JOURNAL OF STATISTICAL MODELING & ANALYTICS (JOSMA)
(ISSN: 2180-3102)
UNIVERSITI MALAYA

Monte Carlo Experiment on the Asymptotic Efficiency of System Estimators under Multicollinearity

Emmanuel A. Oduntan

Department of Mathematics and Statistics, American University of Nigeria. Yola, Adamawa State. Nigeria.

Corresponding Author: emmanuel.oduntan@aun.edu.ng

RESEARCH ARTICLE

Received 1 November 2022
Accepted 20 February 2023
Published 25 October 2023

Abstract

Classical linear regression model assumes that there is no multicollinearity among the explanatory variables in a regression model. Contrary to this assumption, where multicollinearity is perfect, the regression coefficients of the explanatory variables are indeterminable and their standard errors are infinite. On the other hand, where multicollinearity is less than perfect, the regression coefficients, although determinable, possess large standard errors. This implies that the coefficients cannot be estimated with great precision. Hence multicollinearity problem is a major problem in econometric analysis. Using Monte Carlo Simulation, we evaluated the asymptotic efficiency of six estimators (OLS, ILS, 2SLS, 3SLS, LIML and FIML), under different magnitudes of the unintended linear relationship between the exogenous variables. Using the SSR criteria, we found that OLS followed by ILS turned out the best estimates amongst the six estimators under multicollinearity. We also found that with increasing sample size, there is no remarkable asymptotic effect in the performance of the estimators at the levels of multicollinearity.

Keywords: Asymptotic Efficiency, Correlation, Estimators, Monte Carlo, Multicollinearity, Simultaneous Equation

1. Introduction

Multicollinearity or the presence of linear relationship between pairs of exogenous variables is an important problem in econometrics, especially in single-equation model estimation. When multicollinearity is present in a single equation model estimation, a common empirical solution is for a simultaneous equation model to be adopted, if this is admissible to theory. However, we may still be confronted with multicollinearity in the individual equation in the simultaneous system. If the single-equation solution to this problem is adopted, there may be an intolerable rise in the size of the model with the consequent depletion of the number of exogenous variables some of which are usually included for the purpose of studying the effect of shocks and for policy simulation. It is a common practice for researchers, especially when undertaking Monte Carlo studies to resort to the use of randomly generated values of exogenous variables. The values obtained are invariably not examined for multicollinearity before they are used in such studies. Hence, many researchers make the potentially wrong assumption that the randomly generated values are uncorrelated. The results generated on the assumption they are inherently orthogonal, may be misleading. Where the variables are correlated, the multicollinearity problem may occur. In view of the possibility of multicollinearity problem affecting results from such studies, inferences drawn from them may be impaired. Consequently, the conclusions based on such results may be misleading.

In this paper, we considered the comparative efficiency of six estimators under two different magnitudes of the unintended correlation between the exogenous variables using the Monte Carlo approach. Furthermore, the interest of researchers in Monte Carlo studies may be the analysis of the performance of estimators with increasing sample size with a view to validating the asymptotic efficiency of parameter estimates. The asymptotic efficiency of the estimators under multicollinearity is also studies.

Consider the Ordinary Least Square (OLS) estimator is given as:

$$\mathbf{b} = (\mathbf{X}^1\mathbf{X})^{-1}\ \mathbf{X}^1\mathbf{y}$$

and its variance matrix is:

$$\mathbf{Var(b)} = \boldsymbol{\sigma}^2\ (\mathbf{X}^1\mathbf{X})^{-1}$$

The sampling variances **Var(b)** depend not only on the error variances σ^2, but also on the sample values of explanatory variables. Where the explanatory variables are non-orthogonal, the correlation between them will be high. In such a case, the numerical values of the off-diagonal terms will be high.

When the explanatory variables are orthogonal, the coefficients of the **X**'s in the multiple regression equation would be the same as those given by the sample regression of **Y** on each **X** in turn. Orthogonal variables may be set up in experimental designs but such variables are not common in economic data. Increasing correlation between two explanatory variables will results in an increased of numerical values for the off-diagonal terms as well as a dramatic fall in the value of the determinant. This is described as a situation of multicollinearity between the explanatory variables.

Monte Carlo methods are commonly used in studying the small sample properties of estimators. In a given study, they are applicable in the determination of the impact of multicollinearity, serial correlation, sample size, choice amongst alternative estimators, etc., on the different estimators, Wagner (1958), Balder Raj (1980), Kloek and Dijk (1978), Kmenta and Joseph (1963), Nagar (1960), Carlin et al (1992). Monte Carlo approach provides a "laboratory environment" for performing controlled experiments on estimators.

Pingel and Waernbaum (2015) in their study on how the asymptotic efficiency of matching and inverse probability weighting estimators for average causal effects change when the covariates are correlated, stated that "correlation can both increase and decrease the large sample variances of the estimators and that the correlation affects the asymptotic efficiency of the estimators differently with regards to its magnitude and direction". Studying relative asymptotic efficiency of estimators for non-linear time series, Amano (2009), found that through numerical analysis, for important nonlinear time series models, e.g. GARCH, RCA, non-linear AR models, the asymptotic variance of the G estimator proposed by Chindra and Taniguchi (2001) is smaller than that of the conditional least squares (CLR) estimators. Grobler (2018) considered the effect of the choice of weight function on the efficiency of the empirical characteristic function (ECF) estimator. He analyzed the asymptotic efficiency relative (ARE) to maximum likelihood estimators of estimators using a weight function that is either a normal density or a mixture of two normal densities. The study results indicate that ARE is sensitive to the choice of weight function.

In their paper on Asymptotic Efficiency of Maximum Likelihood Estimators Under Misspecified Models, Ghosh and Sing (2012) illustrates when and how maximum likelihood estimators continue to be asymptotically efficient under mis-specified models. They also provided necessary and sufficient conditions under which a subset of the vector of MLE's retains its asymptotic efficiency under misspecified models where the MLE itself is not fully asymptotically efficient. Also, Lahiri, Lee and Cressie (2002), investigated the efficiency of different least squares variogram-parameter estimators in large samples. They posited that "when the number of lags used to define the estimates is chosen to be equal to the number of variogram parameter to be estimated, the ordinary least squares, the weighted least squares and the generalized least squares estimators are all asymptotically efficient". Chambers (2003) examined the effects of temporal aggregation on the asymptotic variance of estimates in cointegrated systems. The study found that estimators based on flow data alone are more efficient than when the data are all stocks and flows. The study further revealed that estimators based on flow data are as efficient as when the data are recorded continuously. Also, Emanuele Dolera, (2022), examined the asymptotic efficiency of point estimators in Bayesian Predictive Inference.

Amano and Taniguchi (2008), from the standpoint of the conditional least squares (CL) estimators proposed by Tjistheim (1986) which are not asymptotically efficient in general, developed a necessary and sufficient condition that CL is asymptotically efficient based on the local asymptotic normality (LAN) approach. Aoki, Bolfarine and Singer (2002) considered measurement error regression models with null intercepts. They derived explicit method of moment estimators (MME) under models with increasing complexity and also evaluated their asymptotic relative efficiencies with respect to the corresponding maximum likelihood estimators (MLE). They identified cases where the MME may be considered as substitutes of the MLE. Takeuchi and Akahira (2003) stated that "the maximum likelihood estimation (MLE) is a third order asymptotically efficient in multi-parameter exponential cases". Other related studies are Wilson (2011), Marlin and Nando de Freitas (2012), Li and Fearnhead (2018), Janon et.al. (2014), Ermakov (2012), and Wilson (2012)].

In this study we evaluated the performance and asymptotic efficiency of six estimation methods under positive multicollinearity using Monte Carlo Simulation. The rest of the paper is organized into three sections. Section 2 presents the Materials and Methods of the study. The findings from our Monte Carlo experiment are presented in Section 3 while Section 4 is the conclusion.

2. Materials and Methods

We adopted the Monte Carlo simulation strategy used by Oduntan and Iyaniwura (2021). This is presented below.

2.1 Theoretical Framework

Consider

$$\mathbf{Y} = \mathrm{F}(\boldsymbol{X}, \boldsymbol{\theta}) + \mathbf{u} \tag{1}$$

where $\mathbf{Y}$ is the vector of endogenous variables, $\boldsymbol{X}$ is the vector of the exogenous variable, $\boldsymbol{\theta}$ is the vector of the structural parameters and $\boldsymbol{u}$ is the vector of the disturbance terms. $\boldsymbol{u} \sim N(0,\ \sigma^2)$ and vector $\boldsymbol{\theta}$ as well as the variance σ^2 are assigned numerical values. Normal deviates are selected on the basis of the assumed σ^2, and used in generating error terms $\boldsymbol{u}$. We selected a random sample of size T for $\boldsymbol{X}$ and computed the numerical values of $\mathrm{F}(\boldsymbol{X}, \theta)$. Computing $\boldsymbol{Y} = \mathrm{F}(\boldsymbol{X}, \boldsymbol{\theta}) + \boldsymbol{u}$, we obtained vector $\boldsymbol{Y}$

Estimate $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is obtained by performing the regression of $\boldsymbol{Y}$ on $\boldsymbol{X}$. Replicating this procedure several times in order to facilitate the construction of the sampling distribution of $\widehat{\boldsymbol{\theta}}$. We used the empirical distribution through this procedure in evaluating the precision of $\widehat{\boldsymbol{\theta}}$ as well as comparisons of the performance of different estimators of $\boldsymbol{\theta}$.

2.2 Empirical Strategy

Consider,

$$y_{1t} = \beta_{12}y_{2t} + \gamma_{11}X_{1t} + \gamma_{12}X_{2t} + u_{1t}$$
$$y_{2t} = \beta_{21}y_{1t} + \gamma_{22}X_{2t} + \gamma_{23}X_{3t} + u_{2t} \tag{2}$$

where $y's$ are the endogenous variables, $x's$ are the exogenous variables and $u's$ are the disturbance terms. In matrix form equation 2 becomes:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}$$

where $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, $\boldsymbol{X} = \begin{bmatrix} 1 & 1 & 1 \\ X_1 & X_2 & X_3 \\ 1 & 1 & 1 \end{bmatrix}$, $\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix}$ and $\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$.

A linear simultaneous equation econometric model containing K predetermined variables $X_1, X_2, \ldots, X_k$, and G endogenous variables $y_1, \ldots, y_G$, the G structural equations at time t may, be expressed as:

$$\beta_{i1}y_{1t} + \beta_{i2}y_{2t} + \ldots\ldots + \beta_{1G}y_{Gt} + \gamma_{i1}x_{1t} \ldots\ldots + \gamma_{ik}x_{kt} = v_{it}$$
$$i = 1,2 \ldots\ldots, G, \quad t = 1,2 \ldots., n \tag{3}$$

Equation 3 can be rewritten as:

$$\boldsymbol{B}\boldsymbol{y}_t + \boldsymbol{\Gamma}X_t = \boldsymbol{u}_t \tag{4}$$

where $\boldsymbol{\Gamma}$ is a $G\ by\ K$ matrix of coefficient of predetermined variables y_t, $\boldsymbol{B}$ is a $G\ by\ G$ matrix of coefficients of current endogenous variables, with $\boldsymbol{X}_t$ and $\boldsymbol{u}_t$ being column vectors of G and K elements respectively.

$$\boldsymbol{B} = \begin{bmatrix} \beta_{11} & \beta_{12} \cdots & \beta_{1G} \\ \beta_{21} & \beta_{22} \cdots & \beta_{2G} \\ & & \vdots \\ \beta_{G1} & \beta_{G2} \cdots & \beta_{GG} \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} \gamma_{11} & \gamma_{21} \cdots & \gamma_{1K} \\ \gamma_{21} & \gamma_{22} \cdots & \gamma_{2K} \\ & & \vdots \\ \gamma_{G1} & \gamma_{G1} \cdots & \gamma_{GK} \end{bmatrix}, \quad \boldsymbol{y_t} = \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{Gt} \end{bmatrix}, \quad \boldsymbol{X}_t = \begin{bmatrix} X_{1t} \\ X_{2t} \\ \vdots \\ X_{Kt} \end{bmatrix}, \boldsymbol{u_t} = \begin{bmatrix} u_{1t} \\ u_{2t} \\ \vdots \\ u_{Gt} \end{bmatrix}$$

The assumptions about the model are:

1. $\boldsymbol{Y}$ is expressed as a linear combination of the explanatory $\boldsymbol{X}$ variables plus a disturbance vector.
2. At any time t,

$$E(\boldsymbol{u_t}) = 0 \tag{5}$$

3. The Variance-Covariance matrix of $\boldsymbol{u}$ is:

$$E(\boldsymbol{u_t}\boldsymbol{u_t'}) = \Sigma = \Omega \otimes l_n \tag{6}$$

$$E(\boldsymbol{u_t u_t'}) = \Sigma = \begin{bmatrix} \begin{bmatrix} u_{11} \\ u_{12} \\ \vdots \\ u_{1n} \\ u_{21} \\ \vdots \\ u_{2n} \\ \vdots \\ u_{G1} \\ \vdots \\ u_{Gn} \end{bmatrix} [u_{11} \; u_{12} \; \dots \; u_{1n} \; u_{21} \; \dots \; u_{2n} \; u_{G1} \; \dots \; u_{Gn}] \end{bmatrix} \quad (7)$$

The elements of the Gn by Gn symmetric matrix in equation 7 are given by:

$$E(u_{it}u_{i^*t^*}) = \begin{cases} \sigma_t^2, & for \;\; i = i^*, \;\; t = t^* \quad (a) \\ 0, & for \;\; i = i^*, \;\; t \neq t^* \quad (b) \\ \sigma_{ii^*}, & for \;\; i \neq i^*, \;\; t = t^* \quad (c) \\ 0, & for \;\; i \neq i^*, \;\; t \neq t^* \quad (d) \end{cases}$$

where, (a), (b), (c), and (d) - implies respectively, existence of homoscedasticity, absence of autocorrelation, covariances at the same periods are equal to σ_{it^*}, and non-contemporaneous covariances are zero.

4. $\boldsymbol{X}$ is a non-stochastic matrix with a full column rank i.e., $\rho(\boldsymbol{X}) = K$
5. $\boldsymbol{u} \backsim NID\,(0, \; \Sigma)$

To obtain the reduced-form equations we solved equation 4 for y_t, i.e.,

$$y_t = -\beta^{-1}\Gamma X_t + \beta^{-1}u_t = \Pi X_t + V_t$$

where $\boldsymbol{\Pi} = -\beta^{-1}\Gamma$, $V_t = \beta^{-1}u_t$, and $(V_tV_t') = \beta^{-1}\Sigma\beta^{-1} = \boldsymbol{\Omega}$.

$(\boldsymbol{\beta}, \; \boldsymbol{\Gamma}, \; \boldsymbol{\Sigma})$ represents the parameters of the structural equations, while $(\boldsymbol{\Pi}, \; \boldsymbol{\Omega})$ are the reduced form equation parameters.

$$\boldsymbol{\Pi} = -\beta^{-\mathbf{1}}\Gamma \text{ and } \boldsymbol{\Omega} = \beta^{-1}\Sigma\beta^{-1'}$$

We adopted the Monte Carlo simulation strategy proposed by Oduntan and Iyaniwura (2021) to obtain a conformable data set. This strategy is set out as follows.

i) Set sample size T at 30.

ii) Assigning arbitrary values to the structural parameters of the model, we have;

$$\beta_{12} = 1.8, \quad \gamma_{11} = 1.2, \quad \gamma_{21} = 0.6$$
$$\beta_{21} = 0.4, \quad \gamma_{12} = 0.5, \quad \gamma_{23} = 1.4 \quad (8)$$

iii) Also, assigning arbitrary numerical values to Ω at any given sample point, we have.

$$\boldsymbol{\Omega} = \begin{bmatrix} 4.5 & 3.0 \\ 3.0 & 3.5 \end{bmatrix} \quad (9)$$

iv) From a pool of uniformly distributed random numbers we selected the values of the predetermined variables X_{1t}, X_{2t} and X_{3t}, such that the correlation coefficients, $r_{(x_1x_2)}, r_{(x_1x_3)}$ and $r_{(x_2x_3)}$ are of the following magnitudes;
   (a) significant at 1% level: designated as high multicollinearity.
   (b) insignificant at 5% level: designated as low multicollinearity.

v) We obtained u_{1t} and u_{2t} via the following process;
   (a) From a large collection of random normal deviates we drew independent series $\varepsilon_t \sim N(0, 1)$, with the size of each pair depending on the desired sample size.

(b) We adopted the method proposed by Nagar (1969) to transformed the series ε_t into a series of random errors that guarantees conformity with the variance-covariance matrix **Ω** predetermined for the model. This procedure is highlighted as follows.
Let **Σ** be decomposed by a matrix **P** as;

$$\boldsymbol{\Sigma} = \boldsymbol{PP}^{/} \tag{10}$$

where

$$\mathbf{P} = \begin{pmatrix} S_{11} & S_{21} \\ 0 & S_{22} \end{pmatrix} \tag{11}$$

Then

$$\begin{aligned} S_{22} &= +\sqrt{\sigma_{22}} \\ S_{21} &= {}^{\sigma_{12}}/_{S_{22}} \\ S_{11} &= +\sqrt{(\sigma_{11} - S_{21}^2)} \end{aligned} \tag{12}$$

We have the two random error series using

$$\begin{aligned} \boldsymbol{u} = \boldsymbol{P}\varepsilon_t &= \begin{pmatrix} u_{t1} \\ u_{t2} \end{pmatrix} \\ &= \boldsymbol{P}\begin{pmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \end{pmatrix} \\ &= \begin{pmatrix} S_{11} & S_{21} \\ 0 & S_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \end{pmatrix} \end{aligned} \tag{13}$$

Hence,

$$\begin{aligned} u_{t1} &= S_{11}\,\varepsilon_{t1} + S_{21}\,\varepsilon_{t2} \\ u_{t2} &= S_{22}\,\varepsilon_{t2} \end{aligned} \tag{14}$$

vi) With the reduced form of the model, we generate our T endogenous variables from the values already obtained for the $X's$ and $U's$ and the values assigned to the structural parameters.
For our two-equation model

$$\begin{aligned} y_{1t} &= \beta_{12}y_{2t} + \gamma_{11}X_{1t} + \gamma_{1t}X_{2t} + u_{1t} \\ y_{2t} &= \beta_{21}y_{1t} + \gamma_{21}X_{2t} + \gamma_{2t}X_{3t} + u_{2t} \end{aligned}$$

Rearranging the model, we have,

$$\begin{aligned} y_{1t} - \beta_{12}y_{2t} - \gamma_{11}X_{1t} - \gamma_{2t}X_{2t} - 0X_{3t} = u_{1t} \\ -y_{2t}y_{1t} + y_{2t} - 0X_{1t} - \gamma_{21}X_{2t} - \gamma_{23}X_{3t} = u_{2t} \end{aligned}$$

Which we rewrite as:

$$\boldsymbol{By}_t + \boldsymbol{\Gamma X}_t = u \tag{15}$$

where

$$\boldsymbol{B} = \begin{bmatrix} 1 & -\beta_{12} \\ -\beta_{21} & 1 \end{bmatrix},\ \boldsymbol{\Gamma} = \begin{bmatrix} -\gamma_{11} & -\gamma_{12} & 0 \\ 0 & -\gamma_{21} & -\gamma_{23} \end{bmatrix},\ \boldsymbol{y}_t = \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix},\ \boldsymbol{X}_t = \begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix},\quad \boldsymbol{u} = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

Rewriting equation (15), we have

$$\boldsymbol{y}_t = -\boldsymbol{B}^{-1}\boldsymbol{\Gamma X}_t + \boldsymbol{B}^{-1}u \tag{16}$$

$$= -\frac{1}{1-\beta_{12}\beta_{21}}\begin{bmatrix} 1 & \beta_{21} \\ \beta_{12} & 1 \end{bmatrix}\begin{bmatrix} -\gamma_{11} & -\gamma_{12} & 0 \\ 0 & -\gamma_{21} & -\gamma_{23} \end{bmatrix}\begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} + \frac{1}{1-\beta_{12}\beta_{21}}\begin{bmatrix} 1 & \beta_{21} \\ \beta_{12} & 1 \end{bmatrix}\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

where $\boldsymbol{B}^{-1} = \frac{1}{1-\beta_{12}\beta_{21}}\begin{bmatrix} 1 & \beta_{21} \\ \beta_{12} & 1 \end{bmatrix}$.

Thus, we have

$$y_{1t} = \left[\frac{\gamma_{11}}{1-\beta_{12}\beta_{21}}\right]\boldsymbol{X}_{1t} + \left[\frac{\gamma_{12} + \beta_{21}\beta_{21}}{1-\beta_{12}\beta_{21}}\right]\boldsymbol{X}_{2t} + \left[\frac{\beta_{12}\beta_{23}}{1-\beta_{12}\beta_{21}}\right]\boldsymbol{X}_{3t} + \left[\frac{\varepsilon_{1t} + \beta_{21}u_{2t}}{1-\beta_{12}\beta_{21}}\right]$$

$$y_{2t} = \left[\frac{\gamma_{11}\beta_{21}}{1-\beta_{12}\beta_{21}}\right]\boldsymbol{X}_{1t} + \left[\frac{\gamma_{11}+\beta_{21}\gamma_{21}}{1-\beta_{12}\beta_{21}}\right]\boldsymbol{X}_{2t} + \left[\frac{\gamma_{23}}{1-\beta_{12}\beta_{21}}\right]\boldsymbol{X}_{3t} + \left[\frac{\beta_{12}u_{1t}+u_{1t}}{1-\beta_{12}\beta_{21}}\right] \quad (17)$$

We used equation 17 to obtain the values of the endogenous variables.

vii) Finally, we conclude the Monte Carlo experiment by estimating the structural parameters using the generated data sets for $y_{1t}, y_{2t}, y_{3t}, X_{1t},\ X_{2t},\ and\ X_{3t}$. The estimators used are ILS (Indirect Least Squares), OLS (Ordinary Least Squares), 3SLS (Three Stage Least Squares), 2SLS (Two Stage Least Squares), LIML (Limited Information Maximum Likelihood) and FIML (Full Information Maximum Likelihood).

3. Findings

Under different magnitudes of multicollinearity, we analyzed the performance of six estimators using the Sum of Squared Residuals (SSR) of parameter estimates. In theory, for just identified equations, the parameter estimates obtained via 2SLS, LIML, ILS and 3SLS should be identical, Johnston (1991) – our results are in agreement with this. The results of 2SLS, LIML and 3SLS yielded virtually identical parameter estimates. The other three estimators (OLS, ILS and FIML) generated virtually different estimates. Hence, four estimators OLS, ILS, FIML and L23 **(**LIML, 2SLS and 3SLS)] shall be compared in this analysis.

Table 1 shows the SSR of estimates for sample size *T* = 20, 30, 40, 50, 60 over 150 replications. The coverage of the sample sizes *T* considered here is believed to be in conformity with the generally acceptable small sample size for empirical studies. The SSR are presented on equation basis under both levels of multicollinearity considered for the estimators. The best estimators are those with the least SSR.

Table 1: SSR of Parameter Estimates over 150 replications

Sample Size T	Multi-collinearity	OLS		L23		ILS		FIML	
		EQ_1	EQ_2	EQ_1	EQ_2	EQ_1	EQ_2	EQ_1	EQ_2
20	Low	101.11	13.80	805077	26758	1877	7195	37390	2573
	High	81.22	56.63	315746	1469	18.33	7062	246232	293.27
30	Low	95.29	103.09	871069	75015283	2399	9110	106671	365188
	High	146.63	23.56	24182206	18442	2379	9157	450827	33554
40	Low	175.90	30.20	10659819	7379	3061	11853	65180910	1.26E+8
	High	187.97	37.32	27743929	271568	2829	11105	27412	4430170
50	Low	202.69	30.00	2821906	16782	3396	13276	8.7E+11	1.95E+8
	High	170.84	41.00	1481103	2025531	3599	14009	24580	225.97
60	Low	294.46	556.19	701312	947817	4049	16049	528438	2378
	High	202.06	203.85	480853	3907	4213	16343	110790	66203

Source: Author's compilation from estimation results

In Table 1 the first entry 101.11 represents the OLS SSR for equation 1 under low multicollinearity for sample size $T = 20$ over 150 replications. The same process is repeated for the 2 equations, over 6 estimators and across 2 levels of multicollinearity. We further deduced that at both levels of multicollinearity, OLS performed best across different sample sizes as compared to other estimators in both equations. While ILS performed next to OLS across different sample sizes for equation 1, it did not produce good performance for equation 2. For equation 1, at both low and high multicollinearity, FIML and L23 turned out large SSR thereby displaying poor performance among the estimators. The SSR values for ILS at both level of multicollinearity increased progressively from sample size 20 to sample size 60. All other estimators did not exhibit any remarkable asymptotic effect.

We evaluate the sensitivity of estimates by the six estimators to increase in the magnitude of multicollinearity. This is repeated for each of the five sample sizes. The objective of this evaluation is to examine the relative performance of the six estimators as different levels of multicollinearity. Thus, the SSR generated by the six methods were evaluated with a view to identifying the best on the basis of the following criteria:

a) Best Performance: The best estimator is indicated followed by the least performers in parenthesis.
b) Estimators with large SSR: These are arbitrarily set at SSR greater than 5000.

On the basis of the above criteria, Table 1 is further summarized as in Table 2.

Table 2: Performance of Estimators, viz-a-viz, Equations and Incidence of Multicollinearity using SSR for different sample sizes over 150 replications

Sample Size	Performance	Multicollinearity (EQ_1)		Multicollinearity (EQ_2)	
		Low	High	Low	High
20	Best Performer	O (L23)	O (L23)	O (L23)	O (L23)
	Large SSR	L23F	L23F	IL23	I
30	Best Performer	O (L23)	O (L23)	O (L23)	O (F)
	Large SSR	L23F	L23F	IL23F	IL23F
40	Best Performer	O (F)	O (L23)	O (F)	O (F)
	Large SSR	L23F	L23F	IL23F	IL23F
50	Best Performer	O (F)	O (L23)	O (F)	O (L23)
	Large SSR	L23F	L23F	IL23F	IL23
60	Best Performer	O (L23)	O (L23)	O (L23)	O (F)
	Large SSR	L23F	L23F	IL23	IF

Source: Author's compilation from Table 1.

Using SSR, we evaluate the sensitivity of six estimators to change in the level of multicollinearity and increasing sample size $T = 20$, 30, 40 and 50 while replication is kept constant at 150. The objective of this evaluation is to examine the comparative response of the six estimators to change in the levels of multicollinearity and increase in sample size (asymptotic effect).

The following observations were made on Table 2. For equation 1, OLS performed best at both levels of multicollinearity and across sample size, all other estimators with the exception of OLS turned out large SSR at the different levels of multicollinearity. Also, for equation 2, OLS performed best at both levels of multicollinearity and over all samples.

Using the SSR criterion, for both equations, at both levels of multicollinearity and across sample sizes; OLS estimator recorded overall best performance among the six estimators considered. Only OLS estimator did not generate large SSR among the estimators. This observation is irrespective of the level of multicollinearity. Hence, asymptotically, OLS estimator performed better than the other estimators.

However, there is no remarkable asymptotic effect occasioned by the different levels of positive multicollinearity.

4. Conclusion

The SSR of parameter estimates criteria was applied to estimators generated over ten (10) experimental scenarios. The scenarios are: levels of positive multicollinearity (2 scenarios), sample size (5 scenarios) and replications (1 scenario). From our results, and as expected for just identified models, 3SLS, 2SLS and LIML yielded virtually identical parameter estimates. OLS, followed by ILS performed best at both levels of multicollinearity while LIML, 2SLS, 3SLS and FIML performed poorly. This is attributed to the fact that intra as well as inter equation effects of multicollinearity worsened the performance of system estimators compared to single-equation estimators which only confronted multicollinearity at inter- equation level in the model. Hence, Monte Carlo studies with intentionally or otherwise inclusion of exogenous variables that are inherently correlated would be based on the performance assessment of the six estimators on a poor foundation. It is our suggestion that such exogenous variables that will be used in simultaneous equation model estimation should be screened for the presence of multicollinearity in order to ensure the application of an appropriate estimator.

Furthermore, there is no remarkable asymptotic effect occasioned by the varied level of positive multicollinearity imposed on the system. This is in agreement with the submission of Johnston (1991) that “increasing the size of a data that is plagued by multicollinearity by adding more of multicollinearity plagued data would not result in any remarkable asymptotic effect in parameter estimation”.

5. Acknowledgement

I acknowledge and appreciate the time and valuable inputs of the reviewers of the manuscript of this paper. Their review and inputs have greatly enhanced the quality to the paper.

6. References

Amano M. T. (2009). Asymptotic Efficiency of Estimating function estimators for nonlinear time series models. *Journal of the Japan Statistical Society*. 39(2).

Amano T. and Taniguchi M. (2008). Asymptotic Efficiency of Conditional Least Squares Estimators for ARCH Models. *Statistics and Probability Letters.* 78(2): 179 – 185.

Aoki R., Bolfarine H. and Singer J, M. (2002). Asymptotic Efficiency of Method of Moments Estimators under Null Intercept Measurement Error Regression Models. *Brazilian journal of Probability and Statistics*.16: 157 – 167.

Baldev, Rag. (1980). A Monte Carlo study of Small Sample properties of Simultaneous Equation Estimators with Normal and Non-Normal Disturbances. *Journal of the American Statistical Association.* 75(369).

Carlin, B. P, Polson, N.G and Stoffer, D. S. (1992). A Monte Carlo Approach to Non-Normal and Non-Linear State-space Modeling. *Journal of the American Statistics Association*. 87(413): 493 – 500.

Chambers M. J. (2003). The Asymptotic Efficiency of Cointegration Estimators under Temporal Aggregation. *Econometric Theory.* 19(1).

Emanuele Dolera, (2022). Asymptotic Efficiency of Point Estimators in Bayesian Predictive Inference. *Mathematics.* 10(7).

Ermakov M. (2012). The Sharp Lower Bound of asymptotic Efficiency of Estimators in the Zone of Moderate Deviation Probabilities. *Electronic Journal of Statistics.* 6:2150 – 2184.

Ghosh M. and Sing J. (2012). Asymptotic Efficiency of Maximum Likelihood Estimators Under Misspecified Models. *Pakistan Journal of Statistics and Operation Research.* 8(3).

Grobler G. L. (2018). Asymptotic relative efficiency of empirical characteristic function estimators with normal weighting. *Annual Proceedings of the South African Statistical Association Conference.* 2018(1): 9 – 16.

Gujarati, Damodar N. (2004). *Basic Econometrics*. Tata McGraw-Hill.

Janon A., Klein T., Lagnox A. and Nodet M. (2018). Asymptotic Normality and Efficiency of Two Sobol Index Estimators. *ESAIM Probability and Statistics.* 18(2014): 342 – 364.

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

Kloek, T. and Van Dlik, H. K. (1978). Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo. *Econometrica*. 46(1): 1 – 19.

Kmenta, J. and Joseh, M. E. (1963). A Monte Carlo Study of Alternative Estimates of the Cobb-Douglas Production Function. *Econometrica*. 31(3): 363 – 390.

Lahiri S. N., Lee Y. and Cressie N. (2002). On Asymptotic Distribution and Asymptotic Efficiency of Least Squares of Spatial Variogram Parameters. *Journal of Statistical Planning and Inference.* 103 (1 – 2).

Li W. and Fearnhead P. (2018). On the Asymptotic Efficiency of approximate Bayesian Computation Estimation. *Biomatrika*. 105(2).

Lurie, D and Hartley, H. O (1972). Machine Generation of order Statistics for Monte Carlo Computations. *The American Statistician*. 26(1): 26 – 27.

Marlin B. and Freitas Nando de (2012). Asymptotic Efficiency of Deterministic Estimators for Discrete Energy-Based Models: Ratio Matching and Pseudo-likelihood, *Computer Science – Machine Learning*, 14, February 2012.

Nagar, A. L. (1960). A Monte Carlo Study of Alternative Simultaneous Equation Estimators. *Econometrica*. 28(3): 573 – 590.

Oduntan, Emmanuel A. and Iyaniwura, J. O. (2021). A Monte Carlo Simulation Framework on the Relative Performance of System Estimators in the Presence of Multicollinearity. *Cogent Social Sciences*. 7: 1926071.

Pingel R. and Waernbaum I. (2015). Effect Of Correlation Covariates on The Asymptotic Efficiency of The Matching and Inverse Probability Weighing Estimators for Causal Inference. *Journal of theoretical and applied Statistics.* 49(4).

Takeuchi K. and Akahira M. (2003). Third order Asymptotic Efficiency and Asymptotic Completeness of Estimators. *Joint Statistical Papers of Akahira and Takeuchi*. 191 – 198.

Wagner, H. M. (1958). A Monte Carlo Study of Estimates of Simultaneous Linear Structural Equations. *Econometrica*. 26: 117 – 133.

Wilson P. W. (2011). Asymptotic properties of some Non-Parametric Hyperbolic Efficiency Estimators. *Exploring Research Frontiers in Contemporary Statistics and Econometrics*. 115 – 150.