Pak. J. Statist.
2025 Vol. 41(3), 339-350

BAYESIAN RECIPROCAL GROUP LASSO COMPOSITE QUANTILE REGRESSION

Saif Hosam Raheem[§], Fadel Hamid Hadi Alhusseini
and **Taha Alshaybawee**
College of Administration and Economics,
University of Al-Qadisiyah, Iraq
Email: [1] saif.hosam@qu.edu.iq
[2] fadel.alhsiny@qu.edu.iq
[3] taha.alshaybawee@qu.edu.iq
[§] Corresponding author

ABSTRACT

This study introduces the Bayesian Reciprocal Group Lasso Composite Quantile Regression (BRGLCQR) model, designed to improve variable selection and predictive accuracy in quantile regression with grouped predictors. Traditional methods for group selection often struggle with handling correlations among predictors, resulting in weak or biased estimates in high-dimensional contexts. By applying a reciprocal group lasso penalty within a Bayesian hierarchical framework, the BRGLCQR model achieves enhanced group-level variable selection and more precise predictions. The model's efficacy is demonstrated through Gibbs sampling, which enables efficient estimation and selection of predictor groups. Comparative analysis using both simulation studies and real-world stock market returns and volatility data reveals that BRGLCQR consistently outperforms existing Bayesian and non-Bayesian methods, achieving lower mean squared prediction errors and higher model selection accuracy MSA. These results highlight BRGLCQR's robust performance, especially in high-dimensional settings with grouped data, positioning it as a valuable tool for complex financial and economic modeling tasks.

KEYWORDS

Bayesian Reciprocal Group Lasso, Composite Quantile Regression, Gibbs Sampler, Grouped Variable Selection, High-Dimensional Data.

1. INTRODUCTION

Quantile regression (QReg), introduced by Koenker and Bassett (1978), has become widely used by statisticians, econometricians, and applied researchers for its flexibility in modeling relationships at different points of a response variable's distribution. QReg has found applications across various fields, such as economics, agriculture, medicine, genetics, and sociology (Alhamzawi, 2013; Koenker, 2005; Yu et al., 2003). Unlike standard regression, QReg captures the effect of predictors across multiple quantiles, making it particularly suitable for analyzing heterogeneous data without assumptions about data distribution (Orsini & Bottai, 2011). Additionally, QReg estimators offer robustness

to outliers, an essential feature when modeling complex response behaviors (Koenker, 2005).

For any quantile τ where $(0 < \tau < 1)$, QReg models the conditional quantile function of the response variable. Koenker and Bassett (1978) formally define QReg through the optimization:

$$\min_{\beta} \sum_{i=1}^{n} \rho_{\tau} (y_i - X_i \beta) \tag{1}$$

where ρ_{τ} represents the quantile loss function, y is the response variable, X is a K-dimensional predictor vector, and β is the coefficient vector. Due to the non-differentiability of the quantile loss function at zero, solutions are obtained using linear programming techniques (Koenker & D'Orey, 1987).

Bayesian methods have been frequently employed to estimate QReg parameters, especially when error distributions follow an asymmetric Laplace distribution (ALD) (Yu & Moyeed, 2001). The Bayesian QReg approach provides reliable estimates, even for smaller sample sizes, by incorporating prior information a valuable asset when analyzing complex structures within high-dimensional data (Koenker, 2005; Alhamzawi, 2013).

Variable selection in high-dimensional QR models is crucial, as selecting the appropriate groups of predictors can improve model interpretability and prediction accuracy. Koenker (2004) introduced regularization in QReg for managing random effects, and the Bayesian Lasso has further extended this concept by applying Laplace priors for sparsity (Wang et al., 2007). However, selecting predictors becomes more complex in cases where predictors are organized into natural groups, as typical variable selection approaches may not fully capture group structures.

In this study, we propose a Bayesian Reciprocal Group Lasso Composite Quantile Regression model (BRGLCQR), which enhances group-level variable selection while allowing smaller, non-zero effects within groups. This model builds upon the principles of composite quantile regression (CQR) by Zou and Yuan (2008), which improves estimation efficiency by simultaneously modeling multiple quantiles and offers greater robustness compared to single quantile approaches. The CQR framework allows for flexibility and efficiency in estimation, making it suitable for datasets where predictors exhibit natural grouping.

Our Bayesian Reciprocal Group Lasso model integrates CQR with reciprocal penalties within a Bayesian hierarchical framework, providing an effective solution for grouped variable selection. Through simulation studies and real-world applications, we demonstrate the proposed model's effectiveness in achieving accurate variable selection and prediction compared to standard methods. While the Bayesian Reciprocal Lasso Composite Quantile Regression (BRLCQR) focuses on individual-level variable selection using reciprocal shrinkage, the proposed BRGLCQR model extends this idea by introducing group-wise shrinkage within a Bayesian framework. This enhancement allows the model to select or discard entire groups of predictors, which is especially useful when variables are naturally clustered or correlated offering improved interpretability and

prediction in grouped high-dimensional settings. The following sections present the model's methodological framework, parameter estimation through MCMC, and a comprehensive assessment using simulated and empirical data.

2. METHODOLOGY

2.1 Bayesian Composite Quantile Regression Model (BCQReg)

Consider a Bayesian framework for composite quantile regression (CQR), which enables simultaneous estimation across multiple quantiles, making it a robust choice for modeling complex data distributions. This method optimizes quantile estimates by incorporating a Bayesian approach, where prior distributions are assigned to model parameters, and posterior distributions are derived using observed data.

The model is formulated as follows:

$$y_i = X_i\beta + \epsilon_i, i = 1,2,\dots,n$$

where y_i is the response variable, X_i is the vector of predictors, β is the vector of parameters to be estimated, and ϵ_i represents the error term, assumed to follow an asymmetric Laplace distribution (ALD), which aligns with the nature of quantile regression.

The asymmetric Laplace distribution (ALD) is key in Bayesian QReg, as it allows the likelihood to focus on specific quantiles. Its probability density function is given by:

$$p(\epsilon|\mu,\sigma,\tau) = \frac{\tau(1-\tau)}{\sigma}\exp\left(-\rho_\tau\left(\frac{\epsilon-\mu}{\sigma}\right)\right) \quad (2)$$

where $\rho_\tau(.)$ is the check function, μ is the location parameter, σ is the scale parameter, and τ represents the quantile level. The check function ρ_τ is used to compute residuals for different quantiles, which are then minimized to estimate the coefficients.

To estimate the parameters β in a composite quantile regression model, we solve:

$$\hat{\beta} = \arg\min_{\beta}\sum_{k=1}^{K}\sum_{i=1}^{n}\rho_{\tau k}\left(y_i - X_i\beta\right) \quad (3)$$

where K represents the number of quantiles considered, and $\rho_{\tau k}$ is the check function at each quantile τk. This approach allows for simultaneous estimation across multiple quantiles, making it robust to outliers and capable of capturing diverse predictor-response relationships.

In a Bayesian setting, the priors for β are incorporated into the estimation, leading to posterior distributions for each parameter. These are typically obtained via Markov Chain Monte Carlo (MCMC) methods, such as Gibbs sampling, where conditional distributions are sampled iteratively to approximate the joint posterior distribution of β.

The use of a normal-exponential mixture further simplifies computation by leveraging the properties of the normal distribution, allowing for efficient estimation of quantiles through simulation methods. By combining CQR with a Bayesian framework, the model achieves a balance between robustness to outliers and computational efficiency, making it suitable for high-dimensional data with potentially heterogeneous structures.

2.2 Bayesian Reciprocal Lasso Composite Quantile Regression (BRLCQR)

The Bayesian Reciprocal Lasso estimator for composite quantile regression can be represented by modifying the quantile regression objective function. Following a similar approach as in Alhamzawi and Mallick (2020), the BRLCQR model solves the following optimization problem:

$$\underbrace{argmin}_{b_\tau,\beta} \sum_{i=1}^{n} \left\{ \sum_{k=1}^{K} \rho_{\tau_k}\left(z_i - b_{\tau_k} - x_i'\beta\right) \right\} + \lambda \sum_{g=1}^{G} \frac{1}{|\beta_g|} I\{\beta_g \neq 0\} \tag{4}$$

where λ is the regularization parameter, corresponding specifically to the reciprocal Lasso penalty. The check function is used instead of the loss function, maintaining the identity of the reciprocal Lasso. This approach addresses the problem of miniaturization, particularly in small-sample settings. The penalty term in equation (4) provides posterior mode estimates when the regression parameters have an Inverse Generalized Gaussian (IGG) distribution (Mallick et al., 2020):

$$\pi(\beta) = \prod_{g=1}^{G} \frac{\lambda}{2|\beta_g|\Gamma(2)} exp\left\{-\frac{\lambda}{|\beta_g|}\right\} I\{\beta_g \neq 0\} \tag{5}$$

To implement this model using Gibbs sampling, we leverage a scale mixture of normal (SMN) approach, as shown by Armagan, Dunson, and Lee (2013), and extended by Mallick, Alhamzawi, and Svetnik (2020). Assuming $\beta \sim N(0, l)I(|\beta| > \eta)$, $l \sim Exp\left(\xi^2/2\right)$, and $\xi \sim Exp(\eta)$, an inverse double exponential distribution for β emerges with a scale parameter $\lambda > 0$ when η follows an Inverse Gamma (2, λ).

Where $u = \frac{1}{\eta}$, $l = (l_1, \ldots, l_G)^T$, and $\xi = (\xi_1, \ldots, \xi_G)^T$. The Bayesian hierarchical formulation includes prior distributions for each parameter, as specified below.

Our Bayesian hierarchical model assumes the following priors for all parameters and latent variables:

$$z_i|x \sim N_n\left(z_i + b_{\tau_k} + x_i'\beta + \vartheta v, 2\sigma v\right),$$

$$\beta|l \sim \prod_{g=1}^{G} N(0, t^2) I\left\{\frac{1}{|\beta_g|} > \frac{1}{u_g}\right\},$$

$$l|\xi \sim \prod_{g=1}^{G} \mathrm{Exp}\left(\xi_g^2\right),$$

$$\xi|u \sim \prod_{g=1}^{G} Exp\left(\frac{1}{u_g}\right),$$

$$u \sim \prod_{g=1}^{G} Gamma\ (2, \lambda),$$

Conditional Posteriors

The conditional posterior distributions are derived as follows:

$$\beta|z_i \sim N_P((X'\Omega^{-1}X + T^{-1})^{-1}X'\Omega^{-1}(z - \vartheta v), (X'\Omega^{-1}X + T^{-1})^{-1}) \prod_{g=1}^{G} I\left\{|\beta_g| > \frac{1}{u_g}\right\},$$

$$v_i^{-1}|z_i \sim Inverse - Gaussian\left(\frac{1}{2}, \frac{1}{\left|z_i + b_{\tau_k} + x_i'\beta\right|}, \frac{1}{2\sigma}\right),$$

$$l^{-1}|z_i \sim \prod_{g=1}^{G} Inverse - Gaussian\left(\frac{1}{2}, \sqrt{\frac{\xi_g^2}{\beta_g^2}}, \xi_g^2\right),$$

$$\xi|z_i \sim \prod_{g=1}^{G} Gamma\left(|\beta_g| + \frac{1}{u_g}\right),$$

$$u|z_i \sim \prod_{g=1}^{G} Exponential(\lambda) I\left\{u_g > \frac{1}{|\beta_g|}\right\},$$

$$\sigma|z_i \sim Inverse - Gamma\left(a + \frac{3n}{2}, b + \frac{1}{4}\left(z_i - b_{\tau_k} - x_i'\beta - \vartheta v\right)' V^{-1}\left(z_i - b_{\tau_k} - x_i'\beta - \vartheta v\right)\right),$$

$$\lambda|z_i \sim Gamma\left(\gamma + 2p, d + \sum_{g=1}^{G} \frac{1}{|\beta_g|}\right),$$

where $L = diag(l_1, \ldots, l_G)$, $\Omega = diag((2\sigma v_1), \ldots, (2\sigma v_n))$, γ, p, and d are fixed hyper parameters.

Algorithm 1: MCMC Sampling for the Bayesian Reciprocal Lasso Composite Quantile Regression (SMN)

1. Input: z, x
2. Initialize: $b_\tau, \beta, \sigma, v, u, \lambda$
3. For $t = 1, \ldots, (t_{max} + t_{burn-in})$ do:
   - sample $v|. \sim \prod_{i=1}^{n} Inverse\ Gaussian\left(\frac{1}{2\sigma}, \frac{1}{\left|z_i - b_{\tau_k} - x_i'\beta\right|}, \frac{1}{2\sigma}\right)$
   - sample $u|. \sim \prod_{g=1}^{G} Exponential\ (\lambda) I\left\{u_g > \frac{1}{|\beta_g|}\right\}$
   - sample $l|. \sim \prod_{g=1}^{G} Inverse - Gaussin\left(\frac{1}{2}, \sqrt{\frac{\xi_k^2}{\beta_k^2}}, \xi_k^2\right)$

- sample $\xi \Big| . \sim \prod_{g=1}^{G} Gamma\left(2, \left(|\beta_g| + \frac{1}{u_g}\right)\right)$
- sample $\beta|$. From a truncated multivariate normal proportional to

$$N_P((X'\Omega^{-1}X + T^{-1})^{-1}X'\Omega^{-1}(z - \vartheta v), (X'\Omega^{-1}X + T^{-1})^{-1}) \prod_{k=1}^{p} I\left\{|\beta_k| > \frac{1}{u_k}\right\},$$

- $\hat{\beta} = \left(\sum_{i=1}^{n}\sum_{k=1}^{K}\frac{x'x}{2\sigma v_i}\right)$ and $\hat{\beta} = \hat{B}\left(\sum_{i=1}^{n}\sum_{k=1}^{K}\frac{\left(x_i\left(z_i - b_{\tau_k} - x_i'\beta - \vartheta v_i\right)\right)}{2\sigma v_i}\right)$
- sample $b_\tau|. \sim N\left(\frac{\sum_{i=1}^{n}\sum_{k=1}^{K}\left(z_i - b_{\tau_k} - x_i'\beta - \vartheta v_i\right)}{\sum_{i=1}^{n} 1/2\sigma v_i}, \frac{1}{\sum_{i=1}^{n} 1/2\sigma v_i}\right)$
- sample $\sigma|. \sim Inverse\ Gamma\left(a, \frac{3n}{2}, b + \frac{1}{4}\sum_{i=1}^{n}\sum_{K=1}^{k}\left(z_i - b_{\tau_k} + x_i'\beta - \vartheta v_i\right)' V^{-1}\sum_{i=1}^{n}\sum_{k=1}^{K}\left(z_i - b_{\tau_k} + x_i'\beta - \vartheta v_i\right)\right)$
- sample $\lambda|. \sim Gamma\left(\gamma + 2p, d + \sum_{g=1}^{G}\frac{1}{|\beta_g|}\right)$
- sample δ_c, with c from 1 to $C-1$, from a uniform distribution over the interval $(min\{min(z_i|y_i = c+1), \delta_{c+1}, \delta_C\}, \max\{max(z_i|y_i = c), \delta_{c-1}, \delta_0\})$.
- Sample z_i, for i from 1 to n, from truncated normal (TN) distribution $TN_{(\delta_{c-1},\delta_c)}\left(b_{\tau_k} + x_i'\beta + \vartheta v, 2\sigma v\right)$.

4. end for

In our implementation of the Gibbs sampler for BRGLCQR, we used 10,000 iterations, discarding the first 2,000 samples as burn-in to reduce the influence of initial values. Convergence diagnostics were performed using trace plots and the Gelman-Rubin statistic, ensuring that all chains stabilized and mixed well across parameters. The final posterior summaries were based on the remaining 8,000 samples, which provided accurate estimates and stable inference across all quantiles.

2.3 Bayesian Reciprocal Group Lasso Composite Quantile Regression (BRGLCQR)

The **(**BRGLCQR**)** model is tailored for variable selection and quantile estimation, particularly beneficial when predictors are organized into groups. This model enhances composite quantile regression by implementing a Bayesian framework to improve group-level variable selection accuracy, which is especially useful in high-dimensional data settings.

The BRGLCQR model seeks to minimize an objective function that integrates a composite quantile loss term with a reciprocal group lasso penalty to promote sparsity across grouped predictors. The optimization objective is defined as:

$$\underbrace{argmin}_{b_\tau,\beta} \sum_{i=1}^{n}\left\{\sum_{k=1}^{K}\rho_{\tau_k}\left(z_i - b_{\tau_k} - x_i'\beta\right)\right\} + \lambda\sum_{g=1}^{G}\frac{1}{|\beta_{I_g}|} I\{\beta_g \neq 0\} \quad (6)$$

where λ is the regularization parameter. $\frac{1}{|\beta_{I_g}|}$ applies the reciprocal group lasso penalty across each group g, enhancing sparsity at the group level by focusing on the inverse magnitude of group norms. This setup is designed to encourage selection of entire groups, either retaining or excluding all predictors in each group based on their collective contribution.

The Bayesian hierarchical framework enables the use of prior distributions on each model parameter, allowing for effective regulation of model complexity through group-wise sparsity. The model for latent variables z_i is formulated as:

$$z_i|x \sim N_n\left(z_i + b_{\tau_k} + x_i'\beta + \vartheta v, 2\sigma v\right),$$

The prior distributions used in our model, including inverse-gamma priors for variance components and exponential priors for scale parameters, were selected based on their conjugacy and tractability within the Gibbs sampler. These priors are commonly used in Bayesian shrinkage models and offer computational efficiency. To assess robustness, we performed sensitivity checks by varying the hyper parameters (e.g., shape and scale of the inverse-gamma priors) and observed that the posterior estimates remained stable, indicating that the model is not overly sensitive to the specific prior choices. For the group coefficients, we apply the following prior:

$$\beta|l \sim \prod_{g=1}^{G} N(0, t^2) I\left\{\frac{1}{|\beta_{I_g}|} > \frac{1}{u_g}\right\},$$

This prior enforces the reciprocal group lasso structure by setting a threshold for the group coefficients. The scale parameters have priors defined as:

$$l|\xi \sim \prod_{g=1}^{G} Exp\left(\xi_g^2\right),$$

$$\xi|u \sim \prod_{g=1}^{G} Exp\left(\frac{1}{u_g}\right),$$

$$u \sim \prod_{g=1}^{G} Gamma\ (2, \lambda),$$

Posterior Distributions

Using Gibbs sampling, we derive the conditional posterior distributions to iteratively update model parameters:

$$\beta|z_i \sim N_P((X'\Omega^{-1}X + T^{-1})^{-1}X'\Omega^{-1}(z - \vartheta v), (X'\Omega^{-1}X + T^{-1})^{-1}) \prod_{g=1}^{G} I\left\{\frac{1}{|\beta_{I_g}|} > u_g\right\},$$

$$v_i^{-1}|z_i \sim Inverse - Gaussian\left(\frac{1}{2}, \frac{1}{\left|z_i + b_{\tau_k} + x_i'\beta\right|}, \frac{1}{2\sigma}\right),$$

$$l^{-1}|z_i \sim \prod_{g=1}^{G} Inverse - Gaussian\left(\frac{1}{2}, \frac{\xi_g^2}{|\beta_{I_g}|}, \xi_g^2\right),$$

$$\xi|z_i \sim \prod_{g=1}^{G} Gamma\left(|\beta_i| + \frac{1}{u_g}\right),$$

MCMC Sampling Algorithm

Algorithm: MCMC Sampling for Bayesian Reciprocal Group Lasso Composite Quantile Regression

1. Input Data z, x
2. Initialize: $(b_\tau, \beta, \sigma, v, u, \lambda)$
3. Iterate For $t = 1, \dots, (t_{max} + t_{burn-in})$:
   - sample $v|. \sim \prod_{i=1}^{n} \; Inverse\ Gaussian\left(\frac{1}{2\sigma}, \frac{1}{\left|z_i - b_{\tau_k} - x_i'\beta\right|}, \frac{1}{2\sigma}\right)$
   - sample $u|. \sim \prod_{g=1}^{G} Exponential\ (\lambda) I\left\{u_g > \frac{1}{|\beta_{I_g}|}\right\}$
   - sample $l|. \sim \prod_{g=1}^{G} Inverse - Gaussin\left(\frac{1}{2}, \frac{\xi_g^2}{|\beta_{I_g}|}, \xi_k^2\right)$
   - sample $\xi \Big|. \sim \prod_{g=1}^{G} Gamma\left(2, \left(|\beta_i| + \frac{1}{u_g}\right)\right)$
   - sample $\beta|.$ From a truncated multivariate normal proportional to
     $N_P((X'\Omega^{-1}X + T^{-1})^{-1}X'\Omega^{-1}(z - \vartheta v), (X'\Omega^{-1}X + T^{-1})^{-1}) \prod_{k=1}^{p} I\left\{|\beta_{I_g}| > \frac{1}{u_g}\right\},$
   - sample $b_\tau|. \sim N\left(\sum_{i=1}^{n}\sum_{k=1}^{K} \frac{\left(z_i - b_{\tau_k} - x_i'\beta - \vartheta v_i\right)}{2\sigma v_i}, \frac{1}{2\sigma v_i}\right)$
   - sample $\sigma|. \sim Inverse\ Gamma\ \left(a, \frac{3n}{2}, b + \frac{1}{4}\sum_{i=1}^{n}\sum_{K=1}^{k}\ \left(z_i - b_{\tau_k} + x_i'\beta - \vartheta v_i\right)' V^{-1} \sum_{i=1}^{n}\sum_{k=1}^{K}\ \left(z_i - b_{\tau_k} + x_i'\beta - \vartheta v_i\right)\right)$
   - sample $\lambda|. \sim Gamma\ \left(\gamma + 2p, d + \sum_{g=1}^{G} \frac{1}{|\beta_{I_g}|}\right)$

End Iteration.

3. SIMULATION STUDIES

To assess the performance of our proposed Bayesian Reciprocal Group Lasso Composite Quantile Regression (BRGLCQR) method, we conducted simulation studies with varying sample sizes (50 and 100) to evaluate accuracy in variable selection and prediction. Our model was compared against two alternative methods: Composite Quantile Regression with Lasso (CQR-Lasso; Zou and Yuan, 2008) and Bayesian Composite

Quantile Regression (BCQR; Alhamzawi and Yu, 2013). The comparisons used two metrics: Mean Squared Prediction Error (MSPE) and Model Selection Accuracy (MSA).

The data was generated according to the following linear model:

$$y_i = X_i'\beta + \epsilon_i, i = 1,2,\dots,n$$

where X_i represents the predictor matrix with relevant predictors and noise variables. Each relevant predictor variable was sampled from a standard normal distribution, $N(0,1)$, and ten additional noise variables were generated independently with a covariance structure $\sum_x$, where $(\sum_x)_{gh} = 0.75^{|g-h|}$. The error term ϵ_i follows a normal distribution $N(0,1)$.

The response variable y was categorized based on specified cut-points to ensure a discrete response, relevant to composite quantile regression. Sample Sizes: 50 and 100 Cut-Point Vector: $\delta = (0.5,2,3.5)$. Relevant Predictors: Three predictors were chosen to be relevant, while ten were treated as noise. Mean Squared Prediction Error (MSPE): This measures the accuracy of predictions generated by each model. Model Selection Accuracy (MSA): This metric evaluates the number of correctly identified relevant predictors versus the number of false inclusions.

Table 1
Simulation Results for Sample Size 50

Method	MSPE (SD)	MSA (SD)
BRGLCQR	0.183 (0.015)	0.925 (0.032)
CQR-Lasso	0.235 (0.023)	0.853 (0.045)
BCQR	0.276 (0.031)	0.812 (0.050)

Table 2
Simulation Results for Sample Size 100

Method	MSPE (SD)	MSA (SD)
BRGLCQR	0.162 (0.012)	0.940 (0.028)
CQR-Lasso	0.221 (0.020)	0.872 (0.038)
BCQR	0.258 (0.029)	0.835 (0.046)

Table 1 and Table 2 demonstrate that the BRGLCQR method consistently achieved superior results in prediction accuracy (indicated by lower MSPE values) and model selection accuracy (shown by higher MSA values) when compared to both CQR-Lasso and BCQR across sample sizes of 50 and 100. These results suggest that the Bayesian Reciprocal Group Lasso Composite Quantile Regression method offers substantial improvements in variable selection accuracy while maintaining strong predictive reliability. The findings imply that BRGLCQR is particularly effective in high-dimensional contexts where predictors are grouped, highlighting its potential as a robust tool for grouped variable selection in complex data structures.

4. REAL DATA

In this section, we apply the proposed Bayesian Reciprocal Group Lasso Composite Quantile Regression (BRGLCQR) model to analyze stock market returns and volatility. The data consists of daily closing prices from the S&P 500 index, spanning from January 1, 2010, to December 31, 2020, which includes 2,758 observations. The data was obtained from Yahoo Finance.

We calculate daily log returns using the following formula:

$$r_i = \log\left(\frac{p_t}{p_{t-1}}\right)$$

where p_t represents the closing price on day t, and p_{t-1} is the closing price from the previous day. To measure volatility, we compute the 10-day rolling standard deviation of returns:

$$\sigma_t = \sqrt{\frac{1}{10}\sum_{i=0}^{9}(r_{t-1} - \bar{r}_t)^2}$$

where $\bar{r}_t$ is the mean return over the past 10 days. This provides us with a measure of market volatility over time.

We use the BRGLCQR model to explore the relationship between daily stock returns and their lagged values, along with lagged volatility measures. The predictors in the model include:

r_{t-1}: Return from the previous day.
r_{t-2}: Return from two days prior.
σ_{t-1}: Volatility from the previous day.
σ_{t-2}: Volatility from two days prior.

To evaluate the performance of our method, we compare BRGLCQR with two established techniques:

1. **Composite Quantile Regression with Lasso (CQR-Lasso):** This method applies Lasso regularization to composite quantile regression, encouraging sparsity in the model.
2. **Bayesian Composite Quantile Regression (BCQR):** A Bayesian version of composite quantile regression, which incorporates prior information into the quantile regression framework.

We assess the performance of these models using two main criteria:

1. **Mean Absolute Prediction Error (MAPE):** Measures the average absolute difference between the predicted and actual returns. A lower MAPE indicates better predictive accuracy.
2. **Model Selection Accuracy (MSA):** Reflects how well each method identifies significant predictors. A higher MSA suggests better variable selection.

We used 10-fold cross-validation to evaluate the performance of each model. The results, averaged across all folds, are presented in Table below.

Table 3
Model Performance Comparison

Method	MAPE (Mean ± SD)	MSA (Mean ± SD)
BRGLCQR	0.015 ± 0.002	0.92 ± 0.03
CQR-Lasso	0.018 ± 0.003	0.85 ± 0.04
BCQR	0.020 ± 0.004	0.80 ± 0.05

The results clearly show that BRGLCQR outperforms both CQR-Lasso and BCQR in terms of predictive accuracy (as indicated by the lower MAPE) and model selection accuracy (as indicated by the higher MSA). Specifically, the BRGLCQR model provides more precise predictions of stock returns and more accurately identifies the important predictors.

This performance advantage suggests that BRGLCQR is particularly well-suited for financial time series data, where relationships between variables can be complex and difficult to capture. The group lasso penalty in our model helps to account for the grouped structure in the data, which may represent factors such as market trends and volatility clusters. This leads to better predictions and a clearer understanding of the key drivers of stock market behavior.

In summary, the BRGLCQR method not only improves prediction accuracy but also enhances interpretability by selecting the most relevant predictors. This makes it a powerful tool for financial modeling, particularly in high-dimensional settings with grouped predictors.

5. CONCLUSIONS

The Bayesian Reciprocal Group Lasso Composite Quantile Regression (BRGLCQR) model presented in this study addresses significant challenges in high-dimensional quantile regression, especially where predictors exhibit natural grouping. The BRGLCQR model's Bayesian framework, combined with a reciprocal penalty approach, enhances both the accuracy of variable selection and prediction, making it particularly effective for data with complex grouped structures. Simulation results and application to stock market returns and volatility data confirm that BRGLCQR consistently outperforms traditional approaches like Composite Quantile Regression with Lasso and Bayesian Composite Quantile Regression in terms of predictive accuracy and model selection. The findings suggest that BRGLCQR provides substantial improvements in capturing essential predictor groups, demonstrating its value as a versatile tool for financial and economic modeling. Future research could explore the extension of this framework to other domains, leveraging its adaptability for high-dimensional, grouped data in diverse applications.

REFERENCES

1. Koenker, R. and Bassett, G. (1978). Regression Quantiles. *Econometrica*, 46(1), 33-50.
2. Alhamzawi, R. and Yu, K. (2013). Conjugate priors and variable selection for Bayesian quantile regression. *Computational Statistics & Data Analysis*, 64, 209-219.
3. Zou, H. and Yuan, M. (2008). Composite quantile regression and the oracle model selection theory. *The Annals of Statistics*, 36(3), 1108-1126.
4. Yu, K. and Moyeed, R.A. (2001). Bayesian quantile regression. *Statistics & Probability Letters*, 54(4), 437-447.
5. Mallick, B.K., Alhamzawi, R. and Svetnik, V. (2020). Bayesian Reciprocal Adaptive Bridge Regression. *Journal of Computational and Graphical Statistics*, 29(3), 642-653.
6. Koenker, R. (2005). *Quantile Regression*. Cambridge Books. Cambridge University Press.
7. Orsini, N. and Bottai, M. (2011). Logistic quantile regression in Stata. *The Stata Journal*, 11(3), 327-344.
8. Armagan, A., Dunson, D.B. and Lee, J. (2013). Generalized double Pareto shrinkage. *Statistica Sinica*, 23(1), 119-143.
9. Wang, H., Li, B. and Leng, C. (2007). Shrinkage tuning parameter selection in regularized quantile regression. *Statistica Sinica*, 17(2), 749-763.
10. Alhamzawi, R. (2016). Bayesian model selection in ordinal quantile regression. *Bayesian Analysis*, 11(1), 1-24.
11. Yu, K., Lu, Z. and Stander, J. (2003). Quantile regression: applications and current research areas. *The Statistician*, 52, 331-350.
12. Koenker, R. and D'Orey, V. (1987). Algorithm AS 229: Computing regression quantiles. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 36(3), 383-393.
13. Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91(1), 74-89.
14. Alhamzawi, R. and Mallick, H. (2022). Bayesian reciprocal LASSO quantile regression. *Communications in Statistics-Simulation and Computation*, 51(11), 6479-6494.
15. Raheem, S.H., Alhusseini, F.H.H. and Alshaybawee, T. (2020). Modelling volatility in financial time series using ARCH models. *International Journal of Innovation, Creativity and Change*, 12(7), 248-261.