Bayesian adaptive Lasso Tobit regression with a practical application

A thesis submitted for the degree of

MASTER OF STATSTICAL SCIENCE

By

Haider Kadhim Abbas Hilali

Supervised by

Dr. Rahim J. Al-Hamzawi

Statistics Department
College of Administration and Economics
University of Al-Qadisiyah

2019

Supervisor's recommendation

I certify that the thesis (Bayesian adaptive Lasso Tobit regression with a practical application) has been under my supervision in the Department of Statistics/College of Administration & Economics/University of Al_Qadisiyah it is a part of the requirements for a master's degree in statistics.

Signature:

Supervisor: Dr. Prof. Rahim J. Al-Hamzawi

Date:

Recommendation of the Chairman of the Graduate Studies Committee

Based on available recommendations, I would like to forward this thesis for discussion .

Signature:

Name:

Chairman of the Higher Studies Committee in department of statistics

Date:

Endorsement of the linguistics expert

This is to hereby that the thesis entitled (**Bayesian adaptive Lasso Tobit regression with a practical application**) has been reviewed in terms of stylistics and linguistics (grammar and spelling). Therefore, after the modification of all recommended notes thesis has become free of all linguistics errors and ready to be defended and used as a scientific method to award the degree of master in statistical science.

Signature : Haidr

Name: Dr. HAIDER H. M.

Date :

The decision if the defense committee

We are the head and members of the defense committee certify that we have been looked at the thesis entitled (Bayesian adaptive Lasso Tobit regression with a practical application) and we have debated the student (Haider Kadhim Abbas Hilali). As a result, the student has defended his thesis and all its content. So that we have found the thesis is worthy to be accepted to award the degree of Master in statistics sciences.

Ass. Prof. Dr. Ali Jawad Kadhim

(The head of the committee)

Ass. Prof. Dr. Ahmad Naeem Flaih

(Member)

Ass. Prof. Dr. Basim Shlaibah Msallam

(Member)

Prof. Dr. Rahim Jabbar Alhamzawi

(Member and supervisor)

Approval of the college committee

The council of the college of administration and economics at the University of Al-qadissiah have been approved on the decision of the defense committee.

Ass. Prof. Dr. Ali Jawad Kadhim

The Dean if the college of administration and economics / Al-qadissiah university

ACKNOWLEGMENTS

First of all, I would like to express my deep thanks to my supervisor Dr. Rahim J. Alhamzawi for his kind support and valuable notes throughout the thesis.

Hence, I must mention my gratitude to the staff of statistics department, especially Dr. Ali Alkenani, Dr. Muhannad Al-Saadony and Dr. Hassan S. Uraibi for their help and assistance throughout my study.

A special word of thanks dedicated to Dr. Taha Alshaybawe and Dr. Fadel Alhusseini for their valuable notes and assistance.

I am so thankful to my dear parents, brother, sisters, wife and my beloved sons.

Finally, I am indebted to my close colleague Luay Al-hemmery, and specials thanks are given to my boss Mr. Safaa Al-Janabi, the manager of Diwaniya agriculture directorate for his encouragement and supports.

DEDICATION

This thesis is dedicated to my beloved family and my dear wife and sons for their continuous supports and always being with me

ABSTRACT

The process of the variable selection refers to the category of problems where one attempt to determine the best subset of the pertinent variables, which can be used to obtain accurate adjustments to the results of a given response variable. Often, when the number of variables is too large, it is difficult to identify important and influential variables on the response variable. For this reason, the variable selection (VS) characteristic was considered very important in the data analysis. Regularization techniques is one fabulous way that has proven effectively for dealing with high dimensional data.

In previous years, statisticians have made great efforts in developing procedures of regularization to solve problems of VS. These procedures auto facilitate for VS by setting specific coefficients to zero and shrinking the coefficients estimates, and provide advantageous estimates even if the model contains a large number of highly correlated variables. Although the regularizations approaches have developed in recent years, these procedures can still be improved.

In this thesis, we have proposed new techniques for model selection in Tobit regression. These techniques are Bayesian Lasso and Bayesian adaptive Lasso in Tobit regression (BLTR, and BALTR). These techniques have many features that give good estimation and VS. Specifically, we have introduced a new hierarchal model for each technique. Then, new Gibbs sampler methods are introduced. We also extended the new approaches by adding the ridge parameter inside the variance-covariance matrix to avoid the singularity in case of multicollinearity or in case the number of predictors greater than the number of observations. A comparison was made with other existing techniques by applying the simulation examples and real data. It is worth mentioning, that the obtained results are promising and encouraging, giving better results compared to the existing methods.

TABLE OF CONTENTS

page

Acknowledgements i

Dedication ii

Abstract iii

Chapter One 1

1-1 Introduction... 1

1-2 Literature review.. 2

Chapter Two (Bayesian Lasso Tobit regression).. 11

2-1 Introduction... 11

2-2 BLTR hierarchy model and prior distribution 13

2-3 Full conditional posterior distributions for BLTR 14

2-4 BLTR computation 15

2-5 BLTR with ridge parameter.. 16

2-6 BLTR with ridge parameter computation.. 17

Chapter Three (Bayesian adaptive Lasso Tobit regression)........................ 19

3-1 Introduction... 19

3-2 BALTR hierarchy model and prior distribution 21

3-3 Full conditional posterior distributions for BALTR.................................... 22

3-4 BALTR computation 23

3-5 BALTR with ridge parameter... 24

3-6 BALTR with ridge parameter computation.. 25

Chapter Four (Simulation study analysis) 27
4-1 Introduction..........27
4-2 Independent and identically distribution (*i.i.d*) random errors examples 27
4-3 Non *i.i.d* (Heterogeneous) random errors examples..........48
Chapter Five (Practical application)..........51
5-1 Introduction..........51
4-2 The independent variables..........52
4-3 Real data results..........53
Chapter Six (Conclusions and recommendations)..........62
6-1 Conclusions..........62
6-2 Recommendations for future research..........62
Bibliography..........64

LIST OF TABLES

Table page

4-1 The coefficients estimate of Simulation example 1, when $\varepsilon_i \sim N(0,1)$ 28

4-2 The coefficients estimate of Simulation example 1, when $\varepsilon_i \sim N(0,4)$ 28

4-3 The coefficients estimate of Simulation example 1, when $\varepsilon_i \sim N(0,9)$ 29

4-4 The MMADs and SDs of Simulation example 1 29

4-5 The coefficients estimate of Simulation example 2, when $\varepsilon_i \sim N(0,1)$ 30

4-6 The coefficients estimate of Simulation example 2, when $\varepsilon_i \sim N(0,4)$ 30

4-7 The coefficients estimate of Simulation example 2, when $\varepsilon_i \sim N(0,9)$ 31

4-8 The MMADs and SDs of Simulation example 2 31

4-9 The coefficients estimate of Simulation example 3, when $\varepsilon_i \sim N(0,1)$ 32

4-10 The coefficients estimate of Simulation example 3, when $\varepsilon_i \sim N(0,4)$ 32

4-11 The coefficients estimate of Simulation example 3, when $\varepsilon_i \sim N(0,9)$ 33

4-12 The MMADs and SDs of Simulation example 3 33

4-13 The coefficients estimate of Simulation example 4, when $\varepsilon_i \sim N(0,1)$ 34

4-14 The coefficients estimate of Simulation example 4, when $\varepsilon_i \sim N(0,4)$ 34

4-15 The coefficients estimate of Simulation example 4, when $\varepsilon_i \sim N(0,9)$ 35

4-16 The MMADs and SDs of Simulation example 4 35

4-17 The coefficients estimate of Simulation example 5, when $\varepsilon_i \sim N(0,1)$ 36

4-18 The coefficients estimate of Simulation example 5, when $\varepsilon_i \sim N(0,4)$ 36

4-19 The coefficients estimate of Simulation example 5, when $\varepsilon_i \sim N(0,9)$ 37

4-20 The MMADs and SDs of Simulation example 5 37

4-21 The coefficients estimate of Simulation example 6, when $\varepsilon_i \sim N(0,1)$ 38

4-22 The coefficients estimate of Simulation example 6, when $\varepsilon_i \sim N(0,4)$ 38

4-23 The coefficients estimate of Simulation example 6, when $\varepsilon_i \sim N(0,9)$ 39

4-24 The MMADs and SDs of Simulation example 6 39

4-25 The coefficients estimate of Simulation example 7, when $\varepsilon_i \sim N(0,1)$ 40

4-26 The coefficients estimate of Simulation example 7, when $\varepsilon_i \sim N(0,4)$ 40

4-27 The coefficients estimate of Simulation example 7, when $\varepsilon_i \sim N(0,9)$ 41

4-28 The MMADs and SDs of Simulation example 7 41

4-29 The coefficients estimate of Simulation example 8, when $\varepsilon_i \sim N(0,1)$ 42

4-30 The coefficients estimate of Simulation example 8, when $\varepsilon_i \sim N(0,4)$ 42

4-31 The coefficients estimate of Simulation example 8, when $\varepsilon_i \sim N(0,9)$ 43

4-32 The MMADs and SDs of Simulation example 8 43

4-33 The coefficients estimate of Simulation example 9, when $\varepsilon_i \sim N(0,1)$ 44

4-34 The coefficients estimate of Simulation example 9, when $\varepsilon_i \sim N(0,4)$ 44

4-35 The coefficients estimate of Simulation example 9, when $\varepsilon_i \sim N(0,9)$ 45

4-36 The MMADs and SDs of Simulation example 9 45

4-37 The coefficients estimate of Simulation example 10, when $\varepsilon_i \sim N(0,1)$ 46

4-38 The coefficients estimate of Simulation example 10, when $\varepsilon_i \sim N(0,4)$ 46

4-39 The coefficients estimate of Simulation example 10, when $\varepsilon_i \sim N(0,9)$ 47

4-40 The MMADs and SDs of Simulation example 10 47

4-41 The coefficients estimate of Simulation example 11 48

4-42 The MMADs and SDs of Simulation example 11 49

4-43 The coefficients estimate of Simulation example 12 49

4-44 The MMADs and SDs of Simulation example 12 50

4-45 The top 11 worthy variables for real data...53

4-46 Coefficients estimations and Credible intervals CIs for real data.................53

LIST OF FIGURES

Figure page

4-1 Histograms of BLTR coefficients estimation 55

4-2 Autocorrelations of BLTR coefficients estimation 56

4-3 Trace plots of BLTR coefficients estimation 57

4-4 Histograms of BALTR coefficients estimation 58

4-5 Autocorrelations of BALTR coefficients estimation 59

4-6 Trace plots of BALTR coefficients estimation 60

CHAPTER ONE

1. Introduction:

Linear statistical models are widely used in biological, agricultural, physical and social sciences, biology, as well as in economics and engineering. They are useful in both planning stages of research and analysis of the resulting data. We are aware of that regression analysis is one of the statistical procedures that illustrate the relationships between explanatory variables and the dependent variable (outcome). When the explanatory variables values are known, then the regression analysis assistance us to predict the values of the outcome variable. Variable selection is a difficult and important problem that is an important goal for many types of statistical modelling. The difficulty of this problem is increased in actual applications when a true model may not exist. Given a dataset, you can fit many models.

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. However, when the independent variables are too large, or the number of observations is less than the number of variables, then it is very difficult to distinguish the independent variables that are important and influential in describing the Tobit regression model, which leads to the instability and overfitting of the model, consequently 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 procedure is known variable selection (VS). Although all VS procedures have evolved in recent years in linear regression, these procedures can still be improved.

In generals, the main objective of this thesis is to present several Bayesian regularization approaches in the framework of Tobit regression. These approaches are Bayesian Lasso and Bayesian adaptive Lasso in Tobit regression (BLTR, and BALTR). These approaches have many features that give a good estimation and get rid of all problems by variable selection.

2. Literature Review:

In many applications, the observations are partially constrained with the dependent variable and not constrained in the other part. This data is 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. Such a model was first proposed in a great exertion via Tobin (1958). Tobin analyzed the dependent variables of the regression model that cannot be negative. Consequently, the Tobit regression process is appropriate to this 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.

Between 1958 "when Tobin's articles appeared" and 1970, the Tobit model was used infrequently in econometric applications. At the same time, many Tobit models and several evaluation procedures have been suggested for these models. Additionally, estimation procedures are now many and varied, but it is hard for statisticians to follow all current models and estimation procedures. 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 (Kleiber et al., 2017), and MCMCtobit function in MCMCpack package (Martin et al., 2018) . The general formula of Tobit regression is

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

where y_i denote the outcome variable of interest for $i = 1, \dots, n$. Here, y_i^* denote a latent response variable as follows

$$y_i^* = x_i'\boldsymbol{\beta} + \varepsilon_i , \qquad (1.1)$$

where x_i' is the $1 \times k$ vector denoting the ith row of the $n \times k$ matrix of predictors $\boldsymbol{X}$, and the vector of predictors $\boldsymbol{\beta}$ is

$$\boldsymbol{\beta} = (\beta_0, \beta_1, \cdots, \beta_k)' ,$$

and ε_i is the error term, $p(\varepsilon_i) = Normal(0, \sigma^2)$, k is a number of independent variables, and n is the number of observations.

However, previous attempts of applying Tobit regression in diverse applications, VS are used in multiple fields. This performance opens the doors for applying this technique in several topics that enable statisticians to analyze the data. Especially, when the data has a wide array of variables. Therewith, occasionally the number of explanatory variables is too large. It is then difficult to know which variables are really important, which are the noise variables. Additionally, to the emergence of a number of problems, when we use some independent variables that are not important in describing the Tobit regression, and this leads to a regression model that will be unstable and weak in terms of prediction. The mechanism of VS is to improve the model prediction, providing faster and lower cost models and a good understanding of the data.

If a model is to be used for prediction, time and fatigue can be saved by measuring only necessary predictors. Frequent predictions will add noise to the estimation of other levels of interests and also leads to the loss of certain degrees of freedom. The selection of predictors among the several potential ones to be included in a model is one of the prime challenges in regression analysis. Therefore, VS procedure is very important in Tobit regression for several reasons, such as:

- Reducing the number of explanatory variables will be so beneficial for many purposes. Especially, economics and statistics.
- Besides performing VS process on a dataset, it is sometimes also useful to looks at variable importance. A high score on variable importance means that variable has a large effect on the responses and small adjustments in that variable value can lead to a large deviation in the response. Variables that score low on variable importance will mostly be removed during VS process, additionally, variables that are not removed from the model will also score high on importance.
- Creating a highly accurate precision model that gives good estimations and high predictions.

Efroymson (1960) proposed an algorithm called stepwise selection technique, considered one of the most widely known and widely used subsets selection. It is defined as an automatic process for selecting models in cases where there are a big number of possible explanatory variables. Then method is implemented mainly in regression analysis. It is a procedure that permits implements in forwards selection

(FS) and backward elimination (BE) in same time, dropping or adding variables at the various steps. This procedure is created two approaches, that is FS and BE.

The FS can be evaluated also by a F test as follows

$$F_j^{+} = max_j \left[\frac{RSS_k - RSS_{k+j}}{S_{k+j}^2}\right] > F_{in} \,. \tag{1.2}$$

And the BE process can be calculated by a F test as follows

$$F_j^{-} = min_j \left[\frac{RSS_{k-j} - RSS_k}{S_k^2}\right] < F_{out} \,. \tag{1.3}$$

where RSS is the residuals sum of squares, k is the number of estimated parameters in the model, S^2 is the estimate error variance of the model, F_{in} and F_{out} are used as stop criterions.

Even if k is less than n, looking at all possible models may not be the best thing to do, then the FS and BE approaches are not warranted to give the best model.

Mallows (1974) proposed a procedure called a Mallows C_k. This procedure named for Colin L. Mallows is applied to evaluated predictors that have been estimated via ordinary least square regression (OLS). It is implemented in the model selection situation, where the number of explanatory variables is obtainable for predicting the response variable. The aim is to get the better model that contains a subset of these predictors. The small values of C_k mean that the model is comparatively accurate. The formula of Mallows procedure is

$$C_k = \frac{RSS(k)}{S^2} - n + 2k \,, \tag{1.4}$$

where the RSS is the residuals sum of squares on a set of training data, n is the samples size of data, k is a number of the covariates in the model, and the S^2 is an estimation of the variance related to every response in the model. It was detecting that the C_k procedure does not give equivalent values to the previous procedure, but the model with the smallest C_k from this meaning will also be the same form with the smallest C_k from the previous meaning. If the model is correct then C_k will tend to be close to or smaller than k. Thus, a simple process of C_k can be utilized to submit the best model. However, the C_k procedure suffers from two major constraints, the first constraint is that C_k approximation is only valid for large sample size, and the second constraint is that C_k cannot deal with complex groups models as in the VS.

Akaike (1974) proposed the Akaike Information Criterion procedure a theoretical approach to information for model selection. The AIC procedure is one of the most common methods of variable selection. The value of an AIC procedure can be used to compare different models. By calculating the value of AIC for all models, the model with the lowest AIC is the best model, the formula of AIC procedure is

$$AIC = -2\, ln\, L + 2k\,. \qquad (1.5)$$

Int terms of the residual sum of squares then AIC formula is

$$AIC = n\, ln(RSS/n) + 2k\,, \qquad (1.6)$$

where L bis the maximum likelihood function of the model (*MLE*), k is the number of estimated parameters in the model, *RSS* is estimated residual of fitted model and n is the sample size, with the note that the error is normal ($i.i.d$).

A small *RSS* results in a lower AIC value and therefore a better model. For best subset selection the AIC offers a measure to compare models of different sizes with each other. In this way, the best model can be found. Though the AIC process suffers from two major constraints, the first constraint is that AIC relies on a weak procedure when k is large, and the second constraint is that AIC is no clear penalty takes in account the number of variables.

Schwarz (1978) proposed the Bayesian information criterion procedure (BIC), this procedure approximated the ln of the marginal density of the data, which is the density of the data unconditional on the parameters. Note that the likelihood is the density of the data conditional on the parameters. Then formula of BIC procedure is

$$BIC = -2\, ln\, L + k\, ln\, n \ . \qquad (1.7)$$

The residual sum of squares (RSS), the BIC formula is

$$BIC = n\, ln(RSS/n) + k\, ln\, n \, . \qquad (1.8)$$

Though then *BIC* and *AIC* arise from two dissimilar procedures, their explanations are alike. Exactly, a BIC procedure, such as the AIC procedure, can be clarified as a measure of model fit plus a penalty for the complexity. While $n \geq 10$ then $ln\,(n) > 2$ and so the penalty term in *BIC* is greater than the penalty term in AIC. In addition, BIC penalizes complex models more than AIC. The BIC usually penalizes free parameters more strongly than the AIC. However, it relies on the size of n and virtual magnitude of n and k. The BIC can be applied to compare models fitting only, then the numerical values of the response variable are identical for all estimates being compared. These compared models need not be overlapped, but that is some problems when using BIC procedure, the first problem is that BIC procedure suffers from the approximation, it is only valid for $n > k$ of parameters in the models. Additionally, the second problem is that BIC cannot handle complex groups of models as in the VS problem in high dimension data.

George and McCulloch (1993) suggested an alternative procedure for using an information criterion for model selection, this procedure is stochastic search VS ($SSVS$), which is feasible specifically to the Bayesian $MCMC$ framework, is $SSVS$ procedure. Starting with the full coefficients model and selecting the mixture prior distributions to removes inefficient coefficients in the regression and making them equal to zero, next that sampling the parameters from the posteriors and estimating the marginal inclusion probability ($MIPs$) for each coefficient using the amount of $SSVS$ samples containing each coefficient, at end estimating the posterior model probabilities ($PMPs$) for each model applying the quantities of $SSVS$ samples consumed in each model. One of the disadvantages of an $SSVS$ procedure takes a longer time to estimate, and the model with the highest posterior model may only be visited a handful of times.

Statistical researchers and data predictors are occasionally faced with more difficulties such as a large number of independent variables or low ratio the number of observations to the number of independent variables, or because of appearing multicollinearity problems, there be many numbers of strategies for statisticians auto utilizes in transacting with highs dimensional data, as well as VS procedures, and data reduction performances. A third family of techniques that have proven beneficial in the context of high dimensional data involves alternative parameter estimation algorithms known as regularization or shrinkage techniques. There be another set of the model estimation process, which can be used in such circumstances. In a sense, these procedures regularize or adjust the imprecise and volatile estimates of the regression coefficients.

Donoho and Johnstone (1994) have first introduced the notion of VS through regularization, it was then developed by Tibshirani (1996). Although these procedures are motivated by high dimensional data, they can also be effectively applied to sparse low to moderate dimensional problems, facilitating applications in a wide range of scientific problems. The general formula model, that illustrates the concept of regularization procedures, can be written as follows

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \ (\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta}) + F_\lambda(\boldsymbol{\beta}), \tag{1.9}$$

where $F_\lambda(.)$ is a function of the model coefficients in terms by a positive penalty parameter.

This parameter has an effective influence in VS, where it controls the degree of deflation applied to estimates. Regularization procedures recognize good values of $\widehat{\boldsymbol{\beta}}$ such that the at most important coefficients accept advanced values, and the minimum important are allocated coefficients at or near 0. Many of these procedures have attracted much attention recently; see, for example, bridge procedure Frank and Friedman (1993), Lasso procedure introduced in order to interpretability of regression models such as LARS (Efron, 2004), LARS process makes available a quick execution of the Lasso solution.

Zou and Hastie (2005) proposed the elastic net procedure to obtain enhanced performance when there is multicollinearity event between variables. In 2006, another regularization procedure introduced by Zou (2006), this procedure is adaptive Lasso regression, authorizing different penalty parameters to different regression coefficients, he proved that his proposed procedure had the characteristics of Oracle mentioned in Fan and Li (2001) that Lasso does not have. Park and Casella (2008) explicated that the parameters of the Lasso procedure can be estimated by the Bayesian pattern. Similarly, from a Bayesian point of view, "Bayesian Lasso regression (Hans, 2009)", "Bayesian adaptive Lasso; iterative adaptive Lasso (Sun, Zou, and Ibrahim, 2009)", "Bayesian adaptive Lasso with non-convex penalization (Griffin and Brown, 2010)", "The Bayesian elastic net (Li and Lin, 2010)", "A Bayesian Lasso via reversible jump MCMC (Chen, Wang, and McKeown, 2011)", "The new Bayesian Lasso (Mallick and Yi, 2014)", "Bayesian adaptive Lasso (Leng, Tran, and Nott, 2014)", and "Bayesian variable selection and estimation for group Lasso (Xu and Ghosh, 2015)".

In this thesis, we proposed two Bayesian regularization procedures, Bayesian Lasso and Bayesian adaptive Lasso in Tobit regression.

CHAPTER TWO

2. BAYESIAN LASSO TOBIT REGRESSION

2.1 Introduction:

The Lasso (least absolute shrinkage and selection operator) model is widely used as a regularization process for coefficient estimation in regression problems. Specifically, Tibshirani (1996) introduced Lasso method 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. Efron (2004) presented an effective algorithm for calculating the Lasso estimates of β_j via the LARS algorithm. The Lasso regression formulated as follows

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{Lasso}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}}(\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta}) + \lambda \sum_{j=1}^{k} |\beta_j| \dots \lambda \geq 0 \qquad (2.1)$$

where λ is a positive penalty parameter.

Tibshirani (1996) and Park and Casella (2008) explicated that the parameters of the Lasso regression can be estimated by the Bayesian pattern. Then Lasso will be taken as posterior mode underneath independent Laplace distribution prior for the β_j. On the other hand, Bayesian Lasso results are superbly like regular Lasso results. Though, the Bayesian Lasso is very simple to execute, and auto generate interval estimates for coefficients, containing the error variances. Following Andrews and Mallows (1974), Park and Casella (2008) represented the prior distribution of $\boldsymbol{\beta}$ as follows

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

$$=\prod_{j=1}^{k}\int_0^{\infty}\frac{1}{\sqrt{2\pi\sigma^2 t_j}}exp\left(-\frac{{\beta_j}^2}{2\sigma^2 t_j}\right)\frac{\lambda^2}{2}exp(-\frac{\lambda^2}{2}t_j)\,dt_j. \qquad (2.2)$$

Hans (2009) introduced a new aspect of other Bayesian remediation of Lasso regression, by using a new Gibbs sampler for Bayesian Lasso regression. Mallick and Yi (2014) proposed a new Bayesian Lasso by using scale mixture of uniform instead of scale mixture of normal distribution that used in Park and Casella (2008). The findings of Mallick and Yi (2014) in his research were very good and proved their efficiency from the previous Bayesian processes used. The good results notified in Mallick and Yi (2014) process encourage us to use the new Bayesian procedure in Tobit regression.

Recently, a new representation of the Laplace density given by Mallick and Yi (2014), this representation provided a different process of Lasso based model by using the scale mixture of a uniform representation of the Laplace density. This representation is written as follows:

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

$$=\prod_{j=1}^{k}\frac{\lambda}{2\sqrt{\sigma^2}}\int_{u_j>\frac{|\beta_j|}{\sqrt{\sigma^2}}}\lambda\ \ e^{-\lambda u_j}\ \ du_j. \qquad (2.3)$$

They pointed out that the posterior distribution $\pi(\boldsymbol{\beta}|\sigma^2,\lambda)$ is similar to the main procedure of Park and Casella (2008), this formulation has gorgeous properties. In addition, Alhamzawi (2018) developed a new Gibbs sampler for Bayesian Lasso via mixture of truncated normal formulation with exponential mixture densities. In this thesis, following Mallick and Yi (2014), we used a new hierarchical representation of BLTR.

2.2 BLTR hierarchy model and prior distributions:

Mallick and Yi (2014) proposed Bayesian Lasso procedure as follows:

$$\frac{\lambda}{2} e^{-\lambda|\beta_j|} = \int_{w>|\beta_j|} \frac{1}{2w} \frac{\lambda^2}{\Gamma(2)}\ w^{2-1}\ e^{-\lambda w} dw \ldots \lambda \geq 0 \qquad (2.4)$$

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

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

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

$$= \int_{z_j>\lambda|\beta_j|} \frac{\lambda}{2}\ e^{-z_j} dz_j \ldots \lambda \geq 0 \qquad (2.5)$$

The hierarchical model of BLTR is

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

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

$$\beta_j|\lambda \sim \text{Uniform}\left(-\frac{1}{\lambda}, \frac{1}{\lambda}\right), \qquad (2.6)$$

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

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

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

where $\boldsymbol{z} = (z_1, \cdots, z_k)$

2.3 Full conditional posterior distributions of BLTR:

The conditional distribution of $\boldsymbol{y}^*$ is follows:

$$y_i^* | \, y_i, \boldsymbol{\beta} \; \sim \begin{Bmatrix} y_i, & if \;\; y_i^* > y^0 \\ N(\boldsymbol{X\beta}, \sigma^2 I_n) I\{y_i^* \leq y^0\}, & otherwise \end{Bmatrix}.$$

where $y^0$is a known censoring point.

Then the conditional posterior distribution of $\boldsymbol{\beta}$ as follows:

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

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

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

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

$$\boldsymbol{\beta}| \, \boldsymbol{y}^*, \boldsymbol{X}, \boldsymbol{z} \sim N_k\left(\widehat{\boldsymbol{\beta}}_{OLS}, \; (\boldsymbol{X}'\boldsymbol{X})^{-1} \, \sigma^2\right) \prod_{j=1}^{k} I\left\{-\frac{z_j}{\lambda} < \beta_j < \frac{z_j}{\lambda}\right\}. \qquad (2.7)$$

The conditional posterior distribution of $\boldsymbol{z}$ as follows:

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

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

$$\boldsymbol{z} \sim \prod_{j=1}^{k} Exponential(1)\, I\{z_j > \lambda|\beta_j|\}\,. \tag{2.8}$$

The conditional posterior distribution of σ^2 as follows:

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

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

$$\sigma^2|\boldsymbol{y}^*, \boldsymbol{X}, \boldsymbol{\beta} \sim \text{InvGamma}\ \left(\frac{n}{2} + a, \frac{1}{2}(\boldsymbol{y}^* - \boldsymbol{X\beta})'(\boldsymbol{y}^* - \boldsymbol{X\beta}) + b\right). \tag{2.9}$$

The conditional posterior distribution of λ as follows:

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

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

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

$$\lambda|\boldsymbol{\beta} \sim\ \text{Gamma}\big((k+h), d\big) \prod_{j'=1}^{k} I\left\{\lambda < \frac{z_j}{|\beta_j|}\right\}, \tag{2.10}$$

where the $\widehat{\boldsymbol{\beta}}_{OLS}$ is ordinary least squares estimators, and $I(.)$ denotes an indicator function.

2.4 BLTR computation:

In the beginning, we specify Gibbs samples for BLTR procedure by initiate with the initial valuations for parameters $\boldsymbol{\beta}$, $\boldsymbol{z}$, λ and σ^2, then we carry out the algorithm as follows:

Algorithm 1 (Sampling in BLTR model).

- **Sampling $\boldsymbol{y}^*$:** We generate $\boldsymbol{y}^*$ latent variable from truncated normal distribution with mean $\boldsymbol{X\beta}$ and variance $\sigma^2 I_n$.

- **Sampling $\boldsymbol{z}$:** We generate z_j as follows $z_j = {z_j}^* + \lambda|\beta_j|$, where ${z_j}^*$ is an exponential distribution.

- **Sampling $\boldsymbol{\beta}$:** We generate $\boldsymbol{\beta}$ coefficients from truncated normal distribution with $(\widehat{\boldsymbol{\beta}}_{OLS},\ (\boldsymbol{X}'\boldsymbol{X})^{-1}\sigma^2)$.

- **Sampling $\boldsymbol{\sigma^2}$:** We generate σ^2 from inverse gamma distribution with shape parameter $\frac{n}{2} + a$ and rate parameter

$$\frac{1}{2}(\boldsymbol{y}^* - \boldsymbol{X\beta})'(\boldsymbol{y}^* - \boldsymbol{X\beta}) + b.$$

- **Sampling $\boldsymbol{\lambda}$:** We generate λ from truncated gamma distribution with shape parameter $k + h$ and rate parameter d.

2.5 BLTR with ridge parameter:

In practice, the above procedure performs very well. However, the above estimator of $\widehat{\boldsymbol{\beta}}_{OLS}$ is highly unsteady in the existence of multicollinearity. In addition, the matrix $\boldsymbol{X}'\boldsymbol{X}$ is singular if explanatory variables k is more than n. Gupta and Ibrahim (2007) have developed a process to deal with these situations. Specifically, they added a ridge parameter within $\boldsymbol{X}'\boldsymbol{X}$ matrix. Following Gupta and Ibrahim (2007), we added the ridge parameter remedy actual challenges that may appear via multicollinearity and overfitting problems.

Now, referring to the equation (2.7) and adding ridge parameter ϑ to the equation, we will get

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

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

$$\boldsymbol{\beta}|\boldsymbol{y}^*,\boldsymbol{X},\boldsymbol{z} \sim N_k(\boldsymbol{\beta_R},(\boldsymbol{X}'\boldsymbol{X}+\vartheta I_k)^{-1}\sigma^2)\prod_{j=1}^{k} I\left\{-\frac{z_j}{\lambda} < \beta_j < \frac{z_j}{\lambda}\right\}. \quad (2.11)$$

where the $\boldsymbol{\beta_R}$ is ridge estimators and $I(.)$ denotes an indicator function.

2.6 BLTR with ridge parameter computation:

We require Gibbs samples for BLTR procedure with ridge parameter by initiate with the initial valuations for parameters $\boldsymbol{\beta}$, $\boldsymbol{z}$, λ and σ^2, then we carry out the algorithm as follows:

Algorithm 2 (Sampling in BLTR model with ridge parameter).

- **Sampling $\boldsymbol{y}^*$:** We generate $\boldsymbol{y}^*$ latent variable from truncated normal distribution with mean $\boldsymbol{X\beta}$ and variance $\sigma^2 I_n$.
- **Sampling $\boldsymbol{z}$:** We generate the z_j as follows $z_j = z_j^* + \lambda|\beta_j|$, where $z_j^*$is an exponential distribution.

- **Sampling $\boldsymbol{\beta}$:** We generate $\boldsymbol{\beta}$ coefficients from truncated normal with mean $\boldsymbol{\beta_R}$ and variance covariance $(\boldsymbol{X}'\boldsymbol{X} + \vartheta I_k)^{-1}\sigma 2$.

- **Sampling $\boldsymbol{\sigma^2}$:** We generate σ^2 from inverse gamma distribution with shape parameter $\frac{n}{2} + a$ and rate parameter

$$\frac{1}{2}(\boldsymbol{y}^* - \boldsymbol{X\beta})'(\boldsymbol{y}^* - \boldsymbol{X\beta}) + b.$$

- **Sampling $\boldsymbol{\lambda}$:** We generate λ from truncated gamma distribution with shape parameter $k + h$ and rate parameter d.

CHAPTER THREE

3. BAYESIAN ADAPTIVE LASSO TOBIT REGRESSION

3.1 Introduction:

A lot of work has been devoted to the development of diverse Bayesian organizational procedures for making a variable selection in linear models. One of these approaches is adaptive Lasso, as a regularization method, evades overfitting penalizing large coefficients. Also, it has the same advantage that Lasso, it can shrink some of the coefficients to exactly zero, giving subsequently a selection of attributes by the regularization. Zou (2006) proposed the adaptive Lasso, who upgraded the Lasso way proposed by Tibshirani (1996), the adaptive Lasso procedure permitting different penalty parameters to different regression coefficients. Zou (2006) proved that his proposed procedure had the characteristics of Oracle mentioned in Fan and Li (2001) that Lasso does not have. Specifically, Zou (2006) indicates that his proposed procedure adopts the correct form of non-zero coefficients with the probability that he tends to one. Park and Casella (2008) suggested the Lasso procedure based on a Bayesian point of sight. Likewise, Mallick and Yi (2014) suggested a new procedure known to be as new Bayesian Lasso regression for VS and coefficient estimation in linear regression. In general, the last procedure observed results display that the Mallick and Yi (2014) procedure applied well compares with other Bayesian and non-Bayesian regression procedures.

The good results reported in Mallick procedure motivate us to suggest a new Bayesian regression procedure. Subsequently, we suggested a Bayesian hierarchical for adaptive Lasso Tobit regression (BALTR), and proposed a new

Gibbs sampler (GS) for BALTR, that is sets up on a theoretical derivation of the Laplace density.

It is well known, that the Lasso procedure gives biased estimates of considerable coefficients, so it might be below the required best level in terms of estimation risks. Zou (2006) evidenced that the Lasso selects the incorrect model with non-fade the probability, despite the sample size and how λ is chosen. The event requires that coefficients not in the model aren't representable by coefficients in the real models. But this event is simply suffering because of the collinearity cases between the coefficients. On the opposite hand, that the Lasso process does not have Oracle properties. So, Zou (2006) suggested the adaptive Lasso technique which gives a consistent model for VS. Therefore, we consider BALTR approach in this thesis, the adaptive Lasso enjoys the Oracle properties by utilizing the adaptably weighted Lasso penalty parameter, and leads to a near minimax good estimators. Additionally, the adaptive Lasso technique needs to initials estimates of the regression coefficients, when a sample sizes is less than of the covariates number, which is mostly not available in the high dimensional data. The estimator of adaptive Lasso is given by

$$\widehat{\boldsymbol{\beta}}_{Alasso} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} (\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta}) + \sum_{j=1}^{k} \lambda_j |\beta_j| \ \ldots \ \lambda_j \geq 0 \,, \qquad (3.1)$$

where varied penalty parameters are utilized for the regression coefficients. Confidently, for the not important explanatory variables, we must place larger penalty λ_j on their matching coefficients.

We propose a BALTR procedure in this thesis for coefficients estimations and VS. We suggest a news practices of the adaptive Lasso form by using the scale mixtures of a uniform representation of the Laplace distribution. Following Mallick & Yi (2014), the Laplace representation in equation (2.4) can adaptive as

$$\frac{\lambda_j}{2} e^{-\lambda_j|\beta_j|} = \int_{w_j>|\beta_j|} \frac{1}{2w_j} \frac{{\lambda_j}^2}{\Gamma 2} {w_j}^{2-1} e^{-\lambda_j w_j} \, dw_j \;\; \dots \; \lambda_j \geq 0$$

In this thesis, we modify the above formula as follows:

Let $z_j = \lambda_j w_j \quad \Rightarrow \quad dz_j = \lambda_j dw_j$ then

$$\frac{\lambda_j}{2} e^{-\lambda_j|\beta_j|} = \frac{\lambda_j}{2} e^{-|\lambda_j\beta_j|},$$

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

$$= \int_{z_j>|\lambda_j\beta_j|} \frac{\lambda_j}{2} e^{-z_j} dz_j \;\; \dots \; \lambda_j \geq 0 \qquad (3.2)$$

In practices, this formula produces more tractable and efficient Gibbs Samples than the previous formula.

3.2 BALTR hierarchy model and Prior Distributions:

By using equation (1.1) and equation (3.2), the Bayesian hierarchical model can be formulated as follows:

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

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

$$\boldsymbol{\beta}|\lambda \sim \prod_{j=1}^{k} \text{Uniform}\left(-\frac{1}{\lambda_j}, \frac{1}{\lambda_j}\right), \qquad (3.3)$$

$$\boldsymbol{z} \sim \prod_{j=1}^{k} \text{Exponential}(1),$$

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

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

where $\boldsymbol{z} = (z_1, \cdots, z_k)$ and $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_k)$.

3.3 Full Conditional Posterior Distributions of BALTR:

Firstly, we can express the joint posterior distribution of all BALTR procedure parameters as follows

The distribution of $\boldsymbol{y}^*$ is follows:

$$y_i^* | \, y_i, \boldsymbol{\beta} \ \sim \begin{Bmatrix} y_i, & if \ \ y_i^* > y^0 \\ N(\boldsymbol{X\beta}, \sigma^2 I_n) I\{y_i^* \leq y^0\}, & otherwise \end{Bmatrix}.$$

The conditional posterior distribution of $\boldsymbol{\beta}$ is follows:

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

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

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

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

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

$$\boldsymbol{\beta} | \boldsymbol{y}^*, \boldsymbol{X}, \boldsymbol{z}, \boldsymbol{\lambda} \sim N_k\left(\widehat{\boldsymbol{\beta}}_{OLS}, (\boldsymbol{X}'\boldsymbol{X})^{-1}\sigma^2\right) \prod_{j=1}^{k} I\left\{-\frac{z_j}{\lambda_j} < \beta_j < \frac{z_j}{\lambda_j}\right\}. \qquad (3.4)$$

The conditional posterior distribution of $\boldsymbol{z}$ as follows:

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

$$\propto \ \pi(\boldsymbol{z}) \ I\{z_j > |\lambda_j \beta_j|\},$$

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

$$\boldsymbol{z} \sim \prod_{j=1}^{k} \text{Exponential}(1)\ I\{z_j > |\lambda_j \beta_j|\}. \tag{3.5}$$

The conditional posterior distribution of σ^2 as follows:

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

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

$$\sigma^2|\boldsymbol{y}^*, \boldsymbol{X}, \boldsymbol{\beta} \sim \text{InvGamma}\left(\frac{n}{2} + a, \frac{1}{2}(\boldsymbol{y}^* - \boldsymbol{X\beta})'(\boldsymbol{y}^* - \boldsymbol{X\beta}) + b\right). \tag{3.6}$$

The conditional posterior distribution of λ as follows:

$$\pi(\lambda_j|\beta_j) \propto \pi(\beta_j|\lambda_j)\, \pi(\lambda_j),$$

$$\propto \lambda_j\, \pi(\lambda_j) I\left\{\lambda_j < \frac{z_j}{|\beta_j|}\right\},$$

$$\propto \lambda_j^{(c+1)-1} exp\{-d\lambda_j\}\, I\left\{\lambda_j < \frac{z_j}{|\beta_j|}\right\},$$

$$\lambda_j|\beta_j \sim \text{Gamma}(c + 1, d) I\left\{\lambda_j < \frac{z_j}{|\beta_j|}\right\}, \tag{3.7}$$

where the $\widehat{\boldsymbol{\beta}}_{OLS}$ is ordinary least squares estimators and $I(.)$ denotes an indicator function.

3.4 BALTR computation:

We specify Gibbs samples for BALTR procedure by initiate with the initial valuations for parameters $\boldsymbol{\beta}$, $\boldsymbol{z}$, $\boldsymbol{\lambda}$ and σ^2, then we carry out the algorithm as follows:

Algorithm 3 (Sampling in BALTR model).

- **Sampling $\boldsymbol{y}^*$:** We generate $\boldsymbol{y}^*$ latent variable from truncated normal distribution with mean $\boldsymbol{X\beta}$ and variance $\sigma^2 I_n$.

- **Sampling $\boldsymbol{z}$:** We generate the z_j as follows $z_j = z_j{}^* + \left|\lambda_j \beta_j\right|$, where $z_j{}^*$is an exponential distribution.

- **Sampling $\boldsymbol{\beta}$:** We generate $\boldsymbol{\beta}$ coefficients from truncated normal distribution with mean $\boldsymbol{\beta_{Ols}}$ and variance covariance $(\boldsymbol{X'X})^{-1}\sigma^2$.

- **Sampling $\boldsymbol{\sigma^2}$:** We generate the σ^2 from the inverse gamma distribution with shape parameter $\frac{n}{2} + a$ and rate parameter

$$\frac{1}{2}(\boldsymbol{y}^* - \boldsymbol{X\beta})'(\boldsymbol{y}^* - \boldsymbol{X\beta}) + b\ .$$

- **Sampling $\boldsymbol{\lambda}$:** We generate the λ_j from truncated gamma distribution with shape parameter $c + 1$ and rate parameter d.

3.5 BALTR with ridge parameter:

For the same reasons mentioned in Chapter 2, exactly in section 2.5, we adding a ridge parameter to remedy actual challenges that may appear with multicollinearity and overfitting problems. Then, referring to the equation (3.4) and add ridge parameter ϑ to the equation as follows:

$$exp\left\{-\frac{1}{2\sigma^2}(-2\boldsymbol{y}^{*\prime}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X}+\vartheta I_k)^{-\mathbf{1}}(\boldsymbol{X}'\boldsymbol{X}+\vartheta I_k)\boldsymbol{\beta} + \boldsymbol{\beta}'(\boldsymbol{X}'\boldsymbol{X}+\vartheta I_k)\boldsymbol{\beta})\right\}\prod_{j=1}^{k} I\left\{|\beta_j| < \frac{z_j}{\lambda_j}\right\},$$

$$\propto exp\left\{-\frac{1}{2\sigma^2}(-2\boldsymbol{\beta_R}'(\boldsymbol{X}'\boldsymbol{X}+\vartheta I_k)\boldsymbol{\beta}+\boldsymbol{\beta}'(\boldsymbol{X}'\boldsymbol{X}+\vartheta I_k)\boldsymbol{\beta})\right\}\prod_{j=1}^{k} I\left\{|\beta_j| < \frac{z_j}{\lambda_j}\right\},$$

$$\boldsymbol{\beta}|\ \boldsymbol{y}^*,\boldsymbol{X},\boldsymbol{z},\boldsymbol{\lambda} \sim N_k(\boldsymbol{\beta_R},(\boldsymbol{X}'\boldsymbol{X}+\vartheta I_k)^{-1}\sigma^2)\prod_{j=1}^{k} I\left\{-\frac{z_j}{\lambda_j} < \beta_j < \frac{z_j}{\lambda_j}\right\}, \qquad (3.8)$$

where the $\boldsymbol{\beta_R}$ is ridge estimators and $I(.)$ denotes an indicator function.

3.6 BALTR with ridge parameter computation:

We specify Gibbs samples for BALTR procedure with ridge parameter by initiate with the initial valuations for parameters $\boldsymbol{\beta}$, $\boldsymbol{z}$, $\boldsymbol{\lambda}$ and σ^2 , then we carry out the algorithm as follows

Algorithm 4 (Sampling in BALTR model with ridge parameter).

- **Sampling $\boldsymbol{y}^*$:** We generate the latent variable $\boldsymbol{y}^*$ from truncated normal distribution with mean $\boldsymbol{X\beta}$ and variance $\sigma^2 I_n$.

- **Sampling $\boldsymbol{z}$:** We generate $\boldsymbol{z}$ as follows $z_j = z_j{}^* + |\lambda_j\beta_j|$, where $z_j{}^*$ is an exponential distribution.

- **Sampling $\boldsymbol{\beta}$:** We generate $\boldsymbol{\beta}$ coefficients from truncated normal distribution with mean $\boldsymbol{\beta_R}$ and variance covariance $(\boldsymbol{X}'\boldsymbol{X}+\vartheta I_k)^{-1}\sigma^2$.

- **Sampling $\boldsymbol{\sigma^2}$:** We generate the σ^2 from the inverse gamma distribution with shape parameter $\frac{n}{2} + a$ and rate parameter

$$\frac{1}{2}(\boldsymbol{y}^* - \boldsymbol{X\beta})'(\boldsymbol{y}^* - \boldsymbol{X\beta}) + b.$$

- **Sampling $\boldsymbol{\lambda}$.:** We generate the λ_j from truncated gamma distribution with shape parameter $c + 1$ and rate parameter d.

CHAPTER FOUR

4. SIMULATION STUDY ANALYSIS

4.1 Introduction:

In this part of this Chapter, we test our proposed procedures and measure its performance compared to previous techniques for estimating parameters and VS in Tobit regression. This test is carried out by applying simulation examples to our procedures BALTR and BLTR, Tobit regression (Tr) by implementing the AER package (Kleiber et al., 2017), Bayesian Tobit regression method (Btr) by implementing the MCMCpack package (Martin et al., 2018), and the Bayesian Tobit Quantile regression method (Btqr) by using the Brq package (Alhamzawi, 2018). All these packages will be implemented in R language. For comparison, we draw 10,000 iterations of the Gibbs sampling, the first 1000 were ruled out as burn in. The procedures are evaluated based on the median of mean absolutes deviations MMAD). The formula of MMAD as follows

$$\text{MMAD} = \text{median}(\text{mean}\ \left(\left| \boldsymbol{X}\widehat{\boldsymbol{\beta}} - \boldsymbol{X}\boldsymbol{\beta}^{\boldsymbol{true}} \right| \right), \qquad (4.1)$$

where the parameter $\widehat{\boldsymbol{\beta}}$ is a vector of estimated coefficients and the parameter $\boldsymbol{\beta}^{\boldsymbol{true}}$ is a vector of true coefficients values in the simulation examples. In this chapter, we set a=b=c=d=0.05, ϑ=0.01 and Tau=0.5. For each simulation study, we run 200 simulations.

4.2 Independents and identically distributed (*i.i.d*) random errors:

Here, a clarification of what we will do in this simulation, that we will create independent variables from the multivariate normal distribution with parameters mean 0, and three values of the variance σ^2.

4.2.1 Simulation example 1:

In this simulations example, we create 7 independent variables with 100 observations, the pair wise correlation between each independent variable equalizes to $0.5^{|i-j|}$, and we set the true regression coefficients as follows:

$$\boldsymbol{\beta}^{true} = (1, 0, 1, 0, 1, 0, 1, 0)'.$$

We simulated y_i^* as follows:

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

where $\varepsilon_i \sim$ Normal $(0,\ \sigma^2)$, and $\sigma^2 \in \{1, 4, 9\}$.

We list the results of regression coefficients estimates as tables below

$\boldsymbol{\beta}^{true}$	BALTR	BLTR	Tr	Btr	Btqr
$\boldsymbol{\beta_0} = \mathbf{1}$	0.98827	0.98864	0.99785	0.96099	0.97584
$\boldsymbol{\beta_1} = \mathbf{0}$	0.00038	0.00045	0.00002	0.00145	0.00296
$\boldsymbol{\beta_2} = \mathbf{1}$	1.01733	1.01547	1.01471	1.03469	1.02866
$\boldsymbol{\beta_3} = \mathbf{0}$	-0.00748	-0.00737	-0.00842	-0.00566	-0.00396
$\boldsymbol{\beta_4} = \mathbf{1}$	1.00599	1.00435	1.00140	1.02812	1.03061
$\boldsymbol{\beta_5} = \mathbf{0}$	-0.00089	-0.00054	-0.00036	-0.00301	0.00045
$\boldsymbol{\beta_6} = \mathbf{1}$	1.00329	1.00104	0.99984	1.02516	1.02201
$\boldsymbol{\beta_7} = \mathbf{0}$	-0.00801	-0.00840	-0.00916	-0.00666	-0.01113

Table 1: The coefficients estimates of Simulation example 1, when $\varepsilon_i \sim \mathrm{N}(0,1)$

$\boldsymbol{\beta}^{true}$	BALTR	BLTR	Tr	Btr	Btqr
$\boldsymbol{\beta_0} = \mathbf{1}$	0.96165	0.96346	0.98341	0.90094	0.92477
$\boldsymbol{\beta_1} = \mathbf{0}$	-0.03419	-0.03484	-0.03525	-0.03771	-0.04608
$\boldsymbol{\beta_2} = \mathbf{1}$	1.03859	1.03449	1.03971	1.07396	1.08750
$\boldsymbol{\beta_3} = \mathbf{0}$	-0.01906	-0.01769	-0.01797	-0.02327	-0.04456
$\boldsymbol{\beta_4} = \mathbf{1}$	0.99276	0.98515	1.00533	1.04831	1.07056
$\boldsymbol{\beta_5} = \mathbf{0}$	-0.00555	-0.00538	-0.00443	0.00100	-0.01794
$\boldsymbol{\beta_6} = \mathbf{1}$	0.96883	0.96457	0.97656	1.00677	1.00723
$\boldsymbol{\beta_7} = \mathbf{0}$	-0.00152	-0.00026	0.00109	-0.00239	-0.01064

Table 2: The coefficients estimates of Simulation example 1, when $\varepsilon_i \sim \mathrm{N}(0,4)$

β^{true}	BALTR	BLTR	Tr	Btr	Btqr
$\beta_0 = 1$	0.93230	0.93821	0.98741	0.85266	0.92306
$\beta_1 = 0$	-0.00241	-0.00409	-0.00274	0.00304	0.00642
$\beta_2 = 1$	0.90639	0.89763	0.96096	1.00688	0.97875
$\beta_3 = 0$	0.06056	0.06073	0.06894	0.06871	0.08918
$\beta_4 = 1$	0.92676	0.91765	0.95287	1.00207	0.97170
$\beta_5 = 0$	-0.00396	-0.00296	-0.00268	-0.00263	0.01202
$\beta_6 = 1$	0.93320	0.92347	0.97901	1.02463	1.01493
$\beta_7 = 0$	0.02254	0.01934	0.01915	0.01989	0.00235

Table 3: The coefficients estimates of Simulation example 1, when $\varepsilon_i \sim \mathrm{N}(0,9)$

Procedure	σ^2	MMAD	SD
BALTR	1	0.253714	0.066260
BLTR		0.254578	0.068390
Tr		0.255161	0.069849
BTr		0.264513	0.075860
BTqr		0.282732	0.077842
BALTR	4	0.486562	0.142146
BLTR		0.487090	0.142711
Tr		0.492903	0.149641
BTr		0.508624	0.168405
BTqr		0.559407	0.166772
BALTR	9	0.688334	0.197845
BLTR		0.699088	0.196353
Tr		0.711708	0.209395
BTr		0.735739	0.231249
BTqr		0.789161	0.230088

Table 4: The MMADs and SDs results of Simulation example 1

4.2.2 Simulation example 2:

This simulation example is same as example above except we create 8 independent variables, and we set the true regression coefficients as follows:

$$\boldsymbol{\beta}^{true} = (1, 2, 1, 0, 0, 2, 0, 0, 0)'.$$

We simulated y_i^* as follows:

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

where $\varepsilon_i \sim$ Normal $(0,\ \sigma^2)$, and $\sigma^2 \in \{1, 4, 9\}$.

We list the results of regression coefficients estimates as tables below

$\boldsymbol{\beta}^{true}$	BALTR	BLTR	Tr	Btr	Btqr
$\boldsymbol{\beta_0 = 1}$	0.97159	0.97221	0.98196	0.94891	0.93512
$\boldsymbol{\beta_1 = 2}$	2.00872	2.00722	2.00367	2.03296	2.00367
$\boldsymbol{\beta_2 = 1}$	1.00763	1.00670	1.00807	1.01861	1.00807
$\boldsymbol{\beta_3 = 0}$	0.03387	0.03542	0.03594	0.03729	0.05146
$\boldsymbol{\beta_4 = 0}$	-0.02673	-0.02681	-0.02934	-0.02641	-0.02389
$\boldsymbol{\beta_5 = 2}$	2.03684	2.03427	2.03763	2.04875	2.05431
$\boldsymbol{\beta_6 = 0}$	-0.01184	-0.01180	-0.01395	-0.00831	-0.01093
$\boldsymbol{\beta_7 = 0}$	0.01131	0.01112	0.01176	0.01293	0.01731
$\boldsymbol{\beta_8 = 0}$	0.02385	0.02455	0.02422	0.02653	0.03331

Table 5: The coefficients estimates of Simulation example 2, when $\varepsilon_i \sim$ N(0,1)

$\boldsymbol{\beta}^{true}$	BALTR	BLTR	Tr	Btr	Btqr
$\boldsymbol{\beta_0 = 1}$	0.97095	0.97731	1.00273	0.89674	0.90440
$\boldsymbol{\beta_1 = 2}$	1.99212	1.98133	1.98334	2.04186	2.02277
$\boldsymbol{\beta_2 = 1}$	0.99698	0.99874	1.00515	1.03967	1.03717
$\boldsymbol{\beta_3 = 0}$	0.01277	0.01409	0.01603	0.00550	-0.01092
$\boldsymbol{\beta_4 = 0}$	0.00038	-0.00033	-0.00220	0.00672	-0.01011
$\boldsymbol{\beta_5 = 2}$	2.01144	1.99696	2.00244	2.06962	2.08351
$\boldsymbol{\beta_6 = 0}$	-0.01746	-0.01740	-0.01891	-0.02984	-0.03428
$\boldsymbol{\beta_7 = 0}$	0.00332	0.00659	0.00527	0.00489	0.02098
$\boldsymbol{\beta_8 = 0}$	-0.01538	-0.01637	-0.01637	-0.01705	-0.03639

Table 6: The coefficients estimates of Simulation example 2, when $\varepsilon_i \sim$ N(0,4)

β^{true}	BALTR	BLTR	Tr	Btr	Btqr
$\beta_0 = 1$	0.88552	0.91542	0.94838	0.73714	0.79963
$\beta_1 = 2$	2.09478	2.06696	2.09071	2.19900	2.16909
$\beta_2 = 1$	0.95556	0.95517	1.00633	1.07622	1.09604
$\beta_3 = 0$	-0.01773	-0.02413	-0.01887	-0.00588	-0.01590
$\beta_4 = 0$	0.06045	0.06462	0.06589	0.06675	0.07159
$\beta_5 = 2$	2.02103	1.98286	2.02437	2.14770	2.15143
$\beta_6 = 0$	-0.04312	-0.04821	-0.04970	-0.05432	-0.04958
$\beta_7 = 0$	0.01631	0.01681	0.02267	0.03698	0.01310
$\beta_8 = 0$	-0.02775	-0.02589	-0.02533	-0.02832	-0.03070

Table 7: The coefficients estimates of Simulation example 2, when $\varepsilon_i \sim \mathrm{N}(0,9)$

Procedure	σ^2	MMAD	SD
BALTR	1	0.2867665	0.0784412
BLTR		0.2878924	0.0794854
Tr		0.2886266	0.0804479
BTr		0.3030122	0.0875418
BTqr		0.3292688	0.0884608
BALTR	4	0.5291207	0.1342116
BLTR		0.5317522	0.1358630
Tr		0.5358101	0.1419345
BTr		0.5736709	0.1478548
BTqr		0.6248725	0.1717885
BALTR	9	0.8161338	0.2344180
BLTR		0.8247406	0.2504589
Tr		0.8639433	0.2594731
BTr		0.9450992	0.3328603
BTqr		0.9945814	0.3460455

Table 8: The MMADs and SDs results of Simulation example 2

4.2.3 Simulation example 3:

In this simulation example, we create 7 independent variables with 200 observations, the pair wise correlation between each independent variable equalizes to $0.5^{|i-j|}$, and we set the true regression coefficients as follows:

$$\boldsymbol{\beta}^{true} = (1, 6, 0, 0, 0, 0, 0, 0)'.$$

We simulated y_i^* as follows:

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

where $\varepsilon_i \sim$ Normal $(0,\ \sigma^2)$, and $\sigma^2 \in \{1, 4, 9\}$.

We list the results of regression coefficients estimates as tables below

$\boldsymbol{\beta}^{true}$	BALTR	BLTR	Tr	Btr	Btqr
$\boldsymbol{\beta_0 = 1}$	0.97368	0.97753	0.98879	0.94602	0.94129
$\boldsymbol{\beta_1 = 6}$	6.01734	6.00915	6.00311	6.04569	6.06105
$\boldsymbol{\beta_2 = 0}$	0.01687	0.02028	0.02032	0.01908	0.00891
$\boldsymbol{\beta_3 = 0}$	-0.02981	-0.03250	-0.03228	-0.03437	-0.03492
$\boldsymbol{\beta_4 = 0}$	0.01732	0.01819	0.01768	0.01843	0.01383
$\boldsymbol{\beta_5 = 0}$	-0.00262	-0.00358	-0.00503	0.00219	0.00320
$\boldsymbol{\beta_6 = 0}$	-0.00378	-0.00401	-0.00218	-0.00925	-0.00812
$\boldsymbol{\beta_7 = 0}$	-0.00518	-0.00384	-0.00430	-0.00595	-0.01012

Table 9: The coefficients estimates of Simulation example 3, when $\varepsilon_i \sim \mathrm{N}(0,1)$

$\boldsymbol{\beta}^{true}$	BALTR	BLTR	Tr	Btr	Btqr
$\boldsymbol{\beta_0 = 1}$	0.94907	0.96559	1.00646	0.88484	0.89730
$\boldsymbol{\beta_1 = 6}$	6.02860	5.99982	5.97664	6.10380	6.09037
$\boldsymbol{\beta_2 = 0}$	-0.02884	-0.03165	-0.03466	-0.02816	-0.01705
$\boldsymbol{\beta_3 = 0}$	0.02693	0.02488	0.02954	0.01865	0.01067
$\boldsymbol{\beta_4 = 0}$	-0.02435	-0.02439	-0.02528	-0.03082	-0.00979
$\boldsymbol{\beta_5 = 0}$	-0.00439	-0.00282	-0.00963	0.01587	0.01298
$\boldsymbol{\beta_6 = 0}$	0.03007	0.03157	0.04076	0.02071	-0.00122
$\boldsymbol{\beta_7 = 0}$	-0.00213	0.00071	0.00041	0.00368	0.02046

Table 10: The coefficients estimates of Simulation example 3, when $\varepsilon_i \sim \mathrm{N}(0,4)$

β^{true}	BALTR	BLTR	Tr	Btr	Btqr
$\beta_0 = 1$	0.85252	0.90898	0.97288	0.78279	0.78323
$\beta_1 = 6$	6.12307	6.04103	6.01833	6.21566	6.25580
$\beta_2 = 0$	-0.00169	-0.01736	-0.02162	0.00461	0.02084
$\beta_3 = 0$	-0.00479	-0.00091	0.00215	-0.02989	-0.04074
$\beta_4 = 0$	0.01931	0.02467	0.02528	0.03211	0.03718
$\beta_5 = 0$	-0.03964	-0.03964	-0.03971	-0.06806	-0.08469
$\beta_6 = 0$	-0.01256	-0.00415	-0.00234	-0.01975	-0.01429
$\beta_7 = 0$	-0.00448	-0.006959	-0.01159	0.00913	0.01829

Table 11: The coefficients estimates of Simulation example 3, when $\varepsilon_i \sim \text{N}(0,9)$

Procedure	σ^2	MMAD	SD
BALTR	1	0.3106147	0.1013127
BLTR		0.3115898	0.1060711
Tr		0.3121467	0.1046592
BTr		0.3161041	0.1115924
BTqr		0.3446810	0.1168242
BALTR	4	0.5711975	0.1695946
BLTR		0.5864465	0.1737943
Tr		0.5941101	0.1866304
BTr		0.6330689	0.1966508
BTqr		0.6741430	0.2176665
BALTR	9	0.8190953	0.2743333
BLTR		0.8360951	0.2539165
Tr		0.8664098	0.2966426
BTr		0.8928808	0.3108317
BTqr		0.9693136	0.3371386

Table 12: The MMADs and SDs results of Simulation example 3

4.2.4 Simulation example 4:

In this simulation example, we create 8 independent variables with 100 observations, the pair wise correlation between each independent variable equalizes to 0,9, and we set the true regression coefficients as follows:

$$\boldsymbol{\beta}^{true} = (1, 4, 0, 0, 0, 0, 0, 0, 0)'.$$

We simulated y_i^* as follows:

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

where $\varepsilon_i \sim$ Normal $(0,\ \sigma^2)$, and $\sigma^2 \in \{1, 4, 9\}$.

We list the results of regression coefficients estimates as tables below

$\boldsymbol{\beta}^{true}$	BALTR	BLTR	Tr	Btr	Btqr
$\boldsymbol{\beta_0} = \mathbf{1}$	0.98066	0.98114	0.99732	0.94336	0.94509
$\boldsymbol{\beta_1} = \mathbf{4}$	4.01580	4.01280	4.00121	4.05204	4.05518
$\boldsymbol{\beta_2} = \mathbf{0}$	-0.02369	-0.02443	-0.02401	-0.02597	-0.02490
$\boldsymbol{\beta_3} = \mathbf{0}$	0.00661	0.00656	0.00627	0.00755	-0.00942
$\boldsymbol{\beta_4} = \mathbf{0}$	-0.01306	-0.01326	-0.01122	-0.01811	-0.01485
$\boldsymbol{\beta_5} = \mathbf{0}$	-0.01593	-0.01743	-0.01530	-0.02417	-0.02233
$\boldsymbol{\beta_6} = \mathbf{0}$	-0.01972	-0.02033	-0.01952	-0.02484	-0.02248
$\boldsymbol{\beta_7} = \mathbf{0}$	0.00007	-0.00030	-0.00002	-0.00108	-0.00336
$\boldsymbol{\beta_8} = \mathbf{0}$	0.00155	0.00184	0.00259	0.00056	0.00530

Table 13: The coefficients estimates of Simulation example 4, when $\varepsilon_i \sim$ N(0,1)

$\boldsymbol{\beta}^{true}$	BALTR	BLTR	Tr	Btr	Btqr
$\boldsymbol{\beta_0} = \mathbf{1}$	0.93806	0.94796	0.98354	0.85672	0.87234
$\boldsymbol{\beta_1} = \mathbf{4}$	4.00038	3.97771	3.95982	4.09213	4.09190
$\boldsymbol{\beta_2} = \mathbf{0}$	-0.05298	-0.05676	-0.05471	-0.06522	-0.05747
$\boldsymbol{\beta_3} = \mathbf{0}$	0.01241	0.01078	0.01139	0.01108	0.01751
$\boldsymbol{\beta_4} = \mathbf{0}$	0.01391	0.01499	0.01537	0.01834	0.01579
$\boldsymbol{\beta_5} = \mathbf{0}$	-0.00205	-0.00121	-0.00191	-0.00161	-0.00616
$\boldsymbol{\beta_6} = \mathbf{0}$	0.00512	0.00353	0.00019	0.01596	0.02037
$\boldsymbol{\beta_7} = \mathbf{0}$	-0.00527	-0.00625	-0.00654	-0.00443	0.00298
$\boldsymbol{\beta_8} = \mathbf{0}$	0.02382	0.02719	0.02781	0.02752	0.01539

Table 14: The coefficients estimates of Simulation example 4, when $\varepsilon_i \sim$ N(0,4)

β^{true}	BALTR	BLTR	Tr	Btr	Btqr
$\beta_0 = 1$	0.85857	0.87405	0.93708	0.76538	0.79680
$\beta_1 = 4$	4.06267	4.03328	4.02374	4.15795	4.14522
$\beta_2 = 0$	0.01158	0.01511	0.01580	0.02550	0.02188
$\beta_3 = 0$	-0.04653	-0.04888	-0.05147	-0.05095	-0.05048
$\beta_4 = 0$	-0.00739	-0.00989	-0.01354	0.00265	0.01444
$\beta_5 = 0$	-0.02513	-0.02520	-0.02375	-0.03634	-0.03140
$\beta_6 = 0$	0.02400	0.01936	0.01766	0.01341	0.01714
$\beta_7 = 0$	0.01162	0.01649	0.01981	0.01195	0.00986
$\beta_8 = 0$	0.01130	0.00759	0.00732	0.01839	0.01591

Table 15: The coefficients estimates of Simulation example 4, when $\varepsilon_i \sim$ N(0,9)

Procedure	σ^2	MMAD	SD
BALTR	1	0.3175153	0.0878509
BLTR		0.3188315	0.0883490
Tr		0.3241813	0.0892745
BTr		0.3317360	0.0943799
BTqr		0.3483748	0.0999571
BALTR	4	0.5533796	0.1576824
BLTR		0.5586609	0.1638874
Tr		0.5673092	0.1745200
BTr		0.5902533	0.1929422
BTqr		0.6342843	0.2009979
BALTR	9	0.7782545	0.2106083
BLTR		0.7823382	0.2187003
Tr		0.8064242	0.2196771
BTr		0.8462888	0.2480293
BTqr		0.8727480	0.2616631

Table 16: The MMADs and SDs results of Simulation example 4

4.2.5 Simulation example 5:

In this example, we set the same number of observations and same true regression coefficients in the example above. But, the pair wise correlation between x_i and $x_j$is high correlation and it is equal to 0,75.

We list the results of regression coefficients estimates as tables below

β^{true}	BALTR	BLTR	Tr	Btr	Btqr
$\beta_0 = 1$	0.96334	0.97432	0.98914	0.93030	0.93054
$\beta_1 = 4$	3.99911	3.97485	3.96787	4.03959	4.04445
$\beta_2 = 0$	0.02810	0.02907	0.02915	0.03780	0.05102
$\beta_3 = 0$	-0.05722	-0.05931	-0.05809	-0.07588	-0.09464
$\beta_4 = 0$	0.03286	0.03630	0.03240	0.05007	0.05885
$\beta_5 = 0$	0.02766	0.02966	0.03094	0.02647	0.02646
$\beta_6 = 0$	-0.03771	-0.04115	-0.03994	-0.04496	-0.05025
$\beta_7 = 0$	0.00514	0.00745	0.00778	0.01035	0.02153
$\beta_8 = 0$	0.01061	0.01078	0.01007	0.01590	0.01057

Table 17: The coefficients estimates of Simulation example 5, when $\varepsilon_i \sim \mathrm{N}(0,1)$

β^{true}	BALTR	BLTR	Tr	Btr	Btqr
$\beta_0 = 1$	0.90221	0.92680	0.98491	0.85950	0.87261
$\beta_1 = 4$	4.09041	4.05841	4.03304	4.14416	4.14184
$\beta_2 = 0$	-0.02764	-0.03752	-0.03607	-0.04203	-0.05109
$\beta_3 = 0$	-0.03041	-0.02766	-0.03004	-0.03646	-0.04094
$\beta_4 = 0$	0.08249	0.07324	0.06840	0.10042	0.13265
$\beta_5 = 0$	-0.02687	-0.03145	-0.02989	-0.04422	-0.06207
$\beta_6 = 0$	0.01350	0.00990	0.00329	0.01418	0.02799
$\beta_7 = 0$	-0.03645	-0.03800	-0.03699	-0.05141	-0.06235
$\beta_8 = 0$	0.04680	0.04654	0.04803	0.05217	0.06563

Table 18: The coefficients estimates of Simulation example 5, when $\varepsilon_i \sim \mathrm{N}(0,4)$

β^{true}	BALTR	BLTR	Tr	Btr	Btqr
$\beta_0 = 1$	0.73119	0.78657	0.94885	0.75659	0.80464
$\beta_1 = 4$	4.12586	4.03465	3.99195	4.14338	4.15168
$\beta_2 = 0$	0.02199	0.02505	0.04213	0.04481	0.00129
$\beta_3 = 0$	-0.02119	-0.04096	-0.07182	-0.04829	-0.00941
$\beta_4 = 0$	0.02292	0.04250	0.06801	0.04011	0.02100
$\beta_5 = 0$	0.05144	0.04283	0.02239	0.04867	0.07637
$\beta_6 = 0$	-0.04245	-0.05081	-0.03979	-0.05118	-0.06790
$\beta_7 = 0$	0.01408	0.00877	-0.00874	0.00103	0.03011
$\beta_8 = 0$	0.73119	0.78657	0.94885	0.75659	0.80464

Table 19: The coefficients estimates of Simulation example 5, when $\varepsilon_i \sim \mathrm{N}(0,9)$

Procedure	σ^2	MMAD	SD
BALTR	1	0.3143854	0.1002000
BLTR		0.3157433	0.1004268
Tr		0.3249725	0.1007086
BTr		0.3270238	0.1140092
BTqr		0.3693706	0.1202663
BALTR	4	0.5635968	0.1528850
BLTR		0.5690248	0.1624616
Tr		0.5784414	0.1919185
BTr		0.5923137	0.1807643
BTqr		0.6501139	0.1794833
BALTR	9	0.7290041	0.2281853
BLTR		0.7558415	0.2599720
Tr		0.8092581	0.2846500
BTr		0.8406544	0.2806326
BTqr		0.8931200	0.2883510

Table 20: The MMADs and SDs results of Simulation example 5

4.2.6 Simulation example 6:

In this simulation example, we create 7 independent variables with 200 observations, the pair wise correlation between each independent variable equalize to $0.5^{|i-j|}$, and we set the true regression coefficients as follows:

$$\boldsymbol{\beta}^{true} = (0, \underbrace{0.7, \dots, 0.7}_{7})'.$$

We simulated y_i^* as follows:

$$y_i^* = 0.7x_{1i} + 0.7x_{2i} + 0.7x_{3i} + 0.7x_{4i} + 0.7x_{5i} + 0.7x_{6i} + 0.7x_{7i} + \varepsilon_i$$

where $\varepsilon_i \sim$ Normal $(0,\ \sigma^2)$, and $\sigma^2 \in \{1, 4, 9\}$.

We list the results of regression coefficients estimates as tables below

$\boldsymbol{\beta}^{true}$	BALTR	BLTR	Tr	Btr	Btqr
$\boldsymbol{\beta_0} = \mathbf{0}$	-0.00121	-0.00423	-0.00533	-0.08591	-0.06556
$\boldsymbol{\beta_1} = \mathbf{0.7}$	0.70947	0.71037	0.71166	0.73317	0.73406
$\boldsymbol{\beta_2} = \mathbf{0.7}$	0.69885	0.69889	0.69910	0.70385	0.69077
$\boldsymbol{\beta_3} = \mathbf{0.7}$	0.70073	0.70289	0.70702	0.72777	0.73125
$\boldsymbol{\beta_4} = \mathbf{0.7}$	0.67094	0.67292	0.67990	0.69661	0.69749
$\boldsymbol{\beta_5} = \mathbf{0.7}$	0.68911	0.68964	0.69058	0.69096	0.68514
$\boldsymbol{\beta_6} = \mathbf{0.7}$	0.69695	0.69789	0.69921	0.72360	0.71674
$\boldsymbol{\beta_7} = \mathbf{0.7}$	0.69993	0.70086	0.70014	0.72937	0.72692

Table 21: The coefficients estimates of Simulation example 6, when $\varepsilon_i \sim$ N(0,1)

$\boldsymbol{\beta}^{true}$	BALTR	BLTR	Tr	Btr	Btqr
$\boldsymbol{\beta_0} = \mathbf{0}$	0.01359	-0.00959	-0.02789	-0.20080	-0.15736
$\boldsymbol{\beta_1} = \mathbf{0.7}$	0.67988	0.68716	0.70428	0.72475	0.71989
$\boldsymbol{\beta_2} = \mathbf{0.7}$	0.68443	0.69440	0.71776	0.74273	0.73490
$\boldsymbol{\beta_3} = \mathbf{0.7}$	0.68594	0.70291	0.73325	0.78796	0.79404
$\boldsymbol{\beta_4} = \mathbf{0.7}$	0.70081	0.67994	0.70662	0.72362	0.73400
$\boldsymbol{\beta_5} = \mathbf{0.7}$	0.69904	0.68332	0.71282	0.76138	0.73504
$\boldsymbol{\beta_6} = \mathbf{0.7}$	0.68678	0.69480	0.70902	0.73962	0.76310
$\boldsymbol{\beta_7} = \mathbf{0.7}$	0.69172	0.67046	0.68789	0.71432	0.69523

Table 22: The coefficients estimates of Simulation example 6, when $\varepsilon_i \sim$ N(0,4)

β^{true}	BALTR	BLTR	Tr	Btr	Btqr
$\beta_0 = 0$	-0.05573	-0.03550	-0.07786	-0.35636	-0.28588
$\beta_1 = 0.7$	0.67484	0.66341	0.70466	0.76623	0.77236
$\beta_2 = 0.7$	0.66696	0.64567	0.74590	0.81375	0.83822
$\beta_3 = 0.7$	0.69203	0.68537	0.69142	0.72902	0.69549
$\beta_4 = 0.7$	0.70936	0.61233	0.69836	0.75535	0.76538
$\beta_5 = 0.7$	0.63479	0.61751	0.69721	0.75672	0.76575
$\beta_6 = 0.7$	0.74763	0.73993	0.77888	0.83389	0.84087
$\beta_7 = 0.7$	0.68780	0.66230	0.65789	0.69270	0.66976

Table 23: The coefficients estimates of Simulation example 6, when $\varepsilon_i \sim \mathrm{N}(0,9)$

Procedure	σ^2	MMAD	SD
BALTR	1	0.276327	0.104663
BLTR		0.277788	0.105090
Tr		0.280375	0.106778
BTr		0.286062	0.127102
BTqr		0.305961	0.126867
BALTR	4	0.545705	0.137516
BLTR		0.545787	0.141699
Tr		0.549043	0.154127
BTr		0.612324	0.198907
BTqr		0.626475	0.200574
BALTR	9	0.785711	0.230850
BLTR		0.805738	0.253614
Tr		0.814186	0.285252
BTr		0.906358	0.377037
BTqr		0.946205	0.386729

Table 24: The MMADs and SDs results of Simulation example 6

4.2.7 Simulation example 7:

This example considers a difficult case model. We create 4 independent variables with 100 observations, and the pair wise correlation between each independent variable is low correlation, it is equalized to -0.25**,** and we set the true regression coefficients as follows:

$$\boldsymbol{\beta}^{\boldsymbol{true}} = (0, 6.8, 6.8, 6.8, 0)'.$$

And we simulated y_i^* as follows:

$$y_i^* = 6.8x_{1i} + 6.8x_{2i} + 6.8x_{3i} + \varepsilon_i$$

where $\varepsilon_i \sim$ Normal $(0,\ \sigma^2)$, and $\sigma^2 \in \{1, 4, 9\}$.

We list the results of regression coefficients estimates as tables below

$\boldsymbol{\beta}^{\boldsymbol{true}}$	**BALTR**	**BLTR**	**Tr**	**Btr**	**Btqr**
$\boldsymbol{\beta_0 = 0}$	-0.03204	-0.02874	-0.02564	-0.04972	-0.04884
$\boldsymbol{\beta_1 = 6.8}$	6.81289	6.80855	6.80866	6.82615	6.82673
$\boldsymbol{\beta_2 = 6.8}$	6.82236	6.81757	6.81833	6.83434	6.82800
$\boldsymbol{\beta_3 = 6.8}$	6.80231	6.79864	6.79951	6.81096	6.79943
$\boldsymbol{\beta_4 = 0}$	-0.00738	-0.00849	-0.00822	-0.00876	-0.01579

Table 25: The coefficients estimates of Simulation example 7, when $\varepsilon_i \sim$ N(0,1)

$\boldsymbol{\beta}^{\boldsymbol{true}}$	**BALTR**	**BLTR**	**Tr**	**Btr**	**Btqr**
$\boldsymbol{\beta_0 = 0}$	-0.01848	-0.01769	-0.00053	-0.07119	-0.05421
$\boldsymbol{\beta_1 = 6.8}$	6.80706	6.80504	6.79841	6.83788	6.83524
$\boldsymbol{\beta_2 = 6.8}$	6.78025	6.77714	6.77373	6.80896	6.79998
$\boldsymbol{\beta_3 = 6.8}$	6.81374	6.81031	6.80598	6.84394	6.84213
$\boldsymbol{\beta_4 = 0}$	-0.01473	-0.01726	-0.01741	-0.01320	-0.01883

Table 26: The coefficients estimates of Simulation example 7, when $\varepsilon_i \sim$ N(0,4)

β^{true}	BALTR	BLTR	Tr	Btr	Btqr
$\beta_0 = 0$	-0.05187	-0.05787	-0.00290	-0.20772	-0.22551
$\beta_1 = 6.8$	6.82950	6.82827	6.80091	6.94322	6.95459
$\beta_2 = 6.8$	6.88990	6.88623	6.86843	6.99837	7.01069
$\beta_3 = 6.8$	6.82117	6.81366	6.79564	6.93702	6.98893
$\beta_4 = 0$	0.03453	0.03309	0.03350	0.04149	0.04776

Table 27: The coefficients estimates of Simulation example 7, when $\varepsilon_i \sim \mathrm{N}(0,9)$

Procedure	σ^2	MMAD	SD
BALTR	1	0.2715830	0.1158086
BLTR		0.2739781	0.1161730
Tr		0.2748512	0.1167151
BTr		0.2755876	0.1208232
BTqr		0.2970783	0.1276834
BALTR	4	0.4730141	0.2025605
BLTR		0.4736110	0.2042301
Tr		0.4852631	0.2043655
BTr		0.4895753	0.2118048
BTqr		0.5720289	0.2340398
BALTR	9	0.7991411	0.3320651
BLTR		0.8110890	0.3479748
Tr		0.8323682	0.3412223
BTr		0.8268126	0.4119798
BTqr		0.9511592	0.4644230

Table 28: The MMADs and SDs results of Simulation example 7

4.2.8 Simulation example 8:

This simulation example is same in the example above except the pair wise correlation between each independent variable is very low correlation, it is equalized to 0.95**,** and we set the true regression coefficients as follows:

$$\boldsymbol{\beta}^{true} = (0, 5.5, 5.5, 5.5, 0)'.$$

We simulated y_i^* as follows:

$$y_i^* = 5.5x_{1i} + 5.5x_{2i} + 5.5x_{3i} + \varepsilon_i$$

where $\varepsilon_i \sim$ Normal $(0,\ \sigma^2)$, and $\sigma^2 \in \{1, 4, 9\}$.

We list the results of regression coefficients estimates as tables below

$\boldsymbol{\beta}^{true}$	**BALTR**	**BLTR**	**Tr**	**Btr**	**Btqr**
$\boldsymbol{\beta_0 = 0}$	0.00060	0.00606	0.01219	-0.02171	-0.02990
$\boldsymbol{\beta_1 = 5.5}$	5.49903	5.49370	5.49235	5.51362	5.51767
$\boldsymbol{\beta_2 = 5.5}$	5.49752	5.49176	5.49292	5.50951	5.50678
$\boldsymbol{\beta_3 = 5.5}$	5.51334	5.50702	5.50574	5.53024	5.53255
$\boldsymbol{\beta_4 = 0}$	-0.00655	-0.00587	-0.00761	-0.00458	-0.00813

Table 29: The coefficients estimates of Simulation example 8, when $\varepsilon_i \sim \text{N}(0,1)$

$\boldsymbol{\beta}^{true}$	**BALTR**	**BLTR**	**Tr**	**Btr**	**Btqr**
$\boldsymbol{\beta_0 = 0}$	-0.02741	-0.02886	-0.05559	-0.09647	-0.09765
$\boldsymbol{\beta_1 = 5.5}$	5.53919	5.53667	5.52496	5.59940	5.60046
$\boldsymbol{\beta_2 = 5.5}$	5.52721	5.52524	5.50943	5.58874	5.59006
$\boldsymbol{\beta_3 = 5.5}$	5.54451	5.54064	5.53180	5.60197	5.62056
$\boldsymbol{\beta_4 = 0}$	0.00502	0.00462	0.00431	0.01049	0.00578

Table 30: The coefficients estimates of Simulation example 8, when $\varepsilon_i \sim \text{N}(0,4)$

β^{true}	BALTR	BLTR	Tr	Btr	Btqr
$\beta_0 = 0$	-0.12842	-0.14020	-0.18551	-0.28450	-0.25737
$\beta_1 = 5.5$	5.53337	5.53558	5.50877	5.64924	5.64082
$\beta_2 = 5.5$	5.50742	5.50453	5.48611	5.62610	5.64194
$\beta_3 = 5.5$	5.57582	5.57051	5.56494	5.67917	5.67718
$\beta_4 = 0$	-0.03505	-0.04467	-0.04179	-0.03981	-0.01707

Table 31: The coefficients estimates of Simulation example 8, when $\varepsilon_i \sim \mathrm{N}(0,9)$

Procedure	σ^2	MMAD	SD
BALTR	1	0.2693833	0.1276641
BLTR		0.2701359	0.1276692
Tr		0.2737591	0.1278145
BTr		0.2809201	0.1340808
BTqr		0.3205587	0.1549083
BALTR	4	0.4916656	0.2144122
BLTR		0.4948648	0.2147106
Tr		0.4962678	0.2168003
BTr		0.5145276	0.2295402
BTqr		0.5344612	0.2361655
BALTR	9	0.6875267	0.3454354
BLTR		0.7362656	0.3522602
Tr		0.7448768	0.3553119
BTr		0.7762767	0.4178073
BTqr		0.7977359	0.4015878

Table 32: The MMADs and SDs results of Simulation example 8

4.2.9 Simulation example 9:

In this simulation example, we create 7 independent variables with 200 observations, and without intercept. The pair wise correlation between each independent variable equalize to $0.8^{|i-j|}$, and we set the true regression coefficients as follows:

$$\boldsymbol{\beta}^{\boldsymbol{true}} = (1, 0, 0, 1, 0, 0, 1)'.$$

We simulated y_i^* as follows:

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

where $\varepsilon_i \sim$ Normal $(0,\ \sigma^2)$, and $\sigma^2 \in \{1, 4, 9\}$.

We list the results of regression coefficients estimates as tables below

$\boldsymbol{\beta}^{\boldsymbol{true}}$	**BALTR**	**BLTR**	**Tr**	**Btr**	**Btqr**
$\boldsymbol{\beta_1} = \mathbf{1}$	0.99602	0.99472	0.99891	0.99938	1.00005
$\boldsymbol{\beta_2} = \mathbf{0}$	0.02253	0.02361	0.02502	0.02872	0.02909
$\boldsymbol{\beta_3} = \mathbf{0}$	-0.01936	-0.02096	-0.02302	-0.01557	-0.01819
$\boldsymbol{\beta_4} = \mathbf{1}$	0.98914	0.98748	0.98879	0.99998	1.01144
$\boldsymbol{\beta_5} = \mathbf{0}$	0.00151	0.00218	0.00401	0.00070	-0.00791
$\boldsymbol{\beta_6} = \mathbf{0}$	0.00593	0.00778	0.00648	0.01289	0.01664
$\boldsymbol{\beta_7} = \mathbf{1}$	1.02899	1.02624	1.02538	1.05374	1.07164

Table 33: The coefficients estimates of Simulation example 9, when $\varepsilon_i \sim$ N(0,1)

$\boldsymbol{\beta}^{\boldsymbol{true}}$	**BALTR**	**BLTR**	**Tr**	**Btr**	**Btqr**
$\boldsymbol{\beta_1} = \mathbf{1}$	1.00121	0.99423	1.00942	1.05071	1.06561
$\boldsymbol{\beta_2} = \mathbf{0}$	-0.01434	-0.02126	-0.01914	-0.03293	-0.03984
$\boldsymbol{\beta_3} = \mathbf{0}$	-0.00154	-0.00022	-0.00004	-0.01048	-0.01597
$\boldsymbol{\beta_4} = \mathbf{1}$	1.04370	1.03332	1.05055	1.10838	1.11395
$\boldsymbol{\beta_5} = \mathbf{0}$	-0.02595	-0.02849	-0.02941	-0.03784	-0.02168
$\boldsymbol{\beta_6} = \mathbf{0}$	0.02130	0.02220	0.01712	0.04521	0.00745
$\boldsymbol{\beta_7} = \mathbf{1}$	0.98648	0.98243	0.99409	1.00535	1.03446

Table 34: The coefficients estimates of Simulation example 9, when $\varepsilon_i \sim$ N(0,4)

$\boldsymbol{\beta}^{true}$	BALTR	BLTR	Tr	Btr	Btqr
$\boldsymbol{\beta_1} = \mathbf{1}$	0.87722	0.87473	0.93722	0.93620	0.96776
$\boldsymbol{\beta_2} = \mathbf{0}$	0.02059	0.02356	0.01815	0.02717	0.04077
$\boldsymbol{\beta_3} = \mathbf{0}$	0.03436	0.03484	0.04302	0.02835	0.02563
$\boldsymbol{\beta_4} = \mathbf{1}$	0.99151	0.97907	1.03864	1.08294	1.06385
$\boldsymbol{\beta_5} = \mathbf{0}$	-0.02858	-0.03023	-0.03350	-0.00132	-0.01357
$\boldsymbol{\beta_6} = \mathbf{0}$	-0.00635	-0.01027	-0.01450	-0.03378	-0.05047
$\boldsymbol{\beta_7} = \mathbf{1}$	1.00904	0.99530	1.05497	1.09108	1.10440

Table 35: The coefficients estimates of Simulation example 9, when $\varepsilon_i \sim \mathrm{N}(0,9)$

Procedure	$\boldsymbol{\sigma^2}$	MMAD	SD
BALTR	1	0.2637977	0.0629574
BLTR		0.2645657	0.0630926
Tr		0.2661152	0.0641458
BTr		0.2690778	0.0672531
BTqr		0.3022255	0.0756868
BALTR	4	0.4451257	0.1621673
BLTR		0.4584184	0.1585630
Tr		0.4645096	0.1603979
BTr		0.4939741	0.1682316
BTqr		0.5428853	0.1769871
BALTR	9	0.6663666	0.1890329
BLTR		0.6716854	0.1890237
Tr		0.7004311	0.1977441
BTr		0.7115157	0.2022168
BTqr		0.7552406	0.2121890

Table 36: The MMADs and SDs results of Simulation example 9

4.2.10 Simulation example 10:

This simulation example is same as example 9 except we set the true regression coefficients as follows:

$$\boldsymbol{\beta}^{true} = (2, \underbrace{0, \dots, 0}_{6})'.$$

We simulated y_i^* as follows:

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

where $\varepsilon_i \sim$ Normal $(0, \sigma^2)$, and $\sigma^2 \in \{1, 4, 9\}$.

We list the results of regression coefficients estimates as tables below

$\boldsymbol{\beta}^{true}$	**BALTR**	**BLTR**	**Tr**	**Btr**	**Btqr**
$\boldsymbol{\beta_1} = \mathbf{2}$	2.02453	2.02054	2.02216	2.03512	2.03164
$\boldsymbol{\beta_2} = \mathbf{0}$	0.01362	0.01412	0.01344	0.01859	0.01118
$\boldsymbol{\beta_3} = \mathbf{0}$	0.00231	0.00285	0.00208	0.00654	0.00891
$\boldsymbol{\beta_4} = \mathbf{0}$	-0.01615	-0.01740	-0.01611	-0.02272	-0.01397
$\boldsymbol{\beta_5} = \mathbf{0}$	0.01777	0.01868	0.01845	0.02269	0.01890
$\boldsymbol{\beta_6} = \mathbf{0}$	0.00883	0.00852	0.00901	0.00715	0.00495
$\boldsymbol{\beta_7} = \mathbf{0}$	-0.00492	-0.00466	-0.00480	-0.00474	-0.00445

Table 37: The coefficients estimates of Simulation example 10, when $\varepsilon_i \sim \mathrm{N}(0,1)$

$\boldsymbol{\beta}^{true}$	**BALTR**	**BLTR**	**Tr**	**Btr**	**Btqr**
$\boldsymbol{\beta_1} = \mathbf{2}$	2.03589	2.01806	2.02225	2.09623	2.09136
$\boldsymbol{\beta_2} = \mathbf{0}$	-0.02655	-0.02897	-0.01920	-0.06953	-0.07637
$\boldsymbol{\beta_3} = \mathbf{0}$	0.01203	0.01314	0.00604	0.03564	0.03165
$\boldsymbol{\beta_4} = \mathbf{0}$	0.00093	0.00088	0.00063	-0.00408	-0.00881
$\boldsymbol{\beta_5} = \mathbf{0}$	-0.00516	-0.00285	0.00208	-0.02818	-0.01613
$\boldsymbol{\beta_6} = \mathbf{0}$	0.00599	0.00667	0.00192	0.03046	0.02646
$\boldsymbol{\beta_7} = \mathbf{0}$	-0.03617	-0.03757	-0.04135	-0.04638	-0.04298

Table 38: The coefficients estimates of Simulation example 10, when $\varepsilon_i \sim \mathrm{N}(0,4)$

β^{true}	BALTR	BLTR	Tr	Btr	Btqr
$\beta_1 = 2$	2.03086	1.98035	2.03613	2.12114	2.11584
$\beta_2 = 0$	-0.00503	-0.00731	-0.01374	-0.00790	-0.01854
$\beta_3 = 0$	0.04881	0.04794	0.06198	0.05704	0.06209
$\beta_4 = 0$	0.01883	0.01793	0.01211	0.06594	0.05424
$\beta_5 = 0$	-0.01938	-0.02196	-0.02338	-0.04635	-0.03894
$\beta_6 = 0$	-0.00672	-0.00744	-0.00609	-0.01412	-0.00675
$\beta_7 = 0$	0.00185	-0.00638	-0.00801	-0.00229	-0.00272

Table 39: The coefficients estimates of Simulation example 10, when $\varepsilon_i \sim \mathrm{N}(0,9)$

Procedure	σ^2	MMAD	SD
BALTR	1	0.2425656	0.0606486
BLTR		0.2432271	0.0698898
Tr		0.2444597	0.0700119
BTr		0.2493422	0.0699531
BTqr		0.2732725	0.0735678
BALTR	4	0.4527618	0.1528419
BLTR		0.4556374	0.1539392
Tr		0.4613127	0.1539059
BTr		0.4755329	0.1625885
BTqr		0.5232845	0.1674268
BALTR	9	0.6200592	0.2003216
BLTR		0.6468788	0.2013237
Tr		0.7054764	0.2016802
BTr		0.7411863	0.2152123
BTqr		0.7909137	0.2284264

Table 40: The MMADs and SDs results of Simulation example 10

4.3 Non-i.i.d (Heterogeneous)random errors:

4.3.1 Simulation example 11:

In this simulation example, we created 100 observations, and 8 independent variables, 5 of these variables are represented as standard normal noise variables, and we set the true regression coefficients as follows:

$$\boldsymbol{\beta}^{true} = (1, 2, 1, 2, 0, 0, 0, 0, 0)'$$

We simulated y_i^* as follows:

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

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

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

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

We list the results of regression coefficients estimates as tables below

$\boldsymbol{\beta}^{true}$	BALTR	BLTR	Tr	Btr	Btqr
$\boldsymbol{\beta_0 = 1}$	0.96936	0.96770	0.98601	0.96044	0.97442
$\boldsymbol{\beta_1 = 2}$	2.01515	2.01446	2.00962	2.02357	2.02289
$\boldsymbol{\beta_2 = 1}$	0.99955	1.00096	0.99772	1.00197	0.99635
$\boldsymbol{\beta_3 = 2}$	2.00402	1.98195	2.01909	2.02554	2.03448
$\boldsymbol{\beta_4 = 0}$	- 0.00316	-0.00300	-0.00252	-0.00489	0.00180
$\boldsymbol{\beta_5 = 0}$	-0.01780	-0.01785	-0.01846	-0.01910	-0.02701
$\boldsymbol{\beta_6 = 0}$	0.00675	0.00713	0.00685	0.00907	0.01659
$\boldsymbol{\beta_7 = 0}$	-0.01106	-0.01131	-0.01095	-0.01031	-0.00753
$\boldsymbol{\beta_8 = 0}$	0.00066	0.00075	0.00121	-0.00144	0.00883

Table 41: The coefficients estimates of Simulation example 11

Procedure	MMAD	SD
BALTR	0.2488675	0.0673076
BLTR	0.2489086	0.0676727
Tr	0.2498349	0.0679520
BTr	0.2498877	0.0689287
BTqr	0.2718577	0.0777695

Table 42: The MMADs and SDs results of Simulation example 11

4.3.2 Simulation example 12:

This simulation example is same as example 11, except the number of independent variables is 7, and without intercept, 4 variables are represented as standards normal noise variables. we set the true regression coefficients as follows:

$$\boldsymbol{\beta}^{true} = (1, 1, 0, 0, 0, 1, 1)'$$

We list the results of regression coefficients estimates as tables below

$\boldsymbol{\beta}^{true}$	BALTR	BLTR	Tr	Btr	Btqr
$\boldsymbol{\beta_1 = 1}$	0.97645	0.97437	0.98484	0.99867	1.02606
$\boldsymbol{\beta_2 = 1}$	0.98797	0.98689	0.98858	1.00582	1.00308
$\boldsymbol{\beta_3 = 0}$	0.07754	0.07911	0.09580	-0.00931	-0.05584
$\boldsymbol{\beta_4 = 0}$	-0.02608	-0.02539	-0.02589	-0.02436	-0.03266
$\boldsymbol{\beta_5 = 0}$	0.01822	0.02086	0.02213	0.02496	0.01569
$\boldsymbol{\beta_6 = 1}$	0.98689	0.98565	0.98829	0.99837	1.02980
$\boldsymbol{\beta_7 = 1}$	1.01277	1.01076	1.01569	1.02484	1.03444

Table 43: The coefficients estimates of Simulation example 12

Procedure	MMAD	SD
BALTR	0.3876545	0.1058635
BLTR	0.3880013	0.1059328
Tr	0.3898011	0.1064951
BTr	0.3975563	0.1119146
BTqr	0.4086063	0.1165248

Table 44: The MMADs and SDs results of Simulation example 12

From all tables of the previous examples, it can be observed that the proposed methods (BLTR and BALTR) generally perform better than the other methods (Tr, Btr, and Btqr) used in the comparisons. The MMADs results of BALTR approach were relatively less than these results of the other procedures and gives the best MMADs and SD most of the times. This indicates the quality of the performance of the BALTR procedure in terms of coefficient estimation and VS.

CHAPTER FIVE

5. PRACTICAL APPLICATION

5.1 Introduction:

In this part of application chapter, and after we have demonstrated the merit of our proposed methods in the simulation study, we applied our proposed methods to the real data and then analyze the results statistically. The identification and detection of the causes about increasing the rate wheat production are one of the priorities of agricultural economist researchers. So, the determination of the real factors to increase wheat crop production among several factors that will help us to predict the rate of increase wheat production in the future. Hence the importance of our new proposed methods, which attempt to identify some variables and to show how strong their impact on the rate of increase in wheat production. We make our methods BALTR and BLTR within the three processes in wheat production data for comparison in terms of accurate prediction and variable selection. The data used in this chapter is taken from the National program for the development of wheat cultivation in Iraq; Qadisiyah governorate board (2017). The wheat production data includes 11 variables within 584 observations, these variables are sorting as follows in table 45. The response variable in this dataset represents the relative increase in wheat yield per dunum, note that each dunum is equal to $2500\, m^2$. R code is available upon request.

5.2 The Independent variables:

- Urea fertilizer (U): The urea fertilizer is a simple fertilizer that supplies the major essential element nitrogen, and the crops need it in larger quantities than any other nutrient.
- Date of sowing (Ds): The date of cultivation of wheat seeds in the field.
- Quantity of sowing seeds (Qs): The quantity of wheat seeds in the field, and this amount is measured in kilograms per dunum.
- Laser field leveling technique (LT): This method is a smoothing procedure and leveling the farm ground. This method offers the potential for water savings.
- Compound fertilizer (NPK): NPK fertilizer is a complex fertilizer containing principally of the three fundamentals nutrients necessary for healthy plants growth (Nitrogen, Phosphorus and Potassium).
- Seed sowing machine technique (SMT): A sowing seed machine is a device that sows the seeds for crops in the soil, then cover the seeds to a nominated typical rate depth.
- Planting successive (SC): Successive planting is a way to extend corps harvest by staggering planting of crops, or planting varieties with staggered maturing dates. In this real dataset, the other corps which planting before sowing wheat seeds is Mung bean corp.
- Herbicide for weed (H): This process contains chemical applications. top control the growth of weeds types.
- High Potassium fertilizer (K): High Potassium is essential for crops health and there must be an adequate supply in the soils to maintain goods growth.

- Micro-Elements fertilizer (ME): Mineral elements that are needed by crops in only trace amounts are known as micro-elements. This fertilizer which helps the growth of crops, such as iron, magnesium, potassium, and manganese.

symbol	Variables description	Rank	Rank description
Y	The outcome variable	Percentage increase of wheat product	
U	Urea fertilizer	Numeral	Quantity in kilogram
Ds	Date of sowing	1	Ideal
		2	Early
		3	Late
Qs	Quantity of sowing	Numeral	Quantity in kilogram
LT	Laser field leveling	1	Unused
		2	Used
NPK	Compound fertilizer	Numeral	Quantity in kilogram
SMT	Sowing seeds machine	1	Unused
		2	Used
SC	Planting successive (Mung bean crop)	1	Planting
		2	Not planting
H	Herbicide for weeds	Numeral	Quantity in milliliter
K	High Potassium	Numeral	Quantity in kilogram
ME	Micro-Elements	Numeral	Quantity int gram

Table 45: The top 11 worthy variables

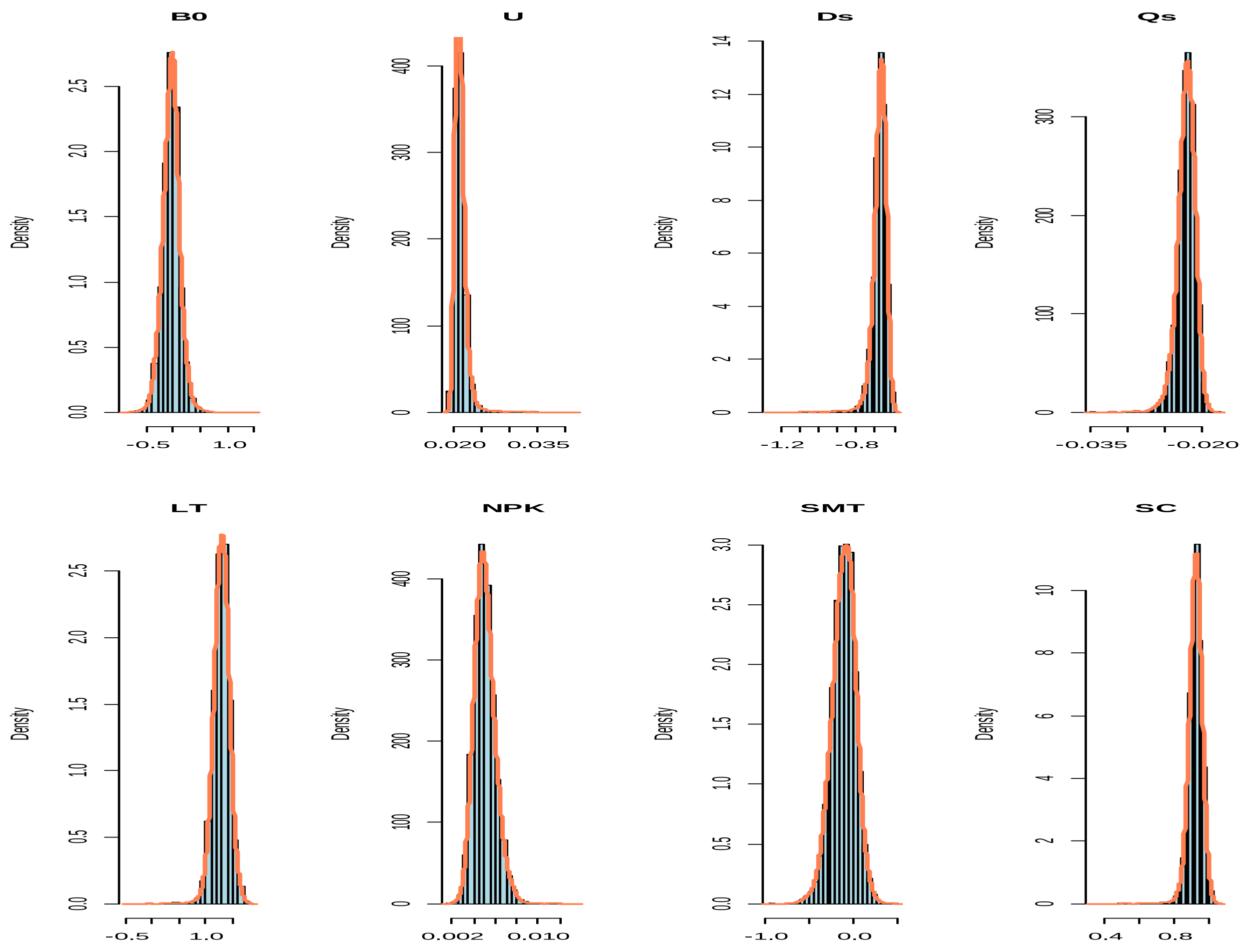
5.3 Real data results:

	BALTR	**BLTR**	**Tr**	**Btr**	**Btqr**
$\boldsymbol{\beta_0}$	-0.0433838	-0.0431762	-0.084922	-0.081976	-1.2196515
L.C I	-0.3578263	-0.3604089	-0.8722770	-0.898573	-1.8071276
U.C I	0.2323458	0.2673627	0.7024328	0.719634	-0.4907951

U	0.0211009	0.0213583	0.020928	0.020974	0.0243195
L.C I	0.0199047	0.0200037	0.0137744	0.013798	0.0176873
U.C I	0.0228502	0.0236142	0.0280822	0.028236	0.0307460
Ds	-0.6671660	-0.6763215	-0.663651	-0.665663	-0.6545110
L.C I	-0.7316255	-0.7574811	-0.7864428	-0.790510	-0.8024576
U.C I	-0.6207167	-0.6239916	-0.5408589	-0.546217	-0.5029025
Qs	-0.0218912	-0.0220591	-0.021788	-0.021790	-0.0057674
L.C I	-0.0242680	-0.0247427	-0.0347353	-0.034747	-0.0177361
U.C I	-0.0199373	-0.0200504	-0.0088414	-0.008437	0.0032729
LT	1.3392113	1.2890701	1.357168	1.357630	1.4284952
L.C I	1.0704020	0.9941205	0.6807916	0.657868	0.4566205
U.C I	1.6029090	1.5630652	2.0335436	2.035161	2.3610716
NPK	0.0049840	0.0050224	0.004908	0.004920	-0.0052290
L.C I	0.0033343	0.0033702	-0.0075619	-0.007545	-0.0164622
U.C I	0.0069746	0.0071492	0.0173782	0.017452	0.0068326
SMT	-0.0936879	-0.1100872	-0.142634	-0.145098	0.2445190
L.C I	-0.3627955	-0.3842552	-0.8377516	-0.839958	-0.6438307
U.C I	0.1158550	0.1310164	0.5524826	0.558794	1.2208572
SC	0.9282265	0.9198140	0.932829	0.930684	1.0000151
L.C I	0.8565598	0.8406293	0.6111322	0.600929	0.6683828
U.C I	0.9953362	0.9880197	1.2545265	1.258660	1.3113546
H	0.0043298	0.0043892	0.004315	0.004323	0.0051035
L.C I	0.0038758	0.0039010	0.0027357	0.002712	0.0037272
U.C I	0.0048939	0.0050331	0.0058946	0.005972	0.0064843
K	0.0327570	0.0328492	0.032676	0.032692	0.0246412
L.C I	0.0321197	0.0321597	0.0254615	0.025595	0.0139049
U.C I	0.0335991	0.0338707	0.0398907	0.039990	0.0360497
ME	0.0062279	0.0062152	0.006228	0.006250	0.0075853
L.C I	0.0060113	0.0059725	0.0045106	0.004517	0.0046838
U.C I	0.0063928	0.0063932	0.0079452	0.007948	0.0101096

Table 46 - Coefficients estimation and Credible intervals CIs (25%, 95%)

In table 46, the results showed the coefficients estimation and credible intervals (low credible interval L.CI and upper credible interval U.CI). The credible interval results of the proposed techniques BLTR and BALTR are narrower than Tr, BTr, BTqr methods, and our proposed methods are including all the estimations of other methods. At the same time, the results described above have shown that the proposed method BALTR is the best technique of all other techniques used.



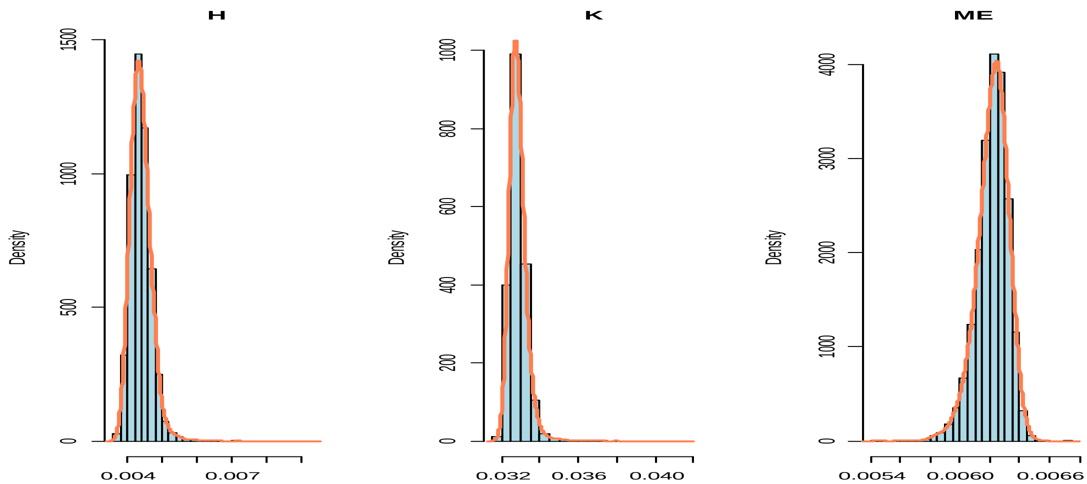
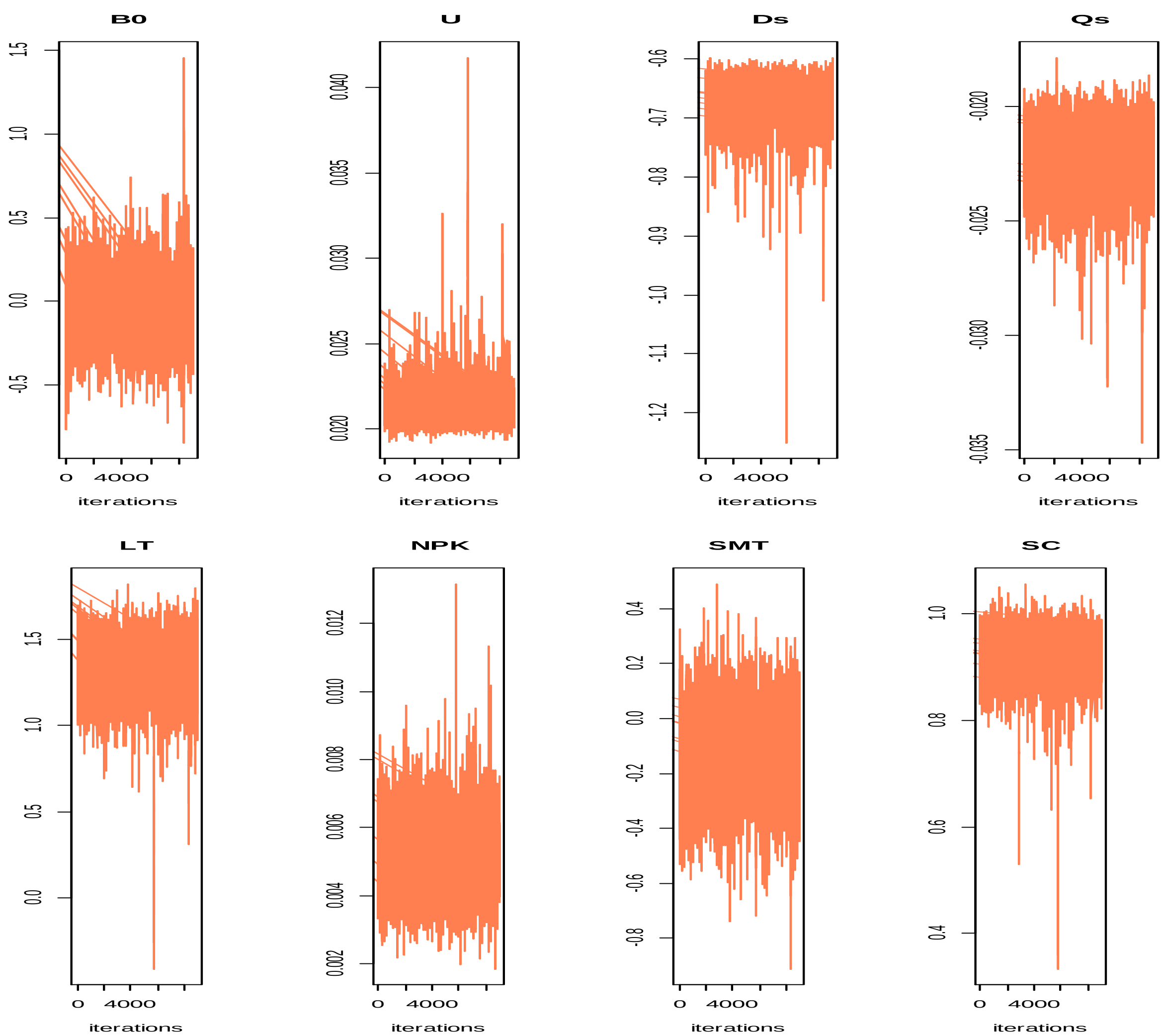


Figure 1- Histograms of BLTR coefficients estimation



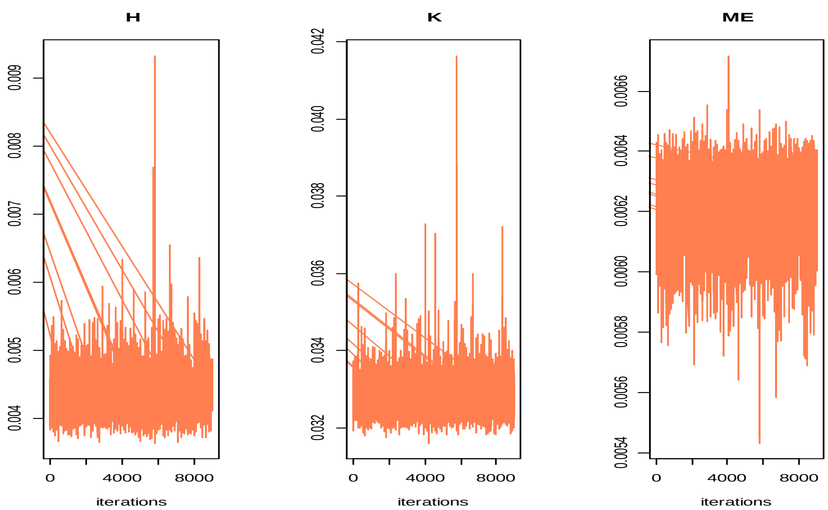
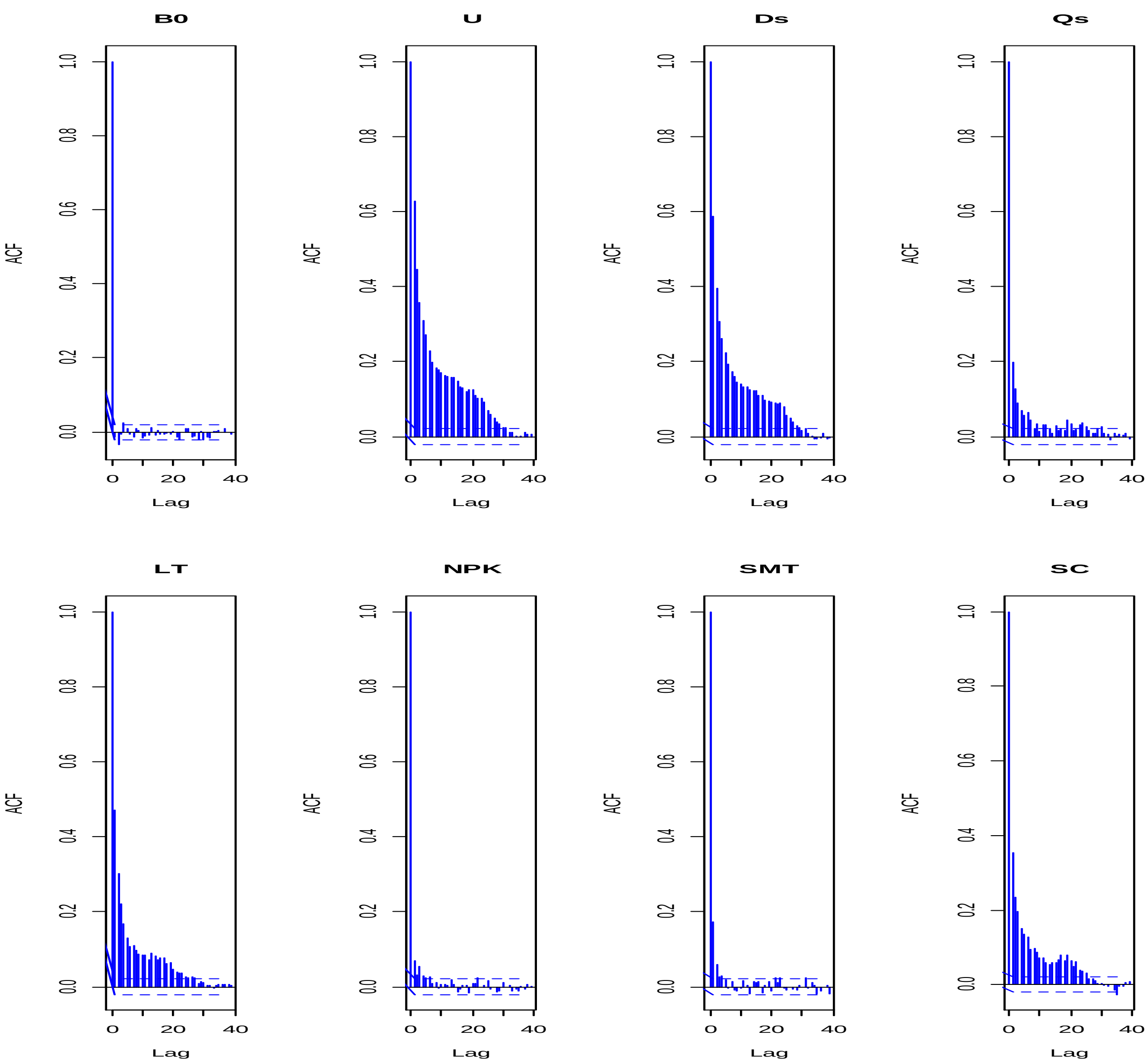


Figure 2- Trace plots of BLTR coefficients estimation



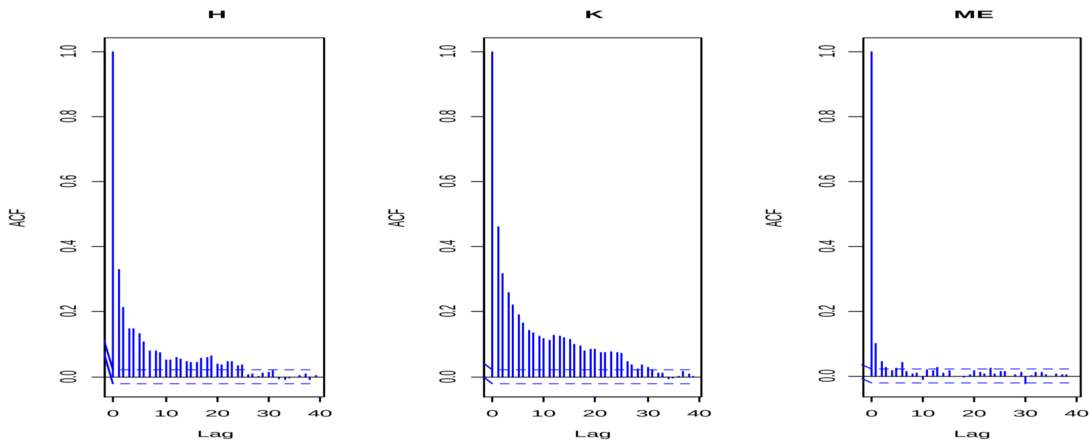
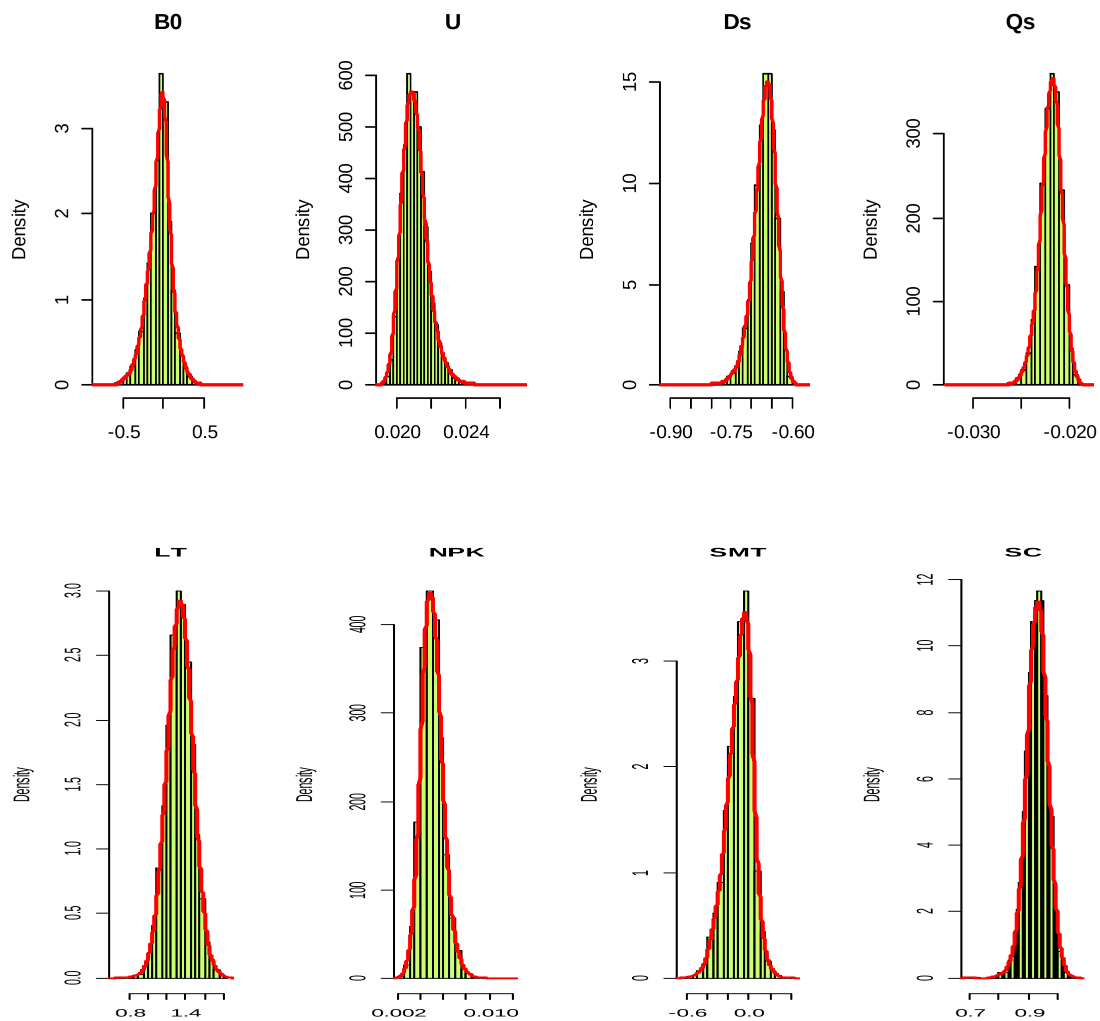


Figure 3- Autocorrelations of BLTR coefficients estimation



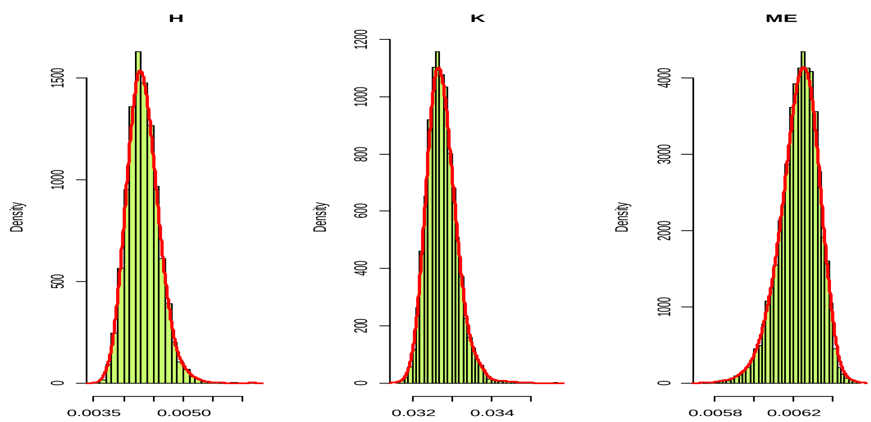
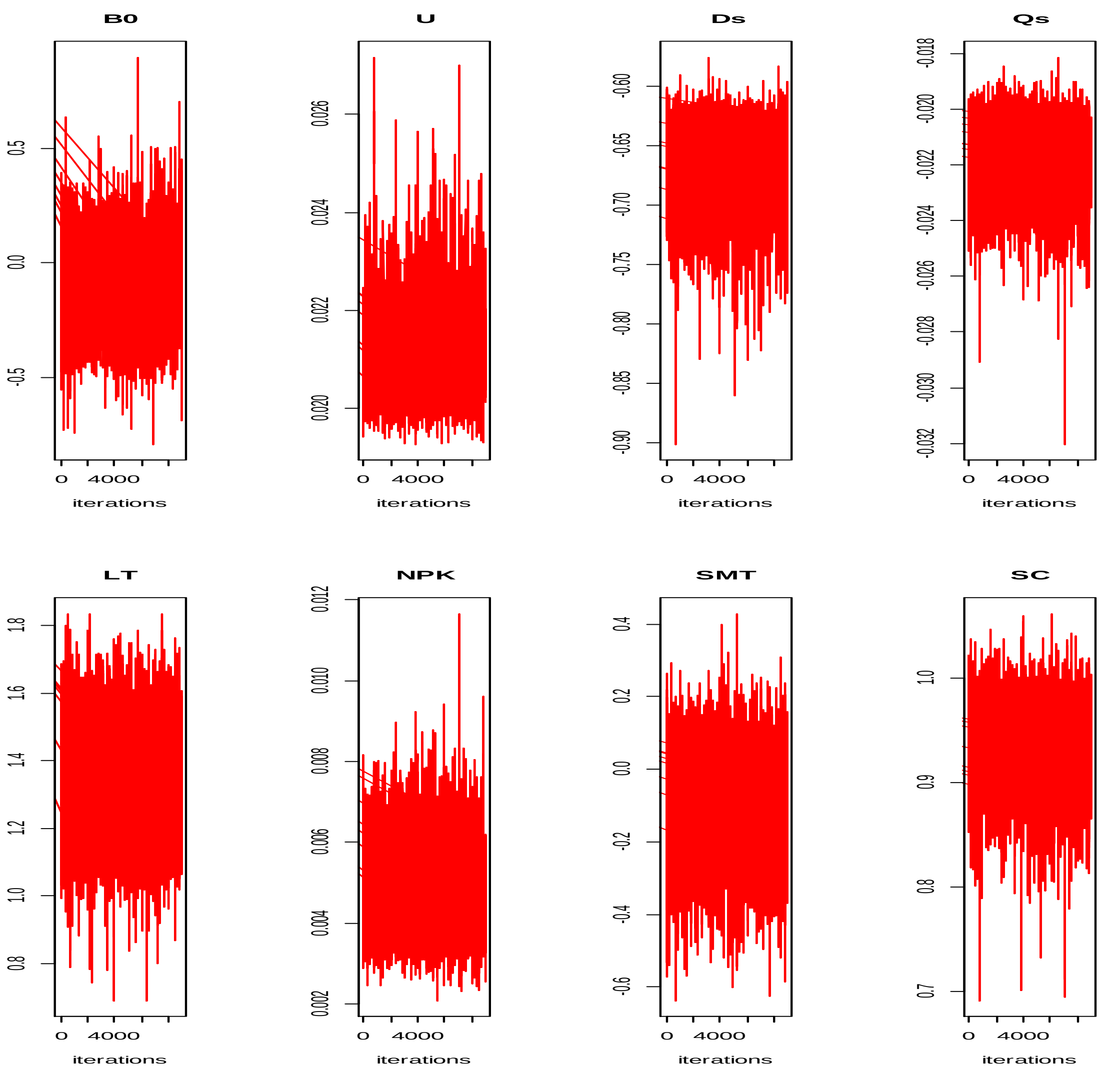


Figure 4- Histograms of BALTR coefficients estimation



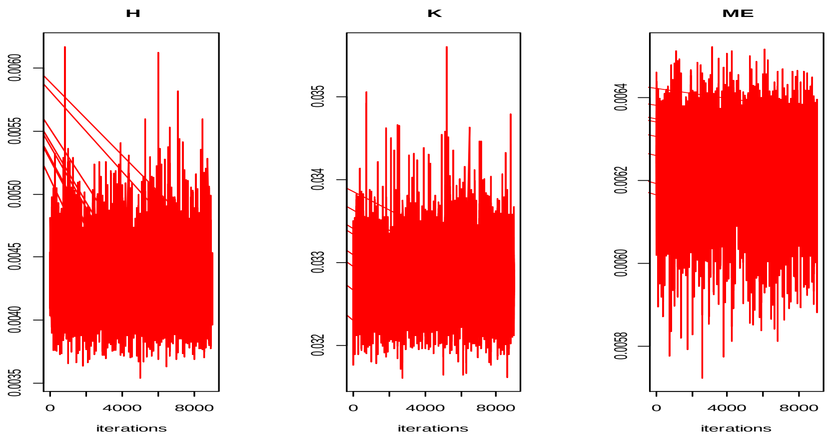
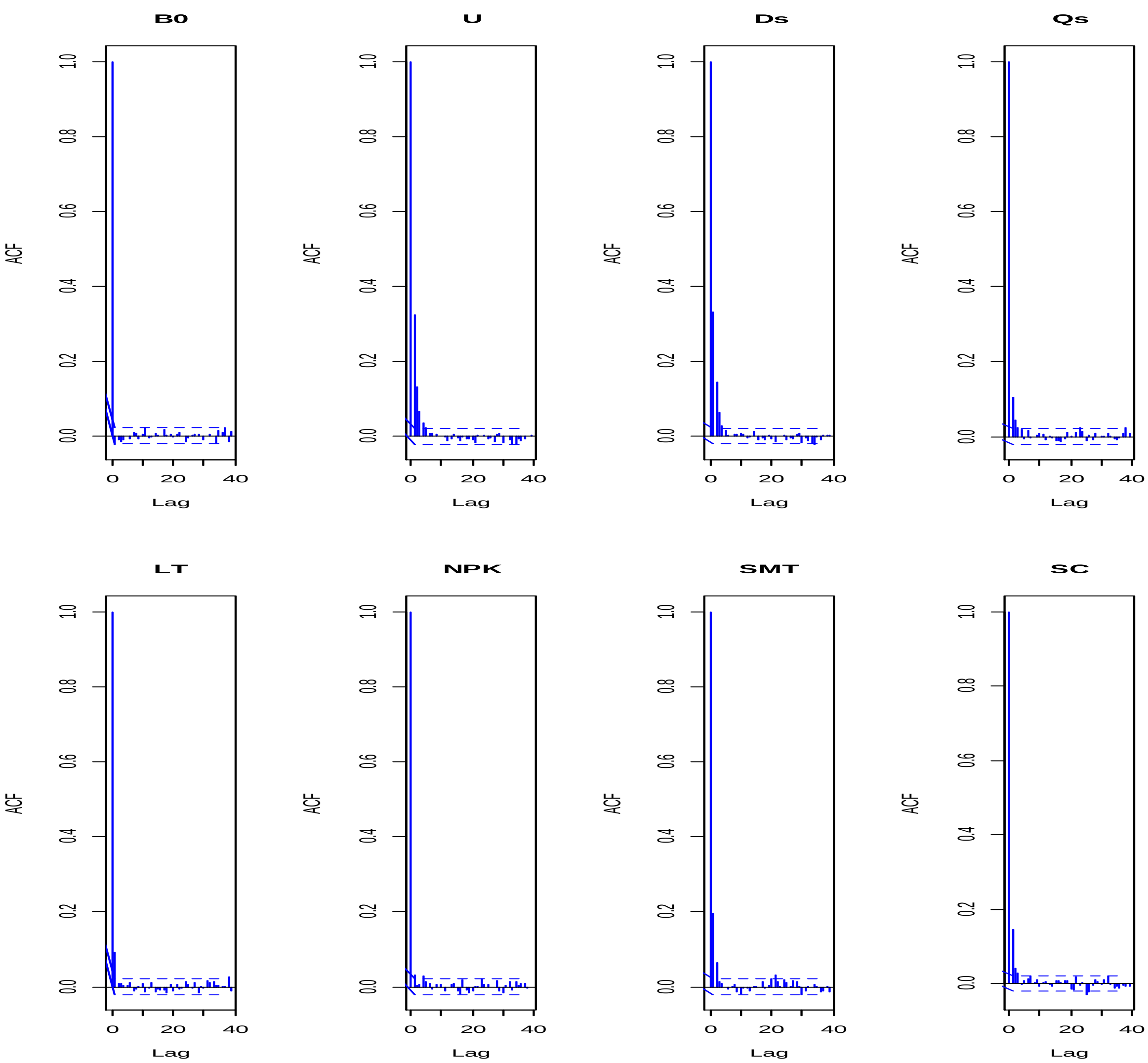


Figure 5- Trace plots of BALTR coefficients estimation



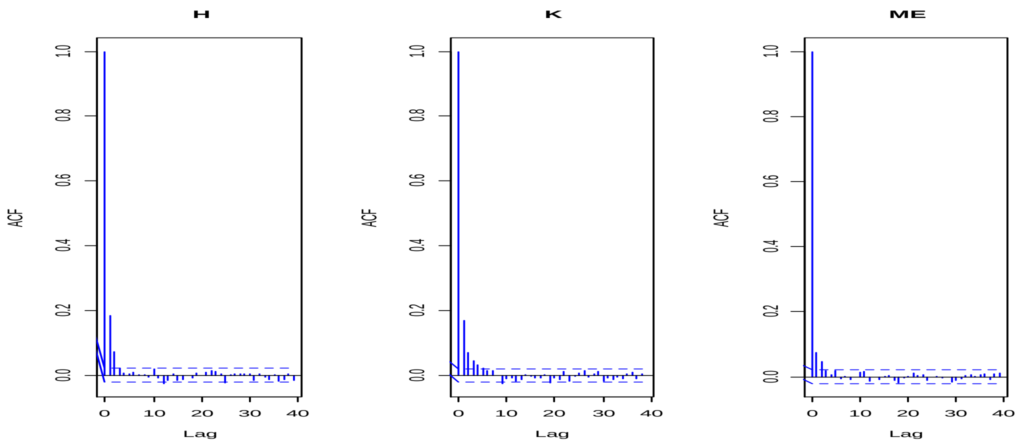


Figure 6- Autocorrelations of BALTR coefficients estimation

The BLTR and BALTR coefficients estimation are based on posteriors samples of 10,000 recurrences.

In figures 1 and 4, which illustrates the histograms of BLTR and BALTR covariates, these histograms displayed that the conditional posteriors of the two methods are stationary for its underlying truncated normal distribution. In figures 2 and 5, showed the trace plots of BLTR and BALTR covariates, these plots show a reasonably good approximation, and the noise has 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 figures 3 and 6, the plots showed the autocorrelations of BLTR and BALTR coefficients, the 10 covariates in these data are highly correlated, and these MCMCs chains in BLTR process were practically well.

CHAPTER SIX

6. CONCLUSION $ RECOMMENDATION

6.1 Conclusion:

This thesis has presented new techniques for model selection of Tobit regression from Bayesian framework, where we suggested BLTR, and BALTR procedures to estimates the coefficients with VS. The proposed procedures depend on the scale mixture uniform as prior distribution. We developed a new Bayesian hierarchical model for BLTR and BALTR procedures. Furthermore, we have provided the Gibbs samples for these procedures. The extension of our procedures has been included in our thesis, where the ridge parameter is added within the variance covariance matrix to prevent the singularity in case of multicollinearity and overfitting problems. We demonstrated the advantages of the new procedures in both simulations and analysis of real data in chapter four. The results showed that our procedures performed well in terms of VS and parameter estimation. In particular, the BALTR technique is absolutely the best of all the procedures mentioned above. Through the conclusions of this thesis, statisticians are assisted by the presence of BALTR technique in statistics, using this new technique to ensure accurate and useful results for the correct prediction.

6.2 Recommendations for future research:

The BLTR and BALTR procedures in this thesis will provide statistical researchers with promising hope, to introduce and extend new procedures for coefficients estimation and VS in the Tobit regression. There are many other probable extensions, such as, using the Bayesian group Lasso in Tobit regression, Bayesian elastic net in Tobit regression, Bayesian Bridge and group Bridge in Tobit regression.

Bibliography

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, *19*(6), 716-723.

Alhamzawi, R. (2012). Package 'Brq'. *Statistical Modelling*, *12*(3), 279-297.

Alhamzawi, R. (2013). Tobit quantile regression with the adaptive lasso penalty. In *The Fourth International Arab Conference of Statistics, 450 ISSN (1681 6870)*.

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

Alhamzawi, R. (2016-1). Bayesian elastic net Tobit quantile regression. *Communications in Statistics-Simulation and Computation*, *45*(7), 2409-2427.

Alhamzawi, R. (2016-2). Bayesian model selection in ordinal quantile regression. *Computational Statistics & Data Analysis*, *103*, 68-78.

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

Alhamzawi, R., & Yu, K. (2012). Variable selection in quantile regression via Gibbs sampling. *Journal of Applied Statistics*, *39*(4), 799-813.

Alhamzawi, R., & Yu, K. (2013). Conjugate priors and variable selection for Bayesian quantile regression. *Computational Statistics & Data Analysis*, *64*, 209-219.

Alhamzawi, R., & Yu, K. (2014). Bayesian Lasso-mixed quantile regression. *Journal of Statistical Computation and Simulation*, *84*(4), 868-880.

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.

Alhamzawi, R., Yu, K., & Benoit, D. F. (2012). Bayesian adaptive Lasso quantile regression. *Statistical Modelling*, *12*(3), 279-297.

Alhusseini, F. H. H. (2017-1). New Bayesian Lasso Composite Quantile Regression. *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, *29*(1), 138-152.

Alhusseini, F. H. H. (2017-2). New Bayesian Lasso in Tobit Quantile Regression. *Romanian Statistical Review Supplement*, *65*(6), 213-229.

Alhusseini, F. H. H., & Georgescu, V. (2018). Bayesian composite Tobit quantile regression. *Journal of Applied Statistics*, *45*(4), 727-739.

Alshaybawee, T., Alhamzawi, R., Midi, H., & Allyas, I. I. (2018). Bayesian variable selection and coefficient estimation in heteroscedastic linear regression model. *Journal of Applied Statistics*, 1-15.

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.

Andrews, D. F., & Mallows, C. L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 99-102.

Benoit, D. F., Alhamzawi, R., & Yu, K. (2013). Bayesian lasso binary quantile regression. *Computational Statistics*, *28*(6), 2861-2873.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research*, *33*(2), 261-304.

Chand, S. (2012). On tuning parameter selection of lasso-type methods-a monte carlo study. In *Applied Sciences and Technology (IBCAST), 2012 9th International Bhurban Conference on* (pp. 120-129). IEEE.

Chen, J., & Chen, Z. (2008). Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*, *95*(3), 759-771.

Chen, X., Wang, Z. J., & McKeown, M. J. (2011). A Bayesian Lasso via reversible-jump MCMC. *Signal Processing*, *91*(8), 1920-1932.

Chib, S. (1992). Bayes inference in the Tobit censored regression model. *Journal of Econometrics*, *51*(1-2), 79-99.

Choy, S. T. B., & Smith, A. F. M. (1997). Hierarchical models with scale mixtures of normal distributions. *Test*, *6*(1), 205-221.

Christian K., & Achim Z. (2008). Applied Econometrics with R. New York: Springer-Verlag. ISBN 978-0-387-77316-2.

Donoho, D. L., & Johnstone, J. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *biometrika*, *81*(3), 425-455.

Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression. *The Annals of statistics*, *32*(2), 407-499.

Efroymson, M. A. (1960). Multiple regression analysis. *Mathematical methods for digital computers*, 191-203.

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, *96*(456), 1348-1360.

Feng, X. N., Wu, H. T., & Song, X. Y. (2017). Bayesian adaptive lasso for ordinal regression with latent variables. *Sociological Methods & Research*, *46*(4), 926-953.

Frank, L. E., & Friedman, J. H. (1993). A statistical view of some chemometrics regression tools. *Technometrics*, *35*(2), 109-135.

George, E. I., & McCulloch, R. E. (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association*, *88*(423), 881-889.

George, E. I., & McCulloch, R. E. (1997). Approaches for Bayesian variable selection. *Statistica sinica*, 339-373.

Griffin, J. E., & Brown, P. J. (2010). Inference with normal-gamma prior distributions in regression problems. *Bayesian Analysis*, *5*(1), 171-188.

Hans, C. (2009). Bayesian lasso regression. *Biometrika*, *96*(4), 835-845.

Huang, J., Ma, S., & Zhang, C. H. (2008). Adaptive Lasso for sparse high-dimensional regression models. *Statistica Sinica*, 1603-1618.

Hwang, J. S., & Hu, T. H. (2015). A stepwise regression algorithm for high-dimensional variable selection. *Journal of Statistical Computation and Simulation*, *85*(9), 1793-1806.

Johnson, V. E., & Rossell, D. (2012). Bayesian model selection in high-dimensional settings. *Journal of the American Statistical Association*, *107*(498), 649-660.

Kundu, D., & Murali, G. (1996). Model selection in linear regression. *Computational statistics & data analysis*, *22*(5), 461-469.

Kyung, M., Gill, J., Ghosh, M., & Casella, G. (2010). Penalized regression, standard errors, and Bayesian lassos. *Bayesian Analysis*, *5*(2), 369-411.

Leng, C., Tran, M. N., & Nott, D. (2014). Bayesian adaptive lasso. *Annals of the Institute of Statistical Mathematics*, *66*(2), 221-244.

Li, H., Zhang, Z., & Zhang, C. (2017). China's intervention in the central parity rate: A Bayesian Tobit analysis. *Research in International Business and Finance*, *39*, 612-624.

Li, J., Das, K., Fu, G., Li, R., & Wu, R. (2010). The Bayesian lasso for genome-wide association studies. *Bioinformatics*, *27*(4), 516-523.

Li, Q., & Lin, N. (2010). The Bayesian elastic net. *Bayesian analysis*, *5*(1), 151-170.

Lincheng, W. Z. W. Y. Z. (2010). A LASSO-Type Approach to Variable Selection and Estimation for Censored Regression Model [J]. *Chinese Journal of Applied Probability and Statistics*, *1*, 008.

Liu, X., Wang, Z., & Wu, Y. (2013). Group variable selection and estimation in the tobit censored response model. *Computational Statistics & Data Analysis*, *60*, 80-89.

Lykou, A., & Ntzoufras, I. (2013). On Bayesian lasso variable selection and the specification of the shrinkage parameter. *Statistics and Computing*, *23*(3), 361-390.

Mallick, H., & Yi, N. (2014). A new Bayesian lasso. *Statistics and its interface*, *7*(4), 571.

Mitchell, T. J., & Beauchamp, J. J. (1988). Bayesian variable selection in linear regression. *Journal of the American Statistical Association*, *83*(404), 1023-1032.

Olagunju, F. I., & Ajiboye, A. (2010). Agricultural lending decision: a tobit regression analysis. *African journal of food, agriculture, nutrition and development*, *10*(5).

Park, T., & Casella, G. (2008). The bayesian lasso. *Journal of the American Statistical Association*, *103*(482), 681-686.

Pötscher, B. M., & Schneider, U. (2009). On the distribution of the adaptive LASSO estimator. *Journal of Statistical Planning and Inference*, *139*(8), 2775-2790.

Raftery, A. E., Madigan, D., & Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, *92*(437), 179-191.

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.

Sun, W., Ibriham, J. G., & Zou, F. (2010). Genome-wide multiple loci mapping in experimental crosses by the iterative adaptive penalized regression. *Genetics*.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267-288.

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(3), 273-282.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica: journal of the Econometric Society*, 24-36.

Wang, H., & Leng, C. (2007). Unified LASSO estimation by least squares approximation. *Journal of the American Statistical Association*, *102*(479), 1039-1048.

West, M. (1987). On scale mixtures of normal distributions. *Biometrika*, *74*(3), 646-648.

Yu, K., & Stander, J. (2007). Bayesian analysis of a Tobit quantile regression model. *Journal of Econometrics*, *137*(1), 260-276.

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.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, *101*(476), 1418-1429.

الخلاصة

ان اجراء اختيار المتغيرات تشير إلى نوع المشاكل التي يحاول فيها المرء تحديد المجموعة الجزئية المثلى للمتغيرات ذات الصلة في النموذج ، والتي يمكن استخدامها للحصول على تعديلات دقيقة على نتائج متغير استجابة معين. في كثير من الأحيان ، عندما يكون عدد المتغيرات كبيرًا جدًا ، يصعب تحديد المتغيرات الهامة والمؤثرة في متغير الاستجابة. لهذا السبب ، اعتبرت اجراء اختيار المتغيرات مهمة جدا في تحليل البيانات.

تعد طرق التنظيم من افضل الطرق التي أثبتت فعاليتها في التعامل مع البيانات عالية الأبعاد ، وفي السنوات السابقة ، بذل الباحثون الإحصائيون جهودًا كبيرة في تطوير طرق التنظيم لحل مشاكل اختيار المتغيرات. فهذه الطرق تسهّل تلقائياً عملية اختيار المتغيرات من خلال تعيين معاملات معيّنة إلى الصفر وتقليص المتبقي ، وتقديم تقديرات مفيدة حتى وان كان النموذج يحتوي على عدد كبير من متغيرات عالية الارتباط. بالرغم من تطوُّر طرق التنظيم في السنوات الأخيرة ، إلا أنه لا يزال بالإمكان تحسين هذه الطرق.

في هذه الرسالة ، اقترحنا طريقتان جديدتان لاختيار النموذج في انحدار توبت ، طريقة انحدار توبت لاسو البيزي وطريقة انحدار توبت لاسو التكيفي البيزي. تمتلك الطريقتان ميزات كثيرة والتي اعطت افضل تقدير مع اختيار المتغيرات المهمة. وعلى وجه التحديد ، قدمنا نموذجًا هرمياً جديداً مع عينات جبس جديدة لكل طريقة.

ثم وسعنا الطريقتان المقترحتان عن طريق إضافة معلمة الحرف (Ridge) داخل مصفوفة التباين لتجنب حالة التفرد في مشكلة التعدد الخطي أو في حالة عدد المتغيرات أكبر من عدد من المشاهدات. وتم إجراء مقارنة مع الطرق السابقة الأخرى من خلال تطبيق أمثلة المحاكاة واستخدام بيانات حقيقية. ومن الجدير بالذكر أن النتائج التي تم الحصول عليها واعدة ومشجعة ، والتي اعطت نتائج أفضل مقارنة بالطرق السابقة.

بسم الله الرحمن الرحيم

قَالُوا۟ سُبْحَـٰنَكَ

لَا عِلْمَ لَنَآ إِلَّا مَا عَلَّمْتَنَآ ۖ

إِنَّكَ أَنتَ ٱلْعَلِيمُ ٱلْحَكِيمُ

صدق الله العظيم

سورة البقرة الآية (32)

وزارة التعليم العالي والبحث العلمي
جامعة القادسية
كلية الادارة والاقتصاد
قسم الإحصاء
الدراسات العليا

انحدار توبت لاسو التكيفي البيزي مع تطبيق عملي

رسالة مقدمة

إلى مجلس كلية الإدارة والاقتصاد - جامعة القادسية

وهي جزء من متطلبات نيل درجة الماجستير في علوم الإحصاء

من الطالب

حيدر كاظم عباس الهلالي

إشراف

الأستاذ الدكتور رحيم جبار الحمزاوي

2019 م