A Monte Carlo simulation framework on the relative performance of system estimators in the presence of multicollinearity

Emmanuel A. Oduntan* and J. O. Iyaniwura

Abstract: The correctness and reliability of findings and recommendations of empirical studies conducted by social and economic researchers depend largely on the efficiency of the econometrics methodologies employed in such studies. Of particular interest are such studies which are centered on the Sustainable Development Goals (SDG) considering the relevance of such studies to the total wellbeing of the world populace. In view of this, there is always a need for theoretical review of econometrics methodologies commonly used by researchers with a view to providing researchers with research updates on the theoretical standing of these methodologies. In this study, we set up a Monte Carlo Experiment (MCE) to evaluate the relative performance of various estimators of a simultaneous equation model in the presence of varied levels of multicollinearity. The model was estimated with a simulated data set of sample size 30 over 100 replications. The parameter estimates obtained from the six estimators considered were evaluated using RMSE criteria. Our result revealed that irrespective of the level of multicollinearity in our model, ILS and OLS yielded best estimates of the parameters. On the contrary, the system estimators all performed poorly in the presence of multicollinearity. Also, 2SLS, LIML and 3SLS estimators yielded virtually identical estimates. By our findings, in the presence of multicollinearity, estimators OLS and ILS performed best and should therefore be preferred above the multi-equation estimators.

Subjects: Economics; Econometrics; Mathematical Economics; Economic Forecasting

ABOUT THE AUTHORS

Professor J. O. Iyaniwura is a Professor of Statistics with a track record spanning over forty years’ experience in teaching and research and with specialty in Econometrics and Time Series Analysis. He has authored many articles in reputable scholarly journals.

Emmanuel A. Oduntan, is an astute academician with experience in teaching and research spanning over 2 decades. His teaching and research interest includes Statistics and Econometrics. Prior to his venture into the academis, he worked extensively in the Nigerian banking industry.

PUBLIC INTEREST STATEMENT

The correctness and reliability of findings from empirical studies conducted by social and economic researchers depend largely on the efficiency of the econometrics methodologies employed in such studies. Of particular interest are such studies, which are centered on the Sustainable Development Goals (SDG) considering the relevance of such studies to the total wellbeing of the world populace. In view of this, there is always a need for continuous contribution to literature on econometrics methodologies commonly used by researchers with a view to providing researchers with research updates on these methodologies. To this end, in this study we set up a Monte Carlo Experiment (MCE) to evaluate the relative performance of various estimators of a simultaneous equation model in the presence of varied levels of multicollinearity.
Keywords: Monte Carlo simulation; multicollinearity; simultaneous equation model; exogenous and endogenous variables; correlation coefficient

1. Introduction

The correctness and reliability of findings and recommendations of empirical studies conducted by social and economic researchers depend largely on the efficiency of the econometrics methodologies employed in such studies. Of particular interest are such studies, which are centered on the Sustainable Development Goals (SDG) considering the relevance of such studies to the total well-being of the world populace. In view of this, there is always a need for theoretical review of econometrics methodologies commonly used by researchers with a view to providing researchers with research updates on the theoretical standing of these methodologies. In this study, we set up a Monte Carlo Experiment (MCE) to evaluate the relative performance of various estimators of a simultaneous equation model in the presence of varied levels of multicollinearity. In the last few decades, Monte Carlo Simulation (MCS) has found extensive use in the fields of operational research and nuclear physics, where there are varieties of problems beyond the available tools of theoretical mathematics. They have been employed sporadically in numerous other fields of science, including Chemistry, Biology, Econometrics and Medicine. Problems handled by Monte Carlo methods are of two types. These are probabilistic or deterministic type according to whether or not they are directly concerned with the behavior and outcome of random processes.

Although the basic procedure of the Monte Carlo method is the manipulation of random numbers, these should not be employed prodigally. Each random number is a potential source of added uncertainty in the final result. Olubusoye (2001) warned researchers using the Monte Carlo method of the problem of using inherently correlated normal deviates to simulate the disturbance term in a simultaneous equation context. Consequently, it is necessary to scrutinize each part of a Monte Carlo experiment to see whether that part cannot be replaced by exact theoretical analysis contributing no uncertainty and also to ensure that experimental results are not confounded by the use of inherently deficient random deviates.

The Monte Carlo experimentalist needs wide experience in the use of some tools of pure mathematics and statistics, especially the theory of probability, in order that he may discern those connections between apparently dissimilar problems which suggest sophisticated Monte Carlo methods. He has to exercise ingenuity in distorting and modifying problems in the pursuit of variance-reducing techniques. He has to be competent at statistical and inferential procedures in order to extract the most reliable conclusions from his artificially generated data. The tools of his trade include computing machinery, and he must be familiar with numerical analysis. As in all experimental works, a feel of the problems is a great and sometimes essential asset.

In this paper, we present a Monte Carlo Simulation (MCS) framework used in studying the relative efficiency of system estimators in the presence of multicollinearity.

2. Literature review

The principle of asymptotic theory is the utilization of distributional properties in large samples to find estimates of unknown population parameters and draw inference about them. Some of the properties include consistency, asymptotic efficiency, and asymptotic normality. In econometrics, attention is often focused on evaluating the behavior of estimators in small sample and asymptotic theory fails most of the time to furnish enough useful information about these estimators. This inevitably creates a problem in trying to choose from among competing estimators.

One way of studying the small sample properties of estimators is to utilize the Monte Carlo method. The Monte Carlo method has been applied not only to the choice of alternative estimators but also in determining the impact of sample size, serial correlation, multicollinearity, and other factors on the different possible estimators in a given study. This approach creates a “laboratory
environment” where controlled experiments on estimators are performed. So far, the sample properties of the various econometric techniques have been studied from simulated data in Monte Carlo studies and not with direct application of the techniques to actual observations. This approach is due to the fact that actual observations on economic variables are usually infected by multicollinearity, autocorrelation, errors of measurement and most other econometric “diseases” and in some cases simultaneously. Studies on small sample properties of estimators are usually based on the assumption of the simultaneous occurrence of all these problems. By Monte Carlo approach, the econometrician can generate data sets and stochastic terms, which are free of all but one of the problems listed above and therefore generate data resembling those obtained from controlled experiments.

Wagner (1958) employed Monte Carlo sampling technique to compare small sample properties of limited information single Equation (LISe), least squares, and instrumental variable estimates. Two versions of two-equation models were used to generate 100 sets of observations over 20 time periods. With these observations and various statistical techniques, estimates of the parameters of an over-identified equation were obtained and compared. The analysis of the results of the 100 sets of data produced for each of the model indicates that the least squares generally give more biases but less variable estimates than the LISe method. Baldev (1980) considered four alternative forms of two-parameter normal and non-normal error distributions. He also reported on a Monte Carlo study of the small-sample properties of OLS, 2SLS, 3SLS, and FIML estimators. On the basis of 1,000 replications of sample size 20 in two experiments on an over-identified model, he found that the small-sample rankings of econometric estimators of both structural coefficients and forecasts of endogenous variables according to parametric and non-parametric measures of bias, dispersion and dispersion including bias do not change for any of the four error distributions considered.

Kloek and Van Dijk (1978) applied Monte Carlo methods to obtain estimates of posterior moments of structural and reduced form parameters of Bayesian simultaneous equation systems. They noted that the Monte Carlo methods allow the analyst to make use of several types of exact and stochastic prior information. The method was also used to compute marginal posterior distributions of some parameter of interest. Lurie and Hartley (1972) employed Monte Carlo methods to handle distributional problems involving order statistics that cannot be treated analytically. Kmenta and Joseph (1963) carried out a Monte Carlo experiment to examine the small sample properties of ordinary least squares, indirect least squares, Huchs and Kleins estimates of the parameters of the Cobb-Douglas production function. A perfectly competitive group of firms in a single industry is considered in nine situations, which differ in the behavior of the disturbances, the variability of inputs and the position of the average firm. In each case, 200 samples of size 20 and 200 samples of size 100 were selected to approximate the sampling distribution of the various estimators.

Nagar (1960) also studied the small sample properties of the simultaneous equation estimators using the Monte Carlo approach. Four methods of estimation considered were least squares, 2SLS, unbiased and minimum second moment. The study showed that the least squares of the four methods possesses the smallest second order sampling moments about the true parameter value in a majority of cases, while 2SLS showed the smallest bias in all cases. Carlin et al. (1992) proposed an adaptive Monte Carlo integration technique known as Gibbs sample as a mechanism for implementing a conceptually and computationally simple solution for multivariate space-modeling, forecasting and smoothening. They allowed for possibilities of non-normal errors and non-linear functions in the state equation, the observational equation or both.

Also, Essi (1997), employed the Monte Carlo approach to examine the consequences of wrong specification of error terms on the robustness of the estimators of non-linear single-equation econometric model. His study focused on the consequences of making a wrong choice
i.e. assuming a multiplicative error-based model instead of an additive error-based model, and vice-versa; in the process of estimating an intrinsically nonlinear Cobb-Douglas model with parameters $\theta_1, \theta_2$ and $\theta_3$. Nwabueze (2000) considered a single-equation study of the choice of estimators of nine autocorrelated models also using the Monte Carlo technique. The study confirmed the least squares as the best out of the four estimators considered. Olubusoye (2001) using a two-equation simultaneous econometric model studied the consequences of the violation of the assumption of zero correlation between pairs of random stochastic terms used in Monte Carlo experiments. He concluded that pairs of normal deviates should be screened and those inherently correlated should not be used in generating stochastic error terms to be used in Monte Carlo experiments.

Despite the widespread use of the Monte Carlo technique, it suffers from two main deficiencies namely: specificity and impression. Specificity means that results obtained are largely dependent on the specification of the model(s) used in a given Monte Carlo experiment while imprecision means that the results are not precise. The impact of these deficiencies can be reduced to a reasonable extent. For instance, Response Surface Methodology may be used to reduce specificity deficiency while Control Variate Technique may be used to reduce the problem of imprecision. The implication of these deficiencies is two-fold. One is that simulation results are most of the time suggestive rather than conclusive since the extent of generalization of sampling experiments compares to their analytical counterparts is smaller. Secondly, Monte Carlo results are not end-product in econometric theory in general, or in the investigation of finite sample distribution in particular, but stands somewhere in-between. Hendry (1984) asserted that: “Monte Carlo experimentation can efficiently complement analysis to establish numerical analytic formulae which jointly summarizes the experimental findings and known analytical results in order to help interpret empirical evidence and to compute outcomes at other points within the relevant parameter space. The accuracy obtainable depends on the given budget constraints and the relevant price of capital labour. The later determining the point at which simulation is substituted for analysis and the former the overall precision of the exercise.”

3. Monte Carlo simulation framework

3.1. Theoretical framework

Assume the econometric model

$$Y = F(X, \theta) + \mu$$

where $\mu N(0, \sigma^2)$ and it also satisfies other classical assumptions for least squares estimation.

Numerical values are assigned to all the parameters embodied in the vector $\theta$. The variance $\sigma^2$ is also assigned a numerical value, and on the basis of the assumed $\sigma^2$, normal deviates are selected and used in generating $\mu$, the disturbance term

A random sample size $T$ of $X$ is selected and the numerical values of $F(X, \theta)$ are computed.

The vector $Y$ is then obtained by computing $F(X, \theta) + \mu$

The regression of $Y$ on $X$ is performed to produce estimate $\hat{\theta}$ of $\theta$. This procedure is replicated many times using the same sample size making it possible for the sampling distribution of $\hat{\theta}$ to be constructed. The empirical distribution so obtained is then utilized in evaluating the precision of $\hat{\theta}$ and in making other comparisons especially of the relative performance of different estimators of $\theta$. 
3.2. Empirical framework

Assume the following model

\[ y_{1t} = \beta_{11} y_{2t} + \gamma_{11} x_{1t} + \gamma_{12} x_{2t} + u_{1t} \]

\[ y_{2t} = \beta_{21} y_{1t} + \gamma_{22} x_{2t} + \gamma_{23} x_{3t} + u_{2t} \]  

(2)

where

\( y \)'s are the endogenous variables

\( x \)'s are the exogenous variables

\( u \)'s are the disturbance terms

Equation (2) may be expressed in matrix form as

\[ y = X\beta + u \]

Where:

\[
\begin{bmatrix}
  y_1 \\
  y_2 
\end{bmatrix} =
\begin{bmatrix}
  1 & 1 & 1 \\
  1 & 1 & 1 
\end{bmatrix}
\begin{bmatrix}
  \beta_1 \\
  \beta_2 \\
  \gamma_1 \\
  \gamma_2 \\
  \gamma_3 
\end{bmatrix} +
\begin{bmatrix}
  u_1 \\
  u_2 
\end{bmatrix}
\]

In the linear simultaneous equations econometric model containing \( G \) endogenous (jointly dependent) variables \( y_1, \ldots, y_G \) and \( K \) predetermined variables \( X_1, X_2, \ldots, X_K \), the structural equations at time \( t \) may in general be written as:

\[ \beta_{1i} y_{1t} + \beta_{2i} y_{2t} + \cdots + \beta_{Gi} y_{Gt} + \gamma_{1i} x_{1t} + \cdots + \gamma_{Ki} x_{Kt} = u_{it} \quad i = 1, \ldots, G, t = 1, \ldots, n \]  

(3)

In matrix form, (3) can be written as:

\[ B y_t + \Gamma X_t = u_t \]  

(4)

where \( B \) is a \( G \times G \) matrix of coefficients of current endogenous variables, \( \Gamma \) a \( G \times K \) matrix of coefficient of predetermined variables \( y_t, X_t \) and \( u_t \) are column vectors of \( G \) and \( K \) elements respectively.

\[
B = 
\begin{bmatrix}
  \beta_{11} & \beta_{12} & \cdots & \beta_{1G} \\
  \beta_{12} & \beta_{22} & \cdots & \beta_{21} \\
  \vdots & \ddots & \ddots & \vdots \\
  \beta_{G1} & \beta_{G2} & \cdots & \beta_{GG} 
\end{bmatrix},
\Gamma = 
\begin{bmatrix}
  y_{11} & y_{12} & \cdots & y_{1K} \\
  y_{21} & y_{22} & \cdots & y_{2K} \\
  \vdots & \ddots & \ddots & \vdots \\
  y_{G1} & y_{G2} & \cdots & y_{GK} 
\end{bmatrix},
\]

\[
y_t =
\begin{bmatrix}
  y_{1t} \\
  y_{2t} \\
  \vdots \\
  y_{Gt} 
\end{bmatrix},
X_t =
\begin{bmatrix}
  X_{1t} \\
  X_{2t} \\
  \vdots \\
  X_{Gt} 
\end{bmatrix},
u_t =
\begin{bmatrix}
  u_{1t} \\
  u_{2t} \\
  \vdots \\
  u_{Gt} 
\end{bmatrix}
\]

The system of equations is complete since the number of endogenous variables is equal to the number of structural equations in the model. The system of equations as written in Equation (4) is a structural specification of the model. Also, the system of equations jointly determines at time \( t \), the \( G \) endogenous variables, in terms of the predetermined variables and the values assumed by the stochastic disturbance term.
Each equation in Equation (4) has G + K parameters $\beta_1, \beta_2, \ldots, \beta_K; \gamma_1, \gamma_2, \ldots, \gamma_K$. In practice, some of the parameters are usually specified to be zero. Otherwise, all the equations in the model will look alike statistically, and no equation could be identified. Constant terms are assumed to be included in the model by specifying one of the predetermined variables to be identically unity.

Major assumptions about the model are:

(1) The first basic assumption of the model is that the vector of sample observations on $Y$ may be expressed as a linear combination of the sample observation as the explanatory $X$ variables plus a disturbance vector.

The assumptions about $u$ are as follows:

1. For an arbitrary period $t$

\[
E(u_t) = 0
\]  
(5)

2. Variance-Covariance matrix of $u$ is defined as:

\[
E(U_tU_t') = \Sigma = \Omega \otimes \mathbb{I}_n
\]  
(6)

\[
E(U_tU_t') = E\left[ \begin{bmatrix} u_{11} \\ u_{11} \\ \vdots \\ u_{1n} \\ u_{21} \\ \vdots \\ u_{2n} \\ \vdots \\ u_{G1} \\ \vdots \\ u_{Gn} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & \ldots & u_{1n} & u_{21} & \ldots & u_{2n} & \ldots & u_{G1} & \ldots & u_{Gn} \end{bmatrix} \right]
\]  
(7)

Equation (7) gives a $G \times G$ symmetric matrix whose elements are given by:

\[
E(U_tU_{t', t}) = \begin{cases} 
\sigma^2_{u}; & \text{for } i = i', t = t' & \text{(a)} \\
0; & \text{for } i = i', t \neq t' & \text{(b)} \\
\sigma_{u}; & \text{for } i \neq i'; t = t' & \text{(c)} \\
0; & \text{for } i \neq i'; t \neq t' & \text{(d)} 
\end{cases}
\]

where,

(a) implies existence of homoscedasticity
(b) implies absence of autocorrelation
(c) implies that covariances of the same periods are equal to $\sigma_u$. This means that contemporaneous covariances in different equations are independent of time.
(d) Implies that covariances at different times in different equations or non-contemporaneous covariances are zero.

1. $X$ matrix is of full column rank i.e. $\rho(X) = K$
2. $X$ is a non-stochastic matrix.
3. The $u$ vector has a multivariate normal distribution, i.e., $\text{UNID}(0, \Sigma)$. 

The reduced-form equations are obtained by solving the structural equation [Equation (4)] for $y_t$, i.e. 

$$y_t = \beta' \Gamma x_t + \beta' u_t = \mathbb{E} x_t + V_t$$

where

$$\mathbb{E} = -\beta' \Gamma$$ is the matrix of reduced form coefficients, and,

$$V_t = \beta' u_t$$ is the residual vector of the reduced form equations with

$$E(V_t V_t) = \beta' \Sigma \beta' = \Omega$$

The parameters of the structural equations

$$(\beta, \Gamma, \Sigma)$$

and the parameters of the reduced form equations

$$(\mathbb{E}, \Omega)$$

These two sets of parameters are related as follows:

$$\mathbb{E} = -\beta' \Gamma$$ and $$\Omega = \beta' \Sigma \beta'$$

The $\mathbb{E}$ matrix is of order $G \times K$ and thus contains $GK$ elements. The $\beta$ and $\Gamma$ matrices contain at most $G^2 + GK$ elements.

The MCS framework presented below was adopted to assemble a data set that is conformable to the assumption specified for the model and the imposed levels multicollinearity:

(i) We set sample sizes $T$ at 30.
(ii) The following numerical values are arbitrarily assigned to each of the structural parameters of the model.

$$\beta_{12} = 1.8, \gamma_{11} = 1.2, \gamma_{21} = 0.6$$

$$\beta_{21} = 0.4, \gamma_{12} = 0.5, \gamma_{23} = 1.4$$

(8)

(i) The following arbitrary values are assigned to each of the elements of the variance-covariance matrix of the disturbance terms of the model at any given sample point.

$$\Omega = \begin{bmatrix} 4.5 & 3.0 \\ 3.0 & 3.5 \end{bmatrix}$$

(9)

(i) Values of the predetermined variables $X_{1t}, X_{2t}$ and $X_{3t}$ were selected from a pool of uniformly distributed random numbers such that the correlation coefficients $r_{x_{1}x_{2}}, r_{x_{2}x_{3}}$ and $r_{x_{1}x_{3}}$ are of the following magnitudes

(a) insignificant at 5% level—Low Multicollinearity
(b) significant at 1% level—High Multicollinearity

Consequently, there are three sets of $X$'s in the low multicollinearity group and another three sets of $X$'s in the high multicollinearity group i.e.,

$$r_{x_{1}x_{2}}, r_{x_{2}x_{3}}$$ and $$r_{x_{1}x_{3}}$$; Low Multicollinearity
Values of the disturbance terms $U_{1t}$ and $U_{2t}$ are specified for each sample points. We employed a two-stage process in generating values of the random disturbance terms from a Normal (0, 1) distribution. The two stages involved here are described below.

(a) Independent series $\varepsilon_t$ of random normal deviates $N(0, 1)$ were drawn from a pool of random normal deviates. The length of each pair is determined by the sample size designed.

(b) The generated series were transformed into a series of random disturbances to guarantee conformity with the variance-covariance matrix $\Omega$ predetermined for the model. The method presented by Nagar (1960) for transformation of independent series of standard random deviates into series of random deviates with zero mean and a specified variance-covariance matrix was used for this purpose. This is described below.

The two random disturbance series are thus formed using

$$u = P\varepsilon_t = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

$$= P \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

$$= \begin{pmatrix} S_{11} & S_{21} \\ 0 & S_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

Hence,

$$u_{1t} = S_{11} \varepsilon_{1t} + S_{21} \varepsilon_{2t}$$

$$u_{2t} = S_{22} \varepsilon_{2t}$$

(vi) The endogenous variables were generated from the values already obtained for the $X$'s and $U$'s and the values assigned to the structural parameters. This is most conveniently done using the reduced form of the model.
\[ y_{1t} = \beta_1 y_{2t} + \gamma_{11} X_{1t} + \gamma_{12} X_{2t} + u_{1t} \]
\[ y_{2t} = \beta_2 y_{1t} + \gamma_{21} X_{2t} + \gamma_{22} X_{3t} + u_{2t} \]

Rearranging the model, we have,
\[ y_{1t} - \beta_{12} y_{2t} - \gamma_{11} X_{1t} - \gamma_{12} X_{2t} - \delta X_{3t} = u_{1t} \]
\[ -y_{2t} y_{1t} + y_{2t} - 0 X_{1t} - \gamma_{21} X_{2t} - \gamma_{22} X_{3t} = u_{2t} \]

This can be written as
\[ B y_t + \Gamma X_t = u \quad (15) \]

Where,
\[ B = \begin{bmatrix} 1 & -\beta_{12} \\ -\beta_{21} & 1 \end{bmatrix}, \Gamma = \begin{bmatrix} -\gamma_{11} & -\gamma_{12} & 0 \\ 0 & -\gamma_{21} & -\gamma_{22} \end{bmatrix}. \]
\[ y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix}, X_t = \begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix}, u = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \]

Rewriting Equation (15), we have
\[ y_t = -B^{-1} \Gamma X_t + B^{-1} u \quad (16) \]

\[ = -\frac{1}{1-\beta_{12} \beta_{21}} \left[ \begin{array}{c} 1 \\ \beta_{12} \end{array} \right] \left[ \begin{array}{c} \beta_{21} \\ 1 \end{array} \right] \begin{bmatrix} -\gamma_{11} & -\gamma_{12} & 0 \\ 0 & -\gamma_{21} & -\gamma_{22} \end{bmatrix} \begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} + \frac{1}{1-\beta_{12} \beta_{21}} \left[ \begin{array}{c} 1 \\ \beta_{12} \end{array} \right] u_t \]

Thus, we have
\[ y_{1t} = \frac{\gamma_{11}}{1-\beta_{12} \beta_{21}} X_{1t} + \left[ \frac{\gamma_{12} \beta_{21}}{1-\beta_{12} \beta_{21}} + \frac{\beta_{12} \beta_{23}}{1-\beta_{12} \beta_{21}} \right] X_{2t} + \left[ \frac{\beta_{12} \beta_{23}}{1-\beta_{12} \beta_{21}} \right] X_{3t} \]
\[ + \left[ \frac{\gamma_{11} \beta_{23}}{1-\beta_{12} \beta_{21}} \right] X_{1t} + \left[ \frac{\gamma_{21} \beta_{23}}{1-\beta_{12} \beta_{21}} \right] X_{2t} + \left[ \frac{\beta_{23}}{1-\beta_{12} \beta_{21}} \right] X_{3t} \]
\[ + \left[ \frac{\beta_{12} u_{1t} + u_{2t}}{1-\beta_{12} \beta_{21}} \right] \quad (17) \]

Equation (17) was used to determine the values of the endogenous variables at each sample point.

(vii) The final stage of this experiment involved estimation of the structural parameters with the aid of the generated data sets for \( y_{1t}, y_{2t}, X_{1t}, X_{2t}, \) and \( X_{3t} \) using Ordinary Least Squares(OLS), Two Stage Least Squares (2SLS), Limited Information Maximum Likelihood (LIMF), Indirect Least Squares(ILL), Three Stage Least Squares (3SLS) and Full Information Maximum Likelihood methods.
Table 1. Average of parameter estimates for $T = 30$, $R = 100$

| Estimator | Multicollinearity | \( \beta_2(1.8) \) | \( \gamma_1(1.2) \) | \( \gamma_2(0.6) \) | \( \beta_2(0.4) \) | \( \gamma_2(0.5) \) | \( \gamma_2(1.4) \) |
|-----------|-------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| Low       |                   | 0.5120              | 3.8172              | 1.1932              | 1.9523              | -0.0252             | -13.0877            |
| OLS       | High              | 0.5207              | 5.336               | 1.1660              | 1.9582              | 0.0132              | -13.1205            |
| Low       |                   | 1.2449              | -18.1709            | 1.0886              | -18.3399            | 184.5604            | 143.8901            |
| LIML      | High              | 5.8651              | -12.5154            | -65.4724            | 2.2097              | -0.0772             | -16.7034            |
| Low       |                   | 1.2449              | -18.1709            | 1.0886              | -18.3399            | 184.5604            | 143.8901            |
| 2SLS      | High              | 5.8651              | -12.5154            | -65.4724            | 2.2097              | -0.0772             | -16.7034            |
| Low       |                   | 1.9541              | -1.3030             | 0.4217              | 0.5636              | -0.1629             | 2.4472              |
| ILS       | High              | 2.1497              | 2.0500              | 0.5748              | 0.3791              | -0.2834             | 0.2515              |
| Low       |                   | 2.1436              | -18.1286            | 1.1170              | -18.3237            | 184.5129            | 143.51              |
| 3SLS      | High              | 5.8651              | -12.5154            | -65.4724            | 2.1503              | -0.3587             | -15.6574            |
| Low       |                   | 0.4381              | 11.0323             | -1.3123             | -2.8065             | 9.8823              | 69.7537             |
| FIML      | High              | 1.7903              | 5.4472              | -18.3167            | 1.6277              | 3.8023              | -9.6204             |

Source: Author’s compilation from TSP.
Table 2. Root Mean Square Error of parameter estimates $T = 30, r = 100$

| Estimator | Level of Multicollinearity | EQUATION1 | EQUATION2 |
|-----------|---------------------------|-----------|-----------|
|           | $\beta_{12}$ | $\gamma_{11}$ | $\gamma_{12}$ | $\beta_{21}$ | $\gamma_{22}$ | $\gamma_{23}$ |
| OLS       | Low          | 1.2881     | 2.7177     | 2.4187 | 1.5524     | 0.6399     | 14.7179     |
|           | High         | 1.2796     | 4.2254     | 0.7746 | 1.5583     | 0.7229     | 14.5329     |
| L23       | Low          | 12.4412    | 270.2402   | 65.7437 | 203.5729 | 1856.869   | 1574.384   |
|           | High         | 51.7673    | 141.3109   | 674.9163 | 2.2227   | 5.3284     | 26.1284     |
| ILS       | Low          | 0.4347     | 17.3921    | 5.9896 | 0.6341     | 3.3932     | 7.3309      |
|           | High         | 1.2087     | 16.0237    | 12.2803 | 1.0190     | 4.6779     | 8.1863      |
| FIML      | Low          | 2.0904     | 22.8463    | 28.6133 | 47.4228    | 96.9632    | 826.26      |
|           | High         | 9.0280     | 52.4830    | 148.228 | 3.9580     | 25.9361    | 53.7650     |

Source: Author's compilation from TSP.
4. Results and discussion
By the Order and Rank conditions of identification, the two equations in the model are exactly identified. This makes the prospect of unique estimates of the parameters realizable. We employed 2 multicollinearity scenarios in studying the relative performance of the estimators. These are:

(i) Low multicollinearity—Where correlation among the exogenous variables are insignificant at 5% level.

(ii) High multicollinearity—Where correlation among the exogenous variables are significant at 1% level.

To empirically utilize the simulation framework, a finite data set of sample size \( T = 30 \) over 100 replications \( (R = 100) \) was simulated to facilitate an exploration of the two multicollinearity scenarios highlighted above. Table 1 highlights averages of the parameter estimates generated for the 100 replications considered for the chosen sample size. The true values of the parameters are presented in parenthesis. Best estimators are those whose average estimates are closest to the true parameter value. For parameter \( \beta_{12} \) whose true value is 1.8, the best estimates are obtained from ILS for both low and high multicollinearity. Similarly, for parameter \( \beta_{21} \) whose true value is 0.4, the best estimates are obtained from ILS for both low and high multicollinearity.

As confirmed by Johnston (1991), when an equation is just identified, estimates of parameters obtained by 2SLS, LIML, ILS and 3SLS should be identical. The results obtained in this study show that 2SLS, LIML and 3SLS estimators yielded virtually identical estimates as can be seen from Table 1. The other three estimators: OLS, ILS and FIML yielded results that are clearly different from those of 2SLS, LIML and 3SLS. Consequently, we classified our original six estimators into 4 categories for the purpose if this analysis. These categories are: OLS—Ordinary Least Squares estimator; L23—Limited Information Maximum Likelihood estimator, Two Stage Least Squares estimator and Three Stage Least Squares estimator; ILS—Indirect Least Squares estimator; and FIML—Full Information Maximum Likelihood estimator. Table 2 presents further evaluation of the parameter estimates obtained through our Monte Carlo Experiment. Root Mean Square Error of Estimates (RMSE) was used to evaluate the relative performance of our estimators in the presence of multicollinearity as stated above. RMSE criterion takes into consideration the bias as well as the variance of estimates in evaluating the performance of the estimators.

For parameter \( \beta_{12} \), in the case of low multicollinearity, ILS performed best with the lowest RMSE of 0.4347 followed by OLS with the RMSE of 1.2881. Also, for high multicollinearity, ILS performed best with the lowest RMSE of 1.2087 followed by OLS with the RMSE of 1.2796. Similar result is obtained for parameter \( \beta_{21} \). For parameter \( \gamma_{11} \), in the case of low multicollinearity, OLS performed best with the lowest RMSE of 2.7177. Also for high multicollinearity, OLS performed best with the lowest RMSE of 4.2254. Similar result like that of \( \gamma_{11} \) was obtained for parameters \( \gamma_{12} \) and \( \gamma_{22} \). For parameter \( \gamma_{23} \), and the case of low multicollinearity, ILS performed best with the lowest RMSE of 7.3309. Also, for high multicollinearity, ILS performed best with the lowest RMSE of 8.1863. In summary, ILS followed by OLS yielded best estimates for \( \beta_{12} \) and \( \beta_{21} \) which are the parameters of the endogenous variables. On the contrary for the parameters of the exogenous variables, OLS is best for \( \gamma_{11} \), \( \gamma_{12} \) and \( \gamma_{22} \) while ILS is best for \( \gamma_{23} \). These observations are surprisingly the same for both low and high measures of multicollinearity.

5. Conclusion and recommendations
We set up a Monte Carlo Experiment (MCE) to evaluate the relative performance of various estimators of a simultaneous equation model in the presence of varied levels of multicollinearity. Our model was estimated with a simulated data set of sample size 30 over 100 replications. The parameter estimates obtained from the six estimators considered were evaluated using Root Means Squares of Estimates (RMSE) criteria. Our result revealed that irrespective of the level of multicollinearity in our model, ILS followed by OLS yielded best estimates for \( \beta_{12} \) and \( \beta_{21} \) which are
the parameters of the endogenous variables. On the contrary for the parameters of the exogenous variables, OLS is best for $\gamma_{12}$, $\gamma_{13}$ and $\gamma_{23}$ while ILS is best for $\gamma_{12}$. However, the system estimators all performed poorly. Also 2SLS, LIML and 3SLS estimators yielded virtually identical estimates. By our findings, in the presence of multicollinearity, estimators OLS and ILS performed best and should therefore be preferred above the multi-equation estimators.

**Funding**
The authors received no direct funding for this research.

**Author details**
Emmanuel A. Oduntan
E-mail: emmanuel.oduntan@covenantuniversity.edu.ng

J. O. Iyaniwura

1 Department of Economics and Development Studies, Covenant University, Ota, Ogun State, Nigeria.

2 Department of Statistics, University of Ibadan, Ibadan, Oyo State, Nigeria.

**Citation information**
Cite this article as: A Monte Carlo simulation framework on the relative performance of system estimators in the presence of multicollinearity, Emmanuel A. Oduntan & J. O. Iyaniwura, Cogent Social Sciences (2021), 7: 1926071.

**References**
Boldey, R. (1980). A Monte Carlo study of small sample property of simultaneous equation estimators with normal and non-normal disturbances. *Journal of the American Statistical Association*, 75(369).

Carlin, B. P., Poislon, N. G., & Stoffer, D. S. (1992). A Monte Carlo approach to non-normal and non-linear state-space modeling. *Journal of the American Statistical Association*, 87(413), 493–500. [https://doi.org/10.1080/01621459.1992.10475231](https://doi.org/10.1080/01621459.1992.10475231)

Essi, N. (1997). The consequences of wrong specification of error terms on the robustness of the estimators of non-linear econometric model [An Unpublished PhD thesis submitted to the department of Statistics]. University of Ibadan.

Hendry, D. F. (1984). The structure of simultaneous equations estimators. *Journal of Econometrics*, 26, 51–88. [https://doi.org/10.1016/0304-4076(76)90017-8](https://doi.org/10.1016/0304-4076(76)90017-8)

Johnston, J. (1991). *Econometrics methods*. McGraw – Hill Book Company.

Kloek, T., & Van Dijk, H. K. (1978). Bayesian estimates of equation system parameters: An application of integration by Monte Carlo. *Econometrica*, 46(1), 1–19. [https://doi.org/10.2307/1913641](https://doi.org/10.2307/1913641)

Kmenta, J., & Joseph, M. E. (1963). A Monte Carlo study of alternative estimates of the Cob-Douglas production function. *Econometrica*, 31(3), 363–390. [https://doi.org/10.2307/1909977](https://doi.org/10.2307/1909977)

Lurie, D., & Hartley, H. O. (1972). Machine generation of order statistics for Monte Carlo Computations. *The American Statistician*, 26(1), 26–27.

Nagar, A. L. (1960). A Monte Carlo study of alternative simultaneous equation estimators. *Econometrica*, 28(3), 573–590. [https://doi.org/10.2307/1910132](https://doi.org/10.2307/1910132)

Nwabueze. (2000). A single-equation study of the choice of estimators of nine autocorrelate models [An unpublished Ph.D thesis submitted to the Department of Statistics]. University of Ibadan.

Wagner, H. M. (1958). A Monte Carlo study of estimates of simultaneous linear structural equations. *Econometrica*, 26(1), 117–133. [https://doi.org/10.2307/1907386](https://doi.org/10.2307/1907386)
