Al-Qadisiyah Journal for Administrative and Economic Sciences ISSN$_{Online}$ : 2312-9883
*QJAE*, Volume 25, Issue 4 (2023) ISSN$_{Print}$ : 1816-9171

# Bayesian Reciprocal Bridge Regression For Ordinal Models With an Application

Rahim J. Al-Hamzawi
Elaa Jaber Jain AL-Badiri
College of Business and Economics, Al-Qadisiyah University
*Corresponding Author: Elaa Jaber Jain AL-Badiri*

**Abstract :** Regularization methods always focus on the selection of variables (vs) and estimation of regression parameters. So it is relied upon (vs) Because it is difficult to identify the important variables in the model, if the number of common variables is very large, and to choose the most effective variables in the model.
In this paper, we proposed a new method for selecting an ordinal model . This method is Bayesian reciprocal bridge regression for the ordinal model (BOrBridge ), We have developed a new hierarchical Bayesian regression model Bayesian reciprocal bridge for the ordinal model (BOrBridge). Which motivates us to suggest a Gibbs sample New to sample parameters from the posteriors . The performance of the proposed approach was examined through simulation studies and real data analysis.
The results show that our proposed method (BOrBridge) After comparing it with the AIC and BIC Identifying the best model in standard ordinal regression works very well This also indicates the convergence of the construct-specific Gibbs samples to the posteriors distribution, was quick and the mixing was good . The research reached important conclusions, represented by the superiority of the proposed method over existing methods in selecting variables and estimating parameters.

**1. INTRODUCTION:** Linear regression is used to identify and describe the relationship between a dependent variable and a set of explanatory (independent) variables (Koenker and Bassett, (1982)). The first to refer to the term regression was the English scientist Galton (14 February 1822–January 1911) in an article he published in the late nineteenth century. It has wide applications in economics and engineering, as well as in various sciences such as agriculture, physics, medicine, and the social sciences (Bark, R., 2004).
The linear regression equation can be written as (Kutner, Nachtsheim, Neter, and Li, 2005). :

$$y_i = X_i^t \boldsymbol{\beta} + \varepsilon_i , \qquad \forall i = 1, \dots, n, \qquad (1)$$

where $y_i = (y_1, \dots, y_n)$ is the $n \times 1$ vector of centered $x_i = (x_1, \dots, x_n)^t$ is the $p \times 1$ matrix of standardized regressions ,$\boldsymbol{\beta}$ is the $p \times 1$ vector of coefficients to be estimated, and the vector of independent and identically distributed normal errors with mean 0 and variance $\sigma^2$ . The linear regression model tasted some conditions, such as the regression model being linear in parameters, the normality of the error distribution, the mean of residuals being zero, and the homoscedasticity (constant variance) of residuals. No autocorrelation of residuals ،the predictors and residuals are uncorrelated $n > p$ . The variability in predictor values is positive. The predictors and responses are specified correctly. There is no perfect multicollinearity ( Berk, R. A. (2004), Peter D. correlated . The variability in predictor values is positive. The predictors and responses are specified correctly. There is no perfect multicollinearity ( Berk, R. A. (2004), Peter D Hoff, (2009)). But there is difficulty in realizing these conditions when there is an increase in the number of variants. A problem appears with estimation, and this needs to be reduced. This is the classic method . we resort to the Bayesian way , where the parameters are random predictors that have a finite distribution
(Turki, 2019). Assuming that these parameters have preliminary information that can be put in the form of a probability distribution (Abboudi, 1996). Suggest (Akaike (1974)) the information criterion. It is one of the most common criteria used to select the model that gives the most accurate description of the data displayed Nishii (1984). can be written asHoff, (2009)). But there is difficulty in realizing these conditions when there is an increase in the number of variants. A problem appears with estimation, and this needs to be reduced. This is the classic method . we resort the Bayesian way , where the parameters are random predictors that have a finite distribution.
(Turki, 2019). Assuming that these parameters have preliminary information that can be put in the form of a probability distribution (Abboudi, 1996). Suggest (Akaike (1974)) the information criterion ($AIC$ ). It is one of the most common

criteria used to select the model that gives the most accurate description of the data displayed. Nishii (1984). can be written as:

$$AIC = -2 \log m + 2p, \qquad (2)$$

where $(m)$ is the probability function calculated using a Markov chain estimator $(AIC)$, and $(p)$ is the total number of model parameters (Mallick, 2015). But of defects $(AIC)$ It is not suitable when the –value n is high (Javed and Mantalos, 2013; Bozdogan, 1987). To get rid of this problem, the researcher tends to use another criterion, such as the Bayesian information criterion($BIC$ ) (Schwarz, 1978).

$$BIC = -2 \log m + plog\, n, \qquad (3)$$

where $(n)$ is the sample size that was suggested to address the $AIC$ issue in the , and the model chosen will be consistent in choosing the right model with a probability of 1 (Javed and Mantalos, 2013; Mallick, 2015). It is known that $AIC\ and\ BIC$ neither works better all the time Spiegelhalter et al. (2002) (George and McCulloch, 1993). Suggested generalizations $AIC$ and $BIC$ For model selection in normal linear hierarchical Bayesian models, the skewness information criterion is used( $DIC$ ) Ando (2007). Dubai Internet City Choosing the best-equipped models despite providing very little information about them.

One of the important models of regression is the ordinal model. It is a type of statistical data in which the common variants are in the form of ordered categories that can be arranged naturally (Zhou, 2006). And it has wide applications in psychology, climate, economics, political economy, social sciences, medicine, and many other sciences (AL-JABRI, 2020).
For example, Students' intelligence level (weak, average, high) in the low, medium, and high categories take the ranking values. Respectively. The high level is not a multiple of the average level. One of the problems with ordinal Regression ordinal is when the number of variants is large (Rahman, 2016). So it must be reduced. While a comparison of the models using the information standard shows the deviation ordinal models can provide a better-fit model compared to the ordinal probability model in classical methods( Rahman (2016)).

When the number of variants in the ordinal is large, we use many methods to reduce it. Most of these methods are organizational methods. Examples of these methods are the bridge penalty (Frank and Friedman 1993). the least absolute shrinkage and selection operator (Lasso) penalty (Tibshirani 1996). The adaptive lasso penalty (Zou 2006). The elastic net penalty (Zou and Hastie 2005). The adaptive elastic net penalty (Zou and Zhang 2009). The smoothly clipped absolute deviation (SCAD) penalty (Fan and Li 2001). The group bridge penalty (Huang et al. 2009). And the adaptive bridge penalty (Park and Yoon 2011).
Finally, in this research, we use a new method, which is the ordinal regression model with penalty functions, such as: (rBridge).

The paper is organized as follows. In Section. 2, we introduce the reciprocal Bayesian bridge regression parts, a regression algorithm for data subject to ordinal models based on the reciprocal Bayesian bridge regression Before th inverse uniform distribution section.3, carried out by means of simulations section.4, via real data example in section.5, We conclude briefly discussing in section.6,

## 2.Mothed:

### 2.1. Reciprocal bridge regression:

Consider reciprocal bridge (r bridge) (Alhamzawi and Mallick 2020;zainab2022). Results from the following organizational problem:
Let's say $\alpha = \frac{1}{2}$

$$\min_\beta (L - X\beta)^t\, (L - X\beta) + \lambda \sum_{j=1}^{p} \frac{1}{|\beta_j|^{\frac{1}{2}}} \mathrm{I}\{\beta_j \neq 0\}, \qquad (4)$$

where $I(.)$ Indicates the function of the pointer $\lambda > 0$ It is the setting parameter that controls the degree of punishment.

### 2.2.Bayesian reciprocal bridge regression for ordinal data

The statistical model for ordinal regression is written as follows:

$$\mathrm{z_i} = \mathrm{X_i^t}\boldsymbol{\beta} + \varepsilon_\mathrm{i}, \qquad \forall i = 1 \dots, n \qquad (5)$$

where $\mathrm{X_i}$ , is $n \times p$ vector of covariates and a vector $\boldsymbol{\beta}$ is $p \times 1$ of unknown parameters . Containing categories or results $\delta$ through the cutoff point vector $\delta$ as follows:

$$\delta_{c-1} < \mathrm{z_i} < \delta_c \rightarrow \mathrm{y_i} = c \qquad \forall i = 1, \dots, n; c = 1, \dots, C \qquad (6)$$

where

$$y_i = \begin{cases} 1\ if\ \delta_0 \ < z_i \ \le \ \delta_1 ; \\ c\ if\ \delta_{c-1} < \ z_i \le \ \delta_c, \qquad c = 2, \dots , C-1 \\ C\ if\ \delta_{C-1} \ < z_i \le \delta_C; \end{cases}$$

where $\delta_0 = -\infty$ and $\delta_c = \infty$ (see Jeliazkov et al., 2008). Following Alhamzawi(2016). The $CDF$ for the $c$ category of the observed response $y_i$ is:

$$p(y_i \le c/z_i , \delta_c) = p(z_i \le \ \delta_c/\beta)$$
$$= p(X_i^t \beta \ \le \ \delta_c)$$
$$= p(\epsilon_i \le \delta_c - X_i^t \beta),$$
$$= \phi(\delta_c - X_i^t \beta \,),$$

where $\phi$ is the standard normal $CDF$. using $\phi$ , we can calculate $p(y_i = c/z_i , \delta_{c-1}, \delta_c)$ as follows :

$$p(y_i = c/z_i , \delta_{c-1}, \delta_c) = p(\delta_{c-1} < z_i \le \delta_c/\beta)$$
$$= \phi(\delta_c - X_i^t \beta) - \phi(\delta_{c-1} - X_i^t \beta).$$

We take into account the following issue with ordinal data:

$$\min_\beta (z - X\beta)^t \, (z - X\beta) + \lambda \sum_{j=1}^{p} \frac{1}{|\beta_j|^{\frac{1}{2}}} I\{\beta_j \ne 0\}, \qquad (7)$$

where $I(.)$Indicates the function . In this research, instead of minimizing problem (7), we solved it by adopting a Bayesian hierarchical model and sampling the regression coefficients using Gibbs' sample, in contrast to the iterative approach to solving (7). which has not yet been proposed .
the equation (7) results in VS if $(0 < \alpha \le 1)$ (Park and Yoon, 2011). The optimal subset selection, the Lasso regression, and the ridge regression, respectively, when $\alpha = 0, \alpha = 1,$ and $\alpha = 2,$ are three specific examples of the aforementioned equation (Mallick and Yi, 2017). Bayesian statistical inference rBridge regression it can be a tuning parameter λ An easy estimator as an automatic by-product of the Markov chain Monte Carlo .
The r bridge penalty in (7) includes Many methods, such as the best selection of a subset,$(\alpha = 0)$ (Song and Liang 2015). And his penalty $(\alpha = 1)$ and reciprocal Ridge (rRidge) penalty $(\alpha = 2)$

**3.priors**

Note the form of in reciprocal bridge (6) Mallick, Alhamzawi, and Svetnik (2020) and Alhamzawi and Mallick (2020). They show that estimates can be interpreted as reciprocal bridge According to ex-post-mode estimates when the regression coefficients are independent and identical to the inverse ex distributions Gaussian (IGG) prior distributions of the form:

$$\pi(\beta) = \frac{\lambda^{1/\frac{1}{2}}}{2\beta^2 \Gamma(\frac{1}{1/2}+1)} \exp\left\{-\frac{\lambda}{|\beta_j|^{\frac{1}{2}}}\right\} I\{\beta_j \ne 0\},$$

$$\pi(\beta) = \frac{\lambda^2}{2\beta^2 \Gamma(3)} \exp\left\{-\frac{\lambda}{|\beta_j|^{\frac{1}{2}}}\right\} \, I\{\beta \ne 0\} \qquad (8)$$

where $\alpha > 0$ is a shape parameter and $\lambda > 0$ is a scale parameter. Gypsum samples can be sampled from this posterior using an extended hierarchy with unified inverse primes on the coefficients and independent gamma rates on their mixing coefficients (Mallick, Alhamzawi, and Svetnik 2020).

$$\frac{\lambda^2}{2\beta^2 \Gamma(3)} \, e^{-\lambda/|\beta|^{-\frac{1}{2}}} = \frac{\lambda^{\frac{1}{2}}}{2\beta^2 \Gamma(3)} \int_{u > |\beta|^{\frac{1}{2}}} \lambda \, e^{-\lambda u} \; du \qquad (9)$$

we assign the following prior distribution for (δ) :

$$p(\delta) = (c-1)! \, (\frac{1}{\delta_{max} - \delta_{min}})^{c-1} I(\delta \in T),$$

where $\delta = (\delta_0, \delta_1, \dots , \delta_c)$ and $T = (\delta_{min}, \delta_1, \dots , \delta_{max}/\delta_{min} < \ \delta_1 < \cdots < \delta_{c-1} < \delta_{max})$.
we assign the following prior distribution for $(u)$ :

$$p(u) = \frac{\lambda^3}{\Gamma(3)} x^2 e^{-x/\lambda},$$

where $\Gamma$ $it\ is\ gamma\ function$ , $and\ \lambda, \left(\frac{1}{\frac{1}{2}}+1\right)$ $positive\ value.$

For $\lambda$ ,we assign a Jeffery prior of the for $\lambda \sim \frac{1}{\lambda}$ , which is a special case of gamma distribution when $a=0, b=0$

## 3.Bayesian heretical Modeling

To proceed with a Bayesian inference, in this section, we use the same prior specifications for $\lambda$ the Non-informational preset fixed marginal scale on $\lambda$ and you have the following Bayesian hierarchical model:

$y_i = c$ if $\delta_{c-1} < z_i \leq \delta_c$

$$z_i \backslash \beta \sim N(X_i^t \beta, 1),$$

$\beta \backslash u \sim \prod_{j=1}^{p} \frac{1}{Uniform(-u_j^2, u_j^2)}$,

$u \backslash \delta \sim \prod_{j=1}^{p} Gamma(2+1, \lambda)$,

$\delta \backslash \lambda \sim (c-1)! \left(\frac{1}{\delta_{max}-\delta_{min}}\right)^{c-1} I(\delta \in T)$,

$\lambda \sim \frac{1}{\lambda}$

the parameters of interest $(\beta, u, \lambda, \delta)$ can be sampled as listed in Algorithm

## 4.MCMC sampling

Algorithm : for the Bayesian reciprocal bridge for Ordinal Model

Using the data augmentation approach as which is described in( Albert and Chib (1993)), Alhamzawi (2016).An Gibbs sampling Method for the ordinal model is designed by updating $\beta, \delta, \lambda, u,$ , from their full conditional distributions .

**Input**: $(\mathbf{y}, \mathbf{x})$

**Initialize** : $(\boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\delta}, \mathbf{z})$

$for\ t = 1, \dots, (t_{max} + t_{burn-in}) do$

1-Sample $\boldsymbol{\beta} \backslash\ . \sim N_p\left(\hat{\beta}_{MLE}, \sigma^2 (X^t X)^{-1}\right) \prod_{j=1}^{p} I\{|\beta_j| > \frac{1}{u_j^2}\}$

2-Sample $\mathbf{u}\ \backslash\ . \sim \prod_{j=1}^{p} Exponetial\ (\lambda) I\{u_j > \frac{1}{|\beta_j|^{\frac{1}{2}}}\}$

3-Sample $\boldsymbol{\lambda} \backslash\ . \sim Gamma\ (a=0, b=0), \sim \frac{1}{\lambda}$

4-Sample $\boldsymbol{\delta_c}$ a uniform distribution on the interval $[\min\{\min(z_i \backslash y_i = c+1), \delta_{c+1}, \delta_{max}\}, \max\{\max(z_i \backslash y_i = c), \delta_{c-1}, \delta_{min}\}]$.

## 5.Simulation studies

Empirical simulation results are presented here to demonstrate the performance of the proposed method. In this section, three simulation studies were used to demonstrate the performance of the proposed method in ordinal models, referred to as "BOrBridge". The performance of our proposed methods BOrBridge is compared with the Akaike's information criterion (AIC; Akaike, 1998) and the Bayesian in- formation criterion (BIC; Schwarz, 1978). We run our proposed Gibbs sampler for 14,000 iterations, after a burn-in period of 2000 iterations. Convergence of the proposed Gibbs sampler was conducted using the multivariate potential scale reduction factor (MPSRF) (Brooks and Gelman, 1998) which is given by (Alhamzawi, 2016):

$$MPSRF = \frac{q-1}{q} + \left(\frac{m+1}{m}\right)\alpha,$$

where $\alpha$ is the largest eigenvalue of the matrix $\Omega^{-1}\Upsilon/q, \varphi$ denote a parameter vector of interest, $\varphi_{st}$denote the $tth$ of the $q$ iterations of $\varphi$ in chain s , for $s = 1, \dots, m$ $(m \geq 2)$ and

$$\Upsilon/q = \frac{1}{m-1}\sum_{s=1}^{m}(\overline{\varphi_s} - \overline{\varphi_{\cdot\cdot}})^2,$$

$$\Omega = \frac{1}{m(q-1)} \sum_{s=1}^{m} \sum_{t=1}^{q} (\varphi_{st} - \overline{\varphi_s})^2,$$

with $\overline{\varphi_s} = \frac{1}{q} \sum_{t=1}^{q} \varphi_{st}$ and $\bar{\varphi} = \frac{1}{mq} \sum_{t=1}^{q} \varphi_{st}$.

### 5.1. Simulation study 1

We generate data from the true model,

$$w_i = \mathbf{x}_i' \, \boldsymbol{\beta} + \varepsilon_i, \qquad \varepsilon_i \sim N(0,1) \;\; i = 1, \ldots, 200 \qquad (10)$$

where $x_i = (1, x_{1i}, x_{2i}, x_{3i}, x_{4i}, x_{5i}, x_{6i}, x_{6i}, x_{7i}, x_{8i}, x_{9i}, x_{10,})'$ and $\beta = (1,1,0,0,0,0,0,0,0,0,0)'$, including the intercept term. This simulation study is corresponding to very sparse case. The predictors are generated independently from $N_{10}(0, \boldsymbol{\Sigma}_x)$with $(\Sigma_x)_{fm}$= $0.75 \; |f - m|$, where $fm$ refers to the $(f, m)^{th}$ entry of

**Table 1:** Number of observations corresponding to the categories of $y$ in Simulation 1

| Simulation 1 | y=1 | y=2 | y=3 | y=4 | y=5 |
|---|---|---|---|---|---|
| | 32 | 68 | 68 | 27 | 5 |

**Table 2:** Comparing average numbers of correct and wrong zeros for the best model in Simulation study 1, averaged over 100 simulations. In the parentheses are standard deviations. Our proposed method BOrBridge is compared with AIC and BIC.

| | Methods | | |
|---|---|---|---|
| | BOrBridge | AIC | BIC |
| correct | 8.42 (0.23) | 6.23 (0.59) | 6.48 (0.59) |
| wrong | 0.00 (0.00) | 0.17 (0.29) | 0.17 (0.13) |

the matrix $\boldsymbol{\Sigma}_x$.

The outcome of interest $\boldsymbol{y}$ were calculated based on the cut-point vector $\boldsymbol{\delta} = (0, 1, 2, 3)'$, yielding five categories which are listed in Table 1. In this simulation study, AIC and BIC are used to select the best model in the standard ordinal regression model. We summarized the results of BOrBridge, AIC and BIC in Table 2. The results show that the proposed methods BOrBridge performs very well compared with AIC and BIC in terms of selecting the correct model. BOrBridge gives more correct zero coefficients than AIC and BIC. Figure (1) shows that the MPSRF for the proposed method BOrBridge becomes stable and close to 1 after about 2000 iterations. This shows that the convergence of our Gibbs sampler to the posterior distribution was quick and the mixing was good.

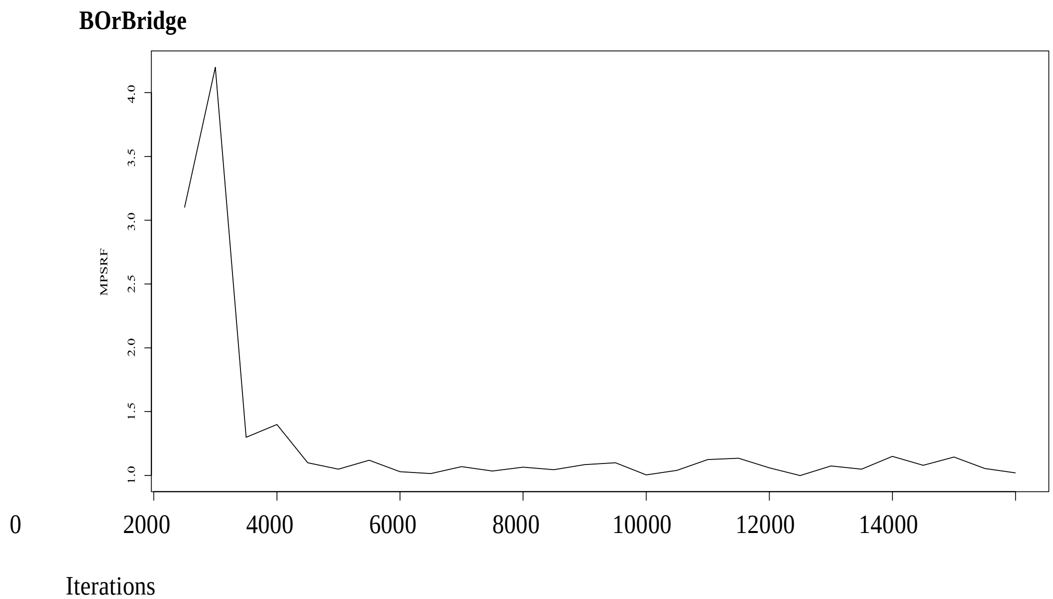


**Figure 1:** MPSRF for the proposed method BOrBridge in Simulation 1.

### 5.2. Simulation study 2

This simulation study is the same as Simulation 1, except that we set $\boldsymbol{\beta}$ =(1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0)′, including the intercept term. This simulation study is corresponding to sparse case. Similar to Simulation 1, the outcome of interest $\boldsymbol{y}$ were calculated based on the cut-point vector $\boldsymbol{\delta}$ = (0, 1, 2, 3)′, yielding five categories which are listed in Table 5.

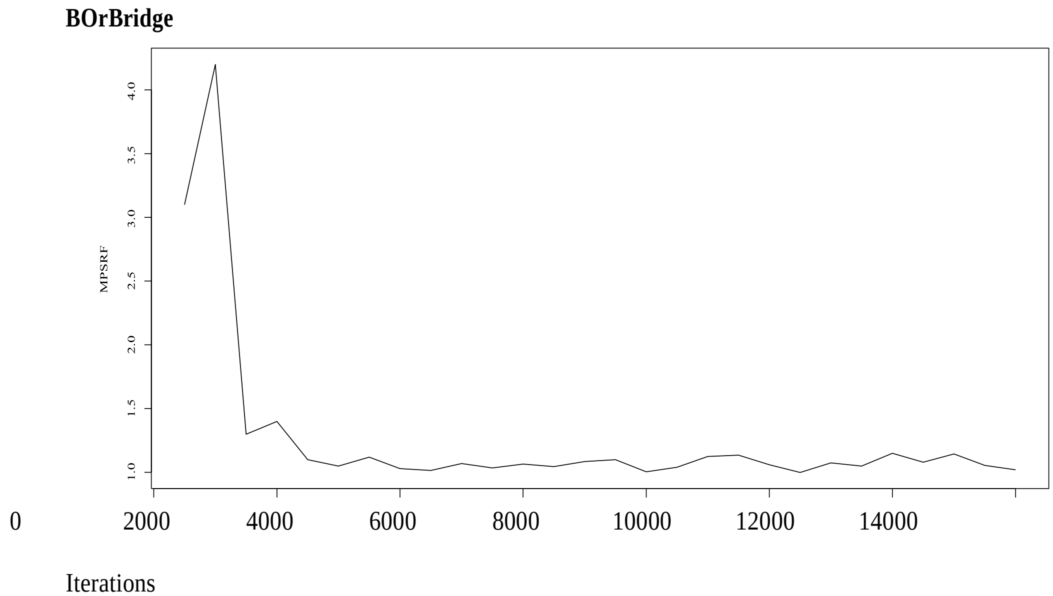


**Figure 1:** MPSRF for the proposed method BOrBridge in Simulation 1.

### 5.3. Simulation study 2

This simulation study is the same as Simulation 1, except that we set $\boldsymbol{\beta}$ =(1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0)′, including the intercept term. This simulation study is corresponding to sparse case. Similar to Simulation 1, the outcome of interest $\boldsymbol{y}$ were calculated based on the cut-point vector $\boldsymbol{\delta}$ = (0, 1, 2, 3)′, yielding five categories which are listed in Table 5.

**Table 3:** Number of observations corresponding to the categories of *y* in Simulation 2

| Simulation 2 | y=1 | y=2 | y=3 | y=4 | y=5 |
|---|---|---|---|---|---|
| 78 | 22 | 22 | 21 | 57 | |

**Table 4:** Comparing average numbers of correct and wrong zeros for the best model in Simulation study 2, averaged over 100 simulations. In the parentheses are standard deviations. Our proposed method BOrBridge is compared with AIC and BIC.

| | Methods | | |
|---|---|---|---|
| | BOrBridge | AIC | BIC |
| correct | 4.33 (0.19) | 3.78 (0.47) | 3.90 (0.72) |
| wrong | 0.00 (0.00) | 0.47 (0.38) | 0.34 (0.32) |

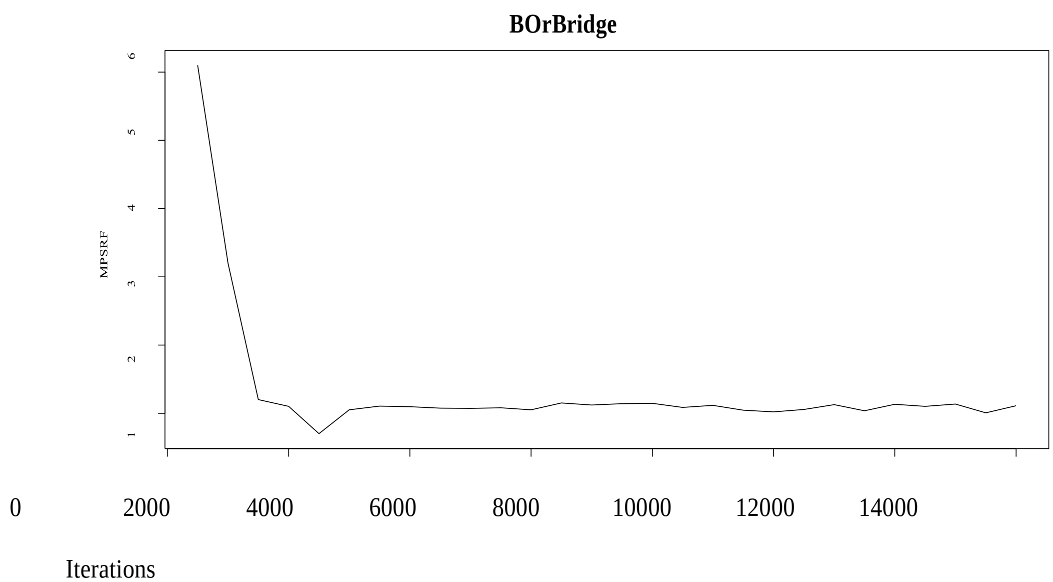


**Figure 2:** MPSRF for the proposed method BOrBridge in Simulation 2.

The results of BOrBridge, AIC and BIC were listed in Table

4 which shows thatthe proposed method BOrBridge perform very well compared with the other methods in the comparison. This table shows that our method gives more correct zero coefficients than AIC and BIC. Figure (2) shows that the MPSRF for the proposed method BOrBridge becomes stable and close to 1 after about 2000 iterations.

**Table 5:** Number of observations corresponding to the categories of $y$ in Simulation 2

| Simulation 2 | y=1 | y=2 | y=3 | y=4 | y=5 |
|---|---|---|---|---|---|
| | 24 | 19 | 26 | 62 | 69 |

**Table 6:** Comparing average numbers of correct and wrong zeros for the best model in Simulation study 3, averaged over 100 simulations. In the parentheses are standard deviations. Our proposed method BOrBridge is compared with AIC and BIC.

| | Methods | | |
|---|---|---|---|
| | BOrBridge | AIC | BIC |
| correct | 6.19 (0.24) | 4.56 (0.71) | 3.17 (0.64) |
| wrong | 0.00 (0.00) | 0.37 (0.26) | 1.42 (0.29) |

### 5.3.Simulation 3

We consider the correct model

$$w_i = x_i'\beta + (1 + x_{2i})\varepsilon_i, \quad \varepsilon_i \sim N(0,1) \quad i = 1, \dots, 200 \qquad (11)$$

where
$x_i = (1, x_{1i}, x_{2i}, x_{3i}, x_{4i}, x_{5i}, x_{6i}, x_{7i}, x_{8i}, x_{9i}, x_{10i})'$ *and* $\boldsymbol{\beta} = (1,1,1,1,0,0,0,0,0,0,0)'$. *Let* $x_{1i}, u_i \sim N(0,1)$, $x_{2i} \sim Uniform\,[0,1]$, $x_{3i} = x_{1i} + x_{2i} + u_i$ and $x_{4i}, x_{5i}, x_{6i}, x_{7i}, x_{8i}, x_{9i}, x_{10i}$ are mutually independent from $N(0,1)$. The outcome of interest **y** were obtained based on the cut-point vector $\boldsymbol{\delta} = (-1,0,1,3)'$, yielding five categories.

The results of this simulation study were summarized in Table 4 which shows that the proposed method BOrBridge perform very well compared with the other method in the comparison. Again, the results show that the proposed method produces more correct zero coefficients than AIC and BIC. Figure (3) shows that the MPSRF for the proposed method BOrBridge becomes stable and close to 1 after about 2000 iterations.

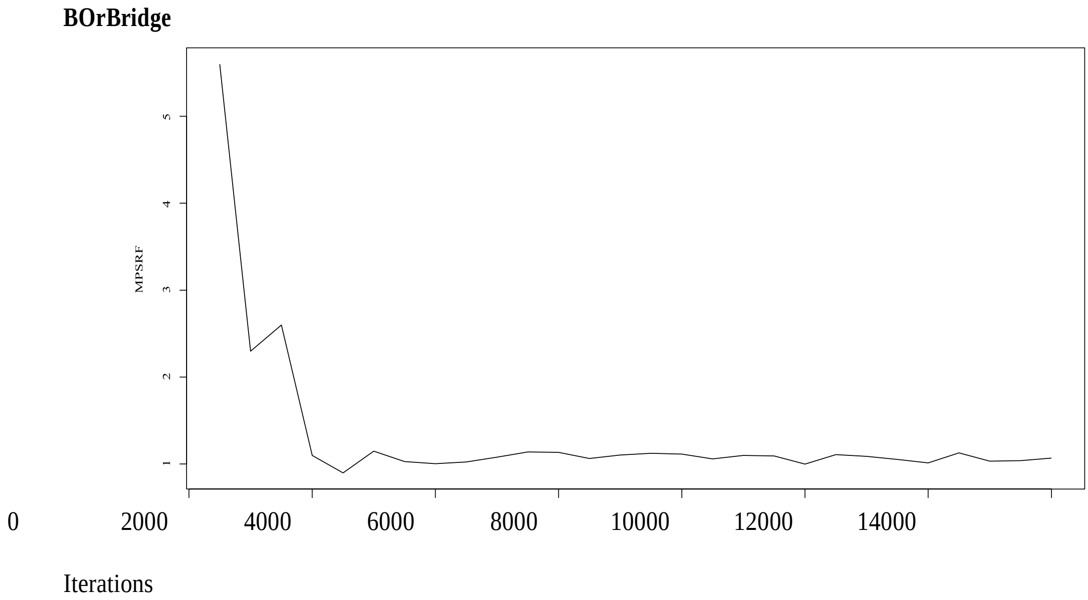


**Figure 3:** MPSRF for the proposed method BOrBridge in Simulation 3

### 5.4. Real Data Anal ysis

#### 5.4.1. Tax Policy

In this section, we compare the performance of the proposed method, BOrBridge on the tax policy data. This data set was analyzed by Rahman (2016) in a study on a subject of extended political debate with respect to the definition of benchmark income, beneficiaries of the tax cuts, and whether it would spur sufficient growth. The dependent variable has three categories. The data consist of 7 covariates: indicator for individual being employed (employed), indicator for household income > $75, 000 (income), individual's highest degree is Bachelors (bachelors), highest degree is Masters (bachelors), Professional or Doctorate (post bachelors), Individual or household owns a computer (computers), individual or household owns a cell phone (cellphone) and Race of the individual is white (white). The results of parameter estimation are summarized in Table 1, which shows that both approaches have very similar results. Again, Figure (1) shows that the MPSRF for the proposed method BOrBridge become stable and close to 1 after about 2000 iterations.

Table 1: Posterior mean (mean) and standard deviation (std) of model parameters for the tax policy application .

| | Method |
|---|---|
| | BOrBridge |
| Intercept | 2.21 (0.39) |
| Employed | 0.17 (0.21) |
| Income | -0.39 (0.27) |
| Bachelors | 0.05 (0.29) |
| Post-bachelors | 0.43 (0.45) |
| Computers | 0.59 (0.37) |
| Cellphone | 0.82 (0.35) |
| White | 0.03 (0.41) |

Hence, both the simulation studies and real data example support the proposedmethod.

**BOrBridge**

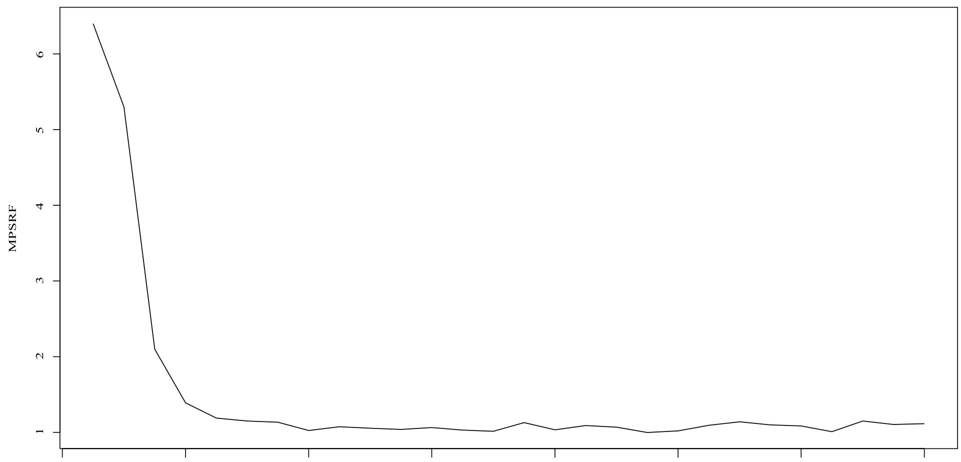


**Figure 1:** MPSRF for the proposed method BOrBridge in Tax Policydata.

### 5.5.Conclusions and Future Research

The research proposes a Bayesian reciprocal bridge estimation method for univariate ordinal models. Specifically, Clear advantages over existing approaches include efficient Gibbs sampler and use of data augmentation to allow ordinal outcome of interest. The main contributions and future research topics are listed below.

#### 5.5.1.Main Contributions

Research introduces the Bayesian inference of regression models for univariate ordinal data, and proposes method that can be extensively utilized in a wide class of applications across disciplines including medicine, biological studies and social sciences.

Bayesian reciprocal Bridge regression for ordinal data. We develop a Gibbs sampler methods for sampling from this posterior distributions using the latent variable inferential framework of Albert and Chib (1993).

The proposed method are applied to three simulation studies and a real data example and compare the results with AIC and BIC. It is found that the performance of the proposed method is better than AIC and BIC.

#### 5.5.2. Recommendations for Future Research

The work proposed in this research opens the door to new research directions for Bayesian regularization in ordinal models. One of these directions, the proposed methods can be extended to ordinal quantile regression models by assuming the error distribution follows the asymmetric Laplace distribution and using the normal exponential mixture representation of the asymmetric Laplace distribution.

**References**

Al-Jabri, D. H. Q., & Al-Hamzawi, R. J. (2020). Bayesian Adaptive Bridge Regression for Ordinal Models with an Application. Iraqi Journal of Science, 170-178

Alhamzawi, R. (2016). Bayesian model selection in ordinal quantile regression. Computational Statistics & Data Analysis, 103, 68-78.

Albert, J. H. and Chib, S. (1993). “Bayesian analysis of binary and polychotomous response data.” Journal of the American statistical Association, 88(422): 669–679

Akaike, H. (1974). A new look at the statistical model identification. IEEE transactions on automatic control, 19(6), 716-723

Ando, T. (2007). Bayesian predictive information criterion for the evaluation of hierarchical Bayesian and empirical Bayes models. Biometrika, 94(2), 443-458Ando, T. (2007). Bayesian predictive information criterion for the evaluation of hierarchical Bayesian and empirical Bayes models. Biometrika, 94(2), 443-458

Alsaadi, Z., & Alhamzawi, R. (2022). Bayesian bridge and reciprocal bridge composite quantile regression. *Communications in Statistics-Simulation and Computation*, 1-18.

Brooks, S. P. and A. Gelman (1998). General methods for monitoring convergence of iterative simulations. *Journal of computational and graphical statistics 7* (4), 434–455.

Berk, R, "Regression Analysis: A constructive critique", Sage Publications, 2004, 44-58

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. Journal of the American statistical Association, 96(456), 1348-1360.

Frank, L. E., & Friedman, J. H. (1993). A statistical view of some chemometrics regression tools. Technometrics, 35(2), 109-135.

Huang, J., Ma, S., Xie, H., & Zhang, C. H. (2009). A group bridge approach for variable selection. Biometrika, 96(2), 339-355.

Hoff, P. D. (2009). A first course in Bayesian statistical methods (Vol. 580). New York: Springer.

Javed, F., & Mantalos, P. (2013). GARCH-type models and performance of information criteria. Communications in StatisticsSimulation and Computation, 42(8), 1917-1933

Koenker, R., & Bassett Jr, G. (1982). Robust tests for heteroscedasticity based on regression quantiles. Econometrical: Journal of the Econometric Society, 43-61

Kutner, M, Nachtsheim, C, Neter, J & Li, W "Applied Linear Statistical Models”, Fifth Edition, McGraw Hill, Boston: United State of American, 2005

Mallick, H., & Yi, N. (2017). Bayesian group bridge for bi-level variable selection. Computational statistics & data analysis, 110, 115- 133

Mallick, H., R. Alhamzawi, and V. Svetnik. 2020. The reciprocal Bayesian lasso. arXiv:2001.08327

Nishii, R. (1984). Asymptotic properties of criteria for selection of variables in multiple regression. The Annals of Statistics, 758-765.

Park, C., & Yoon, Y. J. (2011). Bridge regression: adaptivity and group selection. Journal of Statistical Planning and Inference, 141(11), 3506-3519

Song, Q., & Liang, F. (2015). High-dimensional variable selection with reciprocal l 1-regularization. Journal of the American Statistical Association, 110(512), 1607-1620

.Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B 58:267–88.

Zou, H. (2006). The adaptive lasso and its oracle properties. Journal of the American statistical association 101 (476), 1418–1429 .

Zou, H., & Zhang, H. H. (2009). On the adaptive elastic-net with a diverging number of parameters. Annals of statistics, 37(4), 1733