Int. J. Nonlinear Anal. Appl. In Press, 1–11
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2025.9624

The Bayesian Lasso of quantile structural equation model

Balsam Mustafa Shafeeq[a], Lekaa Ali Muhamed[b,*]

[a] Technical College of Management, Middle Technical University, Baghdad, Iraq

[b] College of Administration and Economics, Department of Statistics, University of Baghdad, Baghdad, Iraq

(Communicated by Zakieh Avazzadeh)

Abstract

Structural equation models have been extensively applied to medical, and social sciences, the most important latent variable models are structural equation models. Structural equation modeling (SEM) is a popular multivariate technique for analyzing the interrelationships between latent variables. In general, structural equation models includes of a measurement equation to characterize latent variables through multiple observable variables and a mean regression type structural equation to investigate how the explanatory latent variables affect the outcomes of interest. In this study, we apply Bayesian least absolute shrinkage and selection operator (Lasso) procedure to conduct estimation in the Quantile SEM, and compare this estimator with estimator of Bayesian Quantile Structural equation model, and apply the use of the Markov chain Monte Carlo (MCMC) method by Gibbs sampler to conduct Bayesian inference. The simulation was implemented assuming-different distributions of the error term for the structural equations model and values of the parameters for small sample size.

Keywords: Bayesian inference, Lasso regression, quantile structural equation model
2020 MSC: 62F15

1 Introduction

1.1 Structural equation modeling (SEM)

Structural equation modelling (SEM) is a versatile class of models that allow for complicated modelling of correlated multivariate data to examine interrelationships between observable and latent variables. Many extensively used statistical models, such as regression, factor analysis, canonical correlations, and analysis of variance and covariance, are included in this class of models, which is well recognized in social and psychological sciences [10].

Most applications of SEMs are related to the study of interrelationships among latent variables. In particular, they are useful for examining the effects of explanatory latent variables on outcome latent variables of interest. In such situations, researchers usually consider what observed variables should be selected from the whole data set for the analysis and how these observed variables are grouped to form latent variables.

The structural equation model consists of two components, as follows:

*Corresponding author
Email addresses: balsammustafa95@mtu.edu.iq (Balsam Mustafa Shafeeq), lekaa.ali.1968@gmail.com (Lekaa Ali Muhamed)

Received: November 2021 *Accepted:* January 2022

1. Let $y_i = (y_{i1}, ..., y_{ip})^T$ be a $p \times 1$ vector representing the ith observation in a random sample of size n, and $\omega_i = (\omega_{i1}, ..., \omega_{iq})^T$ be a $q \times 1$ vector of latent variables with $(q < p)$. The link between y_i and ω_i is defined by the following measurement equation:

$$y_i = Ac_i + \Lambda\omega_i + \varepsilon_i, \qquad i = 1, ..., n \tag{1.1}$$

where $A(p \times r_1)$ and $\Lambda(p \times q)$ are matrices of unknown coefficients, $c_i(r_1 \times 1)$ is a vector of fixed covariates, and $\varepsilon_i(p \times 1)$ is a random vector of error terms.

2. η_i can be assessed in the following structural equation

$$\eta_i = \beta_\tau d_i + \Gamma_\tau \xi_i + \delta_i, \qquad i = 1, ..., n. \tag{1.2}$$

Then the quantile SEM is defined by Equations (1.1) and (1.2) [12]. To analyze the interrelationship among latent variables, let partition $\omega_i = (\eta_i^T, \xi_i^T)^T$, where $\eta_i(q_1 \times 1)$ denote outcome latent variables and $\xi_i(q_2 \times 1)$ is explanatory latent variables. To simplify, we assume that $q_1 = 1$. The primary goal of SEM is to analyze the behaviour of latent variable η_i given the information n contained in a set of explanatory latent variables ξ_i.

The purpose of the measurement equation in an SEM is to relate the latent variables in ω to the observed variables in y. It represents the link between observed and latent variables, through the specified factor loading matrix Λ, the vector of measurement error ε is used to take the residual error into account. The important issue in formulating the measurement equation is to specify the structure of the factor loading matrix Λ, based on the determination of the observed variables in the study. Any element of Λ can be a free parameter or a fixed parameter with a predetermined value.

The positions and the pre-assigned values of the fixed parameters are decided based on the prior knowledge of the observed and latent variables, and they are also related to the interpretations of the latent variables. It can also be known from previous studies [10].

2 Quantile structural equation model (QSEM)

The primary aim of SEM is to analyze the behaviour of the latent variable η_i given the information contained in a set of explanatory latent variables ξ_i. This is done in traditional SEM by calculating the conditional mean of $(\eta_i \backslash \xi_i)$ and fixed covariates $d_i(r_2 \times 1)$ as follows [12]:

$$E(\eta_i \backslash \xi_i, d_i) = Bd_i + \Gamma\xi_i, \qquad i = 1, ..., n \tag{2.1}$$

where $B(q_1 \times r_2)$ and $\Gamma(q_1 \times q_2)$ are the matrices of unknown coefficients to be estimated. The conditional mean does not provide a complete description of the interrelationship among latent variables. A more comprehensive analysis can be achieved from a combination of $Q(\eta_i \backslash \xi_i, d_i)$, the conditional quantile of η_i, under various quantiles $\tau \in (0, 1)$ as follows:

$$Q_\tau(\eta_i \backslash \xi_i, d_i) = B_\tau d_i + \Gamma_\tau \xi_i, \qquad i = 1, ..., n. \tag{2.2}$$

The coefficient matrices B_τ and Γ_τ have a subscript τ because they might not be equal for different quantiles. Unlike in conventional SEMs, here the distribution of δ_i is undefined. The only assumption is that the τ-quantile of δ_i is 0 to guarantee that (2.2) holds. The rest of the paper is organized as follows. In section 2, we present the Quantile Structural equation model (QSEM). In section 3 we present Bayesian inference of QSEM model with display the conditional distributions of parameters and latent variable within the Bayesian analysis, in section 4 we present Regularization technique in Bayesian Quantile SEM (Bayesian lasso) and display the conditional distributions of parameters and latent variable within the Bayesian lasso analysis by using Gibbs sampling. And in section 5, we perform simulation studies to examine the performance of the method used with different error term distributions. We conclude with condensed conclusions in section 6.

3 Bayesian inference for quantile structural equation model

To speed up and increase the performance of the Bayesian method in the analysis of the QSEM model, and for the reasons mentioned previously, this research was based on the proposal of Kozumi and Kobayshi [6] in using the mixed representation of the skewed Laplace distribution (AL) for random error in the model. According to Wang's assumption [12], that will be adopted in this research for the error terms, specifically ϵ_{ij} the kth component of the error terms ϵ_i is distributed $AL(0, \sigma_{yk}, 0.5)$ for measurement equation (1.1) the median regression, and δ_i is distributed $AL(0, \sigma_{yk}, \tau)$ for structural equation (1.2) the τ-quantile regression. Noting that the variables e_{yik} and $e_{\eta i}$ are the nuisance variables for augmenting ϵ_{ij} and δ_i [12].

Let θ_y the unknown parameters in equation (1.1), and θ_ω unknown parameters in equation (1.2), and $\theta = (\theta_y, \theta_\omega)$, then the Bayesian for Quantile SEM by the following hierarchical representation:

$$(y_i/\eta_i, \xi_i, \theta_y, e_{yi}) \overset{ind}{\sim} N_p(Ac_i + \Lambda\omega_i, \Psi_i) \tag{3.1}$$

$$(\eta_i/\xi_i, \theta_\omega, e_{\eta i}) \overset{ind}{\sim} N(B_\tau d_i + \Gamma_\tau \xi_i + \kappa_1 e_{\eta i}, \kappa_2 \sigma_\eta e_{\eta i}). \tag{3.2}$$

$$e_{\eta i} \overset{i.i.d}{\sim} \exp(\sigma_\eta)$$

$$e_{yik} \overset{i.i.d}{\sim} \exp(\sigma_{yk})$$

$$\xi_i \overset{i.i.d}{\sim} N_{q_2}(0, \Phi)$$

where $e_{yik} = (e_{yi1}, ..., e_{yip})^T$, $\Psi_i = diag(8\sigma_{y1}e_{yi1}, ..., 8\sigma_{yp}e_{yip})$, and $e_\eta = (e_{\eta 1}, ..., e_{\eta n})^T$. Let $\Lambda_y = (A, \Lambda) = (\lambda_{ykj})$, and in the structural equation (3.2), the unknown parameters are $\Lambda_{\omega\tau} = (B_\tau, \Gamma_\tau)$. Some elements of θ_y must be fixed for identification purposes, for the measurement equation, an index matrix $M = (I_{ykj})$ as its identification matrix is created as follows [10]: when $I_{ykj} = 1$ if λ_{ykj} is subject to estimation and $I_{ykj} = 0$ if the value of λ_{ykj} for the purpose of identification, is prefixed. The following conjugate prior distribution in Bayesian quantile SEM are:

- For measurement equation as follows:

$$\begin{aligned} \theta_{1yk} &\sim N_{r1+q}(\Lambda_{0yk}, H_{0yk}) \\ \sigma_{yk}^{-1} &\sim \Gamma(a_{0yk}, b_{0yk}) \end{aligned} \tag{3.3}$$

- For structural equation as follows:

$$\begin{aligned} \theta 2_{\omega\tau} &\sim N_{r2+q2}(\Lambda_{0\omega}, H_{0\omega}) \\ \sigma_\eta^{-1} &\sim \Gamma(a_{0\sigma}, b_{0\sigma}) \\ \Phi^{-1} &\sim \text{Wishart}(R_0, \rho_0) \end{aligned} \tag{3.4}$$

  where $(\Lambda_{0yk}, a_{0yk}, b_{0yk}, \Lambda_{0\omega}, a_{0\sigma}, b_{0\sigma})$ are hyperparameters and the positive-definite $H_{0yk}, H_{0\omega}$ are also hyperparameters, Noting that the values are given from previous research or professional knowledge.

Let $Y = (y_1, ..., y_n)$, $C = (c_1, ..., c_n)$, $D = (d_1, ..., d_n)$ and $\Omega = (\omega_1, ..., \omega_n)$ be the matrix of latent variable. Given the complexity of the model, direct inference of the common posterior distribution $p(\Omega, \theta \backslash Y, C, D, e_\eta)$ is difficult and complex. However, the full conditional distributions of the latent variables and all parameters are common. Therefore, the Gibbs sampling method is used as an easy and uncomplicated method in obtaining Bayesian estimators, so that the Gibbs sampling tool can be implemented easily, and a Bayesian estimate is taken for each parameter to be the average of the sample random observations derived from each iteration.

As is well known the Bayesian estimate of parameters are obtained from the joint posterior distribution $p\ (\Omega, \theta \backslash Y, C, D, e_\eta)$ by drawing samples iteratively for parameters and latent variables, each component of the posterior distribution is generated by the Gibbs sampling method from its full conditional posterior distribution in an iteratively. The Bayesian estimates of θ and Ω are taken to be the sample mean of the random observations generated.

As mentioned earlier, the main objective is to use MCMC methods to obtain the Bayesian estimates of θ and Ω, for this reason, a sequence of random observations from the joint posterior distribution $[\theta, \Omega/Y]$ will be generated via the Gibbs sampler which is implemented as follows. At the jth iteration with current value $\theta(j)$ [6]:

a. Generate a random variate $\Omega(j+1)$ from the condition $[\Omega/Y, \theta(j)]$

b. Generate a random variate $\theta(j+1)$ from the condition $[\theta/Y, \Omega(j+1)]$ and return to step a if necessary

Then the full conditional posterior distribution for Bayesian quantile SEM (BQSEM) as follows:

The Gibbs sampling algorithm is implemented with the following full conditional posterior distribution of parameters and latent variable [12].

Let $\theta_y = (A, \Lambda), \theta_\omega = (B_\tau, \Gamma_\tau), u_i = (c_i^T, \omega_i^T)^T, v_i = (d_i^T, \xi_i^T)^T, U = (u_1, ..., u_i)$, where U_k be its submatrix with rows corresponding to $I_{ykj} = 0$ are deleted, $Y_k^* = (y_{1k}^*, ..., y_{nk}^*)$ where

$$Y_k^* = Y_k - \sum_{j=1}^{r_1+q} \lambda_{ykj} u_{ij}(1 - I_{ykj}).$$

1. The full conditional posterior distribution of the latent variable Ω. The y distribution ia as follows:

$$(y_i/\theta 1_y, \eta_i, \xi_i, e_{yi}) \overset{ind}{\sim} N_p(Ac_i + \Lambda\omega_i, \Psi_i)$$
$$p(Y/\theta_y, \eta_i, \xi_i, e_{yi})\eta_i = (\Psi_i)^{\frac{-n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(y_i - \theta 1_y u_i)^T \Psi_e^{-1}(y_i - \theta_y u_i)\right\}. \tag{3.5}$$

It is known that

$$p(\omega_i/(y_i, \theta_y) \propto p(\omega_i/\theta_y)p(y_i/(\omega_i, \theta_y))).$$

Then, the full conditional posterior distribution of the latent variable is

$$(\omega_i \backslash y_i \sigma_{yi} e_{yi} \theta_y \sigma_\eta e_{\eta i} \Lambda_\omega \Phi) \sim N_q(\mu_i, \sum_i^{-1*}) \tag{3.6}$$

where

$$\mu_i = \sum_i^{*-1} \Lambda^T \psi_i^{-1}(y_i - Ac_i) + \sum_i^{*-1}\sum_{\omega i}^{-1} \begin{pmatrix} B_\tau d_i + k_1 e_{\eta i} \\ 0 \end{pmatrix}$$
$$\sum_i^{*} = \sum_{\omega i}^{-1} + \Lambda^T \psi_i^{-1} \Lambda$$
$$\sum_{\omega i} = \begin{pmatrix} \Gamma_\tau \Phi \Gamma_\tau^T + k_2 \sigma_\eta e_{\eta i} & \Gamma_\tau \Phi \\ \Phi \Gamma^T & \Phi \end{pmatrix}$$
$$\psi_i = diag(8\sigma_{y1} e_{yi1}, ..., 8\sigma_{yp} e_{yio})$$

2. The full conditional posterior distribution of the e_{yik} : for $(i = 1, ..., n,\ k = 1, ..., p)$

$$p(e_{yik}^{-1} \backslash y_{ik}, \omega_i, \theta 1_{yk}, \sigma_{yk}) \propto f(y_{ik}, \omega_i, \theta 1_{yk}, \sigma_{yk}) f(e_{yik} \backslash \sigma_{yk})$$
$$p(e_{yik}^{-1} \backslash y_{ik}, \omega_i, \theta 1_{yk}, \sigma_{yk}) \propto \left\{\frac{2\sigma_{yk}^{-1}}{2\pi (e_{yik}^{-1})^3}\right\}^{\frac{1}{2}} \exp\left\{\frac{2\sigma_{yk}^{-1}\left(e_{yik}^{-1} - \frac{4}{|y_{ik} - \theta 1_{yk} u_i|}\right)^2}{2\left[4|y_{ik} - \theta 1_{yk} u_i|^{-1}\right]^2 e_{yik}^{-1}}\right\} \tag{3.7}$$

Thus, the full conditional distribution of e_{yih} is an inverse Gaussian distribution with parameter $(4|y_{ik} - \theta 1_{yk} u_i|^{-1}, 2\sigma_{yk}^{-1})$.

3. The full conditional posterior distribution of the θ_y, for $(k = 1, ..., p)$:

$$p(\theta 1_{yk} \backslash Y, e_{yik}, \sigma_{yk}) \propto \left(\sum_{\theta 1k}^{-1}\right)^{\frac{-1}{2}} \exp\left(-\frac{1}{2}(\theta 1_{yk} - M_{\Lambda k})^T \left(\sum_{\Lambda k}^{-1}\right)^{-1} (\theta 1_{yk} - Mu_{\Lambda k})\right) \tag{3.8}$$

where $Mu_{\Lambda k} = \sum_{\Lambda k}^{-1}\left(H_{0y}^{-1}\Lambda_{oy} + \sum_{i=1}^{n} \frac{y_{ik} u_i}{8\sigma_{yk} e_{yik}}\right)$, and $\sum_{\Lambda k} = H_{0y}^{-1} + \sum_{i=1}^{n} \frac{u_i u_i^T}{8\sigma_{yk} e_{yik}}$. Thus, the full conditional posterior distribution of the θ_y in equation (3.8) is a normal distribution.

4. The full conditional posterior distribution of the σ_{yk}, for $k = 1, ..., p$,

$$p(\sigma_{yk}^{-1} \backslash Y, U, \Lambda_{yk}) \propto (\sigma_{yk}^{-1})^{n + a_{0yk} - 1} \exp\left\{\left(b_{0yk} + \frac{1}{2}\sum_{i=1}^{n}|y_{ik} - \theta 1_{yk} u_i|\right)\sigma_{yk}^{-1}\right\} \tag{3.9}$$

Thus, the full conditional posterior distribution of the σ_{yk} is Gamma distribution $(n + a_{0yk}, b_{0yk} + \frac{1}{2} - \sum_{i=1}^{n}$ $|y_{ik} - \theta 1_{yk} u_i|)$.

5. The full conditional posterior distribution of the Φ:

$$p(\Phi\backslash\Omega_2) \propto p(\Phi)\prod_{i=1}^{n} p(\xi_i\backslash\Phi)$$

$$p(\Phi\backslash\Omega_2) \propto |\Phi|^{-(n+\rho_0+q_2+1)/2}\exp\left\{-\frac{1}{2}tr\left[\Phi^{-1}(\Omega_2\Omega_2^T+R_0^{-1})\right]\right\} \tag{3.10}$$

Since the right-hand side of (3.10) is proportional to the density function of an inverted Wishart distribution [14], it follows that the conditional posterior distribution of $(\Phi\backslash\Omega_2)$ is given by

$$[\Phi\backslash\Omega_2] \sim IW_{q_2}[(\Omega_2\Omega_2^T+R_0^{-1}), n+\rho_0]$$

6. The full conditional posterior distribution of the $e_{\eta i}$: for $(i = 1, ..., n)$

$$p(e_{\eta i}^{-1}\backslash\omega_i,\theta_\omega,\sigma_\eta) \propto f(\omega_i,\theta_\omega,e_{\eta i}^{-1},\sigma_\eta)f(e_{\eta i}^{-1}\backslash\sigma_{yk})$$

$$(e_{\eta i}^{-1}\backslash\omega_i,\theta_\omega,\sigma_\eta) \propto \left\{\frac{\frac{k_2}{4\sigma_\eta}}{2\pi(e_{\eta i}^{-1})^3}\right\}^{\frac{1}{2}}\exp\left\{\frac{\frac{k_2}{4\sigma_\eta}\left(e_{\eta i}^{-1}-\frac{k_2}{2|\eta_i-B_\tau d_i-\Gamma_\tau\xi_i|}\right)^2}{2\left[\frac{k_2}{2|\eta_i-B_\tau d_i-\Gamma_\tau\xi_i|}\right]^2 e_{\eta i}^{-1}}\right\}. \tag{3.11}$$

Thus, the full conditional posterior distribution of the $e_{\eta i}$ is the Inverse Gaussian distribution $\left(\frac{k_2}{2|\eta_i-B_\tau d_i-\Gamma_\tau\xi_i|},\frac{k_2}{4\sigma_\eta}\right)$.

7. The full conditional posterior distribution of the $\theta_{\omega\tau}$

$$p(\theta_{\omega\tau}\backslash\Omega,e_\eta,\sigma_\eta) \propto (\sum_{\theta_\omega}^{-1})^{\frac{-1}{2}}\exp\left(\frac{-1}{2}(\theta_{\omega\tau}-Mu_{\theta\omega})^T(\sum_{\theta_\omega}^{-1})^{-1}(\theta_{\omega\tau}-Mu_{\theta\omega})\right) \tag{3.12}$$

where $Mu_{\theta\omega} = \sum_{\theta 2\omega}^{-1}\left(H_{o\omega}^{-1}\theta_{\omega\tau}+\sum_{i=1}^{n}\frac{(\eta_i-k1e_{\eta i})v_i}{k2\sigma_\eta e_{\eta i}}\right)$ and $\sum_{\theta 2\omega} = H_{o\omega}^{-1}+\sum_{i=1}^{n}\frac{v_iv_i^T}{k2\sigma_\eta e_{\eta i}}$. Thus, the full conditional posterior distribution of the $\theta 2_\omega$ is a normal distribution.

8. The full conditional posterior distribution of the σ_η :

$$p(\sigma_\eta^{-1}\backslash\Omega,\theta 2_{\omega\tau}) \propto (\sigma_\eta^{-1})^{n+a_{0\delta}-1}\exp\left(b_{0\delta}+\sum_{i=1}^{n}\rho_\tau|\eta_i-\theta_{\omega\tau}v_i|\right)\sigma_\eta^{-1} \tag{3.13}$$

Thus, the full conditional posterior distribution of the σ_η is Gamma distribution $(n+a_{0\delta}, b_{0\delta}+\sum_{i=1}^{n}\rho_\tau|\eta_i-\theta_{\omega\tau}v_i|)$.

4 Regularization technique in Bayesian quantile SEM (Bayesian Lasso)

Tibshirani [11] proposed a penalty function for the linear regression model known as Lasso, which is abbreviated for (Least Absolute Shrinkage and Selection Operator) [9]. It is one of the important techniques that were used in estimating the parameters of regression models. This technique is of great importance in controlling the variance of the model parameters and selecting the important variables in the model. It can reach explanatory models, and it is also of great importance in reducing the prediction error [11]. It was proposed to estimate the parameters of the linear regression model and to perform the variable selection simultaneously [1]. The principle of the Lasso method is to reduce the sum of squares of the residuals according to a constraint representing the absolute sum of the coefficients, which are less than a certain constant. For the linear regression model. The Lasso estimator is the solution to the following L_1-penalized least squares problem [8]:

$$\min_\beta\sum_{i=1}^{n}(y_i-x_i^T\beta)^2+\gamma\sum_{j=1}^{p}|\beta_j| \tag{4.1}$$

where $\sum_{j=1}^{p}|\beta_j|$ is penalty function or it is sometimes called Regularization function, $\hat{\beta}_{Lasso} = (\hat{\beta}_1,\hat{\beta}_2,...,\hat{\beta}_p)$. γ is a tuning parameter $(\gamma \geq 0)$ that controls the penalty amount, such that the Lasso estimator is equal to the least squares estimator when $\gamma = 0$ and shrinks towards zero as γ increases.

The Bayesian inference in Lasso technique has gained great interest in recent years in estimating the regression model because of its great importance in achieving the accurate inference of this model, Park and Casella [9], proposed a Bayesian framework of the Lasso (BaLasso), they assumed they considered the error term of the model is follow the normal distribution $(0, \sigma^2)$, they proposed the Bayesian Lasso estimator of β is defined as the posterior mode of β by assuming that conditionally independent double-exponential prior distribution by the following [5]:

$$\pi(\beta/\sigma^2) \prod_{j=1}^{p} \frac{\gamma}{2\sigma} e^{-\frac{\gamma|\beta_j|}{\sigma}}. \tag{4.2}$$

So that produces the same effect in contraction as in the original equation of Lasso, as in equation (4.2). As it is known that in achieving the Bayesian analysis with this technique, the Laplace distribution is assumed independently as a prior distribution of the model parameters. To facilitate Gibbs sampling in Bayesian inference, in most research, the mixed representation of the Laplace function assumed by Andrews and Mallows [2] is used, so that the probability density function of the Laplace distribution is written with a mixed representation of the two distributions (Normal and Exponantial), as follows [3]:

$$\frac{\gamma}{2\sigma} e^{-\gamma|\beta_j|/\sigma} = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2 s_j}} e^{-\beta_j^2/(2\sigma^2 s_j)} \frac{\gamma^2}{2} e^{-\gamma^2 s_j/2} ds_j \tag{4.3}$$

According to the hierarchical formula, β has a normal distribution, as follows:

$$[\beta/\sigma^2, s_j] \sim N_p(0, \sigma^2.s_j)$$

where $s_j \sim$ exponential$(2/\gamma^2)$, s_j is diagonal matrix $(s_1, ..., s_p)$. The tuning parameter $\gamma^2 \sim \Gamma(a_\gamma, b_\gamma)$, when a_γ, b_γ are predefined hyperparameters, where it was specified by Feng et al. [3], we set $a_\gamma = 1$ and $b_\gamma = 0.05$ for obtaining dispersed priors. Based on the previously described hierarchical structure, Blasso and Balasoo may be easily used in more complex models, such as quantile regression models or quantile SEM, to conduct simultaneous estimation and variable selection. Quantile regression was pointed out by Koenker and Bassett Jr [4], where the frequentist approach to the estimation of coefficients is to solve the following optimization problem:

$$\min_\beta \sum_{i=1}^{n} \rho_\tau(y_i - x_i^T\beta) \tag{4.4}$$

where $\rho_\tau(x) = x(\tau - I(x < 0))$ is the quantile loss function. Li and Zhu proposed the regularized quantile regression to achieve estimation and variable selection, which uses the Lasso type penalty function, as follow [8]:

$$\min_\beta \sum_{i=1}^{n} \rho_\tau(y_i - x_i^T\beta) + \gamma \sum_{j=1}^{p} |\beta_j|. \tag{4.5}$$

In a Bayesian quantile regression framework, we need to specify a working likelihood for the model error. According to Yu and Moyeed [13], maximizing the likelihood under ALD error is equivalent to minimizing the objective loss function (4.4) of quantile regression skewed Asymmetric Laplace (ALD) has its probability density function as follows:

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

where μ is the location parameter, σ is the scale parameter and $(0 < \tau < 1)$ is the skewness parameter. According to Yu and Moyeed (2001) implementing Bayesian inference for quantile regression [13], if the error term ε_i are follow $AL(0, \sigma, \tau)$, then the conditional likelihood function for the quantile regression model as follows [3]:

$$L(\beta, \sigma; y, X) = \frac{\tau^n(1-\tau)^n}{\sigma^n} \exp\left\{-\frac{\sum_{i=1}^{n} \rho_\tau(y_i - x_i^T\beta)}{\sigma}\right\}. \tag{4.6}$$

Hence, the minimization problem given by (4.2) is equivalent to maximizing the likelihood function (4.6), for the conditional likelihood function (4.6), we suffer computation difficulty due to the inherent non-differentiability of the QR check function. Nevertheless, Kozumi and Kobayashi [6] proved that the skewed Laplace distribution (4.6) can be viewed as a mixture of normal and exponential distributions as follows [7]:

$$y = \mu + k_1 e + \sqrt{k_2 \sigma e}\varsigma \tag{4.7}$$

where $k_1 = (1-2\tau)/(\tau(1-\tau))$, $k_2 = 2/\tau(1-\tau)$, $\varsigma \sim N[0,1]$, $e \sim \exp(1/\sigma)$. The resulting conditional distribution of y is normal, with a mean $(\mu + k_1 e)$ and variance $(k_2 \sigma e)$. The posterior distribution of β can be expressed as follows:

$$f(\beta/y, X) \propto \pi(\beta) \exp\left\{-\frac{\sum_{i=1}^{n} \rho_\tau(y_i - x_i^T \beta)}{\sigma}\right\} \tag{4.8}$$

where $\pi(\beta)$ is a prior distribution, The prior distribution of β is not unique, but there have been many attempts by researchers, initially Yu and Moyeed [13] employed non-informative prior ($\pi(\beta) \propto 1$) which yielded a proper joint posterior distribution, and the posterior mode of β is also identical to the solution to quantile regression in (4.4), and based on the aforementioned normal mixture representation of (ε_i) Kozumi and Kobayashi [6] specified a conjugate normal prior for β, and the posterior of a normal distribution.

Feng et al. [3] have adopted Li et al. [7] proposing the Bayesian regularized quantile regression by employing the double-exponential prior in equation (4.2), such that the maximization of the posterior of β is equivalent to the minimization of equation (4.5) in Lasso technique, to implement the Gibbs sampling we need to generate the unknowns from the fully conditional posterior distributions. The fully conditional posterior distributions are provided below.

Thus, by using this prior distribution, an easy posterior distribution analysis is obtained, as well as an easy possibility to apply the Gibbs sampling method. Then the Bayesian hierarchical model based on the hierarchical model presented by Feng et al. [3] was used in estimating the parameters of the structural equation as well as the measurement equation within the structural equations model using the Lasso technique, which was explained in this section. The common conjugated prior distributions were used in the Bayesian analysis of the structural equations model, as follows [3, 7]. To simplify the expression of the distributions, we define several notations. For the measurement equation (1.1), we let $\Omega = (\omega_1, ..., \omega_n)$, $\Lambda y = (A, \Lambda) = \{\lambda_{ykj}\}$, and define $L_y = \{l_{ykj}\}$ as its identification matrix. That is, $l_{ykj} = 0$ if the value of λ_{ykj} is prefixed for identification purposes, and $l_{ykj} = 1$ if λ_{ykj} is subject to estimation.

We let $u_i = (c_i^T, \omega_i^T)^T, U = (u_1, ..., u_n)$, and define U_k as the submatrix of U after removing the rows corresponding to $l_{ykj} = 0$. We let $Y * k = (y * k_1, ..., y * k_n)^T$ with

$$y_{ik}^* = y_{ik} - \sum_{j=1}^{r_2+q_2} \lambda_{ykj} u_{ij} (1 - l_{ykj}).$$

For the median regression in measurement equation (1.1), we can be expressed as follows:

$$(y_i/\omega_i, \theta_y, e_{yi}) \overset{ind}{\sim} N_p(Ac_i + \Lambda\omega_i, \Psi_i).$$

To simplify the notations, let $u_i = (c_i^T, \omega_i^T)^T, \theta_y = (A, \Xi)$, $\theta_{yk}T$ be the kth row of $\theta 1_y$ for $k = 1, ..., p$. Then the distribution of $(\gamma_i/\omega_i, \theta_\gamma, e_{\gamma i})$ is in the following form

$$(y_i/\omega_i, \theta_y, e_{yi}) \overset{ind}{\sim} N_p(\theta_i u_i, \Psi_i)$$

$$\begin{aligned}
\theta_{yk} &\sim N(\Lambda_{0yk}, H_{0yk}) \\
e_{yik} &\sim \exp(\sigma_{yk}) \\
\sigma_{yk} - 1 &\sim \Gamma(a_{0\gamma k}, b_{0\gamma k})
\end{aligned}$$

where $a_{0\gamma}, b_{0\gamma}, \Lambda_{0yk}$ and H_{0yk} (positive-definite matrix) are the hyperparameters and $e_{yi} = (e_{yi1}, ..., e_{yip})^T, \Psi_i = diag(8\sigma_{y1}e_{yi1}, ..., 8\sigma_{yp}e_{yip})$ and the structural equation (1.2) with Bayesian Lasso as follow: $\beta_\tau = (\beta_{1\tau}^T, \beta_{2\tau}^T)^T$, $v_i = (d_i^T, \xi_i^T)^T$, $(\eta_i/\xi_i, \theta_{2\omega}, e_{\eta i}) \overset{ind}{\sim} N(\beta_\tau^T v_i, k_1 e_{\eta i}, k_2 \sigma_\eta e_{\eta i})$, $\xi_i \overset{ind}{\sim} N_{q_2}(0, \Phi)$, $\Phi^{-1} \sim \text{Wishart}(R_0, \rho_0)$, $\beta_\tau \sim N_{\gamma_2+q_2}(0, S)$, where $S = diag(s_1, ..., s_{\gamma_2+q_2})$

$$\begin{aligned}
s_j &\sim \exp\left(\frac{2\sigma_\eta}{\gamma^2}\right) \\
\gamma^2 &\sim \Gamma(a_\gamma, b_\gamma) \\
\sigma_\eta^{-1} &\sim \Gamma(\alpha_{0\sigma}, \beta_{0\sigma}) \\
e_{\eta i} &\sim \exp(\sigma_\eta)
\end{aligned}$$

where $a_{0\gamma}, b_{0\gamma}, \Lambda_{0yk}$ and H_{0yk} (positive-definite matrix) are the hyperparameters and $e_\eta = (e_{\eta 1}, ..., e_{\eta n})^T$ Due to the complexity of the model, direct inference of the common posterior distribution $p(\Omega, \theta \backslash Y, C, D, e_\eta)$ is difficult and complex. However, the full conditional distributions of the latent variables and all parameters are common distributions. Therefore, the Gibbs sampling method is used as an easy and uncomplicated way to obtain Bayesian estimates for the parameters and the latent variable, so that the Gibbs sampling tool can be easily implemented, and the Bayesian estimate for each parameter is taken to be the mean of the sample of random observations derived from each iteration.

As is known, a Bayesian estimate for parameters is obtained from the posterior joint distribution $p(\Omega, \theta \backslash Y, C, D, e_\eta)$ by an iterative sampling of the parameters and latent variables, each component of the posterior distribution is generated by the Gibbs sampling method From the conditional complete post hoc distribution iteratively [3]. Bayes estimates for and were taken to be the sample mean for the random observations generated as mentioned in Section 3.

1. The full conditional posterior distribution of the Φ:

$$p(\Phi \Omega_2) \propto p(\Phi) \prod_{i=1}^{n} p(\xi_i \backslash \Phi), \quad p(\Phi \Omega_2) \propto |\Phi|^{-(n+\rho_0+q_2+1)/2} \exp\left\{-\frac{1}{2} tr\left[\Phi^{-1}(\Omega_2 \Omega_2^T + R_0^{-1})\right]\right\}. \tag{4.9}$$

2. The full conditional posterior distribution of σ_η:

$$(\sigma_\eta^{-1} \backslash \Omega, \beta_\tau, s_j, \gamma_j) \sim \Gamma(n + a_\sigma + r_2 + q_2, b_\sigma + \sum_{i=1}^{n} \rho_\tau(\eta_i - \beta_\tau^T v_i) + \frac{\gamma^2}{2} \sum_{j=1}^{r_2+q_2} s_j)$$

3. The full conditional posterior distribution of the $e_{\eta i}$ is a

$$\text{Inverse Gaussian distribution} \left(\frac{k_2}{2|\eta_i - B_\tau d_i - \Gamma_\tau \xi_i|}, \frac{k_2}{4\sigma_\eta}\right) \tag{4.10}$$

4. The full conditional posterior distribution of β_τ:

$$f(\beta_\tau \backslash \Omega, e_\eta, \sigma_\eta) \propto f(\eta_i \backslash \Omega, e_\eta, \sigma_\eta) f(\beta_\tau), \quad f(\beta_\tau \backslash \Omega, e_\eta, \sigma_\eta) \propto N_{\gamma_2+q_2}\left(\sum_\beta^{-1} V E_\sigma^{-1} \Xi^*, \sum_\beta^{-1}\right) \tag{4.11}$$

   where $\sum_\beta^{-1} = (S^{-1} + V E_\sigma^{-1} V^T)^{-1}$.

5. The full conditional posterior distribution of s_j:

$$(s_j^{-1} \backslash \beta_{\tau j}, \gamma, \sigma_\eta) \sim Inverse - Gaussian\left(\frac{\gamma}{\sqrt{\sigma}|\beta_{\tau j}|}, \frac{\gamma^2}{\sigma_\eta}\right) \tag{4.12}$$

6. The full conditional posterior distribution of Υ:

$$f(\gamma^2 \backslash s_j, \sigma_\eta) \propto f(s_i \backslash \sigma_\eta) f(\gamma^2), \quad f(\gamma^2 \backslash s_j, \sigma_\eta) \sim \Gamma\left(a_{0\gamma} + \gamma_2 + q_2, b_{0\gamma} + \frac{\sum_{j=1}^{r_2+q_2} s_j}{2\sigma_\eta}\right). \tag{4.13}$$

5 Simulation study

In this section, we employ simulation to evaluate the Bayesian quantile SEM's empirical performance. We generated the data set from SEM:

$$y_i = Ac_i + \Lambda \omega_i + \varepsilon_i, \quad \eta_i = b_1 d_i + \gamma_1 \xi_{i1} + \gamma_2 \xi_{i2} + \delta_i$$

where $p = 9, q = 3, q_1 = 1, q_2 = 2$ and $\gamma_1 = \gamma_2 = 1$.

The simulation study's main purpose is to estimate the quantile regression coefficients b_1, γ_1 and γ_2 under different quantiles with small sample size and compare them to their theoretical values. We are choose three sample size $n = (25, 50, 100)$ and the quantile we choose $\tau = 0.25, 0.5$ and 0.75. The factor loading matrix Λ has the common non-overlapping structure

$$\Lambda^T = \begin{bmatrix} 1^* & \lambda_{21} & \lambda_{31} & 0^* & 0^* & 0^* & 0^* & 0^* & 0^* \\ 0^* & 0^* & 0^* & 1^* & \lambda_{52} & \lambda_{62} & 0^* & 0^* & 0^* \\ 0^* & 0^* & 0^* & 0^* & 0^* & 0^* & 1^* & \lambda_{83} & \lambda_{93} \end{bmatrix}$$

where the zero and ones marked with are fixed in advance to allow for a clear interpretation of latent variables and model identification, while the other λ_{jk} are unknown parameters. The true vales of parameters λ_{jk} and a_j in the measurement equation are taken to be $\lambda_{21} = \lambda_{31} = \lambda_{52} = \lambda_{62} = \lambda_{83} = \lambda_{93} = 0.7$, then the factor loading matrix Λ will be in the following

$$\Lambda^T = \begin{bmatrix} 1^* & 0.7 & 0.7 & 0^* & 0^* & 0^* & 0^* & 0^* & 0^* \\ 0^* & 0^* & 0^* & 1^* & 0.7 & 0.7 & 0^* & 0^* & 0^* \\ 0^* & 0^* & 0^* & 0^* & 0^* & 0^* & 1^* & 0.7 & 0.7 \end{bmatrix}$$

The true values of parameters in the model are $A = (0.5, \cdots, 0.5)^T$, $b_1 = 0.1$ and $\Gamma_\tau = (\gamma_1, \gamma_2) = (0.2, 0.3)$ and the explanatory latent variable $\xi_i = (\xi_{i1}, \xi_{i2})^T$ is assumed to follow a normal distribution $N(0, \varphi)$ where $\Phi = \begin{bmatrix} 1 & 0.2 \\ 0.2 & 1 \end{bmatrix}$ and the fixed covariates c_{1i} and d_i are independently generated from standard normal distribution $N(0, 1)$. Also, the prior distributions and the hyperparameters are as follows:

For the conjugate prior of $\Lambda_{yk} \sim Nr1 + q(\Lambda_{0yk}, H_{0yk})$, the free elements in the prior mean Λ_{0yk} and H_{0yk} is taken as a diagonal matrix with diagonal elements (10000), As well for structural equation the Λ_ω the prior mean $\Lambda_{0\omega} = (1, 0.7, 0.7)$ and the covariance matrix $H_{0\omega} = 100$ an identity matrix. And for the conjugate inverse gamma prior of $\sigma_{yk} = (a_{0yk}, b_{0yk}) = (1, 1)$ and $\sigma_\eta = (a_{0\sigma}, b_{0\sigma}) = (1, 1)$. For the inverse Wishart prior of Φ, we set $\rho_0 = 1$. For the error terms for SEM ε_i and δ_i, we consider the following different distribution for $k = 1, ..., p$:

(i) ε_{ik}'s and δ_i's follow the normal distribution $N(0, 0.4)$.

(ii) ε_{ik}'s and δ_i's are distributed as the heavy- tailed central t-distribution $t(5)$.

(iii) ε_{ik}'s and δ_i's are distributed as the skewed ln $N(0, 0.35)$.

In the case (i), the normal distribution was chosen for the error terms as it aligns with that of traditional SEM., In the case (ii), the heavy-tailed t-distribution is used to assess the quantile SEM's performance in the presence of outliers in both the observed and latent variables. In Case (iii), the quantile SEM with skewed outcome latent variables is evaluated using a log-normal distribution we run 10,000 iterations with the initial 2,000 observations dropped in the burn-in phase on the basis of 100 replications where the program was written in R language. The performance of the Bayesian quantile structured equation model (QSEM) is assessed using the bias and root mean square error (RMS), where the root mean square error (RMS) is:

$$RMS(\hat{\theta}) = \left\{ \frac{1}{n} \sum_{i=1}^{n} (\hat{\theta} - \theta) \right\}^{\frac{1}{2}}$$

The Bayesian Quantile estimators referred to in section 3 and Bayesian Lasso estimators referred to in section 4 were compared for the regression coefficients in the structural equation (1.1) presented in Table 1, Table 2 and Table 3 for different distributions of error terms. Also, the estimates of the measurement equation coefficients were presented in table 4 within the structural equations model with a sample size of 25.

Table 1: Bayesian estimators of the regression coefficients for structural equation for sample size $n = 25$, $\varepsilon_i \sim N(0, 0.4)$

δ_i		n=25		N (0 0.4)	
		BQSEM		**BQLsso**	
Par	τ	**RMS**	**Bias**	**RMS**	**Bias**
$b_{1\tau}$	0.25	0.06807999	-0.0140986	0.063584	-0.0522042
	0.5	0.09094509	0.08902956	0.009783	-0.0093578
	0.75	0.2809844	0.2372889	0.118061	0.08338740
$\gamma_{1\tau}$	0.25	0.16636208	0.1614523	0.076555	0.07208023
	0.5	0.17667059	0.17661718	0.0930055	0.068673
	0.75	0.1935929	0.1933980	0.09351625	0.05959757
$\gamma_{2\tau}$	0.25	0.57522667	0.5733104	0.39888852	0.07208023
	0.5	0.63500834	0.62789313	0.4341381	0.433329
	0.75	0.7171769	0.7129003	0.47395937	0.46795710

The results of the Quantile structural equation model were compared with Lasso estimators for the structural equation (1.2), which are shown in Tables 1, 2 and 3 for all the hypotheses of error distributions, which are $N(0, 0.4)$, $t(5)$ and $LnN(0, 0.35)$. It has been proven that Lasso estimators are more efficient because it is less MSE and Bias in the

Table 2: Bayesian estimators of the regression coefficients for structural equation for sample size $n = 25$, $\varepsilon_i \sim t(5)$

δ_i		**n=25**		**t(5)**	
		BQSEM		**BQLsso**	
Par	τ	**RMS**	**Bias**	**RMS**	**Bias**
$b_{1\tau}$	0.25	0.07410148	-0.0118373	0.06568740	-0.0511498
	0.5	0.08304162	0.08054925	0.01297118	-0.0124334
	0.75	0.2727498	0.2235936	0.10316629	0.07475327
$\gamma_{1\tau}$	0.25	0.1704115	0.16709373	0.06886815	0.06620751
	0.5	0.1763317	0.17632264	0.09179943	0.06549495
	0.75	0.1869714	0.1860112	0.09142032	0.06389780
$\gamma_{2\tau}$	0.25	0.5614167	0.56016058	0.40046371	0.39962378
	0.5	0.62791284	0.62107868	0.41763562	0.41597434
	0.75	0.7293822	0.7253235	0.47400667	0.46586527

Table 3: Bayesian estimators of the regression coefficients for structural equation for sample size $n = 25, \varepsilon_i \sim \ln\ N(0, 0.35)$

δ_i		**n=25**		$\ln\ N(0, 0.35)$	
		BQSEM		**BQLsso**	
Par	τ	**RMS**	**Bias**	**RMS**	**Bias**
$b_{1\tau}$	0.25	0.06419643	-0.0105932	0.05785819	-0.0484447
	0.5	0.06242667	0.06231965	0.00776956	-0.0074559
	0.75	0.2656497	0.2214283	0.1096436	0.07971700
$\gamma_{1\tau}$	0.25	0.15787750	0.15453646	0.07084877	0.06884314
	0.5	0.16742392	0.16739034	0.087399142	0.06599080
	0.75	0.2001441	0.1999136	0.1066480	0.06435982
$\gamma_{2\tau}$	0.25	0.54994305	0.54731377	0.41025828	0.40896808
	0.5	0.61886252	0.61280847	0.42290513	0.4203174
	0.75	0.7106546	0.7065425	0.4469011	0.43697801

Table 4: Bayesian estimates of the parameters for measurement equation for sample size $n = 25$ with $\varepsilon_i \sim N(0, 0.4)$

n=25	**εi**	**N (0 0.4)**		
		τ=0.25		
	BQSEM		BQLsso	
par	RMS	Bias	RMS	Bias
λ21	0.2723786	0.2251264	0.2895677	0.2428582
λ31	0.1908624	0.1559416	0.2050512	0.1729432
λ52	0.2014524	0.2012309	0.1831696	0.1831486
λ62	0.1380175	0.1253578	0.1215341	0.1122781
λ83	0.2270678	0.2269210	0.1996226	0.1996048
λ93	0.4260815	0.4259355	0.3910151	0.3908882
a_1	0.2772910	-0.199912	0.1777285	-0.1536474
a_2	0.3399343	-0.188530	0.2262240	-0.13007292
a_3	0.3807948	-0.230275	0.2740716	-0.17601410
a_4	0.2747572	-0.128856	0.1947475	-0.0685311
a_5	0.1958347	-0.066282	0.1323766	-0.01356433
a_6	0.3318009	-0.250378	0.2675194	-0.20365583
a_7	0.2989145	-0.239258	0.2156715	-0.18370155
a_8	0.2718342	-0.129267	0.1921324	-0.06936999
a_9	0.2967833	-0.244413	0.2011347	-0.17631868
ϕ11	0.9624390	0.6415850	0.7782706	-0.3077184
ϕ11	0.7286783	-0.254164	0.7529228	-0.2137863
ϕ21	0.7286783	-0.254164	0.7529228	-0.2137863
ϕ22	0.8531774	0.8084972	0.3261442	-0.1074465

n=25	**εi**	**N (0 0.4)**	
		τ=0.5	
BQSEM		BQLsso	
RMS	Bias	RMS	Bias
0.2779994	0.2318954	0.2917239	0.2452200
0.1962811	0.1625731	0.2082480	0.1752191
0.2006136	0.2003969	0.1844621	0.1843456
0.1247432	0.1142498	0.1197662	0.1076939
0.2319345	0.2311487	0.1985172	0.1983994
0.4257695	0.4250122	0.3924662	0.3922321
0.2681579	-0.2647822	0.2197093	-0.20844048
0.2675158	-0.2191904	0.2472033	-0.17777998
0.3170085	-0.2617000	0.2944871	-0.21897541
0.2466052	-0.0988391	0.2351728	-0.10982790
0.1655518	-0.0424665	0.1606776	-0.04971412
0.3038372	-0.2280302	0.3040339	-0.23544626
0.2551209	-0.2264920	0.2439697	-0.19891463
0.2111867	-0.1046550	0.2211651	-0.08757637
0.2469740	-0.222384	0.2386273	-0.19934342
0.9666163	0.6533229	0.8017693	-0.3121501
0.7429847	-0.244579	0.7488955	-0.2177580
0.7429847	-0.244579	0.7488955	-0.2177580
0.8597813	0.8047740	0.3077471	-0.1069108

n=25	**εi**	**N (0 0.4)**	
		τ=0.75	
BQSEM		BQLsso	
RMS	Bias	RMS	Bias
0.2582621	0.2186927	0.2876030	0.2459550
0.1881680	0.1579935	0.2128413	0.1809846
0.1902863	0.1902813	0.1811771	0.1811651
0.1115734	0.1069998	0.1212607	0.1119983
0.2131624	0.2131280	0.1896719	0.1894530
0.4083165	0.4079739	0.3820790	0.3817064
0.3664790	-0.3532595	0.2137805	-0.211589848
0.3771138	-0.3081385	0.2187482	-0.172034278
0.4159396	-0.3443118	0.2660682	-0.212497004
0.2459933	-0.0804526	0.1762964	-0.053776408
0.1690368	-0.0260171	0.1168257	-0.002170028
0.2989621	-0.2103864	0.2501125	-0.192978590
0.2649451	-0.2216838	0.1863530	-0.176248619
0.2291696	-0.1113568	0.1499082	-0.067831678
0.2638958	-0.2304664	0.1798988	-0.172674511
0.9666121	0.6561557	0.8008439	-0.30803225
0.7413527	-0.2430368	0.7625531	-0.20609260
0.7413527	-0.2430368	0.7625531	-0.20609260
0.8666590	0.8214288	0.3253710	-0.07650527

case of $n = 25$ and for the three error distributions. Also, the estimators of the measurement equation coefficients equation (1.1) shown in Table 4, Lasso estimators are the best and most efficient. Also, the Lasso method is preferable in the case of $n = 25$.

6 Conclusion

In this article, the Bayesian Lasso technique has been applied to estimate the parameters of the divisional structural equation model to provide a comprehensive analysis of the interrelationships between the latent variables and compare them with the estimations of the parameters of the Quantile structural equation model.

References

[1] R. Alhamzawi, *Model selection in quantile regression models*, J. Appl. Stat. **42** (2015), no. 2, 445–458.

[2] D.F. Andrews and C.L. Mallows, *Scale mixtures of normal distributions*, J. Royal Stat. Soc.: Ser. B Method. **36** (1974), no. 1, 99–102.

[3] X. Feng, Y. Wang, B. Lub, and X. Song, *Bayesian regularized quantile structural equation models*, J. Multivariate Anal. **154** (2017), 234–248.

[4] R. Koenker and G. Bassett Jr., *Regression quantiles*, Econ.: J. Econ. Soc. **46** (1978), no. 1, 33–50.

[5] I. Komujer, *Quasi-maximum likelihood estimation for conditional quantiles*, J. Econ. **128** (2005), 137–164.

[6] H. Kozumi and G. Kobayashi, *Gibbs sampling methods for Bayesian quantile regression*, J. Stat. Comput. Simulat. **81** (2011), no. 11, 1565–1578.

[7] Q. Li, R. Xi, and N. Lin, *Bayesian regularized quantile regression*, Bayesian Anal. **5** (2010), 533–556.

[8] Y. Li and J. Zhu, *L_1-norm quantile regression*, J. Comput. Graph. Statist. **17** (2008), no. 1, 163–185.

[9] T. Park and G. Casella, *The Bayesian lasso*, J. Amer. Statist. Assoc. **103** (2008), 681–686.

[10] X.Y. Song and S.Y. Lee, *Basic and Advanced Bayesian Structural Equation Modeling: With Applications in the Medical and Behavioral Sciences*, John Wiley and Sons, 2012.

[11] R. Tibshirani, *Regression shrinkage and selection via the lasso*, J. Royal Stat. Soc. Ser. B: Statist. Method. **58** (1996), no. 1, 267–288.

[12] Y. Wang, X.-N. Feng, and X.-Y. Song, *Bayesian quantile structural equation models*, Structural Equ. Model.: Multidiscip. J. **23** (2016), no. 2, 246–258.

[13] K. Yu and R.A. Moyeed, *Bayesian quantile regression*, Statist. Prob. Lett. **54** (2001), 437–447.

[14] A. Zellner, *An Introduction to Bayesian Inference in Econometrics*, John Wiley and Sons, New York, 1971.