International Journal of Mathematics and Statistics Studies
Vol.8, No.2, pp.1-8, July 2020
Published by ECRTD-UK
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

BAYESIAN ESTIMATION OF SEEMINGLY UNRELATED REGRESSION WITH COLLINEAR CATEGORICAL EXPLANATORY VARIABLES

Onatunji Adewale P.
LAUTECH Int'l College, Ogomoso, Oyo State, Nigeria

Adepoju Adadayo A.
Department of Statistics, University of Ibadan, Ibadan, Nigerian

Adesina Oluwaseun A.
Department of Mathematics and Statistics,The Polytechnic, Ibadan, Nigeria

Oni Olanreaju V.
Department of Statistics, Federal College of Animal Health and Production Technology, Ibadan, Nigeria

ABSTRACT: *The efficiency of Seemingly Unrelated Regression (SUR) not only depends on contemporaneous correlation between errors term but also on the degree of collinearity among explanatory variables in the equations of the model. The problem of collinearity consequentially leads to biased estimation, rank deficient, large standard deviations and misleading interpretation of the estimates among others in analysis. This study examines the robustness of the Bayesian estimator to varying degree of correlation among categorical explanatory variables of seemingly unrelated regression model. The result revealed an asymptotic property of the Bayesian method and the best estimates were obtained when sample size, N is large irrespective of the degree of correlation among the regressors.*

KEYWORDS: collinearity, seemingly unrelated regression, posterior standard deviations, Bayesian estimator

INTRODUCTION

Seemingly unrelated regression (SUR) is a system of equations in which there is at least two response variables jointly regressed with predictive variables on the basis of contemporaneous correlation of errors term between them. The efficiency of SUR is not independent of collinearity or non collinearity of predictive variables. Collinearity in the analysis of SUR model practically leads to biased coefficients, large standard errors of regression coefficients and loss of power. Jian and Yong (2019) presented closed-form assumption of the best linear unbiased parameter(BLUP) and best linear unbiased estimator(BLUE) of all unknown parameters in the SUR models; necessary and sufficient conditions for a family of equalities of predictors under SUR. Funda and Fikri(2016) developed a restricted feasible seemingly unrelated estimator(RFSURE) in the estimation problem with linear restrictions and investigated using MCMC. In this study, it was revealed that RFSURE produced smaller mean square estimators(MSEs) as against feasible generalized least squares(FGLS). The simulation study revealed that SUR estimator is efficient at tolerable non–orthogonal correlation points(TNCP) among the explanatory variables which follow a Gaussian distribution(Yahya et al 2008). Olanrewaju et al (2017) reported that SUR model was preferable in the presence of multicollinearity and autocorrelation from Monte Carlo experiment of 1000 trial for all the sample sizes considered. Banterle et al (2018) used a Bayesian variable selection(BVS) in SUR to allow residuals to conditional dependence between response variables and reported a matrix of binary variable selection indicators. However in addition, expensive iterative methods of simulation failed to scale well with the problem of dimension with exact computation. Ando (2012) developed Bayesian method to solve problem of variable selection and model parameter with a large number of predictors in SUR using Direct Monte Carlo(DMC). These are the crucial problems in Bayesian modeling of SUR. DMC performed better than MCMC usually used for Bayesian model estimation. Okewole(2014) investigated the performance of the Bayesian approach in estimating multi-equation models in the presence of multicollinearity and found that this method performed well in the small sample size $N \leq 40$. Adepoju and Ojo(2019) showed that Mante Carlo Intergration(MCI) outperformed the Gibbs sampler(GS) for the different levels of correlation used in the prior covariance and the accuracy of MCI does not depend on the levels of correlation either positive or negative. This study examines the robustness of the Bayesian estimation method to correlation among the explanatory in a SUR model. The rest of the paper is structured as follows: section 2 considers the model specification and brief discussion of the prior, likelihood and posterior distributions, section 3 discusses the derivation of the Bayesian estimator using the MCMC, section 4 presents the analysis, results and discussion and lastly section 5 concludes the paper.

MATERIAL AND METHOD

The Model

The aim of this paper is to carry out a Bayesian method of SUR when the categorical explanatory variables are correlated in at least one of the equations in the model. This study examines the sensitivity of the Bayesian method to varying levels of collinearity among the explanatory variables in equation(1). The asymptotic and consistent properties of the Bayesian method will be investigated via the posterior mean and the posterior standard deviation respectively.
Given the following two–equation seemingly unrelated regression with the regressors in equation(1) assumed to be correlated;

$$\begin{aligned} y_1 &= \beta_{11}x_{11cor} + \beta_{12}x_{12cor} + u_{11} \\ y_2 &= \beta_{21}x_{21} + \beta_{22}x_{12} + u_{21} \end{aligned} \quad (1)$$

where $y = (y_1, y_2)$ is a vector of response variables with β is $p \times 1$ vector of regression coefficients, X is $m \times n$ matrix of categorical explanatory variables and $u = (u_1, u_2)$ is a vector of random errors that assume multivariate normal with covariance matrix $\sum$. Thus, equation 1 is here compactly written as a single equation defined as

International Journal of Mathematics and Statistics Studies
Vol.8, No.2, pp.1-8, July 2020
Published by ECRTD-UK
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

$$y_j = X_j\beta + u_j \qquad (2)$$

where

$$y_j = \begin{pmatrix} y_{1j} \\ y_{2j} \end{pmatrix},\ \mathrm{X_j} = \begin{pmatrix} \mathrm{X}'_{1j} & 0 \\ 0 & \mathrm{X}'_{2j} \end{pmatrix}, \beta = \begin{pmatrix} \beta_{1j} \\ \beta_{2j} \end{pmatrix} and\ \mathrm{u} = \begin{pmatrix} u_{1j} \\ u_{2j} \end{pmatrix}$$

The likelihood, Prior and Posterior distribution

In this study, the posterior distributions of β and $\sum$ are obtained by the likelihood functions conditional on reliable prior (diffused prior and Wishart prior). The likelihood function is designated as $p(y \mid \beta, \Sigma) \sim N(X\beta, \Sigma)$,and its density is given as

$$\begin{aligned} L(y \mid \beta, \Sigma) &= \prod_{i=1}^{N} \frac{|\Sigma|^{1/2}}{(2\pi)^{M/2}} \exp\left[-\frac{1}{2}\left(\mathbf{y} - X_j\beta_j\right)'(\Sigma^{-1} \otimes I)\left(\mathbf{y} - X_j\beta_j\right)\right] \\ &= \frac{|\Sigma|^{1/2}}{(2\pi)^{M/2}} \exp\left[-\frac{1}{2}\sum_{i=1}^{N}\left(\mathbf{y} - X_j\beta_j\right)'(\Sigma^{-1} \otimes I)\left(\mathbf{y} - X_j\beta_j\right)\right] \end{aligned} \qquad (3)$$

The insufficient knowledge of any prior for related categorical covariates gives rise to the use of diffused prior for β while $\sum$ follows Wishart prior as a generalization of gamma distribution for SUR.Independent diffused priors are used for categorical variables. Jeffrey (1961) suggested diffused priors for these variables parameters.

$$p(\beta) \sim const \qquad (4)$$

Inverse-wishart prior is used for the covariance matrix of the error term($\sum$) where A is a scalar and B symmetric and positive definite matrix , $\Sigma \sim IW(A,B)$ for $|\Sigma| > 0$ and with a standard choice of $A = 1$ and $B = diag(d_1,...,d_m)$ as small as d=0.005. The probability distribution function of $\sum$ is then given as

$$p(\Sigma) \propto |\Sigma|^{-A-(K+1)/2} \exp\left(-tr(B\Sigma^{-1})\right) \qquad (5)$$

The posterior distribution

The Bayesian inference is based on the posterior distribution of the model parameter. This is done by combining equation 3, 4 and 5 to give 6, the posterior distribution defined below.

$$p(\beta, \Sigma \mid y) \propto p(y \mid \beta, \Sigma)\, p(\beta)\, p(\Sigma) \qquad (6)$$

International Journal of Mathematics and Statistics Studies
Vol.8, No.2, pp.1-8, July 2020
Published by ECRTD-UK
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

$$p\left(\beta,\Sigma \mid y\right) \propto \left(\frac{|\Sigma|^{1-A-(K+1)/2}}{(2\pi)^{M/2}} \exp-\left[\frac{1}{2}\sum_{i=1}^{N}\left(\mathbf{y}_j - X_j\beta\right)'(\Sigma^{-1}\otimes I)\left(\mathbf{y}_j - X_j\beta\right)\right]\left(tr(B\Sigma^{-1})\right)\right)p\left(\beta\right)$$

The posterior distributions of model parameters β and Σ to be estimated are numerically intractable. Therefore, this problem is addressed by MCMC simulation methods whereby samples are drawn from the sample from the full conditional distribution of the parameter given other information.

The marginal posterior distribution for β

Sample β from, $p\left(\beta|\sigma^2, y\right) \sim N\left\{\left(X^T X\right)^{-1} X^T y, \sigma^2\left(X^T X\right)^{-1}\right\}$ (7)

The marginal distribution of covariance matrix for Σ

Sample Σ from, $p\left(\Sigma|\beta,\sigma^2,\Sigma y\right) \sim IW\left(A+\frac{1}{2}K, B+\frac{1}{2}\|u\|^2\right)$ (8)

Bayesian MCMC

Datasets of sample sizes of 20, 50, 100 , 500 and 1000 with two response variables(y_1 and y_2) and four categorical predictive variables $X_1=\left(\mathrm{x}_{11\mathrm{cor}}\ \mathrm{x}_{12\mathrm{cor}}\right)$ in the first equation and $X_2=\left(\mathrm{x}_{21}\ \mathrm{x}_{22}\right)$ in the second equation were generated from binomial distribution with $\beta_{11}=0.2$, $\beta_{12}=0.7$, $\beta_{21}=0.4$ and $\beta_{22}=0.6$, and random error that follows a normal distribution with mean zero and variance one. The two response variables, y_1 and y_2 were obtained with associated distinct vector of categorical explanatory variables X_1 and X_2 respectively. Let $X^{'}=\left(X_1\ X_2\right)$ denote a vector of explanatory variables, $\beta^{'}=(\beta_1\ \beta_2)$, the vector of regression coefficients and Σ the variance covariance matrix of the error $\left(\varepsilon\right)$. The correlated categorical explanatory variables in the first equation are generated by modifying Gary and Diane method (Gary and Diane, 1975) in the equation below, where x_{11cor} and $\mathrm{x}_{12\mathrm{cor}}$ the two covariates in equation(1) are derived.

$$x_{1jcor} = \mathrm{z}_1\left(1-\rho\right)^{1/2} + \mathrm{z}_2\sqrt{\rho} \quad \text{j=1,2} \tag{9}$$

where z_1 and z_2 are binomially distributed for the specified sample sizes with mean and variance 2 and 0.5 respectively. The $\rho(0.0, 0.1, 0.2, ...0.9, 1.0)$ represents different degrees of collinearity between two categorical explanatory variables of the first equation. To solve this problem MCMC

Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

methods are employed to randomly draw samples from the marginal posterior distribution for the parameters sated above with the aids of BayesX software.

ANALYSIS, RESULTS AND DISCUSSION

Table1.Posterior Means Of Bayesian Seemingly Unrelated Regression Model Under Varying Degree Of Collinearity With Different Sample Sizes Between CEVs x_{11cor} and x_{12cor}

Collinearity		Posterior mean of $\beta_{11} = 0.2$					Posterior mean of $\beta_{12} = 0.7$				
	levels	20	50	100	500	1000	20	50	100	500	1000
$\rho_{x_{11}x_{12}}$	0.0	0.199711	0.199926	0.199832	0.199981	0.200005	0.700223	0.699993	0.700156	0.699938	0.700076
	0.1	0.19974	0.199887	0.199887	0.199979	0.200012	0.700317	0.699971	0.699971	0.699936	0.700046
	0.2	0.199767	0.199881	0.199899	0.19998	0.200013	0.700343	0.699963	0.700093	0.699938	0.700033
	0.3	0.199789	0.199878	0.199917	0.19998	0.200014	0.700361	0.699956	0.700078	0.69994	0.700024
	0.4	0.199807	0.199876	0.199933	0.199981	0.200015	0.700376	0.69995	0.700065	0.699942	0.700016
	0.5	0.19982	0.199874	0.199948	0.199981	0.200016	0.70039	0.699944	0.700053	0.699944	0.700008
	0.6	0.199828	0.19987	0.199965	0.199981	0.200016	0.700402	0.699937	0.70004	0.699946	0.7
	0.7	0.199828	0.199866	0.199984	0.199982	0.200017	0.700412	0.699929	0.700024	0.699948	0.699991
	0.8	0.199815	0.199859	0.199859	0.199982	0.200018	0.70042	0.699917	0.699917	0.699952	0.69998
	0.9	0.199781	0.19985	0.200047	0.199983	0.200019	0.70042	0.699898	0.699974	0.699957	0.699963
	1.0	0.199674	0.199824	0.200138	0.199986	0.200019	0.700394	0.699849	0.69989	0.699978	0.69992

Table 1 shows the posterior means of Bayesian seemingly unrelated regression(BSUR) estimator with varying degree of collinearity(ρ) between CEVs X_{11} and X_{12} in the first equation of SURM of equation(2) for sample sizes 20 to 1000 when the true values are set at $\beta_{11} = 0.2$ and $\beta_{12} = 0.7$. For example, the posterior means when $\rho_{x_{11}x_{12}} = 0.1$ for sample sizes of 20,50, 100, 500 and 1000 are 0.19974(0.700317), 0.199887(0.699971)0.199887(0.699971), and 0.199979(0.700046) respectively, which are close to the true values β_{11} (β_{12}). The posterior means when $\rho_{x_{11}x_{12}} = 0.8$ for all sample sizes considered are 0.199815(0.699917), 0.199859(0.699917), 0.199859(0.699917), 0.199982(0.699952). The results show the asymptotic behaviour of the Bayesian procedure in that, for all the degrees of correlation, the posterior means move toward the true parameter values as the sample size increases from20 to 1000.

International Journal of Mathematics and Statistics Studies
Vol.8, No.2, pp.1-8, July 2020
Published by ECRTD-UK
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

Table2.Posterior Standard Deviation Of Bayesian Seemingly Unrelated Regression Model Under Varying Degree Of Collinearity With Different Sample Sizes Between CEVs x_{11cor} and $\mathrm{x}_{12\mathrm{cor}}$

Collinearity	Posterior Standard Deviation of β_{11}						Posterior Standard Deviation of β_{12}				
	Levels	20	50	100	500	1000	20	50	100	500	1000
	0.0	0.002251	0.0007180	0.0003794	0.0000917	0.0000560	0.001779	0.0006929	0.0003688	0.0000916	0.0000560
	0.1	0.002262	0.0006436	0.0006436	0.0000884	0.0000540	0.001739	0.0007062	0.0007062	0.0000903	0.0000537
$\rho_{x_{11}x_{12}}$	0.2	0.002316	0.0006221	0.0004013	0.0000867	0.0000529	0.001754	0.0007144	0.0003710	0.0000891	0.0000526
	0.3	0.002369	0.0006090	0.0004048	0.0000855	0.0000521	0.00177	0.0007214	0.0003714	0.0000884	0.0000519
	0.4	0.002419	0.0006000	0.0004077	0.0000849	0.0000517	0.001792	0.0007281	0.0003721	0.0000880	0.0000515
	0.5	0.002462	0.0005937	0.0004103	0.0000846	0.0000515	0.001807	0.0007345	0.0003730	0.0000878	0.0000514
	0.6	0.002496	0.0005894	0.0004125	0.0000847	0.0000516	0.001815	0.0007408	0.0003742	0.0000880	0.0000515
	0.7	0.002516	0.0005873	0.0004144	0.0000851	0.0000520	0.001813	0.0007467	0.0003757	0.0000884	0.0000519
	0.8	0.002514	0.0005883	0.0005883	0.0000860	0.0000527	0.001797	0.0007516	0.0007516	0.0000891	0.0000523
	0.9	0.002484	0.0005957	0.0004139	0.0000875	0.0000538	0.001765	0.0007528	0.0003794	0.0000900	0.0000537
	1.0	0.002428	0.0006489	0.0003943	0.0000907	0.0000560	0.001754	0.0007281	0.0003786	0.0000908	0.0000555

Table2 contains posterior standard deviations of Bayesian seemingly unrelated regression(BSUR) methods under varying levels of collinearity(ρ) between CEVs X_{11} and X_{12} in the first equation of equation(2). For example when $\rho_{X_{11}X_{12}} = 0.1$ and sample size ranges from 20 to 1000, the obtained posterior standard deviations of the estimators of β_{11} (β_{12}) are 0.002262(0.001739), 0.0006436(0.0007062), 0.0006436(0.0007062), 0.0000884(0.0000903) and 0.0000884(0.0000537).The posterior standard deviations of β_{11} (β_{12}) are 0.002484(0.001765), 0.0005957(0.0007528), 0.0004139(0.0003794), 0.0000875(0.0000900) and 0.0000875(0.0000537) when $\rho_{X_{11}X_{12}} = 0.9$ for all sample sizes. The results reveal that posterior standard deviations of Bayesian estimator consistently decrease as sample size increases. Also, the posterior standard deviations increase as level of collinearity increases for all sample sizes except for N=20. However, for N=20, the posterior standard deviations only when the degree of correlation $\rho_{X_{11}X_{12}}$ ranges from 0.7 to 1.0. For N=50 and 100 when $\rho_{X_{11}X_{12}} = 0.1$ and 0.8 , the values of posterior standard deviations obtained are similar. Best estimates are obtained when N is large irrespective of the level of correlation between the explanatory variables.

CONCLUSION

This study investigated the behavior of Bayesian method when applied to the seemingly unrelated regression with the regressors assumed to be correlated at different degrees of collinearity(ρ)

from 0.0 to 1.0. The results show that the posterior means are close to the true values at all levels of collinearity as the sample size increases. Moreover, the posterior standard deviations consistently decrease as sample size increases. In addition, in small size when the degree of collinearity is high between CEVs, the posterior standard deviations are considerably reduced. This implies that without loss of efficiency, BSUR estimator performs well at all levels of collinearity of CEVs for the sample sizes considered

Reference

1. Adedayo A.A. and Oluwadare O.O.(2019). A comparative analysis of posterior simulation techniques in the estimation of Bayesian regression model. Estudios De Economia Aplicada Vol.37-1, pg.139-146
2. Ando T.(2012) Bayesian variable selection for the seemingly unrelated regression models with a large number of predictors "J. Japan Statist. Soc.Vol. 41 No. 2 187–203"
3. Banterle M.,L. Bottolo S.R., Ala-Korpela M., Jarvelin M.-R. and Lewin A. (2018). Sparse variable and covariance selection for high-dimensional seemingly unrelated Bayesian regression. https://doi.org/10.1101/467019.
4. Funda E.and Fikri A.(2016).Restricted Estimator in Two Seemingly Unrelated Regression Model. Pak.j.stat.oper.res. Vol.XII No.4 pp579-588
5. Jian H. and Yong Z.(2019).Some remarks on a pair of seeminglyunrelated regression models, Open Math.; 17:979–989. https://doi.org/10.1515/math-2019-0077
6. McDonald G. C. and Diane I. G.(1975). Monte Carlo Evaluation of Some Ridge. Type Estimator, Journal of the America Association, 70:350-416. http://dx.doi.org/10.1080/01621459.1975.10479882
7. Okewole D.M.O.(2014).The Bayesian Approach To Estimation Of Multi-Equation Econometric Models in The Presence of Multicollinearity, A Thesis submitted in the Department of Statistics, University of Ibadan.
8. Olanrewaju. S.O., Yahaya H.U. & Nasiru M.O(2017). Effects of multicollinearity and autocorrelation on someestimators in a system of regression equation. European Journal of Statistics and Probability Vol.5, No.3, Pp.1-15

9. Yahya W. B., Adebayo S. B., Jolayemi E. T., Oyejola B. A. and Sanni O. O. M. (2008) Effects of non-orthogonality on the efficiency of seemingly unrelated regression (SUR) models. InterStat Journal 62J05, 11D04.