AL- Rafidain University College

PISSN: (1681-6870); EISSN: (2790-2293)
Journal of AL-Rafidain University College for Sciences

Available online at: https://www.jrucs.iq

JRUCS
Journal of AL-Rafidain University College for Sciences

The Bayesian Bridge of Quantile Structural Equation Model With Application

Lekaa A. Muhamed	Balsam M. Shafeeq
lekaa.ali.1968@gmail.com	balsammustafa95@mtu.edu.iq
Department of Statistics - College of Administration and Economics - University of Baghdad, Baghdad, Iraq	Department of Management Information Techniques - Technical College of Management-Middle Technical University, Baghdad, Iraq

Article Information

Article History:
Received: December, 19, 2022
Accepted: March, 3, 2023
Available Online: December, 31, 2023

Keywords:
Structural equation models, quantile regression, latent variable, Bayesian Lasso, Bridge technique.

Correspondence:
Lekaa A. Muhamed
lekaa.ali.1968@gmail.com

Abstract

The structural equation model (SEM) is widely recognized as the most important statistical tool for evaluating the interrelationships between latent variables and is one of the latent variable models. As a recent advance, Bayesian quantile SEM provides a comprehensive quantitative assessment of the conditional response-latent variables given both explanatory and latent covariates. In this study, we propose a Bridge technique in Bayesian quantile regression, compare it with a Bayesian least absolute shrinkage and selection operator (Lasso) and perform simultaneous estimation and variable selection in the context of quantile SEM. We suggest using Gibbs samples to perform Bayesian inference. Simulations with different sample sizes show that the proposed method gives good results. The proposed method was applied to a group of patients with Kidney failure disease to study the factors affecting this disease.

https://doi.org/10.55562/jrucs.v54i1.615

Introduction

Structural equation models (SEMs) are widely applied in the fields of social, behavioral and medical sciences to analyze latent variable that are measured by multiple highly correlated covariates. Structural equation modeling (SEM) is a popular multivariate technique for analyzing the interrelationships between latent variables.In classical SEMs, the effects of explanatory latent variables on outcome latent variables are assumed to have predefined parametric forms.

In quantile regression, the conditional quantiles of the response variable are assessed, given the explanatory variables. The main purpose of quantile regression is to obtain a highly comprehensive analysis of the relationship between variables by using different measures of central tendency and statistical dispersion. Although quantile regression has been rarely studied in structural equation models (SEMs).

Quantile regression has emerged as a useful supplement to classical mean regression, In recent years, variable selection based on penalty likelihood methods has used widly. Based on the Gibbs sampling algorithm of asymmetric Laplace distribution, this paper considers the quantile regression with Lasso penalty from a Bayesian point of view with proposed the Bridge technique in Bayesian quantile regression for Structural equation model.

1. Structural equation modeling (SEM)

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

Let $yi = (yi1 ,\ldots, yip)^T$ be a(px1) vector representing the ith observation in a random sample of size n, and $\omega i = (\omega i1,\ldots, \omega ip)^T$ be a (qx1) vector of latent variables with (q<p)

The link between yi and ωi is defined by the following measurement equation:

1) $$yi = Aci + \Lambda\, \omega i + \varepsilon i \quad , \quad i = 1,\ldots, n \tag{1}$$

2) ηi can be assessed in the following structural equation

$$\eta i=\beta\, di + \Gamma\tau\, \xi i + \delta i \qquad i = 1,\ldots, n \tag{2}$$

Then the Structural equation model (SEM) is defined by Equations (1) and (2).

Where A (p × r1) and Λ (p × q) are matrices of unknown coefficients, ci (r1× 1) is a vector of fixed covariates, and εi (p ×1) is a random vector of error terms.To analyze the interrelationship among latent variables, let partition $\omega i = (\eta i, \xi i)^T$,where ηi (q1 ×1) denote outcome latent variables and ξi (q2 × 1) is explanatory latent variables.

To simplify we assume that q1 = 1. The primary goal of SEM is to analyze the behavior of latent variable ηi given the information contained in a set of explanatory latent variables ξi . 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 εi 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 knowledge of the observed variables in the study. Any element of Λ can be a free parameter or fixed parameter with a preassigned value [18].

The positions and the preassigned values of fixed parameter are decided based on the prior knowledge of the observed and latent variables, and they are also related to the interpretations of latent variables. It can also be known from previous studies [18].

In the traditional Structural equation model (SEM), the error term εi is assumed to follow a normal distribution with mean zero. And therefore the conditional mean of yi is assumed to be a linear combination of latent factors ωi and covariates ci, with the error having a normal distribution. While this assumption is common and reasonable in many instances, it may induce biased estimates when the true underlying distribution of εi is highly non-normal, such as skewed. For this, the quantile regression for SEM will be adopte upon, which will be explained in the section (2). The rest of the paper is organized as follows. In section (2), we present Quantile Structural equation model (QSEM), In section(3) we present Bayesian lasso technique in a model,in section (4) we present Lasso technique in Bayesian quantile regression, in section (5) we present Bayesian Modeling for Lasso Quantile Structural Equation Model and and finding the conditional distributions of parameters and latent variable within the Bayesian lasso analysis by using Gibbs sampling, in section (6) we Proposed Bridge technique in Bayesian quantile regression for Structural equation model, in section(7) we present Bayesian model of the Bridge quantile SEM, A simulation study is conducted to evaluate the empirical performance of the proposed method in section 8 and in section 9 we report a real data on the determination of the risk factors of CKD patients. We conclude with brief conclusions in section (10).

2. Quantile Structural equation model (QSEM)

The primary aim of SEM is to analyze the behavior of latent variable ηi given the information contained in a set of explanatory latent variables ξi. This is done in traditional SEM by evaluating the conditional mean of (ηi\ξi) and fixed covariates di ($r_2 \times 1$) as follows [20]:

$$E(\eta i \setminus \xi i , di) = B\, di + \Gamma\, \xi i \qquad i = 1,\ldots, n \tag{3}$$

Where B ($q_1 \times r_2$) and Γ ($q_1 \times q_2$) 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 $Q(\eta i \setminus \xi i, di)$ the conditional quantile of ηi, under some different quantiles $\tau \in (0,1)$, as follows:

$$Q_\tau (\eta_i \setminus \xi_i , d_i) = B_\tau d_i + \Gamma_\tau \xi_i \qquad i = 1,\ldots, n \tag{4}$$

The coefficients matrices B_τ and Γ_τ have a subscript τ because they might not be equal for different quantiles. Thus, the model of quantile structural models is as follows:

1) 1- $y_i = Ac_i + \Lambda \omega_i + \varepsilon_i \quad , \qquad i = 1,\ldots, n$ (5)
2) ηi can be assessed in the following structural equation

$$\eta_i=\beta_\tau d_i + \Gamma_\tau \xi_i + \delta_i \qquad i = 1,\ldots, n \tag{6}$$

Where A ($p \times r_1$) and Λ ($p \times q$) are matrices of unknown coefficients, ci ($r_1\times 1$) is a vector of fixed covariates, and ε_i ($p \times 1$) is a random vector of error terms. Then the Quantile SEM is defined by Equations (5) and (6).

3. Bayesian lasso technique in a model

(Tibshirani 1996) proposed a penalty function for the linear regression model known as Lasso), which is abbreviated for (Least Absolute Shrinkage and Selection Operator). 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 was proposed to estimate the parameters of the linear regression model and to perform the variable selection simultaneously.

The principle of (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:

$$min_\beta \sum_{i=1}^{n}(y_i - x_i^{\mathrm{T}}\beta)^2 + \gamma \sum_{j=1}^{p}|\beta_j| \tag{7}$$

Where $\sum_{j=1}^{p}|\beta_j|$ is penalty function or It is sometimes called Regularization function, $\hat{\beta}_{\text{Lasso}} = (\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_p)$

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 (2008)[16] proposed a Bayesian framework of the Lasso (BLasso), they assumed 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 [10]:

$$\pi(\beta/\sigma^2) = \prod_{j=1}^{p} \frac{\gamma}{2\sigma} e^{-\frac{\gamma|\beta_j|}{\sigma}} \tag{8}$$

So that produces the same effect in contraction as in the original equation of Lasso as in equation (7). 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.In order to facilitate Gibbs sampling in Bayesian inference, in most research, the mixed representation of the Laplace function assumed by Andrews and Mallows 1974 is used, so that the probability density function of the Laplace distribution is written with a mixed representation of the two distributions (Normal & exponential), as follows [10]:

$$\frac{\gamma}{2\sigma} e^{-\gamma|\beta_j|/\sigma} = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2 s_j}} e^{-\beta_j^2/(2\sigma^2 s_j)} \frac{\gamma^2}{2} e^{-\gamma^2 s_j/2} ds_j \tag{9}$$

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 ($2/\Upsilon^2$) S_j is diagonal matrix ($s_1, \ldots, s_p$), The tuning parameter $\gamma^2 \sim Gamma(a1_\gamma, b1_\gamma)$ When $a1_\Upsilon$, $b1_\Upsilon$ are predefined hyperparameters, where it was specified by (Feng et al., 2015, we set $a1_\Upsilon=1$ and $b1_\Upsilon= 0.05$ for obtaining dispersed priors)[10;]

Based on the previously described hierarchical structure, Blasso may be easily used in more complex models, such as quantile regression models or quantile SEM, to conduct simultaneous estimation and variable selection.

4. Lasso technique in Bayesian quantile regression

As we explained in section 3 that the Quantile regression presented by Koenker and Bassett Jr (1978)[11] where the frequentist approach to the estimation of coefficients, is to solve the following optimization problem:

$$\min_{\beta} \sum_{i=1}^{n} \rho_\tau(y_i - x_i^{\mathrm{T}} \beta) \tag{10}$$

Where $\rho_\tau(x) = x(\tau - I(x<0))$ is the quantile loss function (so-called check function)
Li and Zhu (2008) proposed the regularized quantile regression to achieve estimation and variable selection, which uses the Lasso type penalty function, as follow [15]:

$$\min_{\beta} \sum_{i=1}^{n} \rho_\tau (y_i - x_i^{\mathrm{T}} \beta) + \gamma \sum_{j=1}^{p} |\beta_j| \tag{11}$$

In a Bayesian quantile regression framework, we need to specify a working likelihood function for the model. According to Yu and Moyeed (2001)[23], they proposed the Bayesian inference of quantile regression by introducing the ALD as a parametric link between the likelihood function of the quantile regression model is equivalent to minimizing the problem in equation (10). The Asymmetric Laplace (ALD) has its probability density function (pdf) as follows:

$$f(y \backslash \mu, \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp\left\{-\rho_\tau\left(\frac{y-\mu}{\sigma}\right)\right\} \tag{12}$$

Where μ is the location parameter, σ is the scale parameter and τ ($0 < \tau < 1$) is the skewness parameter. According to Yu and Moyeed (2001) implementing Bayesian inference for quantile regression, if the error term εi are follow AL(0, σ, τ), then the likelihood function for the quantile regression model as follows:

$$L(\beta, \sigma; y, X) = \frac{\tau^n (1-\tau)^n}{\sigma^n} exp\left\{- \frac{\sum_{i=1}^{n} \rho_\tau (y_i - x_i^{\mathrm{T}} \beta)}{\sigma}\right\} \tag{13}$$

Hence the solution of equation given by equation (10) is equivalent to maximizing the likelihood function (13), For the likelihood function equation (13), we suffer computation difficulty. Nevertheless, Kozumi and Kobayashi (2011) provided that the skewed Laplace distribution equation (13) can be viewed as a mixture of normal and exponential distribution as follow [15]:

$$y = \mu + k_1 e + \sqrt{k_2 \sigma e}\, \varsigma \tag{14}$$

Where $k_1 = (1-2\tau)/(\tau(1-\tau))$ $\quad$ $k_2 = 2/\tau(1-\tau)$
$\varsigma \sim N[0,1]$, $e \sim \exp(1/\sigma)$
The resulting conditional distribution of y is normal, with a mean ($\mu+k_1e$) and variance ($k_2\sigma e$). The posterior distribution of β can be expressed as follows:

$$f(\beta/y, X) \propto \pi(\beta) exp\left\{-\frac{\sum_{i=1}^{n} \rho_\tau(y_i - x_i^{\mathrm{T}} \beta)}{\sigma}\right\} \tag{15}$$

Where π (β) is a prior distribution, The prior distribution of β is not unique, but there have been many attempts by researchers, initially Yu and Moyeed (2001) 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 equation (10). And based on the aforementioned normal mixture representation of (εi) Kozumi and Kobayashi (2011) specified a conjugate normal prior for β, and the posterior of a normal distribution.

Feng et al.(2017) [15] have adopted Li et al. (2010) [15] proposing the Bayesian regularized quantile regression by employing the double-exponential prior in equation (8), such that the maximization of the posterior of β is equivalent to the minimization of equation (11) in Lasso technique.

In order 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 [10].

5. Bayesian Modeling for Lasso Quantile Structural Equation Model

The Bayesian hierarchical model based on the hierarchical model presented by Feng et al. (2017) 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 section (4).

Based on the hierarchical representation, the Bayesian Lasso can be conveniently implemented in a more complex Bayesian model by simply adding extra steps to the Gibbs sampling.

The common conjugated prior distributions were used in the Bayesian analysis of the structural equations model, as follows [10] [18]: To simplify the expression of the distributions we define several notations:

For the measurement equation (5), we let $\Omega = (\omega_1, \ldots, \omega_n)$, $\Lambda y = (A, \Lambda) = \{\lambda_{ykj}\}$, and define $Ly = \{l_{ykj}\}$ 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 $ui = (c_i^T, \omega_i^T)^T$, $U = (u_1, \ldots, u_n)$, and define U_k as the submatrix of U after removing the rows corresponding to lykj = 0. We let $Y^*_k = (y^*_{1k}, \ldots, y^*_{nk})^T$ with:

$$y_{ik}^{*} = y_{ik} - \sum_{j=1}^{r2+q2} \lambda_{ykj}\, u_{ij}(1 - I_{ykj}) \tag{16}$$

For the median regression in measurement equation (5), we can be expressed as follows:

$$\left(y_i/\omega_i, \theta 1_y, e_{yi}\right) \overset{ind}{\sim} N_p\left(Ac_i + \Lambda\omega_i, \Psi_i\right) \tag{17}$$

To simplify the notations, let $u_i = (ci^T, wi^T)^T$, $\theta 1_y = (A, \Lambda)$, $\theta_{1yk}{}^T$ be the kth row of θ_{1y} for k=1,….,p

Then the equation (5) is in the following form:

$$\left(y_i/\omega_i, \theta 1_y, e_{yi}\right) \overset{ind}{\sim} N_p\left(\theta 1_i u_i, \Psi_i\right) \tag{18}$$

and

$\theta 1_{yk} \sim N_{r1+q}(\Lambda_{0yk}, H_{0yk})$ $e_{yik} \sim \exp(\sigma_{yk})$ $\sigma_{yk}^{-1} \sim \text{Gamma}(a_{0\Upsilon k}, b_{0\Upsilon k})$

Where $a_{0\Upsilon}$, $b_{0\Upsilon}$, Λ_{0yk} and H_{0yk} (positive-definite matrix)are the hyperparameters and

$e_{yi} = (e_{yi1}, \ldots, e_{yip})^T$, $\Psi i = \text{diag}(8\sigma_{y1}e_{yi1}, \ldots, 8\sigma_{yp}e_{yip})$

And the structural equation (6) with Bayesian Lasso as follow:

Let $\beta_\tau = (\beta_{1\tau}^T, \beta_{2\tau}^T)^T$, $v_i = \left(d_i^T, \xi_i^T\right)^T$

$$\left(\eta_i/\xi_i, \theta 2_\omega, e_{\eta i}\right) \overset{ind}{\sim} N\left(\beta_\tau^T v_i + k_1 e_{\eta i}, k_2\sigma_\eta e_{\eta i}\right) \tag{19}$$

$\xi_i \overset{ind}{\sim} N_{q2}(0, \Phi)$ $\Phi^{-1} \sim Wishart(R_0, \rho_0)$ $\beta 1_\tau \sim N_{r2+q2}(0, S)$, where $S = \text{diag}(s_1, \ldots, s_{r2+q2})$

$s_j \sim Exponential(\frac{2\sigma_\eta}{\gamma^2})$ $\gamma^2 \sim Gamma(a1_\gamma, b1_\gamma)$ $\sigma_\eta^{-1} \sim Gamma(a1_\sigma, b1_\sigma)$ $e_{\eta i} \sim exponential\ (\sigma_\eta)$

Where $a_{0\Upsilon}$, $b_{0\Upsilon}$, Λ_{0yk} and H_{0yk}(positive-definite matrix) are the hyperparameters and $e_\eta = (e_{\eta 1}, \ldots, e_{\eta n})^T$.

As is known, a Bayesian estimate for parameters is obtained from the posterior joint distribution p (Ω, θ \ Y, C, D, eη) by 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. Bayes estimates for and were taken to be the sample mean for the random observations generated [10].

Then the full conditional posterior for the latent variable and parameters in QSEM: The full conditional posterior for the parameters of measurement equation (5) was derived, and the full conditional posterior for latent variable and parameters of structural equation (6) as follows:
The Gibbs sampling algorithm is implemented with the following full conditional posterior distribution of parameters and latent variable [20].
Let $\theta 1_y = (A, \Lambda)$, $\theta 2\omega = (B_\tau, \Gamma\tau)$, $u_i = (c_i^T, \omega_i^T)^T$ $v_i = (di^T, \xi_i^T)^T$, $U = (u1, \ldots, ui)$ where U_k be its submatrix with rows corresponding to Iykj =0 are deleted, $Y_k^* = (y_{1k}^*, \ldots, y_{nk}^*)$ where $y_{ik}^* = y_{ik} - \sum_{j=1}^{r_1+q} \lambda_{ykj} u_{ij} (1 - I_{ykj})$

1) The full conditional posterior distribution of the latent variable Ω

$$(y_i/\theta 1_y, \eta_i, \xi_i, , e_{yi}) \overset{ind}{\sim} N_p(Ac_i + \Lambda\omega_i, \Psi_i)$$

$$p(Y/\theta 1_y, \eta_i, \xi_i, e_{yi})\eta_i = (\Psi_i)^{\frac{-n}{2}} exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(y_i - \theta 1_y u_i)^T \Psi_e^{-1}(y_i - \theta 1_y u_i)\right\} \quad (20)$$

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

$$(\omega_i \backslash y_i \sigma_{yi} e_{yi} \theta 1_y \sigma_\eta e_{\eta i} \Lambda_\omega \Phi) \sim N_q(\mu_i, \Sigma_i^{-1*}) \quad (21)$$

Where :

$$\mu_i = \Sigma_i^{*-1} \Lambda^T \psi_i^{-1}(y_i - Ac_i) + \Sigma_i^{*-1} \Sigma_{\omega i}^{-1} \begin{pmatrix} B_\tau d_i + k_1 e_{\eta i} \\ 0 \end{pmatrix}$$

$$\Sigma_i^* = \Sigma_{\omega i}^{-1} + \Lambda^T \psi_i^{-1} \Lambda$$

$$\Sigma_{\omega i} = \begin{pmatrix} \Gamma_\tau \Phi \Gamma_\tau^T + k_2 \sigma_\eta e_{\eta i} & \Gamma_\tau \Phi \\ \Phi \Gamma^T & \Phi \end{pmatrix} \qquad \psi_i = diag(8\sigma_{y1} e_{yi1}, \ldots, 8\sigma_{yp} e_{yio})$$

2) The full conditional posterior distribution of the eyik : for (i=1,…,n , k=1,…,p)

$$p(e_{yik}^{-1} \backslash y_{ik}, \omega_i, \theta 1_{yk}, \sigma_{yk}) \propto f(y_{ik}, \omega_i, \theta 1_{yk}, \sigma_{yk}) f(e_{yik} \backslash \sigma_{yk})$$

$$p(e_{yik}^{-1} \backslash y_{ik}, \omega_i, \theta 1_{yk}, \sigma_{yk}) \propto \left\{\frac{2\sigma_{yk}^{-1}}{2\pi(e_{yik}^{-1})^3}\right\}^{\frac{1}{2}} exp\left\{\frac{2\sigma_{yk}^{-1}\left(e_{yik}^{-1} - \frac{4}{|y_{ik} - \theta 1_{yk} u_i|}\right)^2}{2\left[4|y_{ik} - \theta 1_{yk} u_i|^{-1}\right]^2 e_{yik}^{-1}}\right\} \quad (22)$$

Thus, the full conditional distribution of e_{yih} is an inverse Gaussian distribution $(4|yik - \theta 1_{yk} u_i|^{-1}, 2\sigma^{-1}_{yk})$

3) The full conditional posterior distribution of the θ1y: for (k=1,…,p)

$$p(\theta 1_{yk} \backslash Y, e_{yik}, \sigma_{yk}) \propto (\Sigma_{\theta 1k}^{-1})^{\frac{-1}{2}} \exp\left(\frac{-1}{2}(\theta 1_{yk} - M_{\Lambda k})^T (\Sigma_{\Lambda k}^{-1})^{-1} (\theta 1_{yk} - Mu_{\Lambda k})\right) \quad (23)$$

Where $Mu_{\Lambda k} = \Sigma_{\Lambda k}^{-1}\left(H_{0y}^{-1}\Lambda_{oy} + \sum_{i=1}^{n} \frac{y_{ik} u_i}{8\sigma_{yk} e_{yik}}\right)$

And $\Sigma_{\Lambda k} = H_{0y}^{-1} + \sum_{i=1}^{n} \frac{u_i u_i^T}{8\sigma_{yk} e_{yik}}$ Thus, The full conditional posterior distribution of the $\theta 1_y$ is a normal distribution $(Mu_{\Lambda k}, \Sigma^{-1}_{\Lambda k})$

4) The full conditional posterior distribution of the σyk : for (k=1,…,p)

$$p(\sigma_{yk}^{-1} \backslash Y, U, \Lambda_{yk}) \propto (\sigma_{yk}^{-1})^{n + a_{oyk} - 1} \exp\left\{\left(b_{oyk} + \frac{1}{2}\sum_{i=1}^{n} |y_{ik} - \theta 1_{yk} u_i|\right)\sigma_{yk}^{-1}\right\} \quad (24)$$

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

5) The full conditional posterior distribution of the Φ:

$$p(\Phi\backslash\Omega_2) \propto p(\Phi)\prod_{i=1}^{n} p(\xi_i\backslash\Phi) \tag{25}$$

$$p(\Phi\backslash\Omega_2) \propto |\Phi|^{-(n+\rho_0+q_2+1)/2} exp\left\{-\frac{1}{2}\ tr[\Phi^{-1}(\Omega_2\Omega_2^T + R_0^{-1})]\right\} \tag{26}$$

6) The full conditional posterior distribution of ση

$$\left(\sigma_\eta^{-1} \setminus \Omega, \beta_\tau, s_j, \gamma_j\right) \sim Gamma(n + a1_\sigma + r_2 + q_2, \quad b1_{\boldsymbol{\sigma}} + \sum_{i=1}^{n} \rho_\tau\left(\eta_i - \beta_\tau^{\mathrm{T}} v_i\right) + \frac{\gamma^2}{2} \sum_{j=1}^{r2+q2} s_j)$$

7) The full conditional posterior distribution of the eηi is a

Inverse Gaussian distribution $\left(\frac{k_2}{2|\eta_i - B_\tau d_i - \Gamma_\tau \xi_i|}, \frac{k_2}{4\sigma_\eta}\right)$ (27)

8) The full conditional posterior distribution of βτ:

$$f(\beta 1_\tau \setminus \Omega, e_\eta, \sigma_\eta) \propto f\left(\eta_i \setminus \Omega, e_\eta, \sigma_\eta\right) f(\beta_\tau)$$

$$f(\beta 1_\tau \setminus \Omega, e_\eta, \sigma_\eta) \sim N_{r2+q2}(\Sigma_\beta^{-1} V E_\sigma^{-1} \Xi^*, \Sigma_\beta^{-1}) \tag{28}$$

Where $\Sigma_\beta^{-1} = (S^{-1} + V E_\sigma^{-1} V^{\mathrm{T}})^{-1}$

9) The full conditional posterior distribution of S:

$$\left(s_j^{-1} \setminus \beta_{\tau j}, \gamma, \sigma_\eta\right) \sim Inverse - Gaussian\left(\frac{\gamma}{\sqrt{\sigma}\left|\beta_{\tau j}\right|}, \frac{\gamma^2}{\sigma_\eta}\right) \tag{29}$$

10) The full conditional posterior distribution of ϒ:

$$f(\gamma^2 \setminus s_j, \sigma_\eta) \propto f\left(s_i \setminus \sigma_\eta\right) f(\gamma^2)$$

$$f\left(\gamma^2 \setminus s_j, \sigma_\eta\right) \sim Gamma(a1_\gamma + r_2 + q_2, \quad b1_\gamma + \frac{\sum_{j=1}^{r2+q2} s_j}{2\sigma_\eta} \tag{30}$$

6. Proposed Bridge technique in Bayesian quantile regression (Proposed method)

Frank and Friedman (1993) proposed Bridge regression which is a broad class of the penalized regression method, The Bridge regression of frank and Friedman (1993) estimates linear regression coefficients β through Lq penalized least-square estimates. They achieve [2]:

$$\min_\beta (y - X\beta)^{\mathrm{T}}(y - X\beta) + \lambda \sum_{j=1}^{k} \left|\beta_j\right|^q \tag{31}$$

Where y= (y_1, …, y_n) and X = (x_1,….,x_n). Whereas the subject of interest is the quantile regression presented by Koenker and Bassett (1978), the quantile regression coefficients β can be estimated consistently by solving the following loss objective function:

$$\min_\beta \sum_{i=1}^{n} \rho_\tau(y_i - x_i^{\mathrm{T}}\, \beta) \tag{32}$$

Where $\rho_\tau(u)= u(\tau-I(u<0))$ is the quantile check function, and I(.) is an indicator function, which equals 1 if the argument is true and 0 otherwise [2].

In quantile regression model, to select significant covariates for improving prediction accuracy, we can select different of penalty functions, such as L_1 penalty (Lasso and Adaptive Lasso), L_2 penalty (ridge regression). L_ξ ($0 < \xi < 1$) penalty (Bridge regression), which result in the penalized quantile regression loss function as following:

$$Q(\beta) = Q_\theta(\beta) + p_\gamma(|\beta|) \tag{33}$$

Where $\gamma > 0$ is a penalty parameter, and p_Υ (|β|) is a penalty function on β.

The Bridge penalized regression is an important method that utilizes the L_ξ-norm penalty. In general, the Bridge regularization regression with L_ξ-norm penalty of quantile regression models can be expressed as follows:

$$\min_{\beta} \sum_{i=1}^{n} \rho_\tau(y_i - x_i^{\mathrm{T}} \beta) + \gamma \sum_{i=1}^{p} |\beta_j|^{\xi} \tag{34}$$

Where $\gamma > 0$ is a tuning parameter, and ξ is the penalty exponent of the Lξ-norm penalization function.

The Bridge penalized regression equation (34), penalization function $p_\gamma(|\beta|) = \gamma \sum_{j=1}^{p} |\beta_j|^{\xi}, (\xi \geq 0)$ Includes three popular special cases[2]: the best subset selection if $\xi = 0$, Lasso (Tibshirani 1996) if $\xi = 1$, and ridge regression (Hoerl and Kennard 1970) if $\xi = 2$. (Xu et al. 2010) showed that if $0 < \xi < 1$, the Lξ regularizer equation (34) keep many desirable statistical properties such as sparsity, oracle and unbiasedness [2]. And they detect that $L_{1/2}$ penalty is the most sparse and robust among the Lξ ($1/2 \leq \xi \leq 1$) and has similar properties to the Lξ ($0 < \xi < 1/2$) regularizers.

In the subject of Bayesian analysis, Polson et al. (2014) proposed a Bayesian Bridge estimator for regularized regression, they proposed a Bayesian inference which using a prior for β that decomposes as a product of independent exponential power prior as following [2]:

$$p(\beta \backslash \xi, \gamma) \propto \prod_{j=1}^{p} \exp(-\left|\frac{\beta_j}{\tau}\right|^{\xi}) \quad , \text{where } \tau = \gamma^{-1/\xi} \tag{35}$$

They utilize the Markov Chain Monte Carlo (MCMC) methods to obtain the posterior inference based on two different scale mixture representation of the Bayesian Bridge prior in equation (35), this mixture consists of the scale mixture of normal representation and the scale mixture of triangular representation, Mallick (2016) proposed an efficient Gibbs sampling method for bridge regression utilizing a scale mixture of uniform (SMU) representation of the Bayesian bridge prior, A connection with a particular gamma distribution provides tractable full conditional distributions so that β has a truncated multivariatnormal distribution. Alhamzawi and Algamal (2018) studied Bayesian Bridge Quantile regression with fixed penalty exponent ($\xi = 0.5$) [2]**.**

To conduct the fully Bayesian inference of Bridge regularized model in the equation, we need to specify the prior distribution. The analyze Bayesian Bridge quantile regression in equation (34), If we put an exponential power prior to equation

$$p(\boldsymbol{\beta} \backslash \xi, \gamma) \propto \exp[-\gamma \sum_{j=1}^{k} |\beta_j|^{\xi}] \tag{36}$$

On the regression coefficient and assume that the error term εi follow the ALD [33]**,** then the posterior distribution of β as follows:

$$p(\boldsymbol{\beta} \backslash y, \boldsymbol{X}, \sigma, \gamma) \propto \exp\left\{-\sigma \sum_{i=1}^{n} \rho_\tau(y_i - x_i^{\mathrm{T}} \boldsymbol{\beta}) - \gamma \sum_{J=1}^{k} |\beta_J|^{\xi}\right\} \tag{37}$$

As it was previously mentioned that minimizing the Bridge QR equation (34) is equivalent to maximizing the likelihood in equation (37), to conduct the fully Bayesian inference of Bridge regularized models we need to specify the prior distribution, Mallick and Yi (2016) proposed the scale mixture of uniform distribution, whereas the exponential power distribution can be represented as a scale mixture of uniform distribution as follows [2]:

$$\frac{\gamma^{\frac{1}{\xi}}}{2\Gamma(\frac{1}{\xi}+1)} \exp(-\gamma |t|^{\xi}) = \int_{s>|t|^{\xi}} \frac{\gamma^{\frac{1}{\xi}+1}}{2 s^{\frac{1}{\xi}} \Gamma(\frac{1}{\xi}+1)} s^{\frac{1}{\xi}} e^{-\gamma s}\, ds \tag{38}$$

This representation Mallick and Yi (2016) of the exponential power distribution results in an easy and efficient Gibbs sampler with traceable full conditional posterior distributions, which are used. For the scale parameter, we also assign the gamma prior for the scale parameter σ, to summarize the following is the proposed Bayesian hierarchical representation [2]:

$$y_i = x_i^{\mathrm{T}}\beta + (1-2\tau)e_i + \sqrt{2\sigma_b{}^{-1}e_i}\,\zeta_i$$
$$\zeta \sim \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}\varsigma_i^2)$$
$$e\backslash\sigma_b \sim \prod_{i=1}^{n} \tau(1-\tau)\sigma_b \exp(-\tau(1-\tau)\sigma_b e_i) \tag{39}$$
$$(\beta\backslash s,\xi) \sim Uniform(-s_j^{\frac{1}{\xi}}, s_j^{\frac{1}{\xi}})$$
$$(s\backslash\gamma,\xi) \sim \prod^{n} Gamma(\frac{1}{\xi}+1,\gamma)$$
$$\sigma_b \sim Gamma\ (a1_\sigma, b1_\sigma)$$
$$\gamma \sim Gamma\ (a2_\gamma, b2_\gamma)$$

In addition to the prior distributions of the rest of the parameters of the measurement equation (5), which was mentioned in section (5) (Lasso), For the median regression in measurement equation (5), is as follows [10]

The common conjugated prior distributions were used in the Bayesian analysis of the structural equations model

In this section of the thesis, it was proposed to use the hierarchical Bayesian with Bridge Lasso technique to Quantile Structural Equations model (QSEM)and by mixed representation by Mallick and Yi(2016) and by proposing that the exponent parameter is a constant value equal to (0.5) in Equation (34) and this method is considered one of the commonly used organization methods, but according to the researcher's knowledge it has not been applied within Bayesian analysis quantile Structural Equation Models.

7. Bayesian model of the Bridge quantile SEM

The Bayesian hierarchical model based on the hierarchical model presented by Alhamzawi and Algamal (2018) [2] used in estimating the parameters of the quantile regression by applying Bridge Lasso technique, and applied this technique in structural equation model, which was explained in section (1).

There is no available explicit formula of the joint posterior distribution of the unknown parameters, to update the unknowns from the posterior distribution, we utilize Gibbs sampling.

Let θ = [β1,ω s, γ, σ, ξ] be a vector of unknown parameters. The joint posterior distribution of the unknowns (parameters and latent variables) is not available in explicit form. Given that θ includes multiple components, p(θ|y, x, v) is complicated, Thus, we utilize the Gibbs sampling to update the unknown parameters from the posterior distributions. Based on the hierarchical model and prior specifications, the full conditional posterior for the parameters of measurement equation (5) was derived within the section (5) and the full conditional posterior for latent variable and parameters of structural equation (6) as follows:

1) The full conditional posterior distribution of the Φ as mentioned in the equation (25) is given by: $[\Phi\backslash\Omega_2] \sim IW_{q2}[(\Omega_2\Omega_2^T + R_0^{-1}), n+\rho_0]$

2) Updating β1ω

The full conditional posterior distribution of βτ is truncated normal (TN):

$$f(\beta 1_\omega \setminus \Omega, e_\eta, \sigma_\eta, s_j) \propto f\big(\eta_i \setminus \Omega, e_\eta, \sigma_\eta\big) f(\beta 1_\omega \backslash s_j)$$
$$\propto exp\left\{-\frac{1}{k_1\sigma_\eta e_{\eta i}}\left[\eta_i - \beta_\tau d_i - \Gamma_\tau\xi_i - k_1 e_{\eta i}\right]^T \left(k_1\sigma_\eta e_{\eta i}\right)^{-1}\left[\eta_i - \beta_\tau d_i - \Gamma_\tau\xi_i - k_1 e_{\eta i}\right]\right\} \prod_{j=1}^{r2+q2} I\left|\beta_j\right| < s_j^{\frac{1}{\xi}}$$

then

$$(\beta 1_\omega \setminus \Omega, e_\eta, \sigma_\eta) \sim N(\,\mathrm{M}u_{BB}, \Sigma_{BB}{}^{-1}) \prod_{j=1}^{r2+q2} I\left|\beta_j\right| < s_j^{\frac{1}{\xi}} \tag{40}$$

Where mean: $\mathrm{M}u = (V^T W V)^{-1} V^{\mathrm{T}} W \Xi$,

variance- covariance matrix: $W = diag(\frac{\sigma}{2e_1}, \dots, \frac{\sigma}{2e_n})$
$\tilde{\Xi} = [\eta_1 - (1-2t)e_{\eta 1}, \dots, \eta_n - (1-2t)e_{\eta n}]^T$
and
$\Sigma_{BB}{}^{-1} = (V^T W V)^{-1}$, and $I(\beta_j)$ denotes an indicator function

3) Updating eηi: for i=1,..,n

The full conditional posterior distribution of $e_{\eta i}$
$p(e_{\eta i}^{-1} \backslash \omega_i, \beta 1_\omega, \sigma_\eta) \propto f(\omega_i, \beta 1_\omega, e_{\eta i}^{-1}, \sigma_\eta) f(e_{\eta i}^{-1} \backslash \sigma_{yk})$

$$(e_{\eta i}^{-1} \backslash \omega_i, \beta 1_\omega, \sigma_\eta) \propto \left\{\frac{\frac{\sigma}{2}}{2\pi (e_{\eta i}^{-1})^3}\right\}^{\frac{1}{2}} exp\left\{\frac{\frac{\sigma}{2}(e_{\eta i}^{-1} - \frac{1}{\sqrt{(\eta_i - \beta 1_\omega v_i)^2}})^2}{2\left[\frac{1}{\sqrt{(\eta_i - \beta 1_\omega v_i)^2}}\right]^2 e_{\eta i}^{-1}}\right\} \tag{41}$$

Thus, the full conditional posterior distribution of the $e_{\eta i}$ is a
$(e_{\eta i}^{-1} \backslash \omega_i, \beta 1_\omega, \sigma_\eta) \sim Inverse\ Gaussian(\grave{M}, \dot{\varphi})$
Where $\grave{M} = \frac{1}{\sqrt{(\eta_i - \beta 1_\omega v_i)^2}}$ and $\dot{\varphi} = \frac{\sigma}{2}$

4) Updating ση

The full conditional posterior distribution of σ:

$$(\sigma_\eta \ \backslash \Omega, \beta 1_\omega \ , s_j, \gamma_j) = \sigma_\eta^{a_{BB}-1} \exp(b_{BB} \sigma_\eta) \tag{42}$$

Where

$$a_{BB} = a1_\sigma + \frac{3n}{2} \qquad b_{BB} = b1_\sigma + \frac{\sum_{i=1}^n \rho_\tau (\eta_i - \beta 1_\omega^{\mathrm{T}} v_i - (1-2t)e_i)^2}{2e_i} + t(1-t)e_i)$$

$$(\sigma_\eta \ \backslash \Omega, \beta 1_{\omega_\tau}, s_j, \gamma_j) \sim Gamma(a1_\sigma + \frac{3n}{2}, b1_\sigma + \frac{\sum_{i=1}^n \rho_\tau (\eta_i - \beta 1_\omega^{\mathrm{T}} v_i - (1-2t)e_i)^2}{2e_i} + t(1-t)e_i)$$

5) Updating sj

The full conditional posterior distribution of $s_{j,}$(j=1…r_2+q_2)is a left-truncated exponential distribution given by:
$(s_j \backslash \beta 1_\omega, \gamma) \sim Exp(\gamma) I\{s_j > |\beta 1_\omega|^\xi$
Updating s_j can be done by using the inversion method as follows [2].

1) Update s*j from Exponential (γ)

2) Set $s_j = s_j^* + |\beta 1_j|^\xi$ (43)

$\pi(s_j)$ is a left-truncated exponential distribution, sampling from which is accomplished by using the inverse transformation method with two substeps:

a) Generate $s_i{}^* \sim Exp(\lambda)$ and
b) $s_i = s_i{}^* + |\beta_i|^{n1}$, i = 1, · · · , p

3) Updating γ : The full conditional posterior distribution of the γ

$$f(\gamma \backslash s_j, \sigma_\eta) = \gamma^{a1_\gamma + 2(r2+q2)-1} e^{-(b1_\gamma + \frac{\sum_{j=1}^{r2+q2} |s_j|^\xi}{2\sigma_\eta})\gamma}$$

Thus, the full conditional posterior distribution of γ is
$f(\gamma \backslash s_j, \sigma_\eta) \propto f(s_i \backslash \sigma_\eta) f(\gamma)$

$$f(\gamma \backslash s_j, \sigma_\eta) \sim Gamma(a1_\gamma + 2(r_2 + q_2), \qquad b1_\gamma + \frac{\sum_{j=1}^{r2+q2} |s_j|^\xi}{2\sigma_\eta} \tag{44}$$

8. Simulation studies

The performance of the lasso method and proposed Bridge lasso method was illustrated using simulation studies with different sample sizes, the simulation study's main goal is to estimate the quantile regression coefficients b_1, γ_1 and γ_2

For structural equations model (SEM) at different quantiles:

We consider the quantile structural equations model (QSEM) given by (5) and (6) with p=9 , q=3 , q_1=1, q_2=2 , r1= r2= 1 and n= 50, 100 with three different values of quantity were used, which are (0.25, 0.5, 0.75). For measurement equation (5) the true value was determined as follows:

A= (0.5, …, 0.5) and the fixed covariate c_i , we can generate it by drawing from N(0,1) and the factor loading matrix Λ has the common non-overlapping structure

$\lambda_{21}= \lambda_{31}= \lambda_{52}= \lambda_{62}= \lambda_{83}= \lambda_{93}=$ 0.7, then the factor loading matrix Λ will be in the following, The true values of parameters λjk and aj in the measurement equation are taken to be:

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

For structural equation (6) the true value is:

The true values of parameters $B\tau = (b_1) = (0.1)$, $\Gamma\tau = [\gamma_1, \gamma_2] = [0.1, 0.3]$

And the explanatory latent variable $\xi_i = (\xi_{i1} , \xi_{i2})^T$ is assumed to follow a normal distribution N(0, Φ) where $\Phi = \begin{bmatrix} 1 & 0.2 \\ 0.2 & 1 \end{bmatrix}$ Also the fixed covariate d_i, we can generate by drawing from N(0,1). The distributions for the error term in each equation of the structural equation model (SEM), equation (5) and (6), they are assumed as follows:

- **i.** εi and δi 's follow the normal distribution N (0, 0.4)
- **ii.** εi and δi 's are distributed as the heavy- tailed central t-distribution t(5)

The sampling was carried out using the Gibbs Sampler algorithm with the Metropolis Hastings algorithm (10,000 iterations with the initial 2000 observations dropped in the burn-in) from the posterior conditional distributions and The performance of the proposed method is assessed by the (Bias) and root mean square error (RMS)

The conjugate normal prior of $\theta 1_{yk} \sim N_{r1+q} (\Lambda_{0yk}, H_{0yk})$ where the mean Λ_{0yk} is 0.7 and the covariance matrix H_{0yk} is a diagonal matrix with diagonal element 10. Also $\theta 2_{\omega\tau} \sim N_{r2+q2}(\Lambda_{0\omega}, H_{0\omega})$ where the mean $\Lambda_{0\omega}$ is= [1, 0.4, 0.4] and the covariance matrix $H_{0\omega}$ is an identity matrix. The conjugate inverse gamma prior of $\sigma^{-1}_{yk} \sim$ Gamma ($a0_{yk}$, $b0_{yk}$) and $\sigma^{-1}_{\eta} \sim$ Gamma ($a0_\sigma$, $b0_\sigma$) we specified $a0_{yk}= a0_\sigma=1$, $b0_{yk}= b0_\sigma=1$ and the tuning parameter was mentioned in BLasso and BALsso $\gamma^2 \sim Gamma(a1_\gamma, b1_\gamma)$ we specified a1γ =1 and b1γ= 0.05 ,The inverse Wishart prior of $\Phi^{-1} \sim$ Wishart (R_0,ρ_0) we specified R_0= 5 I_2 and ρ_0=4.

Table (1): shows the Bayesian estimates of the Quantile regression coefficients in the structural equation (6) with the RMS Criteria and the BIAS Criteria, with $\delta_{ik} \sim$ N(0, 0.4) with n=50

δi	n=50 N (0 0.4)				
		BQLsso		BBQSEM	
Par	τ	RMS	Bias	RMS	Bias
B1τ	0.25	0.3203991	0.2034196	0.203687	0.190439
	0.5	0.31898737	0.22608887	0.256814	0.162779
	0.75	0.3414035	0.2797307	0.281436	0.218339
γ1τ	0.25	0.4470690	0.3847818	0.403564	0.335814
	0.5	0.47278105	0.40327755	0.431297	0.342611
	0.75	0.4535363	0.3920687	0.403812	0.312782
γ2τ	0.25	0.6168246	0.5996073	0.605692	0.554468
	0.5	0.56062748	0.55901562	0.561235	0.0.55218
	0.75	0.5800023	0.5789627	0.501237	0.569921

Tables (2): shows the Bayesian of the parameters for measurement equation (5) with n=50, εik~ N(0, 0.4), for τ= 0.25,0.5, 0.75

	n=50 εi N (0 0.4) τ=0.5			
	BQLsso		BBQL	
par	RMS	Bias	RMS	Bias
λ21	0.2632676	0.2613677	0.272241	0.256481
λ31	0.2344360	0.2339240	0.214751	0.214358
λ52	0.1884731	0.1868658	0.182146	0.183200
λ62	0.1899072	0.1665583	18234512	0.161478
λ83	0.1868642	0.1857021	0.184015	0.180245
λ93	0.3090305	0.2980920	0.310254	0.297250
a1	0.19255419	-0.1656226	0.180123	-0.165281
a2	0.20055532	-0.1885822	0.190378	--0.18921
a3	0.19207265	-0.1159183	0.1720.3	-0.118536
a4	0.17305769	-0.1682224	0.171023	-0.182456
a5	0.05667344	-0.0383346	0.023489	-0.05674
a6	0.19528392	-0.1424781	0.169874	-0.153698
a7	0.17555317	-0.1681655	0.185641	-0.200158
a8	0.11800695	-0.1139600	0.124581	-0.169548
a9	0.17118411	-0.1546713	0.165483	-0.20.158
φ11	0.5053664	-0.4728446	0.482681	-0.482364
φ12	0.2484923	-0.2392917	0.668542	-0.260158
φ21	0.2484923	-0.2392917	0.668542	-0.260158
φ22	0.3083302	-0.2864561	0.292456	-0.28934
	n=50 εi N (0 0.4) τ=0.25			
	BQLsso		BBQL	
par	RMS	Bias	RMS	Bias
λ21	0.2593113	0.2570893	0.251187	0.247329
λ31	0.2333437	0.2326273	0.221791	0.225761
λ52	0.1849664	0.1835414	0.166788	0.17758
λ62	0.1740932	0.1530985	0.174100	0.154244
λ83	0.1865514	0.1852896	0.185234	0.182456
λ93	0.3096126	0.2998021	0.300421	0.299801
a1	0.13464383	-0.13463404	0.134100	-0.14632
a2	0.15630991	-0.15326645	0.156247	-0.14525
a3	0.10781653	-0.09395914	0.100451	0.12473
a4	0.1796170	-0.16328014	0.179145	-0.17254
a5	0.02965436	-0.02953254	0.012745	-0.03548
a6	0.17554002	-0.14131672	0.118263	-0.17452
a7	0.17841525	-0.17834769	0.179354	-0.17654
a8	0.13480558	-0.11624865	0.135246	-0.10345
a9	0.15309283	-0.14982493	0.152481	-0.15324
φ11	0.4849325	-0.4530673	0.18114	-0.46351
φ11	0.2459442	-0.2348389	0.25661	-0.24681
φ21	0.2459442	-0.2348389	0.25661	-0.23451
φ22	0.3088758	-0.2826899	0.30547	-0.27485

	n=50 εi N (0 0.4) τ=0.75			
	BQLsso		**BBQL**	
par	RMS	Bias	RMS	Bias
λ21	0.2607131	0.2587737	0.2401269	0.241268
λ31	0.2327074	0.2324960	0.2162255	0.217159
λ52	0.1927438	0.1906758	0.1891023	0.184923
λ62	0.1915750	0.1693943	0.1811726	0.158623
λ83	0.1832952	0.1825820	0.8523691	0.182569
λ93	0.3018115	0.2933194	0.3025832	0.292568
a_1	0.24159014	-0.2375906	0.2421580	-0.300178
a_2	0.24527701	-0.2445472	0.2451178	-0.296483
a_3	0.20127242	-0.1664618	0.2011259	-0.222567
a_4	0.18907386	-0.1766708	0.1886903	-0.202389
a_5	0.04522357	-0.0415459	0.0412586	-0.067123
a_6	0.19082905	-0.1538869	0.1916835	-0.172681
a_7	0.18519245	-0.1851555	0.1862549	-0.230183
a_8	0.14129895	-0.1208219	0.1587738	-0.162884
a_9	0.16517565	-0.1628807	0.1724564	-0.210339
φ11	0.5206929	-0.4862106	0.5158144	-0.473691
φ12	0.2590470	-0.2455439	0.2603473	0.698743
φ21	0.2590470	-0.2455439	0.2603473	0.698743
φ22	0.3207422	-0.2874333	0.3203674	0.669811

Table (3): shows the Bayesian estimates of the Quantile regression coefficients in the structural equation (6) with sample size (n=100) With δik~ N(0, 0.4)

δi	n=100 N (0 0.4)				
		BQLsso		**BBQSEM**	
Par	τ	RMS	Bias	RMS	Bias
B1τ	**0.25**	0.3081547	0.3004298	0.274002	0.252193
	0.5	0.34281405	0.34016101	0.314882	0.312361
	0.75	0.3873789	0.3868042	0.3469142	0.345682
γ1τ	**0.25**	0.3580411	0.3521044	0.3111564	0.301009
	0.5	0.38583457	0.38495291	0.3655319	0.336817
	0.75	0.3857904	0.3848131	0.3425803	0.332961
γ2τ	**0.25**	0.6401845	0.6365365	0.6362781	0.634825
	0.5	0.61508492	0.61270257	0.6110293	0.625404
	0.75	0.6485689	0.6450752	0.6243691	0.643510

Table (4): shows the Bayesian estimates of the Quantile regression coefficients in the structural equation (6) with the sample size(n=50) and δik~ t(5)

δi	n=50 t (5)				
		BQLsso		**BBQSEM**	
Par	**τ**	**RMS**	**Bias**	**RMS**	**Bias**
B1τ	**0.25**	0.05146601	-0.0450126	0.158015	-0.61843
	0.5	0.30357562	0.21341256	0.198237	0.093147
	0.75	0.3355429	0.2732943	0.205831	0.221579
γ1τ	**0.25**	0.07086804	0.06567936	0.059147	0.014583
	0.5	0.47390374	0.40581551	0.371583	0.304811
	0.75	0.4532014	0.3897087	0.312831	0.301585
γ2τ	**0.25**	0.40133747	0.40048270	0.392483	0.421843
	0.5	0.55726037	0.55564117	0.5029716	0.592411
	0.75	0.5886226	0.5875550	0.5097126	0.0.62869

Table (5) shows the Bayesian estimates of the Quantile regression coefficients in the structural equation (6) with the sample size (n=100) and δik~ t(5)

δi	n=100 t (5)				
		BQLSEM		BBQSEM	
Par	**τ**	**RMS**	**Bias**	**RMS**	**Bias**
B1τ	**0.25**	0.0075931	0.2998867	0.111536	-0.62638
	0.5	0.0510356	0.34823695	0.1909731	0.091500
	0.75	0.0845135	0.3841183	0.204897	0.213919
γ1τ	**0.25**	0.024999	0.3457980	0.0152866	0.014754
	0.5	0.38865036	0.38781442	0.3705940	0.3025711
	0.75	0.3866752	0.3855469	0.302587	0.3570094
γ2τ	**0.25**	0.3566731	0.6529677	0.6625849	0.654369
	0.5	0.30595768	0.60361698	0.5019371	0.5912076
	0.75	0.3547294	0.6505621	0.502269	0.0619382

The results showed a better performance of the method proposed (BBQSEM) in the research for estimating the parameters and latent variable within the structural equations Equation (5). Compared to the method BQLsso proposed by Feng et al. (2017) [15], where we found that the performance of this method is better than the other method.

The proposed method (BBQSEM) is better for the quantiles 0.5 and 0.75 in a sample size of 100 and the distributed error term t(5),the comparison of the proposed method with other method (BQLsso).

9. Real application

We applied the proposed method was employed in the research to evaluate the serious of chronic kidney disease (CKD), by the logarithm urinary albumin-creatinine ratio (ACR) and estimated glomerular filtration rate (eGFR), both of which are central for diagnosis and staging of CKD They are the most important basic variables that specialist doctors usually rely on in diagnosing the patient's condition, taking the necessary measures and determining the appropriate treatment. Chronic kidney disease is one of the serious diseases facing the individual, which is a gradual loss of the efficiency of kidney work over long periods and this is often not noticed. This disease has serious complications on the heart and blood vessels, diabetes and anemia, as well as other serious effects, and advanced cases of this disease represent a real threat to a person's life.

The QSEM was used to evaluate the effects of CDK (Chronic kidney disease) on patients in this section .the study’s primary goal was to study potential CKD risk factors such as Hypertension, Glycemia, Obesity and body shape

The studied variables are measured as possible risk factors as follows:

- ξ1: Blood pressure, it contains SBP and DBP
- ξ2: Glycemia, it contains glycated hemoglobin (HbAc) and fasting glucose (FBG)
- ξ3 Body Shape, including weight and height
- η: CKD(Chronic kidney disease), It includes a combination of two factors, albumin-creatinine ratio (ACR) and estimated glomerular filtration rate (eGFR)

A smaller eGFR value or a greater ACR value indicates a more severely impaired kidney function, which may lead to kidney failure.

The clear interpretation of each latent variable suggests the measurement equation (5),
Where $\omega_i = (\eta_i, \xi_{i1}, \xi_{i2}, \xi_{i3})^T$, $A = (a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8)^T$ and the factor loading matrix Λ takes the following non-overlapping structure:

$$\Lambda^T = \begin{bmatrix} 1^* & \lambda_{2\eta} & 0^* & 0^* & 0^* & 0^* & 0^* & 0^* \\ 0^* & 0^* & 1^* & \lambda_{41} & 0^* & 0^* & 0^* & 0^* \\ 0^* & 0^* & 0^* & 0^* & 1^* & \lambda_{62} & 0^* & 0^* \\ 0^* & 0^* & 0^* & 0^* & 0^* & 0^* & 1^* & \lambda_{83} \end{bmatrix}$$

Where the zeros and ones with an asterisk were fixed in advance to identify the model and unify the scale of the latent variable. To assess the effects of latent variables ξ_1 to ξ_3 and fixed covariate "Disease period" (d1) on CKD (η) under different quantiles, the structural equation can be show as follows:

$$\eta_i = b_{1\tau} d_{i1} + \gamma_{1\tau}\xi_{i1} + \gamma_{2\tau}\xi_{i2} + \gamma_{3\tau}\xi_{i3} + \delta_i \quad (45)$$

Where $b_{1\tau}$, $\gamma_{1\tau}$, $\gamma_{2\tau}$ and $\gamma_{3\tau}$, are the regression coefficients under τ -quantile. We collected 77 observations of chronic kidney patients, the estimation was conducted for the quantiles of (0.05, 0.25, 0.3, 0.5, 0.75).

The proposed method was applied in the research to estimate the parameters of the SEM model that was used to analyze the relationship between the latent variables that represent the risk factors for the disease referred to in the previous paragraph. Table (8) shows the estimated values of the model parameters in each quantile:

Table (8): Bayesian estimates of quantile SEM parameters in the analysis of CKD data

	Estimation of the parameters at quantiles				
Parameters	**0.05**	**0.25**	**0.3**	**0.5**	**0.75**
	measurement equation				
$\lambda_{2\eta}$	0.921	0.901	0.933	1.120	1.074
λ_{41}	0.642	0.682	0.676	0.751	0.762
λ_{62}	0.211	0.312	0.3011	0.250	0.399
λ_{83}	0.334	0.294	0.298	3.446	1.226
a1	6.285	4.197	0.510	5.119	6.441
a2	5.221	5.009	4.498	4.827	4.260
a3	3.442	3.110	3.349	4.081	2.116
a4	2.003	2.846	2.199	3.117	2.996
a5	5.113	5.004	5.121	6.113	5.994
a6	2.006	2.887	2.989	3.441	3.007
a7	6.772	8.441	6.998	7.119	6.996
a8	7.836	6.043	6.221	6.245	6.111
	structural equation coefficients (parameter and latent variables)				
$b_{1\tau}$	0.2151	0.1745	0.2011	-0.1254	-0.7451
$\gamma_{1\tau}$	0.4256	0.2561	0.2145	0.2499	0.2400
$\gamma_{2\tau}$	0.1642	0.0542	0.0321	0.02584	0.0141
$\gamma_{3\tau}$	0.0455	0.0566	0.1004	0.1564	0.9836
	Covariance matrix of the latent variable				
ϕ_{11}	0.7871	0.8534	0.8529	0.8751	0.8241
ϕ_{21}	0.1593	0.1856	0.1872	0.1456	0.1431
ϕ_{22}	-0.058	-0.0521	-0.0531	-0.0412	-0.0410
ϕ_{31}	0.8233	0.8801	0.8124	0.9751	0.9102
ϕ_{32}	0.0490	0.0723	0.0426	0.0367	0.0187
ϕ_{33}	0.1552	0.1747	0.1647	0.1743	0.1729

1. Blood pressure is positively correlated with the severity of kidney disease, and the association is stronger in light of higher quantities and this is confirmed by doctors through examinations of patients, and high blood pressure is the main cause of chronic kidney disease, and that poor kidney function, in turn, will increase blood pressure.
2. Glycemia has a weak positive correlation with CKD for lower quantiles, but not for higher quantiles.
3. There is very little association between body shape and chronic kidney disease. This is true, but the severity of the disease and its danger by overweight

To summarize, in order to slow down the progression of CKD it is important to effectively control the risk factors, such as hypertension and hyperglycemia especially for patients with advanced stages of CKD and to reduce weight. Quantile SEM provides a more comprehensive picture of the relationship between CKD development and various risk factors. Results with some medical knowledge provide deeper insight that cannot be achieved with traditional SEMs.

10. Conclusions

In this paper, we introduce a QSEM to provide a comprehensive analysis of the interrelationships among latent variables, and develop a BB approach to make statistical inference on QSEM. In the QSEM, latent variables are imputed through the estimated density function and the linear interpolation method.

The MCMC algorithm is presented to sample observations required for statistical inference by combining the Gibbs sampler and the Metropolis-Hastings algorithm. A simulation study is conducted to investigate the finite sample performance of the proposed methodologies. An example is illustrated the proposed methodologies. The proposed method has the following characteristics. First, the method does not depend on the distribution of random errors and latent variables, and is robust. The empirical likelihood is employed to establish the working likelihoods so that the proposed method is less vulnerable to extreme distributions. Second, the proposed method can estimate parameters in the QSEMs when simultaneously considering multiple quantiles, and can be used to investigate the effect of explanatory latent variables on outcome latent variables under various quantiles. Although we only consider linear quantile for the structural equation model, the above proposed method can be used to analyze a QSEM with nonlinear, nonparametric or semiparametric quantiles.

References

[1] Rahim Alhamzawi & Haithem Taha Mohammad Ali, 2020. "Brq: an R package for Bayesian quantile regression," METRON, Vol. 78(3), pages 313-328, December.

[2] R. Alhamzawi, and Z. Y. Algamal, “Bayesian bridge quantile regression,” Communication in statistics – Simulation and Computation, Vol.0, No. 0, 1-13, 2017.

[3] R. Alhamzawi, and Y. Keming, “Bayesian Lasso-mixed quantile Regression,” Journal of Statistical Computation and Simulation, Vol. 84, no. 4, pp. 868-880, 2012

[4] D. F. Andrews, and C. L Mallows, “Scale mixtures of normal distributions,” Journal of the Royal Statistical Society. Series B (Methodological), Vol. 36, No. 1, 99-102, 1974

[5] D. F. Benoit, and D. Van den Poel, “Binary quantile regression: a Bayesian approach based on the asymmetric Laplace distribution,” Journal of Applied Econometrics, 27(7): 1174-1188, 2012

[6] S. Chand, “Goodness of Fit Tests and Lasso Variable Selection in Time Series Analysis,” Unpublished PhD Thesis, University of Nottingham, MT. United Kingdom, 2011

[7] B. D. Dunson, J, Palomo, and K. Bollen, “Bayesian Structural Equation Modeling,” Statistical and Applied Mathematical Sciences Institute. Technical Report. p.p (1-36), 2005

[8] R. Everett, An introduction to latent variable models, Springer Science & Business Media, 2013

[9] J. Fan, and R.Z. Li, “Variable selection via non-concave penalized likelihood and its oracle properties,” Journal of the American Statistical Association, Vol. 96, pp. 1348–1360, 2001

[10] X. Feng, Y. Wang, B. Lub, and X. Song, “Bayesian regularized quantile structural equation models,” Journal of Multivariate Analysis, Vol. 154, pp. 234–248, 2017

[11] R. Koenker, G.Jr. Bassett, “Regression Quantiles,” Econometrica Journal of the Econometric Society. 46, pp. 33-50, 1978

[12] R. Koenker, Quantile regression, Cambridge University Press, London, 2005

[13] H. Kozumi, and G. Kobayashi, “Gibbs Sampling Methods for Bayesian quantile regression,” Journal of Statistical Computation and Simulation, Vol. 81, No. 11, November, 1565-1578, 2011

[14] S. Y. Lee, Structural Equation Modeling: A Bayesian approach, John Wiley Sons, 2007

[15] Q. Li, R. Xi, and N. Lin, “Bayesian regularized quantile regression,” Bayesian Analysis, Vol. 5, No. 3, pp. 533-556, 2010

[16] T. Park, and G. Casella, "The Bayesian Lasso," Journal of the American Statistical Association, Vol.103, pp.681–686, 2008

[17] L. Sik-Yum, and S. Jian-Qing "Bayesian Analysis of Structural Equation Model With Fixed Covariates, Structural Equation Modeling," A Multidisciplinary Journal, Vol. 7, No.3, pp. 411-430, Lawrence Erlbaum Associates, Inc, 2000

[18] X.Y. Song, and S. Y. Lee, Basic and advanced Bayesian structural equation modeling: With applications in the medical and behavioral sciences, John Wiley Sons, 2012

[19] Q. Tang, H. Zhang, and S. Gong, "Bayesian regularized quantile regression analysis based on asymmetric Laplace distribution," Journal of Applied Mathematics and Physics, Vol. 8, No. 1, pp. 70-84. 2020

[20] Y. Wang, X.N. Feng, and X.Y. Song, "Bayesian Quantile Structural Equation Models", Structural Equation Modeling A Multidisciplinary Journal, Vol. 23, No. 2, pp. 1-13. 2015

[21] Z. Yanqing, and T. Niansheng, "Bayesian Empirical Likelihood Estimation of Quantile Structural Equation Models," J Syst Sci Complex, 30, 122-138, 2017

[22] F. Yanuar, "The Estimation Process in Bayesian Structural Equation Modeling Approach," Journal of Physics: Conference Series 495 (2014) 012047, p.p. 1-10, 2014

[23] K. Yu, and R.A. Moyeed, "Bayesian Quantile regression," Statistical & Probability Letters, 54, 437-447, 2001

مجلة كلية الرافدين الجامعة للعلوم (2023)؛ العدد 54؛ 472 - 489

PISSN: (1681-6870); EISSN: (2790-2293)
مجلة كلية الرافدين الجامعة للعلوم
Available online at: https://www.jrucs.iq

JRUCS
Journal of AL-Rafidain University College for Sciences

AL- Rafidain University College

تقنية الجسر البيزي لأنموذج المعادلة الهيكلية التجزيئي مع التطبيق

لقاء علي محمد	بلسم مصطفى شفيق
lekaa.ali.1968@gmail.com	balsammustafa95@mtu.edu.iq
قسم الإحصاء - كلية الإدارة والاقتصاد - جامعة بغداد، بغداد، العراق	قسم تقنيات المعلوماتية، الكلية التقنية الادارية، الجامعة التقنية الوسطى، بغداد، العراق

المستخلص

يعرف نموذج المعادلات الهيكلية (SEM) على نطاق واسع بأنه أهم أداة إحصائية لتقييم العلاقات المتبادلة بين المتغيرات الكامنة وهو أحد نماذج المتغيرات الكامنة. وكتقدم حديث ، يوفر نموذج المعادلات الهيكلية ضمن الانحدار الكمي البايزي تقييمًا كميًا شاملاً للمتغيرات الشرطية-الكامنة في ضوء كل من المتغيرات التفسيرية والكامنة. في هذه الدراسة ، تم اقتراح تقنية جسر البيزية في الانحدار الكمي ، ونقارنها مع عامل بايزي الأقل انكماشًا مطلقًا واختيار العامل (Lasso) ونقوم بإجراء تقدير متزامن واختيار متغير في سياق نموذج المعادلات الهيكلية (SEM) الكمي. وتم استخدام عينات جيبس لإجراء الاستدلال بايزي. تظهر عمليات المحاكاة بأحجام عينات مختلفة أن الطريقة المقترحة تعطي نتائج جيدة. تم تطبيق الطريقة المقترحة على مجموعة من مرضى الفشل الكلوي لدراسة العوامل المؤثرة على هذا المرض.

معلومات البحث

تواريخ البحث:
تاريخ تقديم البحث: 2022/12/19
تاريخ قبول البحث: 2023/3/3
تاريخ رفع البحث على الموقع: 2023/12/31

الكلمات المفتاحية:
نماذج المعادلات الهيكلية، الانحدار التجزيئي، المتغيرات الكامنة، لاسو البيزي، تقنية الجسر

للمراسلة:
لقاء علي محمد
lekaa.ali.1968@gmail.com

https://doi.org/10.55562/jrucs.v54i1.615