STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. 11, March 2023, pp 445–464.
Published online in International Academic Press (www.IAPress.org)

A Bayesian Semi-parametric Quantile Regression Approach for Joint Modeling of Longitudinal Ordinal and Continuous Responses

Omid Khazaei , Mojtaba Ganjali *, Mojtaba Khazaee

Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran

Abstract   Quantile regression (QR) models are one of the methods for longitudinal data analysis. When responses seem to be skew and asymmetric due to outliers and heavy-tails, QR models may work suitably. This paper developes the semi-parametric quantile regression model for analyzing longitudinal continuous and ordinal mixed responses. The latent variable model and some threshold parameters are used to perform the quantile regression model's ordinal part. The error of the latent variable model has Asymmetric Laplace (AL) distribution. The error term's distribution is assumed to be AL distribution to model the continuous responses. The correlations of longitudinal responses belong to the same individual and those of mixed continuous and ordinal responses are considered using a random-effects approach. The regression spline is used to approximate the non-parametric part of the model. The parameter estimation procedure is performed under a Bayesian paradigm using the Gibbs sampling method. A simulation study is performed to demonstrate the proposed model's performance where the relative biases, standard errors, and root of MSEs of estimated parameters are decreased in the semi-parametric QR joint model when the number of subjects is increased. In our application, it was found that the mother's age and her child's age have significant effects on reading ability, and antisocial behavior depends on the child's gender.

Keywords   Semi-parametric Quantile regression, longitudinal data, joint model, Regression spline, Asymmetric Laplace distribution, Gibbs sampling.

AMS 2010 subject classifications 62G08, 62F15, 62C10, 62J05

DOI:10.19139/soic-2310-5070-1225

1. Introduction

1.1. Motivation

There are correlated continuous and ordinal responses in many longitudinal studies in social, psychiatry, educational, public health, political, and medical studies. Analysis of the joint responses by a model has some potential advantages over using separate models. However, models for response variables of different types (discrete and continuous) are challenging to define and fit. Ordered categorical data are commonly used in scientific fields where respondents express a graduated evaluation on a particular item. Examples include survey responses on opinions, ratings of preference in consumer studies, numerical or verbal rating scale, sensory evaluation on food perception and appreciation, self-evaluation of well-being and life satisfaction, job satisfaction, job classifications by skill levels, economic perceived conditions, etc. In many cases, the ordinal responses may be collected with other responses, such as continuous or count responses. For example, suppose a questionnaire is designed to extract the simultaneous effects of gender, level of education and age on income and life satisfaction. Then, it is clear that the variables of continuous

*Correspondence to: Mojtaba Ganjali (Email: m-ganjali@sbu.ac.ir). Department of Mathematics, Shahid Beheshti University. Daneshjou Boulevard, Shahid Chamran Highway, Tehran, Tehran Province, Iran.

ISSN 2310-5070 (online)   ISSN 2311-004X (print)

response (income) and ordinal response (life satisfaction) are correlated, and their joint modeling will be more helpful.

In some longitudinal surveys, the ordinal responses are observed through a hierarchical questionnaire or interview across time. Consider the following example to clarify the matter. The PIAT data were prepared at two-year intervals between 1986 and 1992 and contains both the mother and her child's interviews. This paper has selected two interesting continuous and ordinal data responses as the primary variables for modeling. The continuous response is the child's reading recognition skill (read), and the ordinal response is the child's antisocial behavior (anti). The study's main idea was to specify the connection between parental response to a child's antisocial behavior and reading recognition skills. There are 221 children and mothers who were participated in the study. Further descriptions of this data set and the mentioned variables are provided in the section 5. The ordinal response is a child's antisocial behavior (anti), and the continuous response is the child's reading recognition skill (read). The antisocial behavior sub-scale consisted of the mother's report on six items that evaluated the child's antisocial behavior over the previous period of three months. The item measured the extent to which the child:

- Cheats or tells lies
- Bullies or is cruel or mean to others
- Does not seem to feel sorry after he/she misbehaves
- Breaks things on purpose or deliberately destroys his/her own or other's things
- Is disobedient at school
- Has trouble getting along with teachers.

The three possible response choices were "not true" (score of 0), "sometimes true" (score of 1), and "often true" (score of 2). These six item scores were summed to compute an overall measure of antisocial behavior, which could range in value from 0 to 12. We categorized the variable to a three-level ordinal response which takes the values 1, 2, and 3 when the sum is equal to 0, larger than 0 but less than or equal to 6 and more than 6, respectively.

Figure 1 displays the box-plots of the continuous response (reading ability) for different levels of ordinal response (antisocial behavior). According to this Figure, it can be found that children with low levels of antisocial behavior have more reading ability than children with high levels of antisocial behavior at all time points. Also, the Reading ability is improving over time. So, these two longitudinal endogenous responses are correlated, and the effects of covariates on them should be taken into account simultaneously.

1.2. Related works

In many studies, interests are often concentrated on the significant effect of predictor variables on specific responses variables over time. Linear regression models are statistical tools used to model the association between variables. Many researchers used Classic Mean Regression (CMR) as a general tool to analyse the data. The principal purpose of CMR is to model the mean of the response variable under a set of predictor variables. Quantile regression (QR) is a type of regression analysis used for data analysis in statistics, biomedical, medicine and econometrics. [28] used QR for the first time, which does not follow the Gaussian assumption for the model error distribution, but formalized asymptotic properties of the least absolute deviations estimator for independent observations. In recent years, many researchers used QR models for longitudinal data analysis. QR models are the quantiles of the response variable given covariates. Quantiles are less sensitive to skewed distributions and outliers. A comprehensive guide of the subject can be found in [30] and [21]. QR estimators determine a linear plan and can be measured efficiently. [30] characterized the finite-sample distributions of regression quantiles, but they are challenging to use for statistical inference. In the last few years, QR models became a more generally utilized method to define a response variable's distribution given covariates. [27] used a quantile method to analyze the nonlinear longitudinal responses. [23] recommended the transformed ordinal QR for single-index models. [1] suggested a method for regularisation in mixed QR models.

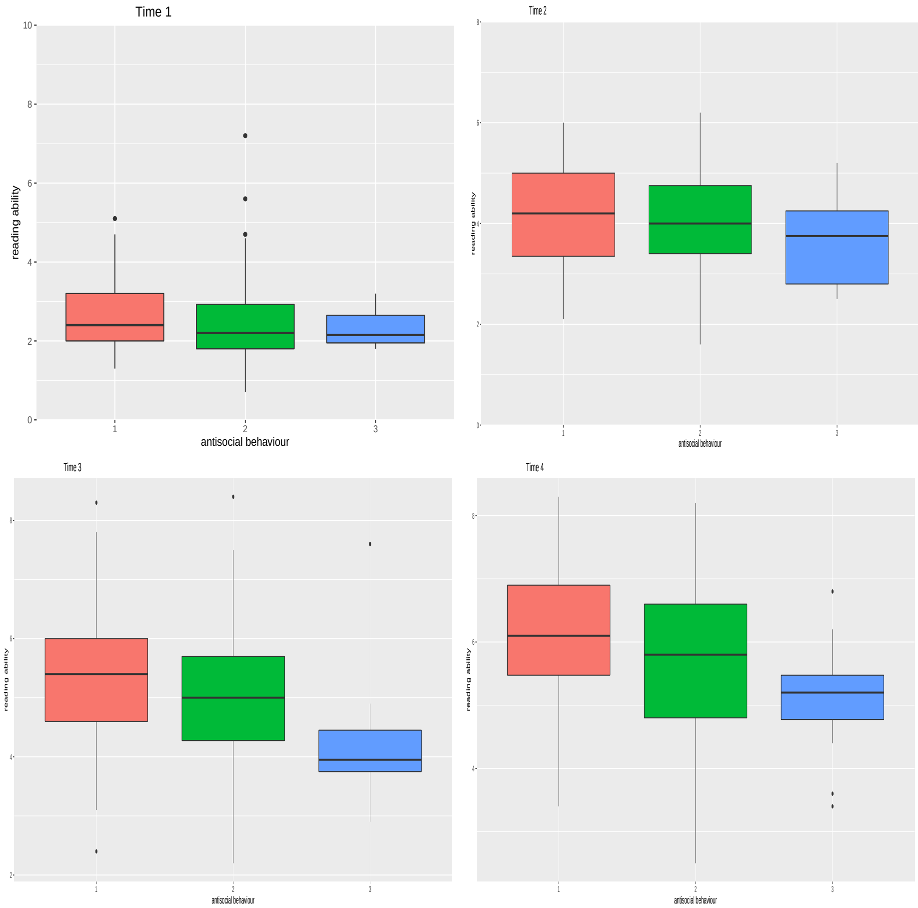

Figure 1. Box-plots of reading ability for different levels of antisocial behavior at four time-points of PIAT data set.

The longitudinal study is often more informative than the classical cross-sectional samples because of containing repeated measurements within each individual over time. The changes of responses and their temporal patterns within each individual usually provide important information of scientific relevance. Since longitudinal study combines the features of cross-sectional sampling and time-series observations, their usefulness goes very high. It is often found in economics, biomedical studies, psychology, social activities and many other scientific areas. In these studies, responses and some covariates are evaluated for independent individuals over time, frequently. There are two approaches for analysing ordered and categorical responses: cumulative link models and the ordered probit models. Cumulative link models are a suitable method for modelling these data because the ordered nature of observation is used, and the flexible regression structure provides good fitting performances. These models with a logit link are generally known as the proportional odds model due to [35]. The cumulative models are also known as ordinal regression models. Other alternatives regarding the logit and probit link functions are ordered logit models and ordered probit models (for more information, see, Ch. 4 of[20]). In its contemporary, regression-based form, [36],[37] proposed the ordered probit model to analyse such responses. The model platform is an underlying random utility model or latent regression model, in which the continuous latent ”measure” is observed in discrete form through a censoring mechanism.

The existence of the mixed correlated discrete and continuous responses for each experimental unit is one of the critical issues in statistical applications. The general location model of [40] is used for joint modelling of mixed categorical and continuous data. In this method, the continuous and categorical variables' joint distribution is decomposed into a marginal multinomial distribution for the categorical variables and a conditional multivariate normal distribution for the continuous variables given the categorical variables. Several models using latent variables have been introduced to analyse multiple non-proportional responses as covariates' functions [44]. [3], [4] and [5] proposed the latent variable model for bivariate and multivariate continuous and ordinal mixed responses. In their models, the ordinal responses are within-correlated and are correlated with continuous response. They modelled ordinal and continuous responses by using and generalising a method similar to the method of [22] and using the idea of latent variables. [34] concentrated on the mixed data analysis, and discussed all of the suitable models that other researchers use. [45] introduced a general joint model for mixed longitudinal set-inflated continuous and set-inflated count responses.

Longitudinal data analysis with the correlation structure within responses may be complicated. In this situation, Joint models can consider suitably for the correlation between the responses. [14] have introduced different producers for joint modeling in their book. One of the methods they use is the random effects approach. As we know, the correlation of longitudinal responses of the same individual is considered using a random-effects approach. The same idea can be used to create longitudinal models by random effects. The random-effects models imply more flexible correlation patterns compared to the other models. The separate random effects are used in the models of the joint model to confirm a longitudinal response dependence. For more information about these, see [38] and [14]. See also [16], [5], [43], [49], [45], [46] and [19] for joint modeling of mixed discrete and continuous longitudinal responses.

Recently, [31] considered a semi-parametric method for the QR model using a Bayesian approach. They suggested the Dirichlet process for the error term in an additive QR construction. [56] used the semi-parametric QR model to decrease the covariates' dimension and keep non-parametric regression flexibility. [7] proposed semi-parametric QR estimation for dynamic models with partially varying coefficients so that the values of some coefficients may be functions of informative covariates. They recommended the estimation of both parametric and non-parametric functional coefficients. [57] used the QR model with a latent variable approach for the ordinal response variable. [24] introduced a multi-index semi-parametric QR model for ordinal data. [8] suggested semi-parametric partial linear quantile regression of longitudinal data with time-varying coefficient and informative observation of time. [25] proposed quantile regression semi-parametric nonlinear mixed-effects models to analyze longitudinal data using a Bayesian approach.

This paper provides a new semi-parametric QR joint model for analyzing mixed longitudinal data using a Bayesian framework where the posterior distribution can be evaluated under various priors. In the proposed approach, we use the random effects to consider the dependence between continuous and ordinal responses. In our method, the Latent variable approach is used for modeling the ordinal response. Also, we use the regression spline method to approximate the model's non-parametric part. We extended a Metropolis-Hastings algorithm within the Gibbs sampling to sample from the posterior distribution.

This paper is organized as follows: In Section 2, we first provide the relation between Asymmetric Laplace distribution and the simple semi-parametric QR (SQR) model. Then a semi-parametric QR approach for analyzing the joint model of longitudinal ordinal and continuous responses is presented. Section 3 specifies prior distributions for unknown parameters and formulates a Bayesian hierarchical model for the semi-parametric QR model defined in Section 2. Section 4 provides some simulation studies to evaluate the performance of the model. The proposed model is used for analyzing a data set (PIAT), in Section 5. Some results and discussion are presented in Section 6.

2. SQR joint models

2.1. Asymmetric laplace distribution and SQR model

Suppose the random variable y has the asymmetric Laplace (AL) distribution, then its probability density function is

$$f(y|\mu,\sigma,\tau)=\frac{\tau(1-\tau)}{\sigma}\exp\left\{-\rho_\tau\left(\frac{y-\mu}{\sigma}\right)\right\}, \tag{1}$$

where μ is a location parameter, σ is scale parameter, τ is skewness parameter, y, $\mu\in R$, $0<\tau<1$ and $\sigma>0$. In equation (1), $\rho_\tau(u)$ is the loss function ([30]) determined by

$$\rho_\tau(u)=u(\tau-I(u<0)) \qquad 0<\tau<1 \tag{2}$$

such that $I(.)$ is the indicator function. In AL distribution $F^{ALD}_{y|x}(\mu)=p(y<\mu)=\tau$, it means that μ is the τth quantile of AL distribution. This feature is one of the attractive properties in the quantile regression model for analyzing data. The ALD has different mixture representations as shown by [32]. We use a mixture demonstration of the standard exponential distribution and standard normal distribution to generate a QR model sampling algorithm. This structure can be found in [32].

Theorem 1
Let U be an exponential variable with mean σ, Z be a standard normal variable and U and Z be independent of each other. If a random variable Y follows the ALD with density (1), then we can represent Y as a location-scale mixture of normals given by

$$Y=\sqrt{aU}Z+bU+\mu \tag{3}$$

where $a=\dfrac{2\sigma}{\tau(1-\tau)}$, $b=\dfrac{1-2\tau}{\tau(1-\tau)}$ and $Y\sim ALD(\mu,\sigma,\tau)$.

A proof for theorem 1 can be found in [33]. The equation (3) can be written in a hierarchical setup as follows

$$\begin{aligned}Y|U&\sim N(bU+\mu,aU),\\ U&\sim exp(\sigma).\end{aligned}$$

This arrangement can be used in performing the Bayesian approach. The mean and variance of AL distribution can be found from (3) as

$$E(Y)=\mu+\frac{\sigma(1-2\tau)}{\tau(1-\tau)},\qquad \mathrm{var}(Y)=\frac{1-2\tau+2\tau^2}{\tau^2(1-\tau)^2}\sigma^2.$$

see more in [55].

Considering the following simple semi-parametric model for the responses y_i,

$$y_i=\boldsymbol{x}_i^T\boldsymbol{\beta}_\tau+\eta_{i,\tau}(.)+\varepsilon_{i,\tau}, \qquad i=1,...,n \tag{4}$$

where $\boldsymbol{x}_i$ is the $k\times 1$ fixed vector of covariates, $\boldsymbol{\beta}_\tau$ is the $k\times 1$ vector of unknown parameters, $\eta_{i,\tau}(.)$ is the non-parametric part of the model and $\varepsilon_{i,\tau}$ is the error of the model. The error distribution is supposed to be unknown in the QR model, and it is restricted to have the τ-th quantile to be equal to zero for $0<\tau<1$.

One way to estimate the fixed regression coefficient in the semi-parametric QR model is to minimize the sum of the asymmetrically weighted absolute errors

$$\min_{\beta_\tau}\sum_{i=1}^n\rho_\tau(y_i-\boldsymbol{x}_i^T\boldsymbol{\beta}_\tau-\eta_{i,\tau}(.)) \tag{5}$$

where $\rho(.)$ was defined in (2). Considering that (5) is not recognizable at the origin, there is no closed-form solution for $\boldsymbol{\beta}$ and a linear programming algorithm can perform the minimization of (5) ([32, 29, 15]). To implement a Bayesian inference with SQR, suppose that the errors in (4), given a known τ, are independent and $\varepsilon_{i,\tau} \sim ALD(0, \sigma, \tau)$. Then a continuous random variable y_i given $\boldsymbol{x}_i$ have $ALD(\mu_i, \sigma, \tau)$ where $\mu_i = \boldsymbol{x}_i^T \boldsymbol{\beta}_\tau - \eta_{i,\tau}(.)$ is the location parameter. The likelihood function of parameters ($\boldsymbol{\beta}_\tau$ and σ) can be composed based on n independent components of $\boldsymbol{y} = (y_1, y_2, ..., y_n)$ as follows:

$$L(\boldsymbol{\beta}_\tau, \sigma|\boldsymbol{y}) = \left(\frac{\tau(1-\tau)}{\sigma}\right)^n exp\left\{-\sum_{i=1}^{n} \rho_\tau(\frac{y_i - \mu_i}{\sigma})\right\}. \tag{6}$$

The maximum likelihood estimates (MLEs) of $\boldsymbol{\beta}_\tau$ and σ are obtained by maximization of Equation (6) with respect to $\boldsymbol{\beta}_\tau$. This is equivalent to minimizing the loss function in (5) with respect to $\boldsymbol{\beta}_\tau$. The connection between AL distribution and Bayesian quantile regression first proceeded in [54].

2.2. Monotone equivariance property of QR models

Authors often transform the variable's scale to get an appropriate model fit or normalize the data. One of the quantiles' features is the monotone equivariance property. See more in, [21]. It means that if we use a monotone transformation g to a random variable, then quantiles of the new random variable are collected using a similar transformation with the quantile of the untransformed random variable. For example, if Q_τ is the τth quantile of the random variable Y then $g(Q_\tau)$ is the τth quantile of $g(Y)$. QR model contains monotone equivariance property as well. It means that if g is a monotone function, then we have

$$Q_\tau\left(g(Y_i)|x_i\right) = g\left(Q_\tau(Y_i|x_i)\right). \tag{7}$$

CMR does not hold the above condition. The QR model can reinterpret by using this feature. For more details, see [21] and [30]. By adopting this feature, a semi-parametric QR joint model can be fitted for ordinal and continuous data.

2.3. Joint model: Notation and model specification

For modeling the mixed longitudinal ordinal and continuous responses, let $\tilde{\boldsymbol{y}} = (\tilde{y}_{ij}, j = 1, ..., n_i, i = 1, ..., n)$ and $\boldsymbol{y} = (y_{ij}, j = 1, ..., n_i, i = 1, ..., n)$, where $\tilde{y}_{ij}$ and y_{ij} denote the ordinal with levels $l_1, l_2, \ldots, l_m$ and continuous longitudinal responses, respectively, for the ith subject measured at the jth time. Suppose that y is relevant to the latent continuous variable, y^*_{ij}, with cut points $l_1 l_2, \ldots, l_m$, as follows

$$\tilde{y}_{ij} = \begin{cases} 1 & l_0 < y^*_{ij} \le l_1 \\ 2 & l_1 < y^*_{ij} \le l_2 \\ \vdots & \vdots \\ m & l_{m-1} < y^*_{ij} \le l_m \end{cases}, \tag{8}$$

where $l_0 = -\infty$ and $l_m = \infty$. The SQR model cannot be straightly used for ordinal data. The monotone equivariant property (7) and the non-decreasing function (8) provide the necessary conditions for modeling ordinal data using the SQR model.

One approach to simultaneously perform the joint model analysis of continuous and ordinal responses is to consider correlations between them. For this purpose, we suggest the following joint model for y_{ij} and y^*_{ij}

$$\begin{aligned} y_{ij} &= \boldsymbol{x}^T_{\boldsymbol{1ij}} \boldsymbol{\beta_\tau} + \boldsymbol{Z}^T_{\boldsymbol{1i}} \boldsymbol{b}^{(1)}_{\boldsymbol{i}} + \eta_\tau(t_{ij}) + \varepsilon^y_{ij}, \\ y^*_{ij} &= \boldsymbol{x}^T_{\boldsymbol{2ij}} \boldsymbol{\alpha_\tau} + \boldsymbol{Z}^T_{\boldsymbol{2i}} \boldsymbol{b}^{(2)}_{\boldsymbol{i}} + \varepsilon^{y^*}_{ij}, \\ \varepsilon^y_{ij} &\sim ALD(0, \sigma_y, \tau_y), \varepsilon^{y^*}_{ij} \sim ALD(0, \sigma_{y^*}, \tau_{y^*}) \end{aligned} \tag{9}$$

where $\boldsymbol{x_{kij}} = (x_{kij1}, ..., x_{kijp})^T$, $k = 1, 2$ denotes the vector of p covariates observed at time t_{ij}, $\boldsymbol{\beta_\tau} = (\beta_1, ..., \beta_p)^T$ and $\boldsymbol{\alpha_\tau} = (\alpha_1, ..., \alpha_p)^T$ are $p \times 1$ vectors of fixed-effect regression coefficients, $\boldsymbol{Z_{ki}} = (z_{ki1}, ..., z_{kiq})^T$, $k = 1, 2$ is a $q \times 1$ vector of covariate connected with random effects, $\boldsymbol{b_i}^{(k)} = (b_{i1}, ..., b_{iq})^T$ is a $q \times 1$ vector of random-effect regression coefficients and $\boldsymbol{b_i} = (\boldsymbol{b_i}^{(1)}, \boldsymbol{b_i}^{(2)})$ such that $\boldsymbol{b_i} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma})$, $\eta_\tau(t_{ij})$ is a nonparametric part of the model, ε^{y}_{ij} and $\varepsilon^{y^*}_{ij}$ are the errors of the model which are independent of each other. In model (9) we assumed $\boldsymbol{b_i}$'s and ε_{ij}'s are independent. Then y_{ij} given $\boldsymbol{b_i}^{(1)}$s and y^*_{ij} given $\boldsymbol{b_i}^{(2)}$s have AL distribution. It means that

$$\begin{aligned} y_{ij}|\boldsymbol{b_i}^{(1)} &\sim ALD(\boldsymbol{x^T_{1ij}}\boldsymbol{\beta_\tau} + \boldsymbol{Z^T_{1i}}\boldsymbol{b_i}^{(1)} + \eta_\tau(t_{ij}), \sigma_y, \tau) \\ y^*_{ij}|\boldsymbol{b_i}^{(2)} &\sim ALD(\boldsymbol{x^T_{2ij}}\boldsymbol{\alpha_\tau} + \boldsymbol{Z^T_{2i}}\boldsymbol{b_i}^{(2)}, \sigma_{y^*}, \tau_{y^*}). \end{aligned} \tag{10}$$

So

$$p(\tilde{y}_{ij} = c) = p(y^*_{ij} \le l_c) = F_{\varepsilon^{y^*}_{ij}}(l_c - \boldsymbol{\alpha_\tau} + \boldsymbol{Z^T_{2i}}\boldsymbol{b_i}^{(2)}) \tag{11}$$

where $F_{\varepsilon^{y^*}_{ij}}$ is the CDF of $AL(0, \sigma_{y^*}, \tau_{y^*})$.

We will fix the variance of y^* to identify the ordinal response in the joint model (9) by considering $\sigma_{y^*} = 1$. We will also fix one of the cut points in (8). Traditionally, the l_0 has to be zero. The more properties of ordinal model's identifiability in the structure of a cut point notion is provided in [26].

To fit model (9), we use the regression spline method to handle the unknown nonparametric part of the model. In this method $\eta_\tau(t_{ij})$ can be approximated with a polynomial splines. Let (a_1, a_2) be an arbitrary interval and $a_1 = \upsilon_0 < \upsilon_1 < ... < \upsilon_{K_n} < \upsilon_{K_n+1} = a_2$ be a partition of (a_1, a_2) into $K_n + 1$ subintervals $I_{nj} = [\upsilon_j, \upsilon_{j+1})$, $j = 0, ..., K_n - 1$ and $I_{nK_n} = [\upsilon_{K_n}, a_2)$,where the number of internal knots K_n increases with the sample size. So within any two neighboring knots, Taylor's expansion up to some degrees is valid. A regression spline is a piecewise polynomial, which is a polynomial of some degrees within any two neighboring knots in I_{nj} for $j = 0, 1, ...K_n$ and is joined together at knots properly but allows discontinuous derivatives at the knots. The key for this method is to approximate $\eta_\tau(t_{ij})$ by a linear combination of a regression spline basis $\boldsymbol{\Psi}_p(t_{ij}) = (\Psi_0(t_{ij}), \Psi_1(t_{ij}), ..., \Psi_{p-1}(t_{ij}))^T$. $\boldsymbol{\Psi}_p(t_{ij})$ can be constructed using any basis such as the truncated power basis ([51, 47, 10]), B-spline basis [9], reproducing kernel Hilbert space basis ([50]), or wavelet basis, between others. When a truncated power basis or a B-spline basis is used in the regression spline, we need to specify the knots and select the number of the basis functions, p, for the proper performance of the regression spline. To select an optimal degree of regression spline and numbers of knots, that is, optimal sizes of p, the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) is often applied, see [52]. For given $\boldsymbol{\Psi}_p(t)$, we can approximate $\eta_\tau(t)$ using the basis vectors. That is,

$$\eta_\tau(t_{ij}) \approx \boldsymbol{\Psi}^T_p(t_{ij})\boldsymbol{\lambda}_p = \sum_{k=0}^{p-1} \lambda_k \Psi_k(t_{ij}) \tag{12}$$

where $\boldsymbol{\lambda}_p = (\lambda_0, \lambda_1, ..., \lambda_{p-1})^T$. Thus, the SQR joint model (9) associated with Equation (12) returns to a parametric QR model. For our model, we set the cubic spline bases with equally spaced knots $K_n = \left[n^{1/2p+3}\right]$ which was used by [53].

The likelihood function of the unknown parameters in model (9) for given τ and $\boldsymbol{y}_i = (y_{i1}, y_{i2}, ..., y_{in_i})^T$ and $\boldsymbol{y}^*_i = \left(y^*_{i1}, y^*_{i2}, ..., y^*_{in_i}\right)^T$ can be written as

$$\begin{aligned} L(\boldsymbol{\beta}_\tau,\sigma,\boldsymbol{\lambda}_p,\boldsymbol{\alpha}_\tau,\boldsymbol{l}|\boldsymbol{y}_i,\boldsymbol{y}_i^*,\boldsymbol{b_i}) &= f(\boldsymbol{y_i},\boldsymbol{y}_i^*|\boldsymbol{\beta}_\tau,\sigma,\boldsymbol{\lambda}_p,\boldsymbol{b_i},\boldsymbol{\alpha}_\tau,\boldsymbol{l}) \\ &= \prod_{j=1}^{n_i} f(y_{ij},y_{ij}^*|\boldsymbol{\beta}_\tau,\sigma,\boldsymbol{\lambda}_p,\boldsymbol{b_i},\boldsymbol{\alpha}_\tau,\boldsymbol{l}), \\ &= \prod_{j=1}^{n_i}\left\{ f(y_{ij}|\boldsymbol{\beta}_\tau,\sigma,\boldsymbol{b}_{\boldsymbol{i}}^{(1)},\boldsymbol{\lambda}_p) f(y_{ij}^*|\boldsymbol{\alpha}_\tau,\boldsymbol{b}_{\boldsymbol{i}}^{(2)},\boldsymbol{l})\right\}, \\ &= \prod_{j=1}^{n_i} f(y_{ij}|\boldsymbol{\beta}_\tau,\sigma,\boldsymbol{b}_{\boldsymbol{i}}^{(1)},\boldsymbol{\lambda}_p) \times \prod_{j=1}^{n_i} f(y_{ij}^*|\boldsymbol{\alpha}_\tau,\boldsymbol{b}_{\boldsymbol{i}}^{(2)},\boldsymbol{l}). \end{aligned} \tag{13}$$

And

$$\begin{aligned} L(\boldsymbol{\beta}_\tau,\sigma,\boldsymbol{\lambda}_p,\boldsymbol{\alpha}_\tau,\boldsymbol{l}|\boldsymbol{Y},\boldsymbol{Y^*},\boldsymbol{b}) &= f(\boldsymbol{Y},\boldsymbol{Y^*}|\boldsymbol{\beta}_\tau,\sigma,\boldsymbol{\lambda}_p,\boldsymbol{b},\boldsymbol{\alpha}_\tau,\boldsymbol{l}) \\ &= \prod_{i=1}^{n} f(\boldsymbol{y_i},\boldsymbol{y}_i^*|\boldsymbol{\beta}_\tau,\sigma,\boldsymbol{\lambda}_p,\boldsymbol{b_i},\boldsymbol{\alpha}_\tau,\boldsymbol{l}) \\ &= \prod_{i=1}^{n}\prod_{j=1}^{n_i} f(y_{ij}|\boldsymbol{\beta}_\tau,\sigma,\boldsymbol{b}_{\boldsymbol{i}}^{(1)},\boldsymbol{\lambda}_p) \times \prod_{i=1}^{n}\prod_{j=1}^{n_i} f(y_{ij}^*|\boldsymbol{\alpha}_\tau,\boldsymbol{b}_{\boldsymbol{i}}^{(2)},\boldsymbol{l}), \end{aligned}$$

where $\boldsymbol{Y} = \left(\boldsymbol{y}_1^T,\boldsymbol{y}_2^T,...,\boldsymbol{y}_n^T\right)^T, \boldsymbol{Y}^* = \left(\boldsymbol{y}_1^{*T},\boldsymbol{y}_2^{*T},...,\boldsymbol{y}_n^{*T}\right)^T$ and $\boldsymbol{b} = \left(\boldsymbol{b}_1^T,\boldsymbol{b}_2^T,...,\boldsymbol{b}_n^T\right)^T$. The third equality in (13) is obtained from the fact that $\boldsymbol{Y}$ and $\boldsymbol{Y}^*$ are independent of each other given $\boldsymbol{b}_i$.

As mentioned in the previous section, to generate a sampling algorithm for the quantile regression, we use a mixture representation based on the exponential and standard normal distributions, which can be detected in [32]. The stochastic representation (see equation 3) in model (9) is used to facilitate some analytical and numerical computations because working with the AL distribution is not easy. By studying the following illustration for the error of model (9), we have

$$\varepsilon_{ij}^y = \sqrt{a_1 u_{ij}}\, z_{ij} + a_2 u_{ij}$$

$$\varepsilon_{ij}^{y^*} = \sqrt{a_1 u_{ij}^*}\, z_{ij}^* + a_2 u_{ij}^*,$$

where $a_1 = \dfrac{2\sigma}{\tau(1-\tau)}$, $a_2 = \dfrac{1-2\tau}{\tau(1-\tau)}$, u_{ij}'s and u_{ij}^*'s have standard exponential distribution with mean σ and 1, respectively. z_{ij}'s and z_{ij}^*'s have standard normal distribution. Therefore, the conditional distribution of ε_{ij}^y given u_{ij} and $\varepsilon_{ij}^{y^*}$ given u_{ij}^* are

$$\varepsilon_{ij}^y\,|u_{ij} \sim N\left(a_2 u_{ij}, a_1 u_{ij}\right)$$

$$\varepsilon_{ij}^{y^*}\,|u_{ij}^* \sim N\left(a_2 u_{ij}^*, a_1 u_{ij}^*\right)$$

then

$$\begin{gathered} y_{ij}|u_{ij},\boldsymbol{b}_i^{(1)} \sim N\left(\mu_{ij}, a_1 u_{ij}\right) \\ y_{ij}^*|u_{ij}^*,\boldsymbol{b}_{\boldsymbol{i}}^{(2)} \sim N\left(\mu_{ij}^*, a_1 u_{ij}^*\right) \Rightarrow \tilde{y}_{ij}|u_{ij}^*,\boldsymbol{b}_{\boldsymbol{i}}^{(2)} \sim DU(\boldsymbol{\varphi}) \end{gathered} \tag{14}$$

where $\mu_{ij} = a_2 u_{ij} + E(y_{ij}|\boldsymbol{b}_{\boldsymbol{i}}^{(1)})$, $\mu_{ij}^* = a_2 u_{ij}^* + E(y_{ij}^*|\boldsymbol{b}_{\boldsymbol{i}}^{(2)})$ and $DU(\boldsymbol{\varphi})$ indicates discrete uniform distribution on vector of parameters $\boldsymbol{\varphi} = (\varphi_{ij,1},\varphi_{ij,2},...,\varphi_{ij,m})$.

The likelihood function of the refreshed parameters for joint model is given by

$$
\begin{aligned}
L\left(\boldsymbol{\beta}_\tau, \sigma, \boldsymbol{\lambda}_p, \boldsymbol{\alpha}_\tau, \boldsymbol{l} | \boldsymbol{Y}, \tilde{\boldsymbol{Y}}, \boldsymbol{b}, \boldsymbol{u}, \boldsymbol{u}^*\right) &= f(\boldsymbol{Y}, \tilde{\boldsymbol{Y}} | \boldsymbol{\beta}_\tau, \sigma, \boldsymbol{\lambda}_p, \boldsymbol{\alpha}_\tau, \boldsymbol{l}, \boldsymbol{b}, \boldsymbol{u}, \boldsymbol{u}^*) \\
&= \prod_{i=1}^{n} \prod_{j=1}^{n_i} f\left(y_{ij}, \tilde{y}_{ij} | \boldsymbol{\beta}_\tau, \sigma, \boldsymbol{\lambda}_p, \boldsymbol{\alpha}_\tau, \boldsymbol{b_i}, \boldsymbol{l}, u_{ij}, u_{ij}^*\right) \\
&= \prod_{i=1}^{n} \prod_{j=1}^{n_i} f\left(y_{ij} \,| \boldsymbol{\beta}_\tau, \sigma, \boldsymbol{\lambda}_p, \boldsymbol{b_i}^{(1)}, u_{ij}\right) \\
&\qquad \times \prod_{i=1}^{n} \prod_{j=1}^{n_i} f\left(\tilde{y}_{ij} | \boldsymbol{\alpha}_\tau, \boldsymbol{b_i}^{(2)}, \boldsymbol{l}, u_{ij}^*\right), \qquad (15)
\end{aligned}
$$

where $\boldsymbol{Y} = \left(\boldsymbol{y}_1^T, \boldsymbol{y}_2^T, ..., \boldsymbol{y}_n^T\right)^T$, $\tilde{\boldsymbol{Y}} = \left(\tilde{\boldsymbol{y}}_1^T, \tilde{\boldsymbol{y}}_2^T, ..., \tilde{\boldsymbol{y}}_n^T\right)^T$, $\boldsymbol{b} = \left(\boldsymbol{b}_1^T, \boldsymbol{b}_2^T, ..., \boldsymbol{b}_n^T\right)^T$, $\boldsymbol{l} = (l_1, l_2, ..., l_m)^T$, $\boldsymbol{u} = \left(\boldsymbol{u}_1^T, \boldsymbol{u}_2^T, ..., \boldsymbol{u}_n^T\right)^T$, $\boldsymbol{u}^* = \left(\boldsymbol{u}_1^{*T}, \boldsymbol{u}_2^{*T}, ..., \boldsymbol{u}_n^{*T}\right)^T$ and

$$
f\left(y_{ij} | \boldsymbol{\beta}_\tau, \sigma, \boldsymbol{b}_i^{(1)}, u_{ij}^*\right) = \frac{1}{\gamma_{ij}\ \sqrt{2\pi}} \exp\left\{ \frac{-1}{2} \left(\frac{y_{ij} - \mu_{ij}}{\gamma_{ij}} \right) \right\}
$$

$$
\begin{aligned}
f\left(\tilde{y}_{ij} | \boldsymbol{\alpha}_\tau, \boldsymbol{b}_i^{(2)}, \boldsymbol{l}, u_{ij}\right) &= p(\tilde{y}_{ij} = k | \boldsymbol{\alpha}_\tau, \boldsymbol{b}_i^{(2)}, \boldsymbol{l}, u_{ij}^*) \\
&= \varphi(ij, k) = \Phi\left(\frac{l_k - \mu_{ij}^*}{\gamma_{ij}^*}\right) - \Phi\left(\frac{l_{k-1} - \mu_{ij}^*}{\gamma_{ij}^*}\right)
\end{aligned}
$$

in which

$$
\gamma_{ij} = \sqrt{\frac{2u_{ij}\sigma_y}{\tau(1-\tau)}}, \ \ \gamma_{ij}^* = \sqrt{\frac{2u_{ij}^*}{\tau(1-\tau)}}.
$$

3. Bayesian inferences

This section performs the Bayesian approach using MCMC methods to estimate the previous section's model parameters. Bayesian approaches provide convenient alternative inference tools for quantile regression. Even though conventional quantile regression does not need any parametric distributional assumptions, a working likelihood is required to carry out Bayesian analysis. Suppose $\boldsymbol{\Theta} = (\boldsymbol{\beta}_\tau, \boldsymbol{\alpha}_\tau, \sigma_y, \boldsymbol{\lambda}_p, \boldsymbol{b}, \boldsymbol{u}, \boldsymbol{u}^*, \boldsymbol{l})$ be the vector of unknown parameters in (15). We allocate independent prior distributions for each elements of $\boldsymbol{\Theta}$. We consider the following priors in the Bayesian hierarchical framework to the SQR model parameters for $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n_i$ and $k = 1, ..., m$ as

$$
\begin{aligned}
& y_{ij} | \boldsymbol{b}_i^{(1)}, u_{ij}, \boldsymbol{\beta}_\tau, \sigma_y \sim N\left(\mu_{ij}, a_1 u_{ij}\right), \\
& \tilde{y}_{ij} | \boldsymbol{b}_i^{(2)}, u_{ij}^*, \boldsymbol{\alpha}_\tau \sim DU\left(\varphi_{ij,1}, \varphi_{ij,2}, ..., \varphi_{ij,m}\right), \\
& \boldsymbol{b}_i | \boldsymbol{\Sigma_b} \sim N\left(\boldsymbol{0}, \boldsymbol{\Sigma_b}\right), \\
& u_{ij} \sim \exp(\sigma), \\
& u_{ij}^* \sim \exp(1), \\
& \boldsymbol{\alpha}_\tau \sim N\left(\boldsymbol{\mu}_\alpha, \boldsymbol{\Omega}_1\right), \qquad (16) \\
& \boldsymbol{\beta}_\tau \sim N\left(\boldsymbol{\mu}_\beta, \boldsymbol{\Omega}_2\right), \\
& \boldsymbol{\lambda}_p \sim N\left(\boldsymbol{\mu}_\lambda, \boldsymbol{\Omega}_3\right), \\
& l_k | l_1, l_2, ..., l_{k-1} \sim TN\left(\mu_{l_k}, \sigma_{l_k}, l_{k-1}, \infty\right), \\
& \boldsymbol{\Sigma_b} \sim IW\left(\boldsymbol{\Omega_4}, \omega_1\right), \\
& \sigma_y \sim IG\left(\omega_2, \omega_3\right),
\end{aligned}
$$

such that μ_{ij}, $\varphi_{ij,k}$ and a_1 are defined in the previous sections and $\omega_r(r = 1, 2, 3)$ are constants. The hyper-parameter matrices $\mathbf{\Omega}_s$ $(s = 1, ..., 4)$ can be assumed to be diagonal for convenient implementation. IG and IW also indicate inverse gamma and inverse Wishart distributions, respectively. An MCMC method, such as the Gibbs sampling algorithm, can perform Bayesian analysis as a numerical method. We need to determine the full conditional distributions of parameters to receive this goal. Since the posterior distributions are analytically intractable, we can use a Metropolis-Hastings algorithm to generate the parameters' full-conditional distributions.

3.1. Convergence diagnostic and Deviance information criteria (DIC)

One of the most critical steps of the Bayesian MCMC algorithm is to determine the chain's convergence and the burn-in period's length, typically assessed using diagnostics tools. For this aim, we use several chains with several initial values and then analyse their behaviours. Therefore chains will be converged when they begin to generate the same values. The Gelman–Rubin diagnostic ([17]) assesses MCMC convergence by analysing the difference between multiple Markov chains. This method evaluated the convergence by comparing the estimated Inter-chain and Intra-chains variances for each model parameter. Significant differences between these variances indicate non-convergence. Let M be the number of chains with length N, although the chains may be of different lengths. The same-length opinion is used for convenience and simplifies the formulas. For a model parameter θ, let $\{\theta_{mt}\}_{t=1}^{N}$ be the mth simulated chain, $m = 1, \ldots, M$. Let $\hat{\theta}_m$ and $\hat{\sigma}_m^2$ be the sample posterior mean and variance of the mth chain, and let the overall sample posterior mean be $\hat{\theta} = (1/M)\sum_{m=1}^{M}\hat{\theta}_m$. The between-chains and within-chain variances are given by

$$B = \frac{N}{M-1}\sum_{m=1}^{M}(\hat{\theta}_m - \hat{\theta})^2$$

$$W = \frac{1}{M}\sum_{m=1}^{M}\hat{\sigma}_m^2$$

Under certain stationarity conditions, the pooled variance is as follows

$$\widehat{V} = \frac{N-1}{N}W + \frac{M+1}{MN}B$$

then

$$\widehat{R} = \frac{\widehat{V}}{W}$$

The chains will be converging If the produced data-size is large. The Intra-chain variance should be relatively small than the Inter-chain variance; thus, R goes to one. In other words, R values greater than one show non-convergence.

DIC is a metric used for model comparison in the Bayesian approach. This criterion is a reduced version of the DIC introduced by [48]. DIC consists of two parts based on the posterior distribution. The first part, the posterior mean, measures the goodness of fit and the second part, the effective number of parameters, is the penalty for increasing model complexity. The DIC formulation is:

$$DIC = \bar{D}(\theta) + p_D,$$

where $p_D = \bar{D}(\theta) - D(\bar{\theta})$, and $D(\theta) = E_{\theta|y} - 2\left[log\left(f(\boldsymbol{y}|\theta)\right)\right]$. $\bar{D}(\theta)$ and $\bar{\theta}$ are the posterior mean estimates of $D(\theta)$ and θ, respectively. Additionally $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ is the complete data with density $f(\boldsymbol{y}\,|\theta)$ and θ is the vector of unknown parameters. We can obtain a Monte Carlo approximation of DIC Utilizing the MCMC samples of the posterior distribution. The lowest value of DIC indicates a better fit of a model to the observations.

4. Simulation

To specify the proposed SQR joint model's performance, we carry out some MCMC simulation studies. Three scenarios for simulation study are considered such that the continuous response distribution is ALD with skew parameters $\tau = 0.25, 0.50$, and 0.75. Each generated dataset is fitted by a semi-parametric Quantile Regression joint Model (SQRJM) for $\tau = 0.25$, $\tau = 0.5$ and $\tau = 0.75$ and later we compare performance of our method with that of a standard method. These datasets are analyzed by a Classic Mean Regression joint Model (CMRJM). The models' performances are evaluated using two criteria: the relative bias (R.Bias) and root of mean-squared errors (RMSE). These are given by

$$R.Bias = \frac{1}{N}\sum_{l=1}^{N}\left(\frac{\hat{\theta}_l - \theta}{\theta}\right)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{l=1}^{N}\left(\hat{\theta}_l - \theta\right)^2}$$

where $\hat{\theta}_l$ is the estimate of θ for the lth sample, and N is the number of simulated samples (N =100 in this section). We run the Gibbs sampler algorithm for 10,000 iterations, and a burn-in period of 5000 iterations are chosen. The last 5000 iterations are considered to get the Bayesian estimation. To fit our suggested model to produce data in a Bayesian structure, we use non-informative prior distributions to estimate parameters by considering $(\boldsymbol{\mu_\beta}, \boldsymbol{\Omega_2}) = (\mathbf{0}_3, 100\boldsymbol{I}_3)$, $(\boldsymbol{\mu_\alpha}, \boldsymbol{\Omega_1}) = (\mathbf{0}_4, 100\boldsymbol{I}_4)$, $(\mu_l, \sigma_l) = (0, 100)$ and $(\omega_2, \omega_3) = (0.01, 0.01)$ in Equation (16). Also, an inverse Wishart prior distribution with parameters $(\Omega_4, \omega_1) = (\boldsymbol{I}_2, 5)$ are considered for $\boldsymbol{\Sigma_b}$. The MCMC method is performed using OpenBUGS software which interacts with R software by R2OpenBUGS package. This software is available online free of cost. See [41] and [48].

In this simulation study, we generate the data set in three situations i.e. the continuous response is assumed to have the AL distribution for $\tau = 0.25, 0.50$, and 0.75. We generate 100 data sets (iterations) based on the above mentioned three cases. Each generated dataset is fitted by SQRJMs for one of $\tau = 0.25$ or $\tau = 0.5$ or $\tau = 0.75$ that the data is generated from and CMRJM. This simulation performs longitudinal studies with $n = 250$ and $n = 500$ subjects that the responses have been measured 4 times. The continuous and ordinal longitudinal responses are sampled from the following SQRJM for $i = 1, 2, ..., n$ and $j = 1, 2, 3, 4$

$$\begin{aligned} y_{ij} &= \beta_0 + \beta_1 x_{ij1} + \beta_2 x_{ij2} + \eta_{ij}(t) + b_{1i} + \varepsilon_{ij}^{y} \\ y_{ij}^{*} &= \alpha_0 + \alpha_1 x_{ij1} + \alpha_2 x_{ij2} + \alpha_3 t_j + b_{2i} + \varepsilon_{ij}^{y^*} \end{aligned} \qquad (17)$$

where $(\beta_0, \beta_1, \beta_2) = (-1, 3, 2), (\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (-1, 1, -1, 1), t_j = j$ and $\eta_{ij}(t)$ is approximated with cubic polynomial $3(1 + t + t^2 + t^3)$. x_{ij1} and x_{ij2} are generated from Bernoulli distribution with $p = 0.5$ and standard normal distribution, respectively. Also ε_{ij}^{y}s and $\varepsilon_{ij}^{y^*}$s sampled from $ALD(0, \sigma_y, \tau)$ and $ALD(0, 1, \tau)$, respectively. The vector of random effects b is sampled from a bivariate normal distribution with mean $\mathbf{0}$ and the following covariance matrix

$$\mathbf{\Sigma_b} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

$(l_0, l_1, l_3) = (0, 10, 20)$ is the vector of cut points that used in the ordinal model. Thus, the ordinal responses in model (17) are sampled according to

$$\tilde{y}_{ij} = \begin{cases} 1 & y_{ij}^{*} \le l_0 \\ 2 & l_0 < y_{ij}^{*} \le l_1 \\ 3 & l_1 < y_{ij}^{*} \le l_2 \\ 4 & y_{ij}^{*} > l_2 \end{cases}.$$

Two parallel MCMC chains with different starting values are performed for 10,000 iterations after a burn-in period of 5000 iterations. In each iteration, the convergence of chains is controlled by using the Gelman-Rubin diagnostic test. If the chains are not converged based on this in each iteration, then the sample is discarded.

Table 1. Results of simulation study from fitting SQRJM(0.25) and CMRJM to generated data by SQRJM(0.25).

			SQRJM(0.25)				CMRJM			
n	parameters	Real	Est.	SD	R.Bias	RMSE	Est.	SD	R.Bias	RMSE
250	β_0	-1	-1.077	0.125	0.077	0.147	1.216	0.333	-2.216	2.241
	β_1	3	2.943	0.187	-0.019	0.195	2.658	0.280	-0.114	0.443
	β_2	2	2.084	0.097	0.042	0.129	1.936	0.143	-0.032	0.156
	α_0	-1	-0.998	0.336	-0.002	0.336	-0.290	7.176	-0.710	7.211
	α_1	1	1.057	0.302	0.057	0.307	-0.356	7.192	-1.356	7.319
	α_2	-1	-0.749	0.145	-0.251	0.290	-0.301	6.895	-0.699	6.931
	α_3	1	0.899	0.143	-0.101	0.175	-0.105	7.227	-1.105	7.311
	σ_{11}	1	0.517	0.173	-0.483	0.513	0.753	0.385	-0.247	0.457
	σ_{12}	0.5	0.302	0.204	-0.396	0.284	0.057	0.913	-0.887	1.015
	σ_{22}	1	0.697	0.432	-0.303	0.527	7.010	0.258	6.010	6.016
	σ_y	1	0.984	0.033	-0.016	0.037	4.015	0.099	3.015	3.016
	l_1	10	9.370	0.427	-0.063	0.762	12.213	0.714	0.221	2.325
	l_2	20	20.056	1.387	0.003	1.388	16.356	1.198	-0.182	3.836
500	β_0	-1	-0.948	0.115	-0.052	0.127	2.007	0.260	-3.007	3.018
	β_1	3	3.011	0.155	0.004	0.156	2.386	0.302	-0.205	0.684
	β_2	2	2.139	0.081	0.070	0.161	1.932	0.141	-0.034	0.156
	α_0	-1	-0.910	0.220	-0.090	0.238	0.716	7.069	-1.716	7.274
	α_1	1	1.213	0.208	0.213	0.298	0.172	6.965	-0.828	7.014
	α_2	-1	-0.897	0.117	-0.103	0.156	0.017	6.994	-1.017	7.067
	α_3	1	0.911	0.092	-0.089	0.128	-0.029	7.187	-1.029	7.261
	σ_{11}	1	1.118	0.193	0.118	0.226	0.913	0.581	-0.087	0.588
	σ_{12}	0.5	0.389	0.183	-0.221	0.214	1.899	1.518	2.799	2.064
	σ_{22}	1	0.403	0.197	-0.597	0.629	2.321	1.073	1.321	1.702
	σ_y	1	1.015	0.025	0.015	0.029	4.330	0.119	3.330	3.332
	l_1	10	9.516	0.301	-0.048	0.570	7.550	1.162	-0.245	2.712
	l_2	20	18.543	0.797	-0.073	1.660	16.134	1.484	-0.193	4.141

Est.: estimate of parameter, S.D.: posterior standard deviation, R.bias: relative biase, RMSE: root of mean-square errors.

Table 1 summarizes simulation results which include estimates of parameters, standard deviations, relative bias and root of MSE (RMSE) for the SQRJM(0.25) and CMRJM with $n = 250$ and $n = 500$ subjects. The relative biases, standard errors, and RMSEs of estimated parameters is decreased in SQRJM(0.25) when the number of subjects is increased. This feature shows that the recommended method provides consistency properties of estimates (see Figure 2). In spite of having low standard deviation among the estimated parameters in CMRJM the performance of the SQRJM is better than the CMRJM because it has lower relative biases and RMSE among the estimated parameters. Table 2 and

Table 2. Results of simulation study from fitting SQRJM(0.5) and CMRJM to generated data by SQRJM(0.5).

			SQRJM(0.50)				CMRJM			
n	parameters	Real	Est.	SD	R.Bias	RMSE	Est.	SD	R.Bias	RMSE
250	β_0	-1	-1.002	0.136	0.002	0.136	-1.183	0.157	0.183	0.241
	β_1	3	2.967	0.193	-0.011	0.195	3.049	0.224	0.016	0.229
	β_2	2	2.007	0.097	0.004	0.097	1.918	0.109	-0.041	0.136
	α_0	-1	-0.996	0.292	-0.004	0.292	-0.542	0.119	-0.458	0.473
	α_1	1	0.945	0.264	-0.055	0.269	0.404	0.111	-0.596	0.606
	α_2	-1	-0.984	0.142	-0.016	0.143	-0.427	0.059	-0.573	0.576
	α_3	1	0.994	0.116	-0.006	0.117	0.442	0.041	-0.558	0.560
	σ_{11}	1	0.942	0.211	-0.058	0.219	0.956	0.283	-0.044	0.286
	σ_{12}	0.5	0.511	0.195	0.022	0.195	0.342	0.098	-0.315	0.186
	σ_{22}	1	0.788	0.355	-0.212	0.414	0.285	0.077	-0.715	0.719
	σ_y	1	1.004	0.034	0.004	0.035	2.899	0.073	1.899	1.900
	l_1	10	10.000	0.607	0.000	0.607	8.605	0.173	-0.140	1.406
	l_2	20	19.257	3.932	-0.037	4.002	17.265	5.157	-0.137	5.853
500	β_0	-1	-0.994	0.136	-0.006	0.136	-1.087	0.161	0.087	0.183
	β_1	3	3.015	0.192	0.005	0.193	3.198	0.223	0.066	0.298
	β_2	2	2.015	0.096	0.008	0.097	2.007	0.108	0.003	0.109
	α_0	-1	-0.960	0.290	-0.040	0.293	-0.165	7.422	-0.835	7.469
	α_1	1	0.974	0.261	-0.026	0.262	0.452	7.425	-0.548	7.445
	α_2	-1	-0.977	0.142	-0.023	0.144	-0.518	7.127	-0.482	7.143
	α_3	1	0.980	0.117	-0.020	0.119	-0.335	7.589	-1.335	7.706
	σ_{11}	1	0.903	0.206	-0.097	0.228	0.477	0.210	-0.523	0.563
	σ_{12}	0.5	0.426	0.192	-0.149	0.205	5.748	8.399	10.496	9.903
	σ_{22}	1	0.745	0.347	-0.255	0.431	1.321	1.916	0.321	1.943
	σ_y	1	1.008	0.035	0.008	0.036	2.990	0.076	1.990	1.991
	l_1	10	9.965	0.603	-0.003	0.604	9.660	2.561	-0.034	2.583
	l_2	20	19.400	4.010	-0.030	4.055	24.060	3.364	0.203	5.273

Est.: estimate of parameter, S.D.: posterior standard deviation, R.bias: relative biase, RMSE: root of mean-square errors.

Table 4 contains estimates of parameters, standard deviations, relative biases and RMSEs estimates by models SQRJM(0.5) and CMRJM and SQRJM(0.75) and CMRJM, respectively. The results from Tables 2 and 4 have interpretation similar to those in Table 1. Figure 2 shows that how the RMSE varies with the variation of the sample size n and consistent estimators in the SQRJM(0.75) model when data are generated by SQRJM(0.75). For all scenarios considered in this simulation study, it is of interest to see that when an estimated bias for an estimator is negative that parameter is underestimated and when an estimated bias for an estimator is positive that parameter is overestimated. To assess the computational time and efficiency, we compute the elapsed time for fitting the different methods. The computation time to estimate for each of the 100 simulated datasets using a computer equipped with an Intel(R) Core (TM) i5-2500 CPU, and 4.00 GB RAM for different models are given in Table 3.

Table 3. The elapsed time (in seconds) for fitting the different models.

models	$n = 250$	$n = 500$
SQRJM(0.25)	1556.34	3326.81
SQRJM(0.50)	1408.61	3234.99
SQRJM(0.75)	1499.79	3292.36
CMRJM	701.56	1754.41

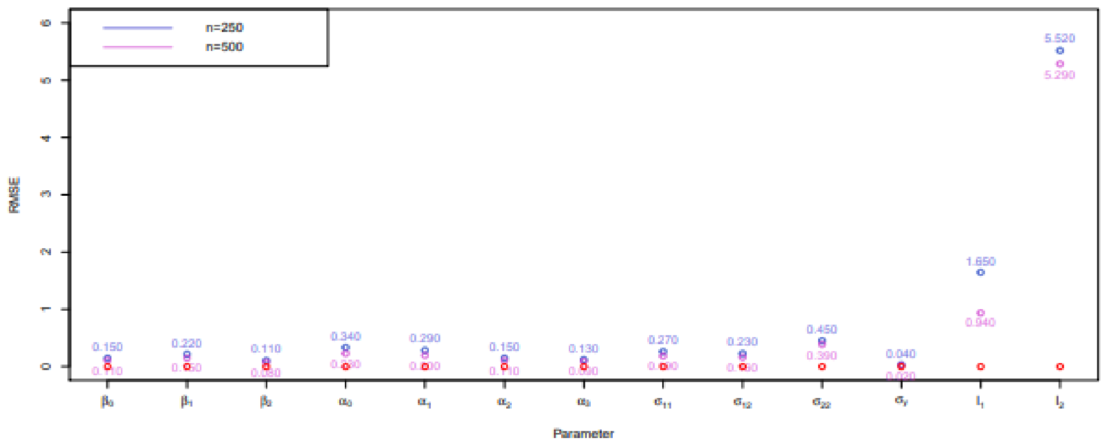

Figure 2. RMSE of parameter estimates in the SQRJM(0.75) model for n = 250 and 500, (parameter estimates are given inside the plot with the sign "∘").

5. Application

In this section, to specify the performance of the proposed SQRM, we analyse a set of real data i.e. the Peabody Individual Achievement Test (PIAT) data set. The PIAT data set was previously analyzed by [43], [39], and [19]. The PIAT data were prepared at two-year intervals between 1986 and 1992 and contains both the mother and her child's interviews. This paper has selected two interesting continuous and ordinal responses from the dataset as the primary variables for modeling. The continuous response is the child's reading recognition skill (read) and the ordinal response is the child's antisocial behavior (anti). The study's main idea was to specify the connection between parental response on a child's antisocial behavior and the child's reading recognition skill.

Children's reading recognition skill was measured by the PIAT reading recognition subtests. The reading recognition subtests measure word recognition and pronunciation ability. These components are considered essential to reading achievement. First, the children read a word silently, then say it aloud. the PIAT Reading Recognition contains 84 items, each of them has four possible response, which increases in difficulty from preschool to high school levels, one of them is the correct response. Skills evaluated containing matching letters, naming names, and reading single words aloud. The reading recognition obtained by summing the total number of correct items for the 84 items subtest and scores could range in value from 0 to 84. The final reading recognition scores were divided by 10. These data consist of 221 children and mothers who had perfect answers at all time. These individuals data have been studied as

Table 4. Results of simulation study from fitting SQRJM(0.75) and CMRJM to generated data by SQRJM(0.75).

			SQRJM(0.75)				CMRJM			
n	parameters	Real	Est.	SD	R.Bias	RMSE	Est.	SD	R.Bias	RMSE
250	β_0	-1	-0.997	0.152	-0.003	0.152	-3.645	0.205	2.645	2.653
	β_1	3	3.045	0.213	0.015	0.218	2.996	0.286	-0.001	0.286
	β_2	2	1.994	0.108	-0.003	0.108	2.003	0.146	0.001	0.146
	α_0	-1	-0.914	0.325	-0.086	0.337	-0.837	1.742	-0.163	1.750
	α_1	1	0.964	0.284	-0.036	0.286	0.283	2.007	-0.717	2.131
	α_2	-1	-0.998	0.153	-0.002	0.153	-0.279	1.651	-0.721	1.801
	α_3	1	0.976	0.123	-0.024	0.126	0.269	1.692	-0.731	1.843
	σ_{11}	1	0.919	0.252	-0.081	0.265	0.611	0.344	-0.389	0.520
	σ_{12}	0.5	0.445	0.223	-0.110	0.230	1.361	2.068	1.722	2.240
	σ_{22}	1	0.742	0.374	-0.258	0.455	6.021	0.971	5.021	5.114
	σ_y	1	1.007	0.035	0.007	0.035	4.224	0.104	3.224	3.226
	l_1	10	10.704	1.488	0.070	1.646	13.301	2.973	0.330	4.442
	l_2	20	16.872	4.545	-0.156	5.517	27.022	5.196	0.351	8.735
500	β_0	-1	-0.971	0.108	-0.029	0.111	-3.639	0.146	2.639	2.643
	β_1	3	2.989	0.151	-0.004	0.151	2.986	0.205	-0.005	0.205
	β_2	2	2.008	0.076	0.004	0.077	2.014	0.105	0.007	0.106
	α_0	-1	-0.952	0.228	-0.048	0.233	-0.796	2.114	-0.204	2.124
	α_1	1	0.983	0.198	-0.017	0.199	0.260	2.355	-0.740	2.468
	α_2	-1	-0.985	0.107	-0.015	0.108	-0.286	2.051	-0.714	2.172
	α_3	1	0.976	0.087	-0.024	0.091	0.216	2.055	-0.784	2.199
	σ_{11}	1	0.956	0.180	-0.044	0.186	0.685	0.313	-0.315	0.444
	σ_{12}	0.5	0.498	0.165	-0.004	0.165	2.089	2.092	3.179	2.627
	σ_{22}	1	0.742	0.293	-0.258	0.391	2.324	0.809	1.324	1.552
	σ_y	1	1.002	0.024	0.002	0.024	4.237	0.119	3.237	3.240
	l_1	10	10.219	0.913	0.022	0.939	11.693	1.076	0.169	2.006
	l_2	20	17.067	4.398	-0.147	5.286	25.511	5.591	0.276	7.851

Est.: estimate of parameter, S.D.: posterior standard deviation, R.bias: relative biase, RMSE: root of mean-square errors.

complete data in this paper. Dataset also includes some covariates which will be used in our analysis: child's gender (Gender, girl=0 and boy=1), mother's age at first time-point (Momage, from age 21 up to age 29), child's age at first time-point (Kidage, from age six up to age 8) and time (Time=$1, 2, 3, 4$ corresponds to the first, second, third and fourth time-points, respectively).

5.1. Model implementation

We use linear combinations of cubic splines with equally spaced knots to approximate the nonparametric functions $\eta_\tau(t)$ (see section 2.2). There are 221 children and mothers, then $K_n = \lceil n^{1/2p+3} \rceil = 1$. The following joint model is considered for reading ability (*read*) and latent variable antisocial behavior (*anti**), for $i = 1, 2, \ldots, 221$ and $j = 1, 2, 3, 4$

$$\begin{aligned} read_{ij} &= \beta_1 Gender_i + \beta_2 Kidage_i + \beta_3 Momage_i + \eta_\tau(t_{ij}) + b_{1i} + \varepsilon_{ij}^{read}, \\ anti_{ij}^* &= \alpha_0 + \alpha_1 Gender_i + \alpha_2 Kidage_i + \alpha_3 Momage_i + \alpha_4 t_{ij} + b_{2i} + \varepsilon_{ij}^{anti}, \end{aligned} \tag{18}$$

where $\boldsymbol{b}_i = (b_{1i}, b_{2i})$ is the normally distributed vector of random effects with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}_b$ and ε_{ij}^{read} and ε_{ij}^{anti} are independently follows as $AL(0, \sigma, \tau)$ and $AL(0, 1, \tau)$, respectively. The continuous covariates ($Momage$ and $Kidage$) are also standardized. $\eta(t_{ij})$ is a non-parametric part of the model approximated by linear combinations of cubic splines. For Bayesian inference, we set $\Psi_0(t) \equiv 1$ and take the truncated power basis in the approximations (12) that suggesting the following function for $\eta_\tau(t)$.

$$\eta_\tau(t) = \lambda_0 + \lambda_1 t + \lambda_2 t^2 + \lambda_3 t^3 + \lambda_4 (t - \upsilon_1)^3.$$

According to equations (11), the corresponding cumulative probabilities of $anti_{ij}$ at the cth ($c = 1, 2$) category are derived from (18) as

$$\begin{aligned} p(anti_{ij} \leq c | \boldsymbol{\alpha}, b_{2i}, l) = F_{AL}(l_c - (&\alpha_0 + \alpha_1 Gender_i + \alpha_2 Kidage_i \\ &+ \alpha_3 Momage_i + \alpha_4 t_j + b_{2i})), \end{aligned} \tag{19}$$

where F_{AL} is the CDF of AL distribution with parameters $(0, 1, \tau)$ and relation between $anti_{ij}^*$ and $anti_{ij}$ is given by

$$anti_{ij} = I_{(-\infty, l_1]}(anti_{ij}^*) + 2I_{(l_1, l_2]}(anti_{ij}^*) + 3I_{(l_2, \infty)}(anti_{ij}^*).$$

The prior distributions used in the Bayesian implementation were similar to the simulation study's prior distributions. We run two parallel MCMC chains with distinctive initial values for 60,000 iterations, and ignore the first 20,000 iterations to do the Bayesian inference. Figure 3 shows the output diagrams of the Gelman-Rubin diagnostics for SQRJM(0.75) of different parameters. Three curves exist in each plot of which one is a horizontal line, which shows the value of one. The curves under the horizontal line display $\hat{V}$ in green color and W in blue color, and the top curve above the horizontal line display the ratio of $\hat{V}$ and W, i.e. $\hat{R}$. As it is seen, when the number of replication increases, $\hat{V}$ and W are steadying, and $\hat{R}$ is conducing to one. Figure 3 shows the convergence of the MCMC chains.
We fit four different models to PIAT data. In order to compare the models, the DIC criterion is used. In the fitted models, we use mean regression and the τth QR model for responses by choosing $\tau = 0.25, 0.50$ and 0.75. The following non-informative prior distributions are employed for a Bayesian implementation.

$$\begin{aligned} \boldsymbol{\beta} &\sim N_3(\mathbf{0}_3, 100\boldsymbol{I}_3), &\quad \boldsymbol{\beta} &= (\beta_1, \beta_2, \beta_3)', \\ \boldsymbol{\lambda} &\sim N_5(\mathbf{0}_5, 100\boldsymbol{I}_5), &\quad \boldsymbol{\lambda} &= (\lambda_0, \lambda_1, \lambda_2, \lambda_3, \lambda_4)', \\ \boldsymbol{\alpha} &\sim N_4(\mathbf{0}_4, 100\boldsymbol{I}_4), &\quad \boldsymbol{\alpha} &= (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4)', \\ \boldsymbol{\Sigma_b} &\sim IW(I_2, 5), \\ \sigma_y &\sim IGamma(0, 01, 0, 01). \end{aligned}$$

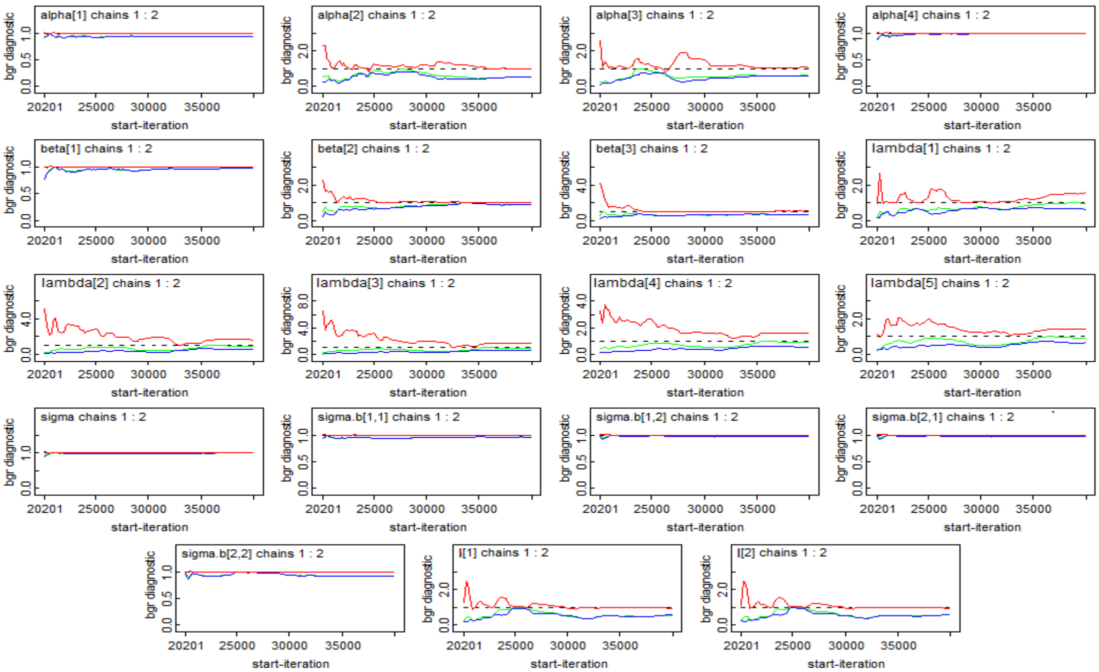

Figure 3. Gelman-Rubin diagnostic plots for all parameters in the SQRJM(0.75).

5.2. Results of data analysis

Bayesian modeling approach in the SQR joint model at the three quantiles of $\tau = 0.25, 0.50$, and 0.75 was used to fit the PIAT data. Results in Table 5 show posterior mean, standard deviation, and Bayesian criteria for comparing models. Results of Table 5 are obtained using SQRJM(0.25), SQRJM(0.50), SQRJM(0.75), and CMRJM. According to the results of Table 5, SQRJM(0.25) has the lowest DIC, and by this continuous model, there are no significant differences between boys' and girls' reading ability. The mother's age and her child's age have significant effects on reading ability. It shows that older mothers can increase their children's reading ability more than younger mothers, and older children also have a higher reading ability than younger children. As we expect, time is significant in continuous models. It means that children's reading ability is developing over time. This result follows what we observed in Figure 1. Additionally, antisocial behavior depends on the child's gender in the ordinal model. Considering the α_1 is significant, this result is obtained that boys more antisocial behavior than girls at all time points. The other covariates (Momage, Kidage and time) do not affect antisocial behavior.

According to DIC values in Table 5, the SQRJM(τ) model is an excellent fit compared to that of the CMRJM model. SQRJM(0.25) provides the lowest DIC value, then SQRJM(0.25) gives the most suitable model in existing models based on DIC values.

6. Conclusion

This paper aimed to use the SQJRM for mixed continuous and ordinal longitudinal data using random effects. However, QR originally does not require any distributional assumption; an AL distribution was allocated to the mixed effect model's error for continuous data in our approach. The QR model is usually used for continuous responses, but it could not be handled straight to the ordinal response. In this

Table 5. Results of analyzing real data from fitting SQRJM($\tau = 0.25, 0.50$ and 0.75) and CMRJM to PIAT data.

	SQRJM($\tau = 0.25$)			SQRJM($\tau = 0.50$)		
parameters	Est.	S.D.	95% PCI	Est.	S.D.	95% PCI
β_1	-0.128	0.117	(-0.322 , 0.066)	-0.109	0.117	(-0.303 , 0.083)
β_2	0.376	0.108	(0.156 , 0.549)	0.385	0.091	(0.235 , 0.537)
β_3	0.108	0.032	(0.051 , 0.153)	0.112	0.030	(0.065 , 0.160)
λ_0	-3.558	1.211	(-5.614 , -1.666)	-4.807	1.321	(-6.858 , -2.474)
λ_1	1.042	0.841	(0.274 , 3.076)	2.094	1.685	(0.175 , 4.701)
λ_2	0.448	0.255	(0.079 , 0.973)	-0.357	0.522	(-1.165 , 0.363)
λ_3	-0.041	0.025	(-0.094 , -0.005)	0.034	0.050	(-0.036 , 0.115)
λ_4	-0.202	0.048	(-0.282 , -0.129)	-0.042	0.083	(-0.155 , -0.014)
σ_y	0.167	0.006	(0.157 , 0.178)	0.227	0.008	(0.214 , 0.241)
α_1	1.316	0.440	(0.604 , 2.047)	1.315	0.449	(0.583 , 2.064)
α_2	0.232	0.349	(-0.287 , 0.811)	0.137	0.338	(-0.487 , 0.661)
α_3	-0.025	0.123	(-0.200 , 0.219)	-0.013	0.123	(-0.217 , 0.189)
α_4	0.098	0.108	(-0.078 , 0.275)	0.098	0.110	(-0.083 , 0.280)
l_1	-0.067	3.008	(-5.280 , 4.899)	-0.433	3.289	(-5.962 , 4.593)
l_2	8.511	3.029	(3.348 , 13.550)	8.152	3.283	(2.673 , 13.240)
σ_{11}	0.647	0.070	(0.541 , 0.770)	0.650	0.072	(0.540 , 0.776)
σ_{12}	-0.622	0.197	(-0.956 , -0.312)	-0.669	0.202	(-1.013 , -0.352)
σ_{22}	6.962	1.461	(4.828 , 9.570)	7.005	1.473	(4.836 , 9.635)
DIC		2144			2247	
	SQRJM($\tau = 0.75$)			CMRJM		
parameter	Est.	S.D.	95% PCI	Est.	S.D.	95% PCI
β_1	-0.069	0.118	(-0.262 , 0.127)	-0.091	0.116	(-0.281 , 0.101)
β_2	0.375	0.096	(0.222 , 0.530)	0.379	0.092	(0.230 , 0.532)
β_3	0.116	0.030	(0.067 , 0.165)	0.116	0.030	(0.067 , 0.167)
λ_0	-3.644	1.024	(-5.331 , -1.854)	-6.700	2.011	(-9.784 , -2.984)
λ_1	1.011	1.184	(-1.442 , 2.682)	5.015	2.985	(-0.796 , 9.836)
λ_2	-0.013	0.369	(-0.540 , 0.748)	-1.249	0.921	(-2.726 , 0.534)
λ_3	0.003	0.036	(-0.067 , 0.054)	0.116	0.086	(-0.048 , 0.253)
λ_4	-0.063	0.060	(-0.180 , -0.026)	-0.096	0.144	(-0.176 , -0.009)
σ_y	0.180	0.007	(0.169 , 0.191)	0.614	0.017	(0.587 , 0.642)
α_1	1.316	0.447	(0.591 , 2.061)	0.497	0.166	(0.226 , 0.772)
α_2	0.135	0.356	(-0.474 , 0.647)	0.087	0.139	(-0.134 , 0.319)
α_3	-0.022	0.109	(-0.187 , 0.166)	-0.013	0.042	(-0.086 , 0.058)
α_4	0.098	0.111	(-0.085 , 0.281)	0.048	0.041	(-0.020 , 0.115)
l_1	-0.665	3.393	(-6.657 , 4.487)	-0.113	1.412	(-2.495 , 2.265)
l_2	7.918	3.395	(1.924 , 13.120)	3.069	1.420	(0.687 , 5.474)
σ_{11}	0.666	0.075	(0.551 , 0.796)	0.631	0.070	(0.524 , 0.752)
σ_{12}	-0.654	0.202	(-0.997 , -0.336)	-0.236	0.073	(-0.360 , -0.120)
σ_{22}	6.993	1.481	(4.811 , 9.628)	1.009	0.193	(0.721 , 1.350)
DIC		2307			2973	

Est.: estimate of parameter, S.D.: posterior standard deviation, PCI: posterior credible interval.

situation, the continuous latent variable mixed-effect model was used for modeling the ordinal responses, such that the latent variable has an asymmetric Laplace distribution given random effects. Consequently, the QR model is performed to ordinal responses using the monotone equivariance feature of quantiles—the connection between these two models being established with some correlated random effects. The non-parametric part of the model is approximated by regression spline. As shown, the suggested model can be fitted under a Bayesian structure via MCMC methods. The simulation was carried out to study the performance of the proposed method. According to this, the performance of the suggested model was good. We earned consistent and desirable results and analysed the PIAT data. The results specified that SQJRM has a better performance than that of the CMRJM. Furthermore, considering the proposed model to analyse data with missing responses is an ongoing research.

REFERENCES

1. A. Beck, and M. Teboulle, Prior elicitation for mixed quantile regression with an allometric model, Environmetrics, vol. 22, no. 7, pp. 911–920, 2011.
2. Bahrami Samani, E., Ganjali, M., and Khodaddadi, A latent variable model for mixed continuous and ordinal responses, Journal of Statistical Theory and Applications, vol. 7, no. 3, pp. 337–349, 2008.
3. Bahrami Samani, E., and Ganjali, M., A multivariate latent variable model for mixed continuous and ordinal responses, World Applied Sciences Journal, vol. 3, no. 2, pp. 294–294, 2008.
4. Bahrami Samani, E., and Ganjali, M., Analyzing mixed correlated continuous and ordinal responses; A latent variable approach, Journal of Applied Statistical Science, vol. 17, no. 4, pp. 1–12, 2009.
5. Bahrami Samani, E., Ganjali, M., and Eftekhari, S., A latent variable model for mixed continuous and ordinal responses with non-ignorable missing responses: Assessing the local influence via covariance structure, Sankhya B, vol. 72, no. 1, pp. 38–57, 2010.
6. Bahrami Samani, E., Ganjali, M., and Amirian, Y., Likelihood estimation for longitudinal zero-inated power series regression models, Journal of Applied Statistics, vol. 39, no. 9, pp. 1965–1974, 2010.
7. Cai, Z. and Xiao, Z., Semiparametric quantile regression estimation in dynamic models with partially varying coefficients, Journal of Econometrics, vol. 167, no. 2, pp. 413–425, 2015.
8. Chen, X., Sun, J. and Liu, L., Semiparametric partial linear quantile regression of longitudinal data with time varying coefficients and informative observation times, Statistica Sinica, vol. 25, no. 4, pp. 1437–1458, 2015.
9. de Boor, C., A Practical Guide to Splines, Springer-Verlag, New York, 1978.
10. Eubank, R.L., NonparametricRegression and Spline Smoothing, Marcel Dekker. New York, 1999.
11. Fan, J., Design-adaptive nonparametric regression, Journal of American Statistical Association, vol. 87, pp. 998–1004, 1992.
12. Fan, J., Local linear regression smoothers and their minimax efficiency, Annals of Statistics, vol. 21, pp. 196–216, 1993.
13. Fan, J. and Gijbels, I, Variable bandwidth and local linear regression smoothers, Annals of Statistics, vol. 20, pp. 2008–2036, 1992.
14. Fitzmaurice, G., Davidian, M., Verbeke, G., and Molenberghs, G., Longitudinal data analysis, Boston, CRC Press, 2008.
15. Fu, L. and Wang, Y.G., Quantile regression for longitudinal data with a working correlation model, Comput.Stat.Data Anal, vol. 56, pp. 2526–2538, 2012.
16. Ganjali, M., A model for mixed continuous and discrete responses with possibility of missing responses, Journal of Sciences Islamic Republic of Iran, vol. 14, no. 1, pp. 53–60, 2003.
17. Gelman, and Rubin, D. B., Inference from iterative simulation using sequences, Statistical Science, vol. 7, pp. 457–511, 1992.
18. Geraci, M., and Bottai, M., Quantile regression for longitudinal data using the asymmetric Laplace distribution, Biostatistics, vol. 8, no. 1, pp. 140–154, 2007.
19. Ghasemzadeh, S., Ganjali, M., and Baghfalaki, T., Bayesian quantile regression for joint modeling of longitudinal mixed ordinal and continuous data, Communications in Statistics-Simulation and Computation, vol. 49, no. 2, pp. 375–395, 2020.
20. Greene W. H., and Hensher D. A., Modeling ordered choices: A primer, Cambridge University Press, 2010.
21. Hao, L. and Naiman, D. Q., Quantile Regression, New York, Sage, 2007.
22. Heckman, J. J. D., Dummy Endogenous variable in a simultaneous Equations system, Econometrica, vol. 46, no. 6, pp. 931–959, 1978.
23. Hong, H.G., He, X., Prediction of functional status for the elderly based on a new ordinal regression model, Journal of American Statistical Association, vol. 105, no. 491, pp. 930–941, 2010.
24. Hong, H.G., Zhou, J., A multi-index model for quantile regression with ordinal data, Journal of Applied Statistics, vol. 40, no. 6, pp. 1231–1245, 2013.
25. Huang, Y., Quantile regression-based Bayesian semiparametric mixed-effects models for longitudinal data with non-normal, missing and mismeasured covariate, Journal of Statistical Computation and Simulation, vol. 86, no. 6, pp. 1183–1202, 2016.

26. Jeliazkov, I., Graves, J., and Kutzbach, M., Fitting and comparison of models for multivariate ordinal outcomes, Advances in Econometrics: Bayesian Econometrics, vol. 23, pp. 115–156, 2008.
27. Karlsson, A., Nonlinear quantile regression estimation of longitudinal data, Communications in Statistics-Simulation and Computation, vol. 37, no. 1, pp. 114–131, 2007.
28. Koenker, R. and Bassett, G., Regression quantiles, Econometrica, vol. 46, pp. 33–50, 1978.
29. Koenker, R., Quantile regression for longitudinal data, J. Multivar. Anal, vol. 91, pp. 74–89, 2004.
30. Koenker, R., Quantile regression, Cambridge, Cambridge university press, 2005.
31. Kottas, A. and Krnjajic, M., Bayesian Semiparametric Modelling in Quantile Regression, Scandinavian Journal of Statistics, vol. 36, pp. 297–319, 2009.
32. Kotz, S., Kozubowski, T.J., Podgorski, K., The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance, Boston:Birkhauser, 2001.
33. Kozumi, H., and Kobayashi, G., Gibbs sampling methods for Bayesian quantile regression, Journal of Statistical Computation and Simulation, vol. 81, no. 11, pp. 1565–1578, 2011.
34. Leon, A. R., and Chough, K. C., Analysis of mixed data: Methods and applications, London: Chapman and Hall/CRC, 2013.
35. McCullagh, P., Regression models for ordinal data (with discussion), Journal of the Royal Statistical Society, Series B, vol. 42, no. 2, pp. 109–142, 1980.
36. McKelvey, R., and Zavoina, W., An IBM Fortran IV Program to perform N-Chotomus multivariate probit aAnalysis, Behavioral Science, vol. 16, no. 2, pp. 186–187, 1971.
37. McKelvey, R., and Zavoina, W., A statistical model for the analysis of ordered level dependent variables, Journal of Mathematical Sociology, vol. 4, pp. 103–120, 1975.
38. Molenberghs, G. and Verbeke, G., Models for discrete longitudinal data, New York, Springer, 2005.
39. Noorian, S., Ganjali, M., and Bahrami Samani, E., A Bayesian test of homogeneity of association parameter using transition modeling of longitudinal mixed responses, Journal of Applied Statistics, vol. 43, no. 10, pp. 1850–1863, 2016.
40. Olkin, L., and Tate, R. F., Multivariate correlation models with mixed discrete and continuous variables, The Annals of Mathematical Statistics, vol. 32, pp. 448–465, 1961.
41. R Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. Available at http://www.R-project.org/, 2013.
42. Samani, E. B., and Ganjali, M., Bayesian multivariate latent variable model for mixed correlated ordinal and continuous responses, Journal of Statistical Research, vol. 43, no. 2, 2009.
43. Samani, E. B., and Ganjali, M., Sensitivity analysis for non-ignorable missing responses with application to multivariate Random effect model, Metron, vol. 69, no. 3, pp. 309–322, 2011.
44. Samuel, M., Ryan, L., and Legler JM., Latent variable models for mixed discreteand continuous outcomes, Journal of the Royal Statistical Society,Series B: Methodological, vol. 59, pp. 667–678, 1997.
45. Sharifian, N., Samani, E. B., and Ganjali, M., Joint modeling for longitudinal set-inflated continuous and count responses, Communications in Statistics - Theory and Methods, vol. 50, no. 5, pp. 1134–1160, 2021. doi:10.1080/03610926.2019.1646768.
46. Sharifian, N., Samani, E. B., and Ganjali, M., Joint model for longitudinal mixture of normal and zero-inflated power series correlated responses, Journal of Biopharmaceutical Statistics, vol. 31, no. 2, pp. 117–140, 2021. doi:10.1080/10543406.2020.1814798.
47. Smith, P.L., Splines as a usefid and convenient statistical tool, American Statistician, vol. 33, pp. 57–62, 2015.
48. Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A., Bayesian measures of model complexity and fit, Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 64, no. 4, pp. 583–639, 2002.
49. Teimourian, M., Baghfalaki, T., Ganjali, M., and Berridge, D., Joint modeling of mixed skewed continuous and ordinal longitudinal responses: a Bayesian approach, Journal of Applied Statistics, vol. 42, no. 10, pp. 2233–2256, 2015.
50. Wahba, G., Spline Models for Observational Data, SIAM, Philadelphia. CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 59, 1990.
51. Wold, S., Spline functions in data analysis, Technometrics, vol. 16, pp. 1–1, 1974.
52. Wu, H., and Zhang, J. T., Nonparametric regression methods for longitudinal data analysis: mixed-effects modeling approaches, John Wiley and Sons, 2006.
53. Xue, L., and A. Qu., Variable selection in high-dimensional varying coefficient models with global optimality, Journal of Machine Learning Research, vol. 13, pp. 1973–1998, 2012.
54. Yu, K., Moyeed, R.A., Bayesian quantile regression, Stat. Probab.Lett, vol. 54, pp. 437–447, 2001.
55. Yu, K., and Zhang, J., A three-parameter asymmetric Laplace distribution and its extension, Communications in Statistics-Theory and Methods, vol. 34, no. 9-10, pp. 1867–1879, 2015.
56. Zhou, L. Huang and M. Li, R., Semiparametric quantile regression with high-dimensional covariates, Statistica Sinica, vol. 22, pp. 1379–1401, 2012.
57. Zhou, L., Conditional quantile estimation with ordinal data, Ph.D. thesis, University of South Carolina, 2010.