

Bayesian Lasso Tobit regression

Haider Kadhim Abbas

Department of Statistics, College of Administration and Economics, University of
Al-Qadisiyah, Iraq.

haiderhilali@gmail.com

Received : 23/12/2018

Revised : \ \

Accepted : 23/1/2019

Available online : 16 /4/2019

Abstract: In the present research, we have proposed a new approach for model selection in Tobit regression. The new technique uses Bayesian Lasso in Tobit regression (BLTR). It has many features that give optimum estimation and variable selection property. Specifically, we introduced a new hierarchical model. Then, a new Gibbs sampler is introduced. We also extend the new approach by adding the ridge parameter inside the variance covariance matrix to avoid the singularity in the case of multicollinearity or in case the number of predictors greater than the number of observations. A comparison was made with other previous techniques applying the simulation examples and real data. It is worth mentioning, that the obtained results were promising and encouraging, giving better results compared to the previous methods.

Keywords: Bayesian Lasso, Ridge parameter, Variable selection (VS), Tobit regression.

1. Introduction:

We are aware of that regression analysis is one of the statistical procedures that illustrate the relationship between explanatory variables and the dependent variables (outcome). If the explanatory variables values are known, then the regression analysis assistance us to predict the values of the dependent outcome variable. In addition, in order to reach accurate results from the studied event, the selected model must correspond the available data as best as possible, and the proposed regression process for the data in question will lead to results that are close to real reality. For each type of data, there is a suitable regression process. For example, if the quantitative data for the dependent outcome variable and the assumptions of the form can be dealt with using the conventional regression procedure, and if binary data is available for the dependent outcome variable, it can be dealt with using the logistic regression procedure. But, if the observations are partially constrained with the dependent variable and not constrained in the other part, these data are called censored data. The application of the conventional regression with this type of data will lead to biased parameters on the one hand, and inconsistent on the other hand. Therefore, it is necessary to determine a regression process that is proportional to this data. So, the Tobit regression process is appropriate to those data, it is elucidating the relationship among the non-negative dependent outcome variable and the independent explanatory variables. The function of Tobit regression is a mixed function, it deals with two-part data, each part of the outcome variable data will take a given distribution. The dependent variable data that equal to zero will take the cumulative distribution function of the normal distribution, and the data is larger than zero will take probability density function distribution. The Tobit regression has been used in many studies, and the statisticians have applied this regression in many packages of different statistical language programs, such as

the Tobit function in AER package and MCMCtobit function in MCMCpack package.

However, when the independent variables are too large or low ratio the number of observations to the number of variables, then it is very difficult to distinguish the independent variables that are important and influential in describing the linear regression model, which leads to the instability of the model, thus the model lacks the validity of the prediction. To get rid of these problems, statisticians resorted to the mechanism of selecting the important and influential variables, while at the same time eliminating as much as possible from the explanatory variables that are not important. This mechanism called by variable selection process (VS).

One of these VS processes is Lasso technique (Tibshriani, 1996) for parameter estimation. Specifically, Tibshriani (1996) introduced Lasso technique in order to interpretability of regression models, and get better prediction accuracy. The aim of the Lasso regression is to obtain a subset of the estimations that reduces the prediction error of the outcome variable, by imposing a constraint on model parameters that cause shrank the unimportant explanatory variables and reduced to zero. Park and Casella (2008) explicated that the parameters of the Lasso regression can be estimated by the Bayesian pattern. Hans (2009) introduced a new aspect of the Bayesian treatment of Lasso regression, by using a news Gibbs sampler for Bayesian Lasso regression. And then, Mallick and Yi (2014) developed the previous methods by using a new method called a new Bayesian Lasso, this proposed procedure for VS and coefficient estimation in linear regression. The findings of Mallick in his research were very good and proved their efficiency from the previous Bayesian processes used. The good results notified in Mallick process encourage us to use the new Bayesian procedure in Tobit regression.

2. Bayesian Lasso Tobit regression:

The general formula of Tobit regression is

$$y_i = \begin{cases} y_i^* & y_i > 0 \\ 0 & \text{otherwise } \leq 0 \end{cases}$$

where y_i is a dependent outcome variable, and y_i^* is a latent variable. Then y_i^* is given by

$$y^* = X\beta + \epsilon \dots (1)$$

where y^* is $nx1$ vector of responses, X is the nxk matrix of regressors, and the predictors β is

$$\beta = (\beta_0, \beta_1, \dots, \beta_{k-1}),$$

ϵ is $nx1$ vector of error and distributed normal $(0, \sigma^2)$, and k is a number of independent variables, and n is the number of observations.

As we highlighted in introduction, the Lasso model are widely used as regularization process for coefficient estimation in regression problems. The Lasso regression formulated as following

$$\hat{\beta}_{lasso} = \underset{\beta}{\operatorname{argmin}} (\mathbf{y} - X\beta)'(\mathbf{y} - X\beta) + \lambda \sum_{j=1}^k |\beta_j| \quad \lambda \geq 0 \dots (2)$$

where λ is a positive penalty parameter. On the other hands, Bayesian inference naturally overcomes this problem by providing a valid measure of standard errors based on the Monte Carlo Markov chain (MCMC) outputs. Results from the Bayesian Lasso are superbly like those from the regular Lasso. Although even though a lot of computationally intensives, the Bayesian Lasso is very simple to execute and automatically provides interval estimates for all parameters, together with the variance error.

Following Andrews and Mallows (1974), Park and Casella (2008) represented the prior distribution of β as follows

$$\pi(\beta | \sigma^2, \lambda) = \prod_{j=1}^k \frac{\lambda}{2\sqrt{\sigma^2}} \exp\left\{-\frac{\lambda|\beta_j|}{\sqrt{\sigma^2}}\right\} =$$

$$\prod_{j=1}^k \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2}l_j} \exp\left(-\frac{\beta_j^2}{2\sigma^2 l_j}\right) \frac{\lambda^2}{2} \exp\left(-\frac{\lambda^2}{2} l_j\right) dl_j$$

Later, a new representation of the Laplace density given by Mallick and Yi (2014) as follows

$$\pi(\beta | \sigma^2, \lambda) = \prod_{j=1}^k \frac{\lambda}{2\sqrt{\sigma^2}} \exp\left\{-\frac{\lambda|\beta_j|}{\sqrt{\sigma^2}}\right\},$$

$$= \prod_{j=1}^k \frac{\lambda}{2\sqrt{\sigma^2}} \int_{u_j > \frac{|\beta_j|}{\sqrt{\sigma^2}}} \lambda e^{-\lambda u_j} du_j \dots (3)$$

They pointed out that the posterior distribution $\pi(\beta | \sigma^2, \lambda)$ is similar to the original Bayesian Lasso procedure of Park and Casella (2008) and the new representation has gorgeous properties. In this research, following Mallick and Yi (2014), we propose a new hierarchical representation of the Bayesian Lasso Tobit regression (BLTR).

2.1 Hierarchical form of BLTR:

Mallick and Yi (2014) presented a new Gibbs sampler by using the following representation:

$$\frac{\lambda}{2} e^{-\lambda|\beta_j|} = \int_{w > |\beta_j|} \frac{1}{2w} \frac{\lambda^2}{\Gamma(2)} w^{2-1} e^{-\lambda w} dw$$

In this research, we adopted the above formula as follows:

$$\text{Let } z_j = \lambda w_j \quad \Rightarrow \quad dz_j = \lambda dw_j \quad \text{then}$$

$$\frac{\lambda}{2} e^{-\lambda|\beta_j|} = \int_{z_j > \lambda|\beta_j|} \frac{\lambda}{2z_j} \frac{\lambda^2}{\Gamma(2)} \left(\frac{z_j}{\lambda}\right)^{2-1} e^{-z_j} \frac{1}{\lambda} dz_j$$

$$= \int_{z_j > \lambda |\beta_j|} \frac{\lambda}{2} e^{-z_j} dz_j \dots (4)$$

Then, the hierarchical model of BLTR is

$$y_i = \begin{cases} y_i^* & y_i > 0 \\ 0 & otherwise \leq 0 \end{cases}$$

$$\mathbf{y}^* | \mathbf{X}, \boldsymbol{\beta}, \sigma^2 \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2 I_n)$$

$$\beta_j | \lambda \sim \text{Uniform}(-\frac{1}{\lambda}, \frac{1}{\lambda}) \quad (5)$$

$$z_j \sim \text{Exp}(1)$$

$$\sigma^2 \sim \text{InvGamma}(a, b)$$

$$\lambda \sim \text{Gamma}(c, d)$$

where $\mathbf{z} = (z_1, \dots, z_k)$

2.2 Full Conditional Posterior Distributions of BLTR:

Firstly, With the properties of Bayesian, we can get a $\boldsymbol{\beta}$ posterior distribution as following

$$\pi(\boldsymbol{\beta} | \mathbf{y}^*, \mathbf{X}, \mathbf{z}) \propto \pi(\mathbf{y}^* | \mathbf{X}, \boldsymbol{\beta}, \sigma^2) \pi(\boldsymbol{\beta} | \mathbf{z})$$

$$\propto \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{y}^* - \mathbf{X}\boldsymbol{\beta})' (\mathbf{y}^* - \mathbf{X}\boldsymbol{\beta}) \right\} \prod_{j=1}^k I \left\{ |\beta_j| < \frac{z_j}{\lambda} \right\}$$

$$\propto \exp \left\{ -\frac{1}{2\sigma^2} (-2\mathbf{y}^{*'} \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}' \mathbf{X}' \mathbf{X}\boldsymbol{\beta}) \right\}$$

$$x \prod_{j=1}^k I \left\{ |\beta_j| < \frac{z_j}{\lambda} \right\}$$

$$\propto \exp \left\{ -\frac{1}{2\sigma^2} (-2\widehat{\boldsymbol{\beta}}' \mathbf{X}' \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}' \mathbf{X}' \mathbf{X}\boldsymbol{\beta}) \right\} \prod_{j=1}^k I \left\{ -\frac{z_j}{\lambda} < \beta_j < \frac{z_j}{\lambda} \right\}$$

$$\boldsymbol{\beta} | \mathbf{y}^*, \mathbf{X}, \mathbf{z} \sim N_k(\widehat{\boldsymbol{\beta}}_{OLS}, (\mathbf{X}' \mathbf{X})^{-1} \sigma^2)$$

$$x \prod_{j=1}^k I \left\{ -\frac{z_j}{\lambda} < \beta_j < \frac{z_j}{\lambda} \right\} \dots (6)$$

Then, the \mathbf{z} , σ^2 , and λ posterior distributions as following

$$\pi(\mathbf{z} | \boldsymbol{\beta}, \lambda) \propto \pi(\mathbf{z}) \pi(\boldsymbol{\beta} | \lambda)$$

$$\propto \prod_{j=1}^k e^{-z_j} I \{ z_j > \lambda |\beta_j| \}$$

$$\mathbf{z} | \boldsymbol{\beta}, \lambda \sim \prod_{j=1}^k \text{Exp}(1) I \{ z_j > \lambda |\beta_j| \} \dots (7)$$

$$\pi(\sigma^2 | \mathbf{y}^*, \mathbf{X}, \boldsymbol{\beta}) \propto \pi(\mathbf{y}^* | \mathbf{X}, \boldsymbol{\beta}, \sigma^2) \pi(\sigma^2)$$

$$\propto (\sigma^2)^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{y}^* - \mathbf{X}\boldsymbol{\beta})' (\mathbf{y}^* - \mathbf{X}\boldsymbol{\beta}) \right\} (\sigma^2)^{-a-1} \exp \left\{ -\frac{b}{\sigma^2} \right\}$$

$$\sigma^2 | \mathbf{y}^*, \mathbf{X}, \boldsymbol{\beta} \sim$$

$$\text{InvGamma} \left(\frac{n}{2} + a, \frac{1}{2} (\mathbf{y}^* - \mathbf{X}\boldsymbol{\beta})' (\mathbf{y}^* - \mathbf{X}\boldsymbol{\beta}) + b \right) \dots (8)$$

$$\pi(\lambda | \boldsymbol{\beta}) \propto \pi(\boldsymbol{\beta} | \lambda) \pi(\lambda)$$

$$\propto \lambda^k \lambda^{c-1} e^{-\lambda d} \prod_{j=1}^k I \left\{ \lambda < \frac{z_j}{|\beta_j|} \right\}$$

$$\propto \lambda^{(k+c)-1} \exp \{ -\lambda d \} \prod_{j=1}^k I \left\{ \lambda < \frac{z_j}{|\beta_j|} \right\}$$

$$\lambda | \boldsymbol{\beta} \propto \text{Gamma}((k+c), d)$$

$$x \prod_{j=1}^k I \left\{ \lambda < \frac{z_j}{|\beta_j|} \right\} \dots (9)$$

In practice, the above Gibbs sampler perform very well. However, the above estimator of $\widehat{\boldsymbol{\beta}}_{OLS}$ is highly unsteady in the existence of

multicollinearity. In addition, the matrix $X'X$ is singular if the number of explanatory variables k is more than the number of observations n . Gupta and Ibrahim (2007) have proposed a method to deal with these situations. Specifically, they added a ridge parameter inside the variance covariance matrix. So, we'll add a ridge parameter to remedy actual challenges that may appears with multicollinearity and overfitting problems.

Now, return to the previous equation (6) and add ridge parameter ϑ to the equation

$$\begin{aligned} & \exp \left\{ -\frac{1}{2\sigma^2} (-2\mathbf{y}^{*'}\mathbf{X}(\mathbf{X}'\mathbf{X} + \vartheta I_k)^{-1}(\mathbf{X}'\mathbf{X} \right. \\ & \quad \left. + \vartheta I_k)\boldsymbol{\beta} + \boldsymbol{\beta}'(\mathbf{X}'\mathbf{X} \right. \\ & \quad \left. + \vartheta I_k)\boldsymbol{\beta} \right\} \prod_{j=1}^k I \left\{ |\beta_j| < \frac{z_j}{\lambda} \right\} \\ & \propto \exp \left\{ -\frac{1}{2\sigma^2} (-2\boldsymbol{\beta}_R'(\mathbf{X}'\mathbf{X} + \vartheta I_k)\boldsymbol{\beta} + \boldsymbol{\beta}'(\mathbf{X}'\mathbf{X} + \right. \\ & \quad \left. \vartheta I_k)\boldsymbol{\beta} \right\} \prod_{j=1}^k I \left\{ -\frac{z_j}{\lambda} < \beta_j < \frac{z_j}{\lambda} \right\} \\ & \boldsymbol{\beta} | \mathbf{y}^*, \mathbf{X}, \mathbf{z} \sim N_k(\boldsymbol{\beta}_R, (\mathbf{X}'\mathbf{X} + \vartheta I_k)^{-1}\sigma^2) \\ & x \prod_{j=1}^k I \left\{ -\frac{z_j}{\lambda} < \beta_j < \frac{z_j}{\lambda} \right\} \dots (10) \end{aligned}$$

where $I(\cdot)$ denotes an indicator function.

2.3 Computation:

In the beginning, we specify Gibbs samples for our procedure BLTR by initiate with the initial valuations for parameters $\boldsymbol{\beta}$, and σ^2 , then we carry out the following procedures:

❖ Sampling \mathbf{z} :

To apply equation (7), Firstly, we produce z_j^* from the standard exponential distribution. After that, we must set $z_j = z_j^* + \lambda|\beta_j|$ by implementing the inversion process.

❖ Sampling $\boldsymbol{\beta}$:

To apply equation (10), we must generate $\boldsymbol{\beta}_R$ coefficients at the beginning of the algorithm, after that, we generate $\boldsymbol{\beta}$ coefficients from truncated normal distribution with the parameters $(\boldsymbol{\beta}_R, (\mathbf{X}'\mathbf{X} + \vartheta I_k)^{-1}\sigma^2)$.

❖ Sampling σ^2 :

To apply equation (8), we sampling σ^2 from the inverse gamma distribution with shape parameters $\frac{n}{2} + a$ and rate parameters $\frac{1}{2}(\mathbf{y}^* - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y}^* - \mathbf{X}\boldsymbol{\beta}) + b$.

❖ Sampling λ :

At last of computation procedure, we apply equation (9) to generate λ from truncated gamma

distribution with shape parameter $k + c$ and rate parameter d .

3. Simulation analysis:

In this chapter of the research, we test our proposed method and measure its performance compared to previous methods for estimating parameters and VS in Tobit regression. This test is carried out by applying simulation examples to the BLTR, Tobit regression process (Tr) by applying the AER package (Kleiber, Zeileis, 2017), Bayesian Tobit regression process (BTR) by applying the MCMCpack package (Martin, Quinn, Park, 2018), the Bayesian Tobit Quantile regression process (BTqr), and Bayesian Lasso Tobit quantile regression process (BLTqr) by applying the Brq package (Alhamzawi, 2018). All these packages will be implemented in R language. The standard comparison of the methods mentioned above is a median of means absolute deviations MMAD for estimated coefficients. The MMAD formula is

$$\text{MMAD} = \text{median}(\text{mean}(|\mathbf{X}\widehat{\boldsymbol{\beta}} - \mathbf{X}\boldsymbol{\beta}^t|))$$

where the parameter $\widehat{\boldsymbol{\beta}}$ is a vector of estimated coefficients and the parameter $\boldsymbol{\beta}^t$ is a vector of true coefficients values in the simulation

examples. This paper, we set $a=b=c=d=0.1$ and $\vartheta=0.01$.

3.1 i.i.d random errors examples:

A clarification of what we will do in this simulation, that we will create 7 independent variables from the normal distribution with parameters $(0, \sigma^2)$.

3.1.1 Example 1:

In this simulation example, we create 100 observations, the pair-wise correlations between each independent variable equalize to $0.5^{|i-j|}$, and we proposed true regression coefficients as follows

$$\beta^t = (0.75, \underbrace{0.75, \dots, 0.75}_7)'$$

And we simulated y_i^* by implementing the equation below

$$y_i^* = 0.75 + 0.75x_{1i} + 0.75x_{2i} + 0.75x_{3i} + 0.75x_{4i} + 0.75x_{5i} + 0.75x_{6i} + 0.75x_{7i} + \varepsilon_i$$

Method	\mathbb{P}^{β}	MMAD	SD
BLTR	1	.2703062	.0975433
Tr		.2721350	.1000219
BTr		.2934124	.1170751
BTqr		.3189504	.1306618
BLTqr		.3199308	.1170562
BLTR	4	.5410527	.1470687
Tr		.5416789	.1471063
BTr		.5560823	.1788521
BTqr		.6188514	.1755557
BLTqr		.5894346	.1280350

Table 1: Example 1 results of MMADs and SDs

3.1.2 Example 2:

This example is same as example above except for the setting that we set the true regression coefficients as follows

$$\beta^t = (1, 1, 2, 0, 1, 2, 0, 1)'$$

And we simulated y_i^* by implementing the equation below

$$y_i^* = 1 + x_{1i} + 2x_{2i} + x_{4i} + 2x_{5i} + x_{7i} + \varepsilon_i$$

Method	\mathbb{P}^{β}	MMAD	SD
BLTR	1	.3081883	.0961881
Tr		.3089278	.0965812
BTr		.3182100	.1002961
BTqr		.3385647	.1115737
BLTqr		.3356854	.1069906
BLTR	4	.6366920	.1454593
Tr		.6474908	.1438293
BTr		.6723154	.1916782
BTqr		.7007120	.1982751
BLTqr		.6723090	.1548616

Table 2: Example 2 results of MMADs and SDs

3.1.3 Example 3:

In this example, the data is same as previous example except for the setting that we set the true regression coefficients as follows

$$\beta^t = (1, 1, 0, 1, 0, 1, 0, 1)'$$

And we simulated y_i^* by implementing the equation below

$$y_i^* = 1 + x_{1i} + x_{3i} + x_{5i} + x_{7i} + \varepsilon_i$$

Method	\mathbb{P}^{β}	MMAD	SD
BLTR	1	0.2566264	0.0696783
Tr		0.2566902	0.0686645
BTr		0.2578269	0.0756091
BTqr		0.2822730	0.0789196
BLTqr		0.2683786	0.0768525
BLTR	4	0.5326660	0.1788678
Tr		0.5334055	0.1736695

BTr		0.5583471	0.2114098
BTqr		0.6365335	0.2154937
BLTqr		0.5643628	0.1859721

Table 3: Example 3 results of MMADs and SDs

3.1.4 Example 4:

In this example, the number of observations is same as previous examples, but the pair-wise correlations between x_i and x_j equal to 0.5, and we proposed true regression coefficients as follows

$$\beta^t = (1, 6, 0, 0, 0, 0, 0, 0)'$$

And we simulated y_i^* by implementing the equation below

$$y_i^* = 1 + 6x_{1i} + \varepsilon_i$$

Method	ρ^{β}	MMAD	SD
BLTR	1	0.3217117	0.1160644
Tr		0.3330405	0.1111610
BTr		0.3250991	0.1318257
BTqr		0.3530526	0.1459358
BLTqr		0.3444863	0.1351279
BLTR	4	0.6028703	0.1810483
Tr		0.6111545	0.1960355
BTr		0.6133544	0.2150937
BTqr		0.6840337	0.2079322
BLTqr		0.6308881	0.1883775

Table 4: Example 4 results of MMADs and SDs

3.1.5 Example 5:

In this example, we proposed the same true regression coefficients in example 4 with low

correlation, the pair-wise correlations between x_i and x_j equal to 0.3.

Method	ρ^{β}	MMAD	SD
BLTR	1	.3045046	.1077820
Tr		.3078803	.1041962
BTr		.3137167	.1174676
BTqr		.3455615	.1187870
BLTqr		.3238524	.1193782
BLTR	4	.6084789	.1946236
Tr		.6196739	.1922582
BTr		.6322880	.2173044
BTqr		.6659606	.2514506
BLTqr		.6103007	.2106599

Table 5: Example 5 results of MMADs and SDs

3.1.6 Example 6:

In this example, we proposed the same true regression coefficients in example 4 with high correlation, the pairwise correlations between x_i and x_j equal to 0.8.

Method	ρ^{β}	MMAD	SD
BLTR	1	0.2977267	0.0989983
Tr		0.3073310	0.1006740
BTr		0.3033420	0.1002597
BTqr		0.3316455	0.1130165
BLTqr		0.3050782	0.1122129
BLTR	4	0.5450217	0.1661782
Tr		0.5640171	0.1665735
BTr		0.5920677	0.1766288
BTqr		0.6169659	0.1918481
BLTqr		0.5463471	0.1691286

Table 6: Example 6 results of MMADs and SDs

From the tables of the previous examples (1,2,3,4,5 and 6), we find that BLTR approach was better than the other methods used in the comparisons. The MMADs results of BLTR procedure were relatively less than these results of the other procedures. This indicates the

quality of the performance of the BLTR process in terms of coefficient estimation and VS.

3.2 Heterogeneous random errors example:

The errors in this example are considered to confirm the performance of BLTR process for VS. We created 100 observations, and 8 independent variables, 5 of these variables are represented as standard normal noise variables. In addition, we set $y_i = \max\{y_i^*, 0\}$

$$\text{Where } y_i^* = x_i' \beta + (1 + x_{3i}) \varepsilon_i,$$

$$\beta^t = (1, 2, 1, 2, 0, 0, 0, 0, 0)',$$

$$\varepsilon_i \sim \text{Normal}(0, 1),$$

$$x_{1i} \sim \text{Normal}(0, 1), x_{3i} \sim \text{Uniform}(0, 1),$$

$$x_{2i} = x_{1i} + x_{3i} + z_i, z_i \sim \text{Normal}(0, 1)$$

Method	MMAD	SD
BLTR	0.24361931	0.06380391
Tr	0.24409931	0.06427611
BTr	0.24687247	0.06492119
BTqr	0.28280945	0.07813537
BLTqr	0.27495522	0.07614802

Table 4: Heterogeneous example results of MMADs and SDs

Table 4 lists MMADs and SDs of heterogeneous example. The results in this example prove that the effectiveness of the BLTR procedure is very good compared to all other procedures (Tr, BTr, BTqr, BLTqr).

4. Real Data:

The identification and detection of the causes about increasing the rate wheat production are one of the priorities of agricultural economist researchers. The identification of the real factors to increase production among several factors has no effect, can help predict the future the rate of

increase in wheat production. Hence the importance of this research, which attempts to identify some variables and to show how strong their impact on the rate of increase in wheat production. We perform BLTR within the four processes in wheat production data for comparison in terms of accurate prediction and VS. The real data in this research are taken from the national wheat development program in Qadisiyah, this data has been documented in 2017. The wheat production data includes 11 variables within 584 observations, these variables are sorting as follows in tables 5.

symbol	Variable description	Rank	Rank description
γ	The outcome variable	Percentage increases of wheat product	
X_1	Urea	Numeric	The quantity in kilogram
X_2	Sowing date	1	Ideal
		2	Early
		3	Late
X_3	The quantity of sowing	Numeric	The quantity in kilogram
X_4	Laser field leveling	1	Not using
		2	Using
X_5	Compound fertilizer	Numeric	The quantity in kilogram
X_6	Sowing seed	1	Not using
	machine	2	Using
X_7	Planting successive Mung bean crop	1	Planting
		2	Not planting
X_8	Herbicide for weeds	Numeric	The quantity in milliliter
X_9	High-Potassium	Numeric	The quantity in kilogram
X_{10}	Micro Elements	Numeric	The quantity in gram

Table 5: The top 11 worthy variables selected by BLTR process

Note: Dunam is equal to 2500 square meters.

β_0	X_1	X_2
-----------	-------	-------

	Estimate (25%, 95%)	Estimate (25%, 95%)	Estimate (25%, 95%)
BLTR	-0.04318 (-0.3604, 0.2674)	0.02136 (0.0200, 0.0236)	-0.67632 (-0.7575, -0.6240)
Tr	-0.08492 (-0.8722, 0.7024)	0.02093 (0.0137, 0.0280)	-0.66365 (-0.7864, -0.5409)
BTr	-0.08198 (-0.8986, 0.7196)	0.02097 (0.014, 0.028)	-0.66566 (-0.791, -0.546)
BTqr	-1.2292 (-1.7985, -0.5204)	0.02444 (0.0181, 0.0306)	-0.65595 (-0.8025, -0.5117)
BLTqr	-0.89837 (-1.5441, -0.1278)	0.02406 (0.0173, 0.0306)	-0.65542 (-0.8053, -0.5068)
	X_3	X_4	X_5
BLTR	-0.02206 (-0.0247, -0.0201)	1.28907 (0.9941, 1.5631)	0.00502 (0.0034, 0.0072)
Tr	-0.02179 (-0.0347, 0.0088)	1.35717 (0.6808, 2.0335)	0.00491 (-0.0076, 0.0174)
BTr	-0.02179 (-0.035, -0.008)	1.35763 (0.658, 2.035)	0.00492 (-0.008, 0.018)
BTqr	-0.00574 (-0.0176, -0.0030)	1.44739 (0.4979, 2.3409)	-0.00528 (0.0167, 0.0069)
BLTqr	-0.00992 (-0.0235, 0.0009)	1.28474 (0.4120, 1.9695)	-0.00308 (-0.0142, 0.0097)
	X_6	X_7	X_8
BLTR	-0.11009 (-0.3843, 0.1310)	0.91981 (0.8406, 0.9880)	0.00439 (0.0039, 0.0050)
Tr	-0.14263 (-0.8376, 0.5525)	0.93283 (0.6111, 1.2545)	0.00432 (0.0027, 0.0059)
BTr	-0.14509 (-0.840, 0.559)	0.93068 (0.601, 1.259)	0.00432 (0.003, 0.006)
BTqr	0.23115 (-0.6298, 1.1661)	0.99091 (0.6657, 1.3012)	0.00512 (0.0037, 0.0066)
BLTqr	0.26520 (-0.2931, 1.0838)	0.95502 (0.6188, 1.2851)	0.00503 (0.0036, 0.0065)
	X_9	X_{10}	
BLTR	0.03285 (0.0322, 0.0339)	0.00622 (0.0060, 0.0064)	
Tr	0.03268 (0.0255, 0.0399)	0.00623 (0.0045, 0.0079)	
BTr	0.03269 (0.026, 0.040)	0.00625 (0.005, 0.008)	
BTqr	0.02432 (0.0136, 0.0353)	0.00763 (0.0048, 0.0102)	
BLTqr	0.02531 (0.0146, 0.0368)	0.00746 (0.0047, 0.0099)	

Table 6: Coefficients estimation and Credible intervals (25%, 95%)

In table 6, the results showed the coefficients estimation and credible intervals (low credible interval LCI 25%, high credible interval HCI 95%). The credible intervals results of BLTR process are narrower than Tr, BTr, BTqr, and BLTqr process, and our proposed method is including all the estimations of other methods.

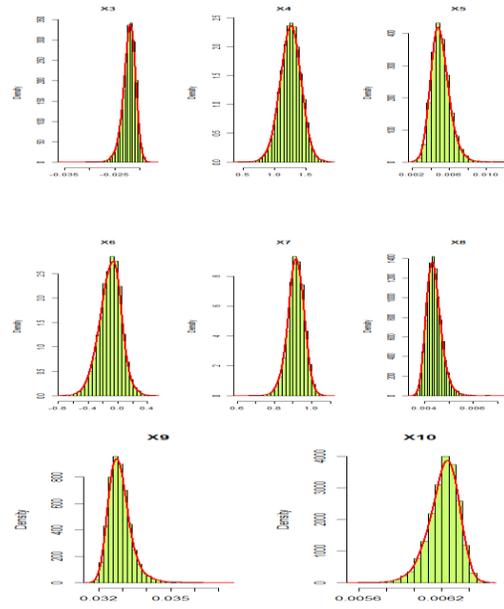
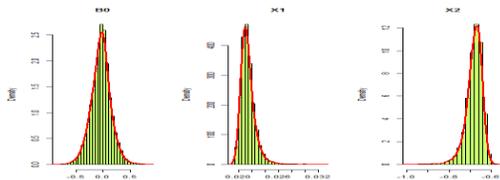


Diagram 1: Histograms of coefficients estimation for BLTR process

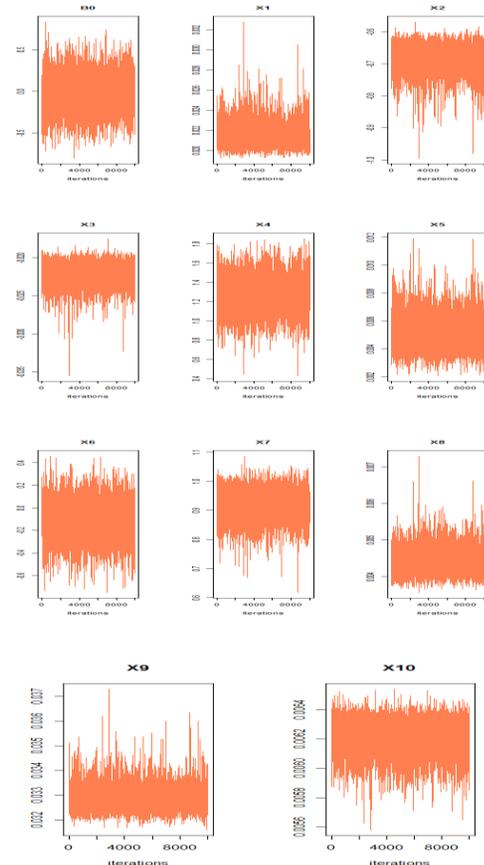


Diagram 2: Trace plots of coefficients estimation for BLTR process

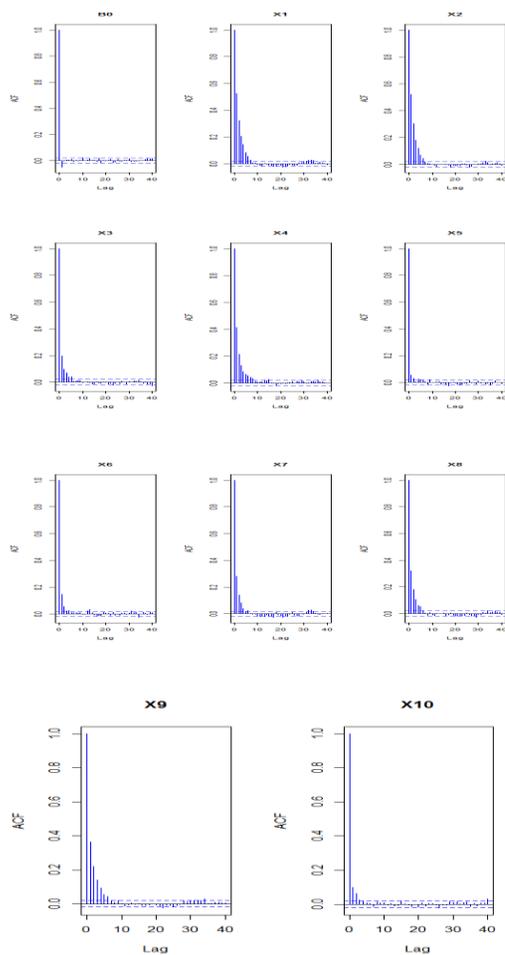


Diagram 3: Autocorrelations of coefficients estimation for BLTR process.

In diagram 1, which illustrates the histograms of BLTR process coefficients estimation, these coefficients are based on posterior samples of 10,000 recurrences. These histograms displayed that the conditional posteriors of BLTR coefficients are stationary for its underlying truncated normal distribution. In diagram 2, showed the BLTR coefficients trace plots, these plots show a reasonably good approximation, and the noise has been significantly deviated and the chain has reached stability and the center remains relatively constant. This means that the

chain is fully mixed and convergent. At last, from diagram 3, the plot showed the BLTR coefficients estimation autocorrelations, the 10 covariates in these data are highly correlated, and these MCMC chains in BLTR process was practically well.

5.Conclusions:

This research has presented a new process for model selection of Tobit regression, we proposed the BLTR method to estimates the coefficient with VS process. We developed new Bayesian hierarchical model for BLTR. In addition, we provided the Gibbs samples for the BLTR method. We demonstrated the advantages of the new procedure in both simulations and analysis of real-data. The results showed that our procedure performed well in terms of VS and parameters estimation.

References:

[1] Alhamzawi, R. (2015). Model selection in quantile regression models. *Journal of Applied Statistics*, 42(2), 445-458.

[2] Alhamzawi, R. (2016). Bayesian elastic net Tobit quantile regression. *Communications in Statistics-Simulation and Computation*, 45(7), 2409-2427.

[3] Alhamzawi, R., & Ali, H. T. M. (2018). Bayesian tobit quantile regression with penalty. *Communications in Statistics-Simulation and Computation*, 47(6), 1739-1750.

[4] Alhamzawi, R., & Taha Mohammad Ali, H. (2018). A new Gibbs sampler for Bayesian lasso. *Communications in Statistics-Simulation and Computation*, 1-17.

- [5] Alhamzawi, R., & Yu, K. (2012). Variable selection in quantile regression via Gibbs sampling. *Journal of Applied Statistics*, 39(4), 799-813.
- [6] Alhamzawi, R., & Yu, K. (2013). Conjugate priors and variable selection for Bayesian quantile regression. *Computational Statistics & Data Analysis*, 64, 209-219.
- [7] Alhamzawi, R., & Yu, K. (2014). Bayesian Lasso-mixed quantile regression. *Journal of Statistical Computation and Simulation*, 84(4), 868-880.
- [8] Alhamzawi, R., & Yu, K. (2015). Bayesian Tobit quantile regression using g-prior distribution with ridge parameter. *Journal of Statistical Computation and Simulation*, 85(14), 2903-2918.
- [9] Alhamzawi, R., Yu, K., & Benoit, D. F. (2012). Bayesian adaptive Lasso quantile regression. *Statistical Modelling*, 12(3), 279-297.
- [10] Alhusseini, F. H. H. (2017). New Bayesian Lasso in Tobit Quantile Regression. *Romanian Statistical Review Supplement*, 65(6), 213-229.
- [11] Alhusseini, F. H. H., & Georgescu, V. (2018). Bayesian composite Tobit quantile regression. *Journal of Applied Statistics*, 45(4), 727-739.
- [12] Andrew D. Martin, Kevin M. Quinn, Jong Hee Park (2011). MCMCpack: Markov Chain Monte Carlo in R. *Journal of Statistical Software*. 42(9): 1-21.
- [13] Andrews, D. F., & Mallows, C. L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 99-102.
- [14] Benoit, D. F., Alhamzawi, R., & Yu, K. (2013). Bayesian lasso binary quantile regression. *Computational Statistics*, 28(6), 2861-2873.
- [15] Chen, X., Wang, Z. J., & McKeown, M. J. (2011). A Bayesian Lasso via reversible-jump MCMC. *Signal Processing*, 91(8), 1920-1932.
- [16] Chib, S. (1992). Bayes inference in the Tobit censored regression model. *Journal of Econometrics*, 51(1-2), 79-99.
- [17] Choy, S. T. B., & Smith, A. F. M. (1997). Hierarchical models with scale mixtures of normal distributions. *Test*, 6(1), 205-221.
- [18] Christian Kleiber and Achim Zeileis (2008). *Applied Econometrics with R*. New York: Springer-Verlag. ISBN 978-0-387-77316-2.
- [19] Dries F. Benoit, Dirk Van den Poel (2018). Brq: An R package for Bayesian Quantile Regression. Working Paper.
- [20] Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression. *The Annals of statistics*, 32(2), 407-499.
- [21] Hans, C. (2009). Bayesian lasso regression. *Biometrika*, 96(4), 835-845.
- [22] Kyung, M., Gill, J., Ghosh, M., & Casella, G. (2010). Penalized regression, standard errors, and Bayesian lassos. *Bayesian Analysis*, 5(2), 369-411.
- [23] Lykou, A., & Ntzoufras, I. (2013). On Bayesian lasso variable selection and the specification of the shrinkage parameter. *Statistics and Computing*, 23(3), 361-390.

- [24] Mallick, H., & Yi, N. (2014). A new Bayesian lasso. *Statistics and its interface*, 7(4), 571.
- [25] Park, T., & Casella, G. (2008). The bayesian lasso. *Journal of the American Statistical Association*, 103(482), 681-686.
- [26] Raheem, S. E., Ahmed, S. E., & Doksum, K. A. (2012). Absolute penalty and shrinkage estimation in partially linear models. *Computational Statistics & Data Analysis*, 56(4), 874-891.
- [27] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267-288.
- [28] Wang, H., & Leng, C. (2007). Unified LASSO estimation by least squares approximation. *Journal of the American Statistical Association*, 102(479), 1039-1048.
- [29] West, M. (1987). On scale mixtures of normal distributions. *Biometrika*, 74(3), 646-648.
- [30] Yu, K., & Stander, J. (2007). Bayesian analysis of a Tobit quantile regression model. *Journal of Econometrics*, 137(1), 260-276.
- [31] Yuan, M., & Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(1), 49-67.

انحدار توبت لاسو البيزي

حيدر كاظم عباس

قسم الاحصاء-كلية الادارة والاقتصاد-جامعة القادسية

الخلاصة :

في هذا البحث ، اقترحنا نهجا جديدا لاختيار النموذج في انحدار توبت. استخدمنا في التقنية الجديدة طريقة Bayesian Lasso في انحدار توبت ، تمتاز هذه الطريقة بالعديد من الميزات التي تعطي التقدير الأمثل واختيار المتغيرات. على وجه التحديد ، قدمنا نموذجا هرميا جديدا. ثم قدمنا عينات جديدة من Gibbs. وقمنا أيضا بتوسيع النهج الجديد عن طريق إضافة معلمة التداخل ridge داخل مصفوفة التباين لتجنب التفرد في حالة تعدد الارتباط بين التغيرات أو في حالة زيادة عدد المعاملات عن عدد المشاهدات. تم إجراء مقارنة مع التقنيات السابقة الأخرى التي باستخدام أمثلة المحاكاة وتحليل البيانات الحقيقية. ومن الجدير بالذكر أن النتائج التي تم الحصول عليها كانت مشجعة وواعدة ، مما أعطت نتائج أفضل بالمقارنة مع الطرق السابقة.