)X_{1t} + \left(\frac{\beta_{23}\beta_{31}\gamma_{12}+\beta_{23}\gamma_{32}}{1-\beta_{12}\beta_{23}\beta_{31}}\right)X_{2t} + \left(\frac{\gamma_{23}+\beta_{23}\gamma_{33}}{1-\beta_{12}\beta_{23}\beta_{31}}\right)X_{3t} + \left(\frac{\beta_{23}\beta_{31}u_{1t}+u_{2t}+\beta_{23}u_{3t}}{1-\beta_{12}\beta_{23}\beta_{31}}\right)$$

$$Y_{3t} = \left(\frac{\beta_{31}\gamma_{11}+\beta_{12}\beta_{31}\gamma_{21}}{1-\beta_{12}\beta_{23}\beta_{31}}\right)X_{1t} + \left(\frac{\beta_{31}\gamma_{12}+\gamma_{32}}{1-\beta_{12}\beta_{23}\beta_{31}}\right)X_{2t} + \left(\frac{\beta_{12}\beta_{31}\gamma_{32}+\gamma_{33}}{1-\beta_{12}\beta_{23}\beta_{31}}\right)X_{3t} + \left(\frac{\beta_{31}u_{1t}+\beta_{12}\beta_{31}u_{2t}+u_{3t}}{1-\beta_{12}\beta_{23}\beta_{31}}\right)$$
(7)

Conforming to all the assumptions specified for the model, our data series are generated as follows.

1. We choose an arbitrary sample sizes n=10, 50, and 100 each; each level of sample size at 50 replications.
2. Arbitrary values were assigned to the structural parameters of the model.

$$\begin{array}{lll} \beta_{12} = 1.5 & \gamma_{11} = 2.5 & \gamma_{12} = 0.4 \\ \beta_{23} = 0.7 & \gamma_{21} = 2.1 & \gamma_{23} = 1.7 \\ \beta_{31} = 0.2 & \gamma_{32} = 1.2 & \gamma_{33} = 3.1 \end{array}$$

3. The variance-covariance matrix of the disturbance terms of the model at any given sample point is assigned arbitrarily to suite the assumptions of variance-covariance matrix, which says that the matrix must be a symmetric matrix and the matrix must be positive semi definite. The matrix is as below;

$$\sum\nolimits_{c} = \begin{pmatrix} 7.0 & 0 & 0 \\ 0 & 4.5 & 0 \\ 0 & 0 & 3.0 \end{pmatrix}$$

4. The exogenous variables X_{1t}, X_{2t}, X_{3t} are generated to follow normal distribution with mean $\overline{X_{1t}} = 10$, $\overline{X_{2t}} = 8.5$, $\overline{X_{3t}} = 17.3$ and variance-covariance matrix $\begin{pmatrix} 8 & 0 & 0 \\ 0 & 9.8 & 0 \\ 0 & 0 & 7.2 \end{pmatrix}$ using R-statistical package.
5. The endogenous variables are then generated from the values already obtained for the $X's$ and $U's$ and the values assigned to the structural parameters. This is done using the reduced form model derived above. i.e, The values earlier generated for X_{1t}, X_{2t}, X_{3t}, u_{1t}, u_{2t}, u_{3t} are then substituted into equation (7) above to obtain numerical values for Y_{1t}, Y_{2t}, Y_{3t}.
6. For the presence of measurement errors; another set of random variables say w_{1t}, w_{2t}, $w_{3t}, h_{1t}, h_{2t}, h_{3t}$ are generated to serve as measurement errors for the variables, $X_{1t}, X_{2t}, X_{3t}, Y_{1t}, Y_{2t}, Y_{3t}$, respectively. They are generated to be uncorrelated with each other and to follow normal distribution with respective means as 2.4, 1.99, 2.8, 4, 2.4, 1.9, and variances 4, 2, 5, 4, 2.4, 3.6 respectively. These generated error variables will be used to contaminate the variables, $X_{1t}, X_{2t}, X_{3t}, Y_{1t}, Y_{2t}, Y_{3t}$, so that the contaminated variables become, $X^e{}_{1t}, X^e{}_{2t}, X^e{}_{3t}, Y^e{}_{1t}, Y^e{}_{2t}, Y^e{}_{3t}$ respectively. Such that when,

i. Only endogenous variables are contaminated by the measurement error; the model becomes

$$y^e{}_{1t} = \beta_{12}y^e{}_{2t} + \gamma_{11}x_{1t} + \gamma_{12}x_{2t} + u_{1t}$$
$$y^e{}_{2t} = \beta_{23}y^e{}_{3t} + \gamma_{21}x_{1t} + \gamma_{23}x_{3t} + u_{2t}$$
$$y^e{}_{3t} = \beta_{31}y^e{}_{1t} + \gamma_{32}x_{2t} + \gamma_{33}x_{2t} + u_{3t}$$

ii. Only the exogenous variables are contaminated by the measurement error, the model becomes

$$y_{1t} = \beta_{12}y_{2t} + \gamma_{11}x^e{}_{1t} + \gamma_{12}x^e{}_{2t} + u_{1t}$$
$$y_{2t} = \beta_{23}y_{3t} + \gamma_{21}x^e{}_{1t} + \gamma_{23}x^e{}_{3t} + u_{2t}$$
$$y_{3t} = \beta_{31}y_{1t} + \gamma_{32}x^e{}_{2t} + \gamma_{33}x^e{}_{2t} + u_{3t}$$

iii. The endogenous and exogenous variables are both contaminated by the measurement error, the model becomes

$$y^e{}_{1t} = \beta_{12}y^e{}_{2t} + \gamma_{11}x^e{}_{1t} + \gamma_{12}x^e{}_{2t} + u_{1t}$$
$$y^e{}_{2t} = \beta_{23}y^e{}_{3t} + \gamma_{21}x^e{}_{1t} + \gamma_{23}x^e{}_{3t} + u_{2t}$$
$$y^e{}_{3t} = \beta_{31}y^e{}_{1t} + \gamma_{32}x^e{}_{2t} + \gamma_{33}x^e{}_{2t} + u_{3t}$$

where $y^e = y + w$ and $x^e = x + h$

Estimation of the parameters of the model is then done with these generated data set of the endogenous and exogenous variables, when they are free from the generated errors, when only the endogenous variables are contaminated, when only the exogenous variables are contaminated, and when both variables are contaminated with the generated error variables $w_{1t},\ w_{2t},\ w_{3t}, h_{1t},\ h_{2t},\ h_{3t}$. The estimation techniques assessed are Ordinary Least Squaes, Indirect Least Squares, Two Stage Least Squares and Three Stage Least Squares. At each experiment, say when none of the variables are contaminated with measurement errors, at n=10 the estimators were assessed and compared using the finite sampling properties of estimators. The finite sampling properties are expressed mathematically as follows;

Bais, $B(\hat{\beta}) = \frac{1}{R}\sum_{j=1}^{R}(\widehat{\beta_J} - \beta) = \bar{\hat{\beta}} - \beta$

Absolute Bais, $AB(\hat{\beta}) = \frac{1}{R}\sum_{j=1}^{R}|\widehat{\beta_J} - \beta|$

Variance, $Var(\hat{\beta}) = \frac{1}{R}\sum_{j=1}^{R}\left(\widehat{\beta_J} - \bar{\hat{\beta}}\right)^2$

Mean Squared Error, $MSE(\hat{\beta}) = \frac{1}{R}\sum_{j=1}^{R}(\widehat{\beta_J} - \beta)^2 = Var(\hat{\beta}) + [B(\hat{\beta})]^2$

For $j = 1,2,3, \dots, R$

The ranking method as adopted by (Ayinde et al 2011), Here, $\hat{\beta}$ represents any of the parameters in the model of equation (1). For each of the estimators, a computer program was written using R software to estimate all the model parameters and to evaluate the criteria. Based on each estimate of the parameter, the estimators were ranked in order of their performance at each criterion. The evaluation of methods was done at two levels-using individual criterions and the totality of all the four criteria. For the first level, the ranks were added for each method and each criterion and for the whole model. Then methods of estimation were ranked by this grand total for each criterion. The method with the least grand total was adjudged the most preferred method and the one with the largest grand total the least preferred. These ranks were added together over all the criteria so as to know how each estimator performs in terms of each parameter in the model. The best estimator in term of the model was identified by further adding all the ranks over the parameters of the model. An estimator is considered best if it has minimum sum of ranks. Here the total of the ranks were used and this will give identical results in terms of ranks if the mean of the ranks had been used.

3. Results

A comparative performance evaluation of the four estimators using the Bias, Absolute Bais, Variance and Mean Squared Error are shown and discussed on Tables 1-5. Bold values in the table indicate the lowest ranks indicating the best estimators.

Table 1: Summary of the total rank of the parameters of the model based on bias criterion

Error measurement	**ESTIMATORS**	**SAMPLE SIZES**		
		n=10	n=50	**n=100**
Free	OLS	31	32	**32**
	ILS	18	21	**27**
	2SLS	17	17	**14.5**
	3SLS	24	20	**17.5**
exogenous	OLS	35	35	**31**
	ILS	18	22	**16**
	2SLS	20	16	**27**
	3SLS	17	17	**16**
endogenous	OLS	17	19	**26**
	ILS	23	23	**21**
	2SLS	18	22	**20**
	3SLS	32	26	**23**
Both variables	OLS	31	35	**35**
	ILS	14	22	**15**
	2SLS	18	15	**21**
	3SLS	27	18	**19**

from Table 1

- When the variables are free from errors, the 2SLS is the best for all sample sizes and it becomes better with increase in sample size. When the sample size is small (n=10), ILS have advantage over 3SLS but with increase in sample sizes, 3SLS performs better than ILS. The least estimator there is OLS at all sample sizes.
- When only the exogenous variables are contaminated, the best estimator is 3SLS when the sample size is small (n=10), but when the sample size is moderate (n=50) the best is 2SLS, then ILS and 3SLS performs alike and are best when the sample is high (n=100). The least is OLS.
- When only the endogenous variables are contaminated, OLS performed best when the sample size is small (n=10) and moderate (n=50), followed by 2SLS. When the sample size is high (n=100), 2SLS performed best.
- When both variables are contaminated, ILS is the best estimator at small and high sample sizes, while at moderate sample size; 2SLS is the best estimation method. The least estimator here is the OLS method.

Table 2: summary of the total rank of the parameters of the model based on absolute bias criterion

Error measurement	**ESTIMATORS**	**SAMPLE SIZES**		
		n=10	n=50	**n=100**
Free	OLS	23	26	**32**
	ILS	21	19	**28**
	2SLS	21	20	**12**
	3SLS	25	25	**18**
exogenous	OLS	35	35	**35**
	ILS	18	21	**16**
	2SLS	18	15	**24**
	3SLS	19	19	**15**
endogenous	OLS	14	20	**26**
	ILS	23	23	**22**
	2SLS	20	20	**19**
	3SLS	33	27	**23**
Both variables	OLS	30	35	**35**
	ILS	17	21	**15**
	2SLS	17	14	**21**
	3SLS	26	20	**19**

From Table 2

- When the variables are free from errors, the Absolute Bias of ILS and 2SLS methods are similar when the sample size is small (n=10), but 2SLS performs better when there is increase in sample size. 3SLS performs poorly but better than 2SLS, since it was getting better with increase in sample size while OLS gets least with increase in sample size.
- When only the exogenous variables are contaminated, ILS and 2SLS are similar in performance when the sample size is small (n=10), but when sample size is moderate (n=50), 2SLS becomes best, followed by 3SLS. Though at high sample size (n=100), 3SLS performs best.
- When only the endogenous variables are contaminated, the estimation methods performed almost exactly as they did on the Bias Criterion.
- When both variables are contaminated, the ILS and 2SLS performed similar at small sample size. When the sample size is moderate, 2SLS performed best. At high sample size, ILS performed best. The least estimation method here is the OLS.

Table 3: Summary of the total rank of the parameters of the model based on variance criterion

Error measurement	**ESTIMATORS**		**SAMPLE SIZES**	
		n=10	n=50	**n=100**
Free	OLS	16	10	**12**
	ILS	28	22	**32**
	2SLS	22	29	**22**
	3SLS	24	29	**24**
exogenous	OLS	11	12	**9**
	ILS	25	28	**31**
	2SLS	18	20	**17**
	3SLS	36	30	**33**
endogenous	OLS	20	15	**20**
	ILS	13	20	**28**
	2SLS	24	22	**19**
	3SLS	33	33	**23**
Both variables	OLS	10	13	**10**
	ILS	26	29.5	**31**
	2SLS	22	19	**18**
	3SLS	32	28.5	**31**

From Table 3

- When the variables are free from errors, the best estimator is OLS in all the sample sizes. 2SLS followed except when the sample size is moderate (n=50) where 2SLS and 3SLS performed alike or similar. The least estimation method in this criterion is ILS.
- When only the exogenous variables are contaminated and for all sample sizes, OLS performs best, followed by 2SLS, then ILS and then 3SLS at all sample sizes.
- When only the endogenous variables are contaminated, ILS did best when the sample size is small (n=10), then at moderate sample size (n=50), OLS performed best, while 2SLS did best at high sample size (n=100).
- When both variables are contaminated, OLS method performed best at all sample sizes, followed by the 2SLS, then the ILS. And the least estimator here is the 3SLS.

Table 4: Summary of the total rank of the parameters of the model based on mean squared error criterion

Error measurement	**ESTIMATORS**	**SAMPLE SIZES**		
		n=10	n=50	**n=100**
Free	OLS	26	26	**28**
	ILS	22	20	**29**
	2SLS	19	19	**14**
	3SLS	23	25	**19**
exogenous	OLS	35	35	**35**
	ILS	16	19	**16**
	2SLS	17	15	**24**
	3SLS	22	21	**15**
endogenous	OLS	15	20	**26**
	ILS	22	23	**22**
	2SLS	20	20	**19**
	3SLS	33	27	**23**
Both variables	OLS	30	35	**35**
	ILS	17	21	**15**
	2SLS	17	15	**21**
	3SLS	26	19	**19**

From Table 4

- When the variables are free from errors, the best estimator here is 2SLS across all sample sizes, and it gets better with increase in sample size, followed by ILS when the sample size is small (n=10) and moderate (n=50). But when the sample size is high (n=100), 3SLS performs better than ILS.
- When only the exogenous variables are contaminated, the best estimator in this criterion are ILS at small sample size (n=10), 2SLS at moderate sample size (n=50), and 3SLS at high sample size (n=100). The least estimator here is OLS.
- When only the endogenous variables are contaminated and at moderate sample size, OLS and 2SLS performed best and similar. At small sample size (n=10), OLS performed best while at high sample size (n=100), 2SLS performed best. The least estimation method at this criterion at small and moderate sample sizes is 3SLS, while performed worst at high sample size.
- When both variables are contaminated, the 2SLS performed best, except when the sample size is high, where the ILS did best. The least estimation method here is OLS.

Table 5: Overall summary of the total rank of the parameters of the model based on all criteria

Error measurement	**ESTIMATORS**		**SAMPLE SIZES**	
		n=10	n=50	**n=100**
Free	OLS	96	94	**104**
	ILS	89	82	**116**
	2SLS	79	85	**62.5**
	3SLS	96	99	**77.5**
exogenous	OLS	116	117	**110**
	ILS	77	90	**79**
	2SLS	72	66	**92**
	3SLS	94	87	**79**
endogenous	OLS	66	74	**98**
	ILS	81	89	**93**
	2SLS	82	84	**77**
	3SLS	131	113	**92**
Both variables	OLS	101	118	**115**
	ILS	74	93.5	**76**
	2SLS	74	63	**81**
	3SLS	111	85.5	**88**

From table 5

- When the variables are free from errors, the best method for estimating the exactly identified simultaneous equation model is 2SLS when the sample size is small (n=10) and high (n=100). When the sample size is moderate (n=50), the best estimator is ILS. The least is OLS.
- When only the exogenous variables are contaminated and at small (n=10) and moderate (n=50) sample sizes, 2SLS is the best estimation method. But at high (n=100) sample size, 3SLS performs best. The least estimation method here is OLS.
- When only the endogenous variables are contaminated and at small sample size, the bet estimator is ILS, but as the sample size increases, the best estimator became 2SLS. The worst estimator here is 3SLS at small and moderate sample sizes while at high sample size, OLS performed least.
- When both variables are contaminated and at small sample size, the ILS and the 2SLS performed best and similar. At moderate sample size, the 2SLS performed best, while at high sample size, the ILS performed best.

4. Conclusion

As confirmed by (Gujarati 2005), errors of measurement affect the parameter estimates of linear regression when the predetermined variables are contaminated and when both dependent and predetermined variables are contaminated, and it does not affect the parameter estimates when only the dependent variables are contaminated. However, the results obtained above show that when a two-way causation occurs in a function, the parameter estimates of the model is affected when any of the variables are contaminated with measurement errors (i.e.: exogenous, endogenous or both). The performance of the four categories of the estimators: OLS, ILS, 2SLS and 3SLS are affected when any or all of the variables in the model are contaminated with measurement errors. This is because the dependent variable Y and independent variables now appear as endogenous variables as well as explanatory variables in other equation(s) of the model.

Among the Four least square estimation techniques studied, findings reveal that when variables are free from errors, the best estimation technique for estimating the exactly identified simultaneous equation model is 2SLS when the sample size is small (n=10) and high (n=100). When the sample size is moderate (n=50), the best estimator is ILS. When only the exogenous variables are contaminated and at small (n=10) and moderate (n=50) sample sizes, 2SLS is the best estimation method. But at high (n=100) sample size, 3SLS performs best. When only the endogenous variables are contaminated and at small sample size, the bet estimator is ILS, but increase in sample size, the best estimator became 2SLS. The worst estimator here is 3SLS at small and moderate sample sizes while at high sample size, OLS performs worst. When both variables are contaminated and at small sample size, the ILS and the 2SLS performed best and similar. At moderate sample size, the 2SLS performed best, while at high sample size, the ILS performed best. The worst estimation technique is OLS across all categories.

References

[1] Agunbiade, D. A.: A Monte-Carlo Approach to the Estimation of a Just Identified simultaneous Three-Equations Model with Three Multicollinear Exogenous variables. Unpublished Ph.D. Thesis, University of Ibadan, Nigeria. (2007):

[2] Agunbiade D. A. and Iyaniwura J. O.): Estimation under multicollinearity: A comparative approach using Monte Carlo methods. Journal of Mathematics and Statistics 6(2) (2010, 183-192

[3] Agunbiade D. A.: Effect of Multicollinearity and the Sensitivity of the Estimation Methods in Simultaneous Equation Model. Journal of Modern Mathematical and Statistics 5(1): (2011): 9-12.

[4] Agunbiade D. A: A note on the effect of the Multicollinearity Phenomenon of a Simultaneous Equation Model. Journal of Mathematical Sciences: Advances and Applications.15,(1), 1-12

[5] Ahmed A. E. and Amt-allateef : Estimation of the Simultaneous Equation model with Measurement Errors. Int. J. Contemp. Math. Sciences, 9(7) 2014, 331-348

[6] Alabi O.O. and Oyejola B.A.: Assessments of Some Simultaneous Equation Estimation Techniques with Normally and Uniformly Distributed Exogenous Variables. Applied Mathematics, 6, (2015) 1902-1912. http://www.scirp.org/journal/am

[7] Ayinde K., Olatunji O., Adewale L. Babatunde R.O and Omotosho O.A: : Effect of equation identification status and correlation between error terms on estimators of a system of Simultaneous Equations Model. American Journal of Scientific and Industrial Research. 2(2): (2011): 184 – 190

[8] Griliches Z.: Errors in variables and other unobservables. Econometrica, 42 (1974):, 197-998.

[9] Gujarati, D. N.: Basis Econometrics. Fifth Edition.. Tata McGraw-Hill Publishing Company Limited. (2005):

[10] Johnson T. L: effect of correlation and identification status on methods of estimating parameters of system of simultaneous equation model. A Ph.D. Thesis submitted to the Department of Statistics, University of Ilorin, Nigeria..(2009):

[11] Koutsoyiannis, A.: Theory of Econometrics; An Introductory Exposition of Econometrics Methods. 2nd Edition;. Palgrave, New York.(2003):

[12] Oduntan A.: Monte Carlo Study of the Problem of Multicollinearity in a Simultaneous Equation Model. An Unpublished Ph.D. Thesis submitted to the Department of Statistics, University of Ibadan, Nigeria. (2004):

[13] Olubusoye, O. E. The Consequences of the Violation of the Assumption of Zero Correlation between Pairs of Random Stochastic Terms Used in Monte Carlo Experiments. An unpublished Ph.D thesis submitted to the Department of Statistics, University of Ibadan, (2001):