CENTRAL ASIAN JOURNAL OF MEDICAL AND NATURAL SCIENCES
https://cajmns.centralasianstudies.org/index.php/CAJMNS
Volume: 05 Issue: 04 | October 2024 ISSN: 2660-4159

Article

Applying the Tobit Quantile Regression Model to Improve the Level of Education in Secondary Schools

Hassanien Adel Salih[1], Haitham Hassoon Majid[2], Maitham Nori Muhsen[3]

1. Lecturer, General Directorate of Education in Najaf Governorate Najaf, Iraq
* Correspondence: hasaneen.adel24@qu.edu.iq
2. Lecturer, General Directorate of Education in Najaf Governorate Najaf, Iraq
* Correspondence: haithamnkhozaee@gmail.com
3. Assistant Lecturer, General Directorate of Education in Najaf Governorate Najaf, Iraq
* Correspondence: maitham.nori@gmail.com

Abstract: This study aimed to raise secondary school education levels in the Iraqi governorate of Najaf by utilizing the Tobit Quantile Regression model. The shortcomings of response variables are frequently overlooked by traditional regression models, which calls for the use of sophisticated techniques like the Bayesian hierarchical model in conjunction with the Lasso methodology. By identifying important variables influencing schooling, this study fills in the gaps in conventional estimation techniques. Data on characteristics including study hours and parental education were provided by a random sample of one hundred pupils. To improve parameter estimate accuracy, the modified Lasso approach with Gibbs Sampler was applied using the R program. The results highlight the important influence of family and economic circumstances on student achievement and point to the need for focused interventions to improve educational outcomes.

Keywords: Lasso technique, Tobit Quantile Regression model, Hierarchical Bayesian model, Gibbs Sampler algorithm

Citation: Salih, H. A., Majid, H. H., & Muhsen, M. N. Applying the Tobit Quantile Regression Model to Improve the Level of Education in Secondary Schools. Central Asian Journal of Medical and Natural Science 2024, 5(4), 117-129.

Received: 3rd Jul 2024
Revised: 10th Jul 2024
Accepted: 17th Jul 2024
Published: 24th Jul 2024

1. Introduction

Regression models are among the most critical probabilistic models applied in finding scientific explanations and predicting the future for many medical, social, educational, and other phenomena. The quantitative regression model is one of the regression models that has proven its effectiveness in explaining the studied phenomena regardless of the distribution of its data, in addition to this regression's advantages, such as depth. Analysis, interpretation, and immunity. The Tobit Quantile Regression model is one of the forms of quantitative regression applied to observational data.

The Lasso technique was introduced [1] and is considered a vital method for estimating regression model parameters. The Lasso technique is essential in controlling the variance of the model parameters and identifying the critical variables included in the model. The Bayesian method of the Lasso technique is also necessary for finding accurate estimates for the Tobit Quantile Regression model.

[2] applied the adaptive lasso technique to estimate the parameters of the Tobit Quantile Regression model and the initial g-pior function to estimate the Tobit Quantile Regression model. The researcher [3] proposed the Laplace distribution independently for the model parameters, except for the β_0 parameters.

In this research, the Bayesian hierarchical method was applied using the Lasso technique modified by Tobit Quantile Regression to improve the level of education in secondary schools.

2. Materials and Methods

The current research aims to identify the most critical factors affecting raising the level of education in secondary schools by applying the modified Lasso technique in the hierarchical Bayesian Tobit Quantile Regression model.

The theoretical side:

Bayesian Hierarchical Model [4]:

The hierarchical Bayesian method is one of the essential and appropriate methods in analyzing many research studies with multiple explanatory variables structurally linked to each other. Researchers sometimes need to gradually formulate the model under study, leading us to the hierarchical Bayesian model in dealing with this method.

The hierarchical Bayesian method is called the multilevel Bayesian method, which is a statistical method that includes a modelling strategy and is characterized by its analysis of elementary functions and hyperparameters, which indicates that the data in the hierarchical models depend on a set of parameters that are at the first level, and an initial distribution is assumed for them, which has other parameters called parameters. The second level: Initial distributions are considered according to the requirements of the study. Therefore, many researchers resort to applying this method because of its importance in dealing with data of a hierarchical nature and obtaining accurate results in its analysis.

On the other hand, the hierarchical Bayesian method is characterized by the possibility of accessing the most significant possible number of parameters, obtaining more accurate predictions than the usual Bayesian method and avoiding the problem of (overfitting). The initial joint distribution can be written as follows:

$$f_3(\theta_1,\theta) = f_1(\theta|\theta_1)f_2(\theta_1) \dots\dots\dots\dots\dots\dots\dots (1)$$

The subsequent distribution can be expressed as follows:

$$\mathrm{p}(\theta_1,\theta|y) \propto f_3(\theta_1,\theta)f(y|\theta_1,\theta) = f_3(\theta_1,\theta)f(y|\theta) \dots\dots\dots\dots (2)$$

The marginal posterior distribution of the hyperparameter θ_1 is as follows:

$$p_1(\theta_1|y) = \frac{p(\theta,\theta_1|y)}{p(\theta|\theta_1,y)} \dots\dots\dots\dots (3)$$

And represents $(p(\theta|\theta_1,y)$ (Conditional subsequent function
(θ $given\ the\ hyperparameter\ \theta_1$).

$$p_1(\theta_1|y) = \int p(\theta,\theta_1|y)d\theta \dots\dots\dots\dots\dots (4)$$

Tobit Quantile Regression Model

Researchers often use regression analysis in research and studies. Still, they encounter some phenomena in which the response variable is of a limited type, and the observations of the response variable increase or decrease from a certain point, and this increase and decrease is unknown to the researchers, so the data for the response variable is recorded.

The Tobit model is one of the essential statistical models developed by researcher Tobin in 1958. This model describes the relationship between the response variable (y_i) and the explanatory variables (X_i) (Al-Bajourdi, 2016: 226). Sometimes, the response variable in a model (Tobit Quantile Regression) is restricted by upper or lower limits.

The researcher presented the Tobit Quantile Regression model for the first time [5]. The Tobit Quantile Regression model emerged as an expansion of the Tobit model. This Tobit Quantile Regression model was used to analyze the relationship between the response and explanatory variables in the entire distribution. Conditional for the response variable. Through this, (Tobit Quantile Regression) is an extension of the Quantile Regression model in obtaining the relationship between the response variable and the explanatory variables if the response variable is of a limited type.

The Tobit Quantile Regression model has added evident scientific importance in many modern applications, as it has proven its efficiency in dealing with survival data and its comparison with the Cox Proportional Hazord model [6].

The Tobit Quantile Regression model can be estimated with the following formula [6]:

$$y_i = \max\{y^0, y_i^*\} \quad \dots\dots\dots\dots\dots (5)$$

$$y_i^* = X_i^{\prime}\beta_p + \varepsilon_i$$

Whereas:

β_p : Parameter vector based on P. (For ease, the vector of parameters (β_p)It will be expressed by symbol (β) And the ratio (1<p<0) corresponding to the division value (q_p)

y^0 : Known observation point. (The point will equal zero, which is called the standard Tobit Quantile Regression model.)

ε_i : Represents the residue.

$$\int_{-\infty}^{0} f_p(\varepsilon_i)d\varepsilon_i = p \quad \dots\dots\dots\dots (6)$$

Or $F_{\varepsilon}^{-1}(P) = 0$

It is worth noting that the usual method of estimating the parameters of the Tobit Quantile Regression model is not based on the standard analysis assumptions used to calculate the ordinary regression model. The usual method of estimating the parameters of the Tobit Quantile Regression model can be written as follows:

$$\hat{\beta} = \arg\min \sum_{i=1}^{n} \rho_p\left(y_i - max\{y^0, x_i^{\prime}\beta\}\right) \dots\dots\dots\dots\dots (7)$$

ρ_p *: It will be defined later as in equations (9 and 10)*

$\hat{\beta}$: Represents a vector of estimated model parameters.

Powell (1986: 143-144) [5] asserts that the estimators of the Tobit Quantile Regression method have the property of consistency and are close to a normal distribution. Despite the critical properties of the Tobit Quantile Regression method estimators, research and studies have proven the usual method for estimating.

The Tobit Quantile Regression model may face the problem of convexity, and the Bayesian method has the advantage of overcoming this problem [6].

As indicated by [5], if the response variable is significantly limited, the usual method is used to estimate the model parameters in the upper sections of the conditional distribution of the response variable (P>0.5) due to the lack of information available in the lower sections in this case.

On the contrary, the Bayesian method for estimating the Tobit Quantile Regression model can deal with the minimal response variable data efficiently. The hierarchical Bayesian method is also important in completing the estimation process by including control points through the possibility of estimating the penalty parameters with the rest of the model parameters at one time, in addition to the possibility of estimating the standard deviation in cases where we are unable to find it using traditional methods [7].

Tobit Quantile Regression Model Estimation Method [11]

The Tobit Quantile Regression model is one of the regression models with a limited response variable and is an extension of the Tobit model [12]. The Tobit Quantile Regression model is also one of the important forms of Quantile Regression models, and it is characterized by its possibility of investigating the relationship between the response variable and the independent explanatory variables in the conditional distribution. These models are of great importance in not being affected by outliers and the problem of heterogeneity of variance, and they also do not need to impose a normal distribution in analyzing regression models.

1. The structure of the Tobit Quantile Regression model [11]:

Assuming that y^0& y^* Random variables are related to the following relationship (Alhamzawi, 2013) :

$y = max\{y^0, y^*\}$ … … … … … (8)

y^0 : Known observation point.

Suppose that $y = (y_1, y_2, \dots, y_n)$ Sample of independent observations associated with explanatory variables $X = (X_1, X_2, \dots, X_n)$ Since X_i It is destined for k of the explanatory variables.

can expressy^* About the following mathematical equation [8]:

$y_i^* = \acute{x}_\iota \beta_p + \varepsilon_\iota$ … … … … (9)

ε_i: Represents the residue.

β_p : Vector of unknown parameters. For ease, it will be expressed as: (β)

The classic method of estimating Tobit Quantile Regression parameters is given by the following equation [2]:

$$min\beta = \sum_{i=1}^{n} \rho_p(y_i - max\{y^0, \acute{x}_i\beta\}) \text{ …. …. …. …. (10)}$$

can be expressed ρ_p In the following format [8]:

$$\rho_p(u) = \begin{cases} pu & if\ u \geq 0 \\ -(1-p)u & if\ u < 0 \end{cases} \text{ …. …. … …. … (11)}$$

The Bayesian method in quantile regression models was used for the first time by [9], assuming that the random error in the linear model of quantile regression follows a skewed Laplace distribution. The researchers [10] also used the Bayesian method. In estimating the Tobit Quantile Regression model by assuming the twisted Laplace (AL) distribution for the random error in the model, there are difficulties resulting from the analysis of subsequent conditional distributions, as the researchers [11] used mixed representation in the twisted Laplace distribution by using Gibbs Sampler algorithm, which facilitated the easy sampling of subsequent conditional distributions. This analysis became used in research on the Bayesian method in quantile regression.

The Latin random variable y^* Equation (2) uses the mixed representation of the error distribution, which can be given by the following equation [13].

$$y_i^* = \acute{x}_i\beta + \xi_1 e_i + T^{-1/2}\xi_2\sqrt{e_i}z_i \dots\dots\dots\dots\dots\dots\dots (12)$$

Where $\xi_1 = \frac{1-2p}{p(1-p)}$, $\xi_2 = \frac{2}{p(1-p)}$

$z_i \sim N(0,1)$, $e_i \sim \exp\left(\frac{1}{\tau}\right)$, $\tau > 0$

2. Modified Lasso technique in the Bayesian Tobit Quantile Regression model:

The Lasso technique in estimating regression model parameters can be applied using the Bayesian method, assuming that the initial distribution of the model parameters is a Laplace distribution [14] and that the probability density function of the Laplace distribution is as follows [15]:

$$f(t) = \frac{\lambda}{2}\exp(-\lambda\,|t|) \dots\dots\dots\dots\dots\dots (13)$$

Where $-\infty < t < \infty$, and $\lambda > 0$ is scale parameter , $\theta = 0$ is location parameter

The researchers [16] also used the initial function for the parameters of the Linear Quantile Regression model with the following equation:

$$f(\beta_k) = \frac{\eta}{2}\exp(-\eta|\,\beta_k|) \dots\dots\dots\dots\dots (14)$$

Where $\eta = \tau\lambda$

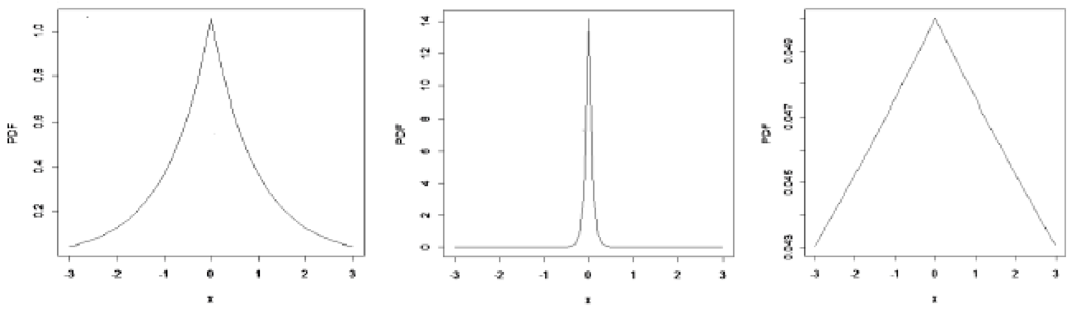
The initial function for each parameter of the Tobit Quantile Regression model was proposed independently as follows:

$$f(\beta_K) = \frac{\eta k}{2\sqrt{c}}\exp\left(\frac{\eta k}{\sqrt{c}}|\beta_K|\right) \dots\dots\dots\dots\dots\dots\dots\dots (15)$$

Since C is a known, pre-calculated value:

$C = \frac{1}{p(1-p)}$, $0 < p < 1$, $\frac{\eta_k}{c} > 0$

1. If the integral of the proposed function in equation (8) is equal to the correct one.
2. The proposed function (8) is always positive.
3. Figure (1) represents the proposed function at different values of (c,η).



c=.9 , η=1   c=0.5 , η=10   c=0.01 , η=0.005

Figure 1. The proposed function at different values of (c,η)

The researchers [17] demonstrated that each ($d \geq 0$) equation is correct since d is a random variable. Note [15]:

$$\frac{d}{2}\exp(-d|S|) = \int_0^\infty \frac{1}{\sqrt{2\pi\alpha}} \exp\left(\frac{s^2}{2\alpha}\right)\frac{d^2}{2}\exp\left(\frac{-d^2}{2}\alpha\right)d\alpha \quad \text{....}(16)$$

Therefore, the initial function of equation (8) can be formulated as follows:

$$f(\beta_k) = \int_0^\infty \frac{1}{\sqrt{2\pi\alpha_k}} exp\left\{-\frac{\beta_k^2}{2\alpha_k}\right\}\frac{\eta_k^2}{2c}\exp\left(\frac{\eta_k^2}{2c}\alpha_k\right)d\alpha_k \ \dots\dots\dots(17)$$

3. MCMC algorithm [14]:

It is called the (MCMC) algorithm, an abbreviation for Markov Chain Monte Carlo, and it is one of the algorithms used in finding complex integrals. This algorithm can examine the subsequent distributions of hierarchical Bayesian models. This algorithm is also one of the most famous algorithms used in research and statistical studies, in addition to other algorithms. (Metropolis Hasting & Gibbs sampling).

In the (MCMC) algorithm, the sampling is performed through the proposed distribution instead of the posterior function if the posterior distribution of the parameter is unknown, so the (MCMC) algorithm is an alternative method for conducting sampling of the proposed distribution, which is very similar to the posterior distribution of the function.

As for the Gibbs Sampler algorithm, this algorithm provides more accessible and faster ways to estimate model parameters than other algorithms, such as the Metropolis Hasting algorithm. The Gibbs Sampler algorithm is also easy to work as it assumes the availability of conditional distributions for the variables during statistical analysis.

Proposed hierarchical Bayesian model

The hierarchical Bayesian model proposed by the researcher is based on the Bayesian hierarchical model presented by the researchers [16] used in estimating the parameters of the linear quantile regression model.

Assuming the following:

$e = (e_1, e_2, \dots, e_n)'$, $\quad z = (z_1, z_2, \dots, z_n)$, $\quad \alpha = (\alpha_1, \alpha_2, \dots, \alpha_p)'$

$y_i = max\{y^0, y_i^*\}$

$$y_i^* = \acute{x}_i\beta + \xi_1 e_i + \tau^{-1/2}\xi_2\sqrt{e_i}z_i$$

$$\frac{e}{\tau} \sim \prod_{i=1}^{n} \tau \exp\left(-\tau e_i\right)$$

$$Z \sim \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_i^2\right)$$

$$\tau \sim \tau^{-1} \qquad \text{........(18)}$$

$$(\beta_k \backslash \alpha_k) \overset{ind}{\sim} \frac{1}{\sqrt{2\pi\alpha_k}} \; exp\left\{-\frac{\beta_k^2}{2\alpha_k}\right\}$$

$$\left(\alpha_k \backslash \eta_k^2\right) \overset{ind}{\sim} \frac{\eta_k^2}{2c} \; exp\left(-\frac{\eta_k^2}{2c}\alpha_k\right)$$

$$\eta_k^2 \overset{ind}{\sim} \frac{d^c}{\Gamma(c)} \; (\eta_k^2)^{c-1} \exp(-d\,\eta_k^2)$$

Assuming that $C = d = 0.1$ As in the model proposed by researchers [16].

Complete posterior conditional distributions of the variables in equation (11) [8]:

1. Complete conditional distribution of the random variable y^***:**

$$f\left(y^* \backslash X, e, \beta, \alpha, \tau, \eta_k^2\right) = \prod_{i=1}^{n} \frac{\tau^{1/2}}{\sqrt{2\pi\xi_2^2 e_i}} exp\{-\tau(y^* - \acute{x}_i\beta - \xi_i e_i)^2 \backslash 2\xi_2^2 e_i\} \ldots.(19)$$

2. Complete conditional distribution of the random variable. e_i**:**

$$f\left(e_i \backslash X, y^*, e_{-i}, \beta, \alpha, \tau, \eta_k^2\right) \propto f\left(y^* \backslash X, e, \beta, \alpha, \tau, \eta_k^2\right) f\left(\frac{e_i}{\tau}\right)$$
$$\propto e_i^{-\frac{1}{2}} exp\left\{-\tau\left(y^* - \acute{x}_i\beta - \xi_1 e_i\right)^2 \backslash 2\xi_2^2 e_i\right\} \exp(-\tau e_i) \ldots(20)$$

$$f\left(e_i \backslash X, y^*, e_{-i}, \beta, \alpha, \tau, \eta_k^2\right) \propto e_i^{-1/2} exp\left\{-\frac{1}{2}\left[\left(\tau\xi_1^2\xi_2^{-2} + 2\tau\right)e_i\right] + \xi_2^{-2}\tau(y_i^* - \acute{x}_i\beta)^2 e_i^{-1}\right\}$$

Note that e_{-i} The exclusion represents the random variable (e) e_i**.**

3. Complete conditional distribution of the random variable.β_k**:**

$$f\left(\beta_k \backslash X, y^*, e, \beta_{-k}, \alpha, \tau, \eta_k^2\right) \propto f\left(y^* \backslash X, e, \beta, \alpha, \tau, \eta_k^2\right) f\left(\frac{\beta_k}{\alpha_k}\right)$$

Note thatβ_{-k} Represents the vector β After exclusionβ_k**.**

$$\propto \exp\left\{-\tau \sum_{i=1}^{n} \left(y_i^* - \acute{x}_i\beta - \xi_1 e_i\right)^2 \backslash 2\xi_2^2 e_i\, exp\left\{-\frac{\beta_k^2}{2\alpha_k}\right\}\right.$$

$$\propto exp\left\{-\frac{1}{2}\left[\tau\xi_2^{-2}\sum_{i=1}^{n} x_{ik}^2 e_i^{-1} + 1/\alpha_k)\beta_k^2\right.\right.$$

$$\left.\left. - 2\sum_{i=1}^{n}\left(\tau w_{ik} x_{ik}\xi_2^{-2} e_i^{-1}\right)\beta_k\right]\right\} \dots\dots(21)$$

$$w_{ik} = y_i^* - \xi_1 e_i - \sum_{j=1, j\neq k}^{p} x_{ij}\beta_j$$

Complete conditional distribution β_k The following equations give the normal mean and variance:

$$Me_k = V_k^2 \xi_2^{-2}\tau \sum_{i=1}^{n} w_{ik}(x_{ik}\backslash e_i)$$

$$V_k^2 = \frac{1}{\xi_2^{-2}\tau\sum_i^n (x_{ik}^2 e_i^{-1}) + \alpha_k^{-1}}$$

4. The complete conditional distribution of the random variable τ**:**

$$f\left(\tau\backslash X, y^*, e, r, \beta_k, \alpha_1, \theta_1, \delta_2, \tilde{\delta}_1\right) \propto f\left(y^*\backslash X, e, \beta, r, \tau, \delta_2, \tilde{\delta}_1\right) f(e_\tau\backslash\tau) f(\tau)$$

$$\propto \tau^{\frac{n}{2}} exp\left\{-\tau\sum_{i=1}^{n}\left(y_i^* - \acute{x}_i\beta - \xi_1 e_i\right)^2\right.$$

$$\left.\backslash 2\xi_2^2 e_i\right\}\tau^n \exp(-\tau\sum\nolimits_{i=1}^{n} e_i)\tau^{-1} \dots\dots\dots.(22)$$

$$\propto \tau^{(3n\backslash 2)-1} exp\left\{-\tau\sum_{i=1}^{n}\left[(y_i^* - \acute{x}_i\beta - \xi_1 e_i)^2\backslash 2\xi_2^2 e_i\right] + e_i)\right\}$$

So, the complete conditional distribution of τ (Gamma distribution).

5. Complete conditional distribution of the random variable.α_k**:**

$$f\left(\alpha_k\backslash X, y^*, e, \alpha_{-k}, \beta, \tau, \eta_k^2\right) \propto f(\beta_k\backslash\alpha_k) f\left(\alpha_k\backslash\eta_k^2\right)$$

$$\propto \frac{1}{\sqrt{2\pi\alpha_k}} exp\left\{-\frac{\beta_k^2}{2\alpha_k}\right\}\exp\left(-\frac{\eta_k^2}{2c}\alpha_k\right)\dots\dots(23)$$

$$f\left(\alpha_k\backslash X, y^*, e, \alpha_{-k}, \beta, \tau, \eta_k^2\right) \propto \frac{1}{\sqrt{\alpha_k}} exp\left\{-\frac{1}{2}\left[\frac{\alpha_k\eta_k^2}{c} + \frac{\beta_k^2}{\alpha_k}\right]\right\}$$

So, the conditional distribution of the random variable α_k Be **(**Generalized inverse Gaussian).

6. Complete conditional distribution of the random variable.η_k^2**:**

$$f\left(\eta_k^2\backslash X, y^*, e, \alpha_{-k}, \beta, \tau\right) \propto f\left(\eta_k^2\right) f\left(\alpha_k\backslash\eta_k^2\right)$$

$$f\left(\eta_k^2\backslash X, y^*, e, \alpha_{-k}, \beta, \tau\right) \propto \frac{d^c}{\Gamma(c)}\left(\eta_k^2\right)^{c-1}\exp\left(-d\,\eta_k^2\right)\frac{\eta_k^2}{2c}\exp\left(-\frac{\eta_k^2}{2c}\alpha_k\right)\dots.(24)$$

$$f\left(\eta_k^2\backslash X, y^*, e, \alpha_{-k}, \beta, \tau\right) \propto \left(\eta_k^2\right)^{c-1}\exp\left(-d\,\eta_k^2\right)\eta_k^2 exp\left(-\frac{\eta_k^2}{2c}\alpha_k\right)$$

$$f\left(\eta_k^2\backslash X, y^*, e, \alpha_{-k}, \beta, \tau\right) \propto (\eta_k^2)\exp\left[-\eta_k^2(d \pm \frac{\alpha_k}{2c})\right]$$

So, the conditional distribution of the random variable η_k^2 It is distribution **(**Gamma**)**.

The practical aspect

The Tobit Quantile Regression Model and the modified Lasso method were applied to improve the level of education in secondary schools in the Najaf Governorate. A random sample of (100) male and female students in the sixth scientific grade was drawn for the academic year 2022/2023. The student's average represented the dependent variable. (y) The independent variables representing the student's grade point average in the fifth academic grade, the level of education of the parents, the number of hours of study, and the number of rooms in the house relative to family members are independent explanatory variables **(**x_1, x_2, x_3, x_4**)**.

The proposed method was used in the research to analyze the relationship between the level of education in secondary schools and a set of explanatory variables that were diagnosed with the help of the specialized supervision department and a group of experienced professors to build a model that explains the relationship between the variables and obtain predictions close to reality.

The sample and its division according to Quantile Regression

The researchers used accurate data from a sample from secondary schools affiliated with the General Directorate of Education in Najaf Governorate. They analyzed the relationship between the educational level of students and the variables that would affect the students.

Data were collected for (100) students in the sixth grade of middle school. The student's families took the data, and the school administrations supervising the students matched the data with each student's school cards on one hand. On the other hand, the data was divided into four divisions within quantitative regression (Quantile Regression) and the conditional distribution of the response variable in the statistical analysis of the relationship (0.25, 0.50, 0.75, 0.95).

Explanatory variables

To study the factors and indicators affecting improving and raising the educational level of secondary school students, it is important to identify the factors surrounding the student capable of raising or lowering his level of study. From the above, independent explanatory variables related to the educational level of secondary school students were identified based on visiting departments. The Directorate of Education has relevant stakeholders, such as the Department of Specialist Supervision, Examinations, and Educational Planning, with expertise and specialization. The factors and explanatory variables that affect the educational level of this important segment have been identified. Four essential variables have been identified as explanatory variables as follows:

Four basic variables were identified as explanatory variables as follows:

X_1 : Student average in the fifth grade of middle school.

X_2 : Parents' education level.

X_3 : Number of student hours of study.

X_4 : The ratio of the number of rooms to the number of family members.

Standards used

The researchers used the criteria of mean error (MAE) and mean square error (MSE) to determine the efficiency of the method applied in the study.

Data check

The researchers used the (Farrar-Gluber test) and the (Golfeld-Quandt test) to examine the data and ensure that there were no problems of multicollinearity or heterogeneity of variance in the sample under study.

1. Testing the multicollinearity problem:

The following hypothesis was tested:

$H_0 : X_j is\ orth. (orthogonal)$

$H_1 : X_j is\ not\ orth. (non - orthogonal)$

The following test statistics were calculated:

$$x_0^2 = -\left\{n - 1 - \frac{1}{6}(2k + 5)\right\} \ln|D| \quad \ldots..(25)$$

$|D| = 0.1186382$

$x_0^2 = 290.0631 > x_{(0.05)}^2 = 33.67$

Therefore, we reject the null hypothesis and conclude that there is a problem of multicollinearity in the data.

2. Testing the problem of heterogeneity of variance:

$H_0 : \sigma_{ui}^2 = \sigma_u^2$

$H_1 : \sigma_{ui}^2 \neq \sigma_u^2$

$n_1 = n_2 = 64$

$S_{e1}^2 = 0.579, \qquad S_{e2}^2 = 0.177$

$$\frac{S_{e2}^2}{S_{e1}^2} = 0.20207 < F_{(0.05)}(1.55793)$$

Therefore, we accept the null hypothesis and conclude that there is no problem of heterogeneity of variance in the data.

3. Results and Discussion

The Tobit Quantile Regression model was applied through the modified method in the Lasso technique, and the parameters of the proposed model were estimated using the (R) program to explain the relationship between the educational level in the secondary stage and the effects represented by the random variables. Note Table 1 shows estimated parameter values for the proposed Tobit Quantile Regression model and the partitioned (p-value) in addition to the mean absolute errors (MAE) and the mean square random error (MSE).

Table 1. The values of the estimated parameters of the (Tobit Quantile Regression) model proposed by the researcher and the divided (p) value in addition to the mean absolute errors (MAE) and the mean square random error (MSE)

Parameter estimation	**BML 0.25**	**BML 0. 5**	**BML 0.75**	**BML 0.95**
Intercept	0.264	0.328	0.164	0.228
The student's average in the previous stage	0.8015	0.7732	0.9015	0.8432
Parents' education level	0.7948	0.8806	0.7548	0.9106
Number of study hours	0.8888	0.7587	0.7788	0.6987
The number of rooms to the number of families	0.8116	0.9099	0.8116	0.8299
Method	MAE			
	P=0.25	P=0.5	P=0.75	P=0.95
BML	0.4017	0.5902	0.3806	0.6598
Method	MSE			
	P=0.25	P=0.5	P=0.75	P=0.95
BML	0.5817	0.7707	0.8006	1.7498

Source: Prepared by the researcher based on student data.

4. Conclusion and Recommendation

Conclusions:

The researchers concluded the following:

1. Economic factors for students and parents are among the most important factors affecting the student's level of education. Housing and the presence of a separate room for the student's study had a major role in raising the student's level of education.
2. Constant pressure from parents to determine the student's future, regardless of his aspirations and interests.
3. The Tobit Quantile Regression model was applied through the modified method in the Lasso technique, and the parameters of the proposed model were estimated using the R program to explain the relationship between educational level in the secondary stage and the actual effects represented by random variables.
4. The hierarchical Bayesian model method using the Lasso technique modified by estimating the variables in the proposed Tobit Quantile Regression model has proven to be of great importance in interpreting and studying the improvement of the educational level of secondary school students.

Recommendations:

In light of the above, some recommendations can be made that would contribute to improving the educational level of secondary school students:

1. The economic problem is one of the most important problems surrounding the student, so the state must allocate cash grants to secondary school students to pass this vital stage.
2. It is necessary to know the student's interests and not interfere with the parents in determining his future. It is also essential to dialogue and discuss with them to find out their desires and aspirations for the future, with the need to respect the

student's decisions and choices away from the previous failure of the parents and the desire to achieve the aspirations that the parents were not able to achieve.

3. Encouraging the student by the school and family throughout the learning process while praising him for the achievements he has achieved effectively so that he remains interested in studying, in addition to making the student realize that we believe in his ability to obtain positive self-esteem without being limited exclusively to exam results and academic grades only.
4. It is necessary to apply the hierarchical Bayesian model using the modified Lasso technique by estimating the variables in the Tobit Quantile Regression model, as the proposed method has proven to be of great importance in explaining the improvement of the educational level of secondary school students.

REFERENCES

[1] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267-288, 1996.

[2] R. Alhamzawi, "Tobit Quantile Regression with Adaptive Lasso Penalty," *The 4th International Scientific Conference of Arab Statistics*, 2013, pp. 450.

[3] R. Alhamzawi, "Bayesian Lasso-Mixed Quantile Regression," *Journal of Statistical Computation and Simulation*, 2014.

[4] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*, 3rd ed. A Chapman & Hall Book, 2014.

[5] J. Powell, "Censored Regression Quantiles," *Journal of Econometrics*, vol. 32, pp. 143-155, 1986.

[6] R. Alhamzawi, "Bayesian Elastic Net Tobit Quantile Regression," *Communication in Statistics-Simulation and Computation*, vol. 45, no. 7, pp. 2409-2427, 2016.

[7] H. Mallick and N. Yi, "A New Bayesian Lasso," *Statistics and Its Interface*, vol. 7, pp. 571-582, 2014.

[8] R. Alhamzawi, "Inference with Three-Level Prior Distributions in Quantile Regression Problems," *Journal of Applied Statistics*, vol. 44, no. 11, pp. 1947-1959, 2017.

[9] K. Yu and R. A. Moyeed, "Bayesian Quantile Regression," *Statistics & Probability Letters*, vol. 54, pp. 437-447, 2001.

[10] K. Yu and J. Stander, "Bayesian Analysis of a Tobit Quantile Regression Model," *Journal of Econometrics*, vol. 137, no. 1, pp. 260-276, 2007.

[11] H. Kozumi and G. Kobayashi, "Gibbs Sampling Methods for Bayesian Quantile Regression," *Journal of Statistical Computation and Simulation*, vol. 81, pp. 1565-1578, 2011.

[12] J. Tobin, "Estimation of Relationships for Limited Dependent Variables," *Econometrica*, vol. 26, no. 1, pp. 24-36, 1958.

[13] W. A. H. Nouri, H. A. Al-Bayati, and S. R. Al-Ani, "Mathematical Statistics," Baghdad, 1981.

[14] A. E. Gelfand and A. F. M. Smith, "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, vol. 85, pp. 398-409, 1990.

[15] F. Alhusseini, "Using the Bayesian Technique to the Tobit Quantile Regression with Scale Mixture Uniform for Selecting Important Variables Affecting Iraqi Investment Banking," *Journal of Quantitative Methods*, vol. 12, no. 1, 2017.

[16] Q. Li, R. Xi, and N. Lin, "Bayesian Regularized Quantile Regression," *International Society for Bayesian Analysis*, vol. 5, pp. 533-556, 2010.

[17] 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, pp. 99-102, 1974.

[18] S. H. Al-Bagouri, "The Problem of Statistics in Sub-Saharan Africa and Analyzing Its Impact on the Effectiveness of Government Performance Using the Tobit Model," *Journal of Economic and Financial Research*, vol. 5, 2016.

[19] F. Alhusseini, "Tobit Quantile Regression and Iraqi Banks' Profit," *Young Economists Journal / Revista Tinerilor Economisti*, vol. 13, no. 27, pp. 141-152, 2016.

[20] R. Koenker and V. D'Orey, "Algorithm AS 229: Computing Regression Quantiles," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 36, pp. 383-393, 1987.