Nig. J. Pure & Appl. Sci. Vol. 39 (Issue 1, 2026)
e-ISSN 2756-4045
(C) 2026 Faculty of Physical Sciences and Faculty of Life Sciences, Univ. of Ilorin, Nigeria
www.njpas.com.ng

Aysmptotic Bahaviour of Bayesian Seemingly Unrelated Regression Estimators with Different Collinearity Structures

Onatunji, A. P. and Adepoju, A.A.

Department of Statistics, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria
Department of Statistics, University of Ibadan, Ibadan, Nigeria

*Date Received:*20-09-2025
*Date Accepted:*13-03-2026
DOI: https://doi.org/10.48198/NJPAS/25.B34

ABSTRACT

Collinearity is a major drawback for efficiency gain of Seemingly Unrelated Regression (SUR) estimators. Any form of correlation between two explanatory variables defined as collinearity causes efficiency loss and unreliable estimates in SUR model. Use of prior information about SUR model parameters to address the problems associated to different Collinearity Structures (CSs), linear, quadratic, ratio and interactive, has not gained adequate attention in literature. This study was aimed at investigating behaviour of SUR estimators with different CSs under prior information as sample size, n increases from 5, 10, 20, 30, 50, 100, 250 and 500. The Bayesian Estimators (BEs) were derived, with analytically intractable Multivariate Normal Gamma (MNG) and Inverse Wishart (IW) non informative and informative priors. Result revealed that posterior standard deviations of SUR model with ratio collinearity decreased faster compared to other CSs as n increases under Non Informative Prior and Informative Prior. Also, The Bayesian estimators for the SUR model can effectively handle collinearity structures when prior information about unknown parameters is carefully incorporated.

Keywords: Multivariate Normal Gamma, Inverse Wishart, out-of- sample information, informative prior and efficiency loss

Introduction

The Seemingly Unrelated Regression (SUR) model is a system of multiple equations with correlated error terms and uncorrelated explanatory variables (Zellner, 1962). There are two main conditions for the use of SUR; to gain efficiency in estimation by relating information on several equations and to impose or test restrictions involving parameters (see Zellner, 1962, Fiebig, 2001, Srivastava and Giles, 1987). Explanatory variables among many equations in SUR systems of equations are frequently correlated in many real-world scenarios, leading to a contemporaneous correlation error terms across the equations. Since its development, both classical and Bayesian approaches have been in use to address efficiency loss in SUR under violation of its conditions. The efficiency of SUR model involved the application of Aitken's Generalized Least Squares (AGLS)(1962), Ordinary Least Squares (OLS)(Baltagi, 1988) and Feasible Generalized Least Squares (FGLS)(Afolayan *et al.,* 2020) and many more in literature. Its efficiency loss or gain depends on the forms of correlation among variables within or across equations. The violation of the uncorrelated assumption of explanatory variables originates from the data, not the SUR model being used. The presence of correlation between explanatory variables defined as collinearity (Neter *et al.,* 1996), has a significant impact on the SUR estimators' efficiency. Silvey (1969) reported loss of information in SUR when

Corresponding Author: Onatunji, A.P.
Department of Statistics, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria Email: aponatunji@lautech.edu.ng

©2026 Creative Common License Attribution 4.0 International

explanatory variables are correlated. This violation assumption among explanatory variables in the various equations of the SUR estimator resulted into to a loss of efficiency (Binkley, 1982). This made the estimated parameters in SUR to be statistically insignificant (Yahya and Olaifa, 2014). Even when collinearity is present, many authors ignore it which consequently leads to misleading statistical interpretation of results as reported by Kristina *et al.* (2016). Relaxation of uncorrelated assumption on explanatory variables in SUR model has done more harm in terms of inefficiencies when the standard errors of its coefficients are large. For a given SUR, the design matrix of multivariate nature of observations is statistically far from unity which causes biases and unreliable estimates as against the true parameters when sample size is varied. Georges *et al*. (2021) presented SUR with Measurement Error (SURME) in the explanatory variables using Bayesian algorithm based on MCMC and Mean Field Variational Bayes (MFVB) approximation. It was reported that MFVB outperformed MCMC by satisfactorily reduced biases of the SURME model. Ojo *et al.* (2021) presented the Bayesian Analysis (BA) of SUR model with independent priors for parameters; it was revealed that SUR produced a smaller Mean Square Error (MSE) as against the classical methods. Therefore, a normal distribution was considered good to generate disturbance for SUR model. Omid and Mousa (2021) developed BA for SUR model under skew-normal distribution assumption for error terms. The study revealed how well BA performed to fit the SUR model data under normal density of error terms as against skewed normal dsitibution. Ojo and Adepoju (2021) examined the efficiency of Bayesian Estimators (BE) of SUR when regressors were correlated and compared the estimates with that of classical estimators of SUR and OLS. In this study, the MSE of BE for SUR reduced faster than OLS as sample size increased. The superiority of SUR over OLS was investigated when correlated categorical explanatory variables existed in the model using Monte Carlo Standard Error (MCSE) criterion based in BA (Onatunji *et al*. 2020). The BA for SUR biomass with informative prior as against SUR additive biomass model for small sample size was studied (Xie *et al.,* 2022). The developed SUR additive biomass models using BA with informative prior was appropriate to estimate the tree biomass when the data available were very small. Gaussian graphical models to model the conditional independencies among the error terms were studied and were suitable forinformation-sharing among the equation-by-equation models and learning some directed edges among the correlated error terms (Salam and Grzegorczyk, 2022). Esser *et al.* (2024) demonstrated the Bayesian Additive Regression Trees (BARTs) suitable for regression and classification analyses involving multiple correlated outcome variables. This proposed multivariate, an extension of BART was suitable for regression and classification analyses involving multiple correlated outcome variables. This SUR assumption violation requires statistical attention for the purpose of ensuring good model fit in research. Despite enough attention given to it by many researchers in literature, there is more to be done on SUR model especially in BA. Also, there is a dearth of study on different Collinearity Structures (CSs) in SUR in the absence of prior information. This consequently makes existing classical approaches typically inappropriate to effectively address the problems of collinearity due to lack of information about the SUR model parameters. The aim of this work is to examine the asymptotic behaviour of BE for SUR under each collinearity structure. Since there is no established rule of thumb to determine the degree of collinearity, it is appropriate to study SUR model with CSs for efficiency loss or gain.

Materials and Method

Development of Seemingly Unrelated Regression Model

Consider M- equation

$$y_i = X_i\theta_i + \varepsilon_i \qquad i=1,2,...,M \tag{1}$$

Where $y_i, i=1,2,...,M$ is aM-dimensional vector of observations on the ith response variable, X_i is a $n \times K_i$ matrix of explanatory variables, K_i - dimensional vector of regression coefficients including constant and with corresponding ε_i, an M-dimensional vector of disturbance terms.
Equation (1) can be compactly written as

$$y = X\theta + \varepsilon \tag{3}$$

Where y is of dimension ($nM\times1$), X is of dimension ($nM\times K$) with $K=\sum_{i=0}^{M}K_i$, $\theta$is ($K\times1$), and ε is ($nM\times1$), such that $\varepsilon\sim N(0,\sum\otimes I_n)$

Then, Variance covariance (V-C) matrix of disturbance terms is

$$\sum=E\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\\ \vdots\\ \varepsilon_M\end{pmatrix}(\varepsilon_1' \quad \varepsilon_2' \quad \cdots \quad \varepsilon_M')=\begin{bmatrix}\sigma_{11}I_n & \sigma_{12}I_n & \cdots & \sigma_{1M}I_n\\ \sigma_{21}I_n & \sigma_{22}I_n & \cdots & \sigma_{2M}I_n\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{M1}I_n & \sigma_{M2}I_n & \cdots & \sigma_{MM}I_n\end{bmatrix}\Rightarrow$$

$$\begin{bmatrix}\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1M}\\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2M}\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{M1} & \sigma_{M2} & \cdots & \sigma_{MM}\end{bmatrix}\otimes I_n \qquad (4)$$

and I_n is a $n\times n$ unit matrix, $\otimes$ denotes kronecker product for which

$\sum=diag(\sigma_{11},\sigma_{22},\cdots\sigma_{MM})$ is $M\times M$ diagonal matrix with positive diagonal element and $\sum=\sigma_{ij}$, $M\times M$ positive definite matrix, $\sum^{-1}=(\sigma_{ij})^{-1}=\sigma^{ij}$

The Likelihood function of untransformed SUR model is given as follows;

$$P(y)=(2\pi\sigma^2)^{\frac{nM}{2}}(det\sum)^{\frac{1}{2}}\left\{-\tfrac{1}{2}(y-X\hat{\theta})'\sum^{-1}(y-X\hat{\theta})\right\} \qquad (5)$$

Equation (3) written in terms of OLS quantities when there exists an $M\times M$ matrix (P) with the property that $P\Omega P'=I_P$. Multiply equation (3) by P

The transformed SUR model becomes

$$y^*=X^*\theta+\mathrm{E}^* \qquad (6)$$

Here, an identical derivative using the transformed model yields a likelihood function of SUR model written in terms of GLS estimators.

$$P(y^*)=\prod_{i=1}^{n}(2\pi\sigma^2)^{\frac{nM}{2}}det(\sum)^{\frac{n}{2}}exp\left\{-\frac{1}{2}(y^*-X^*\hat{\theta})'\sum^{-1}(y^*-X^*\hat{\theta})\right\}$$

$$\propto exp\text{-}\frac{1}{2}\left\{\sum_{i=1}y^*{}_i'\sum^{-1}y^*{}_i\text{-}2\sum_{i=1}X^*\hat{\theta}'\sum^{-1}y^*{}_i\text{-}\sum_{i=1}X^*\hat{\theta}'\sum^{-1}X^*\hat{\theta}\right\}$$

$$=exp\text{-}\frac{1}{2}\left\{\sum_{i=1}y^*{}_i'\sum^{-1}y_i\text{-}2\sum_{i=1}X^*\hat{\theta}'\sum^{-1}y^*{}_i\text{-}\sum_{i=1}X^*\hat{\theta}'\sum^{-1}X^*\hat{\theta}\right\}$$

$$=exp\text{-}\frac{1}{2}\left\{-2X^*\hat{\theta}'\sum^{-1}n\bar{y}^*\text{-}X^*\hat{\theta}'n\sum^{-1}X^*\hat{\theta}\right\}$$

$$=exp\text{-}\frac{1}{2}\left\{-2\theta'\sum_{i=0}^{n}X^*{}'\sum^{-1}y^*\text{-}\theta'\sum_{i=0}^{n}\left(X^*\sum^{-1}X^*{}'\right)\theta\right\}$$

$$=exp\text{-}\frac{1}{2}\left\{-2\theta'b_1\text{-}\theta'A_1\theta\right\} \qquad (7)$$

Where

$$b_1=\sum_{i=0}^{n}X^*{}'\sum^{-1}y,\ A_1=\sum_{i=0}^{n}\left(X^*\sum^{-1}X^*{}'\right) \qquad (8)$$

$$\bar{y}^*=\frac{1}{n}\left(\sum_{i=1}y^*{}_{i1},\cdots,y^*{}_{n1}\right) \qquad (9)$$

Prior Specification for Seemingly Unrelated Regression Model Parameters $(\theta,\sum)$

It is very important to make appropriate prior for SUR model parameters (Gelman, 2006, Adebayo, 2003) In this work, Independent Normal-Wishart (INW) prior, $\sum\sim IW(v_o,S_o)$ which is anextension of Independent Normal-Gamma (INM) prior and multivariate Normal Prior, $\theta\sim MNG(\theta_o,\mathrm{h})$ are incorporated into the SUR model.

Derivatives of Multivariate Normal Gamma Prior, $\theta\sim MN(\theta_o,h)$ ***for Seemingly Unrelated Regression Model Parameters***

The marginal likelihood for Multivariate Normal Gamma Prior, $\theta\sim MN(\theta_o,\mathrm{h})$ for SUR model parameters, θ

$$P(\theta)=(2\pi)^{-\frac{M}{2}}det(\mathrm{h})^{\frac{1}{2}}exp\left[-\frac{1}{2}(\theta-\theta_o)'\mathrm{h}^{-1}(\theta-\theta_o)\right]$$

$$\propto exp\left[-\frac{1}{2}(\theta-\theta_0)'\mathrm{h}^{-1}(\theta-\theta_0)\right]$$

$$\propto exp[-\tfrac{1}{2}\{\theta'\mathrm{h}^{-1}\theta-2\theta'\mathrm{h}^{-1}\theta_0+\theta_0'\mathrm{h}^{-1}\theta_0\}]$$

$$=exp[-\tfrac{1}{2}\{\theta'\mathrm{h}^{-1}\theta-2\,\theta'\mathrm{h}^{-1}\theta_0\}]$$

$$=exp[-\tfrac{1}{2}\{\theta'A_0\theta-2\theta'b_0\}] \qquad (10)$$

Where;

$$A_0=\mathrm{h}^{-1},\ b_0=\mathrm{h}^{-1}\theta_0 \qquad (11)$$

Posterior Distribution

Posterior distribution summarises the data information and prior about the unknown parameters θ and $\sum$. For a SUR model in equation (1), the posterior simulator is a prior form of distribution, confirming that the previously acquired prior is a natural conjugate prior for θ and $\sum$ (see Gary, 2003). The Posterior distribution is derived by multiplying likelihood function of SUR model parameters by a conjugate prior given as;

$$p(\theta,\sum|y^*)=p(y^*|\ \theta,\sum)p(\theta)p(\sum) \qquad (12)$$

Since there is no closed expression for equation (12), then full conditionals for θ and $\sum$ are

derived to a full conditional distribution for SUR model parameter, θ

$$p(\theta|\, y^*,\Sigma) \propto p(y^*|\, \theta,\Sigma)p(\theta,\mathrm{h}) \propto exp\text{-}\frac{1}{2}\{\text{-}2\theta' b_1\text{-}\theta' A_1\theta\}\times exp[\text{-}\frac{1}{2}\{\theta' A_0\theta\text{-}2\theta' b_0\}]$$

$$\propto exp\text{-}\frac{1}{2}[\{\text{-}2\theta' b_1\text{-}\theta' A_1\theta\}+\{\theta' A_0\theta\text{-}2\theta' b_0\}]$$

$$\propto exp\left\{\theta' b_1\text{-}\frac{1}{2}\theta' A_1\theta\text{-}\frac{1}{2}\theta' A_0\theta+\theta' b_0\right\}$$

$$\propto exp\left\{\theta'(b_0+b_1)\text{-}\frac{1}{2}\theta'(A_0+A_1)\theta\right\} \quad (13)$$

Let

$$\wedge_n=A_0+A_1=\Sigma_n^{\text{-}1}+\Sigma_{i=0}^{n}\left(X^*\Sigma^{\text{-}1}X^{*'}\right) \quad (14)$$

$$\vee_n=b_0+b_1=\Sigma_n^{\text{-}1}\theta_0+\Sigma_{i=0}^{n}X^{*'}\Sigma^{\text{-}1}y^* \quad (15)$$

Hence

$$\theta|\, y^*,\Sigma \sim MN(\wedge_n^{\text{-}1}\vee_n, \wedge_n^{\text{-}1})$$

$$\theta|\, y^*,\Sigma \sim MN(\theta_n,\Sigma_n) \quad (17)$$

where;

$$\theta_n=\wedge_n^{\text{-}1}\vee_n=\left(\mathrm{h}^{\text{-}1}+\Sigma_{i=0}^{n}\left(X^*\Sigma^{\text{-}1}X^{*'}\right)\right)^{\text{-}1}(b_0+b_1) \quad (18)$$

$$=\left(\mathrm{h}^{\text{-}1}+\Sigma_{i=0}^{N}\left(X^*\Sigma^{\text{-}1}X^{*'}\right)\right)^{\text{-}1}\left(\mathrm{h}^{\text{-}1}\theta_o+\Sigma_{i=0}^{n}X^{*'}\Sigma^{\text{-}1}y^*\right) \quad (19)$$

$$\Sigma_n=\vee_n=\left(\mathrm{h}^{\text{-}1}+\Sigma_{i=0}^{n}\left(X^*\Sigma^{\text{-}1}X^{*'}\right)\right)^{\text{-}1} \quad (20)$$

Deriving Marginal Likelihood for Seemingly Unrelated Regression Model Parameters

In the SURM model given in equation (1), this study elicits prior distribution for model parameters, □and the precision, $\mathrm{h}=\frac{1}{\Sigma_n}$ which is given by $p(\theta,\mathrm{h})$, this study therefore, do not condition on the data but on parameters. It is easy to write $p(\theta,\mathrm{h}) = p(\theta|\mathrm{h})p(\mathrm{h})$. where conjugate prior involves a Multivariate Normal distribution for $p(\theta|\mathrm{h})$ and a Gamma distribution for $p(\mathrm{h})$. Multivariate Normal Gamma (MNG) prior which when multiplies with likelihood function of a model data gives a conjugate model (Raifa and Sclaife, 1961). The likelihood function of data vector in equation (5) is used to derive BE of SUR model with collinearity structures.

Lemma 1. Let θ and h follow a density function of the natural conjugate prior, Multivariate normal gamma Prior is a joint prior, which is therefore given as;

$$p(\theta,\mathrm{h})=\frac{\frac{\nu_1+n}{\tau^{\frac{2}{2}}}\text{-}1}{(2\pi)^{\frac{n}{2}}|\mathrm{h}^{\text{-}1}|^{\frac{1}{2}}\Gamma\left(\frac{\nu_1}{2}\right)\left(\frac{2\nu_2}{\nu_1}\right)^{\frac{\nu_0}{2}}}\left\{exp\left[\text{-}\frac{1}{2}(\theta\text{-}\theta_o)'\mathrm{h}^{\text{-}1}(\theta\text{-}\theta_o)+\frac{\nu_1}{\nu_2}\right]\right\} \quad (21)$$

that is;

$$\theta|\mathrm{h}\sim MN(\theta_o,\mathrm{h}^{\text{-}1}\Sigma_n) \quad (22)$$

The marginal distribution of h is a Gamma distribution;

$$\mathrm{h} \sim \text{Gamma}\ (\nu_1,\nu_2) \quad (23)$$

Hence, conjugate prior of θ_n and Σ_n is a joint prior. This implies that;

$$p(\theta_n,\Sigma_n)=p(\theta_n|\Sigma_n)p(\Sigma_n) \quad (24)$$

Since the marginal distribution of θ follows a Multivariate normal distribution, then there exists

$$\theta|\, y^*,\Sigma \sim \mathrm{MN}(\theta_n,\Sigma_n) \quad (25)$$

Proposition 1. If $x|\mu_i,\sigma_i^2 \sim \mathrm{MN}\ (\mu_i,\sigma_i^2)$ and $\sigma_i^2 \sim IG(\nu_1,\nu_2)$, then $\theta|\,\theta_n,\Sigma_n \sim \mathrm{MN}\ (\theta_n,\Sigma_n)$ and $\Sigma_n \sim IG(\nu_1,\nu_2)$ iff $\Sigma_n|\,\theta_1,\ldots\theta_n \sim IG(\nu_1+\frac{n}{2},\ \nu_2+\frac{\varepsilon'\varepsilon}{2})$

$$p(\theta|\Sigma_n)=N(\theta_n,\Sigma_n) \quad (26)$$

$$p(\Sigma_n)=Gamma\ (\nu_1,\nu_2) \quad (27)$$

Reparameterise equation (27) in terms of precision, the conjugate prior follows probability density function of a Gamma distribution.

$$\tau \sim Gamma(\nu_1,\nu_2) \quad (28)$$

Therefore,

Full conditional distribution for θ given τ is;

$$\theta|\, y^*,\Sigma \sim MNG(\theta_n,\tau)$$

Derivatives of Inverse Wishart Prior, $\Sigma \sim IW(V_0,S_0^{-1})$ for Variance Covariance Matrix of Associated Error Terms of Seemingly Unrelated Regression Structure

The random matrix $\Sigma_{(p\times p)} = \Sigma_{i=1}^{n} e_i e_i'$ has the Wishart distribution with *"ν_0"* degree of freedom and covariance matrix (S_0) and is denoted by $\Sigma \sim \mathrm{W}(\nu_0, S_0)$. For $\nu_0 \geq M$, the probability density function of Σ is

$$P(\Sigma) \propto det(\Sigma)^{\frac{(\nu_0+M+1)}{2}} exp\left[\text{-}\frac{1}{2}tr(S_0\Sigma^{\text{-}1})\right] \quad (29)$$

where *tr.* is the standard trace function, and with respect to Lebesque measure on the cone of symmetric positive definite matrices.

Full conditional distribution for variance Covariance of error terms of SUR structure, $\sum$

Sum of square of matrix of a collection of multivariate of vectors of error $e_1,...,e_n$ is given by

$$\sum_{i=0}^{n} e_i e_i^{'} = \varepsilon^{'}\varepsilon \tag{30}$$

Where ε is the $n\times p$ matrix whose *ith* row is $e_i^{'}e_i$ is the $p\times 1$ matrix and $e_i e_i^{'}$ is the $p\times p$ matrix

$$e_i e_i^{'} = \begin{pmatrix} e_{1,1}^2 & e_{i,1}e_{i,2} & \cdots & e_{i,1}e_{i,M} \\ e_{i,1}e_{i,2} & e_{1,2}^2 & & e_{i,2}e_{i,M} \\ \vdots & \vdots & & \vdots \\ e_{i,M}e_{i,1} & e_{i,M}e_{i,2} & \cdots & e_{1,M}^2 \end{pmatrix} \tag{31}$$

Also, $e_i^{'}s$ are samples from a distribution with zero mean and the estimate of covariance matrix of error $e_i e_i^{'}/n$ of all of the observations.

Sample covariance matrix

$$\frac{1}{n}\left|\varepsilon^{'}\varepsilon\right|_{j,j} = \frac{1}{n}\sum_{i=1}^{n} e_{i,j}^2 = S_{j,j} = S_j^2 \tag{32}$$

$$\frac{1}{n}\left|\varepsilon^{'}\varepsilon\right|_{j,k} = \frac{1}{n}\sum_{i=1}^{n} e_{i,j}e_{i,k} = S_{j,k} \tag{33}$$

- If $n>M$ and e_i, $i=1,...,\ v_0$ ~i.i.d $MN\ (0,S_o)$, then $\varepsilon^{'}\varepsilon = \sum_{i=0}^{n} e_i e_i^{'}$ will be positive definite and symmetric, with a given positive integer v_o an $M\times M$ random covariance matrix H_o. The population distribution of sum of the squares matrices for $\varepsilon_1^{'}\varepsilon_1,...,\varepsilon_S^{'}\varepsilon_S$ is known as Wishart distribution with parameters (v_o, S_o).
- If $v_o>M$ and e_i, $i=1,...,\ v_0$ ~i.i.d $MN(0,S_o)$, then $\varepsilon^{'}\varepsilon = \sum_{i=0}^{v_o} e_i e_i^{'}$ and setting $\sum = \left(\varepsilon^{'}\varepsilon\right)^{-1}$. The precision matrix $\sum^{-1}$ has Wishart (v_0, S_0^{-1}) distribution.
- The covariance matrix $\sum$ has Inverse-Wishart (v_o, S_o^{-1}) distribution

Hence,

$$p(\sum|\ y^*,\theta) \propto p(y^*/\ \theta,\sum)\ \ p(\sum) \propto det(\sum)^{-\frac{n}{2}} exp\left\{-\frac{1}{2}\left(y-X\hat{\theta}\right)^{'}\sum^{-1}\left(y-X\hat{\theta}\right)\right\} \times det(\sum)^{-\frac{(v_0+M+1)}{2}}$$

$$\propto det(\sum)^{-\frac{n}{2}} exp\left[-\frac{1}{2}tr(T_o\sum^{-1})\right] \times det(\sum)^{-\frac{(v_o+M+1)}{2}} exp\left[-\frac{1}{2}tr(S_o\sum^{-1})\right]$$

$$\propto det(\sum)^{-\frac{(v_0+n+M+1)}{2}} exp\left[-\frac{1}{2}tr(T_o\sum^{-1})\right] \times exp\left[-\frac{1}{2}tr(S_o\sum^{-1})\right]$$

$$\propto det(\sum)^{-\frac{(v_0+n+M+1)}{2}} exp\left[-\frac{\sum^{-1}}{2}tr(S_o+T_o)\right]$$

$$\sum|y,\ \theta \sim IW(\boldsymbol{v_o}+n,|S_o+T_o|^{-1}=T_p) \tag{34}$$

Where

$$T_0 = \left(y^*-X^*\hat{\theta}\right)^{'}\left(y^*-X^*\hat{\theta}\right) \tag{35}$$

$$\sum|y^*,\ \theta \sim IW\left(v_o+n,\left(S_o+\left(y^*-X^*\hat{\beta}\right)^{'}\left(y^*-X^*\hat{\beta}\right)\right)^{-1}\right) \tag{36}$$

$$\sum|y^*,\ \theta \sim IW\left(v_o+n,\ T_p\right) \tag{37}$$

Equation (37) is not analytically possible since $\sum$ is not known, therefore posterior simulator is required to obtain the full conditional distribution for $\theta$and $\sum$ using Monte Carlo Integration (MCI) methods.

Simulation experiments

Simulation experiments were designed to investigate asymptotic behavior of BE for SUR Model with Linear Collinearity (BSURMLC), Quadratic Collinearity (BSURMQC), Ratio Collinearity (BSURMRC) and Interaction Collinearity (BSURMIC) under Non-Informative Prior (N-IP) and Informative Prior (IP). Without loss of generality stimulation experiments are based on a two-equation system. Correlated explanatory variables guided by Yoo *et al.*(2015) are included in equation I of the SUR system of equation, which are drawn using LC, QC, RC and IC

Framework of SUR Model with Collinearity Structures

$$y_1 = \theta_{y_1} + \theta_1 X_{11} + \theta_2 X_{12} + e_1$$

$$y_2 = \theta_{y_2} + \theta_3 X_{21} + \theta_4 X_{22} + e_1$$

such that

(i) The choice of values for $\lambda_1=0.5$ and $\lambda_2=-0.2$, where $u=(u_1,\ u_2)\sim N(0,1)$ (Bayo *et al.*, 2021)

(ii) The uncorrelated and explanatory variables generated in equations, y_1 and y_2 are guided by the procedures using uniform distributions, $x_1, x_2, x_3, X_{21}, X_{22} \sim U(0,1)$ (Kmenta and Gilbert, 1968).

(iii) ε_1 and ε_1 are contemporaneously correlated errors, $\varepsilon \sim N(0,\sum)$ and $\sum=\begin{bmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{12} & \sigma_{11}\end{bmatrix}$ where $\sigma_{11}=\sigma_{22}=1$ and $\sigma_{12}=\sigma_{21}=0.5$ (Wu and Weeks, 2023)

MCMC Simulations and STATA; Software Implementation

Simulations of the Bayesian SUR analysis were performed using Metropolis Hastings with Gibbs (MHG) Sampling Algorithms within STATA Software implementation. The posterior distributions were derived, with analytically intractable Multivariate Normal Gamma (MNG) and Inverse Wishart (IW) priors for model parameters and covariance of error terms respectively.

The derived BSURM with four CSs are $\theta_{BSURMLC}|\, y^*,\sum \sim MNG(\theta_{n,LC.},\tau)$, $\theta_{BSURMQC}|\, y^*,\sum \sim MNG(\theta_{n,QC.},\tau)$, $\theta_{BSURMRC}|\, y^*,\sum \sim MNG(\theta_{n,RC.},\tau)$ and $\theta_{BSURMIC}|\, y^*,\sum \sim MNG(\theta_{n,IC.},\tau)$. Therefore, size of MCMC used is 20,000 with burnin 5,000 when $\theta_n=5$, $n=1,...,6$, $v_o=6\,\tau\sim gamma(2,1)$ for IP values and $\theta_n=0$, $n=1,...,6$, $v_o=0\,\tau\sim gamma(0.1,0.1)$ for NIP values. The results of Bayesian Estimators (BEs) for SUR modelwith Collinearity Structures (CSs) are presented in Tables 1 – 4. The tables contain estimated means and SDs of BEs of a two equation SUR model with different CSs under IP and N-IP; when true values for θ_1, θ_2, θ_3 and θ_4 are 0.2, 0.4, 0.5 and 0.6 respectively with sample sizes, n=5, 10, 20, 30,50, 100, 250 and 500. True values for θ_{y_1}, θ_{y_2} are intercepts, 2 and 1 respectively.

Results

Tables 1 shows that the estimated means of θ_1, θ_2 in equation I of a two equation SUR model under N-IP for BSURMLC are -0.0082, -0.0060; 0.2775, -0.0424; 0.3879 ,0.4781, BSURMQC are 0.0247, 0.0282; 0.2513, -0.5693, BSURMRC are 0.2088 ,0.2928; 0.1930, 0.4532; 0.2000, 0.4000 and BSURMIC are 0.2315, 0.1510; 1.2534, -0.4748; 0.4134, 0.4419 for n=5, 30 and 500 respectively. The results presented in Table 1 reveal the estimated means of θ_1, θ_2 for BSURMRC that are close to true values, 0.2 and 0.4 for n=100, 250 and 500 under N-IP. This implies that estimated means of BSURMRC are unbiased for n=100, 250 and 500 (Weak Law of Large Number (WLLN). Hence, estimated means of the derived BE for SUR model with RC are more reliable than other BE.

Tables 2 shows that the estimated SDs of θ_1, θ_2 in equation I of a two equation SUR model under N-IP for BSURMLC are 0.0632, 0.0666; 0.5075, 0.4487; 0.1827, 0.1376, BSURMQC are 0.1134, 0.1258; 0.5024,0.5081; BSURMRC are 0.0152, 0.2424; 0.0042, 0.0261; 0.4681, 0.3534 and BSURMIC are 0.4514 ,0.3078; 0.4681,0.3538; 0.1680, 0 .1326 for n= 5, 30 and 500 respectively. The results presented in Table 2 reveal that estimated SDs of θ_1 and θ_2 for BSURMLC, BSURMQC, BSURMRC and BSURMIC decrease as n increases from 30-500. The BSURMRC misbehaves for $\theta_2=$ 0. 0379 when n=50 as shown in Table 2. All four derived BE behave well for large sample size. However, the estimated SDs of the BEs for SUR model with ratio collinearity decrease faster than other BE, to effectively minimise efficiency loss for large sample size under N-IP.

Figure 1 shows the systematic pattern of estimators, θ_1 and θ_2 for SUR model with CSs for sample size, n ranges from 5-500. It is evident from this chart that PSD of θ_1 and θ_2 for BSURMRC reduce faster than other BE, except for θ_1 that shows an unclear pattern when n is very small under N-IP.

Tables 3 shows that the estimated means of θ_1, θ_2 in equation I of a two equation SUR model under IP for BSURMLC are 2.1396, 1.3948; 0.9239, -0.2920; 0.4106, 0.4813, BSURMQC are 0.3127, 0.8885; 1.1628, -1.5914; -0.2725, 0.9882, BSURMRC are 0.2134, 0.1560; 0.1938, 0.4452; 0.2000, 0.4000 and BSURMIC are 3.2875, -1.0110; -0.7534, -0.1043; 0.4332, 0.4366 for n=5, 30 and 500 respectively. The results presented in Table 3 reveal the estimated means of θ_1, θ_2 for BSURMRC that are close to true values, 0.2 and 0.4 for n=50, 100, 250 and 500 under IP. This implies that estimated means of BSURMRC are unbiased for n=50, 100, 250 and 500 (WLLN's theorem). This conforms to the obtained results in Table 1. Hence, derived Bayesian estimator for

SUR model with ratio collinearity produce estimated means that are more reliable than other BE under IP.

Table 4 shows that the estimated SDs of θ_1, θ_2 in equation I of a two equation SUR model under IP for BSURMLC are 1.773036, 1.079916; 0.9582599, 0.6825572; 0.1993809, 0.1406341, BSURMQC are 1.798501, 1.686517; 1.128651, 1.088139; 0.4093819, 0.3993413, BSURMRC are 0.01668, 0.2902992; 0.0038352, 0.0240608; 2.41, 0.0003517 and BSURMIC are 0.5379284 ,0.52321; 0.6375828, 0.473223; 0.1823539, 0.1375363 for n= 5, 30 and 500 respectively. Results in Table 4 reveal that the estimated SDs of BSURMQC decrease as n increases from 5-500. The estimated SDs of BSURMLC and BSURMRC and BSURMIC decrease as n increases from 30-500. However, estimated SDs of BSURMQC consistently reduce as against other BE under IP. Hence, BSURMQC produces estimated SDs that are more reliable than other BE to reduce efficiency loss under IP.

Figure 2 shows the systematic pattern of estimators, θ_1 and θ_2 for SUR model with CSs for n ranges from 5-500. It is evident from this chart that PSD of θ_1 and θ_2 for BSURMRC reduced faster than other BE with a little unclear pattern when sample is very small, n<30. However, BSURMQC shows a clear pattern when n ranges from 5-500.

Table 1: Posterior Means of Collinearity Structures in Equation I of a Two Equation Seemingly Unrelated Regression under Noninformative Prior Specification

True value	Estimtors	5	10	20	30	50	100	250	500
θ_1=0.2	BSURMLC	-0.0082	0.3646	0.8937	0.2775	-0.4025	0.2647	0.2373	0.3879
	BSURMQC	0.0246	0.8009	0.7851	0.2513	0.4269	0.2765	0.3002	0.1117
	BSURMRC	0.2088	0.0824	0.1972	0.1930	0.1979	0.1998	0.1999	0.2000
	BSURMIC	0.2315	0.4297	1.2534	-0.4748	-0.3668	0.0496	0.2752	0.4134
θ_2=0.4	BSURMLC	-0.0060	0.3194	-0.4741	-0.0424	0.3920	0.5999	0.5347	0.4781
	BSURMQC	0.0282	0.6311	0.5970	-0.5693	0.2415	-0.1813	0.4028	0.6179
	BSURMRC	0.2928	0.6411	0.4325	0.4532	0.4613	0.4114	0.4025	0.4000
	BSURMIC	0.1510	0.9186	0.6123	-0.1662	0.5741	-0.0488	0.4602	0.4420

Table 2: Posterior Standard Deviations of Collinearity Structures in Equation I of a Two Equation Seemingly Unrelated Regression under Noninformative Prior Specification

True value	Estimator	5	10	20	30	50	100	250	500
θ_1	BSURMLC	0.0632	0.5284	0.4915	0.5075	0.4106	0.3403	0.2487	0.1827
	BSURMQC	0.1134	0.4563	0.4629	0.5024	0.4549	0.4312	0.3814	0.3187
	BSURMRC	0.0152	0.0903	0.0039	0.0042	0.0012	0.0002	0.0000	0.0000
	BSURMIC	0.4514	0.4403	0.5282	0.4681	0.4494	0.3612	0.2424	0.1680
θ_2	BSURMLC	0.0666	0.4996	0.4393	0.4487	0.3438	0.3065	0.1808	0.1376
	BSURMQC	0.1258	0.4565	0.4575	0.5081	0.4423	0.4228	0.3685	0.3126
	BSURMRC	0.2424	0.2080	0.0563	0.0261	0.0379	0.0095	0.0022	0.0003
	BSURMIC	0.3078	0.4077	0.3800	0.3538	0.3038	0.2460	0.1824	0 .1326

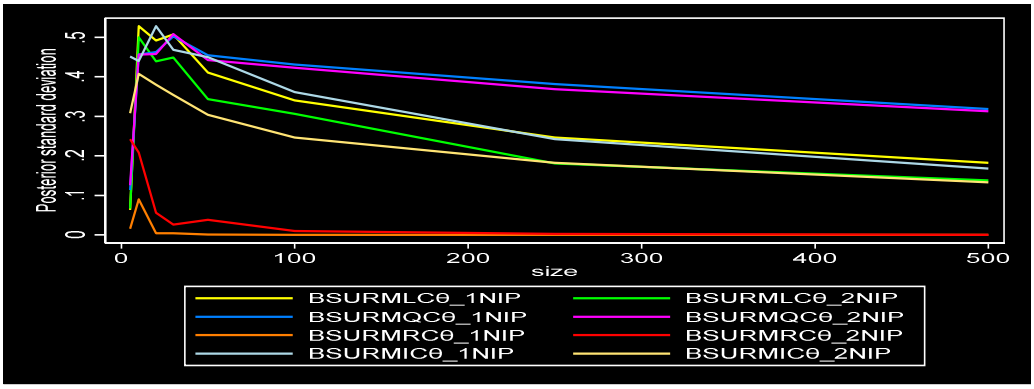

Figure 1: Posterior standard deviation of θ_1 and θ_2 of SUR model with different CSs under N-IP.

Table 3: Posterior Means of Collinearity Structure in Equation I of a Two Equation Seemingly Unrelated Regression under Informative Prior Specification

True value	Estimator	5	10	20	30	50	100	250	500
θ_1=0.2	BSURMLC	-2.1396	2.2002	2.0691	0.9239	-0.7496	0.2947	0.2568	0.4106
	BSURMQC	0.3127	0.9995	1.1954	1.1628	-0.0499	0.3394	-0.0413	-0.2725
	BSURMRC	0.2136	0.1451	0.1994	0.1938	0.1985	0.1999	0.2000	0.2000
	BSURMIC	3.2875	0.4837	2.6963	-0.7534	-0.7970	0.0955	0.2919	0.4332
θ_2=0.4	BSURMLC	1.3948	-1.3435	-1.3324	-0.2920	0.6603	0.7372	0.5750	0.4813
	BSURMQC	0.8885	0.4481	0.2826	-1.5914	0.6195	-0.3103	0.7333	0.9882
	BSURMRC	0.1560	0.4668	0.3847	0.4452	0.4393	0.4092	0.4023	0.4000
	BSURMIC	-1.0110	1.0166	-0.0137	-0.1043	0.7281	-0.1347	0.4676	0.4366

Table 4: Posterior Standard Deviations of Collinearity Structures in Equation I of a Two Equation Seemingly Unrelated Regression under Informative Prior Specifications

True value	Estimator	5	10	20	30	50	100	250	500
θ_1	BSURMLC	1.7730	1.4330	0.6324	0.9583	0.5030	0.4102	0.2722	0.1994
	BSURMQC	1.7985	1.5728	1.3712	1.1287	1.0812	0.8200	0.5848	0.4094
	BSURMRC	0.0167	0.0775	0.0036	0.0038	0.0012	0.0002	0.0000	0.0000
	BSURMIC	0.5379	0.6630	0.7061	0.6376	0.5984	0.4331	0.2707	0.1824
θ_2	BSURMLC	1.0800	1.2968	0.5207	0.6830	0.4102	0.3570	0.1984	0.1406
	BSURMQC	1.6865	1.5174	1.3090	1.0881	1.0343	0.8016	0.5709	0.3993
	BSURMRC	0.2903	0.1795	0.0505	0.0241	0.0367	0.0094	0.0022	0.0004
	BSURMIC	0.52321	0.5988	0.4851	0.4732	0.3624	0.2891	0.1971	0.1375

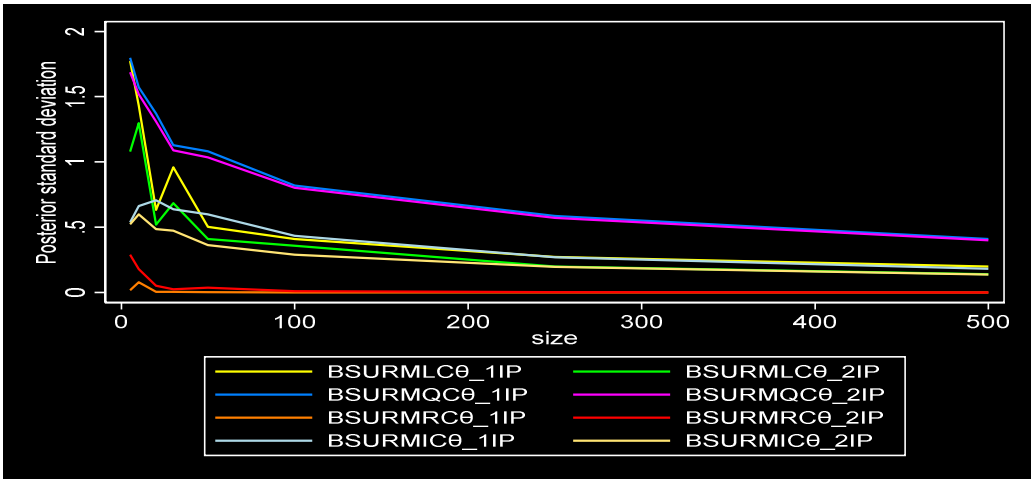

Figure 2: Posterior standard deviation of θ_1 and θ_2 of SUR model with different CSs under IP

Discussion

This study uses an STATA version 14 in which codes were written to implement Metropolis-Hasting algorithm and Gibbs sampling for the derived BE for SUR model with different Collinearity Structures (CSs). The posterior distributions were used to derive BSURMLC, BSURMQC, BSURMRC and BSURMIC, while simulated data validated the derived estimators. The estimated means in equation 1 for SURM with LC, QC and IC are sensitive to N-IP and IP, while that of RC are less sensitive to N-IP and IP. These findings reveal that BSURMRC produce the unbiased estimates compared to other BE when sample size is large under N-IP and IP as shown in Table 1 and Table 3 (Mehta and Swamy, 1975). Also, the estimated SDs reveal the asymptotic behaviour of BE for SUR model with different CSs for N-IP and IP in Table 2 and Table 4.Therefore, estimated SDs of BSURMRC, BSURMQC for N-IP and IP respectively decrease faster than other BE for large sample size (Alaba *et al.,* 2019). In addition, the BSURMQC outperforms other BEs for SUR model with CSs under IP as shown in Table 4. This implies that BSURMQC behaves well for all sample sizes considered (Gary, 2003).

Conclusion

The derived Bayesian estimators for a two-equation seemingly unrelated regression model with four different structures of collinearity validated statistical inference. Thus, these BEs provide alternative approaches to the existing classical estimators which are ineffective due to violation of uncorrelated assumption of SUR model parameters. BEs under IP outperform N-IP when there is SUR assumption violation in one of the two SUR equations. BEs are computational effective to minimise efficiency loss associated with collinearity structures in the SUR model when appropriate prior distributions are specified. This is an appealing property when n increases and explanatory variables are contaminated with collinearity. The derived Bayesian estimators improve the efficiency of the SUR model when priors about the model parameters are incorporated in the SUR model.

Conflict of Interest

Authors declare no conflict of interest.

References

Adebayo, S. B., (2003). Semiparametric Bayesian regression for multivariate response. *Hieronymus, Munich*.

Afolayan, R. B., Banjoko, A. W., Garba, M. K., and Yahya, W. B.. (2020).On Seemingly Unrelated Regression and Single Equation

Estimators Under Heteroscedastic Error and Non-Gaussian Responses. *FUOYEJET, Vol. 5(2), pp 34-42.*

Alaba, O.O., Adepoju, A.A., and Olaomi, O.E. (2019). Seemingly unrelated regression with decomposed variance covariance matrix: a Bayesian approach. *Journal of the Nigerian Association of Mathematical Physics.* Vol. 51, 137 – 144.

Bayo, A. K,, Rafiu, A. B., Funmilayo, A. T., and Oluyemi, O. I. (2021). Investigating the impact of multicollinearity on linear regression estimates. Malaysian Journal of Computing. Vol.6 (1): 698-714.

Baltagi, B.H. (1988).The efficiency of OLS in seemingly unrelated regression model. Econometric Theory, Vol.4(3)

Binkley, J. K. (1982).The effect of variable correlation on the efficiency of seemingly unrelated regression in a two-equation model. *Journal of the American Statistical Association*. Vol. 77, No. 380, pp. 890-895.

Esser, J., Maia, M., Parnell, A. C., Bosmans, J., Dongen, H. V. Klausch, T., and Murphy, K. (2024). Seemingly unrelated Bayesian additive regression trees for cost-effectiveness analyses in healthcare.*arXiv: pp 1-38*.

Fiebig, D.G. (2001). Seemingly unrelated regression. In: Baltagi B.H, Editor. *A companion to theoretical econometrics,* Massachusett: Blackwell Blackwell Publishing 101-121.

Gary, K. (2003). Bayesian Econometrics. John Wiley & Son Ltd, The Atrium, Southerner Gate, Chuchester West Sussex Po19 8SQ, England. ISBN 0-470-84567-8.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical model. Bayesian Analysis, 1(3), 515-533.

Georges B., Anoop, C., Mohammad, A. R., and Shalabh, S. (2021). Seemingly unrelated regression with measurement error: Estimation via Markov chain Monte Carlo and Mean Field Variational Bayes Approximation. The International Journal of Biostatistics. Vol.17 (1):75-97.

Gush, J.K., and Mukerjee, R. (1992). Non informative prior (With discussion). In J,M. Bernardo, J.O Berger A.P David & A.F.M. Smith (Eds), *Bayesian Statistics* 4:195-210.

Kmenta, J., and Gilbert, R. F. (1968). Small Sample Properties of Alternative Estimators of Seemingly Unrelated Regressions. *Journal of the American Statistical Association, Vol. 63, No. 324, pp. 1180-1200.*

Kristina, P. V., MinJae, L., Joseph, B. M., and Mohammad, H.R. (2016). Multicollinearity in regression analyses conducted in epidemiologic studies. *Epidemiology.* Vol. 6(2):1-20.

Lambert, P.C., Sutton, A.J., Burton, P.R., Abrams, K.R., and Jones, D.R. (2005). How vague is vague. A simulation study of the impact of the use of vague prior distribution in MCMC use WinBUGS. *Statistics in Medicine.* Vol. 24, 2401-2428.

Neter, J., Kutner, M. H., Nachtsheim, C. J., and Wasserman W. (1996) Applied Linear of Multivariate Analysis. Vol. 1, pp 1-59.

Ojo, O. O., and Adepoju, A. A. (2021). Bayesian analysis of correlated regressors in seemingly unrelated regression model. *Rattanakosin Journal of Science and Technology,* Vol. 3 (1): 19-30.

Ojo, O. O., Owonipa, O. R. and Enesi, L. O. (2021). A simulation study of Bayesian estimator for seemingly unrelated regression under different distributional assumptions. *Asian Journal of Probability and Statistics.* Vol. 10 (4): 1-8.

Omid, A., and Mousa, G. (2021). On seemingly unrelated regression model with skew error. *Journal of Statistical Theory and Applications.* Vol. 20(1) : 97-110.

Onatunji, A. P., Adepoju, A. A., Adesina, O. A and Oni, O. V. (2020). Bayesian estimation of seemingly unrelated regression with collinear categorical explanatory variables. International Journal of Mathematics and Statistics Studies Vol.8(2) :1-8.

Raifa, H., and Schlaifar, R. (1961). Applied statistical decision theory. *Division of Research, Graduate School of Business Administration, Hrvard University*. Chapter 3, 43-76

Salam, A., and Grzegorczyk, M. (2022). Model averaging for sparse seemingly unrelated regression using Bayesian networks among the errors. *Computational Statistics.* Vol. 38 (7):1-30.

Silvey, S. D. (1969). Multi-collinearity and imprecise estimation. *Journal of Royal Statistical Society, 31: 539 – 552*.

Soffritti, G., and Perrone, G. (2023). Seemingly unrelated clusterwise linear regression for contaminated data. *Statistical Papers.* 64:883–921.

Srivastava, V.K., and Giles, D.E.A. (1987). Seemingly unrelated regression equations models: estimation and inference. *New York: Marcel Dekker*. Chapter 1:1-10.Taylor and Francis.

Tueller, S., and Lubke, G. (2010). Evaluation of structural equation mixture models: parameters estimates and correct class assignment. Structural equation modeling: *A multidisciplinary Journal*. Vol. 17(2):165-192

Wang, H. (2010). Sparse seemingly unrelated regression modelling: Applications in finance and econometrics. *Computational Statistics and Data Analysis 54, 2866–2877.*

Xie, L., Fu L., Widagdo, F. R., Dong, L., and Li, F. (2022). Improving the accuracy of tree biomass estimations for three coniferous tree species in Northeast China. *Trees, 36:451–469.*

Yahya, W.B., and Olaifa, J.B. 2014. A note on ridge regression modeling techniques. *Electronic Journal of Applied Statistical Analysis, 7* (2):*343-361*.

Yoo, W., Mayberry, R., BaeSejong, H.Q., and Lillard J. W. (2015). A study of effects of multicollinearity in the multivariable analysis. *Int J ApplSci Technol*. Vol.4 (2*):1-10*

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests of aggregation bias. *Journal of American Statistical Association, 57, 500-509*.