STATISTICS IN TRANSITION new series, June 2019
Vol. 20, No. 2, pp. 15–32, DOI 10.21307/stattrans-2019-012

GENERALIZED BAYES ESTIMATION OF SPATIAL AUTOREGRESSIVE MODELS

Anoop Chaturvedi[1], **Sandeep Mishra**[2]

ABSTRACT

The spatial autoregressive (SAR) models are widely used in spatial econometrics for analyzing spatial data involving spatial autocorrelation structure. The present paper derives a Generalized Bayes estimator for estimating the parameters of a SAR model. The admissibility and minimaxity properties of the estimator have been discussed. For investigating the finite sample behaviour of the estimator, the results of a simulation study have been presented. The results of the paper are applied to demographic data on total fertility rate for selected Indian states.

Key words: spatial autoregressive model, prior and posterior distributions, generalized Bayes estimator, admissibility and minimaxity; total fertility rate (TFR).

1. Introduction

Spatial data analysis has attracted considerable attention in econometrics literature for modelling data involving spatial dependence. The Spatial Autoregressive (SAR) models assume that the level of response variable depends on the levels of response variable in the neighbouring regions and thus models such spatial spillover effect. Anselin (1988) provided the theoretical aspects of spatial econometrics. Lesage and Pace (2009) discussed various spatial econometric models including SAR model, spatial Durbin model (SDM), and spatial error model (SEM), along with the classical and Bayesian inference procedures for these models and their various applications.

The Bayesian approach involves combining the data distribution embodied in the likelihood function with prior distributions for the parameters assigned by the practitioner, to produce posterior distributions. However, a major drawback of Bayes procedures is lack of robustness with respect to underlying prior assumptions. As mentioned in Berger (1980), the Bayes estimator derived under normal prior has infinite Bayes risk when true prior is Cauchy distribution. One may consider pre-test estimators, but a serious problem with pre-test estimators is that these estimators provide improvement in specific region of parameter space but perform much worse than the usual maximum likelihood estimator

[1] Department of Statistics, University of Allahabad, Allahabad, 211002, India.
E-mail: anoopchaturv@gmail.com. ORCID ID: https://orcid.org/0000-0002-7322-3331.
[2] Department of Statistics, University of Allahabad, Allahabad, 211002, India.
E-mail: sandeepstat24@gmail.com. ORCID ID: https://orcid.org/0000-0001-5631-3354.

(MLE) or least squares estimator outside this region. Rubin (1977) and Berger (1980) demonstrated that the generalized Bayes estimators are a viable alternative to incorporate prior belief and are more robust with respect to underlying prior assumptions. These estimators provide uniform improvement over MLE/least squares estimator and satisfy minimaxity and admissibility properties.

Stein (1973) considered the generalized Bayes estimator for the multivariate normal mean vector under a scale mixture of prior distributions and suggested that the estimator may dominate the James-Stein and positive part James-Stein estimators. Efron and Morris (1976) presented minimax family of estimators for matrix of multivariate normal means in MANOVA model. Berger (1980) provided robust generalized Bayes estimator for multivariate normal mean and obtained confidence region based on generalized Bayes estimator for the mean vector. Brown (1971) derived a powerful condition for the admissibility of generalized Bayes estimators. Berger (1976), and Maruyama (1998) developed classes of admissible minimax generalized Bayes estimators using Brown's (1971) condition. Kubokawa (1991, 1994) showed that the generalized Bayes estimator dominates the usual James-Stein estimator and derived the sufficient dominance condition. Maruyama (1999) considered the extended Stein's prior distribution, which is the scale mixture of multivariate normal distribution, and demonstrated its admissibility and minimaxity. He also showed that the estimator dominates positive part Stein rule estimator. Pal *et al.* (2016) proposed a family of shrinkage estimators for the coefficients vector of a SAR model and investigated its asymptotic properties.

The present paper considers SAR model involving one period lag spatial dependent variable and derives a generalized Bayes estimator for the regression coefficients vector. The admissibility and minimaxity properties of the estimator are investigated. A simulation study has been carried out to assess the finite sample behaviour of the estimator. For illustration purpose, the results of the paper are applied to demographic data on total fertility rate for selected Indian states.

2. The SAR model and estimators

Let us consider the SAR model:

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

where y is $(n \times 1)$ vector of the observations on a dependent variable collected at each of *n* locations, X is $(n \times p)$ matrix of observations on exogenous variables, β is $(p \times 1)$ vector of regression parameters, ρ is the spatial autoregressive parameter, W is known $n \times n$ spatial weight matrix which indicate the potential interaction between contiguous positions and has been standardized to have row sum of unity. This model is termed as spatial autoregressive model as it combines the standard regression model with spatially lagged dependent variable.

When ρ is known, the ordinary least squares (OLS) estimator of β is

$$b(\rho) = (X'X)^{-1}X'(y - \rho W y)$$

$$= \ (X'X)^{-1}X'y(\rho), \tag{2.2}$$

where $y(\rho)$=$y - \rho Wy$.

The OLS estimator can alternatively be written as

$b(\rho) = b - \rho b_w$.

Here $b = (X'X)^{-1}X'y$ and $b_w = (X'X)^{-1}X'Wy$. Further, the maximum likelihood estimator of σ^2 is

$$\tilde{\sigma}^2 = \frac{v}{n}$$

where

$v = \big(y(\rho) - Xb(\rho)\big)'\big(y(\rho) - Xb(\rho)\big) = [y - \rho Wy]'M[y - \rho Wy]$.

Here $M = I_n - X(X'\,X)^{-1}X'$. When ρ is unknown, we replace it by its estimator

$$\hat{\rho} = \frac{y'W'My}{y'W'MWy} \tag{2.3}$$

in (2.2) to obtain feasible least squares estimator of β as

$$b(\hat{\rho}) = b - \hat{\rho} b_w.$$

Then the estimator of σ^2 is

$$\hat{\sigma}^2 = \frac{\hat{v}}{n}, \text{with } \hat{v} = [y - \hat{\rho}Wy]'M[y - \hat{\rho}Wy]. \tag{2.4}$$

3. Generalized Bayes estimator

For obtaining the Generalized Bayes estimator of regression coefficients vector β, let us write the model (2.1) as

$$y(\rho) = X\beta + u, \tag{3.1}$$

where $y(\rho) = y - \rho Wy$. Then the pdf of $y(\rho)$ is given by

$$p(y(\rho)|\beta, \sigma^2) = \frac{1}{(2\pi)^{n/2}\sigma^n} exp\left\{-\frac{1}{2\sigma^2}(y(\rho) - X\beta)'(y(\rho) - X\beta)\right\}. \tag{3.2}$$

We assume that β follows a g-prior $N(0, \sigma^2 gX'X)$, (see Zellner, 1986) with $g = \lambda^{-1}(1 - \lambda), 0 < \lambda < 1$. Hence the pdf of prior distribution of β is given by

$$p(\beta|\sigma^2, \lambda) \propto \sigma^{-p}\left(\frac{\lambda}{(1-\lambda)}\right)^{p/2} exp\left\{-\frac{\lambda}{2\sigma^2(1-\lambda)}\beta'X'X\beta\right\}. \tag{3.3}$$

We take the prior distribution for λ as

$$p(\lambda) \propto \lambda^{-a}(1 - \lambda)^c I_{(0,1)}(\lambda). \tag{3.4}$$

If c > -1, the prior distribution for λ is proper for $a < 1$ and improper for $a \geq 1$. Let us assume σ^2 to be known. The joint density of $(y(\rho), \beta, \lambda)$ is

$p(y(\rho), \beta, \lambda)$

$= p(y(\rho)|\beta, \sigma^2)p(\beta|\sigma^2, \lambda)p(\lambda)$

$$\propto \sigma^{-(n+p)} exp\left\{-\frac{1}{2\sigma^2}(y(\rho)-X\beta)'(y(\rho)-X\beta)\right\}\lambda^{\frac{p}{2}-a}(1-\lambda)^{-\frac{p}{2}+c}$$
$$\times\, exp\left\{-\frac{\lambda}{2\sigma^2(1-\lambda)}\beta'X'X\beta\right\}$$

$$\propto \sigma^{-(n+p)}\lambda^{\frac{p}{2}-a}(1-\lambda)^{-\frac{p}{2}+c}e^{-v/2\sigma^2} exp\left\{-\frac{1}{2\sigma^2}b(\rho)'X'Xb(\rho)\right\}$$
$$\times\, exp\left\{-\frac{1}{2\sigma^2(1-\lambda)}[\beta'X'X\beta-2(1-\lambda)\beta'X'Xb(\rho)]\right\}$$

$$\propto \sigma^{-(n+p)}\lambda^{\frac{p}{2}-a}(1-\lambda)^{-\frac{p}{2}+c}e^{-v/2\sigma^2} exp\left\{-\frac{\lambda}{2\sigma^2}b(\rho)'X'Xb(\rho)\right\}\times$$
$$exp\left\{-\frac{1}{2\sigma^2(1-\lambda)}\left[\big(\beta-(1-\lambda)b(\rho)\big)'X'X\big(\beta-(1-\lambda)b(\rho)\big)\right]\right\}. \quad (3.5)$$

Integrating (3.5) with respect to β, the joint density of $(y(\rho),\lambda)$ is obtained as $p(y(\rho),\lambda)$

$$\propto \sigma^{-(n+p)}\lambda^{\frac{p}{2}-a}(1-\lambda)^{-\frac{p}{2}+c}e^{-v/2\sigma^2} exp\left\{-\frac{\lambda}{2\sigma^2}b(\rho)'X'Xb(\rho)\right\}\times$$
$$\int_{R^p} exp\left\{-\frac{1}{2\sigma^2(1-\lambda)}\left[\big(\beta-(1-\lambda)b(\rho)\big)'X'X\big(\beta-(1-\lambda)b(\rho)\big)\right]\right\}d\beta$$
$$\propto \sigma^{-n}\lambda^{\frac{p}{2}-a}(1-\lambda)^{c}e^{-v/2\sigma^2} exp\left\{-\frac{\lambda}{2\sigma^2}b(\rho)'X'Xb(\rho)\right\}. \quad (3.6)$$

Then the marginal density of $y(\rho)$ is

$$m\big(y(\rho)\big) \propto \sigma^{-n}e^{-v/2\sigma^2}\int_0^1 \lambda^{\frac{p}{2}-a}(1-\lambda)^{c} exp\left\{-\frac{\lambda}{2\sigma^2}b(\rho)'X'Xb(\rho)\right\}d\lambda. \quad (3.7)$$

Further, the posterior expectation of λ given $y(\rho)$ is obtained as

$$E\big(\lambda|y(\rho)\big)=\frac{\int_0^1 \lambda^{\frac{p}{2}-a+1}(1-\lambda)^{c} exp\left\{-\frac{\lambda}{2\sigma^2}b(\rho)'X'Xb(\rho)\right\}d\lambda}{\int_0^1 \lambda^{\frac{p}{2}-a}(1-\lambda)^{c} exp\left\{-\frac{\lambda}{2\sigma^2}b(\rho)'X'Xb(\rho)\right\}d\lambda}$$
$$=\phi_{a,c}\left(\frac{b(\rho)'X'Xb(\rho)}{\sigma^2}\right), \quad (3.8)$$

where

$$\phi_{a,c}\left(\frac{b(\rho)'X'Xb(\rho)}{\sigma^2}\right)$$
$$=\frac{\int_0^1 \lambda^{\frac{p}{2}-a+1}(1-\lambda)^{c} exp\left\{-\frac{\lambda}{2\sigma^2}b(\rho)'X'Xb(\rho)\right\}d\lambda}{\int_0^1 \lambda^{\frac{p}{2}-a}(1-\lambda)^{c} exp\left\{-\frac{\lambda}{2\sigma^2}b(\rho)'X'Xb(\rho)\right\}d\lambda}$$
$$=\frac{\Gamma\left[2-a+\frac{p}{2}\right]\Gamma\left[2-a+c+\frac{p}{2}\right]\ {}_1F_1\left[2-a+\frac{p}{2},3-a+c+\frac{p}{2},-\frac{b(\rho)'X'Xb(\rho)}{2\sigma^2}\right]}{\Gamma\left[1-a+\frac{p}{2}\right]\Gamma\left[3-a+c+\frac{p}{2}\right]\ {}_1F_1\left[1-a+\frac{p}{2},2-a+c+\frac{p}{2},-\frac{b(\rho)'X'Xb(\rho)}{2\sigma^2}\right]}. \quad (3.9)$$

The Kummer confluent hypergeometric function ${}_1F_1[a;c;z]$ used in expression (3.9) is defined as

$$ {}_1F_1[a;c;z] = \sum_{k=0}^{\infty} \frac{(a)_k}{(c)_k} \frac{z^k}{k!}; $$

where $(a)_0 = 1, (a)_k = a(a+1) \dots (a+k-1)$, is the rising factorial.

Then, the generalized Bayes estimator of β is

$$ \hat{\beta}(\rho) = \left[1 - \phi_{a,c}\left(\frac{b(\rho)'X'Xb(\rho)}{\sigma^2}\right)\right] b(\rho). \tag{3.10} $$

If we substitute

$$ \phi_h(w) = w \frac{\int_0^1 \lambda^{\frac{p}{2}+1-a} (1-\lambda)^c e^{-\frac{\lambda w}{2}} d\lambda}{\int_0^1 \lambda^{\frac{p}{2}-a} (1-\lambda)^c e^{-\frac{\lambda w}{2}} d\lambda} = w\phi_{a,c}\left(\frac{b(\rho)'X'Xb(\rho)}{\sigma^2}\right), $$

then the generalized Bayes estimator $\hat{\beta}(\rho)$ can be represented as

$$ \hat{\beta}(\rho) = \left[1 - \frac{\sigma^2}{b(\rho)'X'Xb(\rho)} \phi_h\left(\frac{b(\rho)'X'Xb(\rho)}{\sigma^2}\right)\right] b(\rho). $$

Theorem 1: *Under the loss function*

$$ L(\hat{\beta}, \beta) = \frac{1}{\sigma^2}(\hat{\beta} - \beta)' X'X(\hat{\beta} - \beta) \tag{3.11} $$

the GB estimator $\hat{\beta}(\rho)$ *has finite risk.*

Proof: Let us write

$$ Z = \frac{1}{\sigma}(X'X)^{\frac{1}{2}} b(\rho);\ \theta = \frac{1}{\sigma}(X'X)^{\frac{1}{2}} \beta. $$

Then

$$ R[\hat{\beta}(\rho), \beta] = \frac{1}{\sigma^2} E\left[(\hat{\beta}(\rho) - \beta)' X'X(\hat{\beta}(\rho) - \beta)\right] $$

$$ = E\left[(Z-\theta)'(Z-\theta) + \frac{1}{\|Z\|^2} {\phi_h}^2(\|Z\|^2) - 2\frac{(Z-\theta)'Z\phi_h(\|Z\|^2)}{\|Z\|^2}\right]. $$

Since $Z \sim N(\theta, I_p)$, we have

$$ R[\hat{\beta}(\rho), \beta] = p + E\left[\frac{1}{\|Z\|^2} {\phi_h}^2(\|Z\|^2) - 2\frac{(Z-\theta)'Z\phi_h(\|Z\|^2)}{\|Z\|^2}\right]. $$

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

$$ E\left[\frac{1}{\|Z\|^2} {\phi_h}^2(\|Z\|^2)\right] \le E[\|Z\|^2] = p + \theta'\theta < \infty. $$

Further by Schwarz's inequality

$$ E\left[\frac{(Z-\theta)'Z\phi_h(\|Z\|^2)}{\|Z\|^2}\right] \le \left[E(Z-\theta)'(Z-\theta) E\left\{\frac{\phi_h(\|Z\|^2)^2}{\|Z\|^2}\right\}\right]^{1/2} $$

$$\leq \left[pE[\|Z\|^2]\right]^{1/2}$$
$$= [p(p+\theta'\theta)]^2 < \infty.$$

Hence the risk of $\hat{\beta}(\rho)$ is finite■

Theorem 2: The GB estimator is admissible if and only if $a \leq 2$.

Proof: We have

$$f_R(\|z\|^2) = \int_0^1 e^{-\lambda\|z\|^2/2}\lambda^{\frac{p}{2}-a}(1-\lambda)^c d\lambda$$
$$= 2^{\frac{p}{2}-a+1}\int_0^{\frac{1}{2}} e^{-t\|z\|^2}\, t^{\frac{p}{2}-a}(1-2t)^c dt$$
$$= 2^{\frac{p}{2}-a+1}\int_0^{\infty} e^{-t\|z\|^2}\, t^{\frac{p}{2}-a}(1-2t)^c I_{(0,\frac{1}{2})}(t) dt.$$

Using Tauberlian theorem (see Maruyama, 2000, p. 37), we observe that as $t \to 0$, $t^{\frac{p}{2}-a}(1-2t)^c I_{(0,\frac{1}{2})}(t) \sim t^{\frac{p}{2}-a}$. Hence we have

$$f_h(\|z\|^2) \sim 2^{\frac{p}{2}-a+1}\Gamma\left(\frac{p}{2}-a+1\right)\|z\|^{-2\left(\frac{p}{2}-a+1\right)}. \quad (3.12)$$

Following Maruyama (2000), to show that the GB estimator is admissible it is necessary and sufficient to show that $\int_1^\infty f_h^{-1}(t)\, t^{\frac{p}{2}} dt$ diverges. Using equation (3.12), we have

$$\int_1^\infty f_h^{-1}(t)\, t^{\frac{p}{2}} dt \sim 2^{\frac{p}{2}-a+1}\Gamma^{-1}\left(\frac{p}{2}-a+1\right)\int_1^\infty t^{(-a+1)}\, dt,$$

which diverges as long as $a \leq 2$. This leads to the required result■

Theorem 3: The generalized Bayes estimator $\hat{\beta}(\rho)$ is minimax whenever $3-\frac{p}{2} \leq a \leq \frac{p}{2}+1$.

Proof: Under the loss function (3.11) the difference between the risks of GB estimator $\hat{\beta}(\rho)$ and the OLS estimator $b(\rho)$ is given by

$$R\left[\hat{\beta}(\rho),\beta\right] - R[b(\rho),\beta] = E\left[\frac{1}{\|Z\|^2}\Phi^2{}_h(\|Z\|^2) - 2\frac{(Z-\theta)'Z\phi_h(\|Z\|^2)}{\|Z\|^2}\right].$$

Now

$$E\left[(Z-\theta)'Z\frac{\phi_h(\|Z\|^2)}{\|Z\|^2}\right] = E\left[\frac{\partial}{\partial Z'}\left\{Z\frac{\phi_h(\|Z\|^2)}{\|Z\|^2}\right\}\right]$$
$$= E\left[p\frac{\phi_h(\|Z\|^2)}{\|Z\|^2} - 2\frac{\phi_h(\|Z\|^2)}{\|Z\|^2} + 2\phi'{}_h(\|Z\|^2)\right].$$

Hence

$$R\left[\hat{\beta}(\rho),\beta\right] - R[b(\rho),\beta]$$

$$= E\left[\frac{1}{\|Z\|^2}\phi^2{}_h(\|Z\|^2) - 2(p-2)\frac{\phi_h(\|Z\|^2)}{\|Z\|^2} - 4\phi'{}_h(\|Z\|^2)\right] \tag{3.13}$$

Now we have

$$\frac{\phi_h(w)}{w} = \frac{\int_0^1 \lambda^{\frac{p}{2}+1-a}(1-\lambda)^c e^{-\frac{\lambda w}{2}} d\lambda}{\int_0^1 \lambda^{\frac{p}{2}-a}(1-\lambda)^c e^{-\frac{\lambda w}{2}} d\lambda},$$

so that

$$\frac{\partial}{\partial w}\left[\frac{\phi_h(w)}{w}\right] = \frac{1}{2}\frac{\left\{\int_0^1 \lambda^{\frac{p}{2}+1-a}(1-\lambda)^c e^{-\frac{\lambda w}{2}} d\lambda\right\}^2 - \left\{\int_0^1 \lambda^{\frac{p}{2}+2-a}(1-\lambda)^c e^{-\frac{\lambda w}{2}} d\lambda\right\}\left\{\int_0^1 \lambda^{\frac{p}{2}-a}(1-\lambda)^c e^{-\frac{\lambda w}{2}} d\lambda\right\}}{\left\{\int_0^1 \lambda^{\frac{p}{2}-a}(1-\lambda)^c e^{-\frac{\lambda w}{2}} d\lambda\right\}^2} \tag{3.14}$$

Let us write

$$f_h(\lambda) = \frac{\lambda^{\frac{p}{2}-a}(1-\lambda)^c e^{-\frac{\lambda w}{2}}}{\int_0^1 \lambda^{\frac{p}{2}-a}(1-\lambda)^c e^{-\frac{\lambda w}{2}} d\lambda}, 0 < \lambda < 1.$$

Then

$$\frac{\partial}{\partial w}\left[\frac{\phi_h(w)}{w}\right] = -\frac{1}{2}\left[E_{f_h(\lambda)}(\lambda^2) - \left\{E_{f_h(\lambda)}(\lambda)\right\}^2\right] \le 0.$$

Again

$$\phi'{}_h(w) = \frac{\partial}{\partial w}\left\{w\frac{\phi_h(w)}{w}\right\}$$

$$= \frac{\phi_h(w)}{w} + w\frac{\partial}{\partial w}\left\{\frac{\phi_h(w)}{w}\right\}$$

$$= E_{f_h(\lambda)}(\lambda) - \frac{1}{2}E_{f_h(\lambda)}(\lambda^2) + \frac{1}{2}\left\{E_{f_h(\lambda)}(\lambda)\right\}^2$$

$$\ge \frac{1}{2}\left\{E_{f_h(\lambda)}(\lambda)\right\}^2 \ge 0.$$

Notice that $0 \le \lambda \le 1,$ so that $E_{f_h(\lambda)}(\lambda) - \frac{1}{2}E_{f_h(\lambda)}(\lambda^2) \ge 0.$

We also observe that $\phi_h(w)$ and $\frac{\phi_h(w)}{w}$ are monotone in opposite directions. Therefore we obtain

$$R\left[\hat{\beta}(\rho), \beta\right] - R[b(\rho), \beta]$$

$$\le E\left[\frac{\phi_h(\|Z\|^2)}{\|Z\|^2}\right]E\left[\phi_h(\|Z\|^2) - 2(p-2)\right] - 4E\left[\phi'{}_h(\|Z\|^2)\right]$$

$$\le E\left[\frac{\phi_h(w)}{w}\right]E[\{\phi_h(w) - 2(p-2)\}],$$

which is less than or equal to zero whenever

$$0 \le E[\{\phi_h(w)\}] \le 2(p-2). \tag{3.15}$$

When w is large, we may approximate $\phi_h(w)$ as

$$\phi_h(w) \approx 2\left(\frac{p}{2} - a + 1\right).$$

Further $\phi_h(w)$ is an increasing function of w. Hence, a sufficient dominance condition is

$$0 \leq 2\left(\frac{p}{2} - a + 1\right) \leq 2(p-2),$$

or

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

This leads to the required result■

When ρ and σ^2 are unknown, we replace them by their estimators $\hat{\rho}$ and $\hat{\sigma}^2$ defined in (2.3) and (2.4) respectively to obtain feasible generalized Bayes estimator of β.

4. Simulation study

In this section we carry out a simulation study using R Software to assess the finite sample behaviour of proposed generalized Bayes estimator. The observations on response variable *y* are generated by using the model (2.1). In simulation study we compare the risks of the usual feasible least squares estimator $b(\hat{\rho}) = (X'X)^{-1}X'(y - \hat{\rho}Wy)$ with the following feasible version of GB estimator:

$$\hat{\beta}(\hat{\rho}) = \left[1 - \phi_{a,c}\left(\frac{b(\hat{\rho})'X'Xb(\hat{\rho})}{\hat{\sigma}^2}\right)\right] b(\hat{\rho}).$$

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, say $(w_1, w_2), \dots, (w_{n-1}, w_n)$, are double the weights assigned to the second nearest neighbour values, say, $(w_1, w_3), \dots, (w_{n-2}, w_n)$ and other neighbour weights are taken as zero. The property of weight matrix to be row stochastic is also satisfied. Further, to ensure the stationarity, the values of ρ are selected in the range $\left(\frac{1}{W_{max}}, \frac{1}{W_{min}}\right)$, where W_{max} and W_{min} are, respectively, the maximum and minimum eigen values of *W*. We select *c* = 1, *a* = 0.5 and the results are depicted in figures 1-6. Figures 1 and 2 plot the percentage gain in efficiency of GB estimator over feasible least squares estimator when we vary ρ in the range (-0.95, 0.95). For p=5, $\beta'\beta$=1.525 and for p=10 $\beta'\beta$=1.8819. Further figures 3-6 plot percentage gain in efficiency for variation in $\beta'\beta$ and fixed ρ, *n*, and *p*. The selected values of ρ in figures 3-6 are 0.25 and 0.75, selected values of *n* are 20, 50, 100, 200 and those of *p* are 5 and 10. For each setting of parameters, the experiment is replicated 5000 times. We have used maximum likelihood estimator of ρ for evaluating feasible least squares and feasible GB estimators. The percentage gain in efficiency due to feasible GB estimator $\hat{\beta}(\hat{\rho})$ over feasible least squares estimator $b(\hat{\rho})$ is calculated using formula:

$$\% \text{ gain in efficiency } GE(\hat{\beta}(\hat{\rho})) = \frac{ER(b(\hat{\rho})) - ER(\hat{\beta}(\hat{\rho}))}{ER(b(\hat{\rho}))} \times 100.$$

Empirical risk of the estimator $\hat{\beta}(\hat{\rho})$ based on 5000 replications has been evaluated as

$$ER(\hat{\beta}(\hat{\rho})) = E\left(\left(\hat{\beta}(\hat{\rho}) - \beta\right)'\left(\hat{\beta}(\hat{\rho}) - \beta\right)\right)$$
$$\approx \frac{1}{5000}\sum_{r=1}^{5000}\left(\hat{\beta}(\hat{\rho})_r - \beta\right)'\left(\hat{\beta}(\hat{\rho})_r - \beta\right),$$

where $\hat{\beta}(\hat{\rho})_r$ is the estimated β based on r-th replication.

The main findings of the simulation are as follows:

1. GB estimator performs better than the FLS estimator, in all the selected parametric settings.
2. From Figures 1 and 2 we observe that the percentage gain in efficiency remains almost constant for $\rho < 0$ and then it starts increasing gradually except for *n*= 20, where it increases for $\rho > -0.25$.
3. For *n*=20, *p*=5 and *n*=20, *p*=10, the gain in efficiency is maximum when ρ is close to 0.6 and then again it starts decreasing with increasing ρ.
4. For *n*= 50, *p*= 5, the gain in efficiency increases for $\rho > 0.25$ and for *n*= 50, *p*= 10, it increases for $\rho > 0.5$.
5. For *n*= 100, *p*= 5, the gain in efficiency increases for $\rho > 0.35$ and for *n*= 100, *p*= 10 it increases for $\rho > 0.5$.
6. For *n*= 200, both for *p*= 5, and *p*=10, the gain in efficiency usually keeps on increasing for $\rho > 0.5$.
7. For fixed *n* the gain in efficiency decreases as *p* increases from 5 to 10.
8. Figures 3-6 show that the percentage gain in efficiency increases as the value of ρ increases from 0.25 to 0.75. The gain in efficiency decreases with increasing $\beta'\beta$.
9. For *n*= 20, *p*= 5, 10, $\rho = 0.25$ the percentage gain in efficiency is almost constant for $\beta'\beta > 9$. For $\rho = 0.75$ it gradually decreases with increasing $\beta'\beta$ up to $\beta'\beta = 20$ and, after that, it remains almost constant. For all others combination of parameters the gain in efficiency remains almost constant as long as $\beta'\beta > 3$.

5. Application to TFR Data

In this section we present an application of SAR model for modelling the total fertility rates (TFR) of selected Indian states. We use the causal variables Female literacy rate (FLIT), Headcount poverty ratio (HCPR), and Percentage of urban population (PUP), which control the socio economic conditions influencing TFR, see table 4. For incorporating the influence of spatial structure of states in India, first order spatial autoregressive term is also included. The spatial weight matrix is formed using spatial contiguity matrix. To form the contiguity matrix, define Vij =1 for two spatial units (states in our example) that own a common border of non-zero length, else equal to zero. Since an element is not contiguous or neighbouring to itself, the main diagonal elements of the matrix are zero. The matrix **V** is then scaled to make it row stochastic. Denoting such a standardized first order contiguity matrix by W, its *(i,j)*-th element, say W_{ij} is given by

$$W_{ij} = \frac{w_{ij}'}{\sum_{\substack{j=1 \\ i \neq j}}^{n} w_{ij}'} \tag{2.4}$$

where $W_{ij}' = 1$ if i is linked to j, and 0 otherwise. Moran's I statistic for W is $I = z'Wz/z'z$, where z is $n \times 1$ vector of variables expressed as deviations from the mean. The global Moran's *I* statistic is used to examine the variables in our data set for global autocorrelation. If the observed value of *I* is greater than its expected value, then corresponding observation tend to be surrounded by neighbours with similar values. On the other hand if *I* is less than its expected value, the observation tend to be surrounded by dissimilar values, see Schabenberger and Gotway (2005) for details.

The regression coefficients of fitted SAR model are estimated using feasible LS and feasible GB estimators. In sample predicted values of TFR for different states are also computed based on both the estimators, see table 5. Table 6 gives the estimated coefficients using both of these estimators. The estimated values of spatial autocorrelation coefficient is 0.5923. Table 7 gives the observed and expected value of the Moran's *I* for each of the variables considered in the analysis. We observe that HCPR shows the highest degree of spatial correlation, followed by the FLIT while the PUP shows the lowest degree of spatial autocorrelation among the independent variables. The results from the empirical investigation indicate that the feasible GB estimator performs better than FLS estimator of regression coefficients in terms of predictive efficiency.

6. Concluding remarks

With the objective of achieving robustness with respect to prior distribution and satisfying admissibility and minimaxity properties, we have developed a family of generalized Bayes estimators for the regression coefficients vector of a SAR model. The simulation study has been carried out to examine the efficiency properties of GB estimator and it was observed that GB estimator provides improvement over the usual least squares estimator for a wide range of the parametric settings.

REFERENCES

ANSELIN, L., (1988). Spatial econometrics, methods, and models, Dordrecht: Kluwer Academic.

ANSELIN, L., REY, S., (2010). Perspectives on spatial data analysis, Berlin: Springer, DOI:10.1007/978-3-642-01976-0.

BERGER, J., (1976). Admissible minimax estimation of a multivariate normal mean with arbitrary quadratic loss. Ann. Statist., 4, pp. 223–226.

BERGER, J. O., (1980): Statistical Decision Theory: Foundations, Concepts and Methods, Springer, N.Y.

BROWN, L. D., (1971). Admissible estimators, recurrent diffusions, and insoluble boundary value problems, Ann. Math. Statist., 42, pp. 855–903.

EFRON, B., MORRIS, C., (1976). Families of minimax estimators of the mean of a multivariate normal distribution. Ann. Statist., 4, pp. 11–21.

FOURDRINIER, D., STRAWDERMAN. W. E., WELLS, T., (1998). On the construction of Bayes minimax estimators, Ann. Statist., 26, pp. 660–671.

KUBOKAWA, T., (1991). An approach to improving the James-Stein estimator, J. Multivariate Anal., 36, pp. 121–126.

KUBOKAWA, T., (1994). An unified approach to improving equivariant estimators, Ann. Statist., 22, pp. 290–299.

LESAGE, J. P., (1998): Spatial Econometrics, www.spatial-econometrics.com.

LESAGE, J. P, PACE, R. K., (2009). Introduction to spatial econometrics, Boca Raton, FL: CRC, DOI:10.1201/9781420064254.

MARUYAMA, Y., (1998). A unified and broadened class of admissible minimax estimators of a multivariate normal mean, J. Multivariate Anal., 64, 196–205.

MARUYAMA, Y., (1999). Improving on the James-Stein Estimator, Statistics & Decisions, 17, pp. 137–140.

MARUYAMA, Y., (2000). Minimax admissible estimation of a multivariate normal mean and improvement upon the James-Stein estimator, Ph.D. dissertation, Graduate School of Economics, University of Tokyo.

PAL, A., CHATURVEDI, A., DUBEY, A., (2016). Shrinkage estimation in spatial autoregressive model, Journal of Multivariate Analysis, 143, pp. 362–373.

RUBIN, H. (1977). Robust Bayesian estimation, In Statistical Decision Theov and Related Topics II. (S. S. Gupta and D. S. Moore, eds.), Academic Press.

SCHABENBERGER, O., GOTWAY, C. A., (2005): Statistical Methods for Spatial Data Analysis, Chapman and Hall/CRC: Boca Raton, FL.

STEIN, C., (1973). Estimation of the mean of a multivariate normal distribution, In Proc. Prague Symp. Asymptotic Statist., pp. 345–381.

ZELLNER, A., (1986). On Assessing Prior Distributions and Bayesian Regression Analysis with g-Prior Distributions, In: Goel, P. and Zellner, A., Eds., Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti, Elsevier Science Publishers, Inc., New York, pp. 233–243.

APPENDIX

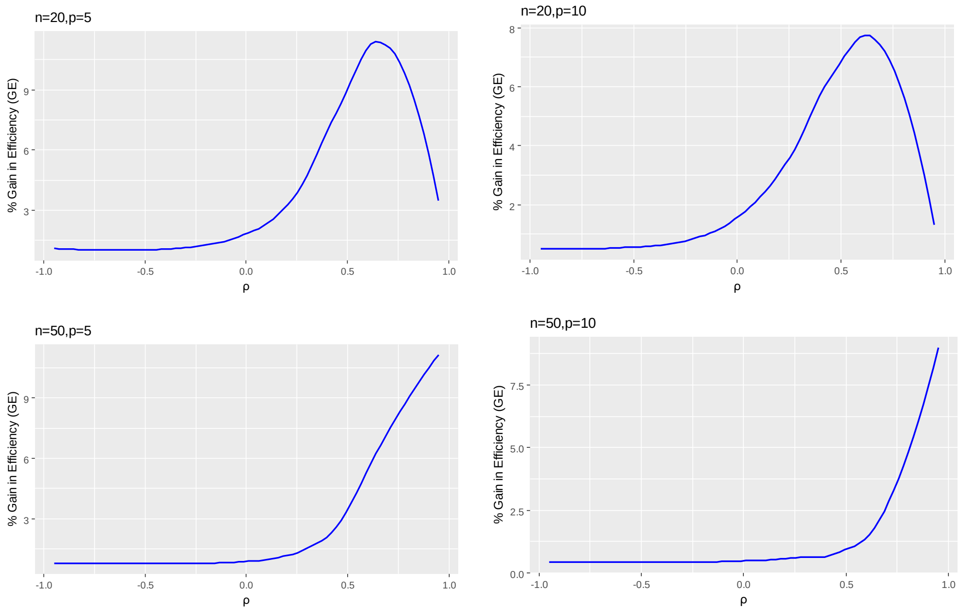

Figure 1. Percentage Gain in efficiency due to change in ρ

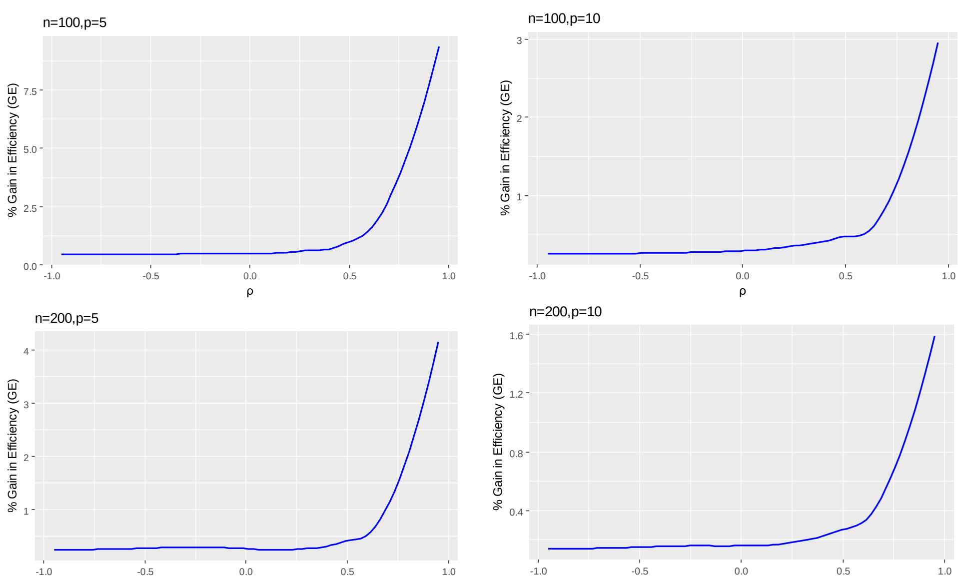

Figure 2. Percentage Gain in efficiency due to change in ρ

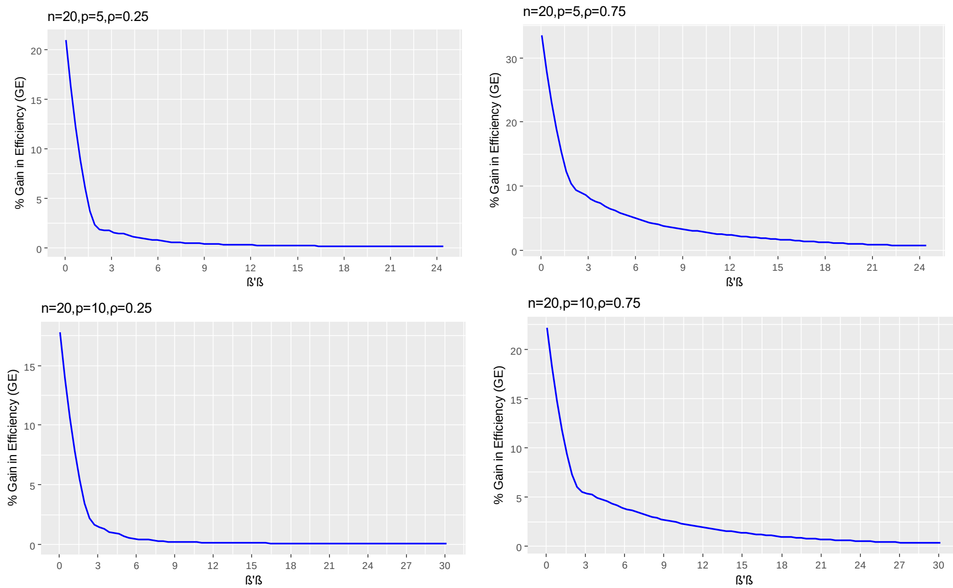

Figure 3. Percentage gain in efficiency due to change in length of parameter β i.e. $\beta'\beta$

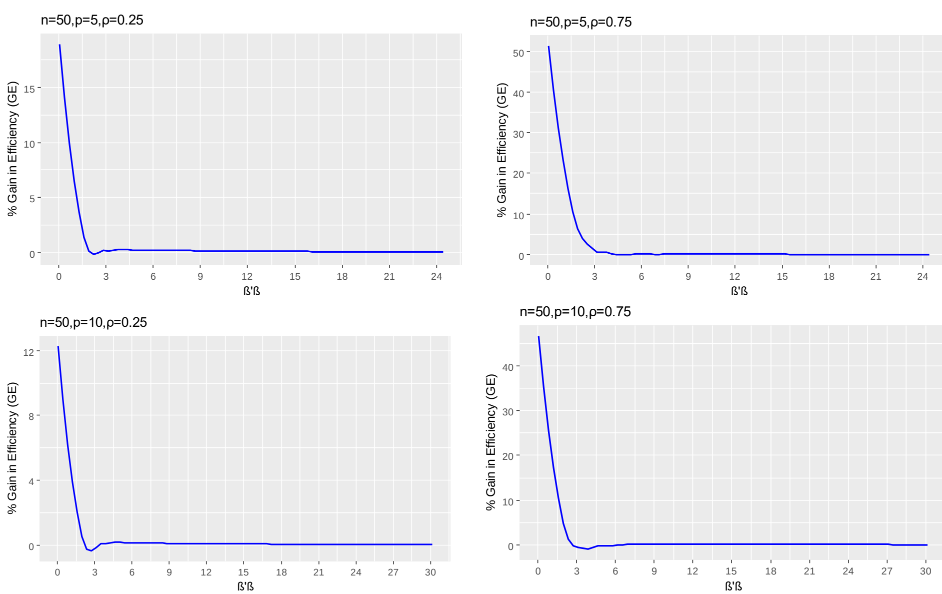

Figure 4. Percentage gain in efficiency due to change in length of parameter β i.e. $\beta'\beta$

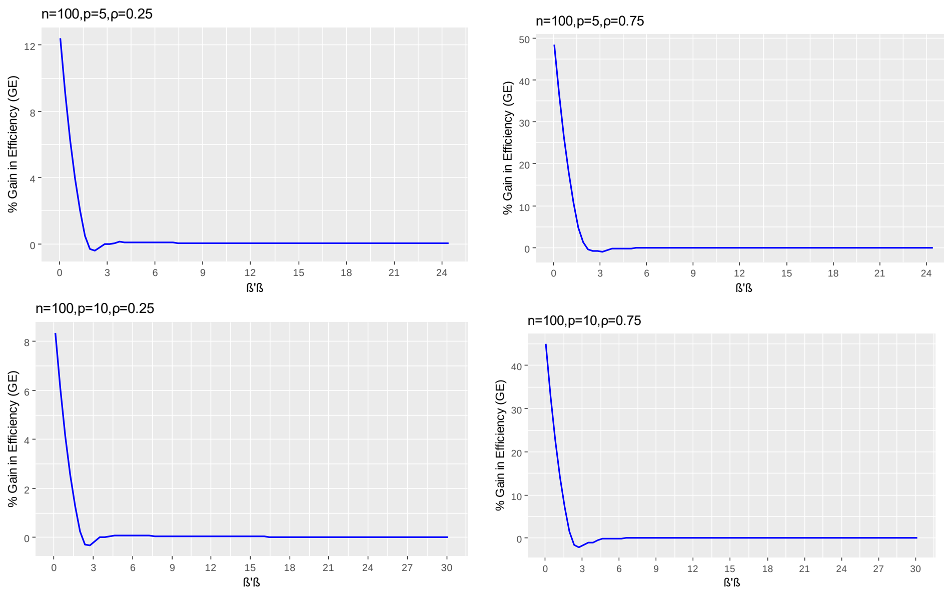

Figure 5. Percentage gain in efficiency due to change in length of parameter β *i.e.* $\beta'\beta$

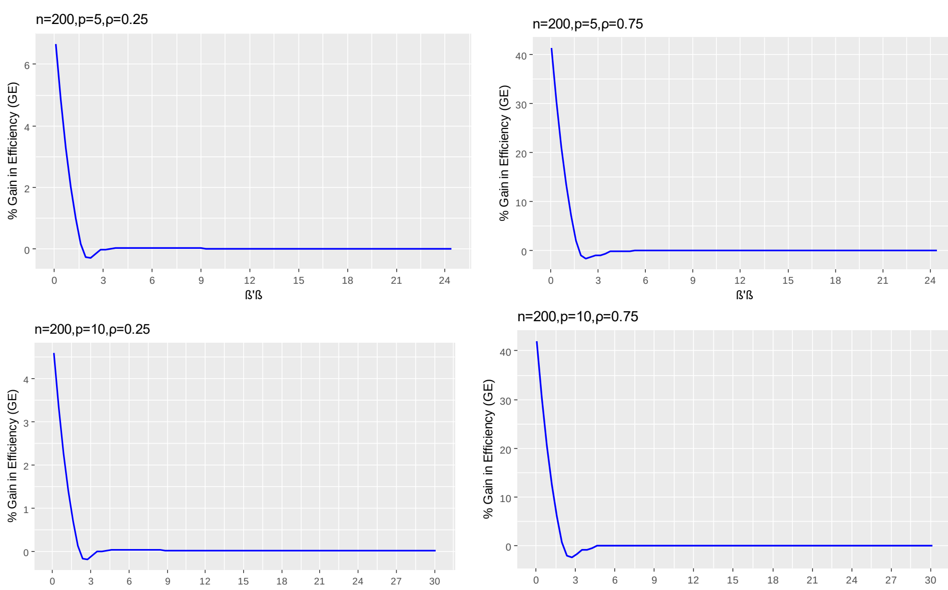

Figure 6. Percentage gain in efficiency due to change in length of parameter β *i.e.* $\beta'\beta$

Table 1. Percentage gain in efficiency due to $\beta'\beta$ for $p = 5$

ρ	$\beta'\beta$	*n=20*	*n=50*	*n=100*	*n=200*
0.25	0.244	13.5009	11.08661	6.037487	2.837103
	0.976	5.328835	2.463001	1.107283	0.476686
	2.196	2.738421	0.815412	0.357004	0.135803
	3.904	1.454371	0.38847	0.158998	0.048988
	6.1	0.778182	0.226422	0.086162	0.019632
	8.784	0.459649	0.148478	0.052888	0.008067
	11.956	0.303455	0.104995	0.03536	0.002975
	15.616	0.216832	0.078216	0.025173	0.000648
	19.764	0.163445	0.060558	0.01869	-0.00044
	24.4	0.127986	0.048275	0.014404	-0.00647
0.75	0.244	26.1813	42.82635	40.14246	31.70275
	0.976	15.00917	17.25584	8.275401	3.31876
	2.196	10.36237	4.615964	1.276183	0.490065
	3.904	7.243892	1.117806	0.354763	0.144584
	6.1	4.940427	0.386443	0.140805	0.054013
	8.784	3.411094	0.190456	0.064894	0.020024
	11.956	2.426919	0.110726	0.031528	0.004777
	15.616	1.603265	0.071166	0.015088	-0.0027
	19.764	1.035082	0.049153	0.00642	-0.00648
	24.4	0.67636	0.035669	0.001649	-0.00831

Table 2. Percentage gain in efficiency due to $\beta'\beta$ for $p = 10$

ρ	$\beta'\beta$	*n=20*	*n=50*	*n=100*	*n=200*
0.25	0.301104	12.03765	6.177598	3.646498	1.938245
	1.204416	5.176314	1.081055	0.63606	0.337801
	2.709936	2.276713	0.405873	0.226304	0.110739
	4.817664	0.912911	0.213313	0.112668	0.049924
	7.527600	0.387297	0.131917	0.067129	0.027217
	10.83974	0.213107	0.089933	0.044548	0.016788
	14.7541	0.137561	0.065157	0.031726	0.011300
	19.27066	0.097152	0.049539	0.023766	0.008085
	24.38942	0.072792	0.038788	0.018471	0.006074
	30.1104	0.056718	0.031372	0.014747	0.004733
0.75	0.301104	15.911	38.03073	33.88339	28.86204
	1.204416	8.84946	8.273935	2.498556	1.544655
	2.709936	6.279212	1.14234	0.438372	0.262278
	4.817664	4.539674	0.360881	0.16237	0.092256
	7.5276	3.2402	0.174061	0.079288	0.042532
	10.83974	2.208562	0.102341	0.044992	0.022133
	14.7541	1.409047	0.067498	0.02815	0.012272
	19.27066	0.867744	0.047913	0.018992	0.00696
	24.38942	0.524181	0.035872	0.01348	0.003907
	30.1104	0.299729	0.027958	0.010022	0.002106

Table 3. Percentage gain in efficiency due to ρ

	ρ	n=20	n=50	n=100	n=200
p=5	-0.95	1.091354	0.7822986	0.4439748	0.2370162
	-0.75	1.038856	0.7635631	0.4476627	0.2456971
	-0.55	1.030078	0.7567423	0.4588296	0.2617378
	-0.35	1.105928	0.7634651	0.4737263	0.2834117
	0.05	2.060012	0.9269042	0.4977228	0.2475341
	0.25	3.778747	1.3376279	0.5968474	0.2441839
	0.45	7.496059	2.5998401	0.9269096	0.3940382
	0.65	12.3777	6.0657974	1.8282938	0.7344557
	0.75	12.3601	9.2943542	3.0362531	1.1202461
	0.95	2.127767	9.4626555	9.7513806	4.6275375
p=10	-0.95	0.505288	0.4321381	0.2548265	0.1388775
	-0.75	0.509345	0.4270315	0.2572827	0.1424235
	-0.55	0.549603	0.4271182	0.2627093	0.1489208
	-0.35	0.667549	0.4339027	0.270695	0.1577049
	0.05	1.821497	0.5037863	0.3031573	0.1600749
	0.25	3.566544	0.6224854	0.3580659	0.1835203
	0.45	6.001154	0.8986891	0.4700185	0.2592248
	0.65	8.427731	1.6429605	0.6812483	0.4055339
	0.75	7.393491	2.8589676	0.8883084	0.5485187
	0.95	0.880642	9.2330294	3.3604948	1.7665873

Table 4. Total fertility rate, Female literacy rate, Headcount poverty ratio, and Percentage of urban population in major states of India

STATE	TFR	FLIT	PUP	HCPR
A.P.	1.8	50.4	27.3	15.8
ASSAM	2.4	54.6	12.9	19.7
BIHAR	4	33.1	10.5	41.4
CHHATTISGARH	2.6	51.9	20.1	40.9
GUJARAT	2.4	57.8	37.4	16.8
HARYANA	2.7	55.7	28.9	10
H.P.	1.9	67.4	9.8	14
J&K	2.4	43	24.8	5.4
JHARKHAND	3.3	38.9	22.2	40.3
KARNATAKA	2.1	56.9	34	25
KERALA	1.9	87.7	26	15
M.P.	3.1	50.3	26.5	38.3
MAHARASTRA	2.1	67	42.4	30.7
ODISHA	2.4	50.5	15	46.4
PUNJAB	2	63.4	33.9	8.4
RAJASTHAN	3.2	43.9	23.4	22.1
TAMIL NADU	1.8	64.4	44	22.5
U.P.	3.8	42.2	20.8	32.8
UTTARAKHAND	2.6	59.6	25.7	39.6
W.B.	2.3	59.6	28	24.7

Sources: (i) TFR from EPWRF (2010-11) (ii) URBAN and FLIT from Census of India (2001) and (iii) POV from Planning Commission (2011).

Table 5. Predicted TFR

STATE	OBSERVED	PFLS	PGB
A.P.	1.8	2.49576	2.47243
ASSAM	2.4	2.68139	2.65632
BIHAR	4	3.60914	3.5754
CHHATTISGARH	2.6	2.94126	2.91376
GUJARAT	2.4	2.37845	2.35621
HARYANA	2.7	2.3918	2.36944
H.P.	1.9	2.17028	2.14999
J&K	2.4	2.47091	2.44781
JHARKHAND	3.3	3.39959	3.36781
KARNATAKA	2.1	2.07161	2.05225
KERALA	1.9	1.22177	1.21034
M.P.	3.1	2.81993	2.79356
MAHARASTRA	2.1	2.17079	2.1505
ODISHA	2.4	3.04555	3.01708
PUNJAB	2	2.10204	2.08239
RAJASTHAN	3.2	2.81922	2.79286
TAMIL NADU	1.8	1.81505	1.79808
U.P.	3.8	3.11339	3.08428
UTTARAKHAND	2.6	2.55796	2.53404
W.B.	2.3	2.76707	2.7412

Table 6. FLS and GB Estimators of Coefficients

Variable	$\widehat{\boldsymbol{\beta}}_{FLS}$	$\widehat{\boldsymbol{\beta}}_{GB}$
Constants	2.42061	2.39798
FLIT	-0.0268	-0.0265
PUP	-0.0041	-0.0041
HCPR	0.00688	0.00682

Table 7. Global Moran's *I* values

Variable	Observed *I*	E[I]
TFR	0.4339011	-0.05263158
FLIT	0.1637386	-0.05263158
PUP	0.1379433	-0.05263158
HCPR	0.3973285	-0.05263158