Bayesian variable selection in linear quantile mixed models for longitudinal data with application to macular degeneration

Yonggang Ji, Haifang Shi

School of Science, Civil Aviation University of China, Tianjin, China

* hfshi@cauc.edu.cn

Abstract

This paper presents a Bayesian analysis of linear mixed models for quantile regression based on a Cholesky decomposition for the covariance matrix of random effects. We develop a Bayesian shrinkage approach to quantile mixed regression models using a Bayesian adaptive lasso and an extended Bayesian adaptive group lasso. We also consider variable selection procedures for both fixed and random effects in a linear quantile mixed model via the Bayesian adaptive lasso and extended Bayesian adaptive group lasso with spike and slab priors. To improve mixing of the Markov chains, a simple and efficient partially collapsed Gibbs sampling algorithm is developed for posterior inference. Simulation experiments and an application to the Age-Related Macular Degeneration Trial data to demonstrate the proposed methods.

Introduction

Quantile regression (QR) in longitudinal or panel data models have received increasing attention in recent years. For example, [1] proposed a general method to estimate the QR coefficients by using a $\ell_1$ penalized likelihood approach. Later [2] generalized the work of [1] to be more flexible model with endogenous variables. [3] offered a estimation of individual effects for panel data QR models, which is widely applied by practitioners because of computationally simple. [4] studied quantile panel models with correlated random effects. [5] proposed a Stochastic Approximation of the EM (SAEM) algorithm to analyze linear quantile mixed regressions (LQMMs) via the asymmetric Laplace distribution. Other related papers include [6–9], among many others.

Like all regression issues, when there are many covariates in longitudinal or panel QR models, variable selection becomes necessary to avoid overfitting and multicollinearity. [10] presented a penalized quantile regression model for random intercept using the Bayesian lasso priors and Bayesian adaptive lasso priors, respectively. [11] also considered a Bayesian Lasso approach to jointly estimate a vector of covariate effects and a vector of random effects by introducing an $\ell_1$ penalty. However, for the random effects, their method mainly focus on the penalty of diagonal elements in the covariance matrix of random effects after a modified
Cholesky decomposition. In addition, they neglect the nonnegative constraint on the diagonal terms in the covariance matrix of random effects after a modified Cholesky decomposition, which can cause non-uniqueness in the decomposition [12]. Then, we first propose a new Bayesian shrinkage approach to solve these problems in the LQMMs. In particular, we develop an extended Bayesian adaptive group lasso which can accommodate the nonnegative constraint of the diagonal terms to attempt to identify the non-zero random effects. Unfortunately, the shrinkage approach need some ad hoc methods to do variable selection similar to [10] and [11] for the well known reason. An alternative to the Bayesian variable selection is to use spike and slab priors. We then develop a general variable selection technic via the Bayesian adaptive lasso and the extended Bayesian adaptive group lasso with spike and slab priors to identify the significant fixed and random effects simultaneously.

The paper is organized as follows. Section 2 presents the reparameterized LQMMs based on a Cholesky decomposition of covariance matrix of random effects. Section 3 describes the structure of our hierarchical Bayesian LQMMs and discuss our prior specifications. We also develop a Partially collapsed Gibbs (PCG) sampling algorithm for posterior inference. In section 4, we will develop a variable selection procedure for both fixed and random effects in a LQMM with spike and slab priors and an efficient PCG. We illustrate the performance of our method by simulation studies and a real data example in section 5 and 6. Results show that the proposed approach performs very well. In Section 7, we conclude the paper.

The re-parameterized LQMMs
Consider \( n \) subjects containing \( n_i \) observations, \((y_{ij}, x_{ij}', z_{ij}')\), for \( j = 1, \ldots, n_i \) and \( i = 1, \ldots, n, N = \sum_{i=1}^{n} n_i \), where \( y_{ij} \) is the \( j \)th response for subject \( i \), \( x_{ij}' \) is the \( j \)th row of a known \( n_i \times p \) matrix \( x_i \) and \( z_{ij}' \) is the \( j \)th row of a known \( n_i \times q \) matrix \( z_i \). The canonical linear mixed model is as follows:

\[
y_{ij} = x_{ij}' \beta + z_{ij}' \alpha_i + \varepsilon_{ij}, \quad \alpha_i \sim N(0, D),
\]

where \( \beta \) is a \( p \times 1 \) vector of unknown population fixed parameters, \( \alpha_i \) is a \( q \times 1 \) vector of unobservable random effects with covariance matrix \( D \). In order to guarantee the positive semidefinite of \( D \), we adopt the Cholesky decomposition to it, resulting in

\[
D = \Gamma \Gamma^T,
\]

where \( \Gamma = (\gamma_{jk}) \) is a \( q \times q \) lower triangular matrix with nonnegative diagonal entries. Similar to [13], the model (1) can be re-expressed as follows:

\[
y_{ij} = x_{ij}' \beta + z_{ij}' \Gamma b_k + \varepsilon_{ij} = x_{ij}' \beta + (b_k^T \otimes z_{ij}') I_q + \varepsilon_{ij}, \quad b_k \sim N(0, I_q).
\]

Here \( \gamma^T = (\gamma_1^T, \ldots, \gamma_J^T) \) is a \( 1 \times q(q + 1)/2 \) vector, where \( \gamma_k \) consists of the non-zero lower triangular elements of the \( k \)th row of \( \Gamma \), i.e., \( \gamma_k = (\gamma_{k1}, \ldots, \gamma_{kk})^T, k = 1, \ldots, q \). \( I_q \) is a transformation matrix with size \( q^2 \times q(q + 1)/2 \) such that \( \text{vec}(\Gamma) = I_q \gamma \). \( I_q \) denotes the \( q \times q \) identity matrix. For example, if \( q = 2 \)

\[
\begin{pmatrix}
\gamma_{11} & 0 \\
\gamma_{21} & \gamma_{22}
\end{pmatrix}.
\]
the transformation matrix $J_2$ would have the following form

$$J_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$ 

With model (2), the linear conditional $\tau$th ($0 < \tau < 1$) mixed quantile estimator can be calculated by

$$\min_{\mathbf{b}, \mathbf{g}} \sum_{i=1}^{n} \sum_{j=1}^{n_i} \rho_{\tau}(y_{ij} - \mathbf{x}_i' \mathbf{b} - (\mathbf{b}_i' \otimes \mathbf{z}_j' \mathbf{1}'), J_2 g).$$

In a Bayesian setup, this is equivalent to assuming that the error terms $e_{ij}$ in (2) follow the asymmetric Laplace distribution (ALD), of which the probability density function is given by

$$f_p(e_{ij}) = \tau (1 - \tau)^{-\sigma} \exp\{-\sigma \rho_{\tau}(e_{ij})\}, \rho_{\tau}(e_{ij}) = e_{ij} \{\tau - I(e_{ij} < 0)\},$$

where $\sigma^{-1}$ and $\tau$ are the scale and skewness parameters, respectively. [14] first used the ALD to establish Bayesian QR in a linear model for independent data and proposed a random walk Metropolis-Hastings algorithm to draw samples. Since then, there has been a great deal of literature on the extension and application of the QR based on the ALD such as [15–18]. However, direct sampling using the ALD is often inconvenient and not easy to generalize to more complex scenarios. We employ a mixture representation of ALD proposed by [19] to facilitate Bayesian inference in LQMMs, which was also used by many authors (e.g., see [11, 20–22]),

$$y_{ij} = \mathbf{x}_i' \mathbf{b} + \mathbf{z}_j' \mathbf{b}_i + \xi_{ij} v_{ij} + \xi_{ij}^{-1/2} \sqrt{v_{ij} \xi_{ij}}, j = 1, \ldots, n_i, i = 1, \ldots, n,$$

where $v_{ij} \sim \exp(1/\sigma)$ has an exponential prior with mean $1/\sigma$, $\xi_{ij} \sim N(0, 1)$ has a standard normal prior and is independent of $v_{ij}$, $\xi_1 \sim \frac{1}{\sigma(1-\tau)}$ and $\xi_2 \sim \frac{1}{\sigma(1+\tau)}$. Let $v = (v_{ij}; i = 1, \ldots, n; j = 1, \ldots, n_i), \xi = (\xi_{ij}; i = 1, \ldots, n; j = 1, \ldots, n_i).$ Hence, based on the Cholesky decomposition and the mixture representation of ALD, we obtain the following hierarchal model,

$$y_{ij} = \mathbf{x}_i' \mathbf{b} + (\mathbf{b}_i' \otimes \mathbf{z}_j') J_2 g + \xi_{ij} v_{ij} + \xi_{ij}^{-1/2} \sqrt{v_{ij} \xi_{ij}},$$

$$v|\sigma \sim \prod_{i=1}^{n} \prod_{j=1}^{n_i} \exp(-\sigma v_{ij}),$$

$$\xi \sim \prod_{i=1}^{n} \prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{v_{ij}^2}{\xi_{ij}}\right).$$

**Shrinkage estimation of linear quantile mixed regression**

We first propose a new Bayesian shrinkage approach in LQMMs by using the Bayesian adaptive lasso and an extended Bayesian adaptive group lasso. A regularized $\tau$th quantile estimates
of the fixed and random effect coefficients can be defined by

$$\arg \min_{\beta} \sum_{i=1}^{n_i} \sum_{j=1}^{y_i} \rho_i(y_{ij} - \mathbf{x}_i^T \mathbf{\beta} - (\mathbf{b}_i^T \otimes \mathbf{z}_j^T) \mathbf{I}_{y_i}) + \sum_{i=1}^{p} \lambda_i^f |\beta_i| + \sum_{k=1}^{q} \lambda_k^r (\gamma_k^T \gamma_k)^{1/2}$$

(6)

s.t. \( \gamma_{ik} \geq 0, k = 1, \ldots q \).

Motivated by [23], we put a Laplace priors distribution of \( \lambda_i^f \), a Laplace priors \( \pi(\gamma_k | \gamma_k^T \gamma_k) = C_i \exp \{ -\sigma \sum_{i=1}^{p} \lambda_i^f |\beta_i| \} \) on \( \beta_i \), where \( C_i = \sigma \lambda_i^f / 2 \). A Laplace priors \( \pi(\gamma_k | \gamma_k^T \gamma_k) = C_2 \exp \{ -\sigma \lambda_k^r (\gamma_k^T \gamma_k)^{1/2} \} \), where \( C_2 = 2^{-(k+1)/2} \Gamma((k+1)/2) / \Gamma(k+1/2) \), and suppose that the error terms \( e_{ij} \) come from the ALD (4). Then, the posterior distribution of \( \mathbf{\beta} \) and \( \gamma \) is proportional to

$$\exp\{-\sigma \rho_i(y_{ij} - \mathbf{x}_i^T \mathbf{\beta} - (\mathbf{b}_i^T \otimes \mathbf{z}_j^T) \mathbf{I}_{y_i}) - \sigma \sum_{i=1}^{p} \lambda_i^f |\beta_i| - \sigma \sum_{k=1}^{q} \lambda_k^r (\gamma_k^T \gamma_k)^{1/2}\}.$$ 

(7)

So it can be easily observed that maximizing the posterior distribution (7) is equivalent to minimizing the target function (6) if we ignore the nonnegative constraints \( \gamma_{ik} \geq 0, k = 1, \ldots q \). However, it is not clear how to incorporate this constraints within the Bayesian Penalized LQMMs framework. Using truncated normal priors, [24] and [25] solved a similar problem on the selection of fixed effects, which is non-group structure in linear and generalized mixed model, respectively. We extend the idea to the Bayesian Group lasso on the shrinkage of random effects with group structure. We specify the conditional prior distributions of \( \gamma_k \) as \( \pi(\gamma_k | \lambda_k^r) = 2C_2 \exp \{ -\sigma \lambda_k^r (\gamma_k^T \gamma_k)^{1/2} \} I(\gamma_k \geq 0) \). To facilitate posterior inference, we utilize a mixture representation to yield

$$\pi(\gamma_k | \lambda_k^r) = 2C_2 \exp \{ -\sigma \lambda_k^r (\gamma_k^T \gamma_k)^{1/2} \} I(\gamma_k \geq 0)$$

$$= \int_{0}^{\infty} \frac{1}{\sqrt{2\pi \eta_k}} \exp \left( -\frac{\gamma_k^T \gamma_k}{2\eta_k} \right) I(\gamma_k \geq 0) \exp \left( -\frac{\lambda_k^r}{2\eta_k} \right) d\eta_k,$$ 

where \( \lambda_k^r = \sigma \lambda_k^r \). Then, \( \gamma_k \) can be expressed hierarchically as

$$\gamma_k | \eta_k \sim N(0, \eta_k I_k) \quad and \quad \eta_k \lambda_k^2 \sim \text{Gamma} \left( \frac{k+1}{2}, \frac{\lambda_k^2}{2} \right).$$ 

(8)

for \( k = 1, \ldots, q \). Moving to the fixed effects parameters, \( \mathbf{\beta} \) can also be written hierarchically as follows

$$\beta_i | t_i \sim N(0, t_i) \quad and \quad t_i \lambda_i^2 \sim \text{Gamma} \left( \frac{1}{2}, \frac{\lambda_i^2}{2} \right),$$ 

(9)

for \( s = 1, \ldots, p \). There are two main ways to estimate the tuning parameter in Bayesian regularization: first, a fully Bayesian method, which specifies a hyperprior on it, such as a conjugate Gamma distribution [23], and second, an empirical Bayesian method, which estimates it using maximum marginal likelihood method [26]. Here we use the fully Bayesian approach and set Gamma priors on the parameters \( \lambda_i^2, \lambda_k^2 \) and \( \sigma \) to complete the Bayesian structure. We develop a Partially collapsed Gibbs (PCG) sampling sampler by integrating the full conditional distribution of \( \mathbf{\beta} \) with respect to \( \{b_i\} \) to sample from posterior distributions. In order to improve the
Next, we present the conditional distribution of unknown parameters. The key idea of the PCG is to replace some of the posterior distributions with marginal posterior distribution while preserving the target distribution. Next, we present the conditional distribution of unknown parameters.

1) Conditional distribution of $\beta$:
Let $T = \text{diag}(t_s, s = 1, \ldots, p)$ and $w_y = (z'_y \Gamma)(z'_y \Gamma)' + \xi_y \sigma^{-1} v_y$. The conditional distribution of $\beta$ is then a $N(\hat{\mu}_\beta, \hat{\Sigma}_\beta)$, where

$$
\hat{\Sigma}_\beta = \left[ \sum_{j=1}^n \sum_{i=1}^{n_j} x_i x_j w_i w_j + T^{-1} \right]^{-1}, \quad \hat{\mu}_\beta = \hat{\Sigma}_\beta \sum_{j=1}^n \sum_{i=1}^{n_j} x_i w_i (y_i - \xi_i v_y).
$$

2) Conditional distribution of $b_i$:
The conditional distribution of $b_i$ is a multivariate normal distribution with mean $\hat{\mu}_b$ and variance $\hat{\Sigma}_b$, where

$$
\hat{\Sigma}_b = \left[ \sum_{j=1}^n \frac{(z'_j \Gamma)(z'_j \Gamma)}{2 \sigma^{-1} v_j} + \frac{I}{q} \right]^{-1}, \quad \hat{\mu}_b = \hat{\Sigma}_b \sum_{j=1}^n \frac{(z'_j \Gamma)(y_j - x'_j \beta - \xi_j v_j)}{2 \sigma^{-1} v_j}.
$$

3) Conditional distribution of $v^{-1}_y$:
The full conditional distribution of $v^{-1}_y$ follows a Inverse Gaussian($\mu', \lambda'$) with p.d.f.

$$
f(x|\mu', \lambda') = \sqrt{\frac{\lambda'}{2\pi x^3}} \exp \left\{ -\frac{\lambda' (x - \mu')^2}{2(\mu')^2 x} \right\}, \quad (10)
$$

Here $\mu' = \sqrt{\frac{\xi_y^2 + \xi_v^2}{\mathbb{E}[e_y^2]}}$, and $\lambda' = \frac{\xi_y^2 + \xi_v^2}{\mathbb{E}[e_y^2]}$, where $e_y = y_i - x'_i \beta - z'_y \Gamma b_i$.

4) Conditional distributions of $\lambda_i^2$ and $\lambda_{2k}^2$:
The full conditional distributions of $\lambda_i^2$ and $\lambda_{2k}^2$ are independent Gamma distributions,

$$
\lambda_i^2 | a_{ki}, b_{ki} \sim \text{Gamma} \left( \frac{1}{2}, \frac{t_i}{2} + b_{ki} \right), \quad \lambda_{2k}^2 | a_{2k}, b_{2k} \sim \text{Gamma} \left( \frac{k+1}{2}, \frac{\eta_k}{2} + b_{2k} \right).
$$

5) Conditional distributions of $t_i$ and $\eta_k$:
The full conditional distributions of $t_i$ and $\eta_k$ are independent Inverse Gaussian,

$$
t_i^{-1} | \lambda_{1i}, \beta_i \sim \text{Inverse Gaussian} \left( \frac{\lambda_{1i}^2}{\beta_i}, \lambda_{1i}^2 \right), \quad \eta_k^{-1} | \lambda_{2k}, \gamma_k \sim \text{Inverse Gaussian} \left( \frac{\lambda_{2k}^2}{\gamma_k^2}, \lambda_{2k}^2 \right).
$$

6) Conditional distribution of $\gamma$:
Denote $F_{ij} = (b'_i \otimes z'_j) F_y i = 1, \ldots, n; j = 1, \ldots, n_i$, and let $F_{ijk}$ be the covariate vector corresponding to $\gamma_k, k = 1, \ldots, q$. The full conditional distribution of $\gamma_k$ is then a truncated normal
distribution $N_k((\bar{\mu}_i + \sum_{j=1}^{\tau_k})I(\gamma_{ik} \geq 0))$ where

\[
\bar{\mu}_i = \sum_{j=1}^{\tau_k} \frac{F_{ik}^T \eta_j}{2\sigma^2 v_j},
\]

\[
\sum_{j=1}^{\tau_k} \frac{F_{ik}^T \eta_j}{2\sigma^2 v_j} + \eta_k I
\]

where $\bar{\gamma}_i = (\gamma_1^T, \ldots, \gamma_{k-1}^T, \gamma_k^T, \ldots, \gamma_{\tau_k}^T)^T$ and $F_{ik}$ represent the covariate matrix corresponding to $\gamma_{ik}$.

(7) Conditional distribution of $\sigma$

The full conditional distribution of $\sigma$ is

\[
\text{Gamma}\left(\frac{3N}{2} + a, \sum_{j=1}^{\tau_k} \frac{(e_j - \bar{\gamma}_j)^2}{2v_j} + \sum_{j=1}^{\tau_k} v_j + b\right).
\]

**Bayesian model selection in LQMMs**

The shrinkage approach proposed in the above section has potential for variable selection in LQMMs. However, the estimators are never exact 0 due to the continuous priors on $\beta$ and $\gamma$. This section develops a general model selection method with spike and slab priors to identify the relative important effects. We first discuss the prior specification and propose the hierarchical model. Subsequently, we consider a Partially collapsed Gibbs sampling algorithm for posterior computation, and derive the relevant conditional distributions.

We assume the following hierarchical Bayesian lasso with independent spike and slab type priors for $\beta$:

\[
\beta_i|t_i, \pi_{t_i} \sim \pi_{t_i} \delta_0(\beta_i) + (1 - \pi_{t_i})N(0, t_i),
\]

\[
t_i|\lambda_{t_{i}}^2 \sim \text{Gamma}\left(1, \frac{\lambda_{t_{i}}^2}{2}\right),
\]

\[
\lambda_{t_{i}}^2 \sim \text{Gamma}(a_{t_{i}}, b_{t_{i}}),
\]

where $\delta_0(\cdot)$ denotes a point mass at 0, $\pi_{t_i}$ is the prior probability of excluding the $i$th fixed effect in the model, which is assigned a beta prior with parameters $a_{t_{i}} = 1$ and $b_{t_{i}} = 1$, resulting in a noninformative uniform prior on (0, 1). Furthermore, we assume the following extended hierarchical adaptive Bayesian group lasso with independent spike and slab type priors for $\gamma$:

\[
\gamma_k|\eta_k, \pi_0 \sim \pi_0 \delta_0(\gamma_k) + (1 - \pi_0)N(0, \eta_k)I(\gamma_{ik} > 0),
\]

\[
\eta_k|\lambda_{\gamma_k}^2 \sim \text{Gamma}\left(k + 1, \frac{\lambda_{\gamma_k}^2}{2}\right),
\]

\[
\lambda_{\gamma_k}^2 \sim \text{Gamma}(a_{\gamma_k}, b_{\gamma_k}),
\]

where $\pi_0$ is the prior probability of excluding the $k$th random effect in the model, which is also assumed to be a noninformative uniform prior on (0, 1). We also present the conditional distribution of unknown parameters.

(1) Conditional distribution of $\beta$:
Let \( \sigma_{\beta_i}^2 \) be
\[
\begin{align*}
\sigma_{\beta_i}^2 &= \left[ \sum_{l=1}^{n_l} \sum_{j=1}^{n_j} x_{ij} w_{ij} x_{ij}^T + I_2 \right]^{-1}, \\
\mu_{\beta_i} &= \sigma_{\beta_i}^2 \sum_{l=1}^{n_l} \sum_{j=1}^{n_j} x_{ij} w_{ij} (y_{ij} - \xi_j v_{ij})
\end{align*}
\]
\( w_{ij} = (z_i \Gamma) (z_i \Gamma)' + \xi_j \sigma^{-1} v_{ij} \). The conditional distribution of \( \beta_i \) is then a spike and slab distribution,
\[
\pi'_0 \delta_0(\beta_i) + (1 - \pi'_0) N(\mu_{\beta_i}, \sigma_{\beta_i}^2),
\]
where
\[
\pi'_0 = \frac{\pi_{0u}}{\pi_{0u} + (1 - \pi_{0u}) \sqrt{2\pi \sigma^2} \exp \left\{ -\frac{\mu_{\beta_i}^2}{2\sigma_{\beta_i}^2} \right\}}.
\]

(2) Conditional distribution of \( b_i \):
The conditional distribution of \( b_i \) is a multivariate normal distribution with mean \( \mu_{b_i} \) and variance \( \Sigma_{b_i} \) where
\[
\begin{align*}
\Sigma_{b_i} &= \left[ \sum_{j=1}^{n_j} (z_j \Gamma)' (z_j \Gamma) + I_2 \right]^{-1}, \\
\mu_{b_i} &= \Sigma_{b_i} \sum_{j=1}^{n_j} (z_j \Gamma) (y_{ij} - x_i \beta - \xi_j v_{ij}) / 2\sigma^{-1} v_{ij}
\end{align*}
\]

(3) Conditional distribution of \( v_{ij}^{-1} \):
The full conditional distribution of \( v_{ij}^{-1} \) follows an Inverse Gaussian(\( \mu', \lambda' \)) with \( \mu' = \sqrt{\frac{\lambda_i + 2\lambda_k}{\lambda_i}} \) and \( \lambda' = \frac{\lambda_i + 2\lambda_k}{\lambda_i} \), where \( \epsilon_{ij} = y_{ij} - x_i \beta - z_j \Gamma b_i \).

(4) Conditional distributions of \( \lambda_{1i} \) and \( \lambda_{2k} \):
The full conditional distributions of \( \lambda_{1i} \) and \( \lambda_{2k} \) are independent Gamma distributions,
\[
\begin{align*}
\lambda_{1i}^2 | t_i, a_{\lambda_i}, b_{\lambda_i} &\sim \text{Gamma} \left( 1 + a_{\lambda_i} t_i + b_{\lambda_i} \right), \\
\lambda_{2k}^2 | \eta_k, a_{\lambda_k}, b_{\lambda_k} &\sim \text{Gamma} \left( k + 1 + a_{\lambda_k} \eta_k + b_{\lambda_k} \right).
\end{align*}
\]

(5) Conditional distributions of \( t_i \) and \( \eta_k \):
\[
\begin{align*}
t_i^{-1} | \lambda_{1i}, \beta_i &\sim \left\{ \begin{array}{ll}
\text{Inverse Gamma} \left( 1, \frac{\lambda_{1i}^2}{2} \right), & \text{if } \beta_i = 0, \\
\text{Inverse Gaussian} \left( \sqrt{\frac{2\lambda_{1i}^2}{\lambda_i}}, \lambda_{1i} \right), & \text{if } \beta_i \neq 0,
\end{array} \right.
\end{align*}
\]
\[
\begin{align*}
\eta_k^{-1} | \lambda_{2k}, \gamma_k &\sim \left\{ \begin{array}{ll}
\text{Inverse Gamma} \left( 1, \frac{\lambda_{2k}^2}{2} \right), & \text{if } \gamma_k = 0, \\
\text{Inverse Gaussian} \left( \sqrt{\frac{2\lambda_{2k}^2}{\lambda_{2k}}}, \lambda_{2k} \right), & \text{if } \gamma_k \neq 0,
\end{array} \right.
\end{align*}
\]

where InverseGamma\((a, b)\) denotes a Inverse Gamma distribution with shape parameter \( a \) and scale parameter \( b \).

(6) Conditional distribution of \( \pi'_{0u} \) and \( \pi_0 \):
The full conditional distributions of $\pi_0'$ and $\pi_0$ are independent Beta distribution.

$$\pi_0'|\beta_i \sim beta(1 + I(\beta_i = 0), 2 - I(\beta_i = 0)),$$

$$\pi_0|\gamma \sim beta(1 + \sum_{k=1}^{q} I(\gamma_k = 0), 1 + q - \sum_{k=1}^{q} I(\gamma_k = 0)).$$

(7) Conditional distribution of $\gamma$:

Denote $F_q = (b_i^T \otimes z_i^q)I_q$, $\bar{\Sigma}_{ik} = \left[\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{F_{ik}F_{j} - F_{ik}F_{jk} - F_{ij}F_{ik}}{2v_{ij0}} + \eta_kI_k\right]^{-1}$, and let $F_{ijk}$ be the covariate vector corresponding to $\gamma_k$. Then a spike and slab distribution, $\mu_k = \hat{\mu}_{ik} + (1 - l_{ik})N(0, \bar{\Sigma}_{ik})I(\gamma_k \geq 0)$, where

$$l_{ik} = \frac{\pi_{ik}}{\pi_{ik} + 2(1 - \pi_{ik})\eta_k^{1/2} |\bar{\Sigma}_{ik}|^{1/2} \Phi(\frac{\hat{\mu}_{ik}}{\pi_{ik}^{1/2}})exp\left\{\frac{\hat{\mu}_{ik}^T \bar{\Sigma}_{ik}^{-1} \hat{\mu}_{ik}}{2}\right\}}.$$

Here, $\Phi(\cdot)$ indicates the cumulative distribution function of the standard normal distribution, $\hat{\mu}_{ik}$ is the $k$th element of $\hat{\mu}_{ik}$, and $\hat{\sigma}_{ik}^2$ is the $k$th diagonal element of $\bar{\Sigma}_{ik}$.

(8) Conditional distribution of $\sigma$:

The full conditional distribution of $\sigma$ is

$$Gamma\left(\frac{3N}{2} + a_s, \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{(y_{ij} - z_i^q x_i^T \beta_i)^2}{2v_{ij}} + \sum_{i=1}^{n} \sum_{j=1}^{n} v_{ij} + b_s\right).$$

**Simulation studies**

In this section we check the performance of the proposed extended Bayesian adaptive group lasso approach (adGL) and the Bayesian adaptive group lasso with independent spike and slab type priors approach (adSpikeGL) through simulations. We also compare them with other models: the PMQ method reported by [10], the BL and BAL methods proposed by [11] and the ordinary quantile regression estimator (QR) focusing solely on independent regression.

The test data are generated from the following model:

$$y_{ij} = x_i^q \beta_j + z_i^q \alpha_i + \epsilon_{ij},$$

for $i = 1, \ldots, 50, j = 1, \ldots, 5$, where $\epsilon_{ij}$ has the $r$th quantile equal to 0, random effects $\alpha_i \sim N(0, D)$, where

$$D = \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix}, D_1 = \begin{pmatrix} 9 & 4.8 & 0.6 \\ 4.8 & 4 & 1 \\ 0.6 & 1 & 1 \end{pmatrix}$$

(11)
Similar formulations of \( \mathbf{D} \) in mixed models have been used by [29] and [30]. We set \( \mathbf{x}'_\eta = (x_{\eta 1}, \ldots, x_{\eta 8}) \) and \( \mathbf{z}'_\eta = \mathbf{x}'_\eta \), where \( x_{\eta k}, k = 1, \ldots, 8 \) are drawn independently from the uniform \([-2, 2]\). The fixed effects parameters \( \beta = (\beta_1, \ldots, \beta_6) \) are set as follows:

- **Simulation 1:** \( \beta = (3, 1.5, 0, 0, 2, 0, 0)' \) to illustrate a sparse case;
- **Simulation 2:** \( \beta = (0.85, 0.85, 0.85, 0.85, 0.85, 0.85)' \) to illustrate a dense case;
- **Simulation 3:** \( \beta = (5, 0, 0, 0, 0, 0)' \) to illustrate a very sparse case.

With each simulation of \( \beta \), we generate data under three different quantiles i.e. \( \tau = 0.1, 0.3 \) and 0.5, and four different error distributions: normal, normal mixture, laplace, and laplace mixture as follows

- (1) \( \epsilon_{ij} \sim \mathcal{N}(0, 9), i = 1, \ldots, n; j = 1, \ldots, n_i; \)
- (2) \( \epsilon_{ij} \sim 0.1\mathcal{N}(0, 1) + 0.9\mathcal{N}(0, 5), i = 1, \ldots, n; j = 1, \ldots, n_i; \)
- (3) \( \epsilon_{ij} \sim \text{Laplace}(0, 3), i = 1, \ldots, n; j = 1, \ldots, n_i; \)
- (4) \( \epsilon_{ij} \sim 0.1\text{Laplace}(0, 1) + 0.9\text{Laplace}(0, \sqrt{5}), i = 1, \ldots, n; j = 1, \ldots, n_i; \)

We consider three sample sizes \( n = 30, n_1 = 5; n = 50, n_1 = 5; n = 100, n_1 = 5 \). Therefore, we have a total of 36 simulation setups for each of the sample sizes. We simulate 100 data sets for simulation setups. Priors for the Bayesian methods are taken to be weak prior. The hyperparameters in the Gamma priors for tuning and scale parameters are all set to be 0.1. We run our PCG sampler for 20000 following initial burn-in 10000 cycles. The convergence of the proposed PCG is monitored by using the multivariate potential scale reduction factor (MPSRF) introduced by [31]. Figs 1–6 show that the MPSRF generally become stable and get close to 1 after about 10000 iterations under normal mixture error distribution, suggesting that the above burn-in size is large enough to ensure the convergence of PCG. The results for other error distributions are similar and omitted.

We use the posterior median estimates of \( \beta \) and \( \mathbf{D} \) as our point estimator, denoted as \( \hat{\beta} \) and \( \hat{\mathbf{D}} \), respectively, and we consider two error measures including the mean absolute deviations (MAD), root mean squared errors (RMSE). To check the performance of the parameter estimation, we also calculate the median of mean square error for \( \beta \left( MME(\hat{\beta}) \right) \) and the median of quadratic loss error for \( \mathbf{D} \left( MMR(\hat{\mathbf{D}}) \right) \). More specifically,

\[
\begin{align*}
\text{MAD} & = \frac{1}{250} \sum_{i=1}^{50} \sum_{j=1}^{5} |y_{ij} - \hat{y}_{ij}|, \\
\text{RMSE} & = \sqrt{\frac{1}{250} \sum_{i=1}^{50} \sum_{j=1}^{5} (y_{ij} - \hat{y}_{ij})^2}, \\
\text{MME}(\hat{\beta}) & = (\hat{\beta} - \beta)^\top (\hat{\beta} - \beta), \\
\text{MMR}(\hat{\mathbf{D}}) & = \text{trace}(\hat{\mathbf{D}} - \mathbf{D})^2/2,
\end{align*}
\]

where \( \hat{y}_{ij} \) is the predicted value of \( y_{ij} \) and the median is taken over the 100 simulations. The results of MAD and RMSE are summarized in Figs 7–12, meanwhile \( MME(\hat{\beta}) \) and \( MMR(\hat{\mathbf{D}}) \) are listed in Tables 1–9. Because BL, BAL, and QR neglect the covariance structure of random effects, we just report the \( MMR(\hat{\mathbf{D}}) \) results of the other three techniques for comparison.

A few observations can be seen from the results. Firstly, it can be seen from Figs 7–12 that the proposed approaches perform well in general, the adGL generally outperforms the other
methods in terms of MAD and RMSE, which indicate explicitly penalizing the off-diagonal covariance components can improve the model selection procedure. Similar findings were also obtained by [32]. On the other hand, the adSpikeGL also has good performance, yet does not perform as well as the adGL. Secondly, Tables 1–9 show that adSpikeGL performs the best.
in terms of $MME(\beta)$ and $MMR(\hat{D})$. Take $n = 50$ for an example, adSpikeGL has the smallest $MME(\beta)$ in 22 out of 36 simulation setups, and the smallest $MMR(\hat{D})$ in 24 out of 36 simulation setups. Though PMQ performs good for $\tau = 0.1$, its performance drops when $\tau = 0.3$ and $\tau = 0.5$, especially in dense case (Simulation 2). Thirdly, Tables 1–9 also display that our
approaches work well in terms of MAD, RMSE, $MME(\hat{\beta})$ and $MMR(\tilde{D})$ even when the true error term is not ALD, indicating that the LQMMs using ALD in model (2) is merely a working model with artificial assumptions that aim to achieve equivalence between the problems in maximizing ALD and minimizing (3). Finally, as we can observe from Figs 7–12 and

Fig 3. MPSRF for adGL in Simulation 3, when normal mixture error distribution.

https://doi.org/10.1371/journal.pone.0241197.g003
Tables 1–9 that, in terms of the four criterions, adGL, adSpikeGl and PMQ all have an overall decreasing trend as the sample size increases, whereas BA, BAL and QR generally fail to estimate accurately, possibly because they almost completely ignore the random effects, which describe within-subject dependence among data, and their performance doesn’t get better as the sample size grows.

https://doi.org/10.1371/journal.pone.0241197.g004

Fig 4. MPSRF for adSpikeGL in Simulation 1, when normal mixture error distribution.
For adSpikeGL, the median posterior probability model (MPPM) criteria [33] and the highest posterior probability model (HPPM) criteria are used for model selection, respectively. Tables 10–14 and 11–15 summarize the number of correct and wrong zero coefficients for all cases based on MPPM and HPPM, respectively. On the whole, adSpikeGL generally yield...
promising results in most cases, even for extreme quantile when $\tau = 0.1$ with notoriously challenging. As the sample size grows from 30 to 100, the number of the wrong zero coefficients goes to zero at all considered quantiles, and the number of the correct zero coefficients also become large at moderate and median quantiles. These all indicates good performance of our method in variable selection.

Fig 6. MPSRF for adSpikeGL in Simulation 3, when normal mixture error distribution.

https://doi.org/10.1371/journal.pone.0241197.g006
Fig 7. Boxplots summarizing the MAD for three sample sizes in Simulation 1. The rows from top to bottom correspond to normal, normal mixture, laplace, and laplace mixture error distributions respectively.

https://doi.org/10.1371/journal.pone.0241197.g007
Fig 8. Boxplots summarizing the RMSE for three sample sizes in Simulation 1. The rows from top to bottom correspond to normal, normal mixture, laplace, and laplace mixture error distributions respectively.

https://doi.org/10.1371/journal.pone.0241197.g008
Fig 9. Boxplots summarizing the MAD for three sample sizes in Simulation 2. The rows from top to bottom correspond to normal, normal mixture, laplace, and laplace mixture error distributions respectively.

https://doi.org/10.1371/journal.pone.0241197.g009
Fig 10. Boxplots summarizing the RMSE for three sample sizes in Simulation 2. The rows from top to bottom correspond to normal, normal mixture, laplace, and laplace mixture error distributions respectively.

https://doi.org/10.1371/journal.pone.0241197.g010
Fig 11. Boxplots summarizing the MAD for three sample sizes in Simulation 3. The rows from top to bottom correspond to normal, normal mixture, laplace, and laplace mixture error distributions respectively.

https://doi.org/10.1371/journal.pone.0241197.g011
Fig 12. Boxplots summarizing the RMSE for three sample sizes in Simulation 3. The rows from top to bottom correspond to normal, normal mixture, laplace, and laplace mixture error distributions respectively.

https://doi.org/10.1371/journal.pone.0241197.g012
Table 1. Median of mean square errors (MME(\(\hat{\beta}\))) and quadratic loss errors (MMR(\(\hat{D}\))) at \(\tau = 0.1\) for Simulation 1. MMR(\(\hat{D}\)) are in parentheses.

| Sample size | Method       | Error distribution |
|-------------|--------------|---------------------|
|             |              | Normal              | Normalmixture | Laplace   | Laplacemixture |
| n = 30      | adGL         | 3.17(11.314)        | 1.783(9.209)  | 4.034(14.133) | 2.273(11.656)  |
|             | adSpikeGL    | 3.535(11.089)       | 1.475(9.02)   | 6.26(20.161)   | 2.409(12.694)   |
|             | PMQ          | 5.017(11.124)       | 4.127(9.899)  | 5.494(14.108)  | 4.364(12.485)   |
|             | BL           | 5.635(-)            | 5.919(-)      | 6.404(-)       | 7.011(-)        |
|             | BAL          | 5.377(-)            | 5.708(-)      | 5.61(-)        | 6.155(-)        |
|             | QR           | 9.964(-)            | 11.953(-)     | 10.803(-)      | 12.641(-)       |
| n = 50      | adGL         | 2.645(9.065)        | 1.508(7.383)  | 3.508(12.752)  | 2.002(9.291)    |
|             | adSpikeGL    | 2.102(9.102)        | 0.944(6.497)  | 3.443(14.761)  | 1.627(9.121)    |
|             | PMQ          | 2.643(8.517)        | 1.796(6.588)  | 3.151(11.198)  | 2.089(8.981)    |
|             | BL           | 4.422(-)            | 5.452(-)      | 4.114(-)       | 5.35(-)         |
|             | BAL          | 4.011(-)            | 4.763(-)      | 3.661(-)       | 4.736(-)        |
|             | QR           | 7.975(-)            | 10.489(-)     | 8.198(-)       | 10.686(-)       |
| n = 100     | adGL         | 2.125(7.452)        | 1.133(5.182)  | 3.094(11.117)  | 1.481(7.841)    |
|             | adSpikeGL    | 1.601(6.809)        | 0.802(4.374)  | 2.692(12.212)  | 1.164(7.423)    |
|             | PMQ          | 1.825(6.597)        | 1.019(4.745)  | 2.526(10.248)  | 1.192(6.914)    |
|             | BL           | 3.627(-)            | 4.737(-)      | 3.462(-)       | 4.885(-)        |
|             | BAL          | 3.2(-)              | 4.125(-)      | 2.923(-)       | 4.21(-)         |
|             | QR           | 7.128(-)            | 9.408(-)      | 7.445(-)       | 9.278(-)        |

https://doi.org/10.1371/journal.pone.0241197.t001

Table 2. Median of mean square errors (MME(\(\hat{\beta}\))) and quadratic loss errors (MMR(\(\hat{D}\))) at \(\tau = 0.3\) for Simulation 1. MMR(\(\hat{D}\)) are in parentheses.

| Sample size | Method       | Error distribution |
|-------------|--------------|---------------------|
|             |              | Normal              | Normalmixture | Laplace   | Laplacemixture |
| n = 30      | adGL         | 0.958(11.408)       | 0.607(9.013)  | 1.447(13.115) | 0.949(10.876)  |
|             | adSpikeGL    | 1.394(7.827)        | 0.85(7.501)   | 2.456(11.005) | 1.572(9.061)   |
|             | PMQ          | 4.366(10.095)       | 3.825(10.022) | 5.127(13.1)   | 4.299(11.501)  |
|             | BL           | 2.838(-)            | 2.763(-)      | 3.785(-)      | 3.697(-)       |
|             | BAL          | 2.576(-)            | 2.641(-)      | 3.761(-)      | 3.615(-)       |
|             | QR           | 3.001(-)            | 3.068(-)      | 3.866(-)      | 3.842(-)       |
| n = 50      | adGL         | 0.543(8.77)         | 0.357(7.495)  | 0.736(10.75)  | 0.516(8.6)     |
|             | adSpikeGL    | 0.337(5.917)        | 0.241(4.7)    | 0.521(7.668)  | 0.432(6.034)   |
|             | PMQ          | 1.395(6.996)        | 0.975(5.921)  | 1.402(8.938)  | 1.441(6.762)   |
|             | BL           | 1.743(-)            | 1.974(-)      | 2.748(-)      | 2.54(-)        |
|             | BAL          | 1.736(-)            | 1.863(-)      | 2.637(-)      | 2.506(-)       |
|             | QR           | 1.978(-)            | 2.062(-)      | 2.862(-)      | 2.606(-)       |
| n = 100     | adGL         | 0.36(6.429)         | 0.191(5.149)  | 0.441(8.083)  | 0.313(6.55)    |
|             | adSpikeGL    | 0.198(4.973)        | 0.133(3.585)  | 0.298(6.177)  | 0.28(3.93)     |
|             | PMQ          | 0.373(5.299)        | 0.305(3.939)  | 0.576(6.255)  | 0.388(4.797)   |
|             | BL           | 1.289(-)            | 1.567(-)      | 1.951(-)      | 2.109(-)       |
|             | BAL          | 1.243(-)            | 1.469(-)      | 1.859(-)      | 1.984(-)       |
|             | QR           | 1.368(-)            | 1.544(-)      | 2.095(-)      | 2.1(-)        |

https://doi.org/10.1371/journal.pone.0241197.t002
Table 3. Median of mean square errors (MME($\hat{\beta}$)) and quadratic loss errors (MMR($\hat{D}$)) at $\tau = 0.5$ for Simulation 1. MMR($\hat{D}$) are in parentheses.

| Sample size | Method   | Error distribution |
|-------------|----------|---------------------|
|             |          | Normal | Normalmixtrue | Laplace | Laplacemixtrue |
| n = 30      | adGL     | 0.924(11.166) | 0.581(8.938) | 1.278(13.494) | 0.862(11.367) |
|             | adSpikeGL| 1.377(7.444) | 0.732(6.637) | 1.736(8.771) | 1.181(7.687)  |
|             | PMQ      | 4.216(10.914) | 3.594(10.373) | 4.813(13.136) | 4.121(10.363) |
|             | BL       | 1.729(-)   | 1.276(-)     | 1.968(-)     | 1.596(-)     |
|             | BAL      | 1.734(-)   | 1.289(-)     | 1.923(-)     | 1.633(-)     |
|             | QR       | 2.032(-)   | 1.517(-)     | 2.223(-)     | 1.88(-)      |
| n = 50      | adGL     | 0.421(8.836) | 0.326(7.586) | 0.605(10.555) | 0.433(8.434) |
|             | adSpikeGL| 0.317(6.044) | 0.238(5.088) | 0.399(6.996) | 0.283(6.042) |
|             | PMQ      | 1.327(7.136) | 0.993(5.648) | 1.324(7.431) | 1.409(6.689) |
|             | BL       | 0.982(-)   | 0.823(-)     | 1.086(-)     | 0.912(-)     |
|             | BAL      | 1.007(-)   | 0.831(-)     | 1.109(-)     | 0.906(-)     |
|             | QR       | 1.157(-)   | 0.923(-)     | 1.294(-)     | 1.102(-)     |
| n = 100     | adGL     | 0.286(6.651) | 0.185(5.065) | 0.311(5.712) | 0.207(6.058) |
|             | adSpikeGL| 0.197(4.55) | 0.128(3.223) | 0.223(5.378) | 0.134(4.589) |
|             | PMQ      | 0.344(5.167) | 0.296(3.711) | 0.472(5.622) | 0.3(4.821)   |
|             | BL       | 0.471(-)   | 0.402(-)     | 0.573(-)     | 0.469(-)     |
|             | BAL      | 0.5(-)     | 0.423(-)     | 0.582(-)     | 0.478(-)     |
|             | QR       | 0.541(-)   | 0.456(-)     | 0.634(-)     | 0.515(-)     |

Next, in order to evaluate the sensitivity of our methods, we consider the following four different priors of tuning and scale parameters for Simulation 1: case(1) $a_{s_i} = b_{s_i} = a_{s_i} = b_{s_i} = 0.1$, $i = 1, 2$, case(2) $a_{s_i} = b_{s_i} = a_{s_i} = b_{s_i} = 0.01$, $i = 1, 2$, case(3) $a_{s_i} = b_{s_i} = a_{s_i} = b_{s_i} = 0.0001$, $i = 1, 2$. From

Table 4. Median of mean square errors (MME($\hat{\beta}$)) and quadratic loss errors (MMR($\hat{D}$)) at $\tau = 0.1$ for Simulation 2. MMR($\hat{D}$) are in parentheses.

| Sample size | Method   | Error distribution |
|-------------|----------|---------------------|
|             |          | Normal | Normalmixtrue | Laplace | Laplacemixtrue |
| n = 30      | adGL     | 3.28(12.016) | 2.091(9.127) | 4.671(14.802) | 2.492(10.868) |
|             | adSpikeGL| 4.802(13.526) | 2.63(8.487) | 7.702(18.516) | 3.663(12.025) |
|             | PMQ      | 2.796(9.328) | 1.919(7.995) | 3.461(11.497) | 2.477(8.897)  |
|             | BL       | 5.359(-)   | 6.499(-)     | 5.447(-)     | 6.328(-)     |
|             | BAL      | 5.214(-)   | 5.98(-)      | 5.255(-)     | 5.842(-)     |
|             | QR       | 9.121(-)   | 11.535(-)    | 10.766(-)    | 11.09(-)     |
| n = 50      | adGL     | 2.639(9.979) | 1.541(7.654) | 3.914(13.945) | 2.098(10.088) |
|             | adSpikeGL| 3.117(9.197) | 1.909(6.214) | 4.98(16.776) | 2.641(10.728) |
|             | PMQ      | 2.19(8.009) | 1.426(6.192) | 2.884(12.507) | 1.769(8.396)  |
|             | BL       | 4.36(-)   | 5.097(-)     | 4.458(-)     | 5.872(-)     |
|             | BAL      | 3.97(-)   | 4.407(-)     | 4.165(-)     | 5.476(-)     |
|             | QR       | 8.105(-)   | 9.584(-)     | 9.089(-)     | 10.387(-)    |
| n = 100     | adGL     | 2.237(7.876) | 1.279(5.596) | 3.163(12.262) | 1.446(7.97)   |
|             | adSpikeGL| 2.364(6.951) | 1.296(5.048) | 3.708(13.513) | 1.772(8.228)  |
|             | PMQ      | 1.796(6.807) | 1.097(4.709) | 2.515(10.873) | 1.247(7.199)  |
|             | BL       | 3.476(-)   | 5.193(-)     | 3.526(-)     | 4.909(-)     |
|             | BAL      | 3.106(-)   | 4.431(-)     | 2.995(-)     | 4.17(-)      |
|             | QR       | 7.104(-)   | 9.498(-)     | 7.716(-)     | 9.371(-)     |

https://doi.org/10.1371/journal.pone.0241197.t003

https://doi.org/10.1371/journal.pone.0241197.t004
### Table 5. Median of mean square errors ($\text{MME}(\hat{\beta})$) and quadratic loss errors ($\text{MMR}(\hat{D})$) at $\tau = 0.3$ for Simulation 2. $\text{MMR}(\hat{D})$ are in parentheses.

| Sample size | Method      | Error distribution |
|-------------|-------------|--------------------|
|             |             | Normal             |
|             |             | Normalmixtrue      |
|             |             | Laplace             |
|             |             | Laplacemixtrue      |
| n = 30      | adGL        | 1.337(11.777)      | 0.948(8.895) | 1.836(13.974) | 1.212(10.574) |
|             | adSpikeGL   | 3.095(10.405)      | 1.849(7.039) | 4.022(13.495) | 2.494(8.874)  |
|             | PMQ         | 1.985(10.15)       | 1.423(7.629) | 2.495(11.006) | 1.765(7.828)  |
|             | BL          | 2.722(-)           | 2.548(-)     | 3.531(-)      | 3.35(-)       |
|             | BAL         | 2.678(-)           | 2.477(-)     | 3.514(-)      | 3.323(-)      |
|             | QR          | 3.2(-)             | 2.831(-)     | 3.648(-)      | 3.539(-)      |
| n = 50      | adGL        | 0.773(9.367)       | 0.51(7.751)  | 1.092(10.676) | 0.837(8.88)   |
|             | adSpikeGL   | 1.78(6.842)        | 1.363(5.378) | 2.569(8.632)  | 1.692(5.98)   |
|             | PMQ         | 1.091(7.026)       | 0.819(5.406) | 1.614(8.468)  | 1.113(6.712)  |
|             | BL          | 1.831(-)           | 1.923(-)     | 2.918(-)      | 2.76(-)       |
|             | BAL         | 1.787(-)           | 1.874(-)     | 2.848(-)      | 2.702(-)      |
|             | QR          | 1.989(-)           | 1.976(-)     | 3.115(-)      | 2.905(-)      |
| n = 100     | adGL        | 0.37(6.871)        | 0.225(5.369) | 0.561(7.846)  | 0.328(6.736)  |
|             | adSpikeGL   | 0.871(4.78)        | 0.546(3.421) | 1.137(5.862)  | 0.789(4.451)  |
|             | PMQ         | 0.515(5.01)        | 0.347(3.739) | 0.703(6.179)  | 0.468(5.179)  |
|             | BL          | 1.253(-)           | 1.61(-)      | 1.952(-)      | 2.022(-)      |
|             | BAL         | 1.183(-)           | 1.483(-)     | 1.848(-)      | 1.956(-)      |
|             | QR          | 1.339(-)           | 1.627(-)     | 2.083(-)      | 2.098(-)      |

### Table 6. Median of mean square errors ($\text{MME}(\hat{\beta})$) and quadratic loss errors ($\text{MMR}(\hat{D})$) at $\tau = 0.5$ for Simulation 2. $\text{MMR}(\hat{D})$ are in parentheses.

| Sample size | Method      | Error distribution |
|-------------|-------------|--------------------|
|             |             | Normal             |
|             |             | Normalmixtrue      |
|             |             | Laplace             |
|             |             | Laplacemixtrue      |
| n = 30      | adGL        | 1.242(11.082)      | 0.903(9.625) | 1.544(13.627) | 1.089(10.873) |
|             | adSpikeGL   | 2.824(8.786)       | 1.797(6.892) | 3.649(10.828) | 2.247(8.175)  |
|             | PMQ         | 1.894(9.806)       | 1.379(7.22)  | 2.419(11.003) | 1.6(7.938)    |
|             | BL          | 1.668(-)           | 1.266(-)     | 2.046(-)      | 1.502(-)      |
|             | BAL         | 1.645(-)           | 1.267(-)     | 2.005(-)      | 1.508(-)      |
|             | QR          | 2.074(-)           | 1.406(-)     | 2.384(-)      | 1.864(-)      |
| n = 50      | adGL        | 0.74(9.271)        | 0.502(7.87)  | 0.953(10.308) | 0.667(8.587)  |
|             | adSpikeGL   | 1.636(6.982)       | 1.288(5.22)  | 2.221(7.434)  | 1.581(6.003)  |
|             | PMQ         | 1.033(6.806)       | 0.802(5.529) | 1.352(7.545)  | 0.99(6.257)   |
|             | BL          | 1.023(-)           | 0.845(-)     | 1.298(-)      | 0.945(-)      |
|             | BAL         | 1.008(-)           | 0.871(-)     | 1.312(-)      | 0.923(-)      |
|             | QR          | 1.121(-)           | 0.959(-)     | 1.528(-)      | 1.007(-)      |
| n = 100     | adGL        | 0.33(6.645)        | 0.213(5.253) | 0.422(7.898)  | 0.265(6.053)  |
|             | adSpikeGL   | 0.843(4.707)       | 0.644(3.548) | 0.948(5.46)   | 0.77(4.39)    |
|             | PMQ         | 0.504(5.029)       | 0.334(3.65)  | 0.58(5.615)   | 0.412(4.861)  |
|             | BL          | 0.51(-)            | 0.363(-)     | 0.57(-)       | 0.429(-)      |
|             | BAL         | 0.536(-)           | 0.389(-)     | 0.554(-)      | 0.43(-)       |
|             | QR          | 0.55(-)            | 0.446(-)     | 0.632(-)      | 0.469(-)      |
Table 7. Median of mean square errors (MME(\(\hat{\beta}\))) and quadratic loss errors (MMR(\(\hat{D}\))) at \(\tau = 0.1\) for Simulation 3. MMR(\(\hat{D}\)) are in parentheses.

| Sample size | Method | Error distribution |
|-------------|--------|--------------------|
|             | Normal | Normalmixtrue | Laplace | Laplacemixtrue |
| n = 30      | adGL   | 2.922(11.216)   | 1.577(9.453) | 3.801(14.215) | 1.962(11.422) |
|             | adSpikeGL | 1.284(11.385) | 0.509(7.829) | 3.764(15.632) | 1.044(10.661) |
|             | PMQ    | 2.045(9.833)   | 1.358(8.179) | 3.375(12.93)  | 1.861(9.465)  |
|             | BL     | 5.425(-)       | 6.815(-)     | 6.204(-)      | 6.326(-)      |
|             | BAL    | 5.426(-)       | 6.208(-)     | 6.357(-)      | 5.936(-)      |
|             | QR     | 9.733(-)       | 11.759(-)    | 11.056(-)     | 10.68(-)      |
| n = 50      | adGL   | 2.584(9.042)   | 1.231(7.481) | 2.929(12.619) | 1.639(9.513)  |
|             | adSpikeGL | 1.545(8.405) | 0.616(7.829) | 2.372(14.632) | 1.079(9.174)  |
|             | PMQ    | 1.931(7.836)   | 1.098(6.659) | 2.336(10.67)  | 1.295(8.154)  |
|             | BL     | 4.411(-)       | 5.516(-)     | 4.747(-)      | 5.774(-)      |
|             | BAL    | 3.947(-)       | 5.067(-)     | 4.248(-)      | 5.313(-)      |
|             | QR     | 8.333(-)       | 10.489(-)    | 9.016(-)      | 11.041(-)     |
| n = 100     | adGL   | 2.09(7.332)    | 1.033(5.836) | 2.619(11.603) | 1.399(7.948)  |
|             | adSpikeGL | 1.538(6.93)  | 0.632(5.038) | 2.104(13.169) | 1.037(9.819)  |
|             | PMQ    | 1.662(6.521)   | 0.842(4.998) | 2.018(10.587) | 1.097(7.047)  |
|             | BL     | 3.592(-)       | 5.101(-)     | 3.939(-)      | 4.656(-)      |
|             | BAL    | 3.106(-)       | 4.404(-)     | 3.438(-)      | 4.084(-)      |
|             | QR     | 6.751(-)       | 9.933(-)     | 8.013(-)      | 9.714(-)      |

https://doi.org/10.1371/journal.pone.0241197.t007

Table 8. Median of mean square errors (MME(\(\hat{\beta}\))) and quadratic loss errors (MMR(\(\hat{D}\))) at \(\tau = 0.3\) for Simulation 3. MMR(\(\hat{D}\)) are in parentheses.

| Sample size | Method | Error distribution |
|-------------|--------|--------------------|
|             | Normal | Normalmixtrue | Laplace | Laplacemixtrue |
| n = 30      | adGL   | 0.693(11.078)   | 0.426(9.632) | 0.935(13.414) | 0.608(11.588) |
|             | adSpikeGL | 0.207(8.688)  | 0.196(6.897) | 0.326(9.916)  | 0.236(8.558)  |
|             | PMQ    | 0.916(9.637)   | 0.757(7.408) | 1.352(11.942) | 1.012(8.875)  |
|             | BL     | 2.838(-)       | 2.522(-)     | 3.815(-)      | 3.22(-)       |
|             | BAL    | 2.659(-)       | 2.599(-)     | 3.799(-)      | 3.089(-)      |
|             | QR     | 2.977(-)       | 2.633(-)     | 3.83(-)       | 3.15(-)       |
| n = 50      | adGL   | 0.489(8.883)   | 0.325(7.26)  | 0.609(10.14)  | 0.479(8.714)  |
|             | adSpikeGL | 0.166(5.899)  | 0.13(5.429)  | 0.265(6.91)   | 0.179(5.89)   |
|             | PMQ    | 0.523(6.773)   | 0.414(5.967) | 0.826(7.842)  | 0.587(6.097)  |
|             | BL     | 1.971(-)       | 2.242(-)     | 2.728(-)      | 2.757(-)      |
|             | BAL    | 1.899(-)       | 2.106(-)     | 2.666(-)      | 2.663(-)      |
|             | QR     | 2.092(-)       | 2.444(-)     | 3.06(-)       | 2.78(-)       |
| n = 100     | adGL   | 0.261(6.807)   | 0.15(5.202)  | 0.394(7.977)  | 0.254(6.402)  |
|             | adSpikeGL | 0.066(4.502)  | 0.031(3.74)  | 0.182(5.552)  | 0.123(4.076)  |
|             | PMQ    | 0.251(5.941)   | 0.152(3.931) | 0.403(5.937)  | 0.289(4.513)  |
|             | BL     | 1.231(-)       | 1.557(-)     | 2.219(-)      | 2.06(-)       |
|             | BAL    | 1.123(-)       | 1.454(-)     | 2.092(-)      | 1.823(-)      |
|             | QR     | 1.338(-)       | 1.717(-)     | 2.297(-)      | 2.1(-)        |

https://doi.org/10.1371/journal.pone.0241197.t008
Table 9. Median of mean square errors (MME(\(\hat{\beta}\))) and quadratic loss errors (MMR(\(\hat{D}\))) at \(\tau = 0.5\) for Simulation 3. MMR(\(\hat{D}\)) are in parentheses.

| Sample size | Method     | Normal | Normalmixtrue | Laplace | Laplacemixtrue |
|-------------|------------|--------|---------------|---------|----------------|
| n = 30      | adGL       | 0.586(11.256) | 0.415(9.501) | 0.822(12.716) | 0.564(11.175)   |
|             | adSpikeGL  | 0.195(7.985)  | 0.172(6.351) | 0.246(9.296)  | 0.183(8.295)    |
|             | PMQ        | 0.912(9.032)  | 0.737(7.37)  | 0.952(11.751) | 0.918(8.996)    |
|             | BL         | 1.733(-)      | 1.303(-)     | 1.955(-)      | 1.401(-)        |
|             | BAL        | 1.729(-)      | 1.282(-)     | 1.953(-)      | 1.445(-)        |
|             | QR         | 2.024(-)      | 1.479(-)     | 2.306(-)      | 1.686(-)        |
| n = 50      | adGL       | 0.428(8.755)  | 0.309(7.215) | 0.453(10.334) | 0.394(8.34)     |
|             | adSpikeGL  | 0.162(5.735)  | 0.127(5.397) | 0.164(6.863)  | 0.101(5.69)     |
|             | PMQ        | 0.447(6.472)  | 0.4(5.933)   | 0.663(7.408)  | 0.404(5.941)    |
|             | BL         | 0.991(-)      | 0.886(-)     | 1.09(-)       | 0.925(-)        |
|             | BAL        | 0.99(-)       | 0.899(-)     | 1.111(-)      | 0.892(-)        |
|             | QR         | 1.193(-)      | 1.006(-)     | 1.314(-)      | 1.02(-)         |
| n = 100     | adGL       | 0.217(6.604)  | 0.143(5.139) | 0.239(7.822)  | 0.166(5.982)    |
|             | adSpikeGL  | 0.068(4.448)  | 0.032(3.543) | 0.053(6.163)  | 0.047(4.268)    |
|             | PMQ        | 0.231(4.655)  | 0.133(3.832) | 0.256(5.194)  | 0.169(5.527)    |
|             | BL         | 0.463(-)      | 0.374(-)     | 0.53(-)       | 0.445(-)        |
|             | BAL        | 0.496(-)      | 0.377(-)     | 0.53(-)       | 0.45(-)         |
|             | QR         | 0.523(-)      | 0.413(-)     | 0.647(-)      | 0.48(-)         |

Table 10. Average numbers of correct and wrong zero coefficients for fixed effects and random effects in terms of MPPM for Simulation 1. Random effects are in parentheses.

| Sample size | \(\tau\) | Error distribution |
|-------------|---------|--------------------|
| n = 30      | p = .1  | correct 4.45(4)     |
|             |         | wrong 0.14(0.4)     |
|             | p = .3  | correct 5.29(4.92)  |
|             |         | wrong 0.12(0.41)    |
|             | p = .5  | correct 5.29(5.42)  |
|             |         | wrong 0.11(0.45)    |
| n = 50      | p = .1  | correct 4.44(4.13)  |
|             |         | wrong 0(0.16)       |
|             | p = .3  | correct 5.48(5.45)  |
|             |         | wrong 0.01(0.21)    |
|             | p = .5  | correct 5.58(5.78)  |
|             |         | wrong 0.0(0.35)     |
| n = 100     | p = .1  | correct 4.53(3.5)   |
|             |         | wrong 0.0(0.08)     |
|             | p = .3  | correct 5.43(5.42)  |
|             |         | wrong 0(0.11)       |
|             | p = .5  | correct 5.4(5.78)   |
|             |         | wrong 0(0.16)       |

https://doi.org/10.1371/journal.pone.0241197.t009

https://doi.org/10.1371/journal.pone.0241197.t010
Table 11. Average number of correct and wrong zero coefficients for fixed effects and random effects in terms of HPPM for Simulation 1. Random effects are in parentheses.

| Sample size | \( r \) | Error distribution |
|-------------|----------|---------------------|
|             |          | Normal | Normalmixtrue | Laplace | Laplacemixtrue |
| n = 30      | \( p = .1 \) | correct | 4.43(4.12) | 4.61(4.83) | 4.5(2.27) | 4.53(3.78) |
|             |          | wrong  | 0.13(0.38) | 0.08(0.35) | 0.22(0.32) | 0.08(0.35) |
|             | \( p = .3 \) | correct | 5.27(5.07) | 5.34(5.52) | 5.28(3.8) | 5.19(4.8) |
|             |          | wrong  | 0.11(0.47) | 0.09(0.33) | 0.16(0.47) | 0.09(0.41) |
|             | \( p = .5 \) | correct | 5.25(5.4) | 5.42(5.8) | 5.33(5.42) | 5.48(5.59) |
|             |          | wrong  | 0.11(0.5) | 0.11(0.43) | 0.17(0.76) | 0.07(0.53) |
| n = 50      | \( p = .1 \) | correct | 4.46(4.13) | 4.53(4.63) | 4.47(2.97) | 4.66(3.87) |
|             |          | wrong  | 0.18(0.13) | 0.11(0.13) | 0.04(0.32) | 0.02(0.29) |
|             | \( p = .3 \) | correct | 5.47(5.42) | 5.37(5.65) | 5.25(4.86) | 5.2(5.46) |
|             |          | wrong  | 0.23(0.14) | 0.11(0.14) | 0.11(0.44) | 0.11(0.37) |
|             | \( p = .5 \) | correct | 5.57(5.76) | 5.43(5.84) | 5.59(5.7) | 5.48(5.89) |
|             |          | wrong  | 0.36(0.18) | 0.11(0.18) | 0.59(0.39) | 0.46(0.3) |
| n = 100     | \( p = .1 \) | correct | 4.54(3.34) | 4.52(3.82) | 4.52(2.44) | 4.67(2.9) |
|             |          | wrong  | 0.07(0.01) | 0(0.01) | 0.1(0.1) | 0(0.06) |
|             | \( p = .3 \) | correct | 5.41(5.43) | 5.46(5.59) | 5.05(5.18) | 4.95(5.68) |
|             |          | wrong  | 0.11(0.01) | 0(0.01) | 0.28(0.0) | 0(0.06) |
|             | \( p = .5 \) | correct | 5.4(5.78) | 5.55(5.92) | 5.6(5.9) | 5.68(5.97) |
|             |          | wrong  | 0.2(0.1) | 0(0) | 0.5(0) | 0(0.13) |

Table 12. Average number of correct and wrong zero coefficients for fixed effects and random effects in terms of MPPM for Simulation 2. Random effects are in parentheses.

| Sample size | \( r \) | Error distribution |
|-------------|----------|---------------------|
|             |          | Normal | Normalmixtrue | Laplace | Laplacemixtrue |
| n = 30      | \( p = .1 \) | correct | 0.11(3.68) | 0.25(4.86) | 0.01(1.85) | 0.21(3.72) |
|             |          | wrong  | 2.93(0.27) | 1.78(0.29) | 3.72(0.29) | 2.64(0.31) |
|             | \( p = .3 \) | correct | 0.87(4.27) | 0.88(5.46) | 0.88(2.32) | 0.64(4.54) |
|             |          | wrong  | 2.54(0.31) | 1.54(0.29) | 3.59(0.24) | 2.27(0.26) |
|             | \( p = .5 \) | correct | 0.94(0.89) | 0.92(5.76) | 0.94(4.05) | 0.92(5.4) |
|             |          | wrong  | 2.27(0.43) | 1.48(0.36) | 3.07(0.48) | 1.99(0.46) |
| n = 50      | \( p = .1 \) | correct | 0.04(3.99) | 0.14(4.71) | 0.04(2.63) | 0.15(3.51) |
|             |          | wrong  | 1.45(0.26) | 1.07(0.12) | 2.33(0.23) | 1.55(0.22) |
|             | \( p = .3 \) | correct | 0.87(5.36) | 0.86(5.64) | 0.61(4.21) | 0.49(5.41) |
|             |          | wrong  | 1.31(0.29) | 1(0.12) | 1.88(0.36) | 1.31(0.26) |
|             | \( p = .5 \) | correct | 0.93(5.61) | 0.94(5.89) | 0.93(5.44) | 0.94(5.91) |
|             |          | wrong  | 1.26(0.35) | 0.93(0.18) | 1.55(0.52) | 1.16(0.39) |
| n = 100     | \( p = .1 \) | correct | 0.01(3.19) | 0.36(3.67) | 0.25(2.25) | 0.02(2.93) |
|             |          | wrong  | 0.5(0.12) | 0.47(0.01) | 0.79(0.16) | 0.43(0.05) |
|             | \( p = .3 \) | correct | 0.9(5.41) | 0.9(5.7) | 0.48(5.05) | 0.32(5.59) |
|             |          | wrong  | 0.5(0.1) | 0.48(0.01) | 0.61(0.27) | 0.35(0.07) |
|             | \( p = .5 \) | correct | 0.93(5.76) | 0.95(5.89) | 0.93(5.91) | 0.91(5.92) |
|             |          | wrong  | 0.51(0.22) | 0.48(0.0) | 0.55(0.4) | 0.3(0.16) |
### Table 13. Average numbers of correct and wrong zero coefficients for fixed effects and random effects in terms of HPPM for Simulation 2. Random effects are in parentheses.

| Sample size | \( r \) | Error distribution |
|-------------|--------|--------------------|
|             |        | Normal | Normalmixtrue | Laplace | Laplacemixtrue |
| \( n = 30 \) | \( p = .1 \) | correct | 0.1(3.66) | 0.25(4.81) | 0.01(1.82) | 0.2(3.71) |
|             |        | wrong  | 2.97(0.3) | 1.64(0.28) | 3.65(0.29) | 2.59(0.32) |
|             | \( p = .3 \) | correct | 0.89(4.32) | 0.88(5.4) | 0.83(2.35) | 0.66(4.63) |
|             |        | wrong  | 2.38(0.34) | 1.5(0.3) | 3.47(0.26) | 2.16(0.34) |
|             | \( p = .5 \) | correct | 0.9(4.95) | 0.94(5.78) | 0.88(4.01) | 0.9(5.46) |
|             |        | wrong  | 2.35(0.51) | 1.5(0.39) | 2.96(0.52) | 1.95(0.49) |
| \( n = 50 \) | \( p = .1 \) | correct | 0.04(4.03) | 0.14(4.72) | 0.04(2.65) | 0.14(3.56) |
|             |        | wrong  | 1.34(0.25) | 0.97(0.13) | 2.24(0.23) | 1.56(0.21) |
|             | \( p = .3 \) | correct | 0.85(5.39) | 0.88(5.67) | 0.6(4.26) | 0.49(5.48) |
|             |        | wrong  | 1.26(0.31) | 0.89(0.11) | 1.79(0.36) | 1.28(0.3) |
|             | \( p = .5 \) | correct | 0.93(5.66) | 0.93(5.89) | 0.93(4.56) | 0.93(5.92) |
|             |        | wrong  | 1.17(0.37) | 0.85(0.18) | 1.47(0.57) | 1.15(0.4) |
| \( n = 100 \) | \( p = .1 \) | correct | 0.01(3.2) | 0(3.66) | 0(2.51) | 0.02(3) |
|             |        | wrong  | 0.45(0.12) | 0.46(0.01) | 0.75(0.16) | 0.34(0.06) |
|             | \( p = .3 \) | correct | 0.9(5.36) | 0.9(5.7) | 0.49(5.11) | 0.32(5.6) |
|             |        | wrong  | 0.46(0.13) | 0.44(0.01) | 0.61(0.3) | 0.32(0.07) |
|             | \( p = .5 \) | correct | 0.93(5.77) | 0.95(5.9) | 0.93(5.93) | 0.91(5.92) |
|             |        | wrong  | 0.44(0.22) | 0.44(0) | 0.54(0.42) | 0.25(0.16) |

https://doi.org/10.1371/journal.pone.0241197.t013

### Table 14. Average numbers of correct and wrong zero coefficients for fixed effects and random effects in terms of MPPM for Simulation 3. Random effects are in parentheses.

| Sample size | \( r \) | Error distribution |
|-------------|--------|--------------------|
|             |        | Normal | Normalmixtrue | Laplace | Laplacemixtrue |
| \( n = 30 \) | \( p = .1 \) | correct | 6.11(4.04) | 6.51(4.91) | 6.18(2.5) | 6.27(4.09) |
|             |        | wrong  | 0(0.43) | 0(0.25) | 0(0.29) | 0(0.43) |
|             | \( p = .3 \) | correct | 6.96(4.8) | 7.24(5.5) | 7(3.77) | 6.97(5.04) |
|             |        | wrong  | 0(0.37) | 0(0.21) | 0(0.39) | 0(0.39) |
|             | \( p = .5 \) | correct | 6.99(5.37) | 7.32(5.83) | 7.13(5.31) | 7.2(5.76) |
|             |        | wrong  | 0(0.47) | 0(0.25) | 0(0.66) | 0(0.55) |
| \( n = 50 \) | \( p = .1 \) | correct | 6.22(4) | 6.1(4.73) | 6.32(3.05) | 6.39(3.83) |
|             |        | wrong  | 0(0.21) | 0(0.12) | 0(0.41) | 0(0.27) |
|             | \( p = .3 \) | correct | 7.12(5.32) | 7.08(5.73) | 6.92(5.07) | 6.91(5.51) |
|             |        | wrong  | 0(0.22) | 0(0.1) | 0(0.46) | 0(0.31) |
|             | \( p = .5 \) | correct | 7.15(5.63) | 7.25(5.88) | 7.41(5.83) | 7.39(5.95) |
|             |        | wrong  | 0(0.38) | 0(0.2) | 0(0.62) | 0(0.37) |
| \( n = 100 \) | \( p = .1 \) | correct | 6.36(3.34) | 6.47(3.73) | 6.49(2.37) | 6.54(2.94) |
|             |        | wrong  | 0(0.05) | 0(0.01) | 0(0.16) | 0(0.11) |
|             | \( p = .3 \) | correct | 7.27(5.29) | 7.42(5.65) | 6.95(5.17) | 6.96(5.52) |
|             |        | wrong  | 0(0.1) | 0(0) | 0(0.3) | 0(0.15) |
|             | \( p = .5 \) | correct | 7.45(5.7) | 7.53(5.89) | 7.61(5.93) | 7.55(5.92) |
|             |        | wrong  | 0(0.16) | 0(0.04) | 0(0.46) | 0(0.14) |

https://doi.org/10.1371/journal.pone.0241197.t014
Table 16. Average numbers of correct and wrong zero coefficients for fixed effects and random effects in terms of HPPM for Simulation 3. Random effects are in parentheses.

| Sample size | r |        | Normal | Normalmixtrue | Laplace | Laplacemixtrue |
|-------------|---|--------|--------|---------------|--------|----------------|
| n = 30      | p = .1 | correct | 6.05(4.11) | 6.48(4.94) | 6.15(2.69) | 6.26(4.14) |
|             |       | wrong  | 0(0.43)   | 0(0.27)    | 0(0.26)   | 0(0.39)    |
|             | p = .3 | correct | 6.92(4.84) | 7.21(5.58) | 6.92(3.94) | 6.89(5.07) |
|             |       | wrong  | 0(0.37)   | 0(0.25)    | 0(0.44)   | 0(0.42)    |
|             | p = .5 | correct | 6.95(5.46) | 7.25(5.83) | 7.09(3.59) | 7.18(5.83) |
|             |       | wrong  | 0(0.51)   | 0(0.27)    | 0(0.75)   | 0(0.59)    |
| n = 50      | p = .1 | correct | 6.2(3.96) | 6.24(4.77) | 6.26(3.07) | 6.38(3.82) |
|             |       | wrong  | 0(0.21)   | 0(0.1)     | 0(0.37)   | 0(0.28)    |
|             | p = .3 | correct | 7.08(5.32) | 6.99(5.76) | 6.92(5.17) | 6.88(5.58) |
|             |       | wrong  | 0(0.23)   | 0(0.11)    | 0(0.51)   | 0(0.33)    |
|             | p = .5 | correct | 7.13(5.63) | 7.25(5.86) | 7.4(5.86)  | 7.38(5.96) |
|             |       | wrong  | 0(0.38)   | 0(0.21)    | 0(0.66)   | 0(0.38)    |
| n = 100     | p = .1 | correct | 6.34(3.33) | 6.48(3.74) | 6.45(2.39) | 6.5(2.94)  |
|             |       | wrong  | 0(0.05)   | 0(0.01)    | 0(0.17)   | 0(0.13)    |
|             | p = .3 | correct | 7.3(5.29) | 7.39(5.64) | 7.01(5.21) | 6.97(5.54) |
|             |       | wrong  | 0(0.1)    | 0(0)       | 0(0.31)   | 0(0.15)    |
|             | p = .5 | correct | 7.47(5.72) | 7.55(5.88) | 7.59(5.96) | 7.53(5.92) |
|             |       | wrong  | 0(0.17)   | 0(0.04)    | 0(0.48)   | 0(0.15)    |

Table 16, it can be observed that the differences between the four different priors is quite small in general. Hence our methods are not sensitivity to the choice of the hyper-parameters for the priors.

Real data example

In this section, we use the Age-Related Macular Degeneration Trial (ARMD) data to assess the behavior of the proposed model under different quantiles. This dataset consists of 4 variables, available in the R package `nlme`. Visual acuity of 188 patients with ARMD were measured at 4 time points, resulting in 752 observations. We consider the relationship between the variable of the visual acuity (VISUAL) and three other explanatory variables, including the value of visual acuity measured at baseline (VISUAL0), the time of the measurement (TIME), the treatment indicator (TREAT). A detailed description is given in [34]. [34] analyzed the data set using a linear mean mixed model and found that there exist possible interactions between time and treat effects, we then consider the following LQMM ($i = 1, \ldots, 188, j = 1, \ldots, 4$):

\[
\text{VISUAL}_{ij} = \beta_0 + \beta_1 \times \text{VISUAL}_0 + \beta_2 \times \text{TIME}_i + \beta_3 \times \text{TREAT}_i + \beta_4 \times \text{TIME}_i \times \text{TREAT}_i + \beta_5 \times \text{VISUAL}_0 + \beta_6 \times \text{TIME}_i + \beta_7 \times \text{TREAT}_i + \beta_8 \times \text{TIME}_i \times \text{TREAT}_i + \varepsilon_{ij},
\]

Here we choose five different values of $r$, i.e. 0.1, 0.3, 0.5, 0.7, 0.9, to thoroughly describe the response distribution. We implement the adGL, adSpikeGL, PMQ, BL, BAL and QR models
on this data set. Table 17 shows the summaries of the MAD and the RMSE, where

\[
\text{MAD} = \frac{1}{752} \sum_{i=1}^{188} \sum_{j=1}^{4} |y_{ij} - \hat{y}_{ij}|, \\
\text{RMSE} = \sqrt{\frac{1}{752} \sum_{i=1}^{188} \sum_{j=1}^{4} (y_{ij} - \hat{y}_{ij})^2}.
\]

The results in Table 17 suggest that PMQ provides smaller MAD and the RMSE than other methods in upper quantiles ($\tau = 0.7$ and $\tau = 0.9$), however, adGL and adSpikeGL perform better in lower and median quantile ($\tau = 0.1$, $\tau = 0.3$ and $\tau = 0.5$). We also report the posterior median estimations and 95% credible intervals for fixed effects and random effect variances in Table, respectively. As in simulations, adGL, adSpikeGL and PMQ are inclined to have similar performance, meanwhile, BAL, BL and QR also are likely to behave similarly, and all methods except for the adSpikeGL do not compress any fixed and random effects to zero.

The results of adSpikeGL in Tables 18 and 19 show that (i) for fixed effects, VISUAL0, TIME and TREAT are all identified as effective variables of Visual acuity at all quantiles except for TREAT at extreme quantile ($\tau = 0.1$), this may be because treatment has no significant effect on the patients with poor visual acuity, (ii) for random effects, VISUAL0 and TREAT are deemed to be significant predictors of Visual acuity at all quantiles, however, TIME to be
### Table 18. Posterior median and 95% credible intervals for the fixed effects of ARMD data.

| Parameter | adGL | adSpikeGL | PMQ | BL | BAL | QR |
|-----------|------|-----------|-----|----|-----|----|
| $\beta_0$ | -1.59 (7.495,1.045) | -0.11 (7.424,0.952) | -0.12 (-2.211,0.796) | -1.18 (-7.773,4.749) | -1.58 (-6.614,3.445) | 2.89 (7.474,5.445) |
| $\beta_1$ | 0.89 (0.834,1.005) | 0.88 (0.841,0.999) | 0.86 (0.831,0.915) | 0.91 (0.803,1.029) | 0.94 (0.849,1.027) | 0.81 (0.753,0.871) |
| $\beta_2$ | -0.21 (-0.27,0.145) | -0.22 (-0.28,0.162) | -0.207 (-0.33,0.139) | -0.222 (-0.27,0.172) | -0.226 (-0.247,0.174) | -0.53 (-0.52,0.13) |
| $\beta_3$ | -0.42 (-3.626,1.11) | 0 (-3.039,0.749) | -0.12 (-1.79,0.695) | -2.49 (-6.143,1.073) | -2.43 (-5.618,1.039) | -4.37 (-4.37,0.13) |
| $\beta_4$ | -0.05 (-0.147,0.04) | 0 (-0.119,0.021) | -0.057 (-0.156,0.029) | -0.065 (-0.138,0.007) | -0.16 (-0.136,0.006) | -0.03 (-0.03,0.006) |

### Table 17. MAD and RMSE for the ARMD data.

| Criterion | Method | $r = 0.1$ | $r = 0.3$ | $r = 0.5$ | $r = 0.7$ | $r = 0.9$ |
|-----------|--------|-----------|-----------|-----------|-----------|-----------|
| **MADs**  | adGL   | 7.140     | 5.306     | 5.078     | 4.935     | 6.805     |
|           | adSpikeGL | 5.931     | 4.459     | 4.717     | 4.970     | 6.321     |
|           | PMQ    | 7.170     | 5.959     | 6.056     | 4.251     | 5.871     |
|           | BL     | 9.138     | 6.011     | 5.166     | 5.936     | 7.961     |
|           | BAL    | 9.189     | 6.052     | 5.211     | 6.009     | 7.978     |
|           | QR     | 16.665    | 10.025    | 8.712     | 9.618     | 13.764    |
| **RMSEs** | adGL   | 9.385     | 7.162     | 7.256     | 7.725     | 9.871     |
|           | adSpikeGL | 8.461     | 6.339     | 6.975     | 7.726     | 9.137     |
|           | PMQ    | 9.490     | 7.820     | 8.370     | 6.870     | 8.909     |
|           | BL     | 12.786    | 8.769     | 8.252     | 9.739     | 12.081    |
|           | BAL    | 12.828    | 8.896     | 8.351     | 9.877     | 12.136    |
|           | QR     | 19.948    | 12.816    | 12.120    | 13.624    | 17.975    |

Insignificant predictor at lower and median quantiles, and (iii) it is worth noticing that the fixed and random effect of the interactions between time and treat effects (TIME*TREAT) is not selected by adSpikeGL across all quantiles, suggesting that the effect of the treatment does not change over time no matter what the patient’s vision is.
Conclusion and discussion

In this work, we presented a novel shrinkage method to linear quantile mixed model for analysing and interpreting longitudinal data. We also consider the variable selection based on a Bayesian adaptive lasso and an extended Bayesian adaptive group lasso with spike and slab priors. Unlike other approaches for LQMMs, which need some ad hoc methods to do model selection, adSpikeGL can directly identify the significant fixed and random effects simultaneously. Several criterions were adopted to evaluate the finite sample performance of the proposed methods thoroughly by simulations and real data example. The results reveal that the proposed methods is very competitive with the existing methods. On the other hand, as the sample size gets large, adGL and adGLSpike generally tend to provide smaller MMAD, MRMSE, and MMR across all scenarios, and adGLSpike also has good performance in variable selection in most cases. The only exception is at $\tau = 0.1$ where the number of the correct zero coefficients of adGLSpike did not show an increasing trend with sample size. The problem can be caused by the drawbacks of the ALD that do not allow the skewness, kurtosis and tails of distribution to vary. To overcome this problem [35] recently introduced a generalized asymmetric Laplace distribution (GALD) to establish QR in linear regression. In the future, we would like to extend the GALD to the linear mixed effect models. Moreover, we would also like to consider some other Bayesian shrinkage priors, such as the double-Pareto

Table 19. Posterior median and 95% credible intervals for the random effect variances of ARMD data.

| $\tau$ | Parameters | adGL | adSpikeGL | PMQ |
|--------|------------|------|-----------|-----|
| 0.1    | $\sigma_{00}$ | 3.182(0.016,14.82) | 91.361(62.646,130.536) | 0.251(0.22,157) |
|        | $\sigma_{11}$ | 0.023(0.007,0.035) | 0(0,0) | 0.029(0.01,0.043) |
|        | $\sigma_{22}$ | 0.076(0.051,0.109) | 0.077(0.058,0.102) | 0.073(0.05,0.106) |
|        | $\sigma_{33}$ | 1.236(0.031,8.159) | 247.861(113.119,422.548) | 1.249(0.002,8.935) |
|        | $\sigma_{44}$ | 0.025(0.001,0.121) | 0(0,0) | 0.065(0.02,0.253) |
| 0.3    | $\sigma_{00}$ | 0.981(0.002,8.525) | 30.736(15.537,56.308) | 0.087(0.12,763) |
|        | $\sigma_{11}$ | 0.008(0.02,0.014) | 0(0,0.003) | 0.007(0.001,0.013) |
|        | $\sigma_{22}$ | 0.072(0.049,0.103) | 0.072(0.053,0.096) | 0.072(0.049,0.102) |
|        | $\sigma_{33}$ | 4.622(0.218,13.325) | 137.31(35.006,217.77) | 20.295(0.957,144.869) |
|        | $\sigma_{44}$ | 0.026(0.002,0.104) | 0(0,0) | 0.007(0.01,0.102) |
| 0.5    | $\sigma_{00}$ | 0.591(0.10,282) | 27.488(15.813,46.503) | 1.387(0.51,397) |
|        | $\sigma_{11}$ | 0.007(0.001,0.014) | 0(0,0) | 0.004(0.01,0.011) |
|        | $\sigma_{22}$ | 0.068(0.046,0.104) | 0.066(0.048,0.099) | 0.067(0.047,0.098) |
|        | $\sigma_{33}$ | 3.363(0.11,9.938) | 81.255(2.149,172.215) | 14.283(0.082,97.629) |
|        | $\sigma_{44}$ | 0.017(0.001,0.153) | 0(0,0) | 0.006(0.01,0.114) |
| 0.7    | $\sigma_{00}$ | 10.218(2.272,24.638) | 143.485(78.765,280.103) | 90.912(28.974,222.612) |
|        | $\sigma_{11}$ | 0.002(0.00,0.007) | 0.013(0.004,0.048) | 0.005(0.002,0.028) |
|        | $\sigma_{22}$ | 0.062(0.04,0.088) | 0.063(0.045,0.087) | 0.064(0.042,0.094) |
|        | $\sigma_{33}$ | 1.795(0.045,9.234) | 54.378(4.052,162.822) | 2.581(0.005,80.103) |
|        | $\sigma_{44}$ | 0.016(0.001,0.093) | 0(0,0) | 0.006(0.01,0.115) |
| 0.9    | $\sigma_{00}$ | 21.142(0.001,40.1) | 162.243(104.606,339.58) | 170.374(92.769,301.983) |
|        | $\sigma_{11}$ | 0.002(0.00,0.03) | 0.013(0.004,0.067) | 0.013(0.003,0.036) |
|        | $\sigma_{22}$ | 0.065(0.039,0.379) | 0.066(0.048,0.089) | 0.074(0.048,0.109) |
|        | $\sigma_{33}$ | 1.168(0.026,8.957) | 50.279(5.654,124.126) | 2.636(0.004,74.622) |
|        | $\sigma_{44}$ | 0.017(0.001,0.389) | 0(0,0) | 0.028(0.01,0.199) |

https://doi.org/10.1371/journal.pone.0241197.t019
prior [36], horseshoe prior [37], the normal-gamma prior [38], etc, though the lasso-type priors is the most widely used shrinkage priors in the literature.

**Supporting information**

S1 File. (RAR)

**Acknowledgments**

We thank the editor, the associate editor and two anonymous referees for their helpful comments which led to a considerable improvement of the original article.

**Author Contributions**

Methodology: Yonggang Ji.

Software: Haifang Shi.

**References**

1. Koenker R. Quantile regression for longitudinal data. Journal of Multivariate Analysis. 2004; 91: 74–89. https://doi.org/10.1016/j.jmva.2004.05.006

2. Harding M, Lamarche C. A quantile regression approach for estimating panel data models using instrumental variables. Economics Letters. 2009; 104:133–135. https://doi.org/10.1016/j.econlet.2009.04.025

3. Canay AI. A simple approach to quantile regression for panel data. Econometrics Journal. 2011; 14:368–386. https://doi.org/10.1111/j.1368-423X.2011.00349.x

4. Graham B, Hahn J, Poirier A, Powell J. A quantile correlated random coefficients panel data model. Journal of Econometrics. 2018; 206:305–335. https://doi.org/10.1016/j.jeconom.2018.06.004

5. Galarza C, Lachos V, Bandyopadhyay D. Quantile regression in linear mixed models: a stochastic approximation EM approach. Statistics and Its Interface. 2017; 10:471–482. https://doi.org/10.4310/SII.2017.v10.n3.a10 PMID: 29104713

6. Chemozhukov V, Fernandez Val I, Hahn J, Newey W. Average and quantile effects in nonseparable panel models. Econometrica. 2013; 81:535–580. https://doi.org/10.3982/ECTA8405

7. Harding M, Lamarche C. Estimating and testing a quantile regression model with interactive effects. Journal of Econometrics. 2014; 178:101–113. https://doi.org/10.1016/j.jeconom.2013.08.010

8. Sriram K, Shi P, Ghosh P. A Bayesian quantile regression model for insurance company costs data. Journal of the Royal Statistical Society, Series A. 2016; 179:177–202. https://doi.org/10.1111/rssa.12111

9. Geraci M, Bottai M. Linear quantile mixed models. Statistics and computing. 2014; 24:461–479. https://doi.org/10.1007/s11222-013-9381-9

10. Aghahosseinmadi A, Mohammadi S. Bayesian analysis of penalized quantile regression for longitudinal data. Statistical Papers. 2017; 58:1035–1053. https://doi.org/10.1007/s00362-015-0737-4

11. Alhamzawi R, Yu K. Bayesian Lasso mixed quantile regression. Journal of Statistical Computation and Simulation. 2014; 84:868–880. https://doi.org/10.1080/00949655.2012.731689

12. Seber GAF. Linear Regression Analysis. New York: Wiley; 1977.

13. Ibrahim JG, Zhu H, Garcia RR, Guo R. Fixed and Random Effects Selection in Mixed Effects Models. Biometrics. 2011; 67:495–503. https://doi.org/10.1111/j.1541-0420.2010.01463.x PMID: 20662831

14. Yu K, Moyeed RA. Bayesian quantile regression. Statistics and Probability Letters. 2001; 54:437–447. https://doi.org/10.1016/S0167-7152(01)00124-9

15. Yu K, Zhang J. A three-parameter asymmetric Laplace distribution and its extension. Communications in Statistician Theory and Methods. 2005; 34:1867–1879. https://doi.org/10.1080/03610920500199018

16. Geraci M, Bottai M. Quantile regression for longitudinal data using the asymmetric Laplace distribution. Biostatistics. 2007; 8:140–154. https://doi.org/10.1093/biostatistics/kxj039 PMID: 16636139
17. Li Q, Xi R, Lin N. Bayesian regularized quantile Regression. Bayesian Analysis. 2010; 5:533–556. https://doi.org/10.1214/10-BA521
18. Alhamzawi R, Yu K. Bayesian adaptive LASSO quantile regression. Statistical Modelling. 2012; 12:279–297. https://doi.org/10.1177/1471082X1101200304
19. Kozumi H, Kobayashi G. Gibbs sampling methods for Bayesian quantile regression. Journal of Statistical Computation and Simulation. 2011; 81:1565–1578. https://doi.org/10.1080/00949655.2010.496117
20. Ji YG, Lin N, Zhang BX. Model selection in binary and tobit quantile regression using the Gibbs sampler. Computational Statistics & Data Analysis. 2012; 56:827–839. https://doi.org/10.1016/j.csda.2011.10.003
21. Alshayba wee T, Midi H, Alhamzawi R. Bayesian elastic net single index quantile regression. Journal of Applied Statistics. 2017; 44:853–871. https://doi.org/10.1080/02664763.2016.1189515
22. Kobayashi G. Bayesian endogenous tobit quantile regression. Bayesian Analysis. 2017; 12:161–191. https://doi.org/10.1214/16-BA996
23. Li Q, Lin N. The Bayesian Elastic Net. Bayesian Analysis. 2010; 5:151–170. https://doi.org/10.1214/10-BA506
24. Chen Z, Dunson D. Random effects selection in linear mixed models. Biometrics. 2003; 59:762–769. https://doi.org/10.1111/j.0006-341X.2003.00089.x PMID: 14969453
25. Kinney SK, Dunson D. Fixed and random effects selection in linear and logistic models. Biometrics. 2008; 63:690–698. https://doi.org/10.1111/j.1541-0420.2007.00771.x
26. van der Pas S, Szabo B, der Vaart v.Aad. Adaptive posterior contraction rates for the horseshoe. Electronic Journal of Statistics. 2017; 11:3196–3225. https://doi.org/10.1214/17-EJS1316
27. van Dyk DA, Park T. Partially collapsed Gibbs samplers: Theory and methods. Journal of the American Statistical Association. 2008; 103:790–796. https://doi.org/10.1198/016214508000000409
28. Liu JS, Wong WH, Kong A. Covariance Structure of the Gibbs Sampler, With Applications to Comparisons of Estimators and Augmentation Schemes. Biometrika. 1994; 81:27–40. https://doi.org/10.1093/biomet/81.1.27
29. Bondell HD, Krishna A, Ghosh SK. Joint Variable Selection for Fixed and Random Effects in Linear Mixed Effects Models. Biometrics. 2010; 66:1069–1077. https://doi.org/10.1111/j.1541-0420.2010.01391.x PMID: 20163404
30. Peng H, Lu Y. Model selection in linear mixed effect models. Journal of Multivariate Analysis. 2012; 109:109–129. https://doi.org/10.1016/j.jmva.2012.02.005
31. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. Journal of computational and graphical statistics. 1998; 7:434–455. https://doi.org/10.1080/10618600.1998.10474787
32. Williams JP, Lu Y. Covariance Selection in the Linear Mixed Effect Model. JMLR: Workshop and Conference Proceedings. 2015; 44:277–291.
33. Barbieri MM, Berger JO. Optimal predictive model selection. The Annals of Statistics. 2004; 32:870–897. https://doi.org/10.1214/009053604000000238
34. Galecki A, Burzykowski T. Linear Mixed-Effects Models Using R. Springer; 2013.
35. Yan Y, Kottas A. A New Family of Error Distributions for Bayesian Quantile Regression. Working Paper. 2017.
36. Armagan T, Dunson D, Lee J. Generalized double Pareto shrinkage. Statistica Sinica. 2013; 23:119–143. PMID: 24478567
37. Carvalho CM, Polson NG, Scott JG. The horseshoe estimator for sparse signals. Biometrika. 2010; 97:465–480. https://doi.org/10.1093/biomet/asq017
38. Griffin JE, Brown PJ. Inference with normal-gamma prior distributions in regression problems. Bayesian Analysis. 2010; 5:171–188. https://doi.org/10.1214/10-BA507