Communications for Statistical Applications and Methods
2021, Vol. 28, No. 4, 315–327

https://doi.org/10.29220/CSAM.2021.28.4.315
Print ISSN 2287-7843 / Online ISSN 2383-4757

Generalized Bayes estimation for a SAR model with linear restrictions binding the coefficients

Anoop Chaturvedi[1,a], Sandeep Mishra[a]

[a]Department of Statistics, University of Allahabad, India

Abstract

The Spatial Autoregressive (SAR) models have drawn considerable attention in recent econometrics literature because of their capability to model the spatial spill overs in a feasible way. While considering the Bayesian analysis of these models, one may face the problem of lack of robustness with respect to underlying prior assumptions. The generalized Bayes estimators provide a viable alternative to incorporate prior belief and are more robust with respect to underlying prior assumptions. The present paper considers the SAR model with a set of linear restrictions binding the regression coefficients and derives restricted generalized Bayes estimator for the coefficients vector. The minimaxity of the restricted generalized Bayes estimator has been established. Using a simulation study, it has been demonstrated that the estimator dominates the restricted least squares as well as restricted Stein rule estimators.

Keywords: spatial autoregressive (SAR) model, generalized Bayes estimator, restricted least squares estimator, Stein rule estimator, minimaxity

1. Introduction

The static spatial econometric models have found wide applications for modelling the spatial spill over, i.e., the dependence of level of response variable on the levels of response variable in the neighbouring regions along with a set of explanatory variables. The classical and Bayesian estimation procedures for different spatial econometric models and their various applications have been extensively discussed in Anselin (1988), Elhorst (2003, 2010, 2014), Anselin *et al.* (2008), and Lesage and Pace (2009), Baltagi (2011), Lee and Yu (2015) to cite a few. While applying Bayesian procedures for the estimation of parameters, a major drawback is the lack of robustness with respect to underlying prior assumptions. For instance, the Bayes estimator derived under normal prior has infinite Bayes risk when true prior is Cauchy distribution, see Berger (1980). Stein (1973) proposed the generalized Bayes (GB) estimator for the multivariate normal mean under a scale mixture of prior distributions and established its dominance over the James-Stein and positive part James-Stein estimators. These estimators are more robust with respect to underlying prior assumptions and satisfy minimaxity and admissibility properties, see Brown (1971) and Rubin (1977). Berger (1980) obtained confidence region for multivariate normal mean based on GB estimator. Using Brown's (1971) condition, Maruyama (1998) developed class of admissible minimax GB estimators. Kubokawa (1991, 1994) established the dominates of GB estimator over the James-Stein estimator. Considering the scale mixture of multivariate normal distribution as prior distribution, Maruyama (1999) derived the

[1] Corresponding author: Department of Statistics, University of Allahabad, Allahabad, 211002, India.
E-mail: sandeepstat24@gmail.com

Published 31 July 2021 / journal homepage: http://csam.or.kr

GB estimator for normal mean vector, established its admissibility and minimaxity and showed that the estimator dominates positive part Stein rule estimator. For estimating the coefficients vector of spatial autoregressive (SAR) model, Pal *et al.* (2016) proposed a family of shrinkage estimators and investigated its asymptotic properties. Recently, Chaturvedi and Mishra (2019) derived a GB estimator for estimating the parameters of a SAR model and investigated the admissibility and minimaxity properties of the estimator. They applied the results to demographic data on total fertility rate for selected Indian states. However, a major drawback of their work is that they assumed the disturbances variance to be known.

For estimating the coefficients vector of a linear regression model in the presence of linear restrictions binding the coefficients, Srivastava and Srivastava (1983, 1984) and Srivastava and Chandra (1991) considered families of improved restricted estimators obtained by mixing SR with restricted least squares. Chaturvedi *et al.* (1996) extended the work of Srivastava and Srivastava (1984) to linear model with non-spherical disturbances. Toutenburg and Shalabh (1996) analysed the performance properties of predictors arising from the methods of restricted regression and mixed regression besides least squares under a target function. Toutenburg and Shalabh (2000) considered the family of Stein rule (SR) estimators proposed by Srivastava and Srivastava (1983) and analysed performance properties of this family when the objective is to predict values outside the sample and within the sample.

Notice that different families of Stein rule restricted regression estimators discussed above are simply formulated by substituting Stein rule for the OLS and are not Stein rule in the true sense. Chaturvedi *et al.* (2001) considered an alternative approach, which utilizes Rao (1973) and used the generalized inverse as a principal tool for estimating coefficients. They proposed a family of shrinkage estimators for the general linear regression model with non-spherical disturbances in the presence of a set of linear restrictions binding the regression coefficients and investigated its asymptotic and finite sample properties.

In Cobb-Douglas production function and its different variants such as trans-log production function, the assumption of constant returns to scale lead to a linear restriction on regression parameters associated with logarithm of labour input and logarithm of capital input. It has been observed by several researchers that the neighbouring production activities show interdependence because of several externalities, and the usual assumption of spatial independence in frontier production functions become inappropriate, see Glass *et al.* (2016), Tsukamoto (2019) and the references cited their in. For modelling such kind of spatial spill over along with the phenomena of constant return to scale, an SAR model with appropriate set of linear restrictions binding the coefficients provides appropriate alternative.

The present paper considers SAR model with exact linear restrictions binding the regression parameters and derives a class of restricted GB estimator for the coefficients vector. Instead of applying GB procedure directly to the restricted regression model, which may not yield an estimator satisfying the restrictions, we followed an alternative approach based on the transformation used by Chaturvedi *et al.* (2001) to derive the GB estimator. The resulting estimator retains the flavour of the GB concept, yet gives rise to an estimator that satisfies the linear constraints. The minimaxity of restricted GB estimator has been established and dominance over the usual restricted least squares estimator and restricted Stein rule estimator have been demonstrated under a quadratic loss structure. For investigating the finite sample behaviour of the class of estimators, a simulation study has been carried out. The findings of the simulation show that the proposed class of restricted GB estimators performs superior to the usual restricted least squares estimator over a wide range of parameters.

2. Generalized Bayes estimator under linear restrictions

2.1. The spatial autoregressive model

Let us consider the spatial autoregressive (SAR) model

$$y = \rho W y + X\beta + u, \quad u \sim N\left(0, \sigma^2 I_n\right), \tag{2.1}$$

where y is an $(n \times 1)$ vector of the sample observations on a dependent variable collected at each of n locations, X is a $(n \times p)$ matrix of observations on p exogenous variables with $\text{rank}(X) = p\ (< n)$, β is a $(p \times 1)$ vector of regression parameters, ρ is the spatial autoregressive parameter, W is known $n \times n$ spatial weight matrix which has been standardized to have row sum of unity. Suppose the prior information available in the form of m exact linear restrictions binding the regression

$$r = R\beta, \tag{2.2}$$

where $r : m \times 1$ is a vector with known elements, and $R : m \times p$ is a matrix of rank $m\ (< p)$ with known elements. For obtaining the restricted regression estimator for β under the restrictions (2.2) let us transform the model (2.1) as

$$\tilde{y} = \rho W y + \tilde{X}\mu + u, \tag{2.3}$$

where

$$\tilde{y} = yN, \quad \tilde{X} = XN,\ N = I_p - R'(RR')^{-1}R, \quad \mu = N\beta, \quad \tilde{\mu} = R'\left(RR'\right)^{-1} r.$$

We write

$$v(\rho) = [\tilde{y} - \rho W y]' M \left[\tilde{y} - \rho W y\right], \quad M = I_n - \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{+}\tilde{X}'.$$

When ρ is unknown, we replace it by its estimator

$$\hat{\rho} = \frac{y'W'M\tilde{y}}{y'W'MWy} \tag{2.4}$$

in (2.2) to obtain feasible restricted least squares estimator of μ.

2.2. Class of generalized Bayes estimators

For obtaining the GB estimator of regression coefficients vector under linear restrictions, we write the model (2.3) as

$$y^*(\rho) = \tilde{X}\mu + u, \tag{2.5}$$

where $y^*(\rho) = \tilde{y} - \rho W y$. Then the pdf of $y^*(\rho)$ is given by

$$p\left(y^*(\rho)|\mu, \sigma^2\right) = \frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n} \exp\left\{-\frac{1}{2\sigma^2}\left(y^*(\rho) - \tilde{X}\mu\right)'\left(y^*(\rho) - \tilde{X}\mu\right)\right\}. \tag{2.6}$$

The prior distribution for β is taken as a g-prior $N\left(\tilde{\mu}, \sigma^2 g X'X\right)$, with $g = (1-\lambda)/\lambda,\ (0 < \lambda < 1)$. Since the prior mean of μ is 0 and prior covariance matrix of μ is

$$E\left(\mu\mu'\right) = \frac{1-\lambda}{\lambda}\sigma^2\tilde{X}'\tilde{X}.$$

We take the prior distribution of μ as normal with mean vector 0 and covariance matrix $\{(1-\lambda)/\lambda\}\sigma^2\tilde{X}'\tilde{X}$. However, rank of $\tilde{X}'\tilde{X}$ is $q = p - m$, implying that the prior distribution of μ is degenerated and concentrated on a lower dimensional Euclidean space $\mathcal{R}^q$. Hence, the prior pdf of μ is given by

$$p\left(\mu|\sigma^2,\lambda\right) \propto \sigma^{-q}\left(\frac{\lambda}{1-\lambda}\right)^{\frac{q}{2}} \exp\left\{-\frac{\lambda}{2\sigma^2(1-\lambda)}\mu'\tilde{X}'\tilde{X}\mu\right\}. \tag{2.7}$$

Further, we assume that

$$p(\lambda) \propto \lambda^{-a}(1-\lambda)^c, \tag{2.8}$$

and σ has the improper prior distribution

$$p(\sigma) \propto \frac{1}{\sigma}, \quad 0 < \sigma < \infty.$$

Then the joint pdf of $(y^*(\rho), \lambda)$ is obtained as

$$p(y^*(\rho),\lambda) \propto \frac{\lambda^{\frac{q}{2}-a}(1-\lambda)^c}{\left(v(\rho) + \lambda\hat{\mu}(\rho)'\tilde{X}'\tilde{X}\hat{\mu}(\rho)\right)^{\frac{n}{2}}}. \tag{2.9}$$

The marginal density of $y^*(\rho)$ is

$$\begin{aligned}
m^*(y^*(\rho)) &\propto \int_0^1 \frac{\lambda^{\frac{q}{2}-a}(1-\lambda)^c}{\left(v(\rho) + \lambda\hat{\mu}(\rho)'\tilde{X}'\tilde{X}\hat{\mu}(\rho)\right)^{\frac{n}{2}}}d\lambda \\
&\propto \int_0^1 \frac{\lambda^{\frac{q}{2}-a}(1-\lambda)^c}{\left(v(\rho) + \hat{\mu}(\rho)'\tilde{X}'\tilde{X}\hat{\mu}(\rho)\right)^{\frac{n}{2}}}\left[1-(1-\lambda)\frac{\hat{\mu}(\rho)'\tilde{X}'\tilde{X}\hat{\mu}(\rho)}{\left(v(\rho) + \lambda\hat{\mu}(\rho)'\tilde{X}'\tilde{X}\hat{\mu}(\rho)\right)}\right]^{-\frac{n}{2}}d\lambda \\
&\propto \sum_{j=0}^{\infty}\gamma_j\frac{\left(\hat{\mu}(\rho)'\tilde{X}'\tilde{X}\hat{\mu}(\rho)\right)^j}{\left(v(\rho) + \hat{\mu}(\rho)'\tilde{X}'\tilde{X}\hat{\mu}(\rho)\right)^{\frac{n}{2}+j}}\frac{\Gamma\left(\frac{q}{2}-a+1\right)\Gamma(\mathrm{j}+\mathrm{c}+1)}{\Gamma\left(\frac{q}{2}-a+j+c+2\right)},
\end{aligned} \tag{2.10}$$

where

$$\gamma_j = \frac{\Gamma\left(\frac{n}{2}+j\right)}{\Gamma\left(\frac{n}{2}\right)j!}. \tag{2.11}$$

We write $\delta(\rho) = v(\rho)(1 + w(\rho))$ with $w(\rho) = \hat{\mu}(\rho)'\tilde{X}'\tilde{X}\hat{\mu}(\rho)/v(\rho)$. Then the posterior expectation of λ given $y^*(\rho)$ is

$$\begin{aligned}
E(\lambda|y^*(\rho)) &= \frac{\sum_{j=0}^{\infty}\gamma_j\delta(\rho)^{-\left(j+\frac{n}{2}\right)}\int_0^1\lambda^{\frac{q}{2}-a+1}(1-\lambda)^{j+c}d\lambda}{\sum_{j=0}^{\infty}\gamma_j\delta(\rho)^{-\left(j+\frac{n}{2}\right)}\int_0^1\lambda^{\frac{q}{2}-a}(1-\lambda)^{j+c}d\lambda} \\
&= \frac{{}_2F_1\left(\frac{n}{2},c+1,\frac{q}{2}-a+c+3,\frac{1}{\delta(\rho)}\right)\Gamma\left(\frac{q}{2}-a+2\right)\Gamma\left(\frac{q}{2}-a+c+2\right)}{{}_2F_1\left(\frac{n}{2},c+1,\frac{q}{2}-a+c+2,\frac{1}{\delta(\rho)}\right)\Gamma\left(\frac{q}{2}-a+1\right)\Gamma\left(\frac{q}{2}-a+c+3\right)} \\
&= \phi_{ac}(w(\rho))\ \text{(say)}.
\end{aligned} \tag{2.12}$$

Here, the confluent hypergeometric function ${}_2F_1\,(a, b; c; z)$ is defined as

$${}_2F_1\,(a, b; c; z) = \sum_{j=0}^{\infty} \frac{(a)_j (b)_j}{(c)_j} \frac{z^j}{j!} .$$

Then the GB estimator of μ is

$$\hat{\mu}_G(\rho) = [1 - \phi_{ac}(w(\rho))]\,\hat{\mu}\,(\rho)\,. \tag{2.13}$$

Hence, the restricted GB estimator of β is given by

$$\hat{\beta}_{RG}\,(\rho) = \hat{\mu}_G(\rho) + \tilde{\mu}. \tag{2.14}$$

Let us write

$$\begin{aligned}\phi_r(w(\rho)) &= w(\rho)\frac{\int_0^\infty \int_0^1 \lambda^{\frac{q}{2}-a+1}(1-\lambda)^c \sigma^{-(n+1)} \exp\left\{-\frac{v(\rho)}{2\sigma^2}\right\} \exp\left\{-\frac{\lambda v(\rho) w(\rho)}{2\sigma^2}\right\} d\lambda d\sigma}{\int_0^\infty \int_0^1 \lambda^{\frac{q}{2}-a}(1-\lambda)^c \sigma^{-(n+1)} \exp\left\{-\frac{v(\rho)}{2\sigma^2}\right\} \exp\left\{-\frac{\lambda v(\rho) w(\rho)}{2\sigma^2}\right\} d\lambda d\sigma} \\ &= w(\rho)\phi_{a,c}(w(\rho)).\end{aligned} \tag{2.15}$$

Then we can express the GB estimator $\hat{\mu}(\rho)$ as

$$\hat{\mu}_G(\rho) = \left[1 - \frac{v(\rho)}{\hat{\mu}(\rho)'\tilde{X}'\tilde{X}\hat{\mu}(\rho)}\phi_r(w(\rho))\right]\hat{\mu}(\rho).$$

The restricted GB estimator $\hat{\beta}_{RG}(\rho)$ is given by

$$\hat{\beta}_{RG}(\rho) = \left[1 - \frac{v(\rho)}{\hat{\mu}\,(\rho)'\tilde{X}'\tilde{X}\hat{\mu}(\rho)}\phi_r(w(\rho))\right]\hat{\mu}(\rho) + \tilde{\mu}. \tag{2.16}$$

3. Minimaxity conditions

Notice that the rank of matrix $\tilde{X}$ (or $\tilde{X}'\tilde{X}$) is $q = p - m$. Thus, we can find $p \times p$ orthogonal matrix P, such that

$$P'\tilde{X}'\tilde{X}P = \begin{bmatrix} \Lambda_q & 0 \\ 0 & 0 \end{bmatrix} = \operatorname{diag}\left(\lambda_1, \lambda_2, \ldots, \lambda_q, 0, \ldots, 0\right), \tag{3.1}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_q$ are eigen values of $\tilde{X}'\tilde{X}$. Define $P = [\ P_1 \quad P_2\ ]$, where P_1 is $q \times p$ and P_2 is $m \times p$. Then

$$P'\tilde{X}'\tilde{X}P = \begin{pmatrix} P_1'\tilde{X}'\tilde{X}P_1 & P_1'\tilde{X}'\tilde{X}P_2 \\ P_2'\tilde{X}'\tilde{X}P_1 & P_2'\tilde{X}'\tilde{X}P_2 \end{pmatrix}. \tag{3.2}$$

Equations (3.1) and (3.2) together imply that $P_1'\tilde{X}'\tilde{X}P_2 = 0$, $P_2'\tilde{X}'\tilde{X}P_2 = 0$. Now

$$P'\hat{\mu}(\rho) = \begin{pmatrix} P_1'\hat{\mu}(\rho) \\ P_2'\hat{\mu}(\rho) \end{pmatrix} = \begin{pmatrix} \Lambda_q^{-1} P_1'\tilde{X}'\tilde{y}(\rho) \\ 0 \end{pmatrix} = \begin{pmatrix} \eta \\ 0 \end{pmatrix}. \tag{3.3}$$

Then $\eta \sim N(P_1'\beta, \sigma^2 \Lambda_q{}^{-1})$. We write

$$Z = \frac{1}{\sigma}\Lambda_q^{-1}\eta, \quad \theta = \frac{1}{\sigma}\Lambda_q^{-1}P_1'\beta, \quad \omega(\rho) = \frac{v(\rho)}{\sigma^2}.$$

Then $Z \sim N(\theta, I_q)$, and $\omega(\rho) \sim \chi^2(n-p)$ independently of Z. Let us consider the quadratic loss function

$$L\left(\hat{\beta},\beta\right) = \frac{1}{\sigma^2}\left(\hat{\beta}-\beta\right)' X'X\left(\hat{\beta}-\beta\right). \tag{3.4}$$

Theorem 1. *Under the loss function (3.4), the restricted GB estimator $\hat{\beta}_{RG}(\rho)$ has finite risk.*

Proof: We observe that

$$\left(\hat{\beta}_{RG}(\rho)-\beta\right)' X'X\left(\hat{\beta}_{RG}(\rho)-\beta\right) = \left(\hat{\mu}_G(\rho)-\mu\right)'\tilde{X}'\tilde{X}\left(\hat{\mu}_G(\rho)-\mu\right).$$

Hence, under the loss function (3.4), the risk of GB estimator $\hat{\beta}_{RG}(\rho)$ is given by

$$\begin{aligned} R\left[\hat{\beta}_{RG}(\rho),\beta\right] &= E\left[(Z-\theta)'(Z-\theta) - 2\omega(\rho)\frac{(Z-\theta)'Z\phi_r(w(\rho))}{Z'^{Z} + \frac{\omega(\rho)^2}{Z'Z}\phi_r^2(w(\rho))}\right] \\ &= q - 2E\left[\omega(\rho)\frac{(Z-\theta)^{'Z\phi_r(w(\rho))}}{\|Z\|^2}\right] + E\left[\frac{\omega(\rho)^2}{\|Z\|^2}\phi_r{}^2(w(\rho))\right]. \end{aligned} \tag{3.5}$$

We observe that $0 \le \phi_r(w(\rho)) \le w$, so that

$$E\left[\frac{\omega(\rho)^2}{\|Z\|^2}\phi_r{}^2(w(\rho))\right] \le E\left[\|Z\|^2\right] = q + \|\theta\|^2 < \infty.$$

Further by Schwarz's inequality

$$\begin{aligned} E\left[\omega(\rho)\frac{(Z-\theta)'Z\phi_r(w(\rho))}{\|Z\|^2}\right] &\le \left[E(Z-\theta)'(Z-\theta)\,E\left\{\omega(\rho)^2\frac{\phi_r{}^2(w(\rho))}{\|Z\|^2}\right\}\right]^{\frac{1}{2}} \\ &\le \left[qE\left[\|Z\|^2\right]\right]^{\frac{1}{2}} \\ &= \left[q\left(q+\theta'^{\theta}\right)\right]^2 < \infty, \end{aligned} \tag{3.6}$$

which proves the required result. □

Since the estimators with uniformly smallest risk under a loss function usually do not exist, one of the possible criteria for selecting an estimator is minimaxity, which minimizes the maximum risk. Since the restricted OLS estimator is minimax, any estimator uniformly dominating it is also minimax. The next theorem proves the minimaxity of GB estimator utilizing this logic.

Theorem 2. *The restricted GB estimator is minimax whenever*

$$\left(3-\frac{p-m}{2}\right) + \frac{(p-m)(p-m-2)}{n+p-m-2} \le a \le \frac{p-m}{2} + 1. \tag{3.7}$$

Proof: Under the loss function (3.4) the difference between the risks of restricted GB estimator $\hat{\beta}_{RG}(\rho)$ and the OLS estimator $b_R(\rho)$ is given by

$$\begin{aligned}
R\left(\hat{\beta}_{RG}(\rho),\beta\right) - R\left(b_R(\rho),\beta\right) &= E\left[\frac{\omega(\rho)^2}{\|Z\|^2}\phi_r^2(w(\rho)) - 2\omega(\rho)\frac{(Z-\theta)'Z\phi_r\left(w\left(\rho\right)\right)}{\|Z\|^2}\right] \\
&= E\left[\frac{\omega(\rho)^2}{\|Z\|^2}\phi_r^2(w(\rho))\right] - 2E\left[\frac{\partial}{\partial Z'}\left\{Z\frac{\omega(\rho)\phi_r(w(\rho))}{\|Z\|^2}\right\}\right] \\
&= E\left[\frac{\omega(\rho)^2}{\|Z\|^2}\phi_r^2(w(\rho)) - 2(q-2)\frac{\omega(\rho)\phi_r(w(\rho))}{\|Z\|^2} - 4\phi_r'\left(w\left(\rho\right)\right)\right]. \qquad (3.8)
\end{aligned}$$

For notational convenience we write $w(\rho) \equiv w$ and $v(\rho) \equiv v$. Let us write the density function

$$g_w(\lambda) = \frac{\frac{\left\{\lambda^{\frac{q}{2}-a}(1-\lambda)^c\right\}}{\{(1+\lambda w)\}^{\frac{n}{2}}}}{\int_0^1 \frac{\left\{\lambda^{\frac{q}{2}-a}(1-\lambda)^c\right\}}{\{(1+\lambda w)\}^{\frac{n}{2}}}d\lambda}, \quad 0 < \lambda < 1.$$

Then

$$\frac{\partial}{\partial w}\left[\frac{\phi_r(w)}{w}\right] = -\frac{n}{2}\left[E_{g_w}\left(\frac{\lambda^2}{1+\lambda w}\right) - E_{g_w}\left(\frac{\lambda}{1+\lambda w}\right)E_{g_w}(\lambda)\right].$$

Since both λ and $\lambda/(1+\lambda w)$ are monotone increasing functions of λ, we get $E_{g_w}[\lambda^2/(1+\lambda w)] \geq E_{g_w}[\lambda/(1+\lambda w)]E_{g_w}(\lambda)$. This implies that $(\partial/\partial w)[\phi_r(w)/w] \leq 0$. Again

$$\begin{aligned}
\phi_r'(w) &= \frac{\phi_r(w)}{w} + w\frac{\partial}{\partial w}\left\{\frac{\phi_r(w)}{w}\right\} \\
&= E_{g_w}(\lambda) - \frac{w\,n}{2}\left[E_{g_w}\left(\frac{\lambda^2}{1+\lambda w}\right) - E_{g_w}\left(\frac{\lambda}{1+\lambda w}\right)E_{g_w}(\lambda)\right].
\end{aligned}$$

Using integration by parts and after few algebraic manipulations, we get

$$\phi_r'(w) = c\left[E_{g_w}\left(\frac{\lambda^2}{1-\lambda}\right) - E_{g_w}\left(\frac{\lambda}{1-\lambda}\right)E_{g_w}(\lambda)\right]. \qquad (3.9)$$

Since $\lambda/(1-\lambda)$ and λ are monotone increasing functions of λ, we obtain

$$E_{g_w}\left(\frac{\lambda^2}{1-\lambda}\right) \geq E_{g_w}\left(\frac{\lambda}{1-\lambda}\right)E_{g_w}(\lambda).$$

Further $0 \leq \lambda \leq 1$, implies that $E_{f_r(\lambda)}(\lambda) - E_{f_r(\lambda)}(\lambda^2) \geq 0$. If $\varphi(\chi_r^2)$ is a function of Chi-square variate with r degrees of freedom, then $E[\chi_r^2\varphi(\chi_r^2)] = rE[\varphi(\chi_{r+2}^2)]$, and $\phi_r(w)$ and $\phi_r(w)/w$ are monotone in opposite directions. Hence, we obtain

$$\begin{aligned}
R\left(\hat{\beta}_{RG}(\rho),\beta\right) - R\left(b_R(\rho),\beta\right) &\leq E\left[\omega^2\phi_r(w) - 2(q-2)\omega\right]E\left[\frac{\phi_r(w)}{\|Z\|^2}\right] \\
&\leq (n-q)(n-q+2)E\left[\phi_r(w)\right] \\
&\quad - 2(q-2)E\left[\frac{\phi_r(w)}{\|Z\|^2}\right]. \qquad (3.10)
\end{aligned}$$

Since $\phi_r(w)$ is an increasing function of w, an upper bound for $\phi_r(w)$ can be obtained when w is large. For large w, we can approximate

$$\int_0^\infty \int_0^1 \lambda^{\frac{q}{2}-a+1}(1-\lambda)^c \sigma^{-(n+1)} \exp\left\{-\frac{v}{2\sigma^2}\right\} \exp\left\{-\frac{\lambda v w}{2\sigma^2}\right\} d\lambda d\sigma$$
$$\approx \Gamma\left(\frac{q}{2}-a+2\right)\Gamma\left(\frac{n-q}{2}+a-2\right)\frac{1}{2}\left(\frac{2}{vw}\right)^{\left(\frac{q}{2}-a+2\right)}\left(\frac{2}{v}\right)^{\left(\frac{n-q}{2}+a-2\right)},$$
$$\int_0^\infty \int_0^1 \lambda^{\frac{q}{2}-a}(1-\lambda)^c \sigma^{-(n+1)} \exp\left\{-\frac{v}{2\sigma^2}\right\} \exp\left\{-\lambda\frac{vw}{2\sigma^2}\right\} d\lambda d\sigma$$
$$\approx \Gamma\left(\frac{q}{2}-a+1\right)\Gamma\left(\frac{n-q}{2}+a-1\right)\frac{1}{2}\left(\frac{2}{vw}\right)^{\left(\frac{q}{2}-a+1\right)}\left(\frac{2}{v}\right)^{\left(\frac{n-q}{2}+a-1\right)}.$$

So that, $\phi_r(w)$ can be approximated as

$$\phi_r(w) \approx \frac{\left(\frac{q}{2}-a+1\right)}{\left(\frac{n-q}{2}+a-2\right)}.$$

Thus, a sufficient dominance condition is

$$0 \leq \frac{(q2-a+1)}{\left(\frac{n-q}{2}+a-2\right)} \leq \frac{2\,(q-2)}{n-q+2}. \tag{3.11}$$

Now $(q/2-a+1) \geq 0$ implies that $a \leq q/2+1$. Further

$$\frac{\left(\frac{q}{2}-a+1\right)}{\left(\frac{n-q}{2}+a-2\right)} \leq \frac{(q-2)}{\frac{n-q}{2}+1},$$

implies that

$$a \geq \left(3-\frac{q}{2}\right)+\frac{q\,(q-2)}{n+q-2}.$$

Hence, we obtain the required sufficient minimaxity condition for the GB estimator. For large n, the minimaxity condition (3.7) reduces to

$$3-\frac{q}{2} \leq a \leq \frac{q}{2}+1.$$

□

4. Simulation study

For studying the finite sample risk performance of restricted GB estimator, we carry out the simulation study using R Software. The observations on response variable y are generated by using the SAR model

$$y = \rho W y + X\beta + u,$$

where $u \sim N(0, \sigma^2 I_n)$.

In simulation study we compare the risks of feasible version of restricted feasible generalized Bayes estimator ($\hat{\beta}_{RG}$) with the usual restricted feasible least squares (RFLS) estimator $b_R(\hat{\rho}) = \hat{\mu}(\hat{\rho}) + \tilde{\mu}$.
We consider the following restricted feasible generalized Bayes (RFGB) estimator

$$\hat{\beta}_{RG}(\hat{\rho}) = \left[1 - \phi_{a,c}\left(\frac{\hat{\mu}(\hat{\rho})'\tilde{X}'\tilde{X}\hat{\mu}(\hat{\rho})}{\widehat{\sigma^2}}\right)\right]\hat{\mu}(\hat{\rho}) + \tilde{\mu}.$$

For comparison purpose, we also consider the restricted feasible Stein-rule (RFSR) estimator

$$\hat{\beta}_{SRG}(\hat{\rho}) = \left[1 - \frac{q-2}{n-q+2}\frac{v(\hat{\rho})}{\hat{\mu}(\hat{\rho})'\tilde{X}'\tilde{X}\hat{\mu}(\hat{\rho})}\right]\hat{\mu}(\hat{\rho}) + \tilde{\mu}.$$

The matrix X has been generated from multivariate normal distribution MVN[(1, 3, 5, 4, 7, 5, 6, 4, 7, 4)', diag(0, 1.6, 0.7, 3.2, 1.5, 1, 2.8, 2, 1.4, 2.2)]. In the weight matrix W, the weights assigned to nearest neighbour values are twice the weights assigned to the second nearest neighbour values and other neighbour weights are taken as zero. For ensuring the property that the weight matrix is row stochastic, initially we take $w_{i,i+1} = 2$, $w_{i,i+2} = 1$ and all other weights as zero. Then, we divide each element of the selected matrix by the corresponding row sum (which is 3 in our case), so that the sum of elements of each row of W is one. The values of ρ are selected in the range $(1/W_{\max}, 1/W_{\min})$, where $W_{\max}$ and $W_{\min}$ are the maximum and minimum eigen value of W. The coefficients vector β is selected so that it follows the linear restriction $\beta_2 + \beta_3 = 1$. The simulation study has been carried out for the value of parameter $c = 1$, $a = 0.5$ and the results are depicted in Figures 1–4 (supplementary material) and Tables 1–4. Figure 1 / Table 1 shows the percentage gain in efficiency of RFGB estimator over RFLS estimator $b_R(\hat{\rho})$ for $\beta'\beta = 1.745$ when $p = 5, \beta'\beta = 2.1019$ when $p = 10$ and different values of ρ in the range (-0.95, 0.95). Figure 2 / Table 2 shows the percentage gain in efficiency of RFGB estimator over RFSR estimator for $\beta'\beta = 1.745$ when $p = 5, \beta'\beta = 2.1019$ when $p = 10$ and different values of ρ. Figure 3 / Table 3 depicts the percentage gain in efficiency of RFGB estimator over RFLS estimator for $\beta'\beta = 6.29945$ when $p = 5, \beta'\beta = 6.810156$ when $p = 10$ and different values of ρ. Figure 4 / Table 4 shows percentage gain in efficiency of RFGB estimator over RFSR estimator for $\beta'\beta = 6.29945$ when $p = 5, \beta'\beta = 6.810156$ when $p = 10$ and different values of ρ. For each setting of parameters, the experiment is replicated 5,000 times. We have used maximum likelihood estimator of ρ for evaluating RFLS and RFGB estimators.

The empirical mean squared error (EMSE) of any estimator $\hat{\delta}$ of δ is defined as

$$\text{EMSE}\left(\hat{d}\right) = \frac{1}{5000}\sum_{5000}^{j=1}\left(\hat{d}_{(j)} - d\right)'\left(\hat{d}_{(j)} - d\right),$$

where $\hat{d}_{(j)}$ is the estimator of d for the $j-^{th}$ replication. Further, for two estimators $\hat{\delta}$ and $\tilde{\delta}$ of δ, the percentage gain in efficiency of estimator $\hat{\delta}$ over $\tilde{\delta}$ is defined as

$$\%\text{GE} = \frac{\text{EMSE}\left(\tilde{\delta}\right) - \text{EMSE}\left(\hat{\delta}\right)}{\text{EMSE}\left(\hat{\delta}\right)} \times 100.$$

The percentage gains in efficiency of RFGB estimator over RFLS estimator and RFSR estimator are tabulated in Tables 1–4 for different values of n, p, $\beta'\beta$, and ρ.

Table 1: Percentage gain in efficiency of RFGB estimator over RFLS with changing ρ

	ρ	n			
		20	50	100	200
$p = 5$, $\beta'\beta = 1.745$	−0.95	87.652465	91.009943	92.980846	94.559246
	−0.75	86.816747	90.111241	92.401142	94.213808
	−0.55	85.892035	88.73986	91.359453	93.488869
	−0.35	84.820146	86.521877	88.886109	90.992238
	0.05	81.963294	77.318969	70.855288	60.480241
	0.25	80.147649	70.92225	60.433685	49.559817
	0.45	78.307783	66.797965	59.82711	55.822992
	0.65	76.988437	66.766127	63.831039	62.507009
	0.75	76.787609	67.94647	66.155969	65.095276
	0.95	78.395001	72.513772	71.750403	69.452348
$p = 10$ $\beta'\beta = 2.1019$	−0.95	93.550795	93.051869	94.905196	95.744847
	−0.75	93.349477	92.602204	94.295924	95.352843
	−0.55	93.060344	91.738786	93.034165	94.243812
	−0.35	92.688015	90.354913	90.549374	91.213383
	0.05	91.700071	85.674783	79.68406	68.717889
	0.25	91.131797	82.945035	75.356991	67.159497
	0.45	90.614072	81.302088	76.089054	74.992764
	0.65	90.308848	81.728405	79.49427	80.835451
	0.75	90.319279	82.760914	81.376767	82.826556
	0.95	91.106411	87.054551	85.109257	85.691465

Table 2: Percentage gain in efficiency of RFGB estimator over RFSR estimator with changing (ρ)

	ρ	n			
		20	50	100	200
$p = 5$ $\beta'\beta = 1.745$	−0.95	82.923094	88.84156	91.77296	93.625716
	−0.75	81.973603	87.814849	91.080396	93.197345
	−0.55	80.887227	86.163016	89.742814	92.180977
	−0.35	79.598078	83.438066	86.495885	88.61416
	0.05	76.078141	72.28575	62.832244	41.066217
	0.25	73.881609	64.988875	51.951933	35.561192
	0.45	71.764217	60.832483	54.189638	50.187771
	0.65	70.533891	61.660405	60.108793	59.593103
	0.75	70.548218	63.466654	63.134815	62.817361
	0.95	72.83734	69.540567	70.141644	68.755346
$p = 10$ $\beta'\beta = 2.1019$	−0.95	63.784184	88.609957	93.509003	95.092562
	−0.75	63.431254	87.973216	92.718718	94.62062
	−0.55	62.724952	86.680816	91.063367	93.26248
	−0.35	61.687076	84.614498	87.794245	89.476292
	0.05	58.198412	78.066565	74.062091	60.816082
	0.25	55.953324	74.742842	69.65128	61.821461
	0.45	54.063984	73.209599	71.549233	72.366209
	0.65	54.045426	74.23149	76.008274	79.182432
	0.75	55.197672	75.55216	78.235483	81.385816
	0.95	59.635264	81.133665	82.263486	84.352788

From the numerical results, we draw the following conclusions:

1. The RFGB estimator outperforms the restricted RFLS for all the selected parametric settings and restricted RFSR estimators in most of the selected parametric settings. The exceptional parametric values for which RFSR estimator dominates RFGB estimator are $n = 200$, $p = 5, \rho = 0.05, 0.25$.

2. From Figure 1 (supplementary material) and Table 1, we observe that for both $n = 20$ and $n =$

Table 3: Percentage gain in efficiency over FLS with changing ρ

	ρ	n			
		20	50	100	200
$p=5$, $\beta'\beta = 6.29945$	−0.95	88.250714	85.331541	88.357915	88.237576
	−0.75	87.280955	84.165149	86.821451	86.933414
	−0.55	86.086996	81.941857	83.574269	83.424748
	−0.35	84.660808	78.541708	77.84365	76.181153
	0.05	81.075901	68.320117	57.696297	45.772802
	0.25	79.04567	62.868848	48.676744	34.593974
	0.45	77.193568	59.21098	46.204104	36.247221
	0.65	76.078811	58.552617	49.679866	44.080881
	0.75	76.063934	59.350188	52.759539	48.219291
	0.95	78.286289	64.428599	61.277897	54.571567
$p=10$, $\beta'\beta = 6.810156$	−0.95	91.035077	76.382531	81.356926	82.434587
	−0.75	91.054064	76.329566	79.529575	80.673056
	−0.55	91.015573	75.528495	76.211548	76.29519
	−0.35	90.835716	74.147884	71.258038	67.899974
	0.05	90.195831	70.663943	59.092948	41.111684
	0.25	89.827964	69.515726	56.170427	40.60945
	0.45	89.541704	69.690507	57.862524	51.794662
	0.65	89.510794	71.929057	63.310748	63.604441
	0.75	89.694623	74.097215	66.934067	68.359051
	0.95	90.946863	82.663914	75.578299	75.556115

Table 4: Percentage gain in efficiency over SR estimator with changing ρ

	ρ	n			
		20	50	100	200
$p=5$, $\beta'\beta = 6.29945$	−0.95	82.263614	75.762608	81.211744	79.709125
	−0.75	81.273469	74.47021	78.716413	77.191035
	-0.55	79.950618	71.365079	72.864911	69.425762
	-0.35	78.326638	66.463705	61.947065	51.094821
	0.05	74.264449	53.03775	25.743741	−31.970086
	0.25	72.063518	47.744477	18.176582	−27.214677
	0.45	70.215858	46.239419	25.182331	5.164118
	0.65	69.394017	48.748694	37.69438	29.676562
	0.75	69.667279	51.031933	44.066218	38.18759
	0.95	72.708087	58.349346	57.17885	48.729913
$p=10$, $\beta'\beta = 6.810156$	−0.95	50.916155	53.247087	71.684003	75.747382
	−0.75	52.027709	54.314679	68.846558	73.036032
	−0.55	53.116361	53.985354	63.563583	65.938524
	−0.35	53.667719	52.820521	55.74647	51.497067
	0.05	53.709882	50.586393	39.290724	6.996654
	0.25	53.341976	51.173714	38.519666	16.873713
	0.45	52.968889	53.917705	44.651203	40.432942
	0.65	53.120161	59.268304	54.438822	58.156965
	0.75	54.289198	63.114882	59.832638	64.383209
	0.95	59.266461	75.654172	71.174588	73.133859

50, the percentage gain in efficiency of RFGB estimator over RFLS estimator is maximum at $\rho = -0.95$, and decreases gradually up to $\rho \approx 0.65$ ($n = 20$, p =5 and 10), and $\rho \approx 0.5$ ($n = 50$, $p = 5, 10$) and then starts increasing. Further, for $n = 100$ and 200, the percentage gain in efficiency of RFGB estimator over RFLS estimator is maximum at $\rho = -0.95$, decreases gradually up to $\rho \approx 0.45$ ($n = 100$, $p = 5$), $\rho \approx 0.35$ ($n = 100$, $p = 10$), $\rho \approx 0.25$ ($n = 200$, $p = 5$) and $\rho \approx 0.15$ ($n = 200$, p =10), then starts increasing with increasing ρ.

3. It can be seen from Figure 2 (supplementary material) and Table 2 that for both $n = 20$ and 50 the percentage gain in efficiency of RFGB estimator over RFSR estimator is maximum at $\rho = -0.95$ and decreases gradually up to $\rho \approx 0.65$ ($n = 20$, $p = 5$), $\rho \approx 0.55$ ($n = 20$, $p = 10$), $\rho \approx 0.5$ ($n = 50$, $p = 5$) and $\rho \approx 0.45$ ($n = 50$, $p = 10$), then starts increasing with increasing ρ. For $n = 100$ and 200, the percentage gain in efficiency of RFGB estimator over RFSR estimator is maximum at $\rho = -0.95$, decreases gradually up to $\rho \approx 0.35$ ($n = 100$, $p = 5$), $\rho \approx 0.3$ ($n = 100$, $p = 10$), $\rho \approx 0.25$ ($n = 200$, $p = 5$) and $\rho \approx 0.15$ ($n = 200$, $p = 10$), then starts increasing with increasing ρ.
4. For fixed n, in all the cases, the gain in efficiency increases as p increases from 5 to 10.
5. For fixed n and p, as $\beta'\beta$ increases the gain in efficiency usually decreases.
6. An interesting observation for sample size $n = 50$ is that for $p = 5$, the gain in efficiency of RFGB estimator over both RFLS estimator and RFSR estimator usually lower for $\rho > 0$ than the corresponding gain in efficiency for $\rho < 0$ whereas for $p = 10$, the gain in efficiency for $\rho > 0$ is more than that corresponding to negative value of ρ.
7. For a fixed parametric setting, the gain in efficiency of RFGB estimator over RFLS estimator is more than corresponding gain in efficiency of RFGB estimator over RFSR estimator.

5. Concluding remarks

With the objective of achieving robustness with respect to prior distribution and satisfying minimaxity property, a family of RFGB estimators for the regression coefficients vector of a SAR model has been derived in the presence of a set of linear restrictions binding the coefficients vector. The results of the simulation study show that the RFGB estimator outperforms both, the RFLS estimator and RFSR estimator over a wide range of parametric settings. The work on extending the results of the paper for panel data spatial autoregressive models, which also incorporates spatial autoregressive stochastic frontier model for spatio-temporal data, is in progress.

6. Acknowledgements

The authors are grateful to the Editor and three anonymous referees for their useful comments and suggestions.

References

Anselin L (1988). *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Boston, USA.

Anselin L, Le Gallo J, and Jayet H (2008). Spatial panel econometrics, *The Econometrics of Panel Data, Fundamentals and Recent Developments in Theory and Practice*(3rd ed), Springer, Berlin.

Baltagi BH (2011). Spatial panels,*The Handbook of Empirical Economics and Finance*(pp.435–454), Chapman and Hall, New York.

Berger JO (1980). *Statistical Decision Theory: Foundations, Concepts and Methods*, Springer, Berlin.

Brown LD (1971). Admissible estimators, recurrent diffusions, and insoluble boundary value problems, *Annals of Mathematical Statistics*, **42**, 85–90.

Chaturvedi A, Tran VH, and Shukla G (1996). Improved estimation in the restricted regression model with non-spherical disturbances, *Journal of Quantitative Economics*, (12), 115–123.

Chaturvedi A, Wan ATK, and Singh SP (2001). Stein-rule restricted regression estimator in a linear regression model with non-spherical disturbances, *Communications in Statistics Theory & Methods*, **30**, 55–68.

Chaturvedi A and Mishra S (2019). Generalized Bayes estimation of spatial autoregressive models, *Statistics in Transition*, **20**.

Elhorst JP (2003). Specification and estimation of spatial panel data models, *International Regional Science Review*, **26**, 244–268.

Elhorst JP (2010). Spatial panel data models, *Handbook of Applied Spatial Analysis*(377–407), Springer, Berlin.

Elhorst JP (2014). *Spatial Econometrics: from Cross-Sectional Data to Spatial Panels*, Springer, Berlin.

Glass AJ, Kenjegalieva K, and Sickles RC (2016). A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers, *Journal of Econometrics*, **190**, 289–300.

Kubokawa T (1991). An approach to improving the James-Stein estimator, *Journal of Multivariate Analysis*, **36**, 121–126.

Kubokawa T (1994). A unified approach to improving equivariant estimators, *Annals of Statistics*, **22**, 290–299.

Lee LF and Yu J(2015). Spatial panel data models, *The Oxford Handbook of Panel Data*(pp.363–401), Oxford University Press, New York.

LeSage JP and Pace RK (2009). *Introduction to Spatial Econometrics*, Boca Raton, New York.

Maruyama Y (1998). A unified and broadened class of admissible minimax estimators of a multivariate normal mean, *Journal of Multivariate Analysis*,**64**, 196–205.

Maruyama, Y (1999). Improving on the James-Stein estimator, *Statistics & Decisions*, **17**, 137–140.

Pal A, Dubey A and Chaturvedi A (2016). Shrinkage estimation in spatial autoregressive model, *Journal of Multivariate Analysis*,**143**, 362–373.

Rao CR (1973) *Linear Statistical Inference and Its Applications*(2nd ed), Wiley, New York.

Rubin H (1977). Robust Bayesian estimation. In *Proceeding of the Statistical Decision Theory and Related Topics II*, Indiana, USA.

Srivastava VK and Srivastava AK (1983). Improved estimation of coefficients in regression models with incomplete prior information, *Biometrical Journal*, **25**, 775–782.

Srivastava VK and Srivastava AK (1984). Stein-rule estimators in restricted regression models, *Estadistica*, **36**, 89–98.

Srivastava AK and Chandra R (1991). *Improved Estimation of Restricted Regression Model When Disturbances are not Necessarily Normal*, *Sankhyā*, B. 53, 119-133.

Stein C (1973). *Estimation of the mean of a multivariate normal distribution*(pp 345–381), CA, USA .

Tsukamoto T(2019). A spatial autoregressive stochastic frontier model for panel data incorporating a model of technical inefficiency, *Japan & The World Economy*, **50**, 66–77.

Toutenburg H and Shalabh (1996). Predictive performance of the methods of restricted and mixed regression estimators, *Biometrical Journal*, **38**, 951–959.

Toutenburg H and Shalabh (2000). Improved predictions in linear regression models with stochastic linear constraints, *Biometrical Journal*,**42**, 71–86.

Received October 25, 2020; Revised February 19, 2021; Accepted March 12, 2021