BAYESIAN ELASTIC NET REGRESSION

BY

MISS KANYALIN JIRATCHAYUT

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY IN STATISTICS
(INTERNATIONAL PROGRAM)
DEPARTMENT OF MATHEMATICS AND STATISTICS
FACULTY OF SCIENCE AND TECHNOLOGY
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2014

BAYESIAN ELASTIC NET REGRESSION

BY

MISS KANYALIN JIRATCHAYUT

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY IN STATISTICS
(INTERNATIONAL PROGRAM)
DEPARTMENT OF MATHEMATICS AND STATISTICS
FACULTY OF SCIENCE AND TECHNOLOGY
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2014

THAMMASAT UNIVERSITY
FACULTY OF SCIENCE AND TECHNOLOGY

DISSERTATION

BY

MISS KANYALIN JIRATCHAYUT

ENTITLED

BAYESIAN ELASTIC NET REGRESSION

was approved as partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Statistics (International Program)

on March 20, 2015

Chairman ______________________________
(Associate Professor Jirawan Jitthavech, Ph.D.)

Member and Advisor ______________________________
(Associate Professor Chinnaphong Bumrungsup, Ph.D.)

Member Kamon Budsaba
(Associate Professor Kamon Budsaba, Ph.D.)

Member Rattana Lerdsuwansri
(Assistant Professor Rattana Lerdsuwansri, Ph.D.)

Member ______________________________
(Assistant Professor Ramidha Srihera, Ph.D.)

Dean P. Sermsuk
(Associate Professor Pakorn Sermsuk, M.Sc.)

Dissertation Title	BAYESIAN ELASTIC NET REGRESSION
Author	Miss Kanyalin Jiratchayut
Degree	Doctor of Philosophy in Statistics (International Program)
Department/Faculty/University	Department of Mathematics and Statistics Faculty of Science and Technology Thammasat University
Dissertation Advisor	Associate Professor Chinnaphong Bumrungsup, Ph.D.
Academic Year	2014

Abstract

We propose the method for estimating the value of the L_2 penalty parameter, λ_2, of elastic net linear regression model using Bayesian analysis. The value of λ_2 is specified through Bayes factor. We compare the performance of the value of λ_2 based on Bayes factor to the value of λ_2 chosen by 10-fold cross-validation method. Simulation studies and real data examples show that the value of λ_2 based on Bayes factor performs better in prediction accuracy. The value of λ_2 based on Bayes factor can be used for adaptive elastic net estimator where the adaptive weight is included in the L_1 penalty. We study the performance of two adaptive elastic net estimation methods where the adaptive weights are constructed using elastic net and least squares estimators. Simulation studies show that two adaptive weights perform differently. When the elastic net estimator is used, the adaptive elastic net performs best in estimation accuracy and variable selection performance. If the least squares estimator is used, the adaptive elastic net has the prediction performance better than using the other adaptive weight. We study the performance of the Bayesian variable selection for elastic net linear regression model (BVS) using two different priors: the penalty parameters λ_1 and λ_2 are estimated by the 10-fold cross-validation, and the penalty parameter λ_2 is based on Bayes factor.

The variable selection result of BVS differs from elastic net. The BVS performs both variable selection and group selection where the pair of predictors which are highly correlated with the response variable is included into the optimal model whereas some pair of predictors which are highly correlated with the response variable is excluded from the elastic net model. The BVS is more parsimonious than the elastic net. For BVS method, the prior for the penalty parameters λ_1 and λ_2 estimated by the 10-fold cross-validation method is the best.

Keywords: adaptive elastic net, adaptive weight, Bayes factor, Bayesian analysis, elastic net, L_2 penalty, shrinkage, variable selection

Acknowledgements

First of all, I would especially like to thank my advisor, Associate Professor Dr. Chinnaphong Bumrungsup, for his supervision, helpful discussion, suggestions and attention throughout this research. I would also like to thank the committee members, Associate Professor Dr. Jirawan Jitthavech, Associate Professor Dr. Kamon Budsaba, Assistant Professor Dr. Rattana Lerdsuwansri, and Assistant Professor Dr. Ramidha Srihera, for their useful comments and suggestions.

I would like to thank my office, Faculty of Science and Art, Burapha University, Chantaburi Campus, for supporting by the scholarship for three academic years (2010 – 2012), and MATLAB2012a software (provided by Dr. Sarun Pibanchon) – the most important tool in my research. I would also like to extend my thanks to my friend, Miss. Kitiya Harinpholsit, and my colleague at Burapha University. They always cheer me up.

I would like to acknowledge the financial support provided by Thammasat University Research Fund under the TU Research Scholar, Contract No.30/2557. This grant provided the funding for my research.

Finally, I would like to express my deep gratitude to my parents and my brother for their love and support. My lovely father and mother have been a constant source of my inspirations throughout my life.

Kanyalin Jiratchayut

Table of Contents

List of tables

List of flowcharts and figures

Chapter 1

Introduction

1.1 Problem statement

Regression analysis is a statistical methodology that utilizes the relation between predictor variables and response variable for analyzing, modeling and prediction. Consider a linear regression model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}\,, \tag{1.1}$$

where $\mathbf{y}$ is an $n \times 1$ vector of response variable,
$\mathbf{X}$ is an $n \times p$ matrix of predictor variables,
$\boldsymbol{\beta}$ is an $p \times 1$ vector of parameter of regression coefficients,
$\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of random errors,
p is the number of predictors,
n is the number of observations.

The errors are assumed to be independent identically normally distributed random variable with mean 0 and finite variance σ^2. Without loss of generality, we assume the response is centered and the predictors are standardized, so the intercept is not included in the regression function.

Multiple linear regression model is the model with more than one predictor variables. When the predictor variables are correlated among themselves, intercorrelation or multicollinearity among them is said to exist. Multicollinearity between the predictor variables or the case where the number p of predictors is larger than the number n of observations affects the least squares estimator. Multicollinearity between the predictor variables makes a near-linear dependence between the predictor variables $\mathbf{X}$, this results $\mathbf{X}$ to be non-full rank matrix and make the determinant of $\mathbf{X}^T\mathbf{X}$ to be very close to zero. The method of least squares is often used to estimate the parameters in a linear regression model. The solution $\widehat{\boldsymbol{\beta}}_{LS} = (\mathbf{X}^T\mathbf{X}\,)^{-1}\mathbf{X}^T\mathbf{y}$ would usually be solvable. The variance of the estimator $var\big(\widehat{\boldsymbol{\beta}}_{LS}\big) =$

$(\mathbf{X}^T\mathbf{X})^{-1}\sigma^2$ is large when there is multicollinearity among the predictor variables $\mathbf{X}$. This leads to unreliable estimates of the regression coefficients. With such a large number of predictors, multicollinearity problem may exist. Also, with a large number of predictors there is often desired to select a smaller subset that not only fits as well as the full set of variables, but also contains the more important predictors (Kyung, Gill, Ghosh, & Casella, 2010).

To deal with multicollinearity problem, many penalized regression methods are developed for coefficient estimation and variable selection. The penalized regression is developed from least squares method with penalty function to discover relevant explanatory factors and to get higher prediction accuracy in linear regression. The examples of penalized regression are ridge regression (Hoerl & Kennard, 1970a, 1970b), the lasso (Tibshirani, 1996), and elastic net (Zou & Hastie, 2005).

Ridge regression is one of several methods that have been proposed to remedy multicollinearity problems by modifying the method of least squares to allow biased estimators of the regression coefficients. Ridge regression was proposed by Hoerl and Kennard (1970a, 1970b). Ridge regression imposes L_2 penalty and shrinks the regression coefficients towards zero. Ridge estimators are sometimes referred to as shrinkage estimators. However, it does not set any coefficients to zero and hence does not give an easily interpretable model.

The least absolute shrinkage and selection operator or lasso technique was proposed by Tibshirani (1996). The lasso is a penalized least squares method imposing an L_1 penalty on the regression coefficients. It shrinks some coefficients and sets others to 0, providing a form of variable selection. Hence, the lasso retains the good features of both subset selection and ridge regression. The lasso has some limitations: (a) in the $p > n$ case, the lasso can only select at most n variables out of p candidates, (b) the lasso lacks the ability to reveal the grouping information (Celeux, Anbari, Marin, & Robert, 2012; Zou & Hastie, 2005), and (c) for usual $n > p$ situations, if there are high correlations between the predictors, it has been empirically observed that the prediction performance of the lasso is dominated by ridge regression (Tibshirani, 1996, cited by Zou & Hastie, 2005). Scenario (a) and (b) make the lasso be an inappropriate variable selection method in some situations.

Classification and regression problems with large numbers of candidate predictor variables ($p > n$) occur in a wide variety of scientific fields, for example, the regression problem in microarray gene expression data. A typical microarray dataset has thousands of predictors (genes) and less than 100 samples. For those genes sharing the same biological "pathway", the correlation between them can be high. This means those genes are forming a group. The ideal gene selection method shall be able to do two things: eliminate the trivial genes, and automatically include whole groups into the model once one gene among them is selected. Hesterberg, Choi, Meier, and Fraley (2008) claimed that the goals in model selection include (1) the accurate predictions, (2) interpretable models – determining which predictors are meaningful, and (3) stability – small changes in the data should not result in large changes in either using the subset of predictors, the associated coefficients, or the predictions.

The lasso has some drawbacks for microarray classification and gene selection. Zou and Hastie (2005) proposed a new regularization technique which called the elastic net to solve the regression problem in microarray genes expression data. The elastic net simultaneously performs automatic variable selection and continuous shrinkage, it can select groups of correlated variables and overcomes the difficulty of $p > n$. The elastic net is based on a combination of the ridge (L_2) and the lasso (L_1) penalties. The elastic net estimator is defined as follows:

$$\widehat{\boldsymbol{\beta}}_{\text{elastic net}} = (1+\lambda_2)\left\{\arg\min_{\boldsymbol{\beta}}\left[\|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2+\lambda_2\|\boldsymbol{\beta}\|_2^2+\lambda_1\|\boldsymbol{\beta}\|_1\right]\right\}, \qquad (1.2)$$

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are the penalty parameters, $\|\boldsymbol{\beta}\|_1 = \sum_{j=1}^{p}|\beta_j|$ is the L_1 norm of $\boldsymbol{\beta}$, and $\|\boldsymbol{\beta}\|_2^2 = \sum_{j=1}^{p}\beta_j^2$ is the L_2 norm of $\boldsymbol{\beta}$. The L_1 part of the elastic net performs automatic variable selection, while the L_2 part stabilizes the solution parts and, hence, improves the prediction. However, when the correlations among the predictors become high, the elastic net can significantly improve the prediction accuracy of the lasso.

The elastic net has good performance. However, it does not enjoy the oracle properties (consistency in variables selection and asymptotic normality). Zou and Zhang (2009), and Ghosh (2011) proposed two adaptive elastic net estimators which have the oracle property. These two adaptive elastic net estimators are different.

Zou and Zhang (2009) proposed the adaptive elastic net using the elastic net estimator to construct the adaptive weight. This adaptive elastic net has the oracle property and outperforms the elastic net. The adaptive elastic net estimator proposed by Zou et al. (2009) is defined as follows:

$$\widehat{\boldsymbol{\beta}}_{\text{AENET2009}} = (1+\lambda_2)\left\{\arg\min_{\boldsymbol{\beta}}\left[\|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda_2\|\boldsymbol{\beta}\|_2^2 + \lambda_1 \textstyle\sum_{j=1}^{p} \widehat{w}_j\left|\beta_j\right|\right]\right\}. \tag{1.3}$$

The adaptive weight $\widehat{w}_j = \left(\left|\hat{\beta}_j(\text{elastic net})\right|\right)^{-\gamma}$, $j = 1, \ldots, p$, where γ is a positive constant.

Ghosh (2011) proposed the adaptive elastic net using the least squares estimator to construct the adaptive weight. This method has good performance on grouped selection and model complexity than the elastic net. The adaptive elastic net estimator proposed by Ghosh (2011) is defined as follows:

$$\widehat{\boldsymbol{\beta}}_{\text{AENET2011}} = (1+\lambda_2)\left\{\arg\min_{\boldsymbol{\beta}}\left[\|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda_2 \textstyle\sum_{j=1}^{p}\left(\left|\beta_j\right| + \frac{\lambda_1}{2\lambda_2}\widehat{w}_j\right)^2 - \frac{(\lambda_1)^2}{4\lambda_2}\sum_{j=1}^{p}\widehat{w}_j^2\right]\right\}. \tag{1.4}$$

The adaptive weight vector $\widehat{\mathbf{w}} = 1/\left|\widehat{\boldsymbol{\beta}}_{least\ square}\right|^{\gamma}$, for some $\gamma > 0$. The adaptive elastic net proposed by Ghosh (2011) has good performance on grouped selection and model complexity than the elastic net.

There is no any comparative study between two adaptive elastic net methods proposed by Ghosh (2011) and Zou and Zhang (2009).

The penalty parameters (λ_1 and λ_2) control the amount of shrinkage imposed on the coefficients, where some weak effects are forced to be exactly zero if the shrinkage level is large enough. Hence, the penalty parameters (λ_1 and λ_2) can be named the shrinkage parameters. If the value of λ_1 and λ_2 are too small, then no shrinkage will be performed. If the value of λ_1 and λ_2 are too high, then all coefficients will be shrunk to zero. Hence, the penalty parameters (λ_1 and λ_2) are important for the elastic net and adaptive elastic net estimators.

For elastic net method proposed by Zou and Hastie (2005), the penalty parameters λ_1 and λ_2 are selected by K-fold cross-validation method. To avoid intensive computation, a grid of values for λ_2 is first specified. Zou et al. (2005) suggested to pick a relatively small grid value of λ_2. They used (0, 0.01, 0.1, 1, 10, 100, 1000). For each λ_2, a 10-fold cross-validation is then used to choose λ_1. The chosen λ_2 is the one giving the smallest cross-validation error. The adaptive elastic net method proposed by Ghosh (2011) used the same method of Zou et al. (2005) for choosing the value of λ_1 and λ_2. Zou and Zhang (2009) selected the penalty parameters by using Bayesian Information Criterion (BIC) which applied from Zou, Hastie, and Tibshirani (2007) but they did not give the detail of the process for choosing the value of λ_1 and λ_2.

K-fold cross-validation is the primary method used for estimating a penalty parameter in the lasso, the elastic net, and the adaptive elastic net techniques. The idea of cross-validation is to split the data into k roughly equal parts, using a portion of the data to build (or "train") the model and the remainder to validate the model. For each training set, the researcher fits the model with a penalty parameter λ and computes its cross-validation error. The researcher does the same process for many values of λ and chooses the value of λ that gives the smallest cross-validation error. The different number of k may cause the different value of λ for the same dataset. Hence, the number of subset k affects the penalty parameter. It is interesting to wonder about what is the appropriate k-fold cross-validation. The case $k = n$ is known as leave-one-out cross validation. With $k = n$, the cross-validation estimator is approximately unbiased for the true (expected) prediction error, but can have high variance because the n "training set" are so similar to one another. With $k = 5$, cross-validation has lower variance but bias could be a problem. Hastie, Tibshirani, and Friedman (2009) suggested that typical choices of k are 5 or 10. Hyndman (2010) suggested the researcher should beware of looking at statistical tests after selecting variables using cross-validation since the tests do not take into account of the variable selection that has taken place and so the *p*-value can mislead. Cross-validation can be misused when there are exact duplicate observations (two or more observations with equal values for all predictors and for the response variable).

Bayesian Information Criterion (BIC) for lasso was developed by Zou, Hastie, and Tibshirani (2007). Given a model fitting method δ, let $\hat{\boldsymbol{\mu}} = \delta(\mathbf{y})$ represent its fit and $\mathbf{y}$ is a vector of response variable. Following the usual definition of BIC proposed by Schwarz 1978, Zou, et al. (2007) proposed BIC for the lasso as

$$\mathrm{BIC}(\hat{\boldsymbol{\mu}}) = \frac{\|\mathbf{y}-\hat{\boldsymbol{\mu}}\|^2}{n\sigma^2} + \frac{\log(n)}{n}\widehat{df}(\hat{\boldsymbol{\mu}}).$$

Zou, et al. (2007) suggested using BIC to select the penalty parameter of the lasso. Zou and Zhang (2009) used BIC to select the penalty parameters λ_1 and λ_2 of the adaptive elastic net.

The BIC was first developed by Schwarz (1978), who gave a Bayesian argument for adopting it. For model selection purposes, BIC is asymptotically consistent as a selection criterion. What this means is that given a family of models, including the true model, the probability that BIC will select the correct model approaches one as the sample size $n \to \infty$. On the other hand, for finite samples, BIC often chooses models that are too simple, because of its heavy penalty on complexity (Hastie, Tibshirani, & Friedman, 2009). Asymptotically, for linear models minimizing BIC is equivalent to leave–υ–out cross-validation when $\upsilon = n\{1 - 1/[\log(n) - 1]\}$ (Shao, 1997, cited by Hyndman, 2010). BIC can be used to obtain a rough approximation of the log-Bayes factor (Kass, & Raftery, 1995, cited by Ntzoufras, 2009).

Bayes factor proposed by Kass and Raftery (1995) is a quantity for comparing models and for testing hypotheses in the Bayesian framework. Bayes factor has played a major role in assessing the goodness of fit of the competing models (Ando, 2010). Bayes factor provides the Bayesian solution to the question, “What evidence do the data provide for one model against another, competing model?” expressed as a ratio of posterior probabilities (Raftery, 1999).

Lykou and Ntzoufras (2012) proposed a Bayesian implementation of the lasso regression that accomplishes both shrinkage and variable selection. They proposed the method for estimating the penalty parameter λ in the lasso technique by using the behavior of Bayes factors and trying to avoid large and small λ values that may overshrink the coefficients or activate the Lindley-Bartlett paradox. The penalty parameter λ was associated with the values of Pearson and partial correlation at the limits between significance and insignificance as defined by Bayes factors. This method differs from the k-fold cross-validation method and BIC.

For elastic net and adaptive elastic net methods, the penalty parameter λ_2 is the penalty parameter of the L_2 part which stabilizes the solution part of the elastic net and adaptive elastic net estimates. Zou and Hastie (2005) suggested to specific the value of λ_2 in the first step of the elastic net procedure. Hence, the value of penalty parameter λ_2 plays an important role for elastic net and adaptive elastic net methods.

After study and analyze the literature review, the above ideas motivate the first new idea in this thesis to propose the new method for estimating the penalty parameter (λ_2) of the elastic net and the adaptive elastic net linear regression models by using Bayes factor and Pearson correlation between y and the candidate predictor x_j.

The penalized regression approaches such as the lasso and the elastic net can lead to finding smaller groups of variables with good prediction accuracy since their process shrunk some coefficients to zero, providing a form of variable selection. Celeux, Anbari, Marin, and Robert (2012) showed that the Bayesian variable selection methods perform better variable selection efficiency than the lasso and the elastic net. In the Bayesian framework, the model selection problem is transformed to the form of parameter estimation, rather than searching for the single optimal model, a Bayesian will attempt to estimate the posterior probability of all models within the considered class of models (O'Hara, & Sillanpää, 2009). In principle, the Bayesian approach for selecting a model is to choose the model with the largest posterior probability among a set of candidate models (Ando, 2010).

This motivates the second new idea in this thesis to combine Bayesian variable selection method with elastic net linear regression model where the penalty parameter (λ_2) is specified by the method proposed in this thesis.

1.2 Research objectives

In this thesis, there are two new ideas and two comparative studies which are the following.

1. To propose the new method for estimating the value of the penalty parameter (λ_2) of the elastic net and two adaptive elastic net linear regression models by using Bayes factor and Pearson correlation between y and the candidate predictor x_j .

2. To combine a Bayesian variable selection with elastic net linear regression model where the prior distribution of $\boldsymbol{\beta}$ is a compromise between normal and double exponential distribution.

3. To perform the comparative study between the new method which is proposed in the first objective of this thesis and the 10-fold cross-validation method for estimating the value of the penalty parameter (λ_2) of the elastic net and adaptive elastic net linear regression models. (The reason for using 10-fold cross-validation method is that it was used by Zou and Hastie (2005) for choosing the penalty parameters of the elastic net regression.)

4. To perform the comparative study between the Bayesian variable selection for elastic net linear regression model which is proposed in the second objective of this thesis and elastic net linear regression model (Zou & Hastie, 2005) where the penalty parameter (λ_2) is estimated by the new method presented in the first objective of this thesis and the 10-fold cross-validation method.

1.3 Research hypotheses

The hypotheses for this study are as follows:

1. Using the penalty parameter (λ_2) associated with Bayes factor, the elastic net and adaptive elastic net estimators have mean squared error less than those estimators derived by using the value of λ_2 obtained from 10-fold cross-validation method do.

2. Using the penalty parameter (λ_2) associated with Bayes factor, the elastic net and adaptive elastic net linear regression models have the variable selection performance – the number of zero coefficients that are correctly estimated by zero (C) is maximum, whereas the number of nonzero coefficients that are incorrectly estimated by zero (IC) is minimum – better than the elastic net and adaptive elastic net models derived by using the value of λ_2 obtained from 10-fold cross-validation method do.

3. Using the penalty parameter (λ_2) associated with Bayes factor, the derived elastic net model is significance model as defined by Bayes factor and this model has prediction error less than the elastic net model derived by using the value of λ_2 obtained from 10-fold cross-validation method does.

4. For linear regression analysis of microarray classification and gene selection where the prior distribution of $\boldsymbol{\beta}$ is a compromise between normal and double exponential distribution, the combination of Bayesian variable selection and elastic net model performs the variable selection efficiency – the optimal model exhibits group selection – better than the classical elastic net model does.

1.4 Scope of the study

The scope of the study includes the following.

1. The proposed new ideas are used for analysis a multiple linear regression where independent variables show degree of multicollinearity.

2. In this thesis, we assume the errors are independently and identically normally distributed with zero mean and finite variance σ^2.

3. In this thesis, we consider the linear regression model when the number p of predictors less than the number n of observations ($p < n$).

4. We use the simulated datasets and two real datasets to access the performance of the proposed methods. The simulated datasets are generated by the simulation methods proposed by Lykou and Ntzoufras (2012) and Zou and Zhang (2009). The real datasets are the diabetes dataset and prostate cancer data used in elastic net literatures and related methods.

5. Tools for computational method are MATLAB 2012a software and the gcdnet R package (Yang & Zou 2012).

1.5 Research advantages

The advantages of this study are the following.

1. To be helpful for choosing the penalty parameter (λ_2) of the elastic net linear regression models.

2. To be helpful for choosing the penalty parameter (λ_2) of two adaptive elastic net linear regression models when the adaptive weights depend on least squares estimate and elastic net estimate.

3. This thesis provides Bayesian variable selection for elastic net linear regression model as an alternative method for variable selection in linear regression analysis of microarray classification and genes selection when the prior distribution of $\boldsymbol{\beta}$ is a compromise between normal and double exponential distribution.

4. This thesis can be used for regression analysis in the other dataset where the predictors show degree of multicollinearity and the regression model encourages a grouping effect – strongly correlated predictors tend to be in or out of the model together, e.g., microarray gene expression data, the Near-Infrared (NIR) Spectroscopy of biscuit dough data (cookie-dough data), image processing, or document classification.

1.6 Abbreviations

AENET2011	adaptive elastic net method proposed by Ghosh (2011)
AENET2009	adaptive elastic net method proposed by Zou and Zhang (2009)
BIC	Bayesian Information Criterion
CV	cross-validation
elastic net	elastic net method proposed by Zou and Hastie (2005)
lasso	least absolute shrinkage and selection operator method proposed by Tibshirani (1996)
MSE	mean squared error
PE	prediction error
λ_2BF	λ_2 is based on Bayes factor
λ_2CV	λ_2 is estimated by 10-fold cross-validation method
w	adaptive weight
OLS	adaptive weight is constructed using ordinary least squares estimator
RENETCV01	adaptive weight is constructed using rescaled elastic net estimator with α = 0.1 and the value of λ_2 is estimated by 10-fold cross-validation method
RENETCV05	adaptive weight is constructed using rescaled elastic net estimator with α = 0.5 and the value of λ_2 is estimated by 10-fold cross-validation method
RENETCV09	adaptive weight is constructed using rescaled elastic net estimator with α = 0.9 and the value of λ_2 is estimated by 10-fold cross-validation method
RENETBF01	adaptive weight is constructed using rescaled elastic net estimator with α = 0.1 and the value of λ_2 is based on Bayes factor
RENETBF05	adaptive weight is constructed using rescaled elastic net estimator with α = 0.5 and the value of λ_2 is based on Bayes factor

RENETBF09	adaptive weight is constructed using rescaled elastic net estimator with α = 0.9 and the value of λ_2 is based on Bayes factor
BVSCV	Bayesian variable selection for elastic net linear regression model where the penalty parameters λ_1 and λ_2 are estimated by the 10-fold cross-validation method
BVSBF	Bayesian variable selection for elastic net linear regression model where the penalty parameter λ_2 is based on Bayes factor
ENETCV	elastic net method where the penalty parameter λ_2 is estimated by the 10-fold cross-validation method with $\alpha = 0.5$
ENETBF	elastic net method where the penalty parameter λ_2 is based on Bayes factor with $\alpha = 0.5$

1.7 Decision criteria

The decision criterions used in this thesis are the following.

1. For each estimator $\hat{\boldsymbol{\beta}}$, its estimation accuracy is measured by the mean square error ($MSE(\hat{\boldsymbol{\beta}})$) defined as $E\left[(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})^T(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\right]$.

2. The variable selection performance is gauged by (C, IC), where C is the number of zero coefficients that are correctly estimated by zero and IC is the number of nonzero coefficients that are incorrectly estimated by zero.

3. The prediction accuracy is measured by the prediction error (PE) defined as $E(\mathbf{y} - \hat{\mathbf{y}})^2$ where $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$.

Chapter 2

Literature review

2.1 Background of elastic net method

Regression analysis is used for prediction, variable selection and coefficient estimation in many research areas. The least squares estimator is an unbiased estimator which commonly used for parameter estimation. Multicollinearity between the predictor variables or the case where the number p of predictors is larger than the number n of observations affects the least squares estimator. (1) When multicollinearity between the predictor variables occurred, the least squares method cannot give a unique solution since $\mathbf{X}$ is not of full rank and variance of the estimator is large. It leads to unreliable estimates of the regression coefficients since they have large variances and covariances, so interpreting and using a fitted regression model are not satisfactory. (2) With such a large number of predictors, there might exist problems among predictor variables, in particular, there could be a problem with multicollinearity. Also, with a large number of predictors there is often a desire to select a smaller subset that not only fits as well as the full set of variables, but also contains the more important predictors (Kyung, Gill, Ghosh, & Casella, 2010).

To deal with multicollinearity problem, many penalized regression methods are developed for coefficient estimation and variable selection. The penalized regression is developed from least squares method with penalty function to discover relevant explanatory factors and to get higher prediction accuracy in linear regression. The examples of penalized regression are ridge regression (Hoerl & Kennard, 1970a, 1970b), the lasso (Tibshirani, 1996), and elastic net (Zou & Hastie, 2005). Penalized regression approaches have been used in cases where $p < n$, and in the case with $p > n$. The penalized regression, and its accompanying variable selection features, can lead to finding smaller groups of variables with good prediction accuracy (Kyung et al., 2010). Penalized regression is a regularization method for regression models. Regularized estimation of the parameter vector $\boldsymbol{\beta}$ is obtained by minimizing the penalized least squares criterion. This procedure is regression parameter estimates based on regularization procedure. Regularized optimization as a form of estimation has attracted broad interest; the goal of such procedures is generally to improve on predictions based on ordinary least squares by shrinking parameter estimates toward zero. Certain regularization methods allow some parameters to be set equal to zero, providing a method for identifying important variables. Parameter estimates for linear regression in this framework are typically

$$\widehat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}}[\|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda\,\mathcal{J}(\boldsymbol{\beta})]\,, \tag{2.1}$$

for some nonnegative penalty function $\mathcal{J}$ and regularization parameter $\lambda \geq 0$ (Hans, 2011). In some papers, the regularization parameter λ can be named as tuning parameter, shrinkage parameter, or penalty parameter.

Hoerl and Kennard (1970a, 1970b) proposed ridge regression to achieve better prediction in the face of multicollinearity. Ridge regression imposes the L_2 penalty and shrinks the regression coefficients towards zero. Ridge regression minimizes the residual sum of squares subject to $\sum_{j=1}^{p}\beta_j^2 \leq t$ (L_2-norm). Assuming that the response is centered and the predictors are standardized, ridge estimate is defined by

$$\widehat{\boldsymbol{\beta}}_{ridge} = \arg\min_{\boldsymbol{\beta}}\left(\|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda\sum_{j=1}^{p}\beta_j^2\right). \tag{2.2}$$

The second term is called ridge penalty. The ridge penalty is a L_2-norm of $\boldsymbol{\beta}$ or $\|\boldsymbol{\beta}\|_2^2 = \sum_{j=1}^{p}\beta_j^2$ so called L_2 penalty with the penalty parameter $\lambda \geq 0$. Usually $\mathbf{y}$ is only centered in ridge regression, but it could be scaled, too, if desired (Draper & Smith, 1998).

Suppose $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\boldsymbol{I}_n)$ and the parameter β_j are each distributed as $N(0,\tau^2)$, independently of one another, with τ^2 and σ^2 assumed known. Then the ridge estimate is the mode of the posterior distribution of $\boldsymbol{\beta}$ with $\lambda = \sigma^2/\tau^2$.

Ridge regression yields a biased estimator with the variance smaller than least squares method. Therefore, better estimation can be achieved on the average in terms of mean squares error with a little sacrifice of bias, and prediction can be improved overall (Fu, 1998). Ridge regression is a continuous process that shrinks coefficients and hence is more stable: however, it does not set any coefficients to 0. For the problem of multicollinearity, ridge regression improves the prediction performance, but it cannot produce a model with only the relevant predictors or a parsimonious model since all predictors are kept in ridge regression model (Ghosh, 2011; Kyung, Gill, Ghosh, & Casella, 2010; Tibshirani, 1996; Zou & Hastie, 2005).

Tibshirani (1996) proposed the least absolute shrinkage and selection operator or lasso technique. The lasso is a penalized least squares procedure that minimizes the residual sum of squares subject to the non-differentiable constraint expressed in terms of the L_1 norm of the coefficients ($\sum_{j=1}^{p}|\beta_j| \leq t$). Assuming that the response is centered and the predictors are standardized, the lasso estimator is given by

$$\widehat{\boldsymbol{\beta}}_{lasso} = \arg\min_{\boldsymbol{\beta}}\left(\|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda\sum_{j=1}^{p}|\beta_j|\right). \tag{2.3}$$

The lasso penalty is a L_1-norm of $\boldsymbol{\beta}$ or $\|\boldsymbol{\beta}\|_1 = \sum_{j=1}^{p}|\beta_j|$ so called L_1 penalty with the penalty parameter $\lambda \geq 0$. The case $\lambda = 0$ results in $\widehat{\boldsymbol{\beta}}_{lasso} = \widehat{\boldsymbol{\beta}}_{OLS}$, the ordinary least squares estimate, and λ sufficiently large shrinks $\widehat{\boldsymbol{\beta}}_{lasso}$ to zero. The L_1 lasso penalty $\sum_{j=1}^{p}|\beta_j|$ makes the solutions nonlinear in the y_i, and there is no closed form expression for the lasso estimator. Computing the lasso solution is a quadratic programming problem with linear inequality constraints.

The lasso can do both continuous shrinkage and variable selection. Penalized regression methods for simultaneous variable selection and coefficient estimation, especially those based on the lasso of Tibshirani (1996), have received a great deal of attention in recent years, mostly through frequentist models (Kyung, Gill, Ghosh, & Casella, 2010).

The lasso has a Bayesian interpretation. Tibshirani (1996) noted that the lasso estimate can be derived as the mode of the posterior distribution of $\boldsymbol{\beta}$ when independent double exponential prior distribution are used for $\boldsymbol{\beta}$. Suppose $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\boldsymbol{I}_n)$. The density of the double exponential (or Laplace) distribution for $\beta_j \sim DE(\mu, b)$ is given by

$$f(\beta_j|\mu, b) = \frac{1}{2b}\exp\left(-\frac{|\beta_j-\mu|}{b}\right), \tag{2.4}$$

with mean μ and variance $2b^2$. Thus, a prior $\beta_j \sim DE(0, \sigma^2/\lambda)$ will produce a posterior distribution that will be maximized under the lasso estimate ($\widehat{\boldsymbol{\beta}}_{lasso}$).

For the computation of the lasso, Efron, Hastie, Johnstone, and Tibshirani (2004) proposed least angle regression selection (LARS) for a model selection algorithm. They showed that with a simple modification, the LARS algorithm implements the lasso, and one of the advantages of LARS is the short computation time compared to other methods. For efficiently selecting the optimal fit, the effective degrees of freedom of the lasso were studied by Efron et al. (2004). They discovered that the size of the active set (the indices corresponding to covariates to be chosen) can be used as a measure of the degrees of freedom, which changes, not necessarily monotonically, along the solution paths of LARS. Zou, Hastie, and Tibshirani (2007) improved this and showed that the number of nonzero coefficients is an unbiased estimate for degrees of freedom of the lasso. In addition, Zou et al. (2007) showed that the unbiased estimator is asymptotically consistent, thus various model selection criteria can be used with the LARS algorithm for the optimal lasso fit. The original LARS method is for linear regression where the regression function has a linear relationship to the predictors. The general form of LARS can be extended for fitting linear model with nonlinear relationships between the response variable y and predictor variables x such as polynomial, splines, and kernels regression. Thus, LARS algorithm allows the lasso can be applied for various linear models, for example; multiple linear regression, generalized linear models, Cox proportional hazards models, polynomial regression, and kernels regression (Hesterberg, Choi, Meier, & Fraley, 2008).

The lasso has some limitations: (a) In the $p > n$ case, the lasso can only select at most n variables out of p candidates. (b) The lasso lacks the ability to reveal the grouping information (Celeux, Anbari, Marin, & Robert, 2012; Zou, & Hastie, 2005). and (c) For usual $n > p$ situations, if there are high correlations between the predictors, it has been empirically observed that the prediction performance of the lasso is dominated by ridge regression (Tibshirani, 1996, cited by Zou & Hastie, 2005). Scenario (a) and (b) make the lasso be an inappropriate variable selection method in some situations.

Classification and regression problems with large numbers of candidate predictor variables ($p > n$) occur in a wide variety of scientific fields, for example, the regression problem in microarray gene expression data. A typical microarray dataset has thousands of predictors (genes) and less than 100 samples. For those genes sharing the same biological “pathway”, the correlation between them can be high. This means those genes are forming a group. Qualitative speaking, a regression method exhibits the *grouping effect* if the regression coefficients of a group of highly correlated variables tend to be equal (up to a change of sign if negatively correlated) (Zou & Hastie, 2005). The ideal gene selection method shall be able to do two things: eliminate the trivial genes, and automatic include whole groups into the model once one gene among them is selected. Hesterberg, Choi, Meier, and Fraley (2008) claimed that the goals in model selection include (1) the accurate predictions, (2) interpretable models – determining which predictors are meaningful, and (3) stability – small changes in the data should not result in large changes in either using the subset of predictors, the associated coefficients, or the predictions.

2.2 Elastic net

The lasso has some drawbacks for microarray classification and gene selection. Zou and Hastie (2005) proposed a new regularization technique which called the elastic net to solve the regression problem in microarray genes expression data. The elastic net simultaneously does automatic variable selection and continuous shrinkage, it can select groups of correlated variables and overcomes the difficulty of $p > n$. The elastic net is based on a combination of the ridge (L_2) and the lasso (L_1) penalties. The elastic net is defined in two stages. Assuming that the response is centered and the predictors are standardized, naïve elastic net estimate is first found via

$$\widehat{\boldsymbol{\beta}}_{\text{Naïve elastic net}} = \arg\min_{\boldsymbol{\beta}}[\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda\, \mathcal{J}(\boldsymbol{\beta})] \tag{2.5}$$

with the elastic net penalty $\mathcal{J}(\boldsymbol{\beta}) = \alpha\|\boldsymbol{\beta}\|_2^2 + (1-\alpha)\|\boldsymbol{\beta}\|_1$, $\lambda = \lambda_1 + \lambda_2$,and let $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$. The elastic net procedure can be viewed as a penalized least squares method. The elastic net penalty $\mathcal{J}(\boldsymbol{\beta}) = \alpha\|\boldsymbol{\beta}\|_2^2 + (1-\alpha)\|\boldsymbol{\beta}\|_1$ is a convex combination of the lasso and ridge penalties. When $\alpha = 1$, the naïve elastic net becomes simple ridge regression. If $\alpha = 0$ then $\lambda_2 = 0$, the naïve elastic net becomes the lasso. Solving $\widehat{\boldsymbol{\beta}}$ in equation (2.5) is equivalent to the optimization problem

$$\widehat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}}\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2, \text{ subject to } \alpha\|\boldsymbol{\beta}\|_2^2 + (1-\alpha)\|\boldsymbol{\beta}\|_1 \leq t \text{ for some } t.$$

The elastic net has the Bayesian connection. The elastic net penalty corresponds to a new prior given by

$$p_{\lambda,\alpha}(\boldsymbol{\beta}) = c(\lambda,\alpha)exp\{-\lambda[\alpha\|\boldsymbol{\beta}\|_2^2 + (1-\alpha)\|\boldsymbol{\beta}\|_1]\}, \tag{2.6}$$

which is a compromise between the Gaussian and Laplacian priors. Zou and Hastie (2005) pointed out that the elastic net estimator can be viewed as the Bayes posterior mode of $\boldsymbol{\beta}$ under the prior in (2.6).

The final elastic net estimate is taken to be a rescaled version of the naïve estimate,

$$\widehat{\boldsymbol{\beta}}_{\text{elastic net}} = (1+\lambda_2)\widehat{\boldsymbol{\beta}}_{\text{Naïve elastic net}} \tag{2.7}$$

The scaling was introduced by Zou et al. (2005) to reduce perceived overshrinkage of the naïve estimate. Hence, the *elastic net estimator* is defined as follows:

$$\widehat{\boldsymbol{\beta}}_{\text{elastic net}} = (1+\lambda_2)\left\{\arg\min_{\boldsymbol{\beta}}[\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda_2\|\boldsymbol{\beta}\|_2^2 + \lambda_1\|\boldsymbol{\beta}\|_1]\right\}, \tag{2.8}$$

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are the penalty parameters, $\|\boldsymbol{\beta}\|_1 = \sum_{j=1}^{p}|\beta_j|$ is the L_1 norm of $\boldsymbol{\beta}$, and $\|\boldsymbol{\beta}\|_2^2 = \sum_{j=1}^{p}\beta_j^2$ is the L_2 norm of $\boldsymbol{\beta}$. The L_1 part of the elastic net performs automatic variable selection, while the L_2 part stabilizes the solution parts and, hence, improves the prediction. The comparative study of Zou et al. (2005) showed that the elastic net has better performance than the lasso, and the elastic net produces a sparse model with good prediction accuracy (minimum median mean squared error), while encouraging a grouping effect. Furthermore, the elastic net appears to perform well on microarray data in terms of the misclassification error, and it does automatic gene selection. When the correlations among the predictors become high, the elastic net can significantly improve the prediction accuracy of the lasso.

The penalty parameters (λ_1 and λ_2) control the amount of shrinkage imposed on the coefficients, where some weak effects are forced to be exactly zero if the shrinkage level is large enough. Hence, the penalty parameters (λ_1 and λ_2) can be named the shrinkage parameters. If the value of λ_1 and λ_2 are too small, then no shrinkage will be performed. If the value of λ_1 and λ_2 are too high, then all coefficients will be shrunk to zero. Hence, the penalty parameters (λ_1 and λ_2) are important for the elastic net estimator.

For elastic net method proposed by Zou and Hastie (2005), the penalty parameters λ_1 and λ_2 are selected by K-fold cross-validation method. To avoid intensive computation, a grid of values for λ_2 is first specified. Zou et al. (2005) suggested to pick a relatively small grid value of λ_2. They used (0, 0.01, 0.1, 1, 10, 100, 1000). For each λ_2, a 10-fold cross-validation is then used to choose λ_1. The chosen λ_2 is the one giving the smallest cross-validation error.

The elastic net inherits sparsity property of the lasso, so the degree of freedom of the elastic net is the number of nonzero coefficients. There is no closed form expression for the elastic net estimator. For the computation of the elastic net, Zou and Hastie (2005) provided the algorithm LARS-EN to solve the elastic net efficiently based on LARS of Efron, Hastie, Johnstone, and Tibshirani (2004). Since LARS-EN is applied from LARS of Efron et al. (2004), hence the elastic net can be extended to generalized linear models. The elastic net becomes popular. Later, Friedman, Hastie, and Tibshirani (2010) developed the cyclical coordinate descent algorithms for estimation of generalized linear models with elastic net penalties.

2.3 Adaptive elastic net

The elastic net has good performance. However, it does not enjoy the oracle properties (consistency in variables selection and asymptotic normality). Zou and Zhang (2009) and Ghosh (2011) proposed two adaptive elastic net estimators which have the oracle property (The oracle property of two adaptive elastic net estimators show in Appendix A). These two adaptive elastic net estimators are different.

2.3.1 Adaptive elastic net (Zou and Zhang, 2009)

Zou and Zhang (2009) proposed the adaptive elastic net using the elastic net estimator to construct the adaptive weight. This adaptive elastic net is designed for high-dimensional data analysis. It has the oracle property and outperforms the elastic net. Assuming that the response is centered and the predictors are standardized, the adaptive elastic net estimator proposed by Zou et al. (2009) is defined as follows:

$$\widehat{\boldsymbol{\beta}}_{\mathrm{AENET2009}} = (1+\lambda_2)\left\{\arg\min_{\boldsymbol{\beta}}\left[\|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda_2\|\boldsymbol{\beta}\|_2^2 + \lambda_1 \textstyle\sum_{j=1}^{p} \widehat{w}_j \left|\beta_j\right|\right]\right\}. \tag{2.9}$$

The adaptive weight $\widehat{w}_j = \left(\left|\hat{\beta}_j(\text{elastic net})\right|\right)^{-\gamma}$, $j = 1, \dots, p,$ where γ is a positive constant. Zou and Zhang (2009) used the concept of Zou, Hastie, and Tibshirani (2007) to define the degree of freedom for their adaptive elastic net, so the degree of freedom is the number of nonzero coefficients (this similar to the elastic net method). Zou et al. (2009) selected the penalty parameters by using Bayesian Information Criterion (BIC) which applied from Zou et al. (2007) but they did not give the detail of the process for choosing the value of λ_1 and λ_2. Zou et al. (2009) considered the problem of model selection and estimation in situations where the number of parameters diverges with the sample size. The estimation accuracy was measured by the mean square error. The variable selection performance was gauged by (C, IC), where C is the number of zero coefficients that are correctly estimated by zero and IC

is the number of nonzero coefficients that are incorrectly estimated by zero. The comparative study showed that the adaptive elastic net of Zou et al. (2009) has best performance and deals with the multicollinearity problem better than the other-like methods such as the lasso, the elastic net, the adaptive lasso proposed by Zou, (2006), and the smoothly clipped absolute deviation penalty (SCAD) proposed by Fan and Li (2001).

2.3.2 Adaptive elastic net (Ghosh, 2011)

Ghosh (2011) proposed the adaptive elastic net using the least squares estimator to construct the adaptive weight. This method has good performance on grouped selection and model complexity than the elastic net. Assuming that the response is centered and the predictors are standardized, the adaptive elastic net estimator proposed by Ghosh (2011) is defined as follows:

$$\hat{\boldsymbol{\beta}}_{\mathrm{AENET2011}} = (1+\lambda_2)\left\{\arg\min_{\boldsymbol{\beta}}\left[\|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda_2 \sum_{j=1}^{p}\left(|\beta_j| + \frac{\lambda_1}{2\lambda_2}\hat{w}_j\right)^2 - \frac{(\lambda_1)^2}{4\lambda_2}\sum_{j=1}^{p}\hat{w}_j^2\right]\right\}. \tag{2.10}$$

The adaptive weight vector $\hat{\mathbf{w}} = 1/\left|\hat{\boldsymbol{\beta}}_{least\ square}\right|^{\gamma}$, for some $\gamma > 0$. Ghosh (2011) used the concept of Efron, Hastie, Johnstone, and Tibshirani (2004) and Zou, Hastie, and Tibshirani (2007) to defined the degree of freedom for this adaptive elastic net as

$df(\hat{\mathbf{y}})$ = number of nonzero coefficients selected by lasso
+ number of other predictors not elected by lasso but
having "high correlation" with lasso selected predictors.

Ghosh (2011) reformulated the LARS-EN algorithm for this adaptive elastic net method. The adaptive elastic net method proposed by Ghosh (2011) used the same method of Zou and Hastie (2005) for selection the value of λ_1 and λ_2. The relative prediction error $(E\,[(\hat{\mathbf{y}}-\mathbf{X}\boldsymbol{\beta})^2]/\sigma^2)$ was used to measure the performance. The comparative study showed that the adaptive elastic net of Ghosh (2011) has better performance for group selection than other method such as ordinary least squares method, ridge, the lasso, adaptive lasso, and elastic net method.

There is no any comparative study between two adaptive elastic net methods proposed by Ghosh (2011) and Zou and Zhang (2009).

For penalized regression, doing statistical inference for the coefficients is difficult. The bootstrap methods and the approximate covariance matrix are used for standard error estimation in penalized regression models. A number of people have considered the standard error estimation for the lasso estimator such as Fan and Li (2001); Knight and Fu (2000); Osborne, Presnell and Turlach (2000);Tibshirani (1996). However, they fail to provide reasonable standard error estimates for the parameter estimated to be zero (Kyung, Gill, Ghosh, & Casella, 2000; Li & Lin, 2010; Park & Casella, 2008). This causes some researchers adopt the lasso and related methods into the Bayesian framework.

The penalized regression such as the lasso and elastic net have Bayesian interpretation, so many researchers adopted the lasso and elastic net in the Bayesian frameworks over the past years.

Park and Casella (2008) illustrated the Bayesian lasso regression by adopting the double exponential prior as a mixture of normal and exponential prior. They extended the Bayesian lasso regression model to account for model uncertainty in the hyperparameters by placing prior distribution on σ^2 and τ^2. They used conjugate normal priors for the regression parameters and independent exponential prior on the variances. They suggested Gibbs sampling for choosing the penalty parameter λ of the lasso with the Laplace prior in the hierarchical model. This Gibbs sampling was applied from Casella (2001) who proposed a Monte Carlo EM algorithm that complements a Gibbs sampler and provides marginal maximum likelihood estimates of hyperparameters. The method for choosing λ proposed by Park et al. (2008) were not intended for direct application to the ordinary lasso but could aid in choosing λ for the lasso. This Bayesian lasso provided the interval estimates (Bayesian credible intervals) that can guide variable selection. However, this approach did not directly implement covariate selection but performed only shrinkage of the regression coefficients towards zero. Park et al. (2008) obtained point estimates of the regression coefficients using the median of the posterior distribution, but they did not address prediction of future observations. The posterior median and 95% credible intervals for the regression parameters were used to measure the performance. The comparative study of Park et al. (2008) showed that results from this Bayesian lasso are strikingly similar to those from the ordinary lasso.

Hans (2009) proposed Bayesian lasso regression and a new Gibbs sampling for Bayesian lasso regression. He imposed directly the double exponential prior on the lasso regression coefficients and a gamma prior on the shrinkage parameter. Emphasis was placed on point estimation using the posterior mean, which facilitated prediction of future observations via the posterior predictive distribution. The average test errors were computed to measure the predictive performance. The comparative study of Hans (2009) showed that the standard lasso prediction method does not necessary agree with model-based, Bayesian predictions.

Model uncertainty of the Bayesian lasso regression (Hans, 2009) was addressed in Hans (2010) by computing exactly the marginal posterior probabilities for small model spaces. He described methods that can be used to evaluate the posterior distribution over the space of all possible regression models for Bayesian lasso regression and described how the marginal likelihood can be accurately computed when the number of predictor is not too large. He handled the cases of large model spaces by imposing a mixture of a mass at zero and of a double exponential prior and estimated the posterior inclusion probabilities by using a Gibbs sampler. Running times were used to discuss the performance of the proposed method. The comparative study of Hans (2010) showed that this Gibb sampling approach is similar in spirit to the stochastic search variable selection methods.

Bornn, Gottardo, and Doucet (2010) discussed that the form of prior distribution which combine and compromise between Laplace and Gaussian prior support the Bayesian elastic net to reveal the grouping effect – the ability of the model to ensure highly correlated variables are assigned similar regression coefficients. They used the maximum a posteriori (MAP) estimate which obtained using Newton-

Raphson to perform cross-validation to obtain approximately optimal choices of the penalty parameters of the elastic net. The mean squared prediction error was used to measure the performance.

Kyung, Gill, Ghosh, and Casella (2010) proposed hierarchical models and Gibbs sampling to get estimators and valid Bayesian standard errors of prediction for generalized lasso estimators such as lasso, group lasso, fused lasso, and elastic net. The penalty parameters were choosing by Gibbs sampling approach proposed by Park and Casella (2008). They were not specifically interested in variable selection, but rather in accurate prediction and determining which predictors are meaningful. The median mean squared error and average mean squared error were used to measure the performance. The comparative study of Kyung et al. (2010) showed that the Bayesian hierarchical lassos performed as well as, or better than, the LARS fit in most of the examples. For the method of choosing the penalty parameters λ, the result found that putting λ into the Gibbs sampler seemed to be as effective as choosing it by cross-validation.

Li and Lin (2010) proposed the Bayesian method to solve the elastic net model using a Gibbs sampler where the prior information was a compromise between normal and double exponential priors and the penalty parameters were chosen through Gibbs sampling method that maximizes the data marginal likelihood. Li et al. (2010) chose the penalty parameters λ_1 and λ_2 simultaneously by a Monte Carlo EM algorithm proposed by Casella (2001). Li et al. (2010) suggested two variable selection criterions. The first criterion was called the credible interval criterion. A predictor x_j is excluded if the credible interval of β_j covers 0 and is retained otherwise. The receiver operating characteristic (ROC) curve was used to measure the variable selection accuracy. The second criterion was called the scaled neighborhood criterion. They considered the posterior probability of β_j in the interval $\left[-\sqrt{var(\beta_j|\mathbf{y}}, \sqrt{var(\beta_j|\mathbf{y}}\right]$. A predictor is excluded if the posterior probability exceeds a certain probability threshold and is retained otherwise. The median of the prediction mean squared errors was used to compare the prediction accuracy. They compared the performance of the proposed Bayesian elastic net with the Bayesian lasso (Park & Casella, 2008), elastic net (Zou & Hastie, 2005), and the lasso

(Tibshirani, 1996) in prediction accuracy and variable selection. The comparative study of Li et al. (2010) showed that the Bayesian elastic net behaves comparably in prediction accuracy but performs better in variable selection.

Hans (2011) developed Bayesian connection to the elastic net regression. Two characterizations of the class of prior distributions of $\boldsymbol{\beta}$ were introduced: a properly normalized, direct characterization, which was shown to be conjugate for linear regression models, and an alternate representation as a scale mixture of normal distributions. Hans (2011) gave the advantages of his proposed prior. Since the prior was shown to be conjugate, resulting in a direct characterization of the posterior distribution of $\boldsymbol{\beta}$. The prior was also shown to be represented as a scale mixture of normal distributions, so this representation allowed for connections to be made to other Bayesian regression shrinkage priors. Posterior inference of $\boldsymbol{\beta}$ was achieved using Markov chain Monte Carlo (MCMC) methods. Hans (2011) did not propose the new method for estimation the penalty parameters. The 10-fold cross-validation was used to choose the values of penalty parameters λ_1 and λ_2 as the classical method proposed by Zou and Hastie (2005). The comparative study of five estimators were performed: $\hat{\boldsymbol{\beta}}_{\mathrm{B}}$ – Bayesian elastic net estimator, $\hat{\boldsymbol{\beta}}_{\mathrm{BMA}}$ – the model average posterior mean of $\boldsymbol{\beta}$, $\hat{\boldsymbol{\beta}}_{\mathrm{E}}$ – the usual rescaled elastic net estimate of Zou et al. (2005), $\hat{\boldsymbol{\beta}}_{\mathrm{BL}}$ – the posterior median of $\boldsymbol{\beta}$ under the Bayesian lasso model of Park and Casella (2008), and $\hat{\boldsymbol{\beta}}_{\mathrm{LS}}$ – the least squares estimate of $\boldsymbol{\beta}$ based on a subset of the predictors that was chosen via stepwise model selection. The predictive performance was achieved via the prediction mean squared error (MSE), $(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T \mathbf{V} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ where $\mathbf{V}$ is the population covariance matrix of $\mathbf{X}$. The comparative study of Hans (2011) showed that the two Bayesian elastic net estimators $\hat{\boldsymbol{\beta}}_{\mathrm{B}}$ and $\hat{\boldsymbol{\beta}}_{\mathrm{BMA}}$ perform well, providing prediction MSEs that are comparable to – and in most cases better than – the traditional elastic net estimate $\hat{\boldsymbol{\beta}}_{\mathrm{E}}$. The Bayesian lasso prediction errors were dominated by the Bayesian elastic net prediction errors in the simulation study, providing evidence that orthant (truncated) normal prior proposed by Hans (2011) allowed for more flexible parameter estimation (and thus prediction) compared with its limiting case, the double-exponential prior. Hans (2011) investigated the model-averaged estimator further in a highly correlated, high-dimensional example. He suggested that when there is considerable uncertainty about model specification (e.g., when the predictor variables were all very highly correlated) prediction under the elastic net framework can be improved by averaging over the model space posterior distribution rather than focusing on a single point estimate corresponding to a posterior mode. Hans (2011) did not consider the question of variable selection directly in this article.

Lykou and Ntzoufras (2012) proposed a Bayesian implementation of the lasso regression that accomplishes both shrinkage and variable selection. They proposed the method for estimating the penalty parameter λ in the lasso technique by using the behavior of Bayes factors. The penalty parameter λ was associated with the values of Pearson and partial correlation at the limits between significance and insignificance as defined by Bayes factors. In this way, a meaningful interpretation of λ is achieved that leads to a simple specification of this parameter. Moreover, they used these values to specify the parameters of a gamma hyperprior distribution for λ. The parameters of the hyperprior are elicited such that appropriate levels of practical significance of the Pearson correlation are achieved and, at the same time, the prior support of λ values that activate the Lindley-Bartlett paradox (when λ is small) or lead to over-shrinkage (when λ is large) of model coefficients is avoided. Lindley-Bartlett paradox is a situation which the Bayesian methods and significance tests support different hypotheses. The behavior of Lindley's paradox was reported by Lindley (1957) and Bartlett (1957) commented on Lindley's paradox. This paradox is the sensitivity of Bayes factor to the prior distribution. To avoid activating the Lindley-Bartlett paradox, the improper prior cannot be used and a large prior variance must be carefully selected (Ntzoufras, 2009).

The comparative study of Lykou et al. (2012) showed that their proposed method work efficiently for selection the important predictors and identifying the best model with increased precision. Their proposed method can identify correctly both the important predictors with high posterior inclusion probabilities (higher than 0.5) and the non-important ones with low posterior probabilities (away from 0.5). The method proposed by Lykou et al. (2012) differs from the *k*-fold cross-validation method and BIC.

After study and analyze the literature review, the above ideas motivate the first new idea in this thesis to propose the new method for estimating the penalty parameter (λ_2) of the elastic net and the adaptive elastic net linear regression models by using Bayes factor and Pearson correlation between y and the candidate predictor x_j.

The penalized regression approaches such as the lasso and the elastic net can lead to finding smaller groups of variables with good prediction accuracy since their process shrunk some coefficients to zero, providing a form of variable selection. Celeux, Anbari, Marin, and Robert (2012) showed that the Bayesian variable selection methods perform better variable selection efficiency than the lasso and the elastic net. In the Bayesian framework, the model selection problem is transformed to the form of parameter estimation, rather than searching for the single optimal model, a Bayesian will attempt to estimate the posterior probability of all models within the considered class of models (O'Hara, & Sillanpää, 2009).

The fully Bayesian approach to variable selection is as follows (George, 1999; cited by George, 2000). For a given set of models $M_1, \ldots, M_{2^p}$, where M_γ corresponds to the γth subset of $\mathbf{X}_1, \ldots, \mathbf{X}_p$, one puts priors $\pi(\boldsymbol{\beta}_\gamma | M_\gamma)$ on the parameters of each M_γ and a prior on the set of models $\pi(M_1), \ldots, \pi(M_{2^p})$. Selection is then based on the posterior model probabilities $\pi(M_\gamma|\mathbf{Y})$, which are obtained in principle by Bayes's theorem (George, 2000). Insofar as the priors $\pi(\boldsymbol{\beta}_\gamma | M_\gamma)$ and $\pi(M_\gamma)$ provide an initial representation of model uncertainty, the model posterior $\pi(M_\gamma|\mathbf{Y})$ provide a complete representation of post-data model uncertainty that can be used for a variety of inferences and decisions. By treating $\pi(M_\gamma|\mathbf{Y})$ as a measure of the "truth" of model M_γ , a natural and simple strategy for model selection is to choose the most probable M_γ , the model for which $\pi(M_\gamma|\mathbf{Y})$ is largest (Clyde & George, 2004b).

In principle, the Bayesian approach for selecting a model is to choose the model with the largest posterior probability among a set of candidate models (Ando, 2010). Berger and Pericchi (2001, p.141) commented on Bayesian model selection as follows. Often, one is constrained to select a single model that will be used for subsequent prediction and, somewhat surprisingly, it is not always optimal to select the model with the largest posterior probability. The largest posterior probability model is optimal under very general conditions if only two models are being entertained (Berger, 1999) and is often optimal for variable selection in linear models having orthogonal design matrices (Clyde & George, 2000, cited by Berger & Pericchi, 2001). For other cases, such as in nested linear models (e.g. polynomial regression), Barbieri and Berger (2001, 2004) defined and illustrated that the optimal single model for prediction is the median probability model. However, in problems where no single model stands out, it may be preferable to report a set of models with high posterior probability along with their probabilities to convey the model uncertainty (Clyde & George, 2004b).

The amount of literature on Bayesian variable selection is quite enormous. Some of papers which propose procedures related to hierarchical Bayes model include George and McCulloch (1993, 1997) who concentrated on univariate Gaussian regression.

George and McCulloch (1993) proposed SSVS (stochastic search variable selection) procedure that uses probabilistic considerations for selecting promising subsets of $\mathbf{X}_1, \ldots, \mathbf{X}_p$ for further consideration. The promising subsets of predictors can be identified as those with higher posterior probability. SSVS is based on embedding the entire regression setup in a hierarchical Bayes mixture model where latent variables are used to identify subset choices. To avoid the overwhelming problem of calculating the posterior probabilities of all 2^p subsets, SSVS proceeds by using Gibbs sampling to indirectly sample from the posterior distribution.

George and McCulloch (1997) described, compared and applied a variety of approaches to Bayesian variable selection in normal linear regression models which included the nonconjugate stochastic search variable selection (George & McCulloch, 1993) as well as conjugate formulations which allow for analytical simplification. These approaches used hierarchical mixture priors to describe the uncertainty present in variable selection problems. Hyperparameter settings which base selection on practical significance, and the implications of using mixtures with point priors were discussed. Conjugate versions of these priors were shown to yield posterior expressions which can sometimes be sequentially computed using efficient updating schemes. The methods of George and McCulloch (1993, 1997) were applied in the other papers. For example, Brown, Vannucci, and Fearn (1998) developed multivariate Bayesian variable selection which is a generalization of George and McCulloch (1993, 1997). Brown, Vannucci, and Fearn (2002) proposed Bayes model averaging incorporating variable selection for prediction.

A general characteristic of hierarchical modeling is that it improves the robustness of the resulting Bayes estimators: while still incorporating prior information, the estimators are also well performing from a frequentist point of view (minimaxity and admissibility), although these two requirements are usually difficult to reconcile (Robert, 2001). The hierarchical structure simplifies both the interpretation and the computation of the model since the corresponding posterior distribution is simplified, resulting in conditional distributions of simpler form. This allows for the implementation of simpler Gibbs-based sampling schemes (Ntzoufras, 2009).

Clyde (2001, p.123), in paper of Chipman, George, and McCulloch (2001), discussed the Bayesian variable selection for the case number of predictor (p) greater than the sample size (n) such as experiments using gene-array technology. She commented that identifying which genes are associated with outcomes (or responses) is a challenging problem for Bayesian model selection, from both a computational standpoint, as well as the choice of prior distribution. Chipman et al. (2001, p.132) suggested to the topic of Clyde (2001) that, for fast Bayesian computational method in high dimensional problem, it may be useful to combine heuristic strategies with Bayesian methods.

Kyung, Gill, Ghosh, and Casella (2010), Li and Lin (2010), and Hans (2011) proposed Bayesian method to solve elastic net model using Gibbs sampling. They focused on the elastic net estimator and the posterior probability of $\boldsymbol{\beta}$. They did not study the posterior model probability of candidate models or Bayesian variable selection. Hans (2011) chose the values of penalty parameters λ_1 and λ_2 by 10-fold cross-validation as the classical method proposed by Zou and Hastie (2005). Kyung, et al. (2010) and Li et al. (2010) put the penalty parameters λ_1 and λ_2 into the Gibbs sampler. They assigned the gamma prior distribution to the penalty parameters (λ_1

and λ_2) and used the Gibbs sampling as suggested by Park and Casella (2008) and the Monte Carlo EM algorithm (Casella, 2001) to estimate λ_1 and λ_2. Kyung, et al. (2010) gave a comment on the method for choosing the tuning parameter λ, and showed that putting λ into the Gibbs sampler seems to be as effective as choosing it by cross-validation.

This motivates the second new idea in this thesis to combine Bayesian variable selection method with elastic linear regression model where the penalty parameter (λ_2) is specified by the method proposed in this thesis.

2.4 K-fold cross-validation

K-fold cross-validation is the primary method used for estimating a penalty parameter in the lasso, the elastic net, and the adaptive elastic net techniques. The idea of cross-validation is to split the data into k roughly equal parts, using a portion of the data to build (or "train") the model and the remainder to validate the model. For each training set, the researcher fits the model with a penalty parameter λ and computes its cross-validation error. The researcher does the same process for many values of λ and chooses the value of λ that gives the smallest cross-validation error. The steps are the following (Hastie, Tibshirani, & Friedman, 2009):

- Divide the data into k roughly equal parts, for example:

1	2	…	$k-1$	k
validation	train	train	train	train

- Let m be the validation set, fit the model with parameter λ using the training set (the remaining $k-1$ set), giving $\hat{\beta}_m(\lambda)$ and compute its error in predicting the mth part (validation set):

$$E_m(\lambda) = \sum_{i\,\epsilon\,m\text{th part}} \left(y_i - x_i\hat{\beta}_m(\lambda)\right)^2 \qquad (2.11)$$

- Repeat step 1 for $m = 1,2, \dots, k$.

- Compute the cross-validation error:

$$CV(\lambda) = \frac{1}{k}\sum_{m=1}^{k} E_m(\lambda) \tag{2.12}$$

- Do this for many value of λ and choose the value of λ that makes $CV(\lambda)$ smallest.

The different number of k may cause the different value of λ for the same dataset. Hence, the number of subset k affects to the penalty parameter. It is interesting to wonder about what is the appropriate k-fold cross-validation. The case $k = n$ is known as leave-one-out cross validation. With $k = n$, the cross-validation estimator is approximately unbiased for the true (expected) prediction error, but can have high variance because the n "training set" are so similar to one another. With $k = 5$, cross-validation has lower variance but bias could be a problem. Hastie, Tibshirani, and Friedman (2009) suggested that typical choices of k are 5 or 10. Hyndman (2010) suggested the researcher should beware of looking at statistical tests after selecting variables using cross-validation since the tests do not take into account of the variable selection that has taken place and so the p-value can mislead. Cross-validation can be misused when there are exact duplicate observations (two or more observations with equal values for all predictors and for the response variable).

2.5 Bayesian Information Criterion (BIC) for lasso

Bayesian Information Criterion (BIC) for lasso was developed by Zou, Hastie, and Tibshirani (2007). Given a model fitting method δ, let $\hat{\boldsymbol{\mu}} = \delta(\mathbf{y})$ represent its fit and $\mathbf{y}$ is a vector of response variable. Following the usual definition of BIC proposed by Schwarz (1978), Zou, et al. (2007) proposed BIC for the lasso as

$$\mathrm{BIC}(\hat{\boldsymbol{\mu}}) = \frac{\|\mathbf{y}-\hat{\boldsymbol{\mu}}\|^2}{n\sigma^2} + \frac{\log(n)}{n}\widehat{df}(\hat{\boldsymbol{\mu}}). \tag{2.13}$$

Zou, et al. (2007) suggested using BIC to select the penalty parameter of the lasso. For each value of the tuning (penalty) parameter λ considered, the BIC for the resulting model is computed. The model corresponding to the smallest BIC is chosen. Zou and Zhang (2009) used BIC (Zou, et al., 2007) to select the penalty parameters of the adaptive elastic net.

The BIC was first developed by Schwarz (1978), who gave a Bayesian argument for adopting it. For model selection purposes, BIC is asymptotically consistent as a selection criterion. What this means is that given a family of models, including the true model, the probability that BIC will select the correct model approaches one as the sample size $n \rightarrow \infty$. On the other hand, for finite samples, BIC often chooses models that are too simple, because of its heavy penalty on complexity (Hastie, Tibshirani, & Friedman, 2009). Asymptotically, for linear models minimizing BIC is equivalent to leave–υ–out cross-validation when $\upsilon = n\{1 - 1/[\log(n) - 1]\}$ (Shao, 1997, cited by Hyndman, 2010). BIC can be used to obtain a rough approximation of the log-Bayes factor under a wide family of prior distributions (Kass, & Raftery, 1995, cited by Ntzoufras, 2009).

Berger and Pericchi (2001) commented on BIC as the follows. The BIC approximation has the advantages of simplicity and an (apparent) freedom from prior assumptions. However, it is valid only for nice problems. For example, Shibata (1981), (cited by Berger et al., 2001), showed that BIC, a popular approximate Bayesian procedure, is not optimal for certain situations in which the true model is not in the candidate set. Stone (1979), (cited by Berger, Ghosh, & Mukhopadhyay, 2003), had observed that BIC can be inconsistent when the dimension of the parameter goes to infinity. Berger, Ghosh, et al. (2003) developed generalized Bayes information criterion (GBIC) that does not depend on the prior and valid for the situation considered in Stone (1979).

Clyde and George (2004b) commented on BIC as follows. One of the difficulties with using BIC, however, is determining the effective sample size n in non-independent settings, such as hierarchical models (Pauler, 1998; Pauler, Wakefield & Kass, 1999, cited by Clyde et al. 2004b). BIC is also not appropriate in problems where the number of parameters increases with the sample size or other irregular asymptotics prevails (Berger, Ghosh, & Mukhopadhyay, 2003, cited by Clyde et al. 2004b).

2.6 Bayes factor

The Bayes factor is a quantity for comparing models and for testing hypotheses in the Bayesian framework. Bayes factor has played a major role in assessing the goodness of fit of the competing models (Ando, 2010).

The Bayesian approach to hypothesis testing was developed by Jeffreys (1961) as a major part of his program for scientific inference. Although Jeffreys called his methods "significance tests", apparently borrowing the term from Fisher, this is misleading, because Jeffreys's perspective and goals were quite different. Jeffreys was concerned with the comparison of predictions made by two competing scientific theories. In his approach, statistical models are introduced to represent the probability of the data according to each of the two theories, and Bayes's theorem is used to compute the posterior probability that one of the theories is correct (Kass & Raftery, 1995).

Following Good (1958) and Jeffreys (1961), Kass and Raftery (1995) considered the Bayes factor on its own ground.

Bayes factor proposed by Kass et al. (1995) is the posterior odds of one hypothesis when the prior probabilities of the two hypotheses are equal.

Suppose there is data $\mathbf{D}$, assumed to have arisen under one of the two hypotheses H_1 and H_2 according to a probability density $p(\mathbf{D}|H_1)$ or $p(\mathbf{D}|H_2)$. Given a prior probabilities $p(H_1)$ and $p(H_2) = 1 - p(H_1)$, the data produce a posterior probabilities $p(H_1|\mathbf{D})$ and $p(H_2|\mathbf{D}) = 1 - p(H_1|\mathbf{D})$. From Bayes's theorem, the posterior probability $p(H_k|\mathbf{D})$ is

$$p(H_k|\mathbf{D}) = \frac{p(\mathbf{D}|H_k)p(H_k)}{p(\mathbf{D}|H_1)p(H_1)+p(\mathbf{D}|H_2)p(H_2)} \quad ; \; k = 1,2. \qquad (2.14)$$

So that

$$PO_{12} = \frac{p(H_1|\mathbf{D})}{p(H_2|\mathbf{D})} = \frac{p(\mathbf{D}|H_1)}{p(\mathbf{D}|H_2)} \times \frac{p(H_1)}{p(H_2)} = B_{12} \times \frac{p(H_1)}{p(H_2)} \qquad (2.15)$$

$$B_{12} = \frac{p(\mathbf{D}|H_1)}{p(\mathbf{D}|H_2)} \text{ is the Bayes factor.} \tag{2.16}$$

Thus, in words,

Posterior odds = Bayes factor × prior odds ,

and the Bayes factor is the ratio of the posterior odds of H_1 to its prior odds, regardless of the value of the prior odds. If the odds are greater than one, we choose the hypothesis H_1. Otherwise, we choose H_2.

In the simplest case, when the two hypotheses are single distributions with no free parameters (the case of "simple versus simple" testing), B_{12} is the likelihood ratio. In other cases, when there are unknown parameters under either or both of the hypotheses, the Bayes factor is still given by (2.16), and, in a sense, it continues to have the form of a likelihood ratio. Then, however, the densities $p(\mathbf{D}|H_k)$ (k = 1, 2) are obtained by integrating (not maximizing) over the parameter space, so that in Equation (2.16),

$$p(\mathbf{D}|H_k) = \int p(\mathbf{D}|\theta_k, H_k)\pi(\theta_k|H_k)d\theta_k \ , \tag{2.17}$$

where θ_k is the parameter under H_k, $\pi(\theta_k|H_k)$ is its prior density, and $p(\mathbf{D}|\theta_k, H_k)$ is the probability density of $\mathbf{D}$ given the value of θ_k, or the likelihood function of θ. (θ_k can be a vector with dimension d_k.)

The quantity $p(\mathbf{D}|H_k)$ given by Equation (2.17) is the marginal probability of the data, because it is obtained by integrating the joint density of $(\mathbf{D}, \theta_k)$ given $\mathbf{D}$ over θ_k. It is also the predictive probability of the data; that is, the probability of seeing the data that actually were observed, calculated before any data became available. It is also sometimes called a marginal likelihood, or an integrated likelihood. Note that, as in computing the likelihood ratio statistics but unlike in some other applications of likelihood, all constants appearing in the definition of the likelihood $p(\mathbf{D}|\theta_k, H_k)$ must be retained when computing B_{12}. In fact, B_{12} is closely related to the likelihood ratio statistic, in which the parameters θ_k are eliminated by maximization rather than by integration (Kass & Raftery, 1995, p.176).

If we consider the model comparison as a hypothesis testing problem where interest lies in evaluating the null hypothesis H_0 (corresponding to a model m_0) against the alternative H_1 (corresponding to a model m_1), then both the posterior model odds PO_{10} and the corresponding Bayes factor B_{10} evaluate the evidence against the null hypothesis, which familiar to classical significance tests. On the other hand, PO_{01} and B_{01} evaluate the evidence in favor of the null hypothesis, which is not feasible in classical significance tests. To summarize, using posterior model odds and Bayes factors, we can (1) evaluate the evidence in favor of the null hypothesis, (2) compare two or more non-nested models, (3) draw inferences without ignoring model uncertainty, and (4) determine which set of explanatory variables gives better predictive results (Ntzoufras, 2009).

Jeffreys (1961) developed a scale to judge the evidence in favor of or against H_0 brought by data, outside a true decision-theoretic setting (Robert, 2001). Jeffreys (1961) suggested interpreting B_{10} in half-units on the $\log_{10}$ scale. Table 2.1 gives Jeffreys' scale.

Table 2.1 Jeffreys' scale of evidence for Bayes factor

$\log_{10}(B_{10})$	B_{10}	Evidence against H_0
$0 < \log_{10}(B_{10}) < 0.5$	$1 < B_{10} < 3.2$	Not worth more than a bare mention (poor)
$0.5 < \log_{10}(B_{10}) < 1$	$3.2 < B_{10} < 10$	Substantial
$1 < log_{10}(B_{10}) < 2$	$10 < B_{10} < 100$	Strong
$\log_{10}(B_{10}) > 2$	$B_{10} > 100$	Decisive

Kass and Raftery (1995) considered twice the natural logarithm of the Bayes factor, which is on the same scale as the familiar deviance and likelihood ratio test statistics. Rounding and using 20 rather than 10 as the requirement for strong evidence, they gave the guidelines for interpreting the Bayes factor as in Table 2.2.

Table 2.2 Bayes factor interpretation according to Kass and Raftery (1995)

$2\log_e(B_{10})$	B_{10}	Evidence against H_0
$0 < 2\log_e(B_{10}) < 2$	$1 < B_{10} < 3$	Not worth more than a bare mention
$2 < 2\log_e(B_{10}) < 6$	$3 < B_{10} < 20$	Positive
$6 < 2\log_e(B_{10}) < 10$	$20 < B_{10} < 150$	Strong
$2\log_e(B_{10}) > 10$	$B_{10} > 150$	Very strong

Kass et al. (1995) suggested that, from their own experience, these categories seem to furnish appropriate guidelines. The data provide substantial evidence in favor of H_1 for B_{10} > 3. The evidence becomes stronger for higher values of the Bayes factor (and posterior odds).

Raftery (1999, pp.423-424) described the reasons for using the Bayes factor in the following. Bayes factor provide the Bayesian solution to the question, "What evidence do the data provide for one model against another, competing model?" expressed as a ratio of posterior probabilities. This leads to several desirable properties. (1) The first is that the hypothesis-testing procedure defined by choosing the model with the higher posterior probability minimizes the total error rate; that is, the sum of Type I and Type II error rates (Jeffreys, 1961, pp.396-97; Kass, 1991, cited by Raftery, 1999). Note that frequentist statisticians sometimes recommend reducing the significance level in tests when the sample size is large; the Bayes factor does this automatically. (2) The second property is as follows. When there is model uncertainty (i.e., doubt about which is the best model to use), the Bayesian solution to inference about quantities of interest is Bayesian model averaging; that is, averaging the posterior densities of the quantity of interest across the different models, with weights proportional to their posterior probabilities (derived from the Bayes factors). Madigan and Raftery (1994) have shown theoretically that this leads to optimal predictive performance, and this has been verified on real data in a series of studies summarized by Raftery, Madigan, and Volinsky (1995) (cited by Raftery, 1999).

Chapter 3

Research methodology

There are four objectives in this thesis. Two new ideas are proposed in the first and second objectives, and the last two objectives are the comparative study to analyze the performance of two proposed methods:

(1) To propose the new method for estimating the value of the penalty parameter (λ_2) of the elastic net and two adaptive elastic net linear regression models by using Bayes factor and Pearson correlation between y and the candidate predictor x_j .

(2) To combine a Bayesian variable selection with elastic net linear regression model where the prior distribution of $\boldsymbol{\beta}$ is a compromise between normal and double exponential distribution.

(3) To perform the comparative study between the new method which is proposed in the first objective of this thesis and the 10-fold cross-validation method for estimating the value of the penalty parameter (λ_2) of the elastic net and adaptive elastic net linear regression models. (The reason for using 10-fold cross-validation method is that it was used by Zou and Hastie (2005) for choosing the penalty parameters of the elastic net regression.)

(4) To perform the comparative study between the Bayesian variable selection for elastic net linear regression model which is proposed in the second objective of this thesis and elastic net linear regression model (Zou & Hastie, 2005) where the penalty parameter (λ_2) is estimated by the new method presented in the first objective of this thesis and the 10-fold cross-validation method.

In this thesis, the research methodology is divided into three sections.

3.1 Process of the method for estimating the value of the penalty parameter λ_2.

3.2 Process of the Bayesian variable selection for elastic net linear regression model.

3.3 Comparative study.

3.1 Process of the method for estimating the value of the penalty parameter $\boldsymbol{\lambda_2}$

The first objective of this thesis is to propose the method for estimating the value of the penalty parameter (λ_2) of the elastic net and two adaptive elastic net linear regression models by using Bayes factor and Pearson correlation between y and the candidate predictor x_j. The steps are the following.

In this process, the response and the predictor variables are transformed by the correlation transformation.

Step 1: Inclusion parameter

We begin by indexing each candidate model with one binary vector, $\boldsymbol{\gamma} = \left(\gamma_1, \ldots, \gamma_j, \ldots, \gamma_p\right)^T$. An element γ_j takes value 0 or 1 depending on whether or not the *j*th predictor is excluded from the model.

Let γ_j be an inclusion parameter; $\boldsymbol{\gamma} = \left(\gamma_1, \ldots, \gamma_j, \ldots, \gamma_p\right)^T$

$$\gamma_j = \begin{cases} 1 & ;\ \text{the parameter } j \text{ is in the regression model} \\ 0 & ;\ \text{the parameter } j \text{ is not in the regression model} \end{cases} \tag{3.1}$$

Hence, there are 2^p possible models $M_1, \ldots, M_{2^p}$ where $M_{\boldsymbol{\gamma}}$ corresponds to the $\boldsymbol{\gamma}$th subset of $\mathbf{X}_1, \ldots, \mathbf{X}_p$. Each submodel is of the form

$$M_{\boldsymbol{\gamma}}: \mathbf{y} = \mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \boldsymbol{\varepsilon}\,, \tag{3.2}$$

where $\mathbf{X}_{\boldsymbol{\gamma}}$ is an $n \times q_{\boldsymbol{\gamma}}$ design matrix whose columns correspond to the $\boldsymbol{\gamma}$th subset of $\mathbf{X}_1, \ldots, \mathbf{X}_p$, $\boldsymbol{\beta}_{\boldsymbol{\gamma}}$ is a $q_{\boldsymbol{\gamma}} \times 1$ vector of regression coefficients for the $\boldsymbol{\gamma}$th subset, $q_{\boldsymbol{\gamma}} \equiv \boldsymbol{\gamma}^T\mathbf{1}$ denotes the size of the $\boldsymbol{\gamma}$th subset, and $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$.

The linear regression model for each submodel of the form (3.2) is

$$\mathbf{y}|\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma} \sim N_n(\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2\mathbf{I}). \tag{3.3}$$

Step 2: Prior distribution for $\boldsymbol{\beta}$, σ^2, and $\boldsymbol{\gamma}$

In the process for estimating the value of λ_2, we use the hierarchical prior models in the form of submodel (3.2) as follows:

$$\mathbf{y}|\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma} \sim N_n(\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2\mathbf{I}),$$

$$\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^2, \boldsymbol{\gamma} \sim N_{q_{\boldsymbol{\gamma}}}(\mathbf{0}, \frac{\sigma^2}{\lambda_2}\mathbf{I}), \tag{3.4}$$

$$\sigma^2|\boldsymbol{\gamma} = \sigma^2 \sim \text{inverse gamma}\left(\frac{\nu}{2}, \frac{\nu\xi}{2}\right),$$

$$p(\boldsymbol{\gamma}) = \left(\frac{1}{2}\right)^p.$$

The reasons for using the prior in (3.4) are described in the following.

Prior distribution for β

For elastic net and adaptive elastic net estimate, the penalty parameter λ_2 is the L_2 part which is ridge penalty when $\lambda_1 = 0$.

Suppose $\mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ and each parameter β_j is distributed as $N(0, \tau^2)$, independently of one another, with τ^2 and σ^2 assumed known. Then the ridge estimate is the mode of the posterior distribution of $\boldsymbol{\beta}$ with $\lambda = \sigma^2/\tau^2$ (Hastie, Tibshirani, & Friedman, 2009). Hence, we assume

$$\beta_j|\sigma^2 \sim N\left(0, \frac{\sigma^2}{\lambda_2}\right) \text{ for } j = 1, \dots, p.$$

The prior distribution of $\boldsymbol{\beta}$ in matrix form is

$$\boldsymbol{\beta}|\sigma^2 \sim N_p(\mathbf{0}, \frac{\sigma^2}{\lambda_2}\mathbf{I}). \tag{3.5}$$

Prior distribution for σ^2

Assume σ^2 has inverse gamma prior distribution as suggested by Chipman, George and McCulloch (2001), George and McCulloch (1997), Hans (2011) and Park and Casella (2008). The inverse gamma prior for σ^2 would maintain conjugacy which gives the posterior distribution in the closed form.

$$\sigma^2 \sim \text{inverse gamma}\left(\frac{\nu}{2}, \frac{\nu\xi}{2}\right), \tag{3.6}$$

(which is equivalent to $\nu\xi/\sigma^2 \sim \chi^2_\nu$).

The prior (3.6) is determined by the hyperparameters ν and ξ which must be specified for implementations. Chipman et al. (2001, p.81) noted that (3.6) corresponds to the likelihood information about σ^2 provided by ν independent observations from a $N(0,\xi)$ distribution. Thus, ξ may be thought of as a prior estimate of σ^2 and ν may be thought of as the prior sample size associated with this estimate. By using the data and treating $S^2_{\mathbf{y}}$, the sample variance of $\mathbf{y}$, as a rough upper bound for σ^2, a simple default strategy is to choose ν small, say around 3, and ξ near $S^2_{\mathbf{y}}$. One might also go a bit further, by treating S^2_{FULL}, the traditional unbiased estimate of σ^2 based on saturated model, as a rough lower bound for σ^2, and then choosing ξ and ν so that (3.6) assigns substantial probability to the interval $(S^2_{FULL}, S^2_{\mathbf{y}})$. Alternatively, the explicit choice of ξ and ν can be avoided by using $p(\sigma^2) \propto 1/\sigma^2$, the limit of (3.6) as $\nu \to 0$, a choice recommended by Smith and Kohn (1996) and Fernandez, Ley and Steel (2001). This prior corresponds to the uniform distribution on $\log\sigma^2$, and is invariant to scale change in **y** (Chipman, George & McCulloch, 2001, p.81).

In this thesis, we choose $\nu = 3$ as suggested by Chipman et al. (2001) and choose $\xi = S^2_{FULL}$ (the traditional unbiased estimate of σ^2 based on saturated model). This choice of $\xi = S^2_{FULL}$ was used by George and McCulloch (1997).

Prior distribution for $\boldsymbol{\gamma}$

For the specification of the model space prior, most Bayesian variable selection implementations have used independence priors of the form

$$p(\boldsymbol{\gamma}) = \prod_{j=1}^{p} w_j^{\gamma_j} (1 - w_j)^{1-\gamma_j} \ . \tag{3.7}$$

Under this prior, each $\mathbf{X}_j$ enters the model independently of the other coefficients, with probability $p(\gamma_j = 1) = 1 - p(\gamma_j = 0) = w_j$. Smaller w_j can be used to down-weight $\mathbf{X}_j$ which are costly or of less interest (Chipman, George & McCulloch, 2001, pp.76-77).

In this thesis, we set $w_j = 1/2$ which yields the uniform prior

$$p(\boldsymbol{\gamma}) = \left(\frac{1}{2}\right)^{p}. \tag{3.8}$$

Chipman, George and McCulloch (2001, p.79) suggested that Bayes factor (corresponding to selection under uniform priors) would be preferable for pairwise model comparison.

Step 3: Posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$

Let $f(\boldsymbol{\gamma}|\mathbf{y})$ be the posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$.

Using the hierarchical prior models described in step 2 (3.4), we find $f(\boldsymbol{\gamma}|\mathbf{y})$.

Step 4: Bayes factor

For the Bayes factor used in this thesis, we define the hypotheses associated with the Bayes factor as

Hypotheses:

H_0: Reduced model (M_R)

versus

H_1: Full model (M_F),

where

M_R (Reduced model) is the linear regression model with the predictors $\mathbf{X}_{\boldsymbol{\gamma}}$,

M_F (Full model) is the linear regression model with the predictors $\mathbf{X}_{\boldsymbol{\gamma}}$ of the reduced model and additional predictor $\mathbf{X}_j$.

Suppose $\mathbf{X}_{\boldsymbol{\gamma}}$ be the predictor variables of the reduced model, so

$$M_R: \mathbf{y} = \mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \boldsymbol{\varepsilon} \,,$$
$$M_F: \mathbf{y} = \mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \mathbf{X}_j\boldsymbol{\beta}_j + \boldsymbol{\varepsilon} \,.$$

The example of hypotheses for Bayes Factor shows in Appendix B.

Let BF_{10} be the Bayes factor for comparing the evidence of model M_F versus model M_R.

In this thesis, we use the uniform prior $p(\boldsymbol{\gamma}) = \left(\frac{1}{2}\right)^p$. Thus, the prior model probabilities $f(M_F)$ and $f(M_R)$ are equal for all competing models. Hence,

$$\mathrm{BF}_{10} = \frac{f(M_F|\mathbf{y})}{f(M_R|\mathbf{y})} = \frac{\text{posterior model probability of } M_F \text{ given } \mathbf{y}}{\text{posterior model probability of } M_R \text{ given } \mathbf{y}} \tag{3.9}$$

If the BF_{10} is greater than one, we choose the hypothesis H_1. Otherwise, we choose H_0.

In thesis, we use Bayes factor interpretation as follows:

Table 3.1 Bayes factor interpretation

BF_{10}	Evidence against H_0
$1 < \mathrm{BF}_{10} < 3$	Negligible
$3 < \mathrm{BF}_{10} < 20$	Positive
$20 < BF_{10} < 150$	Strong
$\mathrm{BF}_{10} > 150$	Very strong

Table 3.1 shows the Bayes factor interpretation according to Kass and Raftery (1995).

Step 5: Specification of the penalty parameter λ_2 based on Bayes factor

Step 5.1:

We define the level $\mathrm{BF}_{10} = 1$ be the limit between significance and non-significance, and find the value of λ_2 which produce $\mathrm{BF}_{10} = 1$.

Step 5.2:

We define the level $\mathrm{BF}_{10} = 3$ be the level of practical significance and find the value of λ_2 which produce $\mathrm{BF}_{10} = 3$.

Step 5.3:

At the levels of Bayes factor ($\mathrm{BF}_{10} = 1$, $\mathrm{BF}_{10} = 3$), we can identify the set of predictors in the following.

Step 5.3.1: The value of $\mathrm{BF}_{10} = 1$ is used for identification the set of significance predictors ($\mathrm{BF}_{10} > 1$) and non-significance predictors ($\mathrm{BF}_{10} < 1$).

- Set of significance predictors is $\{X_j | X_j \in H_1 \text{ and } \mathrm{BF}_{10} > 1\}$.
- Set of non-significance predictors is $\{X_j | X_j \in H_1 \text{ and } \mathrm{BF}_{10} < 1\}$.

Step 5.3.2: The value of $\mathrm{BF}_{10} = 3$ is used for identification the set of *practical significance* predictors. We give the name of "the set of *practical significance* predictors" to be "the set of *important* predictors".

- Set of important predictors is $\{X_j |\, X_j \in\ H_1$ and $\mathrm{BF}_{10} > 3\}$.
- Set of non-important predictors is $\{X_j |\, X_j \in\ H_1$ and $1 < \mathrm{BF}_{10} < 3\}$.

Using Step 5, we obtain the value of λ_2 associated with the level of Bayes factor ($\mathrm{BF}_{10} = 1$, $\mathrm{BF}_{10} = 3$) when $\lambda_1 = 0$.

Step 6: Specification of the penalty parameter λ_2 for elastic net regression model

Step 6.1:

Using the value of λ_2 obtained from Step 5, we find the posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$, $g(\boldsymbol{\gamma}|\mathbf{y})$, when the hierarchical model prior for elastic net linear regression models are as follows.

$$\mathbf{y}|\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma} \sim N_n(\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2\mathbf{I}),$$

$$\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^2, \mathbf{D}_\tau, \boldsymbol{\gamma} \sim N_{q_\gamma}(\mathbf{0}, \sigma^2\mathbf{D}_\tau)\ ,$$

where $\mathbf{D}_\tau$ is a diagonal matrix with diagonal elements $\left(\tau_j^{-2} + \lambda_2\right)^{-1}$,

(3.10)

$$\tau_j^2 \sim Exponential\left(\frac{\lambda_1^2}{2}\right), j = 1,2,3, \dots, p,$$

$$\sigma^2|\boldsymbol{\gamma} = \sigma^2 \sim \text{inverse gamma}\left(\frac{\nu}{2}, \frac{\nu\xi}{2}\right),$$

$$p(\boldsymbol{\gamma}) = \left(\frac{1}{2}\right)^p.$$

For checking the validity of λ_1 and λ_2 used in Step 6.1, we use the relation

$$\alpha = \lambda_2/(\lambda_1 + \lambda_2)\ ;\ \alpha \in (0,1) \tag{3.11}$$

Hence, the value of λ_1 and λ_2 must satisfy $\alpha \in (0,1)$.

Step 6.2: Bayes factor for elastic net linear regression model ($\mathrm{BF}_{elastic\ net}$)

We define the hypotheses associated with the Bayes factor for elastic net linear regression model as

Hypotheses:

H_0: Reduced model $M_{R(elastic\ net)}$

versus

H_1: Full model $M_{F(elastic\ net)}$,

where

The reduced model $M_{R(elastic\ net)}$ is the elastic net linear regression model with the predictors $\mathbf{X}_{\boldsymbol{\gamma}}$.

The full model $M_{F(elastic\ net)}$ is the elastic net linear regression model with the predictors $\mathbf{X}_{\boldsymbol{\gamma}}$ of the reduced model and additional predictor $\mathbf{X}_j$.

Let $g\left(M_{R(elastic\ net)}|\mathbf{y}\right)$ be the posterior model probability of M_R given $\mathbf{y}$ using the elastic net prior (3.10).

Let $g\left(M_{F(elastic\ net)}|\mathbf{y}\right)$ be the posterior model probability of M_F given $\mathbf{y}$ using the elastic net prior (3.10).

Let $\mathrm{BF}_{elastic\ net}$ be the Bayes factor for comparing the evidence of model $M_{F(elastic\ net)}$ versus model $M_{R(elastic\ net)}$.

Thus, the Bayes factor for elastic net linear regression model is

$$\mathrm{BF}_{elastic\ net} = \frac{g(M_{F(elastic\ net)}|\mathbf{y})}{g(M_{R(elastic\ net)}|\mathbf{y})} \tag{3.12}$$

Step 6.3:

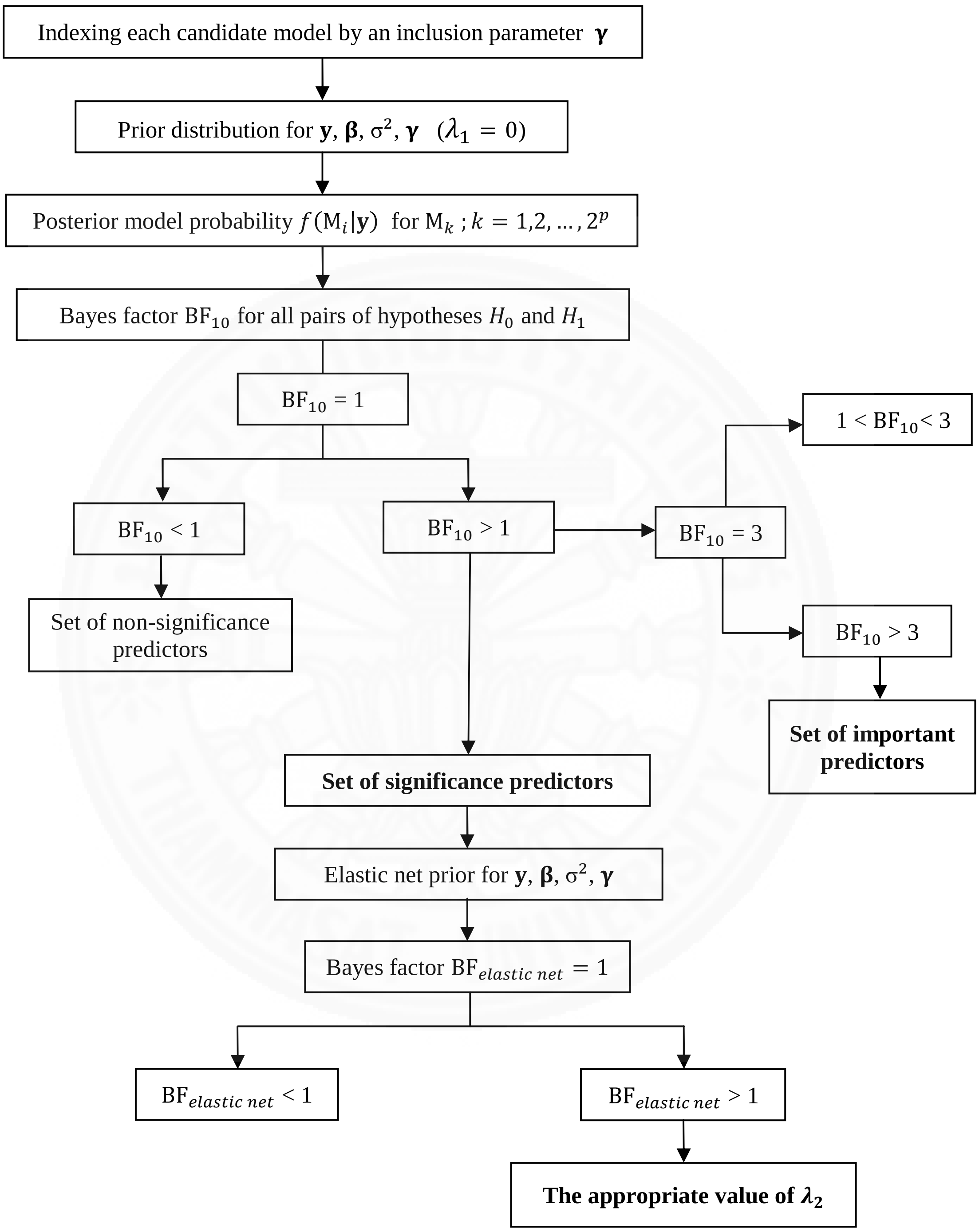
Using Bayes factor interpretation in Table 3.1, we observe the behavior of λ_2 as follows:

- The value of λ_2 which produce $\mathrm{BF}_{elastic\ net} > 1$,
- The value of λ_2 which produce $\mathrm{BF}_{elastic\ net} \in (1,3)$,
- The value of λ_2 which produce $\mathrm{BF}_{elastic\ net} > 3$.

Step 6.4:

By considering the value of λ_2 which associated with the level of $\mathrm{BF}_{elastic\ net} > 1$, we obtain the appropriate value of λ_2 for elastic net regression model.

Flowchart 3.1 Method for estimating the value of λ_2



3.2 Process of the Bayesian variable selection for elastic net linear regression model

The second objective of this thesis is to combine a Bayesian variable selection with elastic net linear regression model where the prior distribution of $\boldsymbol{\beta}$ is a compromise between normal and double exponential distribution.

In this thesis, we combine the Bayesian variable selection technique proposed by George and McCulloch (1997) and discussed in Chipman, George and McCulloch (2001) with elastic net regression model. We adopt a hierarchical Bayes framework of George and McCulloch (1997), but with new prior specifications for $\boldsymbol{\beta}$ and σ^2. The steps are the following.

In this process, the response and the predictor variables are standardized.

Step 1: Inclusion parameter

We begin by indexing each candidate model with one binary vector, $\boldsymbol{\gamma} = \left(\gamma_1, \dots, \gamma_j, \dots, \gamma_p\right)^T$. An element γ_j takes value 0 or 1 depending on whether or not the jth predictor is excluded from the model.

Let γ_j be an inclusion parameter; $\boldsymbol{\gamma} = \left(\gamma_1, \dots, \gamma_j, \dots, \gamma_p\right)^T$

$$\gamma_j = \begin{cases} 1 & ;\ \text{the parameter } j \text{ is in the regression model} \\ 0 & ;\ \text{the parameter } j \text{ is not in the regression model} \end{cases} \tag{3.13}$$

Hence, there are 2^p candidate models $M_1, \dots, M_{2^p}$ where $M_{\boldsymbol{\gamma}}$ corresponds to the $\boldsymbol{\gamma}$th subset of $\mathbf{X}_1, \dots, \mathbf{X}_p$. Each submodel is of the form

$$M_{\boldsymbol{\gamma}}: \mathbf{y} = \mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \boldsymbol{\varepsilon}\ , \tag{3.14}$$

where $\mathbf{X}_{\boldsymbol{\gamma}}$ is an $n \times q_{\boldsymbol{\gamma}}$ design matrix whose columns correspond to the $\boldsymbol{\gamma}$th subset of $\mathbf{X}_1, \dots, \mathbf{X}_p$, $\boldsymbol{\beta}_{\boldsymbol{\gamma}}$ is a $q_{\boldsymbol{\gamma}} \times 1$ vector of regression coefficients for the $\boldsymbol{\gamma}$th subset, $q_{\boldsymbol{\gamma}} \equiv \boldsymbol{\gamma}^T \mathbf{1}$ denotes the size of the $\boldsymbol{\gamma}$th subset, and $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2 \mathbf{I})$.

The linear regression model for each submodel of the form (3.14) is

$$\mathbf{y} | \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma} \sim N_n(\mathbf{X}_{\boldsymbol{\gamma}} \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2 \mathbf{I}). \tag{3.15}$$

Step 2: Prior distribution for $\boldsymbol{\beta}$, σ^2, $\boldsymbol{\gamma}$

In the process of the Bayesian variable selection, we use the hierarchical model prior for elastic net linear regression models in the form of submodel (3.14) which are the following.

$\mathbf{y} | \boldsymbol{\beta}_{\gamma}, \sigma^2, \boldsymbol{\gamma} \sim N_n(\mathbf{X}_{\boldsymbol{\gamma}} \boldsymbol{\beta}_{\gamma}, \sigma^2 \mathbf{I})$,

$\boldsymbol{\beta}_{\boldsymbol{\gamma}} | \sigma^2, \mathbf{D}_{\tau}, \boldsymbol{\gamma} \sim N_{q_{\gamma}}(\mathbf{0}, \sigma^2 \mathbf{D}_{\tau})$,

where $\mathbf{D}_{\tau}$ is a diagonal matrix with diagonal elements $\left(\tau_j^{-2} + \lambda_2\right)^{-1}$,

(3.16)

$\tau_j^2 \sim Exponential\left(\frac{\lambda_1^2}{2}\right)$, $j = 1,2,3, \dots, p$,

$\sigma^2 | \boldsymbol{\gamma} = \sigma^2 \sim$ inverse gamma $\left(\frac{\nu}{2}, \frac{\nu\xi}{2}\right)$,

Prior distribution of γ_j is $\gamma_j \sim \mathrm{Bernoulli}(w_j)$ and let $w_j = 0.5$ for all j .

The reasons for using the prior distribution in (3.16) are described as follows.

Prior distribution of β

In this thesis, we use the prior distribution of $\boldsymbol{\beta}$ that associated with elastic net estimate.

$$f\left(\boldsymbol{\beta}\middle|\sigma^2,\tau_1^2,\dots,\tau_p^2\right) \propto$$
$$\prod_{j=1}^{p}\frac{\sqrt{\lambda_2}}{\sqrt{2\pi\sigma^2}}exp\left[-\frac{\lambda_2}{2\sigma^2}\beta_j^2\right]\times\int_0^\infty\frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}}exp\left[-\frac{\beta_j^2}{2\sigma^2\tau_j^2}\right]\frac{\lambda_1^2}{2}exp\left[-\frac{\lambda_1^2\tau_j^2}{2}\right]d\tau_j^2.$$

Thus, $\beta_j \sim N\left(0,\sigma^2\left(\tau_j^{-2}+\lambda_2\right)^{-1}\right), j = 1,\dots,p.$ (3.17)

The prior in (3.17) was proposed by Kyung, Gill, Ghosh, and Casella (2010) and they gave the hierarchical model prior for $\boldsymbol{\beta}$ as

$\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^2,\mathbf{D}_\tau,\boldsymbol{\gamma} \sim N_{q_\gamma}(\mathbf{0},\sigma^2\mathbf{D}_\tau)$,

where $\mathbf{D}_\tau$ is a diagonal matrix with diagonal elements $\left(\tau_j^{-2}+\lambda_2\right)^{-1}$,

(3.18)

$\tau_j^2 \sim Exponential\left(\frac{\lambda_1^2}{2}\right), j = 1,2,3,\dots,p,$

The prior distribution of $\boldsymbol{\beta}$ in (3.17) differs from the Bayesian variable selection technique proposed by George and McCulloch (1997). The proof of (3.17) and (3.18) shows in Appendix D.

Prior distribution of σ^2

In this thesis, we assume σ^2 has inverse gamma prior distribution as suggested by Chipman, George and McCulloch (2001), George and McCulloch (1997), Hans (2011) and Park and Casella (2008). The inverse gamma prior for σ^2 would maintain conjugacy which gives the posterior distribution in the closed form.

$$\sigma^2 \sim \text{inverse gamma}\left(\frac{\nu}{2}, \frac{\nu\xi}{2}\right), \tag{3.19}$$

(which is equivalent to $\nu\xi/\sigma^2 \sim \chi^2_\nu$). The prior in (3.19) is the same as the prior (3.6) in page 40.

In this thesis, we choose $\nu = 3$ as suggested by Chipman et al (2001) and choose $\xi = S^2_{FULL}$ (the traditional unbiased estimate of σ^2 based on saturated model). The choice of $\xi = S^2_{FULL}$ was used by George and McCulloch (1997). But they suggested to choose ν to be the prior sample size associated with the estimate of ξ. This differs from the prior $\nu = 3$ used in this thesis.

Prior distribution for $\boldsymbol{\gamma}$

For the specification of the model space prior, most Bayesian variable selection implementations have used independence priors of the form

$$p(\boldsymbol{\gamma}) = \prod_{j=1}^{p} w_j^{\gamma_j} (1 - w_j)^{1-\gamma_j} \ . \tag{3.20}$$

Under this prior, each $\mathbf{X}_j$ enters the model independently of the other coefficients, with probability $p(\gamma_j = 1) = 1 - p(\gamma_j = 0) = w_j$. Smaller w_j can be used to down-weight $\mathbf{X}_j$ which are costly or of less interest (Chipman, George & McCulloch, 2001, pp.76-77).

In this thesis, we set $w_j = 1/2$ which yields the uniform prior

$$p(\boldsymbol{\gamma}) = \left(\frac{1}{2}\right)^p. \tag{3.21}$$

The uniform distribution over models, $w = 1/2$, was initially considered by many as the natural noninformative choice. However, in the context of variable selection, the uniform distribution over models induces a Binomial distribution on the model size $p_{\boldsymbol{\gamma}}$, with prior expectation that half of the variables will be included (Clyde & George, 2004b, p.85). The uniform prior in (3.21) was used in the papers of George and McCulloch (1997), Lykou and Ntzoufras (2012), and the others.

For high dimensional problems one cannot specify the prior probability of each $\boldsymbol{\gamma}$ directly, and practical implementations of Bayesian selection have usually made prior assumptions that the presence or absence of a variable is independent of the presence or absence of the other variables. As a special case of this, the uniform prior distribution over models is appealing in that posterior probabilities of models depend only on the likelihood (Clyde, 2001).

For each model of the form (3.14), the posterior probability of all 2^p candidate models can be derived in the following steps.

Step 3: Posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$

Step 3.1:

Let $g(\boldsymbol{\gamma}|\mathbf{y})$ be the posterior probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$.

Using the hierarchical prior models which described in Step 2, we find the posterior probability $g(\boldsymbol{\gamma}|\mathbf{y})$.

Step 3.2:

The posterior probability $g(\boldsymbol{\gamma}|\mathbf{y})$ and the prediction error (PE) for all 2^p candidate models are computed.

Step 3.3:

By considering $g(\boldsymbol{\gamma}|\mathbf{y})$ for all possible 2^p models, the models which have high posterior probability $g(\boldsymbol{\gamma}|\mathbf{y})$ are considered and reported.

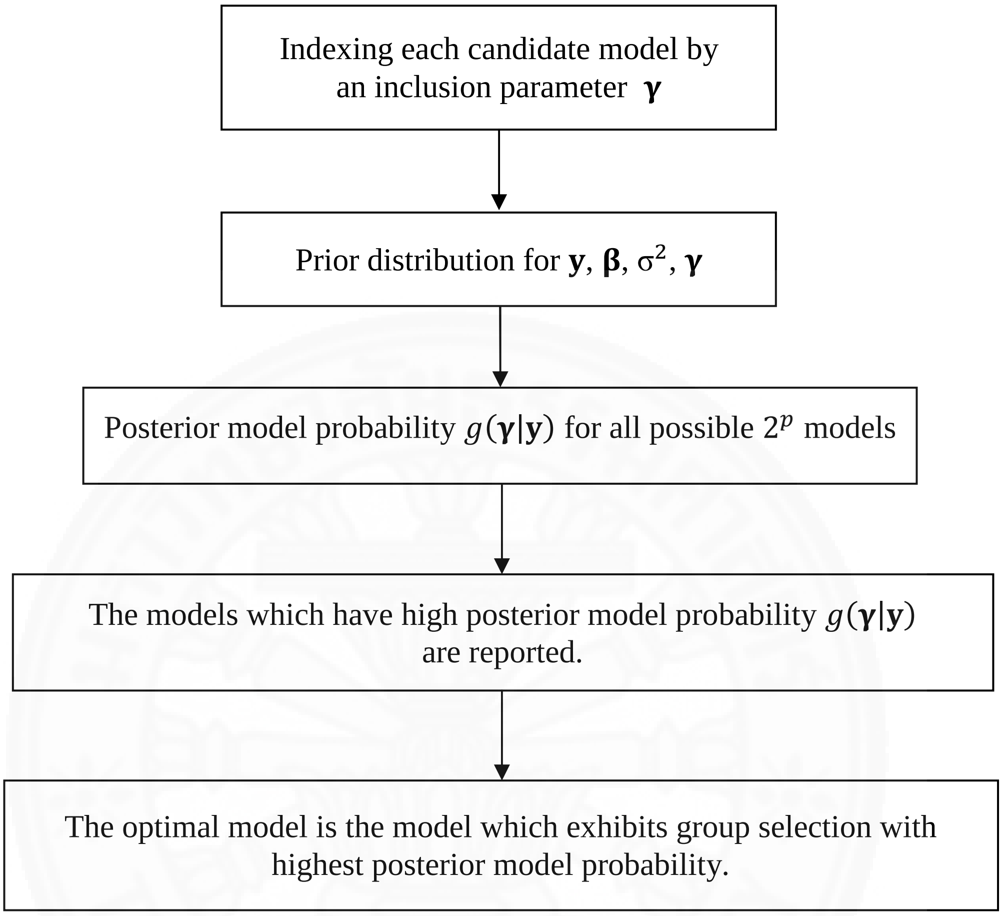
Step 4: Bayesian variable selection for elastic net linear regression model

We observe the behavior of models derived in step 3.3 and choose the optimal model which has the following properties:

(a) The model exhibits group selection – the whole groups of correlated predictors are included into the model or excluded from the model together.

(b) The model has highest posterior model probability $g(\boldsymbol{\gamma}|\mathbf{y})$.
If there are many models have equal posterior probability, the optimal model is the model which has minimum prediction error (PE).

Flowchart 3.2 Bayesian variable selection for elastic net linear regression model



3.3 Comparative study

In this thesis, the third and fourth objectives are the comparative study.

Objective 3: To perform the comparative study between the new method which is proposed in the first objective of this thesis and the 10-fold cross-validation method for estimating the value of the penalty parameter (λ_2) of the elastic net and adaptive elastic net linear regression models. (The reason for using 10-fold cross-validation method is that it was used by Zou and Hastie (2005) for choosing the penalty parameters of the elastic net regression.)

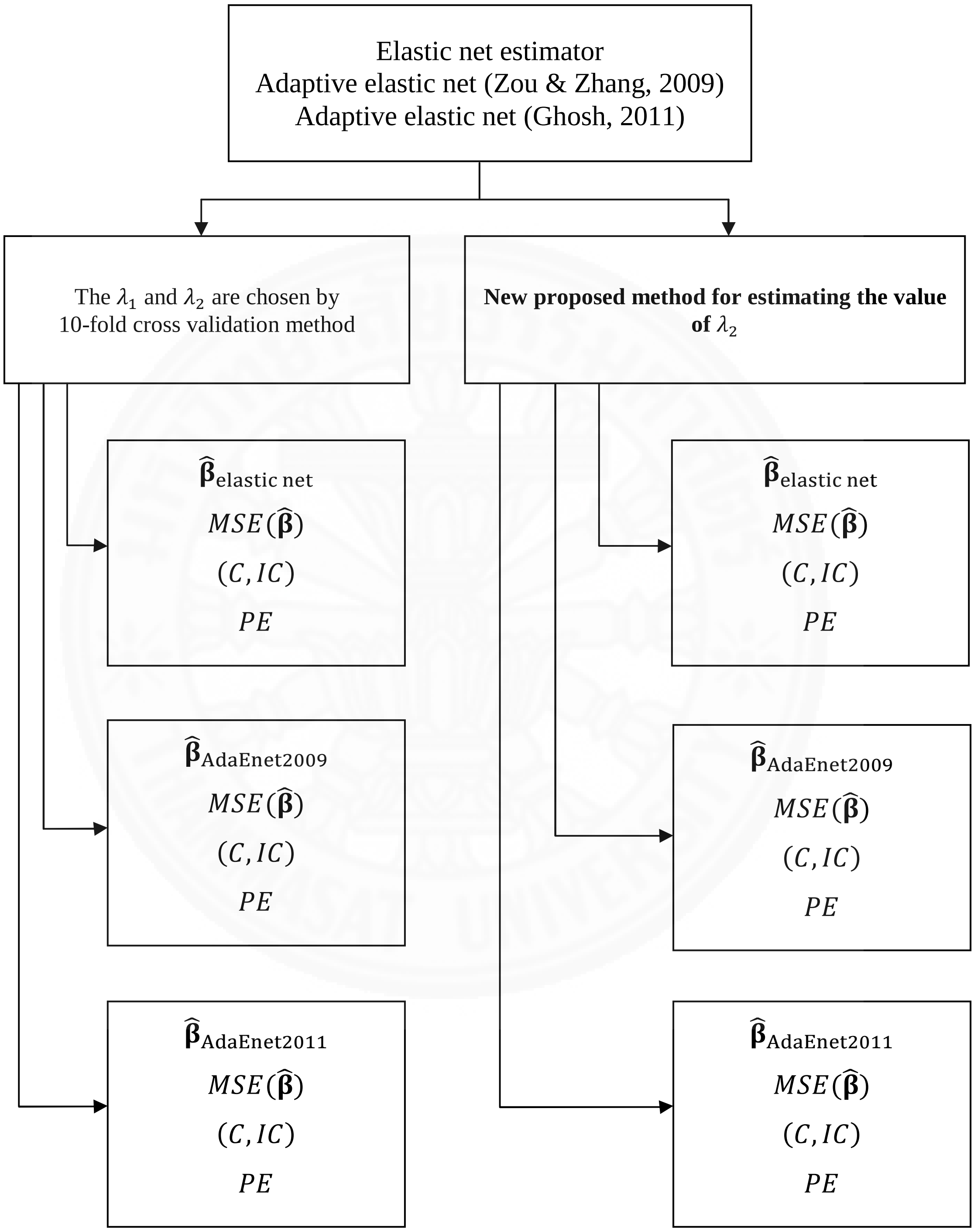
Objective 4: To perform the comparative study between the Bayesian variable selection for elastic net linear regression model which is proposed in the second objective of this thesis and elastic net linear regression model (Zou & Hastie, 2005) where the penalty parameter (λ_2) is estimated by the new method presented in the first objective of this thesis and the 10-fold cross-validation method.

Flowchart 3.3 and Flowchart 3.4 describe the comparative study of the objective 3 and objective 4, respectively.

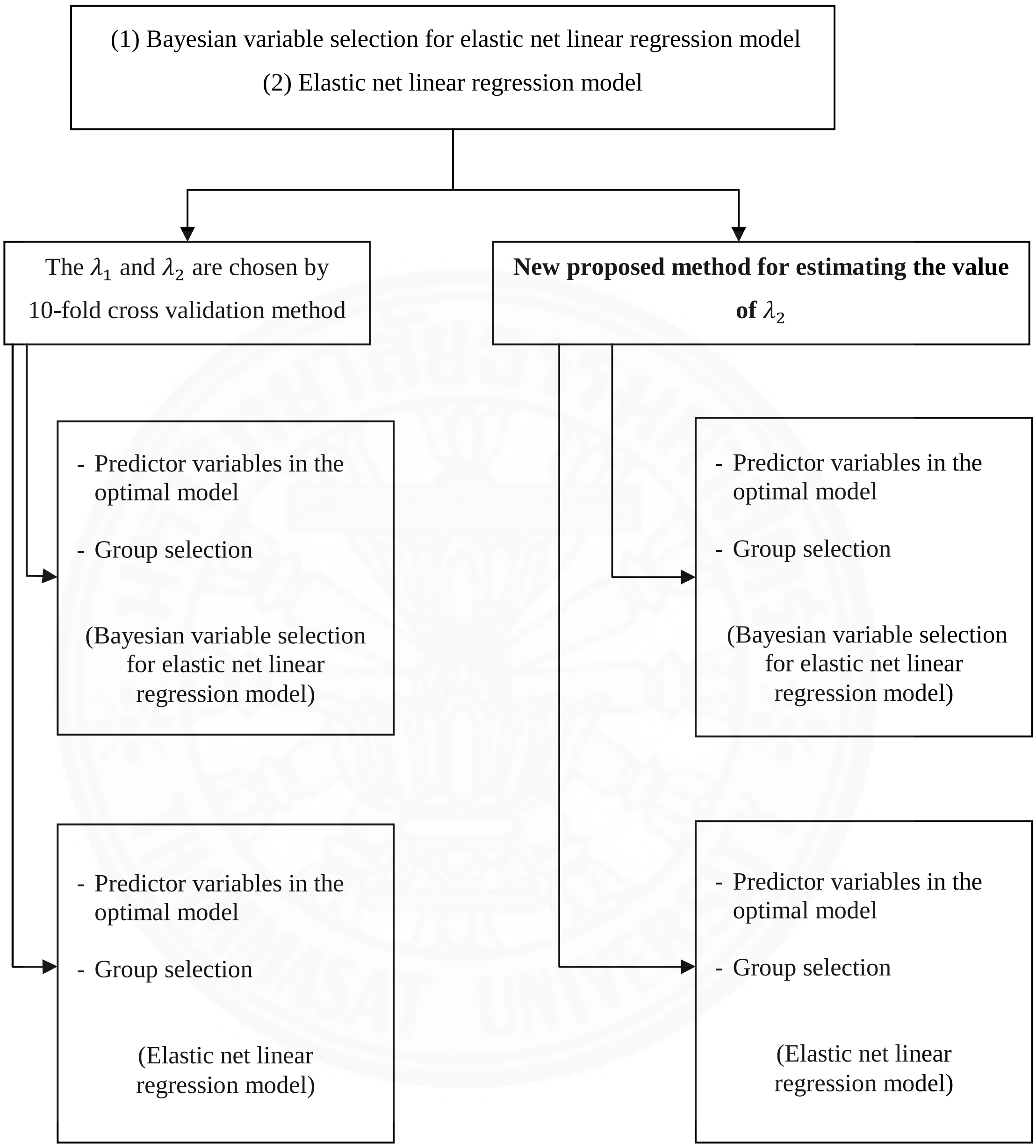
For the elastic net and adaptive elastic net estimators, we assume the response is centered and the predictors are standardized. Hence, the intercept is not included in the regression function.

We use the simulated datasets and two real datasets to access the performance of the proposed methods. The details of simulation data and real data are described in Section 3.3.1 and Section 3.3.2, respectively.

Flowchart 3.3 Comparative study of the objective 3



Flowchart 3.4 Comparative study of the objective 4



In this thesis, we use the simulated datasets and two real datasets to access the performance of the proposed methods. The details of simulation data and real data are described as follows.

3.3.1 Simulation data

The simulated datasets are generated by two simulation methods which proposed by Lykou and Ntzoufras (2012) and Zou and Zhang (2009). Each dataset is repeated 100 times.

Dataset 1: The datasets are simulated by the simulation method proposed by Lykou and Ntzoufras (2012) which adopted from the simulation design of Nott and Kohn (2005). This dataset has different correlated between predictor variables.

Simulation method for dataset 1

The dataset consists of 15 predictor variables of 50 observations each. The first 10 predictors follow independent standard normal distribution and the last 5 predictors are generated as follows,

$$(x_{11}, \dots, x_{15}) = (x_1, \dots, x_5) \times (0.3,0.5,0.7,0.9,1.1)^T \times (1,1,1,1,1) + \boldsymbol{E},$$

where $\boldsymbol{E}$ consists of 5 independent $N(0,1)$ random variables. Under this design, the last five predictors are highly correlated, while there are small to moderate correlations between x_j, $j = 1, \dots, 5$ and $x_{11}, \dots, x_{15}$. The response variable is generated as

$$y = 2x_1 - x_5 + 1.5x_7 + x_{11} + 0.5x_{13} + \varepsilon,$$

where $\varepsilon \sim N(0,2.5^2)$. This set of simulated data comprises of predictor variables that are correlated with each other. The simulation method is repeated 100 times.

Flowchart 3.5 Simulation method for dataset 1

$p = 15, n = 50$

↓

$x_j \sim N(0,1)\, ,\, j = 1,2,\ldots,10$

↓

$(x_{11}, \ldots, x_{15}) = (x_1, \ldots, x_5) \times (0.3,0.5,0.7,0.9,1.1)^T \times (1,1,1,1,1) + \boldsymbol{E}$,

where $\boldsymbol{E}$ consists of 5 independent $N(0,1)$ random variables

↓

$\varepsilon \sim N(0,2.5^2)$

↓

$y = 2x_1 - x_5 + 1.5x_7 + x_{11} + 0.5x_{13} + \varepsilon$

Dataset 2: The datasets are simulated by the simulation method proposed by Zou and Zhang (2009). This simulation method sets the number of parameters (p) depend on the sample size (n).

Simulation method for dataset 2

Let $p = p_n = \left[4n^{1/2}\right] - 5$ for n = 100, 200, 400.

The data is generated from the linear regression model

$$\mathbf{y} = \mathbf{X}^{\mathrm{T}}\boldsymbol{\beta} + \boldsymbol{\varepsilon}\ ,$$

where $\mathbf{y}$ is an $n \times 1$ vector of response variable.

$\boldsymbol{\beta}$ is an $p \times 1$ vector of parameter of regression coefficients.

$\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of random errors, $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$, σ = 6.

$\mathbf{X} = \left[\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p\right]^T$; $\mathbf{X}_j$ is an $n \times 1$ vector of the jth predictor variables.

$\mathbf{X}$ follows a p-dim multivariate normal distribution with zero mean and covariance $\boldsymbol{\Sigma}$, $\mathbf{X} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$, where the covariance matrix $\boldsymbol{\Sigma}$ has the entry

$$\boldsymbol{\Sigma}_{j,k} = corr(j,k) = \rho^{|j-k|}\ , 1 \le k\ , j \le p\ .$$

In this thesis, we set ρ = 0.5 and ρ = 0.75.

Let $\mathbf{1}_q$ denotes a $q \times 1$ vector of 1's, and $\mathbf{0}_{p-3q}$ denotes a $(p - 3q) \times 1$ vector of 0's.

Let the true coefficients are $\boldsymbol{\beta} = \left(3 \cdot \mathbf{1}_q, 3 \cdot \mathbf{1}_q, 3 \cdot \mathbf{1}_q, \mathbf{0}_{p-3q}\right)^T$ and $q = [p_n/9]$.

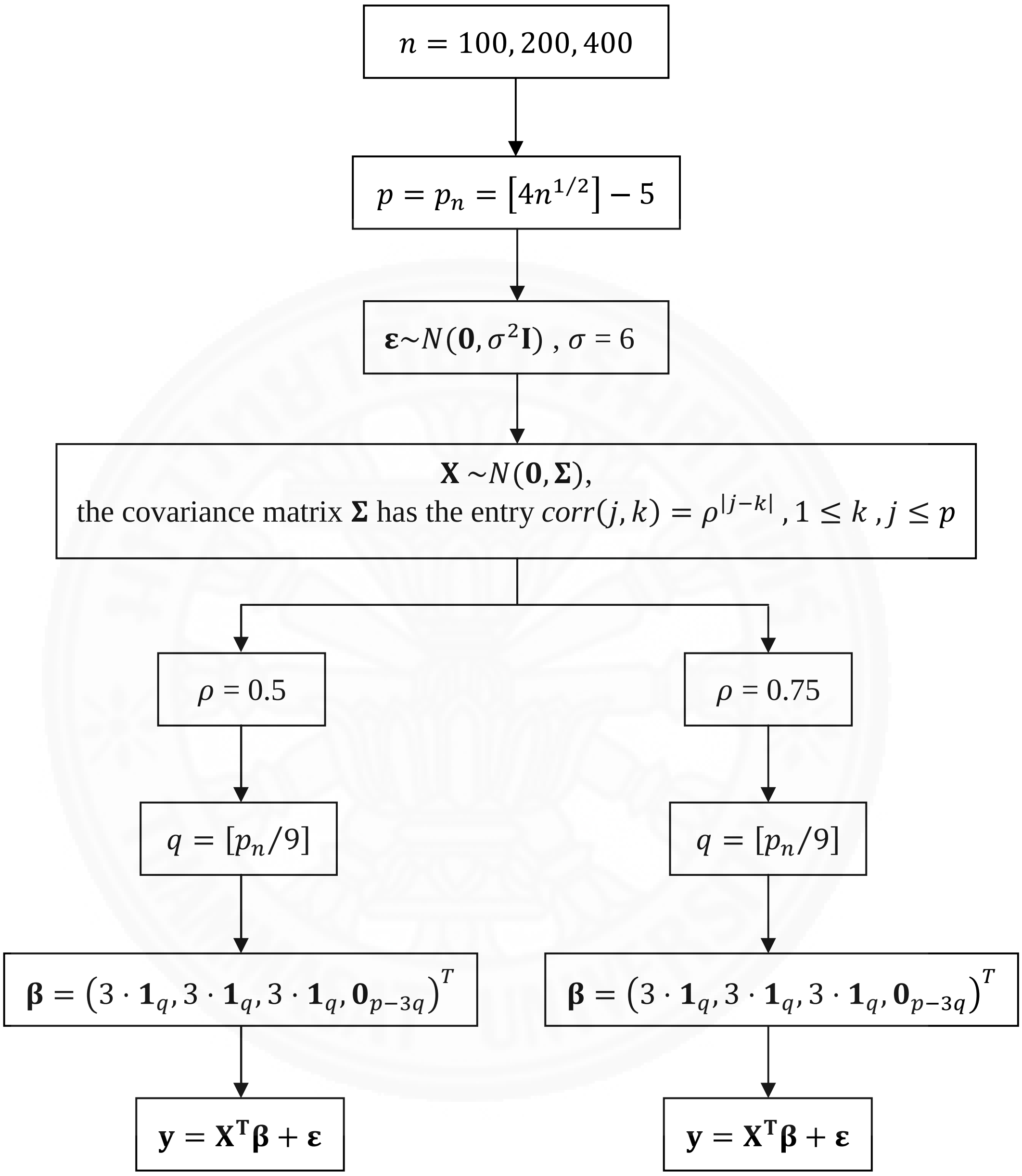
We are interested in the sparse modeling problem where the true model has a sparse representation (i.e., some components of $\boldsymbol{\beta}$ are exactly zero).

Let $\mathcal{A} = \{ j : \beta_j \neq 0, j = 1, 2, \dots, p \}$. The size of $\mathcal{A}$ is the intrinsic dimension of the underlying model. We set $|\mathcal{A}| = 3q$ for this simulation method.

For dataset 2, there are six cases for combination of n = 100, 200, 400 and ρ = 0.5, 0.75.

The simulation method is repeated 100 times.

Flowchart 3.6 Simulation method for dataset 2



3.3.2 Real data

In this thesis, we use two real datasets to illustrate the efficiency of the proposed methods. The two datasets are the diabetes dataset and prostate cancer data used in elastic net literature and related methods.

Dataset 3: Diabetes dataset ($p = 10, n = 442$)

The diabetes data is a data from Efron, Hastie, Johnstone, and Tibshirani, (2004). The response variable is a quantitative measure of disease progression one year after baseline for 442 diabetes patients. The dataset contains 10 baseline predictor variables: age, sex, body mass index (bmi), average blood pressure (bp), and six blood serum measurements (tc, ldl, hdl, tch, ltg, glu). The statisticians were asked to construct a model that predicted response y from predictors x_1, x_2,…, x_{10}. Two hopes were evident here, that the model would produce accurate baseline predictions of response for future patients and that the form of the model would suggest which predictors were important factors in disease progression (Efron, Hastie, Johnstone, & Tibshirani, 2004).

The diabetes dataset was analyzed by Efron, Hastie, Johnstone, and Tibshirani (2004); Hans (2009, 2010); Li and Lin (2010); Lykou and Ntzoufras (2012); Park and Casella (2008).

Dataset 4: Prostate cancer data ($p = 8, n = 97$)

The prostate cancer data is a data from a prostate cancer study of Stamey, Kabalin, Mcneal, Johnstone, Freiha, Redwine, and Yang (1989). These data come from a study that examined the correlation between the level of prostate specific antigen and a number of clinical measures in 97 men who were about to receive a radical prostatectomy.

The response variable is the logarithm of prostate specific antigen (lpsa). The predictor variables are eight clinical measures:

- the logarithm of cancer volume (lcavol),
- the logarithm of prostate weight (lweight),
- age,
- the logarithm of the amount of benign prostatic hyperplasia (lbph),
- seminal vesicle invasion (svi),
- the logarithm of capsular penetration (lcp),
- the Gleason score (gleason), and
- the percentage Gleason score 4 or 5 (pgg45).

The prostate cancer data was analyzed by Bornn, Gottardo, and Doucet (2010); Ghosh (2011); Hans (2011); Kyung, Gill, Ghosh, and Casella (2010); Li and Lin (2010); Tibshirani (1996); Zou and Hastie (2005).

Chapter 4

Results

We propose two new ideas in this thesis. The result is divided into two parts:

Part I: Method for estimating the value of the penalty parameter λ_2 based on Bayes factor

4.1 Posterior probability

4.1.1 Posterior model probability of γ given y when $\lambda_1 = 0$

4.1.2 Posterior model probability of γ given y for elastic net linear regression model

4.2 Bayes factor used in the method for estimating the value of the penalty parameter λ_2

4.2.1 Bayes factor for simple linear regression model

4.2.2 Bayes factor for multiple linear regression model

4.2.3 Bayes factor for elastic net linear regression model

4.3 Method for estimating the value of the penalty parameter λ_2 based on Bayes factor

4.4 Experimental result

4.4.1 Simulation study I

4.4.2 Simulation study II

4.4.3 Result of adaptive elastic net estimates for simulation study II

4.5 Real data example

4.5.1 Real data example for estimating the value of the penalty parameter λ_2 based on Bayes factor

4.5.2 Real data example for adaptive elastic net estimates

Part II: Bayesian variable selection for elastic net linear regression model

4.6 Bayesian variable selection for elastic net linear regression model where the penalty parameters λ_1 and λ_2 are estimated by the 10-fold cross-validation method

4.7 Bayesian variable selection for elastic net linear regression model where the penalty parameter λ_2 is based on Bayes factor

The details are as follows.

Part I: Method for estimating the value of the penalty parameter $\boldsymbol{\lambda_2}$ based on Bayes factor

Using the process of the method for estimating the value of the penalty parameter λ_2 described in Chapter 3.1 and Flowchart 3.1, the results are as follows.

4.1 Posterior probability

4.1.1 Posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$ when $\boldsymbol{\lambda_1 = 0}$

In Chapter 3, we use the hierarchical prior models prior (3.4) as follows.

$\mathbf{y}|\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma} \sim N_n(\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2\mathbf{I}),$

$\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^2, \boldsymbol{\gamma} \sim N_{q_{\boldsymbol{\gamma}}}(\mathbf{0}, \frac{\sigma^2}{\lambda_2}\mathbf{I}),$

$\sigma^2|\boldsymbol{\gamma} = \sigma^2 \sim \text{inverse gamma}\left(\frac{\nu}{2}, \frac{\nu\xi}{2}\right),$

$p(\boldsymbol{\gamma}) = \left(\frac{1}{2}\right)^p.$

Let $f(\boldsymbol{\gamma}|\mathbf{y})$ be the posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$.
Using the prior in (3.4), we obtain the posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$.

$$f(\boldsymbol{\gamma}|\mathbf{y}) = f(\mathbf{y}|\boldsymbol{\gamma})f(\boldsymbol{\gamma}) \tag{4.1}$$

where
$$f(\mathbf{y}|\boldsymbol{\gamma}) \propto \left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right|^{-\frac{1}{2}}(\lambda_2)^{\frac{q_{\boldsymbol{\gamma}}}{2}}\left(\nu\xi + S_{\boldsymbol{\gamma}}^2\right)^{-\frac{(n+\nu)}{2}}, \tag{4.2}$$

and
$$S_{\boldsymbol{\gamma}}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_{\boldsymbol{\gamma}}\left[\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right]^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}. \tag{4.3}$$

The proof of (4.2) and (4.3) are described in Appendix F.

4.1.2 Posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$ for elastic net linear regression model

In Chapter 3, we use the hierarchical model prior for elastic net linear regression models as follows.

$$\mathbf{y}|\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma} \sim N_n(\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2\mathbf{I}),$$

$$\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^2, \mathbf{D}_\tau, \boldsymbol{\gamma} \sim N_{q_{\boldsymbol{\gamma}}}(\mathbf{0}, \sigma^2\mathbf{D}_\tau) \ ,$$

where $\mathbf{D}_\tau$ is a diagonal matrix with diagonal elements $\left(\tau_j^{-2} + \lambda_2\right)^{-1}$,

$$\tau_j^2 \sim Exponential\left(\frac{\lambda_1^2}{2}\right), j = 1,2,3, \dots, p,$$

$$\sigma^2|\boldsymbol{\gamma} = \sigma^2 \sim \text{inverse gamma}\left(\frac{\nu}{2}, \frac{\nu\xi}{2}\right),$$

$$p(\boldsymbol{\gamma}) = \left(\frac{1}{2}\right)^p.$$

Let $g(\boldsymbol{\gamma}|\mathbf{y})$ be the posterior probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$.

We obtain the posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$.

$$g(\boldsymbol{\gamma}|\mathbf{y}) \propto |\mathbf{D}_\tau|^{-\frac{1}{2}}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_\tau^{-1}\right|^{-\frac{1}{2}}\left(\nu\xi + S_{\boldsymbol{\gamma}\mathbf{D}}^2\right)^{-\frac{(n+\nu)}{2}}, \tag{4.4}$$

where $$S_{\boldsymbol{\gamma}\mathbf{D}}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_{\boldsymbol{\gamma}}\left[\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_\tau^{-1}\right]^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}. \tag{4.5}$$

The proof of (4.4) and (4.5) are described in Appendix G.

4.2 Bayes factor used in the method for estimating the value of the penalty parameter $\boldsymbol{\lambda_2}$

For the Bayes factor used in this thesis, we define the hypotheses associated with the Bayes factor as

Hypotheses:

H_0: Reduced model (M_R)

versus

H_1: Full model (M_F),

where

M_R (Reduced model) is the linear regression model with the predictors $\mathbf{X}_{\boldsymbol{\gamma}}$,

M_F (Full model) is the linear regression model with the predictors $\mathbf{X}_{\boldsymbol{\gamma}}$ of the reduced model and additional predictor $\mathbf{X}_j$.

Suppose $\mathbf{X}_{\boldsymbol{\gamma}}$ be the predictor variables of the reduced model, so

M_R: $\mathbf{y} = \mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \boldsymbol{\varepsilon}$,

M_F: $\mathbf{y} = \mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \mathbf{X}_j\boldsymbol{\beta}_j + \boldsymbol{\varepsilon}$.

According to the process described in Section 3.1, the results of Bayes factor for simple and multiple linear regression models are as follows.

4.2.1 Bayes factor for simple linear regression model

Hypotheses:

H_0: M_R is the null (or constant) model where $\mathbf{y}|\boldsymbol{\beta}, \sigma^2, M_R \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$.

versus

H_1: M_F is a simple linear regression model $\mathbf{y} = \mathbf{X}_j\boldsymbol{\beta}_j + \boldsymbol{\varepsilon}$

Let BF_{10} be the Bayes factor for comparing the evidence of model M_F versus model M_R.

The Bayes factor for simple linear regression model is

$$\text{BF}_{10} = \frac{f(\mathbf{y}|M_F)}{f(\mathbf{y}|M_R)} \quad , \tag{4.6}$$

where $f(\mathbf{y}|M_R)$ be the posterior model probability of the reduced model M_R given $\mathbf{y}$, and $f(\mathbf{y}|M_F)$ be the posterior model probability of the full model M_F given $\mathbf{y}$.

The posterior probabilities $f(\mathbf{y}|M_R)$ and $f(\mathbf{y}|M_F)$ are derived from equation (4.2) and (4.3) as follows. From equation (4.2) and (4.3), the posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$ is

$$f(\mathbf{y}|\boldsymbol{\gamma}) \equiv \left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_\gamma}\right|^{-\frac{1}{2}}(\lambda_2)^{\frac{q_\gamma}{2}}\left(\nu\xi + S_{\boldsymbol{\gamma}}^2\right)^{-\frac{(n+\nu)}{2}} ,$$

and
$$S_{\boldsymbol{\gamma}}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_{\boldsymbol{\gamma}}\left[\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_\gamma}\right]^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}.$$

Thus, the posterior probabilities $f(\mathbf{y}|M_R)$ and $f(\mathbf{y}|M_F)$ are

$$f(\mathbf{y}|M_R) = |\lambda_2|^{-\frac{1}{2}}(\nu\xi + \mathbf{y}^T\mathbf{y})^{-\frac{(n+\nu)}{2}}, \tag{4.7}$$

$$f(\mathbf{y}|M_F) = (\lambda_2)^{\frac{1}{2}}\left|\sum_{i=1}^n x_{ij}^2 + \lambda_2\right|^{-\frac{1}{2}}\left\{\nu\xi + \left[\sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n x_{ij}y_i\right)^2\left(\sum_{i=1}^n x_{ij}^2 + \lambda_2\right)^{-1}\right]\right\}^{-\frac{(n+\nu)}{2}} \tag{4.8}$$

To derive the Bayes factor associated with Pearson correlation, the response and the predictor variables are transformed by the correlation transformation. Thus, the simplified version of (4.8) is

$$f(\mathbf{y}|M_F) = \frac{(\lambda_2)^{\frac{1}{2}}}{\left(\sum_{i=1}^{n} x_{ij}^2+\lambda_2\right)^{\frac{1}{2}}}\left\{\nu\xi + \left(\sum_{i=1}^{n} x_{ij}^2\right)\left(\sum_{i=1}^{n} y_i^2\right)\left[\frac{1}{\sum_{i=1}^{n} x_{ij}^2} - \left(r_{x_j y}\right)^2\left(\sum_{i=1}^{n} x_{ij}^2 + \lambda_2\right)^{-1}\right]\right\}^{-\frac{(n+\nu)}{2}}, \tag{4.9}$$

where $r_{x_j y} = \frac{\sum_{i=1}^{n} x_{ij} y_i}{\sqrt{\sum_{i=1}^{n} x_{ij}^2 \sum_{i=1}^{n} y_i^2}}$ is the sample estimate of the Pearson correlation coefficient between y and the candidate predictor X_j.

Thus, the Bayes factor for simple linear regression model is

$$\mathrm{BF}_{10} = \frac{f(\mathbf{y}|M_F)}{f(\mathbf{y}|M_R)}$$

$$\mathrm{BF}_{10} = \frac{(\lambda_2)\left(\sum_{i=1}^{n} x_{ij}^2+\lambda_2\right)^{-\frac{1}{2}}\left\{\nu\xi+\left(\sum_{i=1}^{n} x_{ij}^2\right)\left(\sum_{i=1}^{n} y_i^2\right)\left[\frac{1}{\sum_{i=1}^{n} x_{ij}^2}-\left(r_{x_j y}\right)^2\left(\sum_{i=1}^{n} x_{ij}^2+\lambda_2\right)^{-1}\right]\right\}^{-\frac{(n+\nu)}{2}}}{\left(\nu\xi+\sum_{i=1}^{n} y_i^2\right)^{-\frac{(n+\nu)}{2}}} \tag{4.10}$$

Since $\sum_{i=1}^{n} x_{ij}^2 = 1$ and $\sum_{i=1}^{n} y_i^2 = 1$, the simplify version of (4.10) is

$$\mathrm{BF}_{10} = \frac{(\lambda_2)(1+\lambda_2)^{-\frac{1}{2}}\left\{\nu\xi+\left[1-\left(r_{x_j y}\right)^2(1+\lambda_2)^{-1}\right]\right\}^{-\frac{(n+\nu)}{2}}}{(\nu\xi+1)^{-\frac{(n+\nu)}{2}}} \tag{4.11}$$

The Bayes factor for simple regression model in (4.10) and (4.11) are expressed in terms of the penalty parameter λ_2 and $r_{x_j y}$, i.e., the sample estimate of the Pearson correlation coefficient between y and X_j.

4.2.2 Bayes factor for multiple linear regression model

Hypotheses:

H_0: Reduced model M_R: $\mathbf{y} = \mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \boldsymbol{\varepsilon}$,

versus

H_1: Full model M_F: $\mathbf{y} = \mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \mathbf{X}_j\boldsymbol{\beta}_j + \boldsymbol{\varepsilon}$,

where $\mathbf{X}_{\boldsymbol{\gamma}}$ be the predictor variables of the reduced model.

Let BF_{10} be the Bayes factor for comparing the evidence of model M_F versus model M_R.

The Bayes factor for multiple linear regression model is

$$\mathrm{BF}_{10(Multiple\ model)} = \frac{f(\mathrm{y}|M_F)}{f(\mathrm{y}|M_R)} \quad , \tag{4.12}$$

where $f(\mathbf{y}|M_R)$ be the posterior model probability of the reduced model M_R given $\mathbf{y}$, and $f(\mathbf{y}|M_F)$ be the posterior model probability of the full model M_F given $\mathbf{y}$.

The posterior probabilities $f(\mathbf{y}|M_R)$ and $f(\mathbf{y}|M_F)$ are derived from equation (4.2) and (4.3) as follows. From equation (4.2) and (4.3), we define

$$f(\mathbf{y}|\boldsymbol{\gamma}) \equiv \left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right|^{-\frac{1}{2}} (\lambda_2)^{\frac{q_{\boldsymbol{\gamma}}}{2}} \left(\nu\xi + S_{\boldsymbol{\gamma}}^2\right)^{-\frac{(n+\nu)}{2}} ,$$

and

$$S_{\boldsymbol{\gamma}}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_{\boldsymbol{\gamma}}\left[\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right]^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}.$$

Thus, the posterior probabilities $f(\mathbf{y}|M_R)$ and $f(\mathbf{y}|M_F)$ are

$$f(\mathbf{y}|M_R) = (\lambda_2)^{\frac{q_{M_R}}{2}} \left|\mathbf{X}_{M_R}^T\mathbf{X}_{M_R} + \lambda_2\mathbf{I}_{q_{M_R}}\right|^{-\frac{1}{2}} \left(\nu\xi + S_{M_R}^2\right)^{-\frac{(n+\nu)}{2}}, \tag{4.13}$$

where $$S_{M_R}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_{M_R}\left[\mathbf{X}_{M_R}^T\mathbf{X}_{M_R} + \lambda_2\mathbf{I}_{q_{M_R}}\right]^{-1}\mathbf{X}_{M_R}^T\mathbf{y}. \tag{4.14}$$

$$f(\mathbf{y}|M_F) = (\lambda_2)^{\frac{q_{M_F}}{2}}\left|\mathbf{X}_{M_F}^T\mathbf{X}_{M_F} + \lambda_2\mathbf{I}_{q_{M_F}}\right|^{-\frac{1}{2}}\left(\nu\xi + S_{M_F}^2\right)^{-\frac{(n+\nu)}{2}}, \tag{4.15}$$

where $$S_{M_F}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_{M_F}\left[\mathbf{X}_{M_F}^T\mathbf{X}_{M_F} + \lambda_2\mathbf{I}_{q_{M_F}}\right]^{-1}\mathbf{X}_{M_F}^T\mathbf{y}. \tag{4.16}$$

4.2.3 Bayes factor for elastic net linear regression model

We define the hypotheses associated with the Bayes factor for elastic net linear regression model ($\mathrm{BF}_{elastic\ net}$) as

Hypotheses:

H_0: Reduced model $M_{R(elastic\ net)}$: $\mathbf{y} = \mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \boldsymbol{\varepsilon}$,

versus

H_1: Full model $M_{F(elastic\ net)}$: $\mathbf{y} = \mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \mathbf{X}_j\boldsymbol{\beta}_j + \boldsymbol{\varepsilon}$

Let $\mathrm{BF}_{10(elastic\ net)}$ be the Bayes factor for comparing the evidence of model $M_{F(elastic\ net)}$ versus model $M_{R(elastic\ net)}$.

The Bayes factor for elastic net linear regression model is

$$\mathrm{BF}_{10(elastic\ net)} = \frac{g\left(M_{F(elastic\ net)}|\mathbf{y}\right)}{g\left(M_{R(elastic\ net)}|\mathbf{y}\right)} \tag{4.17}$$

where $g\left(M_{R(elastic\ net)}|\mathbf{y}\right)$ be the posterior model probability of the reduced model M_R given $\mathbf{y}$, and $g\left(M_{F(elastic\ net)}|\mathbf{y}\right)$ be the posterior model probability of the full model M_F given $\mathbf{y}$.

The posterior model probabilities $g\left(M_{R(elastic\ net)}\middle|\mathbf{y}\right)$ and $g\left(M_{F(elastic\ net)}\middle|\mathbf{y}\right)$ are derived from equation (4.4) and (4.5) as follows.

From equation (4.4) and (4.5), we define

$$g(\boldsymbol{\gamma}|\mathbf{y}) \equiv |\mathbf{D}_{\tau}|^{-\frac{1}{2}}\left|\mathbf{X}_{\boldsymbol{\gamma}}^{T}\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right|^{-\frac{1}{2}}\left(\nu\xi + S_{\boldsymbol{\gamma}\mathbf{D}}^{2}\right)^{-\frac{(n+\nu)}{2}},$$

where $$S_{\boldsymbol{\gamma}\mathbf{D}}^{2} = \mathbf{y}^{T}\mathbf{y} - \mathbf{y}^{T}\mathbf{X}_{\boldsymbol{\gamma}}\left[\mathbf{X}_{\boldsymbol{\gamma}}^{T}\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right]^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^{T}\mathbf{y}.$$

Thus, the posterior model probabilities $g\left(M_{R(elastic\ net)}\middle|\mathbf{y}\right)$ and $g\left(M_{F(elastic\ net)}\middle|\mathbf{y}\right)$ are

$$g\left(M_{R(elastic\ net)}\middle|\mathbf{y}\right) \equiv |\mathbf{D}_{\tau}|^{-\frac{1}{2}}\left|\mathbf{X}_{M_R}^{T}\mathbf{X}_{M_R} + \mathbf{D}_{\tau}^{-1}\right|^{-\frac{1}{2}}\left(\nu\xi + S_{M_R\mathbf{D}}^{2}\right)^{-\frac{(n+\nu)}{2}} \tag{4.18}$$

where $$S_{M_R\mathbf{D}}^{2} = \mathbf{y}^{T}\mathbf{y} - \mathbf{y}^{T}\mathbf{X}_{M_R}\left[\mathbf{X}_{M_R}^{T}\mathbf{X}_{M_R} + \mathbf{D}_{\tau}^{-1}\right]^{-1}\mathbf{X}_{M_R}^{T}\mathbf{y}. \tag{4.19}$$

$$g\left(M_{F(elastic\ net)}\middle|\mathbf{y}\right) \equiv |\mathbf{D}_{\tau}|^{-\frac{1}{2}}\left|\mathbf{X}_{M_F}^{T}\mathbf{X}_{M_F} + \mathbf{D}_{\tau}^{-1}\right|^{-\frac{1}{2}}\left(\nu\xi + S_{M_F\mathbf{D}}^{2}\right)^{-\frac{(n+\nu)}{2}} \tag{4.20}$$

where $$S_{M_F\mathbf{D}}^{2} = \mathbf{y}^{T}\mathbf{y} - \mathbf{y}^{T}\mathbf{X}_{M_F}\left[\mathbf{X}_{M_F}^{T}\mathbf{X}_{M_F} + \mathbf{D}_{\tau}^{-1}\right]^{-1}\mathbf{X}_{M_F}^{T}\mathbf{y}. \tag{4.21}$$

4.3 Method for estimating the value of the penalty parameter λ_2 based on Bayes factor (λ_2BF)

According to the Flowchart 3.1 described in Chapter 3, we summarize the method for estimating the value of the penalty parameter λ_2 based on Bayes factor (λ_2BF) as follows.

Step 1: For all pairs of hypotheses, we compute the Bayes factor for multiple linear regression model, $\mathrm{BF}_{10(Multiple\ model)}$, described in Section 4.2.2.

From (4.12), the Bayes factor for multiple linear regression model is

$$\mathrm{BF}_{10(Multiple\ model)} = \frac{f(\mathbf{y}|M_F)}{f(\mathbf{y}|M_R)}$$

To find the posterior probabilities $f(\mathbf{y}|M_F)$ and $f(\mathbf{y}|M_R)$, the shrinkage parameter λ_2 is replaced by θ_{M_F} and θ_{M_R}, respectively.

Thus, the posterior probabilities $f(\mathbf{y}|M_R)$ and $f(\mathbf{y}|M_F)$ are

$$f(\mathbf{y}|M_R) = \left(\theta_{M_R}\right)^{\frac{q_{M_R}}{2}} \left|\mathbf{X}_{M_R}^T\mathbf{X}_{M_R} + \theta_{M_R}\mathbf{I}_{q_{M_R}}\right|^{-\frac{1}{2}} \left(\nu\xi + S_{M_R}^2\right)^{-\frac{(n+\nu)}{2}}, \quad (4.22)$$

where $$S_{M_R}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_{M_R}\left[\mathbf{X}_{M_R}^T\mathbf{X}_{M_R} + \theta_{M_R}\mathbf{I}_{q_{M_R}}\right]^{-1}\mathbf{X}_{M_R}^T\mathbf{y}. \quad (4.23)$$

$$f(\mathbf{y}|M_F) = \left(\theta_{M_F}\right)^{\frac{q_{M_F}}{2}} \left|\mathbf{X}_{M_F}^T\mathbf{X}_{M_F} + \theta_{M_F}\mathbf{I}_{q_{M_F}}\right|^{-\frac{1}{2}} \left(\nu\xi + S_{M_F}^2\right)^{-\frac{(n+\nu)}{2}}, \quad (4.24)$$

where $$S_{M_F}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_{M_F}\left[\mathbf{X}_{M_F}^T\mathbf{X}_{M_F} + \theta_{M_F}\mathbf{I}_{q_{M_F}}\right]^{-1}\mathbf{X}_{M_F}^T\mathbf{y}. \quad (4.25)$$

The shrinkage parameter $\lambda_2 \geq 0$. Hence, $\theta_{M_F} \geq 0$ and $\theta_{M_R} \geq 0$.

In this thesis, we choose the value of θ_M by the method proposed by Hoerl, Kennard , and Baldwin (1975) as follows.

The value of θ_M is

$$\theta_M = \frac{q_M S_M^2}{[\hat{\boldsymbol{\beta}}_{LS(\mathrm{M})}]^T[\hat{\boldsymbol{\beta}}_{LS(\mathrm{M})}]}, \tag{4.26}$$

where

(1) q_M is the number of parameters in the model M (not counting the intercept term).

(2) S_M^2 is the residual mean square in the analysis of variance table obtained from the standard least squared fit of the model M.

(3) $\hat{\boldsymbol{\beta}}_{LS(M)}$ is the least squared estimator of the parameter in model M.

(Hoerl, Kennard , & Baldwin, 1975, cited by Draper & Smith, 1998, pp.390-391.)

Step 2: At the levels of Bayes factor ($\mathrm{BF}_{10} = 1$, $\mathrm{BF}_{10} = 3$), we can identify the set of predictors in the following.

Step 2.1: The value of $\mathrm{BF}_{10} = 1$ is used for identification the set of significance predictors ($\mathrm{BF}_{10} > 1$) and non- significance predictors ($\mathrm{BF}_{10} < 1$).

- Set of significance predictors is $\{X_j |\, X_j \in\ H_1 \text{ and } \mathrm{BF}_{10} > 1\}$.
- Set of non-significance predictors is $\{X_j |\, X_j \in\ H_1 \text{ and } \mathrm{BF}_{10} < 1\}$.

Step 2.2: The value of $\mathrm{BF}_{10} = 3$ is used for identification the set of *practical significance* predictors. We give the name of "the set of *practical significance* predictors" to be "the set of *important* predictors".

- Set of important predictors is $\{X_j |\, X_j \in\ H_1 \text{ and } \mathrm{BF}_{10} > 3\}$.
- Set of non-important predictors is $\{X_j |\, X_j \in\ H_1 \text{ and } 1 < \mathrm{BF}_{10} < 3\}$.

Step 3: Specification of the penalty parameter λ_2 based on Bayes factor.

Step 3.1: If the Bayes factor $\mathrm{BF}_{10(Multiple\,model)} > 1$, the model M_F is a significance model and θ_{M_F} associated with this model is considered to be the choice of the value of λ_2. There are many models M_F in the set $\mathrm{BF}_{10(Multiple\,model)} > 1$, so there are many θ_{M_F} associated with these models. The value θ_{M_F} of the model M_F which has highest posterior model probability $f(\mathbf{y}|M_F)$ (the posterior model probability using the prior in (3.4)) is selected to be the penalty parameter λ_2 associated with Bayes factor and this λ_2 is called λ_2BF.

Step 3.2: For checking the validity of λ_2BF to fit the elastic net model, the Bayes factor $\mathrm{BF}_{elastic\,net}$ is computed. The appropriate value of λ_2BF should give the Bayes factor $\mathrm{BF}_{elastic\,net} > 1$.

Hence, the appropriate value of λ_2BF is the value θ_{M_F} of the submodel which has highest posterior model probability $f(\mathbf{y}|M_F)$ (the posterior model probability using the prior in (3.4)).

4.4 Experimental result

4.4.1 Simulation Study I

Simulation dataset 1

The 100 simulation datasets are generated using the simulation design proposed by Lykou and Ntzoufras (2012), which consists of 15 predictor variables of 50 observations each. This set of simulated data comprises of predictor variables that are correlated with each other. The simulation method is repeated 100 times. This dataset has different correlated between predictors. The last five predictors are highly correlated, while there are small to moderate correlations between x_j, $j = 1, \ldots, 5$ and $x_{11},\ldots,x_{15}$. Appendix K shows true values of the Pearson correlation and the partial correlation coefficients for the simulation dataset 1.

4.4.1.1 Numerical study of Bayes factor for simple linear regression model

In this section, we use 100 simulation datasets to study the behavior of the Pearson correlation coefficient ($corr(y, x_j)$) between **y** and the candidate predictor $\mathbf{X}_j$ at the level of Bayes factor.

Figure 4.1 and Figure 4.2 display graphs between BF_{10} against λ_2 for several values of $corr(y, x_j)$ where the hypotheses H_1 is simple regression model versus H_0 is constant model. Table 4.1 and Table 4.2 show their numerical results.

According to the Bayes factor interpretation in Table 3.1, we observe the behavior of λ_2 associated with $corr(y, x_j)$ as follows:

- The value of λ_2 which produce $\mathrm{BF}_{10} > 1$,
- The value of λ_2 which produce $BF_{10} \in (1,3)$,
- The value of λ_2 which produce $BF_{10} > 3$.

The results are as follows.

Figure 4.1 and Figure 4.2 display the behavior of $corr(y, x_j)$ at the level of Bayes factor as follows:

- For small value of $corr(y, x_j)$, $corr(y, x_j) \to 0$, the larger value of λ_2 gives $BF_{10} > 1$.
- For large value of $corr(y, x_j)$, $corr(y, x_j) \to 1$, the small value of λ_2 gives the large value of BF_{10} .
- At the same value of λ_2 and $corr(y, x_j) > corr(y, x_k)$, the predictor x_j gives BF_{10} more than the predictor x_k does.

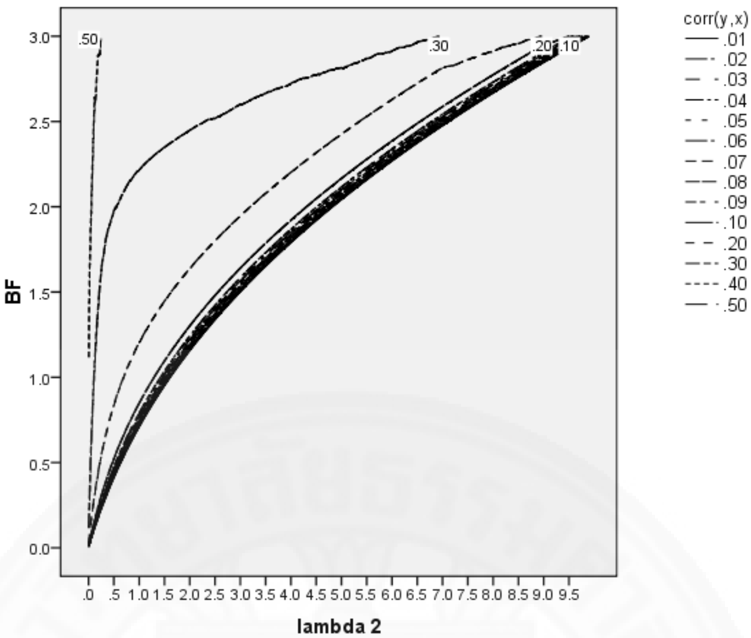


Figure 4.1 Graph between BF_{10} against λ_2 for several values of $corr(y, x_j) \geq 0$

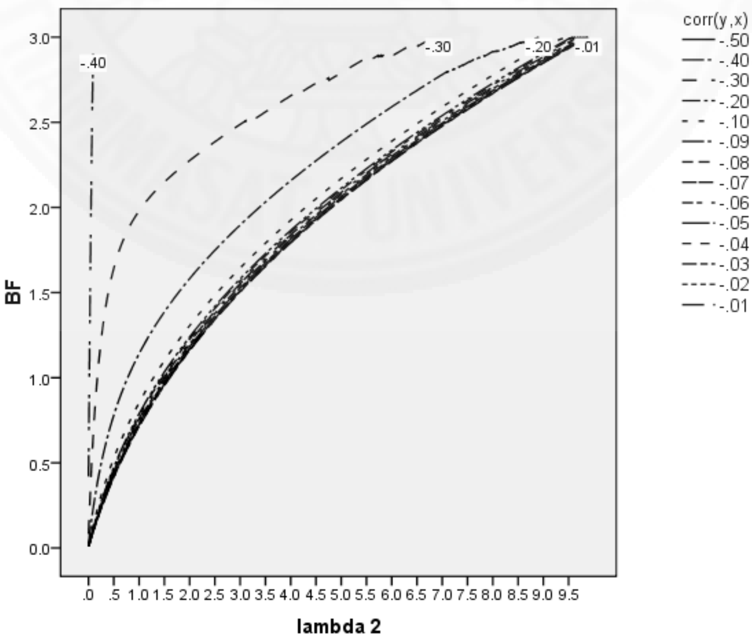


Figure 4.2 Graph between BF_{10} against λ_2 for several values of $corr(y, x_j) < 0$

Table 4.1 The value of λ_2 associated with the value of $corr(y, x_j)$ which produce the Bayes factor $BF_{10} \in (1,3)$

$corr(y, x_j)$	The value of λ_2 which produce $BF_{10} \in (1,3)$				
	minimum	maximum	mean	Std.dev.	N
-0.5	0.01	0.01	0.010		1
-0.4	0.03	0.10	0.065	0.0237	16
-0.3	0.08	6.94	3.092	1.7332	11165
-0.2	0.41	8.92	4.548	2.1823	35727
-0.1	1.15	9.50	5.298	2.3127	40809
-0.09	1.42	9.59	5.502	2.3472	7317
-0.08	1.46	9.65	5.543	2.3546	8156
-0.07	1.49	9.71	5.598	2.3644	12285
-0.06	1.53	9.77	5.640	2.3714	9857
-0.05	1.56	9.82	5.685	2.3795	7418
-0.04	1.58	9.85	5.714	2.3856	9916
-0.03	1.60	9.88	5.736	2.3891	7448
-0.02	1.61	9.89	5.751	2.3908	10766
-0.01	1.62	9.90	5.760	2.3932	16580
0.01	1.62	9.90	5.760	2.3931	19895
0.02	1.61	9.89	5.749	2.3908	15735
0.03	1.60	9.88	5.736	2.3888	14067
0.04	1.58	9.85	5.708	2.3841	9084
0.05	1.56	9.82	5.682	2.3786	14007
0.06	1.53	9.77	5.645	2.3731	8220
0.07	1.49	9.72	5.603	2.3659	9834
0.08	1.46	9.66	5.554	2.3563	18773
0.09	1.42	9.59	5.491	2.3447	10558
0.1	1.14	9.51	5.308	2.3152	74474
0.2	0.37	8.96	4.415	2.1581	161718
0.3	0.05	6.93	2.761	1.6661	78620
0.4	0.01	0.25	0.070	0.0500	923
0.5	0.01	0.01	0.010	0.0000	5

Table 4.2 The value of λ_2 associated with the value of $corr(y, x_j)$ which produce the Bayes factor $BF_{10} > 3$

$corr(y, x_j)$	The value of λ_2 which produce $BF_{10} > 3$ [a]				
	minimum	maximum	mean	Std.dev.	N
-0.5	0.02	0.02	0.020	.	1
-0.4	0.11	0.11	0.110	0.0000	2
-0.3	3.13	6.95	5.804	0.9968	20
-0.2	7.11	8.93	8.261	0.5115	48
-0.1	9.01	9.51	9.303	0.1436	51
-0.09	9.54	9.60	9.572	0.0249	9
-0.08	9.60	9.66	9.626	0.0171	10
-0.07	9.67	9.72	9.698	0.0152	16
-0.06	9.74	9.78	9.752	0.0127	12
-0.05	9.79	9.83	9.811	0.0145	9
-0.04	9.83	9.86	9.851	0.0108	12
-0.03	9.87	9.89	9.879	0.0078	9
-0.02	9.89	9.90	9.897	0.0048	13
-0.01	9.91	9.91	9.910	0.0000	20
0.01	9.90	9.91	9.910	0.0020	24
0.02	9.89	9.90	9.895	0.0051	19
0.03	9.86	9.89	9.878	0.0101	17
0.04	9.83	9.86	9.842	0.0098	11
0.05	9.79	9.83	9.807	0.0121	17
0.06	9.74	9.78	9.760	0.0170	10
0.07	9.67	9.73	9.706	0.0211	12
0.08	9.60	9.67	9.640	0.0233	23
0.09	9.52	9.60	9.557	0.0221	13
0.1	8.98	9.52	9.315	0.1612	93
0.2	6.96	8.97	8.059	0.5873	221
0.3	0.29	6.94	4.501	1.9552	181
0.4	0.02	0.26	0.065	0.0558	202
0.5	0.01	0.02	0.010	0.0015	231
0.6	0.01	0.01	0.010	0.0000	158
0.7	0.01	0.01	0.010	0.0000	28
0.8	0.01	0.01	0.010	0.0000	3

[a] The minimum value of λ_2 which produce $BF_{10} > 3$ since the computer program stop running when the value of $BF_{10} > 3$.

4.4.1.2 Bayes factor for multiple linear regression model

We follow the Step 1 and Step 2 of the method for estimating the value of the penalty parameter λ_2 based on Bayes factor described in Section 4.3.

For all pairs of hypotheses, we compute the Bayes factor for multiple linear regression model, $\mathrm{BF}_{10(Multiple\ model)}$. At the levels of Bayes factor $\mathrm{BF}_{10} = 1$, and $\mathrm{BF}_{10} = 3$, the set of significance and important predictors are identified as follows.

- Set of significance predictors is $\{X_j |\, X_j \in\ H_1 \text{ and } \mathrm{BF}_{10} > 1\}$.
- Set of important predictors is $\{X_j |\, X_j \in\ H_1 \text{ and } \mathrm{BF}_{10} > 3\}$.

The results of the set of significance and important predictors for 100 simulation datasets show in Appendix H.

The result reveals that the predictors x_1, x_5, x_7, x_{11}, and x_{13} which are actually used to generated the response variable are indicated as the significance predictors with high frequency, whereas the other predictors are less frequency. The predictors x_1, x_{11}, x_{13}, and x_7 are the important predictors with very high frequency in descending order. The predictors x_1, x_7, x_{11}, and x_{13} are the predictors which have high correlation coefficients with the response variable y.

4.4.1.3 The value of the penalty parameter λ_2 based on Bayes factor

In this section, we find the value of λ_2 based on Bayes factor for simulation dataset. The 100 datasets are generated using the simulation design proposed by Lykou and Ntzoufras (2012). The dataset consists of 15 predictor variables of 50 observations each. Using the process for estimating the value of the penalty parameter λ_2 based on Bayes factor (λ_2BF) described in Section 4.3 Step 3, the value of λ_2BF is the value θ_{M_F} of the submodel M_F which has highest posterior model probability $f(\mathbf{y}|M_F)$ (the posterior model probability using the prior in (3.4)). The values of λ_2BF are computed for 100 simulation datasets.

4.4.1.4 Numerical study for checking the validity of λ_2BF to fit the elastic net model

In this section, we correct the validity of λ_2BF derived from Section 4.4.1.3 by using Bayes factor $\text{BF}_{elastic\ net}$. For naïve elastic net estimate, the relationship between the shrinkage parameters is $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$ where $\alpha \in (0,1)$, so the $\text{BF}_{elastic\ net}$ is computed using different value of α. Table 4.3 shows summary of the $\text{BF}_{elastic\ net}$ for 100 simulation datasets. The average of prediction errors (PE) of the naïve elastic net estimates with λ_2BF are reported, the numbers in parenthesis are standard error (of PE) estimated using the bootstrap with B = 500 resampling on the 100 prediction errors.

For all values of α, the simulation result reveals that the λ_2BF gives $\text{BF}_{elastic\ net} > 1$, whereas $\text{BF}_{elastic\ net} > 3$ is derived using the small value of α $(0.01 \leq \alpha \leq 0.05)$.

Hence, the appropriate value of λ_2BF is the value θ_{M_F} of the submodel M_F which has highest posterior model probability $f(\mathbf{y}|M_F)$ (the posterior model probability using the prior in (3.4)).

Table 4.3 Summary $\mathrm{BF}_{elastic\ net}$ for 100 simulation datasets

α	$1 < \mathrm{BF}_{elastic\ net} < 3$	$3 < \mathrm{BF}_{elastic\ net} < 20$	$\mathrm{BF}_{elastic\ net} > 20$	PE [a]
0.9	100 datasets	-	-	4.2591 (0.0933)
0.8	100 datasets	-	-	4.2749 (0.1049)
0.7	100 datasets	-	-	4.2968 (0.1006)
0.6	100 datasets	-	-	4.3275 (0.0987)
0.5	100 datasets	-	-	4.3742 (0.0998)
0.4	100 datasets	-	-	4.4501 (0.1050)
0.3	100 datasets	-	-	4.5850 (0.1018)
0.2	100 datasets	-	-	4.8654 (0.1163)
0.1	100 datasets	-	-	5.7001 (0.1470)
0.05	100 datasets	16 datasets	-	7.3519 (0.2160)
0.04	100 datasets	51 datasets	1 dataset	8.2403 (0.2458)
0.03	100 datasets	94 datasets	4 datasets	9.7310 (0.3064)
0.02	100 datasets	98 datasets	38 datasets	12.5960 (0.3609)
0.01	100 datasets	98 datasets	92 datasets	17.6766 (0.3818)

[a] PE and standard error (of PE) of the naïve elastic net estimates using λ_2BF.

4.4.1.5 Result of elastic net method for simulation study I

This section shows the result of naïve elastic net estimates for simulation study I where the value of λ_2 is based on Bayes factor (λ_2BF) and the value of λ_2 is chosen by 10-fold cross validation method (λ_2CV). The elastic net method is implemented using lasso command of MATLAB2012a software. The 10-fold cross-validation (CV) method for tuning the penalty parameters (λ_1 and λ_2) is CV random partition using MATLAB2012a software. The value of λ estimated by 10-fold CV method is the λ with minimum mean prediction squared error as calculated by CV.

For naïve elastic net estimate, the relationship between the shrinkage parameters is $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$ where $\alpha \in (0,1)$. Thus, we study the performance of the elastic net with variety values of α. Appendix I shows results of the naïve elastic net estimates ($\hat{\beta}$) using 10 different values of α.

The decision criterions are the following.

1. For each estimator $\widehat{\boldsymbol{\beta}}$, its estimation accuracy is measured by the mean square error ($MSE(\widehat{\boldsymbol{\beta}})$) defined as $E\left[\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\right)^T\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\right)\right]$.

2. The variable selection performance is gauged by (C, IC), where C is the number of zero coefficients that are correctly estimated by zero and IC is the number of nonzero coefficients that are incorrectly estimated by zero.

3. The prediction accuracy is measured by the prediction error (PE) defined as $E(\mathbf{y} - \hat{\mathbf{y}})^2$ where $\hat{\mathbf{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}}$.

For each value of α, the average of $MSE(\widehat{\boldsymbol{\beta}})$, PE, C, and IC are computed based on 100 datasets. The standard errors of $MSE(\widehat{\boldsymbol{\beta}})$ and PE are estimated using the bootstrap with B = 500 resampling from the 100 $MSE(\widehat{\boldsymbol{\beta}})$'s, and 100 PE's, respectively.

Table 4.4 shows result of naïve elastic net estimates for simulation dataset 1 where the penalty parameters λ_2 are estimated by λ_2CV and λ_2BF.

Table 4.4 Model selection and fitting results of the naïve elastic net estimates for simulation dataset 1 where the penalty parameters λ_2 are estimated by λ_2CV and λ_2BF based on 100 datasets [a]

α	Method for estimation the shrinkage parameter λ_2							
	λ_2CV				**λ_2BF**			
	PE	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE	$MSE(\hat{\boldsymbol{\beta}})$	C	IC
0.9	4.7364 (0.1110)	0.1739 (0.0187)	0.53	0.07	4.2591 (0.0925)	0.1858 (0.0089)	0.13	0.04
0.8	4.7545 (0.1171)	0.1735 (0.0248)	1.22	0.13	4.2749 (0.0944)	0.1780 (0.0086)	0.32	0.05
0.7	4.7669 (0.1112)	0.1550 (0.0122)	1.81	0.23	4.2968 (0.0973)	0.1696 (0.0084)	0.55	0.07
0.6	4.8117 (0.1217)	0.1819 (0.0204)	2.47	0.29	4.3275 (0.0995)	0.1609 (0.0080)	0.84	0.09
0.5	4.8143 (0.1263)	0.1788 (0.0260)	2.92	0.37	4.3742 (0.1029)	0.1514 (0.0075)	1.25	0.11
0.4	4.8442 (0.1268)	0.1762 (0.0202)	3.49	0.37	4.4501 (0.1092)	0.1417 (0.0067)	1.81	0.19
0.3	4.8483 (0.1300)	0.1723 (0.0223)	4.04	0.50	4.5850 (0.1028)	0.1335 (0.0063)	2.69	0.28
0.2	4.9220 (0.1100)	0.1636 (0.0210)	4.70	0.58	4.8654 (0.1217)	0.1320 (0.0061)	4.02	0.43
0.1	4.8910 (0.1256)	0.1733 (0.0268)	5.00	0.64	5.7001 (0.1472)	0.1525 (0.0067)	6.14	0.92
0.09	4.9151 (0.1247)	0.1718 (0.0254)	5.12	0.68	5.8816 (0.1560)	0.1585 (0.0074)	6.55	1.02
0.08	4.8919 (0.1213)	0.1720 (0.0263)	5.02	0.71	6.1082 (0.1674)	0.1662 (0.0079)	6.94	1.13
0.07	4.8459 (0.1186)	0.1581 (0.0152)	4.93	0.69	6.4014 (0.1720)	0.1760 (0.0079)	7.30	1.21
0.06	4.8836 (0.1288)	0.1888 (0.0278)	4.95	0.69	6.7972 (0.1985)	0.1898 (0.0095)	7.75	1.38
0.05	4.8858 (0.1179)	0.1737 (0.0226)	5.08	0.69	7.3519 (0.2223)	0.2094 (0.0097)	8.15	1.52
0.04	4.8900 (0.1275)	0.1760 (0.0233)	5.16	0.78	8.2403 (0.2690)	0.2418 (0.0116)	8.71	1.69
0.03	4.9169 (0.1205)	0.1676 (0.0238)	5.36	0.74	9.7310 (0.2931)	0.2971 (0.0128)	9.19	2.12
0.02	4.8752 (0.1151)	0.1751 (0.0274)	5.11	0.73	12.5960 (0.3705)	0.3943 (0.0121)	9.65	2.92
0.01	4.9237 (0.1168)	0.1761 (0.0242)	5.42	0.77	17.6766 (0.3707)	0.5385 (0.0061)	9.97	4.63

[a] The numbers in parenthesis are the corresponding standard errors of PE and $MSE(\hat{\boldsymbol{\beta}})$ estimated using the bootstrap with B=500 resampling from the 100 PE's and 100 $MSE(\hat{\boldsymbol{\beta}})$'s, respectively. For simulation dataset 1, the true value of C is 10.

The result reveals that λ_2BF has the prediction performance better than λ_2CV when $0.2 \leq \alpha \leq 0.9$. Using λ_2BF, the prediction error tends to be large when α is close to zero.

For estimation accuracy, λ_2BF has $MSE(\hat{\boldsymbol{\beta}})$ less than λ_2CV when $0.08 \leq \alpha \leq 0.6$. The λ_2BF has minimum $MSE(\hat{\boldsymbol{\beta}})$ when α = 0.2, 0.3.

Hence, the λ_2BF has the prediction performance and estimation accuracy better than λ_2CV when $0.2 \leq \alpha \leq 0.6$.

The variable selection performance is gauged by (C, IC), the result reveals that the λ_2CV has the variable selection performance better than the λ_2BF.

Nevertheless, the λ_2BF has the variable selection performance, C is close to true value of C, better than the λ_2CV when α is small ($0.01 \leq \alpha \leq 0.1$).

4.4.2 Result of elastic net method for simulation study II

This section shows the result of naïve elastic net estimates where the penalty parameters λ_2 are estimated by λ_2BF and λ_2CV for simulation dataset 2.

Simulation dataset 2:

The datasets are simulated by the simulation method proposed by Zou and Zhang (2009). This simulation method sets the number of parameters (p) depend on the sample size (n).

Let $p = p_n = \left[4n^{1/2}\right] - 5$ for n = 100, 200, 400.

The data is generated from the linear regression model

$$\mathbf{y} = \mathbf{X}^{\mathbf{T}}\boldsymbol{\beta} + \boldsymbol{\varepsilon}\ ,$$

where $\mathbf{y}$ is an $n \times 1$ vector of response variable.

$\boldsymbol{\beta}$ is an $p \times 1$ vector of parameter of regression coefficients.

$\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of random errors, $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$, σ = 6.

$\mathbf{X} = \left[\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_p\right]^T$; $\mathbf{X}_j$ is an $n \times 1$ vector of the jth predictor variables.

$\mathbf{X}$ follows a p-dim multivariate normal distribution with zero mean and covariance $\boldsymbol{\Sigma}$, $\mathbf{X} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$, where the covariance matrix $\boldsymbol{\Sigma}$ has the entry

$$\boldsymbol{\Sigma}_{j,k} = corr(j,k) = \rho^{|j-k|}\ , 1 \le k\ , j \le p\ .$$

In this thesis, we set ρ = 0.5 and ρ = 0.75.

Let $\mathbf{1}_q$ denotes a $q \times 1$ vector of 1's, and $\mathbf{0}_{p-3q}$ denotes a $(p - 3q) \times 1$ vector of 0's.

Let the true coefficients are $\boldsymbol{\beta} = \left(3 \cdot \mathbf{1}_q, 3 \cdot \mathbf{1}_q, 3 \cdot \mathbf{1}_q, \mathbf{0}_{p-3q}\right)^T$ and $q = [p_n/9]$.

Let $\mathcal{A} = \left\{ j :\ \beta_j \neq 0, j = 1, 2, \ldots, p \right\}$. The size of $\mathcal{A}$ is the number of non-zero coefficients which are used to generate the response variable of the model. For this simulation method, the size of $\mathcal{A}$ is denoted by $|\mathcal{A}| = 3q$.

There are six cases for combination of n = 100, 200, 400 and ρ = 0.5, 0.75. The simulation method is repeated 100 times. Appendix M shows the sample correlation coefficients for the structure of the simulation dataset 2.

The decision criterions are the following.

1. For each estimator $\widehat{\boldsymbol{\beta}}$, its estimation accuracy is measured by the mean square error ($MSE(\widehat{\boldsymbol{\beta}})$) defined as $E\left[\left(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}\right)^T\left(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}\right)\right]$.

2. The variable selection performance is gauged by (C, IC), where C is the number of zero coefficients that are correctly estimated by zero and IC is the number of nonzero coefficients that are incorrectly estimated by zero.

3. The prediction accuracy is measured by the prediction error (PE) defined as $E(\mathbf{y}-\hat{\mathbf{y}})^2$ where $\hat{\mathbf{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}}$.

For each α, the average of $MSE(\widehat{\boldsymbol{\beta}})$, PE, C, and IC are computed based on 100 datasets. The standard errors of $MSE(\widehat{\boldsymbol{\beta}})$ and PE are estimated using the bootstrap with B = 500 resampling from the 100 $MSE(\widehat{\boldsymbol{\beta}})$'s, and 100 PE's, respectively.

Table 4.5 – Table 4.7 show the model selection and fitting results of the naïve elastic net estimates for simulation dataset 2 with different values of α. The numbers in parenthesis are the corresponding standard errors of PE and $MSE(\widehat{\boldsymbol{\beta}})$ estimated using the bootstrap with B = 500 resampling from the 100 PE's, and 100 $MSE(\widehat{\boldsymbol{\beta}})$'s, respectively.

Prediction accuracy

For the case n = 100, 200 and ρ = 0.5, the result reveals that λ_2BF has the prediction performance better than λ_2CV when $0.06 \le \alpha \le 0.9$.

For the case n = 100, 200 and ρ = 0.75, the result reveals that λ_2BF has the prediction performance better than λ_2CV when $0.04 \le \alpha \le 0.9$.

For the case n = 400, the result reveals that λ_2BF has the prediction performance better than λ_2CV when $0.04 \le \alpha \le 0.9$.

When α is close to zero, the method λ_2CV has the prediction performance better than the λ_2BF does.

Estimation accuracy

For almost cases, the result reveals that λ_2CV performs the estimation accuracy better than λ_2BF does. When α is small (α = 0.02, 0.03), the λ_2BF has the estimation accuracy better than λ_2CV does.

For the case n = 100, 200 and ρ = 0.75, the result reveals that λ_2BF has the $MSE(\hat{\boldsymbol{\beta}})$ less than λ_2CV when α = 0.02, 0.03.

For the case n = 100, 200 and ρ = 0.5, the result reveals that λ_2BF has the $MSE(\hat{\boldsymbol{\beta}})$ less than λ_2CV when α = 0.03, 0.04, 0.05.

For the case n = 400, the result reveals that λ_2BF has the $MSE(\hat{\boldsymbol{\beta}})$ less than λ_2CV when α = 0.02, 0.03.

Variable selection performance

The result reveals that the λ_2CV has the variable selection performance better than the λ_2BF does. For small value of α (α is close to zero), the λ_2BF has the value C tends to the true value of C better than the λ_2CV does.

Table 4.5 Model selection and fitting results of naïve elastic net estimates for simulation dataset 2: n = 100 and p_n = 35

n = 100, p_n = 35, ρ = 0.5												
ρ	$\lvert\mathcal{A}\rvert$	Truth		α	λ_2CV				λ_2BF			
		C	IC		PE	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE	$MSE(\hat{\boldsymbol{\beta}})$	C	IC
0.5	9	26	0	0.9	27.0197 (0.4712)	0.3476 (0.0116)	1.25	0	23.3527 0.4215)	0.7631 (0.0224)	0.26	0
				0.8	27.3919 (0.5048)	0.3190 (0.0098)	2.82	0	23.3792 (0.3803)	0.7469 (0.0215)	0.49	0
				0.7	27.6590 (0.5507)	0.3032 (0.0093)	4.50	0	23.4165 (0.3985)	0.7268 (0.0212)	0.77	0
				0.6	28.2900 (0.5340)	0.2769 (0.0097)	6.82	0	23.4717 (0.3931)	0.7014 (0.0200)	1.16	0
				0.5	28.6017 (0.5310)	0.2649 (0.0113)	8.82	0	23.5589 (0.4174)	0.6680 (0.0196)	1.65	0
				0.4	29.1849 (0.5599)	0.2465 (0.0110)	11.06	0	23.7110 (0.4057)	0.6218 (0.0179)	2.26	0
				0.3	29.7014 (0.5257)	0.2379 (0.0109)	13.57	0	24.0115 (0.4244)	0.5548 (0.0167)	3.51	0
				0.2	29.7958 (0.5952)	0.2433 (0.0119)	15.38	0	24.7206 (0.4334)	0.4537 (0.0139)	5.79	0
				0.1	30.2554 (0.5788)	0.2441 (0.0120)	17.60	0	27.1133 (0.4779)	0.2989 (0.0125)	11.76	0
				0.09	30.4487 (0.6386)	0.2339 (0.0099)	18.04	0	27.6419 (0.4904)	0.2807 (0.0120)	12.82	0
				0.08	30.2920 (0.6345)	0.2476 (0.0127)	17.92	0	28.2887 (0.5066)	0.2630 (0.0115)	14.06	0
				0.07	30.5599 (0.5508)	0.2417 (0.0116)	18.38	0	29.0935 (0.5255)	0.2466 (0.0111)	15.42	0
				0.06	30.3694 (0.6509)	0.2514 (0.0130)	18.20	0	30.1075 (0.5470)	0.2324 (0.0106)	16.95	0
				0.05	30.0570 (0.5936)	0.2626 (0.0135)	18.08	0	31.4413 (0.5579)	0.2220 (0.0098)	19.14	0
				0.04	30.4318 (0.6136)	0.2564 (0.0134)	18.71	0	33.3806 (0.6844)	0.2204 (0.0091)	20.96	0
				0.03	30.2753 (0.5783)	0.2610 (0.0135)	18.63	0	36.5138 (0.7253)	0.2393 (0.0088)	23.48	0.01
				0.02	30.3068 (0.6351)	0.2735 (0.0172)	18.98	0	43.1768 (0.8871)	0.3232 (0.0117)	25.29	0.03
				0.01	30.2835 (0.6037)	0.2658 (0.0140)	19.00	0	72.1491 (2.2532)	0.7680 (0.0285)	25.94	0.69
n = 100, p_n = 35, ρ = 0.75												
ρ	$\lvert\mathcal{A}\rvert$	Truth		α	λ_2CV				λ_2BF			
		C	IC		PE	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE	$MSE(\hat{\boldsymbol{\beta}})$	C	IC
0.75	9	26	0	0.9	28.2136 (0.4770)	0.3010 (0.0129)	2.14	0	23.2912 (0.3677)	1.4239 (0.0483)	0.24	0
				0.8	28.4762 (0.5303)	0.2843 (0.0140)	4.30	0	23.3149 (0.3845)	1.3957 (0.0454)	0.44	0
				0.7	28.9119 (0.4746)	0.2666 (0.0173)	6.73	0	23.3477 (0.3959)	1.3608 (0.0489)	0.71	0.01
				0.6	29.4100 (0.5178)	0.2438 (0.0179)	9.28	0	23.3953 (0.3796)	1.3163 (0.0442)	1.04	0.03
				0.5	29.8884 (0.5297)	0.2239 (0.0158)	11.98	0	23.4690 (0.3887)	1.2578 (0.0444)	1.44	0.03
				0.4	30.2770 (0.4994)	0.2110 (0.0137)	14.56	0	23.5919 (0.4130)	1.1776 (0.0441)	2.31	0.04
				0.3	30.7705 (0.4875)	0.2146 (0.0155)	17.14	0	23.8238 (0.3991)	1.0618 (0.0414)	3.36	0.06
				0.2	30.6925 (0.4935)	0.2371 (0.0174)	18.69	0.01	24.3525 (0.3913)	0.8825 (0.0379)	5.55	0.05
				0.1	30.9058 (0.5215)	0.2809 (0.0189)	20.43	0.02	26.0741 (0.4032)	0.5825 (0.0305)	10.76	0.05
				0.09	31.1063 (0.5396)	0.2778 (0.0174)	20.98	0.01	26.4299 (0.4073)	0.5456 (0.0289)	11.78	0.04
				0.08	31.0708 (0.5248)	0.2966 (0.0204)	20.82	0.02	26.8578 (0.4324)	0.5077 (0.0283)	12.56	0.04
				0.07	30.9802 (0.4698)	0.3116 (0.0227)	21.05	0.01	27.3832 (0.4320)	0.4700 (0.0261)	13.93	0.05
				0.06	31.1741 (0.5264)	0.3140 (0.0221)	21.46	0.04	28.0007 (0.4376)	0.4348 (0.0274)	15.45	0.05
				0.05	31.0699 (0.4493)	0.3234 (0.0205)	21.47	0.03	28.7370 (0.4628)	0.4021 (0.0238)	17.21	0.05
				0.04	30.8949 (0.5077)	0.3441 (0.0205)	21.44	0.04	29.6760 (0.4708)	0.3727 (0.0240)	18.86	0.05
				0.03	30.8157 (0.4875)	0.3605 (0.0233)	21.36	0.06	31.0206 (0.4897)	0.3507 (0.0210)	21.01	0.06
				0.02	30.8951 (0.4950)	0.3686 (0.0210)	21.63	0.08	33.3505 (0.5347)	0.3462 (0.0209)	23.38	0.07
				0.01	30.8513 (0.4806)	0.4132 (0.0269)	21.66	0.07	40.9169 (0.9822)	0.4196 (0.0217)	25.54	0.19

Table 4.6 Model selection and fitting results of naïve elastic net estimates for simulation dataset 2: n = 200 and p_n = 51

n = 200, p_n = 51, ρ = 0.5

ρ	$\lvert\mathcal{A}\rvert$	Truth		α	λ_2CV				λ_2BF			
		C	IC		PE	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE	$MSE(\hat{\boldsymbol{\beta}})$	C	IC
0.5	15	36	0	0.9	28.3981 (0.3431)	0.1940 (0.0048)	1.76	0	26.3703 (0.3098)	0.3395 (0.0075)	0.13	0
				0.8	28.5592 (0.3387)	0.1814 (0.0046)	3.59	0	26.3796 (0.3078)	0.3340 (0.0076)	0.35	0
				0.7	28.8540 (0.3537)	0.1691 (0.0046)	5.97	0	26.3935 (0.3155)	0.3272 (0.0075)	0.65	0
				0.6	29.1381 (0.3517)	0.1567 (0.0046)	8.64	0	26.4151 (0.2952)	0.3185 (0.0075)	1.13	0
				0.5	29.4533 (0.3756)	0.1451 (0.0042)	11.68	0	26.4508 (0.3037)	0.3068 (0.0077)	1.91	0
				0.4	29.8231 (0.3635)	0.1360 (0.0043)	15.07	0	26.5157 (0.3017)	0.2906 (0.0073)	2.81	0
				0.3	30.1617 (0.3671)	0.1290 (0.0043)	18.04	0	26.6495 (0.3141)	0.2665 (0.0068)	4.34	0
				0.2	30.5984 (0.3711)	0.1259 (0.0048)	21.45	0	26.9924 (0.3119)	0.2276 (0.0059)	7.24	0
				0.1	30.8230 (0.3596)	0.1258 (0.0048)	24.35	0	28.3058 (0.3297)	0.1618 (0.0048)	15.22	0
				0.09	30.9338 (0.4049)	0.1254 (0.0049)	24.73	0	28.6078 (0.3613)	0.1536 (0.0049)	16.82	0
				0.08	30.8147 (0.3607)	0.1261 (0.0042)	24.56	0	28.9835 (0.3447)	0.1452 (0.0048)	18.49	0
				0.07	30.9431 (0.3822)	0.1261 (0.0048)	25.22	0	29.4668 (0.3560)	0.1370 (0.0047)	20.19	0
				0.06	30.9112 (0.3852)	0.1277 (0.0047)	25.20	0	30.1043 (0.3428)	0.1293 (0.0044)	22.39	0
				0.05	30.9377 (0.3633)	0.1282 (0.0043)	25.56	0	30.9635 (0.3758)	0.1233 (0.0044)	25.14	0
				0.04	31.0337 (0.3759)	0.1281 (0.0047)	25.98	0	32.1606 (0.3791)	0.1207 (0.0043)	28.11	0
				0.03	30.9982 (0.3943)	0.1302 (0.0049)	26.15	0	34.1056 (0.4076)	0.1255 (0.0044)	31.41	0
				0.02	30.9534 (0.3943)	0.1316 (0.0051)	26.09	0	38.1089 (0.4803)	0.1543 (0.0054)	34.54	0
				0.01	30.9319 (0.3871)	0.1327 (0.0050)	26.23	0	55.1290 (0.8797)	0.3336 (0.0109)	35.94	0.02

n = 200, p_n = 51, ρ = 0.75

ρ	$\lvert\mathcal{A}\rvert$	Truth		α	λ_2CV				λ_2BF			
		C	IC		PE	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE	$MSE(\hat{\boldsymbol{\beta}})$	C	IC
0.75	15	36	0	0.9	30.0407 (0.3708)	0.1865 (0.0063)	2.73	0	26.9033 (0.3273)	0.6741 (0.0177)	0.27	0
				0.8	30.3418 (0.3610)	0.1687 (0.0062)	6.03	0	26.9161 (0.3165)	0.6617 (0.0169)	0.62	0
				0.7	30.6939 (0.3814)	0.1528 (0.0062)	9.58	0	26.9346 (0.3096)	0.6464 (0.0174)	0.97	0
				0.6	31.0237 (0.3534)	0.1417 (0.0067)	13.23	0	26.9625 (0.3036)	0.6268 (0.0166)	1.38	0
				0.5	31.4804 (0.3823)	0.1310 (0.0058)	17.33	0	27.0076 (0.3198)	0.6007 (0.0175)	2.03	0
				0.4	31.8323 (0.3675)	0.1267 (0.0070)	21.04	0	27.0871 (0.3135)	0.5646 (0.0155)	3.14	0
				0.3	32.1215 (0.3709)	0.1246 (0.0066)	24.29	0	27.2466 (0.3398)	0.5116 (0.0166)	4.83	0
				0.2	32.3657 (0.3673)	0.1319 (0.0062)	26.83	0	27.6306 (0.3271)	0.4277 (0.0147)	8.07	0
				0.1	32.4528 (0.3603)	0.1586 (0.0078)	29.03	0	28.9601 (0.3265)	0.2878 (0.0126)	15.48	0
				0.09	32.4124 (0.3588)	0.1647 (0.0083)	28.95	0	29.2411 (0.3284)	0.2710 (0.0130)	17.05	0
				0.08	32.3310 (0.3779)	0.1687 (0.0083)	29.08	0	29.5719 (0.3070)	0.2543 (0.0119)	18.61	0
				0.07	32.3409 (0.3801)	0.1750 (0.0083)	29.18	0	29.9589 (0.3472)	0.2384 (0.0113)	20.52	0
				0.06	32.3340 (0.3683)	0.1801 (0.0081)	29.30	0	30.4266 (0.3410)	0.2233 (0.0112)	22.52	0
				0.05	32.2802 (0.3879)	0.1884 (0.0085)	29.33	0	30.9956 (0.3377)	0.2098 (0.0100)	24.73	0
				0.04	32.2772 (0.3787)	0.1941 (0.0103)	29.62	0	31.7217 (0.3360)	0.1990 (0.0099)	27.41	0
				0.03	32.3147 (0.3559)	0.1990 (0.0094)	29.82	0	32.7258 (0.3594)	0.1918 (0.0087)	30.10	0
				0.02	32.1887 (0.3809)	0.2106 (0.0096)	29.62	0	34.4767 (0.3771)	0.1928 (0.0093)	33.23	0
				0.01	32.2076 (0.3716)	0.2187 (0.0104)	29.80	0	40.4271 (0.5470)	0.2383 (0.0089)	35.64	0.01

Table 4.7 Model selection and fitting results of naïve elastic net estimates for simulation dataset 2: n = 400 and p_n = 75

$\boldsymbol{n}$ **= 400, $\boldsymbol{p_n}$ = 75, $\boldsymbol{\rho}$ = 0.5**												
ρ	$\lvert\mathcal{A}\rvert$	Truth		α	λ_2CV				λ_2BF			
		C	IC		PE	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE	$MSE(\hat{\boldsymbol{\beta}})$	C	IC
0.5	24	51	0	0.9	29.7342 (0.2344)	0.1124 (0.0021)	1.98	0	28.6442 (0.2195)	0.1678 (0.0033)	0.22	0
				0.8	29.9357 (0.2417)	0.1041 (0.0023)	4.32	0	28.6470 (0.2291)	0.1660 (0.0031)	0.49	0
				0.7	30.0597 (0.2493)	0.0970 (0.0019)	7.40	0	28.6513 (0.2339)	0.1638 (0.0031)	0.98	0
				0.6	30.3179 (0.2544)	0.0892 (0.0020)	11.00	0	28.6582 (0.2288)	0.1609 (0.0031)	1.54	0
				0.5	30.6452 (0.2618)	0.0811 (0.0021)	15.57	0	28.6699 (0.2248)	0.1570 (0.0031)	2.22	0
				0.4	30.9591 (0.2519)	0.0747 (0.0020)	20.14	0	28.6921 (0.2411)	0.1514 (0.0030)	3.10	0
				0.3	31.2381 (0.2626)	0.0699 (0.0019)	24.60	0	28.7402 (0.2357)	0.1428 (0.0027)	4.64	0
				0.2	31.5680 (0.2551)	0.0660 (0.0017)	29.46	0	28.8723 (0.2247)	0.1277 (0.0026)	7.77	0
				0.1	31.6470 (0.2607)	0.0662 (0.0019)	32.67	0	29.4628 (0.2376)	0.0965 (0.0021)	15.87	0
				0.09	31.7481 (0.2753)	0.0653 (0.0020)	33.42	0	29.6174 (0.2378)	0.0917 (0.0022)	17.36	0
				0.08	31.7449 (0.2682)	0.0657 (0.0019)	33.95	0	29.8161 (0.2338)	0.0866 (0.0021)	19.45	0
				0.07	31.8506 (0.2893)	0.0649 (0.0018)	34.71	0	30.0764 (0.2399)	0.0811 (0.0019)	22.03	0
				0.06	31.8116 (0.2718)	0.0655 (0.0020)	34.73	0	30.4254 (0.2410)	0.0754 (0.0019)	25.30	0
				0.05	31.9557 (0.2607)	0.0647 (0.0018)	35.91	0	30.9026 (0.2380)	0.0698 (0.0018)	29.11	0
				0.04	31.8538 (0.2731)	0.0658 (0.0019)	35.60	0	31.5871 (0.2544)	0.0648 (0.0016)	33.83	0
				0.03	31.8563 (0.2614)	0.0659 (0.0018)	35.91	0	32.6799 (0.2632)	0.0618 (0.0017)	39.25	0
				0.02	31.7761 (0.2814)	0.0664 (0.0019)	35.69	0	34.8073 (0.2924)	0.0652 (0.0017)	46.18	0
				0.01	31.8163 (0.2749)	0.0667 (0.0019)	36.05	0	42.1288 (0.4126)	0.1075 (0.0028)	50.75	0
$\boldsymbol{n}$ **= 400, $\boldsymbol{p_n}$ = 75, $\boldsymbol{\rho}$ = 0.75**												
ρ	$\lvert\mathcal{A}\rvert$	Truth		α	λ_2CV				λ_2BF			
		C	IC		PE	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE	$MSE(\hat{\boldsymbol{\beta}})$	C	IC
0.75	24	51	0	0.9	31.2760 (0.2286)	0.1174 (0.0027)	4.34	0	29.2677 (0.2259)	0.3445 (0.0068)	0.38	0
				0.8	31.4973 (0.2576)	0.1062 (0.0025)	8.87	0	29.2735 (0.2247)	0.3391 (0.0066)	0.73	0
				0.7	31.7710 (0.2440)	0.0951 (0.0025)	14.12	0	29.2823 (0.2152)	0.3323 (0.0069)	1.21	0
				0.6	32.0469 (0.2591)	0.0870 (0.0025)	19.43	0	29.2961 (0.2148)	0.3235 (0.0067)	1.85	0
				0.5	32.3491 (0.2344)	0.0807 (0.0022)	24.68	0	29.3196 (0.2215)	0.3118 (0.0065)	2.70	0
				0.4	32.6519 (0.2501)	0.0776 (0.0025)	29.57	0	29.3627 (0.2273)	0.2955 (0.0060)	4.20	0
				0.3	32.7587 (0.2390)	0.0789 (0.0023)	33.61	0	29.4533 (0.2189)	0.2711 (0.0054)	6.42	0
				0.2	33.0183 (0.2323)	0.0826 (0.0025)	37.82	0	29.6874 (0.2169)	0.2315 (0.0053)	10.71	0
				0.1	33.1332 (0.2419)	0.0960 (0.0033)	40.53	0	30.5516 (0.2175)	0.1631 (0.0041)	21.54	0
				0.09	33.0487 (0.2551)	0.0995 (0.0032)	40.52	0	30.7440 (0.2307)	0.1543 (0.0041)	23.57	0
				0.08	33.0701 (0.2497)	0.1007 (0.0035)	40.92	0	30.9765 (0.2196)	0.1454 (0.0041)	25.77	0
				0.07	33.0901 (0.2514)	0.1029 (0.0033)	41.19	0	31.2627 (0.2411)	0.1365 (0.0038)	28.04	0
				0.06	33.0349 (0.2526)	0.1057 (0.0033)	41.08	0	31.6087 (0.2271)	0.1281 (0.0037)	31.19	0
				0.05	33.0648 (0.2486)	0.1078 (0.0033)	41.49	0	32.0283 (0.2409)	0.1206 (0.0036)	34.66	0
				0.04	33.0284 (0.2537)	0.1113 (0.0036)	41.62	0	32.5531 (0.2322)	0.1142 (0.0034)	38.46	0
				0.03	33.0120 (0.2446)	0.1137 (0.0035)	41.69	0	33.2632 (0.2436)	0.1095 (0.0035)	42.62	0
				0.02	33.1257 (0.2544)	0.1159 (0.0036)	42.49	0	34.4910 (0.2413)	0.1078 (0.0035)	47.19	0
				0.01	32.8665 (0.2389)	0.1221 (0.0038)	41.34	0	38.2721 (0.2797)	0.1226 (0.0036)	50.62	0

4.4.3 Result of adaptive elastic net estimates for simulation study II

In this section, we study the performance of two adaptive elastic net methods proposed by Ghosh (2011) and Zou and Zhang (2009) where the penalty parameters λ_2 are estimated by λ_2BF and λ_2CV.

Zou and Zhang (2009) proposed the adaptive elastic net using the elastic net estimator to construct the adaptive weight. Ghosh (2011) proposed the adaptive elastic net using the least squares estimator to construct the adaptive weight. There is no any comparative study between two adaptive elastic net methods proposed by Ghosh (2011) and Zou and Zhang (2009).

In this thesis, the adaptive weights are constructed using four estimators:

- Ordinary least squares estimator,
- Rescaled elastic net estimator with α = 0.1,
- Rescaled elastic net estimator with α = 0.5,
- Rescaled elastic net estimator with α = 0.9.

To construct the adaptive weight, we choose $\gamma > \frac{2\nu}{1-\nu}$ where $\lim_{n\to\infty} \frac{\log p}{\log n} = \nu$ as suggested by Zou and Zhang (2009). Thus, the value $\gamma = 3$ is used for fitting the adaptive elastic net of the simulation dataset 2. These datasets are the situations where the number of parameters diverges with the sample size.

In this section, the naïve adaptive elastic net estimates are fitted using the same shrinkage values (λ_1 and λ_2) of the elastic net method associated with each dataset with α = 0.9, 0.5, 0.1, 0.05, 0.01. The elastic net method is implemented using lasso command of MATLAB2012a software. The 10-fold cross-validation CV method for tuning the penalty parameters (λ_1 and λ_2) is CV random partition using MATLAB2012a software. The value of λ estimated by 10-fold CV method is the λ with minimum mean prediction squared error as calculated by CV. The adaptive elastic net method is implemented using the **gcdnet** R package (Yang & Zou 2012). Both the lasso command of MATLAB2012a software and the **gcdnet** R package solve the elastic net based on the cyclical coordinate descent algorithms proposed by Friedman, Hastie, and Tibshirani (2010).

Table 4.8 – Table 4.22 show the model selection and fitting results of the naïve adaptive elastic net estimates with different α and different adaptive weights for six cases of simulation dataset 2. For each value of α, the average of PE, $MSE(\hat{\boldsymbol{\beta}})$, C, and IC are computed based on 100 datasets. The numbers in parenthesis are the corresponding standard errors of PE and $MSE(\hat{\boldsymbol{\beta}})$ estimated using the bootstrap with B = 500 resampling from the 100 PE's, and 100 $MSE(\hat{\boldsymbol{\beta}})$'s, respectively.

The abbreviations used in the adaptive elastic net method are as follows:

λ_2BF	λ_2 is based on Bayes factor
λ_2CV	λ_2 is estimated by 10-fold cross-validation method
AENET2011	adaptive elastic net method proposed by Ghosh (2011)
AENET2009	adaptive elastic net method proposed by Zou and Zhang (2009)
w	adaptive weight
OLS	adaptive weight is constructed using ordinary least squares estimator
RENETCV01	adaptive weight is constructed using rescaled elastic net estimator with α = 0.1 and the value of λ_2 is estimated by 10-fold cross-validation method
RENETCV05	adaptive weight is constructed using rescaled elastic net estimator with α = 0.5 and the value of λ_2 is estimated by 10-fold cross-validation method
RENETCV09	adaptive weight is constructed using rescaled elastic net estimator with α = 0.9 and the value of λ_2 is estimated by 10-fold cross-validation method
RENETBF01	adaptive weight is constructed using rescaled elastic net estimator with α = 0.1 and the value of λ_2 is based on Bayes factor

RENETBF05	adaptive weight is constructed using rescaled elastic net estimator with α = 0.5 and the value of λ_2 is based on Bayes factor
RENETBF09	adaptive weight is constructed using rescaled elastic net estimator with α = 0.9 and the value of λ_2 is based on Bayes factor

4.4.3.1 Comparative study between the penalty parameter λ_2 estimated by 10-fold cross-validation method and λ_2 based on Bayes factor of the adaptive elastic net methods

Prediction performance

For almost cases of the adaptive elastic net methods, the λ_2BF has the prediction performance better than the λ_2CV does. The result reveals that the λ_2BF improves the prediction accuracy of the adaptive elastic net AENET2009 and AENET2011. This is expected according to the L_2 part of the adaptive elastic net offers the same kind of contribution as the naïve elastic net estimate. When α is small (α = 0.01, 0.05), the λ_2CV has the prediction performance better than the λ_2BF does. This prediction performance is the same as the result of the elastic net estimate.

Estimation accuracy

For almost cases, the result reveals that the λ_2CV has the estimation accuracy better than the λ_2BF does.

Variable selection performance

For the adaptive elastic net method, the result reveals that the λ_2CV has the variable selection performance better than the λ_2BF does. When α = 0.01 and the correlation is moderate, the λ_2BF has the variable selection performance, C is close to true value of C, better than the λ_2CV does.

4.4.3.2 Comparative study between elastic net and adaptive elastic net methods

Estimation accuracy

When the sample size is large (n = 200, 400) and the correlation is moderate, the AENET2009 performs the estimation accuracy better than the elastic net and AENET2011 do. The AENET2009 has the best performance when the sample is large (n = 400). This is the same as result of the comparative study of Zou and Zhang (2009). The comparative study of Zou and Zhang (2009) showed that the adaptive elastic net of Zou et al. (2009) has best performance and deals with the multicollinearity problem better than the other methods such as the lasso, the elastic net, the adaptive lasso proposed by Zou, (2006), and the smoothly clipped absolute deviation penalty (SCAD) proposed by Fan and Li (2001). The adaptive elastic net is designed for high-dimensional data analysis (Zou & Zhang, 2009, p.2). For n = 100 and α is small (α = 0.01, 0.05), the elastic net method performs the estimation accuracy better than the AENET2009, and AENET2011. For the comparative study between the elastic net and the adaptive elastic net methods, the λ_2BF and λ_2CV perform the same results of the estimation accuracy.

Variable selection performance

The AENET2009 performs the variable selection performance better than the AENET2011 and the elastic net methods do. For the λ_2CV, the adaptive weight RENETCV01 has the best performance. The adaptive weight RENETBF01 is the best when the value of λ_2 is based on Bayes factor.

Prediction performance

The prediction performance of the elastic net and AENET2011 is very close when α = 0.9, 0.5. The elastic net has the prediction performance better than AENET2011 and AENET2009 with α = 0.9. The AENET2011 is the best when α = 0.1, 0.05, 0.01. For every combination of (n, p, ρ, α), the AENET2009 performs the prediction performance worse than the elastic net and AENET2011 do. The adaptive weights RENETCV09 and RENETBF09 have the prediction performance as well as the elastic net and AENET2011 do.

4.4.3.3 Comparative study between different adaptive weights

Estimation accuracy

The AENET2009 performs the estimation accuracy better than the AENET2011 does.

Variable selection performance

The AENET2009 performs the variable selection performance better than the AENET2011 does.

Prediction performance

The AENET2011 performs the prediction performance better than the AENET2009 does. For the AENET2009, the adaptive weights RENETCV09 and RENETBF09 have the prediction performance better than the other adaptive weights where the value of penalty parameter is chosen by the same method.

The AENET2009 performs the estimation accuracy and the variable selection performance better than the AENET2011 does. For the AENET2009, the adaptive weights RENETCV09 and RENETBF09 have the prediction performance better than the other adaptive weights. Hence, the AENET2009 where adaptive weight is constructed using rescaled elastic net with α = 0.9 (RENETCV09 or RENETBF09) performs the estimation accuracy, the variable selection performance, and the prediction performance best.

Table 4.8 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 100, p_n = 35 and α = 0.9

n = 100, p_n = 35 and α = 0.9									
ρ	$\lvert\mathcal{A}\rvert$	Truth		λ_2	Model	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE
		C	IC						
0.5	9	26	0	λ_2CV	Elastic net	0.3476 (0.0116)	1.25	0	27.0197 (0.4712)
					AENET2011 w=OLS	0.3539 (0.0147)	9.81	0.06	27.1799 (0.4964)
					AENET2009 w=RENETCV01	0.2103 (0.0109)	21.64	0	30.6020 (0.5880)
					AENET2009 w=RENETCV05	0.2665 (0.0107)	17.19	0	28.7832 (0.5456)
					AENET2009 w=RENETCV09	0.3094 (0.0107)	12.72	0	27.7533 (0.5240)
0.5	9	26	0	λ_2BF	Elastic net	0.7631 (0.0224)	0.26	0	23.3527 (0.4215)
					AENET2011 w=OLS	0.7684 (0.0243)	5.88	0.06	23.4635 (0.4164)
					AENET2009 w=RENETBF01	0.4788 (0.0191)	15.96	0	25.6448 (0.4590)
					AENET2009 w=RENETBF05	0.7194 (0.0221)	7.37	0	23.5889 (0.4176)
					AENET2009 w=RENETBF09	0.7414 (0.0210)	6.23	0	23.4902 (0.4056)
0.75	9	26	0	λ_2CV	Elastic net	0.3010 (0.0129)	2.14	0	28.2136 (0.4770)
					AENET2011 w=OLS	0.3529 (0.0181)	8.83	0.11	28.1311 (0.4549)
					AENET2009 w=RENETCV01	0.1768 (0.0148)	23.14	0.05	31.5281 (0.4973)
					AENET2009 w=RENETCV05	0.2221 (0.0133)	19.46	0	30.1384 (0.4857)
					AENET2009 w=RENETCV09	0.2643 (0.0126)	14.53	0	28.9949 (0.4551)
0.75	9	26	0	λ_2BF	Elastic net	1.4239 (0.0483)	0.24	0	23.2912 (0.3677)
					AENET2011 w=OLS	1.4558 (0.0491)	4.76	0.10	23.3313 (0.3840)
					AENET2009 w=RENETBF01	0.9932 (0.0471)	14.35	0.10	24.9234 (0.4117)
					AENET2009 w=RENETBF05	1.3819 (0.0482)	6.51	0.10	23.4230 (0.3710)
					AENET2009 w=RENETBF09	1.4103 (0.0468)	5.35	0.10	23.3593 (0.3913)

Table 4.9 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 100, p_n = 35 and α = 0.5

n = 100, p_n = 35, and α = 0.5									
ρ	$\|\mathcal{A}\|$	Truth		λ_2	Model	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE
		C	IC						
0.5	9	26	0	λ_2CV	Elastic net	0.2649 (0.0113)	8.82	0	28.6017 (0.5310)
					AENET2011 w=OLS	0.3286 (0.0183)	15.24	0.08	27.8393 (0.5022)
					AENET2009 w=RENETCV01	0.1847 (0.0127)	23.53	0.02	31.4825 (0.5548)
					AENET2009 w=RENETCV05	0.2115 (0.0125)	21.12	0	30.0953 (0.5515)
					AENET2009 w=RENETCV09	0.2434 (0.0121)	18.46	0	28.9537 (0.5193)
0.5	9	26	0	λ_2BF	Elastic net	0.6680 (0.0196)	1.65	0	23.5589 (0.4174)
					AENET2011 w=OLS	0.6910 (0.0249)	9.64	0.06	23.9652 (0.4067)
					AENET2009 w=RENETBF01	0.3776 (0.0167)	19.11	0.01	26.9992 (0.4633)
					AENET2009 w=RENETBF05	0.6171 (0.0217)	11.58	0.02	24.3040 (0.4264)
					AENET2009 w=RENETBF09	0.6501 (0.0223)	10.44	0.04	24.0907 (0.3907)
0.75	9	26	0	λ_2CV	Elastic net	0.2239 (0.0158)	11.98	0	29.8884 (0.5297)
					AENET2011 w=OLS	0.4387 (0.0287)	13.71	0.28	28.0819 (0.4469)
					AENET2009 w=RENETCV01	0.1778 (0.0180)	24.50	0.09	32.1040 (0.4735)
					AENET2009 w=RENETCV05	0.1891 (0.0183)	22.45	0.02	30.9953 (0.5119)
					AENET2009 w=RENETCV09	0.2191 (0.0165)	19.95	0	29.9943 (0.4850)
0.75	9	26	0	λ_2BF	Elastic net	1.2578 (0.0444)	1.44	0.03	23.4690 (0.3887)
					AENET2011 w=OLS	1.4101 (0.0552)	7.68	0.21	23.5639 (0.3620)
					AENET2009 w=RENETBF01	0.8493 (0.0458)	16.95	0.12	25.7318 (0.4146)
					AENET2009 w=RENETBF05	1.2729 (0.0496)	9.37	0.16	23.7813 (0.3795)
					AENET2009 w=RENETBF09	1.3160 (0.0522)	8.41	0.16	23.6443 (0.4104)

Table 4.10 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 100, p_n = 35 and α = 0.1

n = 100, p_n = 35, and α = 0.1									
ρ	$\lvert\mathcal{A}\rvert$	Truth C	Truth IC	λ_2	Model	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE
0.5	9	26	0	λ_2CV	Elastic net	0.2441 (0.0120)	17.60	0	30.2554 (0.5788)
					AENET2011 w=OLS	0.4370 (0.0259)	17.86	0.15	27.2171 (0.4477)
					AENET2009 w=RENETCV01	0.2315 (0.0165)	24.25	0.07	30.9759 (0.5438)
					AENET2009 w=RENETCV05	0.2361 (0.0150)	22.97	0.01	29.9071 (0.5384)
					AENET2009 w=RENETCV09	0.2669 (0.0160)	21.27	0	28.8579 (0.5144)
0.5	9	26	0	λ_2BF	Elastic net	0.2989 (0.01250	11.76	0	27.1133 (0.4779)
					AENET2011 w=OLS	0.5144 (0.0247)	16.09	0.12	26.0361 (0.4324)
					AENET2009 w=RENETBF01	0.2656 (0.0163)	22.81	0.07	29.6286 (0.5456)
					AENET2009 w=RENETBF05	0.4198 (0.0201)	17.74	0.07	26.7502 (0.4124)
					AENET2009 w=RENETBF09	0.4498 (0.0215)	16.90	0.07	26.3562 (0.4302)
0.75	9	26	0	λ_2CV	Elastic net	0.2809 (0.0189)	20.43	0.02	30.9058 (0.5215)
					AENET2011 w=OLS	0.8516 (0.0509)	15.82	0.65	26.5794 (0.4538)
					AENET2009 w=RENETCV01	0.3443 (0.0266)	25.06	0.23	31.2873 (0.4739)
					AENET2009 w=RENETCV05	0.2829 (0.0219)	24.00	0.07	30.6306 (0.4295)
					AENET2009 w=RENETCV09	0.3180 (0.0252)	22.49	0.04	29.6471 (0.4347)
0.75	9	26	0	λ_2BF	Elastic net	0.5825 (0.0305)	10.76	0.05	26.0741 (0.4032)
					AENET2011 w=OLS	1.2666 (0.0569)	12.64	0.49	24.5743 (0.4124)
					AENET2009 w=RENETBF01	0.6445 (0.0387)	20.54	0.28	27.6939 (0.4272)
					AENET2009 w=RENETBF05	1.0317 (0.0528)	14.58	0.29	25.1131 (0.4279)
					AENET2009 w=RENETBF09	1.0932 (0.0504)	13.63	0.31	24.8522 (0.3939)

Table 4.11 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 100, p_n = 35 and α = 0.05

n = 100, p_n = 35, and α = 0.05									
ρ	$\|\mathcal{A}\|$	Truth		λ_2	Model	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE
		C	IC						
0.5	9	26	0	λ_2CV	Elastic net	0.2626 (0.0135)	18.08	0	30.0570 (0.5936)
					AENET2011 w=OLS	0.4935 (0.0275)	18.00	0.17	27.0820 (0.4480)
					AENET2009 w=RENETCV01	0.2488 (0.0168)	24.36	0.07	30.8963 (0.5301)
					AENET2009 w=RENETCV05	0.2576 (0.0170)	23.01	0.01	29.8361 (0.4747)
					AENET2009 w=RENETCV09	0.2928 (0.0182)	21.33	0.01	28.7267 (0.4765)
0.5	9	26	0	λ_2BF	Elastic net	0.2220 (0.0098)	19.14	0	31.4413 (0.5579)
					AENET2011 w=OLS	0.4513 (0.0245)	18.26	0.15	27.3977 (0.4392)
					AENET2009 w=RENETBF01	0.2448 (0.0152)	23.93	0.07	30.7372 (0.5062)
					AENET2009 w=RENETBF05	0.3593 (0.0193)	20.05	0.09	28.1082 (0.4352)
					AENET2009 w=RENETBF09	0.3823 (0.0216)	19.34	0.09	27.7489 (0.4598)
0.75	9	26	0	λ_2CV	Elastic net	0.3234 (0.0205)	21.47	0.03	31.0699 (0.4493)
					AENET2011 w=OLS	1.0182 (0.0534)	16.13	0.77	26.3585 (0.4136)
					AENET2009 w=RENETCV01	0.4112 (0.0286)	25.19	0.27	31.3058 (0.4354)
					AENET2009 w=RENETCV05	0.3370 (0.0252)	24.17	0.08	30.6088 (0.4266)
					AENET2009 w=RENETCV09	0.3725 (0.0272)	22.85	0.04	29.6644 (0.4437)
0.75	9	26	0	λ_2BF	Elastic net	0.4021 (0.0238)	17.21	0.05	28.7370 (0.4628)
					AENET2011 w=OLS	1.1921 (0.0561)	14.55	0.70	25.3387 (0.4299)
					AENET2009 w=RENETBF01	0.6083 (0.0385)	21.75	0.37	28.6943 (0.4510)
					AENET2009 w=RENETBF05	0.9334 (0.0496)	16.67	0.39	26.0638 (0.4176)
					AENET2009 w=RENETBF09	0.9894 (0.0507)	15.86	0.40	25.7593 (0.4205)

Table 4.12 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 100, p_n = 35 and α = 0.01

n = 100, p_n = 35, and α = 0.01									
ρ	$\lvert\mathcal{A}\rvert$	Truth		λ_2	Model	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE
		C	IC						
0.5	9	26	0	λ_2CV	Elastic net	0.2658 (0.0140)	19.00	0	30.2835 (0.6037)
					AENET2011 w=OLS	0.5364 (0.0316)	18.21	0.20	27.1256 (0.4695)
					AENET2009 w=RENETCV01	0.2670 (0.0177)	24.42	0.08	31.0054 (0.5280)
					AENET2009 w=RENETCV05	0.2732 (0.0171)	23.19	0.02	29.8948 (0.4920)
					AENET2009 w=RENETCV09	0.3070 (0.0172)	21.57	0.02	28.8213 (0.4639)
0.5	9	26	0	λ_2BF	Elastic net	0.7680 (0.0285)	25.94	0.69	72.1491 (2.2532)
					AENET2011 w=OLS	0.4566 (0.0243)	23.25	0.25	31.8730 (0.5203)
					AENET2009 w=RENETBF01	0.3285 (0.0197)	25.55	0.28	33.4044 (0.4883)
					AENET2009 w=RENETBF05	0.3623 (0.0201)	24.26	0.30	32.1010 (0.4648)
					AENET2009 w=RENETBF09	0.3741 (0.0216)	23.93	0.30	31.8741 (0.4802)
0.75	9	26	0	λ_2CV	Elastic net	0.4132 (0.0269)	21.66	0.07	30.8513 (0.4806)
					AENET2011 w=OLS	1.2582 (0.0623)	16.09	0.89	26.1670 (0.4537)
					AENET2009 w=RENETCV01	0.4980 (0.0325)	25.19	0.37	31.3404 (0.4720)
					AENET2009 w=RENETCV05	0.4196 (0.0268)	24.22	0.09	30.5144 (0.4619)
					AENET2009 w=RENETCV09	0.4639 (0.0353)	22.87	0.07	29.5573 (0.4481)
0.75	9	26	0	λ_2BF	Elastic net	0.4196 (0.0217)	25.54	0.19	40.9169 (0.9822)
					AENET2011 w=OLS	1.0668 (0.0512)	19.44	1.28	28.6360 (0.4224)
					AENET2009 w=RENETBF01	0.7146 (0.0384)	24.38	0.94	31.4223 (0.4704)
					AENET2009 w=RENETBF05	0.8520 (0.0427)	21.51	1.00	29.4566 (0.4390)
					AENET2009 w=RENETBF09	0.8856 (0.0426)	20.75	1.01	29.1339 (0.4334)

Table 4.13 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 200, p_n = 51 and α = 0.9

n = 200, p_n = 51, and α = 0.9									
ρ	$\|\mathcal{A}\|$	Truth C	Truth IC	λ_2	Model	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE
0.5	15	36	0	λ_2CV	Elastic net	0.1940 (0.0048)	1.76	0	28.3981 (0.3431)
					AENET2011 w=OLS	0.1705 (0.0050)	18.65	0	28.9504 (0.3383)
					AENET2009 w=RENETCV01	0.1028 (0.0040)	31.56	0	31.5120 (0.3558)
					AENET2009 w=RENETCV05	0.1300 (0.0043)	26.13	0	30.2015 (0.3414)
					AENET2009 w=RENETCV09	0.1539 (0.0047)	21.61	0	29.3792 (0.3482)
0.5	15	36	0	λ_2BF	Elastic net	0.3395 (0.0075)	0.13	0	26.3703 (0.3098)
					AENET2011 w=OLS	0.3203 (0.0079)	11.55	0	26.5750 (0.3277)
					AENET2009 w=RENETBF01	0.2259 (0.0072)	23.19	0	27.7681 (0.3384)
					AENET2009 w=RENETBF05	0.3072 (0.0083)	13.13	0	26.6610 (0.3088)
					AENET2009 w=RENETBF09	0.3152 (0.0076)	11.87	0	26.5905 (0.3223)
0.75	15	36	0	λ_2CV	Elastic net	0.1865 (0.0063)	2.73	0	30.0407 (0.3708)
					AENET2011 w=OLS	0.1887 (0.0076)	16.32	0.03	30.2291 (0.3579)
					AENET2009 w=RENETCV01	0.0893 (0.0051)	33.40	0	33.2574 (0.3627)
					AENET2009 w=RENETCV05	0.1163 (0.0053)	29.24	0	32.1318 (0.3523)
					AENET2009 w=RENETCV09	0.1440 (0.0059)	23.38	0	31.1602 (0.3565)
0.75	15	36	0	λ_2BF	Elastic net	0.6741 (0.0177)	0.27	0	26.9033 (0.3273)
					AENET2011 w=OLS	0.6670 (0.0185)	8.87	0.01	26.9949 (0.3290)
					AENET2009 w=RENETBF01	0.4386 (0.0183)	22.04	0	28.4009 (0.3249)
					AENET2009 w=RENETBF05	0.6380 (0.0188)	10.96	0	27.0717 (0.3266)
					AENET2009 w=RENETBF09	0.6526 (0.0165)	9.30	0	27.0152 (0.3208)

Table 4.14 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 200, p_n = 51 and α = 0.5

n = 200, p_n = 51, and α = 0.5									
ρ	$\lvert\mathcal{A}\rvert$	Truth		λ_2	Model	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE
		C	IC						
0.5	15	36	0	λ_2CV	Elastic net	0.1451 (0.0042)	11.68	0	29.4533 (0.3756)
					AENET2011 w=OLS	0.1272 (0.0043)	26.70	0	30.0657 (0.3534)
					AENET2009 w=RENETCV01	0.0823 (0.0031)	34.39	0	32.4601 (0.3442)
					AENET2009 w=RENETCV05	0.0940 (0.0036)	31.90	0	31.5969 (0.3397)
					AENET2009 w=RENETCV09	0.1069 (0.0039)	29.47	0	30.8325 (0.3555)
0.5	15	36	0	λ_2BF	Elastic net	0.3068 (0.0077)	1.91	0	26.4508 (0.3037)
					AENET2011 w=OLS	0.2593 (0.0072)	18.67	0	27.2650 (0.3047)
					AENET2009 w=RENETBF01	0.1667 (0.0059)	28.11	0	29.0348 (0.3299)
					AENET2009 w=RENETBF05	0.2418 (0.0068)	20.19	0	27.4650 (0.3265)
					AENET2009 w=RENETBF09	0.2521 (0.0073)	19.16	0	27.3173 (0.3074)
0.75	15	36	0	λ_2CV	Elastic net	0.1310 (0.0058)	17.33	0	31.4804 (0.3823)
					AENET2011 w=OLS	0.1802 (0.0099)	24.05	0.05	30.8438 (0.3566)
					AENET2009 w=RENETCV01	0.0866 (0.0051)	35.07	0	33.7111 (0.3689)
					AENET2009 w=RENETCV05	0.0929 (0.0056)	33.33	0	33.1468 (0.3616)
					AENET2009 w=RENETCV09	0.1071 (0.0057)	30.82	0	32.4236 (0.3521)
0.75	15	36	0	λ_2BF	Elastic net	0.6007 (0.0175)	2.03	0	27.0076 (0.3198)
					AENET2011 w=OLS	0.5955 (0.0198)	14.26	0.04	27.3704 (0.3291)
					AENET2009 w=RENETBF01	0.3464 (0.0158)	26.38	0	29.3466 (0.3388)
					AENET2009 w=RENETBF05	0.5456 (0.0190)	16.70	0.01	27.5986 (0.3205)
					AENET2009 w=RENETBF09	0.5684 (0.0188)	15.12	0.01	27.4608 (0.3124)

Table 4.15 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 200, p_n = 51 and α = 0.1

n = 200, p_n = 51, and α = 0.1									
ρ	$\lvert\mathcal{A}\rvert$	Truth		λ_2	Model	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE
		C	IC						
0.5	15	36	0	λ_2CV	Elastic net	0.1258 (0.0048)	24.35	0	30.8230 (0.3596)
					AENET2011 w=OLS	0.1315 (0.0058)	30.20	0	30.2848 (0.3252)
					AENET2009 w=RENETCV01	0.0960 (0.0046)	35.09	0	32.1598 (0.3404)
					AENET2009 w=RENETCV05	0.0997 (0.0041)	34.13	0	31.6210 (0.3299)
					AENET2009 w=RENETCV09	0.1078 (0.0048)	32.62	0	31.0502 (0.3510)
0.5	15	36	0	λ_2BF	Elastic net	0.1618 (0.0048)	15.22	0	28.3058 (0.3297)
					AENET2011 w=OLS	0.1603 (0.0058)	27.78	0	29.2982 (0.3277)
					AENET2009 w=RENETBF01	0.1145 (0.0049)	33.43	0	31.0853 (0.3208)
					AENET2009 w=RENETBF05	0.1476 (0.0055)	28.81	0	29.6398 (0.3169)
					AENET2009 w=RENETBF09	0.1541 (0.0056)	28.17	0	29.4219 (0.3081)
0.75	15	36	0	λ_2CV	Elastic net	0.1586 (0.0078)	29.03	0	32.4528 (0.3603)
					AENET2011 w=OLS	0.3035 (0.0160)	27.71	0.17	30.4843 (0.3439)
					AENET2009 w=RENETCV01	0.1691 (0.0096)	35.48	0.02	33.0853 (0.3246)
					AENET2009 w=RENETCV05	0.1494 (0.0079)	34.84	0	32.7702 (0.3121)
					AENET2009 w=RENETCV09	0.1566 (0.0080)	33.59	0	32.2369 (0.3483)
0.75	15	36	0	λ_2BF	Elastic net	0.2878 (0.0126)	15.48	0	28.9601 (0.3265)
					AENET2011 w=OLS	0.4413 (0.0185)	22.82	0.08	28.8197 (0.3267)
					AENET2009 w=RENETBF01	0.2488 (0.0136)	31.56	0.05	31.2446 (0.3400)
					AENET2009 w=RENETBF05	0.3701 (0.0166)	25.03	0.05	29.3389 (0.3368)
					AENET2009 w=RENETBF09	0.3934 (0.0163)	23.76	0.05	29.0831 (0.3301)

Table 4.16 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 200, p_n = 51 and α = 0.05

n = 200, p_n = 51, and α = 0.05									
ρ	$\|\mathcal{A}\|$	Truth		λ_2	Model	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE
		C	IC						
0.5	15	36	0	λ_2CV	Elastic net	0.1282 (0.0043)	25.56	0	30.9377 (0.3633)
					AENET2011 w=OLS	0.1384 (0.0057)	30.63	0	30.3253 (0.3338)
					AENET2009 w=RENETCV01	0.1024 (0.0048)	35.22	0	32.1677 (0.3455)
					AENET2009 w=RENETCV05	0.1053 (0.0047)	34.32	0	31.6383 (0.3521)
					AENET2009 w=RENETCV09	0.1131 (0.0051)	32.95	0	31.0913 (0.3259)
0.5	15	36	0	λ_2BF	Elastic net	0.1233 (0.0044)	25.14	0	30.9635 (0.3758)
					AENET2011 w=OLS	0.1355 (0.0056)	30.71	0	30.3794 (0.3246)
					AENET2009 w=RENETBF01	0.1073 (0.0046)	34.63	0	31.7728 (0.3313)
					AENET2009 w=RENETBF05	0.1259 (0.0051)	31.86	0	30.6821 (0.3333)
					AENET2009 w=RENETBF09	0.1298 (0.0053)	31.21	0	30.5043 (0.3270)
0.75	15	36	0	λ_2CV	Elastic net	0.1884 (0.0085)	29.33	0	32.2802 (0.3879)
					AENET2011 w=OLS	0.3634 (0.0168)	27.83	0.20	30.3748 (0.3412)
					AENET2009 w=RENETCV01	0.2050 (0.0114)	35.52	0.07	33.0088 (0.3314)
					AENET2009 w=RENETCV05	0.1806 (0.0090)	34.86	0	32.6701 (0.3125)
					AENET2009 w=RENETCV09	0.1877 (0.0099)	33.57	0	32.1453 (0.3362)
0.75	15	36	0	λ_2BF	Elastic net	0.2098 (0.0100)	24.73	0	30.9956 (0.3377)
					AENET2011 w=OLS	0.3946 (0.0185)	26.05	0.14	29.7251 (0.3229)
					AENET2009 w=RENETBF01	0.2421 (0.0141)	33.24	0.07	31.9622 (0.3392)
					AENET2009 w=RENETBF05	0.3244 (0.0163)	28.00	0.09	30.2738 (0.3179)
					AENET2009 w=RENETBF09	0.3418 (0.0164)	27.08	0.09	30.0207 (0.3247)

Table 4.17 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 200, p_n = 51 and α = 0.01

n = 200, p_n = 51, and α = 0.01									
ρ	$\lvert\mathcal{A}\rvert$	Truth		λ_2	Model	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE
		C	IC						
0.5	15	36	0	λ_2CV	Elastic net	0.1327 (0.0050)	26.23	0	30.9319 (0.3871)
					AENET2011 w=OLS	0.1465 (0.0061)	30.88	0	30.3441 (0.3253)
					AENET2009 w=RENETCV01	0.1096 (0.0049)	35.28	0	32.1736 (0.3530)
					AENET2009 w=RENETCV05	0.1119 (0.0048)	34.28	0	31.6274 (0.3307)
					AENET2009 w=RENETCV09	0.1194 (0.0049)	33.01	0	31.0925 (0.3463)
0.5	15	36	0	λ_2BF	Elastic net	0.3336 (0.0109)	35.94	0.02	55.1290 (0.8797)
					AENET2011 w=OLS	0.1478 (0.0085)	35.16	0.05	32.4494 (0.3361)
					AENET2009 w=RENETBF01	0.1388 (0.0077)	35.86	0.04	32.8692 (0.3306)
					AENET2009 w=RENETBF05	0.1383 (0.0078)	35.34	0.03	32.5339 (0.3614)
					AENET2009 w=RENETBF09	0.1390 (0.0079)	35.26	0.02	32.4706 (0.3381)
0.75	15	36	0	λ_2CV	Elastic net	0.2187 (0.0104)	29.80	0	32.2076 (0.3716)
					AENET2011 w=OLS	0.4244 (0.0211)	27.87	0.21	30.3313 (0.3272)
					AENET2009 w=RENETCV01	0.2423 (0.0132)	35.53	0.07	32.9971 (0.3323)
					AENET2009 w=RENETCV05	0.2151 (0.0110)	34.89	0	32.6585 (0.3368)
					AENET2009 w=RENETCV09	0.2212 (0.0109)	33.77	0	32.1513 (0.3533)
0.75	15	36	0	λ_2BF	Elastic net	0.2383 (0.0089)	35.64	0.01	40.4271 (0.5470)
					AENET2011 w=OLS	0.4511 (0.0228)	32.23	0.50	32.3879 (0.3054)
					AENET2009 w=RENETBF01	0.3536 (0.0180)	35.46	0.29	33.3437 (0.3188)
					AENET2009 w=RENETBF05	0.3825 (0.0211)	33.55	0.37	32.6195 (0.3294)
					AENET2009 w=RENETBF09	0.3888 (0.0210)	33.09	0.38	32.4793 (0.3131)

Table 4.18 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 400, p_n = 75 and α = 0.9

n = 400, p_n = 75, and α = 0.9									
ρ	$\lvert\mathcal{A}\rvert$	Truth		λ_2	Model	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE
		C	IC						
0.5	24	51	0	λ_2CV	Elastic net	0.1124 (0.0021)	1.98	0	29.7342 (0.2344)
					AENET2011 w=OLS	0.0831 (0.0021)	31.39	0	30.6585 (0.2381)
					AENET2009 w=RENETCV01	0.0517 (0.0017)	46.31	0	32.5244 (0.2590)
					AENET2009 w=RENETCV05	0.0638 (0.0018)	40.95	0	31.6515 (0.2415)
					AENET2009 w=RENETCV09	0.0752 (0.0020)	34.66	0	30.9963 (0.2409)
0.5	24	51	0	λ_2BF	Elastic net	0.1678 (0.0033)	0.22	0	28.6442 (0.2195)
					AENET2011 w=OLS	0.1504 (0.0032)	19.35	0	28.8792 (0.2295)
					AENET2009 w=RENETBF01	0.1173 (0.0029)	31.16	0	29.5361 (0.2406)
					AENET2009 w=RENETBF05	0.1467 (0.0031)	20.66	0	28.9226 (0.2231)
					AENET2009 w=RENETBF09	0.1493 (0.0033)	19.46	0	28.8824 (0.2350)
0.75	24	51	0	λ_2CV	Elastic net	0.1174 (0.0027)	4.34	0	31.2760 (0.2286)
					AENET2011 w=OLS	0.1011 (0.0028)	29.05	0	31.7842 (0.2450)
					AENET2009 w=RENETCV01	0.0514 (0.0018)	48.69	0	34.1032 (0.2445)
					AENET2009 w=RENETCV05	0.0625 (0.0020)	44.54	0	33.4186 (0.2478)
					AENET2009 w=RENETCV09	0.0776 (0.0022)	38.13	0	32.6319 (0.2402)
0.75	24	51	0	λ_2BF	Elastic net	0.3445 (0.0068)	0.38	0	29.2677 (0.2259)
					AENET2011 w=OLS	0.3235 (0.0073)	16.29	0	29.4157 (0.2246)
					AENET2009 w=RENETBF01	0.2225 (0.0062)	33.60	0	30.3866 (0.2326)
					AENET2009 w=RENETBF05	0.3114 (0.0067)	18.88	0	29.4875 (0.2080)
					AENET2009 w=RENETBF09	0.3189 (0.0068)	16.84	0	29.4321 (0.2121)

Table 4.19 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 400, p_n = 75 and α = 0.5

n = 400, p_n = 75, and α = 0.5									
ρ	$\lvert\mathcal{A}\rvert$	Truth		λ_2	Model	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE
		C	IC						
0.5	24	51	0	λ_2CV	Elastic net	0.0811 (0.0021)	15.57	0	30.6452 (0.2618)
					AENET2011 w=OLS	0.0536 (0.0016)	44.15	0	32.2205 (0.2482)
					AENET2009 w=RENETCV01	0.0412 (0.0012)	49.99	0	33.4250 (0.2473)
					AENET2009 w=RENETCV05	0.0447 (0.0014)	48.16	0	32.9821 (0.2616)
					AENET2009 w=RENETCV09	0.0490 (0.0015)	46.04	0	32.5572 (0.2693)
0.5	24	51	0	λ_2BF	Elastic net	0.1570 (0.0031)	2.22	0	28.6699 (0.2248)
					AENET2011 w=OLS	0.1121 (0.0028)	30.66	0	29.6565 (0.2356)
					AENET2009 w=RENETBF01	0.0829 (0.0024)	40.28	0	30.6566 (0.2294)
					AENET2009 w=RENETBF05	0.1070 (0.0030)	32.12	0	29.7788 (0.2450)
					AENET2009 w=RENETBF09	0.1103 (0.0028)	31.09	0	29.6860 (0.2250)
0.75	24	51	0	λ_2CV	Elastic net	0.0807 (0.0022)	24.68	0	32.3491 (0.2344)
					AENET2011 w=OLS	0.0782 (0.0027)	41.00	0	32.6870 (0.2333)
					AENET2009 w=RENETCV01	0.0512 (0.0017)	50.48	0	34.4035 (0.2207)
					AENET2009 w=RENETCV05	0.0530 (0.0018)	49.33	0	34.1404 (0.2268)
					AENET2009 w=RENETCV09	0.0572 (0.0018)	47.01	0	33.7126 (0.2264)
0.75	24	51	0	λ_2BF	Elastic net	0.3118 (0.0065)	2.70	0	29.3196 (0.2215)
					AENET2011 w=OLS	0.2639 (0.0069)	26.27	0	29.9153 (0.2266)
					AENET2009 w=RENETBF01	0.1629 (0.0046)	40.55	0	31.3677 (0.2326)
					AENET2009 w=RENETBF05	0.2438 (0.0061)	28.83	0	30.1069 (0.2170)
					AENET2009 w=RENETBF09	0.2545 (0.0064)	27.25	0	29.9855 (0.2255)

Table 4.20 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 400, p_n = 75 and α = 0.1

n = 400, p_n = 75, and α = 0.1									
ρ	$\|\mathcal{A}\|$	Truth		λ_2	Model	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE
		C	IC						
0.5	24	51	0	λ_2CV	Elastic net	0.0662 (0.0019)	32.67	0	31.6470 (0.2607)
					AENET2011 w=OLS	0.0530 (0.0018)	47.84	0	32.3976 (0.2555)
					AENET2009 w=RENETCV01	0.0466 (0.0014)	50.66	0	33.1104 (0.2519)
					AENET2009 w=RENETCV05	0.0472 (0.0015)	50.08	0	32.9208 (0.2502)
					AENET2009 w=RENETCV09	0.0491 (0.0015)	49.09	0	32.6705 (0.2317)
0.5	24	51	0	λ_2BF	Elastic net	0.0965 (0.0021)	15.87	0	29.4628 (0.2376)
					AENET2011 w=OLS	0.0652 (0.0021)	44.35	0	31.5734 (0.2526)
					AENET2009 w=RENETBF01	0.0546 (0.0018)	48.21	0	32.3318 (0.2502)
					AENET2009 w=RENETBF05	0.0631 (0.0020)	45.14	0	31.6889 (0.2435)
					AENET2009 w=RENETBF09	0.0644 (0.0020)	44.70	0	31.6087 (0.2337)
0.75	24	51	0	λ_2CV	Elastic net	0.0960 (0.0033)	40.53	0	33.1332 (0.2419)
					AENET2011 w=OLS	0.1175 (0.0044)	45.26	0	32.7074 (0.2318)
					AENET2009 w=RENETCV01	0.0920 (0.0032)	50.80	0	33.9059 (0.2154)
					AENET2009 w=RENETCV05	0.0874 (0.0029)	50.42	0	33.7965 (0.2215)
					AENET2009 w=RENETCV09	0.0883 (0.0031)	49.48	0	33.5688 (0.2186)
0.75	24	51	0	λ_2BF	Elastic net	0.1631 (0.0041)	21.54	0	30.5516 (0.2175)
					AENET2011 w=OLS	0.1643 (0.0050)	39.32	0	31.5063 (0.2319)
					AENET2009 w=RENETBF01	0.1171 (0.0037)	47.67	0	32.9532 (0.2188)
					AENET2009 w=RENETBF05	0.1485 (0.0047)	41.33	0	31.8226 (0.2335)
					AENET2009 w=RENETBF09	0.1549 (0.0047)	40.27	0	31.6559 (0.2364)

Table 4.21 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 400, p_n = 75 and α = 0.05

n = 400, p_n = 75, and α = 0.05									
ρ	$\lvert\mathcal{A}\rvert$	Truth		λ_2	Model	$MSE(\hat{\boldsymbol{\beta}})$	C	IC	PE
		C	IC						
0.5	24	51	0	λ_2CV	Elastic net	0.0647 (0.0018)	35.91	0	31.9557 (0.2607)
					AENET2011 w=OLS	0.0543 (0.0019)	48.31	0	32.4576 (0.2521)
					AENET2009 w=RENETCV01	0.0487 (0.0014)	50.74	0	33.0947 (0.2521)
					AENET2009 w=RENETCV05	0.0490 (0.0015)	50.29	0	32.9453 (0.2336)
					AENET2009 w=RENETCV09	0.0505 (0.0017)	49.38	0	32.7068 (0.2608)
0.5	24	51	0	λ_2BF	Elastic net	0.0698 (0.0018)	29.11	0	30.9026 (0.2380)
					AENET2011 w=OLS	0.0565 (0.0017)	47.34	0	32.2383 (0.2397)
					AENET2009 w=RENETBF01	0.0509 (0.0015)	49.72	0	32.7822 (0.2614)
					AENET2009 w=RENETBF05	0.0551 (0.0016)	47.82	0	32.3402 (0.2466)
					AENET2009 w=RENETBF09	0.0559 (0.0018)	47.58	0	32.2732 (0.2499)
0.75	24	51	0	λ_2CV	Elastic net	0.1078 (0.0033)	41.49	0	33.0648 (0.2486)
					AENET2011 w=OLS	0.1355 (0.0050)	45.42	0	32.6837 (0.2145)
					AENET2009 w=RENETCV01	0.1068 (0.0038)	50.83	0	33.8774 (0.2060)
					AENET2009 w=RENETCV05	0.1007 (0.0033)	50.44	0	33.7552 (0.2047)
					AENET2009 w=RENETCV09	0.1014 (0.0032)	49.53	0	33.5185 (0.2035)
0.75	24	51	0	λ_2BF	Elastic net	0.1206 (0.0036)	34.66	0	32.0283 (0.2409)
					AENET2011 w=OLS	0.1441 (0.0046)	43.21	0	32.2573 (0.2246)
					AENET2009 w=RENETBF01	0.1156 (0.0039)	49.23	0	33.4040 (0.2221)
					AENET2009 w=RENETBF05	0.1321 (0.0045)	44.94	0	32.5685 (0.2042)
					AENET2009 w=RENETBF09	0.1360 (0.0045)	43.95	0	32.4097 (0.2264)

Table 4.22 Model selection and fitting results of naïve adaptive elastic net estimates for simulation dataset 2: n = 400, p_n = 75 and α = 0.01

n = 400, p_n = 75, and α = 0.01									
ρ	$\lvert\mathcal{A}\rvert$	Truth		λ_2	Model	$MSE(\widehat{\boldsymbol{\beta}})$	C	IC	PE
		C	IC						
0.5	24	51	0	λ_2CV	Elastic net	0.0667 (0.0019)	36.05	0	31.8163 (0.2749)
					AENET2011 w=OLS	0.0568 (0.0019)	48.25	0	32.4293 (0.2406)
					AENET2009 w=RENETCV01	0.0511 (0.0016)	50.77	0	33.0829 (0.2538)
					AENET2009 w=RENETCV05	0.0513 (0.0016)	50.20	0	32.9206 (0.2384)
					AENET2009 w=RENETCV09	0.0528 (0.0017)	49.35	0	32.6909 (0.2405)
0.5	24	51	0	λ_2BF	Elastic net	0.1075 (0.0028)	50.75	0	42.1288 (0.4126)
					AENET2011 w=OLS	0.0577 (0.0020)	50.66	0	33.1598 (0.2459)
					AENET2009 w=RENETBF01	0.0571 (0.0020)	50.94	0	33.2145 (0.2681)
					AENET2009 w=RENETBF05	0.0570 (0.0020)	50.78	0	33.1750 (0.2506)
					AENET2009 w=RENETBF09	0.0570 (0.0020)	50.73	0	33.1639 (0.2725)
0.75	24	51	0	λ_2CV	Elastic net	0.1221 (0.0038)	41.34	0	32.8665 (0.2389)
					AENET2011 w=OLS	0.1555 (0.0053)	45.13	0.01	32.6244 (0.2160)
					AENET2009 w=RENETCV01	0.1216 (0.0041)	50.79	0	33.8519 (0.2259)
					AENET2009 w=RENETCV05	0.1146 (0.0038)	50.40	0	33.7230 (0.2249)
					AENET2009 w=RENETCV09	0.1154 (0.0038)	49.45	0	33.4806 (0.2089)
0.75	24	51	0	λ_2BF	Elastic net	0.1226 (0.0036)	50.62	0	38.2721 (0.2797)
					AENET2011 w=OLS	0.1870 (0.0072)	49.32	0.08	33.7238 (0.2295)
					AENET2009 w=RENETBF01	0.1656 (0.0069)	50.84	0.03	34.0165 (0.2163)
					AENET2009 w=RENETBF05	0.1701 (0.0067)	49.87	0.03	33.7845 (0.2214)
					AENET2009 w=RENETBF09	0.1713 (0.0068)	49.73	0.03	33.7383 (0.2164)

4.5 Real data examples

4.5.1 Real data examples for estimating the value of the penalty parameter λ_2 based on Bayes factor

In this section, we apply two real datasets to illustrate the efficiency of the method for estimating the value of the penalty parameter λ_2 based on Bayes factor. The two datasets are the diabetes data and prostate cancer data used in elastic net literature and related methods.

4.5.1.1 Diabetes Data ($p = 10, n = 442$)

The response variable (y) is a quantitative measure of disease progression one year after baseline for 442 diabetes patients. The dataset contains 10 baseline predictor variables: AGE, SEX, body mass index (BMI), average blood pressure (BP), and six blood serum measurements: tc(S1), ldl(S2), hdl(S3), tch(S4), ltg(S5), glu(S6). Appendix L shows the values of the Pearson correlation and the partial correlation coefficients for diabetes data.

The value of the penalty parameter λ_2 based on Bayes factor

Using the method for estimating the value of the penalty parameter λ_2 based on Bayes factor described in Section 4.3,

$$\lambda_2\text{BF} = 0.0128$$

is the penalty parameter λ_2 based on Bayes factor for the diabetes data.

By the method for estimating the value of the penalty parameter λ_2 based on Bayes factor and the Flowchart 3.1 described in Chapter 3 and Section 4.3, we can identify the set of predictors in the following.

Set of significance predictors is $\{X_j |\, X_j \in\, H_1$ and $\mathrm{BF}_{10} > 1\}$.

Set of important predictors is $\{X_j |\, X_j \in\, H_1$ and $\mathrm{BF}_{10} > 3\}$.

Table 4.23 shows the results of the set of significance and important predictors for diabetes data.

Table 4.23 Set of significance and important predictors for diabetes data

Diabetes data										
Predictor variables [a]	BMI	S5	BP	S4	S3	S6	S1	AGE	S2	SEX
Significance predictors [b]	BMI, BP, S5, S3, S4, S6, SEX, S2, S1, AGE									
Important predictors [c]	BMI, S5, BP, S3, S4, S6, S1, SEX, S2, AGE									

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,10$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table 4.24 Summary $\mathrm{BF}_{elastic\ net}$ for diabetes data

α	$1 < \mathrm{BF}_{elastic\ net} < 3$	$3 < \mathrm{BF}_{elastic\ net} < 20$	$\mathrm{BF}_{elastic\ net} > 20$	PE [d]
0.9	✓	-	-	2868.0250
0.8	✓	-	-	2868.0850
0.7	✓	-	-	2868.1630
0.6	✓	-	-	2868.2670
0.5	✓	-	-	2868.4140
0.4	✓	-	-	2868.6380
0.3	✓	-	-	2869.0170
0.2	✓	-	-	2869.8030
0.1	✓	-	-	2872.3810
0.05	✓	-	-	2874.8430
0.04	✓	✓	-	2875.4690
0.03	✓	✓	-	2876.7220
0.02	✓	✓	✓	2880.0060
0.01	✓	✓	✓	2893.6340

[d] PE of the naïve elastic net estimates using λ_2BF.

The results reveal that the λ_2BF gives $\mathrm{BF}_{elastic\ net} > 1$ for all α whereas $\mathrm{BF}_{elastic\ net} > 3$ is derived with the small value of α $(0.01 \leq \alpha \leq 0.04)$. The result $\mathrm{BF}_{elastic\ net}$ for diabetes data is similar to the result of $\mathrm{BF}_{elastic\ net}$ for simulation dataset 1; i.e., $\mathrm{BF}_{elastic\ net} > 3$ is derived when α is close to zero.

Table 4.25 and Table 4.26 show the results of the naïve elastic net estimates for diabetes data where the shrinkage parameters λ_2 are estimated by λ_2CV and λ_2BF, respectively. The prediction error (PE) is computed for each value of α. In this thesis, the CV method is CV random partition. This causes the different values of (λ_1, λ_2) at each value of α.

For $0.1 \leq \alpha \leq 0.3$ and $0.5 \leq \alpha \leq 0.9$, the λ_2BF has the prediction performance better than the λ_2CV does. Using the λ_2BF, the prediction error tends to be large when α is small. Using the λ_2BF with α = 0.01, the predictors AGE, ldl, and tch are excluded. This variable selection result is the same as the result of Li and Lin (2010) when the variable selection criterion of Li and Lin (2010) is the scaled neighborhood criterion.

Table 4.25 Naïve elastic net estimates of diabetes data using λ_2CV

The naïve elastic net estimates ($\hat{\beta}$) with different α values, and λ_2 is estimated by λ_2CV														
α	λ_2CV	λ_1	Predictor variables										df	PE
			AGE	BMI	BP	S1	S2	S3	S4	S5	S6	SEX		
0.9	0.0537	0.0059	-0.0080	5.4561	1.0692	-0.1798	-0.0654	-0.6519	4.1757	43.6550	0.3303	-20.9670	10	2880.171
0.8	0.0605	0.0151	-0.0050	5.4289	1.0637	-0.1642	-0.0766	-0.6649	4.1560	43.0250	0.3342	-20.7470	10	2881.734
0.7	0.0677	0.0290	-0.0015	5.4001	1.0576	-0.1504	-0.0858	-0.6759	4.1302	42.4330	0.3379	-20.5040	10	2883.446
0.6	0.0362	0.0241	-0.0121	5.5233	1.0800	-0.2357	-0.0172	-0.6030	4.1335	45.7490	0.3160	-21.4150	10	2876.303
0.5	0.0736	0.0736	0	5.3747	1.0506	-0.1382	-0.0907	-0.6865	4.0241	41.9670	0.3391	-20.1910	9	2885.256
0.4	0.0030	0.0045	-0.0313	5.6152	1.1110	-0.8008	0.4859	0.0295	5.5871	61.2020	0.2855	-22.6540	10	2861.407
0.3	0.0315	0.0735	-0.0088	5.5383	1.0786	-0.2502	0	-0.5932	3.9812	46.3570	0.3087	-21.3800	9	2875.600
0.2	0.0387	0.1549	0	5.5059	1.0662	-0.2104	-0.0251	-0.6380	3.6790	45.1630	0.3084	-20.8860	9	2878.200
0.1	0.0633	0.5699	0	5.3863	1.0217	-0.1136	-0.0648	-0.7369	2.5814	42.2340	0.3026	-18.7210	9	2890.408
0.01	0.0036	0.3579	-0	5.6538	1.0717	-0.2454	0	-0.6085	3.2822	47.8290	0.2651	-21.0770	8	2876.946

Table 4.26 Naïve elastic net estimates of diabetes data using λ_2BF

The naïve elastic net estimates ($\hat{\beta}$) with different α values and λ_2 is estimated by λ_2BF														
α	λ_2BF	λ_1	Predictor variables										df	PE
			AGE	BMI	BP	S1	S2	S3	S4	S5	S6	SEX		
0.9	0.0128	0.0014	-0.0243	5.6048	1.1007	-0.4581	0.1766	-0.3636	4.6361	52.2420	0.2975	-22.2460	10	2868.025
0.8	0.0128	0.0032	-0.0241	5.6050	1.1005	-0.4558	0.1746	-0.3659	4.6295	52.1870	0.2974	-22.2390	10	2868.085
0.7	0.0128	0.0055	-0.0238	5.6053	1.1002	-0.4529	0.1719	-0.3690	4.6210	52.1170	0.2972	-22.2290	10	2868.163
0.6	0.0128	0.0085	-0.0234	5.6056	1.0999	-0.4490	0.1684	-0.3731	4.6096	52.0230	0.2969	-22.2160	10	2868.267
0.5	0.0128	0.0128	-0.0229	5.6061	1.0994	-0.4436	0.1634	-0.3788	4.5937	51.8910	0.2967	-22.1980	10	2868.415
0.4	0.0128	0.0192	-0.0221	5.6068	1.0987	-0.4355	0.1560	-0.3874	4.5699	51.6930	0.2962	-22.1710	10	2868.638
0.3	0.0128	0.0299	-0.0207	5.6081	1.0975	-0.4219	0.1436	-0.4017	4.5301	51.3640	0.2955	-22.1260	10	2869.017
0.2	0.0128	0.0512	-0.0180	5.6105	1.0952	-0.3949	0.1189	-0.4303	4.4507	50.7050	0.2939	-22.0370	10	2869.803
0.1	0.0128	0.1152	-0.0100	5.6177	1.0881	-0.3138	0.0447	-0.5161	4.2123	48.7290	0.2895	-21.7670	10	2872.381
0.01	0.0128	1.2672	-0	5.5551	0.9960	-0.1148	-0	-0.8154	0	45.329	0.2163	-17.4370	8	2893.634

4.5.1.2 Prostate cancer data ($p = 8, n = 97$)

The prostate cancer data is a data from a prostate cancer study of Stamey, Kabalin, Mcneal, Johnstone, Freiha, Redwine, and Yang (1989). The response variable (y) is the logarithm of prostate specific antigen (lpsa). The predictor variables are eight clinical measures:

X1 - the logarithm of cancer volume (lcavol),

X2 - the logarithm of prostate weight (lweight),

X3 - age,

X4 - the logarithm of the amount of benign prostatic hyperplasia (lbph),

X5 - seminal vesicle invasion (svi),

X6 - the logarithm of capsular penetration (lcp),

X7 - the Gleason score (gleason), and

X8 - the percentage Gleason score 4 or 5 (pgg45).

Appendix L shows the values of the Pearson correlation and the partial correlation coefficients for prostate cancer data.

The value of the penalty parameter λ_2 based on Bayes factor

Using the method for estimating the value of the penalty parameter λ_2 based on Bayes factor described in Section 4.3,

$$\lambda_2\text{BF} = 0.0286$$

is the penalty parameter λ_2 based on Bayes factor for the prostate cancer data.

By the method for estimating the value of the penalty parameter λ_2 based on Bayes factor and the Flowchart 3.1 described in Chapter 3, we can identify the set of predictors in the following.

Set of significance predictors is $\{X_j | \, X_j \in \, H_1 \text{ and } \text{BF}_{10} > 1\}$.

Set of important predictors is $\{X_j | \, X_j \in \, H_1 \text{ and } \text{BF}_{10} > 3\}$.

Table 4.27 shows the results of the set of significance predictors and important predictors for prostate cancer data.

Table 4.27 Set of significance and important predictors for prostate cancer data

Prostate cancer data								
Predictor variables [a]	X1	X5	X6	X2	X8	X7	X4	X3
Significance predictors [b]	X1, X2, X5, X6, X8, X7							
Important predictors [c]	X1, X5, X2, X6, X8, X7							

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \dots ,8$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table 4.28 Summary $\mathrm{BF}_{elastic\ net}$ for prostate cancer data

α	$1 < \mathrm{BF}_{elastic\ net} < 3$	$3 < \mathrm{BF}_{elastic\ net} < 20$	$\mathrm{BF}_{elastic\ net} > 20$	PE [d]
0.9	✓	-	-	0.4459
0.8	✓	-	-	0.4475
0.7	✓	-	-	0.4504
0.6	✓	-	-	0.4548
0.5	✓	-	-	0.4583
0.4	✓	-	-	0.4655
0.3	✓	-	-	0.4805
0.2	✓	-	-	0.5042
0.1	✓	-	-	0.6089
0.05	✓	✓	-	0.9127
0.04	✓	✓	-	1.0860
0.03	✓	✓	✓	1.3187
0.02	✓	✓	✓	1.3187
0.01	✓	✓	✓	1.3187

[d] Prediction error (PE) of the naïve elastic net estimates using λ_2BF.

The results reveal that the λ_2BF gives $\mathrm{BF}_{elastic\ net} > 1$ for all α whereas $\mathrm{BF}_{elastic\ net} > 3$ is derived with the small value of α ($0.01 \le \alpha \le 0.05$). The result $\mathrm{BF}_{elastic\ net}$ for prostate cancer data is similar to the result of $\mathrm{BF}_{elastic\ net}$ for simulation dataset 1; i.e., $\mathrm{BF}_{elastic\ net} > 3$ is derived when α is close to zero.

Table 4.29 and Table 4.30 show the results of the naïve elastic net estimates for prostate cancer data where the shrinkage parameters λ_2 are estimated by λ_2CV and λ_2BF, respectively. The prediction error (PE) is computed for each value of α. In this thesis, the CV method is CV random partition. This causes the different values of (λ_1, λ_2) at each value of α.

For α = 0.8, 0.9, the λ_2BF has the prediction performance better than the λ_2CV does. Using the λ_2BF, the prediction error tends to be large when α is small. For α = 0.8, 0.9 where the λ_2BF has the prediction performance better than the λ_2CV, all predictors are included in the optimal model. This variable selection result is the same as the naïve elastic net of Zou and Hastie (2005).

Table 4.29 Naïve elastic net estimates of prostate cancer data using λ_2CV

The naïve elastic net estimates ($\hat{\beta}$) with different α values and λ_2 is estimated by λ_2CV												
α	λ_2CV	λ_1	Predictor variables								df	PE
			X1	X2	X3	X4	X5	X6	X7	X8		
0.9	0.0314	0.0035	0.5223	0.6079	-0.0178	0.0882	0.7031	-0.0579	0.0521	0.0036	8	0.4462
0.8	0.0816	0.0204	0.4654	0.5611	-0.0103	0.0679	0.6144	0	0.0404	0.0026	7	0.4596
0.7	0.0047	0.0020	0.554	0.6169	-0.0201	0.0939	0.7420	-0.0910	0.0478	0.0042	8	0.4441
0.6	0.0083	0.0055	0.5416	0.6086	-0.0185	0.0897	0.7139	-0.0704	0.0436	0.0038	8	0.4451
0.5	0.0224	0.0224	0.4973	0.5699	-0.0117	0.0717	0.6117	0	0.0238	0.0025	7	0.4555
0.4	0.0286	0.0429	0.4869	0.5263	-0.0054	0.0528	0.5828	0	0.0033	0.0022	7	0.4656
0.3	0.0127	0.0296	0.5009	0.5566	-0.0098	0.0662	0.6006	0	0.0128	0.0024	7	0.4576
0.2	0.0067	0.0269	0.5055	0.5631	-0.0108	0.0692	0.6038	0	0.0138	0.0024	7	0.4562
0.1	0.0007	0.0061	0.5489	0.6089	-0.0188	0.0904	0.7188	-0.0759	0.0406	0.0039	8	0.4447

Table 4.30 Naïve elastic net estimates of prostate cancer data using λ_2BF

The naïve elastic net estimates ($\hat{\beta}$) with different α values and λ_2 is estimated by λ_2BF												
α	λ_2BF	λ_1	Predictor variables								df	PE
			X1	X2	X3	X4	X5	X6	X7	X8		
0.9	0.0286	0.0032	0.5256	0.6092	-0.0181	0.0889	0.7078	-0.0616	0.0520	0.0037	8	0.4459
0.8	0.0286	0.0072	0.5172	0.6007	-0.0167	0.0852	0.6833	-0.0446	0.0461	0.0034	8	0.4475
0.7	0.0286	0.0123	0.5064	0.5899	-0.0148	0.0803	0.6517	-0.0228	0.0386	0.0030	8	0.4504
0.6	0.0286	0.0191	0.4946	0.5755	-0.0124	0.0740	0.6167	-0	0.0291	0.0025	7	0.4548
0.5	0.0286	0.0286	0.4915	0.5558	-0.0096	0.0655	0.6031	-0	0.0188	0.0024	7	0.4583
0.4	0.0286	0.0430	0.4869	0.5263	-0.0054	0.0528	0.5827	0	0.0033	0.0022	7	0.4656
0.3	0.0286	0.0668	0.4774	0.4852	-0	0.0329	0.5510	0	0	0.0016	5	0.4805
0.2	0.0286	0.1146	0.4638	0.4410	-0	0.0062	0.4815	0	0	0.0008	5	0.5042
0.1	0.0286	0.2578	0.4154	0.1950	0	0	0.2763	0	0	0	3	0.6089

4.5.2 Real data examples for adaptive elastic net estimates

In this section, we apply two real datasets to illustrate the prediction performance of the two adaptive elastic net methods proposed by Ghosh (2011) and Zou and Zhang (2009) where the penalty parameters λ_2 are estimated by λ_2BF and λ_2CV.

In this thesis, we study four adaptive weights:

- Ordinary least squares estimator,
- Rescaled elastic net estimator using α=0.1,
- Rescaled elastic net estimator using α=0.5,
- Rescaled elastic net estimator using α=0.9.

To construct the adaptive weight, we choose $\gamma > \frac{2\nu}{1-\nu}$ where $\lim_{n\to\infty}\frac{\log p}{\log n} = \nu$ as suggested by Zou and Zhang (2009). The adaptive elastic net estimates are fitted using the same shrinkage values (λ_1 and λ_2) of the elastic net method associated with each dataset when α = 0.1 – 0.9.

4.5.2.1 Diabetes data

In this section, we apply the adaptive elastic net estimates with different adaptive weights to diabetes data. For diabetes data, the value $\gamma = 1.3$ is used for computing the adaptive weight of the adaptive elastic net. Table 4.31 – Table 4.38 show the results of the naïve adaptive elastic net estimates for diabetes data where the penalty parameters λ_2 are estimated by λ_2BF and λ_2CV.

For $0.1 \le \alpha \le 0.3$ and $0.5 \le \alpha \le 0.9$, the λ_2BF has the prediction performance better than the λ_2CV does.

When the penalty parameter λ_2 is estimated by λ_2CV with $0.3 \le \alpha \le 0.9$, the AENET2011 performs the prediction performance better than the AENET2009 does. For the AENET2009 with α = 0.1, 0.2, 0.3, 0.5 where the penalty parameter λ_2 is estimated by λ_2CV, the adaptive weight RENETCV09 has the prediction performance better than the other adaptive weights.

When the penalty parameter λ_2 is estimated by λ_2BF with $0.2 \leq \alpha \leq 0.9$, the AENET2011 performs the prediction performance better than the AENET2009 does. For $0.2 \leq \alpha \leq 0.9$, the adaptive weight RENETBF09 has the prediction performance better than the other adaptive weights.

Table 4.31 Naïve adaptive elastic net estimates of diabetes data where the adaptive weight is OLS and λ_2 is estimated by λ_2CV

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = OLS, and λ_2 is estimated by λ_2CV														
α	λ_2CV	λ_1	Predictor variables										df	PE
			AGE	BMI	BP	S1	S2	S3	S4	S5	S6	SEX		
0.9	0.0537	0.0059	0	5.4605	1.0696	-0.1724	-0.0734	-0.6588	4.2433	43.1634	0.3316	-21.0211	9	2880.714
0.8	0.0605	0.0151	0	5.4398	1.0673	-0.1473	-0.0943	-0.6796	4.3174	41.8819	0.3400	-20.8425	9	2883.033
0.7	0.0677	0.0290	0	5.4203	1.0659	-0.1210	-0.1161	-0.7008	4.4279	40.3683	0.3501	-20.6652	9	2885.969
0.6	0.0362	0.0241	0	5.5436	1.0855	-0.1926	-0.0615	-0.6423	4.4096	43.4318	0.3252	-21.5831	9	2878.730
0.5	0.0736	0.0736	0	5.4247	1.0721	-0.0695	-0.1616	-0.7442	4.7600	36.9364	0.3702	-20.5848	9	2892.505
0.4	0.0030	0.0045	0	5.6219	1.1064	-0.7783	0.4623	0.0055	5.6460	60.2804	0.2827	-22.7569	9	2861.870
0.3	0.0315	0.0736	0	5.6024	1.1011	-0.1047	-0.1469	-0.7257	4.8039	38.9676	0.3417	-21.8237	9	2884.591
0.2	0.0387	0.1549	0	5.6304	1.1175	0	-0.2475	-0.8109	5.6299	32.1020	0.3804	-21.8161	8	2898.593
0.1	0.0633	0.5699	0	5.7828	1.1993	0.3276	-0.5552	-1.0649	9.2654	4.2853	0.5547	-22.0500	9	3026.493

Table 4.32 Naïve adaptive elastic net estimates of diabetes data where the adaptive weight is OLS and λ_2 is estimated by λ_2BF

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = OLS, and λ_2 is estimated by λ_2BF														
α	λ_2BF	λ_1	Predictor variables										df	PE
			AGE	BMI	BP	S1	S2	S3	S4	S5	S6	SEX		
0.9	0.0128	0.0014	-0.0060	5.6064	1.0976	-0.4539	0.1716	-0.3691	4.6568	52.0057	0.2952	-22.3033	10	2868.235
0.8	0.0128	0.0032	0	5.6084	1.0969	-0.4462	0.1640	-0.3775	4.6668	51.7020	0.2950	-22.3215	9	2868.510
0.7	0.0128	0.0055	0	5.6109	1.0975	-0.4363	0.1547	-0.3876	4.6760	51.3357	0.2960	-22.3245	9	2868.808
0.6	0.0128	0.0085	0	5.6143	1.0983	-0.4231	0.1422	-0.4011	4.6883	50.8473	0.2974	-22.3286	9	2869.218
0.5	0.0128	0.0128	0	5.6189	1.0993	-0.4046	0.1247	-0.4200	4.7054	50.1635	0.2993	-22.3342	9	2869.816
0.4	0.0128	0.0192	0	5.6259	1.1010	-0.3769	0.0985	-0.4484	4.7312	49.1379	0.3022	-22.3427	9	2870.767
0.3	0.0128	0.0299	0	5.6376	1.1037	-0.3306	0.0548	-0.4957	4.7740	47.4285	0.3070	-22.3569	9	2872.491
0.2	0.0128	0.0512	0	5.6590	1.1092	-0.2508	-0.0194	-0.5784	4.8501	44.3357	0.3165	-22.3933	9	2876.022
0.1	0.0128	0.1152	0	5.7258	1.1258	0	-0.2556	-0.8347	5.1354	34.7424	0.34515	-22.4957	8	2890.820

Table 4.33 Naïve adaptive elastic net estimates of diabetes data where the adaptive weight is RENETCV09 and λ_2 is estimated by λ_2CV

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = RENETCV09, and λ_2 is estimated by λ_2CV														
α	λ_2CV	λ_1	Predictor variables										df	PE
			AGE	BMI	BP	S1	S2	S3	S4	S5	S6	SEX		
0.9	0.0537	0.0059	0	5.4623	1.0624	-0.1351	-0.1029	-0.7020	4.0079	42.6058	0.3298	-20.9391	9	2882.001
0.8	0.0605	0.0151	0	5.4435	1.0495	-0.0616	-0.1617	-0.7779	3.7650	40.6681	0.3353	-20.6418	9	2886.138
0.7	0.0677	0.0290	0	5.4246	1.0335	0	-0.2092	-0.8394	3.5666	38.9844	0.3419	-20.3290	8	2890.367
0.6	0.0362	0.0241	0	5.5542	1.0555	0	-0.2184	-0.8709	3.2562	40.1933	0.3184	-21.2274	8	2885.969
0.5	0.0736	0.0736	0	5.4230	0.9997	0	-0.2038	-0.8348	3.6166	38.4832	0.3577	-20.0363	8	2892.913
0.4	0.0030	0.0045	0	5.6349	1.0992	-0.6222	0.3245	-0.1805	5.0119	56.7482	0.2828	-22.6332	9	2864.295
0.3	0.0315	0.0736	0	5.6016	1.0223	0	-0.2151	-0.8768	3.1981	40.0315	0.3295	-21.1957	8	2886.585
0.2	0.0387	0.1549	0	5.6148	0.9601	0	-0.2064	-0.8690	3.2934	39.1751	0.3582	-20.7237	8	2890.377
0.1	0.0633	0.5699	0	5.7196	0.6611	0	-0.1655	-0.8474	3.5909	35.4558	0.4875	-18.7734	8	2930.957

Table 4.34 Naïve adaptive elastic net estimates of diabetes data where the adaptive weight is RENETBF09 and λ_2 is estimated by λ_2BF

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = RENETBF09, and λ_2 is estimated by λ_2BF														
α	λ_2BF	λ_1	Predictor variables										df	PE
			AGE	BMI	BP	S1	S2	S3	S4	S5	S6	SEX		
0.9	0.0128	0.0014	-0.0099	5.6064	1.0976	-0.4467	0.1654	-0.3773	4.6192	51.9010	0.2955	-22.2818	10	2868.383
0.8	0.0128	0.0032	0	5.6086	1.0952	-0.4299	0.1495	-0.3965	4.5865	51.4382	0.2942	-22.2968	9	2868.907
0.7	0.0128	0.0055	0	5.6112	1.0946	-0.4083	0.1297	-0.4202	4.5383	50.8835	0.2947	-22.2822	9	2869.502
0.6	0.0128	0.0085	0	5.6147	1.0937	-0.3796	0.1033	-0.4519	4.4741	50.1439	0.2953	-22.2627	9	2870.326
0.5	0.0128	0.0128	0	5.6196	1.0925	-0.3393	0.0665	-0.4961	4.3841	49.1085	0.2962	-22.2354	9	2871.538
0.4	0.0128	0.0192	0	5.6269	1.0907	-0.2789	0.0111	-0.5626	4.2492	47.5554	0.2975	-22.1945	9	2873.481
0.3	0.0128	0.0299	0	5.6268	1.0881	-0.2565	0	-0.6003	3.9638	46.9815	0.2991	-22.1761	8	2874.257
0.2	0.0128	0.0512	0	5.6325	1.0826	-0.1748	-0.0606	-0.7103	3.4232	44.8796	0.3024	-22.1158	9	2877.335
0.1	0.0128	0.1152	0	5.6525	1.0693	0	-0.1988	-0.9308	2.5799	39.996	0.3139	-21.9883	8	2885.315

Table 4.35 Naïve adaptive elastic net estimates of diabetes data where the adaptive weight is RENETCV05 and λ_2 is estimated by λ_2CV

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = RENETCV05, and λ_2 is estimated by λ_2CV														
α	λ_2CV	λ_1	Predictor variables										df	PE
			AGE	BMI	BP	S1	S2	S3	S4	S5	S6	SEX		
0.9	0.0537	0.0059	0	5.4626	1.0612	-0.1510	-0.0898	-0.6842	4.0721	42.9686	0.3306	-20.9478	9	2881.405
0.8	0.0605	0.0151	0	5.4446	1.0464	-0.0980	-0.1317	-0.7376	3.9130	41.4948	0.3373	-20.6621	9	2884.668
0.7	0.0677	0.0290	0	5.4287	1.0264	-0.0347	-0.1808	-0.8010	3.7044	39.7990	0.3449	-20.3291	9	2888.927
0.6	0.0362	0.0241	0	5.5545	1.0502	-0.0747	-0.1549	-0.7854	3.5484	41.9443	0.3212	-21.2593	9	2882.833
0.5	0.0736	0.0736	0	5.4379	0.9784	0	-0.2038	-0.8333	3.5934	38.6564	0.3620	-19.9281	8	2893.479
0.4	0.0030	0.0045	0	5.6306	1.0987	-0.6862	0.3814	-0.1048	5.2405	58.3020	0.2828	-22.6605	9	2863.175
0.3	0.0315	0.0736	0	5.6182	0.9997	0	-0.2150	-0.8754	3.1645	40.2272	0.3342	-21.0731	8	2887.067
0.2	0.0387	0.1549	0	5.6491	0.9129	0	-0.2062	-0.8660	3.2275	39.5779	0.3679	-20.4713	8	2892.298
0.1	0.0633	0.5699	0	5.8385	0.4942	0	-0.1654	-0.8362	3.3955	36.8353	0.5216	-17.9102	8	2953.496

Table 4.36 Naïve adaptive elastic net estimates of diabetes data where the adaptive weight is RENETBF05 and λ_2 is estimated by λ_2BF

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = RENETBF05, and λ_2 is estimated by λ_2BF														
α	λ_2BF	λ_1	Predictor variables										df	PE
			AGE	BMI	BP	S1	S2	S3	S4	S5	S6	SEX		
0.9	0.0128	0.0014	-0.0089	5.6065	1.0974	-0.4459	0.1646	-0.3783	4.6160	51.8788	0.2953	-22.2839	10	2868.410
0.8	0.0128	0.0032	0	5.6087	1.0951	-0.4281	0.1479	-0.3987	4.5781	51.3958	0.2942	-22.2953	9	2868.955
0.7	0.0128	0.0055	0	5.6114	1.0944	-0.4052	0.1270	-0.4240	4.5239	50.8109	0.2947	-22.2797	9	2869.588
0.6	0.0128	0.0085	0	5.6151	1.0935	-0.3747	0.0992	-0.4578	4.4516	50.0309	0.2953	-22.2588	9	2870.465
0.5	0.0128	0.0128	0	5.6202	1.0922	-0.3320	0.0603	-0.5050	4.3504	48.9389	0.2962	-22.2296	9	2871.758
0.4	0.0128	0.0192	0	5.6278	1.0903	-0.2679	0.0018	-0.5759	4.1986	47.3011	0.2975	-22.1858	9	2873.840
0.3	0.0128	0.0299	0	5.6260	1.0875	-0.2534	0	-0.6080	3.8743	46.9447	0.2989	-22.1715	8	2874.364
0.2	0.0128	0.0512	0	5.6365	1.0814	-0.1353	-0.0959	-0.7553	3.2959	43.9381	0.3026	-22.0861	9	2878.825
0.1	0.0128	0.1152	0	5.6545	1.0682	0	-0.2007	-0.9279	2.6333	39.9726	0.3143	-21.9823	8	2885.295

Table 4.37 Naïve adaptive elastic net estimates of diabetes data where the adaptive weight is RENETCV01 and λ_2 is estimated by λ_2CV

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = RENETCV01, and λ_2 is estimated by λ_2CV														
α	λ_2CV	λ_1	Predictor variables										df	PE
			AGE	BMI	BP	S1	S2	S3	S4	S5	S6	SEX		
0.9	0.0537	0.0059	0	5.4653	1.0587	-0.1346	-0.1035	-0.7021	4.0068	42.6076	0.3307	-20.9197	9	2882.068
0.8	0.0605	0.0151	0	5.4510	1.0404	-0.0603	-0.1629	-0.7782	3.7624	40.6749	0.3375	-20.5936	9	2886.345
0.7	0.0677	0.0290	0	5.4386	1.0161	0	-0.2095	-0.8373	3.5714	39.0474	0.3463	-20.2412	8	2890.755
0.6	0.0362	0.0241	0	5.5667	1.0403	0	-0.2186	-0.8689	3.2607	40.2503	0.3222	-21.1477	8	2886.191
0.5	0.0736	0.0736	0	5.4578	0.9560	0	-0.2044	-0.8295	3.6287	38.6405	0.3685	-19.8173	8	2894.476
0.4	0.0030	0.0045	0	5.6374	1.0962	-0.6225	0.3250	-0.1799	5.0080	56.7713	0.2835	-22.6168	9	2864.293
0.3	0.0315	0.0736	0	5.6404	0.9758	0	-0.2159	-0.8704	3.2123	40.2075	0.3412	-20.9489	8	2887.895
0.2	0.0387	0.1549	0	5.6948	0.8633	0	-0.2081	-0.8559	3.3222	39.5385	0.3824	-20.2150	8	2895.504
0.1	0.0633	0.5699	0	5.9960	0.3179	0	-0.1708	-0.8046	3.6873	36.7075	0.5725	-17.0289	8	2990.229

Table 4.38 Naïve adaptive elastic net estimates of diabetes data where the adaptive weight is RENETBF01 and λ_2 is estimated by λ_2BF

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = RENETBF01, and λ_2 is estimated by λ_2BF														
α	λ_2BF	λ_1	Predictor variables										df	PE
			AGE	BMI	BP	S1	S2	S3	S4	S5	S6	SEX		
0.9	0.0128	0.0014	0	5.6099	1.0946	-0.4051	0.1286	-0.4270	4.4605	50.8919	0.2938	-22.2784	9	2869.565
0.8	0.0128	0.0032	0	5.6162	1.0929	-0.3363	0.0679	-0.5067	4.2185	49.2382	0.2941	-22.2298	9	2871.562
0.7	0.0128	0.0055	0	5.6228	1.0906	-0.2574	0	-0.6002	3.8995	47.3605	0.2945	-22.1736	8	2874.086
0.6	0.0128	0.0085	0	5.6318	1.0876	-0.1508	-0.0922	-0.7262	3.4768	44.8165	0.2949	-22.0975	9	2877.906
0.5	0.0128	0.0128	0	5.6450	1.0834	0	-0.2236	-0.9030	2.9038	41.2073	0.2956	-21.9902	8	2884.106
0.4	0.0128	0.0192	0	5.6469	1.0814	0	-0.2229	-0.9034	2.9064	41.1242	0.2971	-21.9840	8	2884.149
0.3	0.0128	0.0299	0	5.6500	1.0781	0	-0.2216	-0.9041	2.9107	40.9856	0.2996	-21.9738	8	2884.230
0.2	0.0128	0.0512	0	5.6563	1.0714	0	-0.2191	-0.9055	2.9193	40.7085	0.3047	-21.9532	8	2884.432
0.1	0.0128	0.1152	0	5.6750	1.0516	0	-0.2117	-0.9097	2.9450	39.8772	0.3197	-21.8916	8	2885.352

4.5.2.2 Prostate cancer data

In this section, we apply the adaptive elastic net estimates with different adaptive weights to prostate cancer data. For prostate cancer data, the value $\gamma = 1.7$ is used for computing the adaptive weight of the naïve adaptive elastic net. Table 4.39 – Table 4.46 show the results of the naïve adaptive elastic net estimates for prostate cancer data where the penalty parameters λ_2 are estimated by λ_2BF and λ_2CV.

For α = 0.8, 0.9, the λ_2BF has the prediction performance better than the λ_2CV does.

For prostate cancer data, the prediction performance of the AENET2011 and AENET2009 is very close. The AENET2011 performs the prediction performance better than the AENET2009 does. For the AENET2009 with λ_2BF, the adaptive weight RENETBF09 has the prediction performance better than the other adaptive weights.

The adaptive elastic net method is designed for high dimensional data analysis. The prostate cancer data contains eight predictors and 97 observations ($p = 8, n = 97$). For prostate cancer data, the adaptive elastic net methods (AENET2009 and AENET2011) are more parsimonious than the elastic net.

Table 4.39 Naïve adaptive elastic net estimates of prostate cancer data where the adaptive weight is OLS and λ_2 is estimated by λ_2CV

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = OLS, and λ_2 is estimated by λ_2CV												
α	λ_2CV	λ_1	Predictor variables								df	PE
			X1	X2	X3	X4	X5	X6	X7	X8		
0.9	0.0314	0.0035	0.5037	0.6406	0	0	0.6687	0	0	0	3	0.4809
0.8	0.0816	0.0204	0.4530	0.5691	0	0	0.6600	0	0	0	3	0.4911
0.7	0.0047	0.0020	0.5197	0.6534	0	0	0.6659	0	0	0	3	0.4802
0.6	0.0083	0.0055	0.5113	0.6397	0	0	0.6647	0	0	0	3	0.4806
0.5	0.0224	0.0224	0.4746	0.5763	0	0	0.6572	0	0	0	3	0.4865
0.4	0.0286	0.0429	0.4357	0.5025	0	0	0.6460	0	0	0	3	0.5011
0.3	0.0127	0.0296	0.4662	0.5529	0	0	0.6526	0	0	0	3	0.4895
0.2	0.0067	0.0269	0.4736	0.5636	0	0	0.6536	0	0	0	3	0.4875
0.1	0.0007	0.0061	0.5143	0.6398	0	0	0.6632	0	0	0	3	0.4804

Table 4.40 Naïve adaptive elastic net estimates of prostate cancer data where the adaptive weight is OLS and λ_2 is estimated by λ_2BF

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = OLS, and λ_2 is estimated by λ_2BF												
α	λ_2BF	λ_1	Predictor variables								df	PE
			X1	X2	X3	X4	X5	X6	X7	X8		
0.9	0.0286	0.0032	0.5055	0.6425	0	0	0.6687	0	0	0	3	0.4808
0.8	0.0286	0.0072	0.4984	0.6285	0	0	0.6666	0	0	0	3	0.4816
0.7	0.0286	0.0123	0.4895	0.6105	0	0	0.6637	0	0	0	3	0.4830
0.6	0.0286	0.0191	0.4775	0.5866	0	0	0.6599	0	0	0	3	0.4855
0.5	0.0286	0.0286	0.4606	0.5530	0	0	0.6545	0	0	0	3	0.4905
0.4	0.0286	0.0430	0.4355	0.5026	0	0	0.6465	0	0	0	3	0.5011
0.3	0.0286	0.0668	0.3935	0.4187	0	0	0.6331	0	0	0	3	0.5268
0.2	0.0286	0.1146	0.3095	0.2508	0	0	0.6063	0	0	0	3	0.6087
0.1	0.0286	0.2578	0.0332	0	0	0	0.5227	0	0	0	2	1.0300

Table 4.41 Naïve adaptive elastic net estimates of prostate cancer data where the adaptive weight is RENETCV09 and λ_2 is estimated by λ_2CV

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = RENETCV09, and λ_2 is estimated by λ_2CV												
α	λ_2CV	λ_1	Predictor variables								df	PE
			X1	X2	X3	X4	X5	X6	X7	X8		
0.9	0.0314	0.0035	0.5035	0.6409	0	0	0.6682	0	0	0	3	0.4810
0.8	0.0816	0.0204	0.4517	0.5707	0	0	0.6568	0	0	0	3	0.4914
0.7	0.0047	0.0020	0.5195	0.6536	0	0	0.6655	0	0	0	3	0.4802
0.6	0.0083	0.0055	0.5110	0.6402	0	0	0.6638	0	0	0	3	0.4806
0.5	0.0224	0.0224	0.4731	0.5782	0	0	0.6536	0	0	0	3	0.4868
0.4	0.0286	0.0429	0.4328	0.5062	0	0	0.6391	0	0	0	3	0.5021
0.3	0.0127	0.0296	0.4642	0.5555	0	0	0.6478	0	0	0	3	0.4899
0.2	0.0067	0.0269	0.4718	0.5659	0	0	0.6492	0	0	0	3	0.4878
0.1	0.0007	0.0061	0.5139	0.6403	0	0	0.6622	0	0	0	3	0.4805

Table 4.42 Naïve adaptive elastic net estimates of prostate cancer data where the adaptive weight is RENETBF09 and λ_2 is estimated by λ_2BF

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = RENETBF09, and λ_2 is estimated by λ_2BF												
α	λ_2BF	λ_1	Predictor variables								df	PE
			X1	X2	X3	X4	X5	X6	X7	X8		
0.9	0.0286	0.0032	0.5053	0.6427	0	0	0.6682	0	0	0	3	0.4808
0.8	0.0286	0.0072	0.4981	0.6290	0	0	0.6655	0	0	0	3	0.4816
0.7	0.0286	0.0123	0.4888	0.6114	0	0	0.6618	0	0	0	3	0.4830
0.6	0.0286	0.0191	0.4765	0.5879	0	0	0.6569	0	0	0	3	0.4857
0.5	0.0286	0.0286	0.4592	0.5550	0	0	0.6502	0	0	0	3	0.4909
0.4	0.0286	0.0430	0.4332	0.5056	0	0	0.6399	0	0	0	3	0.5019
0.3	0.0286	0.0668	0.3900	0.4233	0	0	0.6229	0	0	0	3	0.5288
0.2	0.0286	0.1146	0.3035	0.2586	0	0	0.5889	0	0	0	3	0.6143
0.1	0.0286	0.2578	0.0215	0	0	0	0.4836	0	0	0	2	1.0629

Table 4.43 Naïve adaptive elastic net estimates of prostate cancer data where the adaptive weight is RENETCV05 and λ_2 is estimated by λ_2CV

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = RENETCV05, and λ_2 is estimated by λ_2CV												
α	λ_2CV	λ_1	Predictor variables								df	PE
			X1	X2	X3	X4	X5	X6	X7	X8		
0.9	0.0314	0.0035	0.5035	0.6392	0	0	0.6646	0	0	0	3	0.4810
0.8	0.0816	0.0204	0.4517	0.5611	0	0	0.6366	0	0	0	3	0.4930
0.7	0.0047	0.0020	0.5196	0.6526	0	0	0.6634	0	0	0	3	0.4802
0.6	0.0083	0.0055	0.5111	0.6374	0	0	0.6578	0	0	0	3	0.4807
0.5	0.0224	0.0224	0.4733	0.5672	0	0	0.6300	0	0	0	3	0.4883
0.4	0.0286	0.0429	0.4331	0.4851	0	0	0.5940	0	0	0	3	0.5072
0.3	0.0127	0.0296	0.4645	0.5407	0	0	0.6162	0	0	0	3	0.4923
0.2	0.0067	0.0269	0.4722	0.5524	0	0	0.6200	0	0	0	3	0.4898
0.1	0.0007	0.0061	0.5140	0.6373	0	0	0.6555	0	0	0	3	0.4806

Table 4.44 Naïve adaptive elastic net estimates of prostate cancer data where the adaptive weight is RENETBF05 and λ_2 is estimated by λ_2BF

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = RENETBF05, and λ_2 is estimated by λ_2BF												
α	λ_2BF	λ_1	Predictor variables								df	PE
			X1	X2	X3	X4	X5	X6	X7	X8		
0.9	0.0286	0.0032	0.5054	0.6406	0	0	0.6646	0	0	0	3	0.4809
0.8	0.0286	0.0072	0.4982	0.6243	0	0	0.6573	0	0	0	3	0.4818
0.7	0.0286	0.0123	0.4889	0.6033	0	0	0.6480	0	0	0	3	0.4836
0.6	0.0286	0.0191	0.4765	0.5754	0	0	0.6355	0	0	0	3	0.4871
0.5	0.0286	0.0286	0.4592	0.5363	0	0	0.6182	0	0	0	3	0.4938
0.4	0.0286	0.0430	0.4333	0.4776	0	0	0.5921	0	0	0	3	0.5080
0.3	0.0286	0.0668	0.3900	0.3798	0	0	0.5486	0	0	0	3	0.5431
0.2	0.0286	0.1146	0.3034	0.1842	0	0	0.4616	0	0	0	3	0.6554
0.1	0.0286	0.2578	0.0049	0	0	0	0.1952	0	0	0	2	1.2114

Table 4.45 Naïve adaptive elastic net estimates of prostate cancer data where the adaptive weight is RENETCV01 and λ_2 is estimated by λ_2CV

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = RENETCV01, and λ_2 is estimated by λ_2CV												
α	λ_2CV	λ_1	Predictor variables								df	PE
			X1	X2	X3	X4	X5	X6	X7	X8		
0.9	0.0314	0.0035	0.5039	0.6398	0	0	0.6675	0	0	0	3	0.4809
0.8	0.0816	0.0204	0.4544	0.5647	0	0	0.6528	0	0	0	3	0.4913
0.7	0.0047	0.0020	0.5198	0.6529	0	0	0.6650	0	0	0	3	0.4802
0.6	0.0083	0.0055	0.5118	0.6384	0	0	0.6625	0	0	0	3	0.4806
0.5	0.0224	0.0224	0.4764	0.5711	0	0	0.6487	0	0	0	3	0.4868
0.4	0.0286	0.0429	0.4391	0.4926	0	0	0.6295	0	0	0	3	0.5019
0.3	0.0127	0.0296	0.4685	0.5460	0	0	0.6413	0	0	0	3	0.4899
0.2	0.0067	0.0269	0.4758	0.5572	0	0	0.6431	0	0	0	3	0.4878
0.1	0.0007	0.0061	0.5148	0.6383	0	0	0.6607	0	0	0	3	0.4805

Table 4.46 Naïve adaptive elastic net estimates of prostate cancer data where the adaptive weight is RENETBF01 and λ_2 is estimated by λ_2BF

The naïve adaptive elastic net estimates ($\hat{\beta}$) using different α values. Weight = RENETBF01, and λ_2 is estimated by λ_2BF												
α	λ_2BF	λ_1	Predictor variables								df	PE
			X1	X2	X3	X4	X5	X6	X7	X8		
0.9	0.0286	0.0032	0.5215	0.5523	0	0	0.6107	0	0	0	3	0.4834
0.8	0.0286	0.0072	0.5342	0.4257	0	0	0.5364	0	0	0	3	0.4937
0.7	0.0286	0.0123	0.5507	0.2628	0	0	0.4409	0	0	0	3	0.5176
0.6	0.0286	0.0191	0.5726	0.0457	0	0	0.3135	0	0	0	3	0.5683
0.5	0.0286	0.0286	0.5782	0	0	0	0.1323	0	0	0	2	0.6067
0.4	0.0286	0.0430	0.5541	0	0	0	0	0	0	0	1	0.6449
0.3	0.0286	0.0668	0.4734	0	0	0	0	0	0	0	1	0.6905
0.2	0.0286	0.1146	0.3121	0	0	0	0	0	0	0	1	0.8354
0.1	0.0286	0.2578	0	0	0	0	0	0	0	0	0	1.3187

Part II: **Bayesian variable selection for elastic net linear regression model (BVS)**

This section shows the result of the Bayesian variable selection for elastic net linear regression model where the prior distribution of $\boldsymbol{\beta}$ is a compromise between normal and double exponential distribution. Using the process of the Bayesian variable selection for elastic net linear regression model (BVS) described in Chapter 3.2 and Flowchart 3.2, the results are as follows.

From equation (4.4) and (4.5), the posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$ is

$$g(\boldsymbol{\gamma}|\mathbf{y}) \propto |\mathbf{D}_\tau|^{-\frac{1}{2}}\left|\mathbf{X}_\gamma^T\mathbf{X}_\gamma + \mathbf{D}_\tau^{-1}\right|^{-\frac{1}{2}}\left(\nu\xi + S_{\boldsymbol{\gamma}\mathbf{D}}^2\right)^{-\frac{(n+\nu)}{2}}, \tag{4.27}$$

where $$S_{\boldsymbol{\gamma}\mathbf{D}}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_\gamma\left[\mathbf{X}_\gamma^T\mathbf{X}_\gamma + \mathbf{D}_\tau^{-1}\right]^{-1}\mathbf{X}_\gamma^T\mathbf{y}. \tag{4.28}$$

For all possible 2^p models, the posterior model probability $g(\boldsymbol{\gamma}|\mathbf{y})$ are computed. The optimal model is the model which has highest posterior probability $g(\boldsymbol{\gamma}|\mathbf{y})$.

In this thesis, we propose the method for estimating the value of the penalty parameter λ_2 based on Bayes factor (λ_2BF). Thus, there are two methods for estimating the value of the penalty parameter λ_2 for the prior distribution of $\boldsymbol{\beta}$.

(a) The penalty parameter λ_2 is estimated by the 10-fold cross-validation method (CV).

(b) The penalty parameter λ_2 is based on Bayes factor (BF).

Hence, we study the BVS with different priors as follows:

(1) The Bayesian variable selection for elastic net linear regression model where the penalty parameters λ_1 and λ_2 are estimated by the 10-fold cross-validation method (BVSCV).

(2) The Bayesian variable selection for elastic net linear regression model where the penalty parameter λ_2 is based on Bayes factor (BVSBF).

For BVSBF, the penalty parameter λ_1 is derived from the relationship between λ_1 and λ_2: $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$.

For the process of Bayesian variable selection for elastic net linear regression model studied in this thesis, we set $\alpha = 0.5$, i.e., $\lambda_1 = \lambda_2$.

The 100 simulation datasets are used for analysis the performance of the Bayesian variable selection for elastic net linear regression model. These datasets are generated using the simulation design for dataset 1 described in Chapter 3. Each dataset consists of 15 predictor variables of 50 observations ($p = 15, n = 50$). Under this design, the last five predictors are highly correlated, while there are small to moderate correlations between x_j, $j = 1, \dots, 5$ and $x_{11}, \dots, x_{15}$. Appendix K shows true values of the Pearson correlation and the partial correlation coefficients for these simulation data.

The predictors x_{12}, x_{14}, and x_{15} are highly correlated with each other and theirs true coefficients are equal. Hence, three pairs of correlated predictors considered in group selection are the pair of predictors (x_{12}, x_{14}), (x_{12}, x_{15}), and (x_{14}, x_{15}).

For BVS method, the result shows the predictor variables in the optimal model which has highest posterior model probability. Appendix J shows the variable selection results of BVSCV and BVSBF of 100 simulation datasets. The BVSCV is compared to elastic net method where the penalty parameter λ_1 and λ_2 are estimated by the 10-fold cross-validation method with $\alpha = 0.5$ (ENETCV). The BVSBF is compared to elastic net method where the penalty parameter λ_2 is based on Bayes factor with $\alpha = 0.5$ (ENETBF). The BVSCV and BVSBF are analyzed using the same values of λ_1 and λ_2 used in ENETCV and ENETBF, respectively. Variable selection results of BVS and elastic net are summarized in Table 4.47 – 4.49. We summarize the behavior of BVSCV and BVSBF as follows.

4.6 Bayesian variable selection for elastic net linear regression model where the penalty parameter λ_1 and λ_2 are estimated by the 10-fold cross-validation method (BVSCV)

By considering the result of 100 simulation datasets, we observe the behavior of the process of BVSCV and discuss it with ENETCV. The results are as follows.

Variable selection

BVSCV contains the predictors which are highly correlated with the response variable more than ENETCV does. The BVSCV is more parsimonious than ENETCV.

Group selection: The model exhibits group selection if the group of correlated predictors is included into the model or excluded from the model together once one predictor among them is selected.

For this dataset, three pairs of correlated predictors are considered in group selection, i.e., the pair of predictors (x_{12}, x_{14}), (x_{12}, x_{15}), and (x_{14}, x_{15}). For BVSCV, the pair of predictors which are highly correlated with the response variable is included into the optimal model. For ENETCV, the group selection results are different: some pair of predictors which are highly correlated with the response variable is excluded from the elastic net model. There are 33 datasets which the BVSCV and ENETCV perform the same group selection results.

4.7 Bayesian variable selection for elastic net linear regression model where the penalty parameter λ_2 is based on Bayes factor (BVSBF)

By considering the result of 100 simulation datasets, we observe the behavior of the process of BVSBF and discuss it with ENETBF. The results are as follows.

Variable selection

BVSBF contains the predictors which are highly correlated with the response variable more than ENETBF does. The BVSBF is more parsimonious than ENETBF.

Group selection: The model exhibits group selection if the group of correlated predictors is included into the model or excluded from the model together once one predictor among them is selected.

For this dataset, three pairs of correlated predictors are considered in group selection, i.e., the pair of predictors (x_{12}, x_{14}), (x_{12}, x_{15}), and (x_{14}, x_{15}). For BVSBF, the pair of predictors which are highly correlated with the response variable is included into the optimal model. For ENETBF, the group selection results are different: some pair of predictors which are highly correlated with the response variable is excluded from the elastic net model. There are 61 datasets which the BVSBF and ENETBF perform the same group selection results.

Table 4.47 Frequency of predictors in the optimal model of BVS and elastic net methods based on 100 datasets

Predictors [a]	Variable selection methods			
	BVSCV	ENETCV	BVSBF	ENETBF
x_1	100	100	100	100
x_{11}	100	100	100	100
x_{13}	96	86	99	95
x_{14}	90	71	98	90
x_{15}	88	59	99	81
x_{12}	87	61	97	87
x_7	92	100	89	100
x_4	65	73	83	82
x_3	58	78	71	90
x_5	21	76	50	94
x_2	38	71	55	85
x_6	32	78	36	89
x_8	34	76	33	90
x_{10}	30	73	34	88
x_9	24	70	36	93

[a] Predictor variables are in descending order according to the average of correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$ based on 100 datasets.

Table 4.48 Number of predictors in the optimal model of BVS and elastic net methods

Method	Number of predictors in the optimal model
BVSCV	9.55 (0.18)
BVSBF	10.80 (0.15)
ENETCV	11.72 (0.17)
ENETBF	13.64 (0.11)

The average of number of predictors is computed based on 100 datasets. The numbers in parenthesis are the corresponding standard errors (of the number of predictors in the optimal model) estimated using the bootstrap with B=500 resampling from the 100 values of number of predictors in the optimal models.

Table 4.49 Degree of freedom of the naïve elastic net with variety values of α

α	Degree of freedom of the naïve elastic net	
	The penalty parameter λ_1 and λ_2 are estimated by the 10-fold cross-validation method (λ_2CV)	The penalty parameter λ_2 is based on Bayes factor method (λ_2BF)
0.9	14.40 (0.08)	14.83 (0.04)
0.8	13.65 (0.12)	14.63 (0.06)
0.7	12.96 (0.14)	14.38 (0.07)
0.6	12.24 (0.19)	14.07 (0.09)
0.5	**11.71 (0.17)**	**13.64 (0.11)**
0.4	11.14 (0.22)	13.00 (0.12)
0.3	10.46 (0.22)	12.03 (0.16)
0.2	9.72 (0.25)	10.55 (0.17)
0.1	9.36 (0.26)	7.94 (0.18)
0.09	9.20 (0.25)	7.43 (0.18)
0.08	9.27 (0.25)	6.93 (0.18)
0.07	9.38 (0.28)	6.49 (0.17)
0.06	9.36 (0.28)	5.87 (0.18)
0.05	9.23 (0.27)	5.33 (0.17)
0.04	9.06 (0.28)	4.60 (0.16)
0.03	8.90 (0.23)	3.69 (0.18)
0.02	9.16 (0.26)	2.43 (0.15)
0.01	8.81 (0.27)	0.40 (0.08)

At each value of α, the average of degree of freedom of the naïve elastic net is computed based on 100 datasets. The numbers in parenthesis are the corresponding standard errors (of degree of freedom) estimated using the bootstrap with B=500 resampling from the 100 values of degree of freedom.

4.8 Posterior probability of the optimal model derived from Bayesian variable selection for elastic net linear regression model

The BVSCV and BVSBF have different priors for values of λ_1 and λ_2. The prior of BVSCV is derived from the penalty parameters λ_1 and λ_2 of the elastic net where the values of λ_1 and λ_2 are estimated by the 10-fold cross-validation method with $\alpha = 0.5$. For BVSBF, the penalty parameter λ_2 is derived from the method for estimating the value of the penalty parameter λ_2 based on Bayes factor (λ_2BF) proposed in the first objective of this thesis and the associated value of λ_1 is derived from the relationship between λ_1 and λ_2: $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$. In Chapter 4 Part I, the result shows that λ_2BF = θ_{M_F} of the submodel which has highest posterior model probability under the elastic net prior when $\lambda_1 = 0$. Thus, the λ_2BF is derived from ridge prior.

Posterior probability of the optimal model derived from BVS methods are shown in Figure 4.3 – Figure 4.5 and Table 4.49. The optimal model derived from BVSCV has average of posterior probability more than the BVSBF does. There are 92 datasets of the optimal model of the BVSCV have the maximum posterior probability more than the BVSBF do.

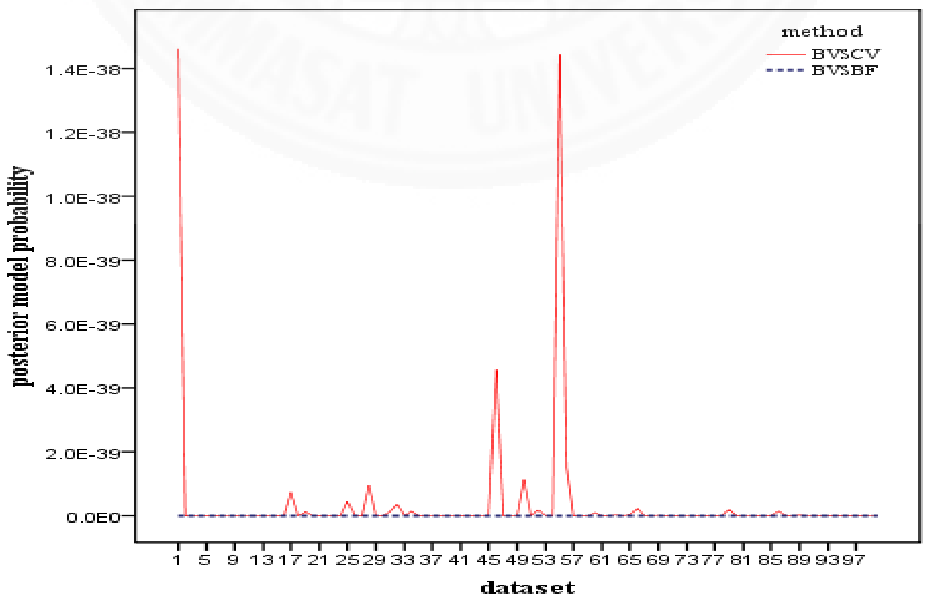


Figure 4.3 Posterior probabilities of the optimal model derived from BVS

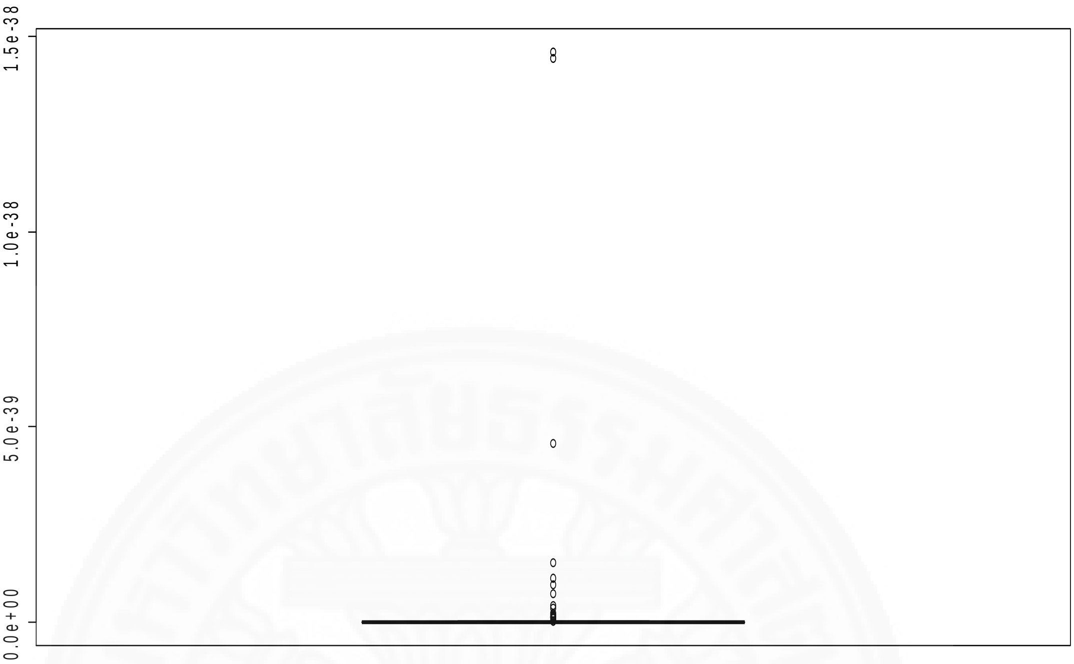


Figure 4.4 Boxplot of the posterior probabilities of the optimal model derived from BVSCV

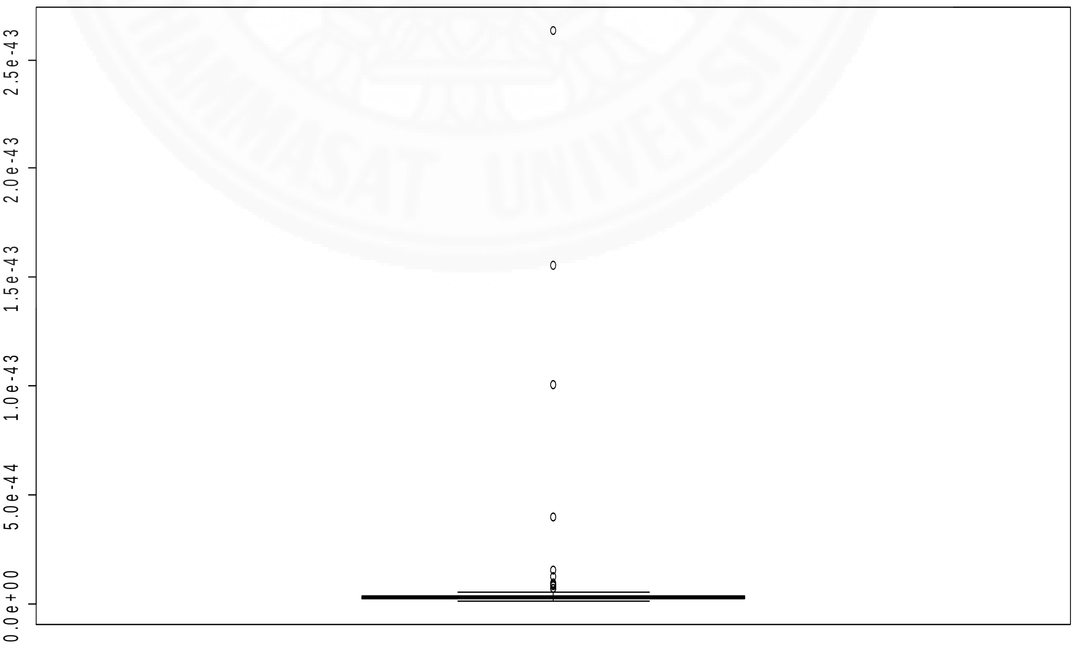


Figure 4.5 Boxplot of the posterior probabilities of the optimal model derived from BVSBF

Table 4.50 Maximum posterior probabilities of the optimal model of BVSCV and BVSBF for 100 datasets

Dataset	Maximum posterior probabilities of the optimal model	
	BVSCV	BVSBF
1	1.461872E-38	2.707062E-45
2	5.199319E-44	1.875833E-45
3	8.946127E-44	4.995641E-45
4	8.197793E-42	2.535883E-45
5*	1.306272E-45	1.839685E-45
6	1.950081E-44	3.016918E-45
7	9.943054E-44	1.933431E-45
8	8.64485E-43	2.229908E-45
9	7.258095E-42	2.869629E-45
10	4.243489E-45	2.189378E-45
11	8.457695E-42	1.867549E-45
12	1.125819E-42	2.462249E-45
13	6.585411E-45	2.655616E-45
14	1.62583E-42	2.629052E-45
15	7.322368E-45	3.36889E-45
16	8.926086E-44	3.602886E-45
17	7.341503E-40	4.825042E-45
18	4.946682E-44	1.678632E-45
19	1.18172E-40	2.739267E-45
20	2.069106E-42	1.153888E-45
21*	5.074237E-45	1.522133E-44
22	2.163992E-42	2.319427E-45
23	4.170968E-44	1.580964E-45
24	3.231266E-45	2.574703E-45
25	4.356681E-40	6.877465E-45
26	5.428501E-43	2.499132E-45
27*	1.784078E-45	1.943375E-45
28	9.515256E-40	2.737848E-45
29	1.125451E-44	3.39633E-45
30	2.77551E-45	2.091888E-45
31	1.19721E-40	8.162128E-45
32	3.482547E-40	1.429904E-45
33	2.99387E-44	3.202629E-45
34	1.430547E-40	1.251609E-44
35	6.501039E-45	1.883072E-45
36	3.790597E-43	2.525786E-45
37	8.259437E-43	3.468347E-45
38	5.09297E-42	3.180263E-45
39*	9.667087E-46	1.557844E-43
40	5.659833E-43	2.401667E-45
41	4.867275E-44	2.383895E-45
42	7.802377E-43	1.61934E-45
43	1.565756E-44	2.721493E-45
44	1.969323E-42	2.887343E-45
45	2.302726E-44	6.917403E-45
46	4.574766E-39	3.384733E-45
47	8.944176E-45	3.783937E-45
48	7.508842E-42	4.229539E-45
49*	1.191701E-45	2.312638E-45
50	1.128007E-39	3.499586E-45

* BVSCV has posterior probability less than BVSBF.

Table 4.50 Maximum posterior probabilities of the optimal model of BVSCV and BVSBF for 100 datasets (Cont.)

Dataset	Maximum posterior probabilities of the optimal model	
	BVSCV	BVSBF
51*	2.736513E-45	3.933065E-45
52	1.729925E-40	3.967795E-45
53	8.895349E-43	2.993622E-45
54*	1.090198E-45	4.97758E-45
55	1.444009E-38	3.682619E-45
56	1.519577E-39	9.647499E-45
57	2.365784E-45	2.099913E-45
58	1.499596E-42	5.321825E-45
59	5.983339E-44	3.595775E-45
60	9.804069E-41	5.019045E-45
61	4.593277E-45	1.468218E-45
62	5.288706E-45	1.379667E-45
63	3.154805E-41	4.233897E-45
64	7.79469E-45	4.447572E-45
65	3.676584E-41	4.777808E-45
66	2.245505E-40	1.902441E-45
67	3.537973E-45	2.529096E-45
68	3.103817E-44	1.93237E-45
69	1.648352E-41	3.524741E-45
70	1.008567E-43	2.854943E-45
71	4.598153E-42	2.660968E-45
72	4.178906E-42	2.635376E-43
73	5.979312E-43	2.515383E-45
74	9.620667E-45	2.887981E-45
75	4.641803E-43	2.169769E-45
76	1.920327E-44	3.515238E-45
77	3.222024E-45	2.654502E-45
78	5.128969E-43	2.804265E-45
79	1.848643E-40	1.788321E-45
80	4.573276E-43	3.682132E-45
81	1.248732E-42	3.344088E-45
82	7.664798E-42	2.877663E-45
83	1.340023E-44	2.332661E-45
84*	1.338772E-45	1.735667E-45
85	4.915736E-43	1.004047E-43
86	1.439643E-40	3.63191E-45
87	7.300998E-45	4.112071E-45
88	1.408163E-42	3.968023E-44
89	2.402119E-41	5.201882E-45
90	2.124046E-42	2.189246E-45
91	1.338194E-41	2.212755E-45
92	1.158261E-43	3.405918E-45
93	1.373065E-44	2.143973E-45
94	2.682875E-43	2.175463E-45
95	1.296078E-44	2.025773E-45
96	4.928713E-44	2.265348E-45
97	2.373405E-43	9.042362E-45
98	6.600024E-44	3.491189E-45
99	2.656884E-42	1.698268E-45
100	7.751215E-42	1.886067E-45
average	**4.016595E-40**	8.810760E-45

* BVSCV has posterior probability less than BVSBF.

4.9 Prediction performance

The elastic net is based on a combination of the ridge (L_2) and the lasso (L_1) penalties, so it does both parameter estimation and variable selection. The L_1 part of the elastic net performs variable selection and the L_2 part stabilizes the solution part. When $\alpha = 1$ ($\lambda_1 = 0$), the naïve elastic net becomes ridge regression. BVS performs only variable selection result. To study the prediction performance of the optimal model derived from BVS, ridge regression is used for parameter estimation of the BVS. Table 4.51 shows average of prediction error of the optimal model derived from BVS and elastic net for simulation dataset 1. The result reveals that BVSCV has prediction accuracy as well as BVSBF does. The ENETCV minimizes a cross-validated prediction error, so it has less prediction error. For ENETBF, the value of λ_2 is estimated by λ_2BF. The result in Chapter 4 Part I show that the method of λ_2BF performs the prediction performance better than cross-validation method, so ENETBF has the best prediction performance.

Table 4.51 Prediction performance of the optimal model derived from BVS and elastic net methods

Method	Prediction error (PE)
BVSCV	5.0643 (0.1220)
BVSBF	4.9506 (0.1174)
ENETCV	4.8143 (0.1263)
ENETBF	4.3742 (0.1029)

The average of PE is computed based on 100 datasets. The numbers in parenthesis are the corresponding standard errors (of PE) estimated using the bootstrap with B=500 resampling from the 100 PE's.

4.10 Real data examples

In this section, we apply two real datasets to illustrate the efficiency of BVSCV and BVSBF. The two datasets are the diabetes dataset and prostate cancer data used in elastic net literature and related methods.

4.10.1 Diabetes data ($p = 10, n = 442$)

The response variable (y) is a quantitative measure of disease progression one year after baseline for 442 diabetes patients. The dataset contains 10 baseline predictor variables: AGE, SEX, body mass index (BMI), average blood pressure (BP), and six blood serum measurements: tc(S1), ldl(S2), hdl(S3), tch(S4), ltg(S5), glu(S6). Appendix L shows the values of the Pearson correlation and the partial correlation coefficients for diabetes data. Table 4.52 shows variable selection results for diabetes data using BVSCV and BVSBF.

Table 4.52 Variable selection results for diabetes data using BVSCV and BVSBF

Diabetes data											Prediction error
Predictor variables [a]	BMI	S5	BP	S4	S3	S6	S1	AGE	S2	SEX	
λ_1 and λ_2 are estimated by the 10-fold cross-validation method (λ_1 = 0.0736 and λ_2 = 0.0736)											
BVSCV	BMI	S5	BP	S4	S3	S6	-	-	-	SEX	2901.647
ENETCV $\alpha = 0.5$	BMI	S5	BP	S4	S3	S6	S1	-	S2	SEX	2878.887
λ_2 is based on Bayes factor (λ_1 = 0.0128 and λ_2BF = 0.0128)											
BVSBF	BMI	S5	BP	S4	S3	S6	S1	AGE	S2	-	2970.086
ENETBF $\alpha = 0.5$	BMI	S5	BP	S4	S3	S6	S1	AGE	S2	SEX	2868.415

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,10$.

The result reveals that BVSCV performs group selection and variable selection efficiency better than elastic net. For BVSCV, the optimal model contains seven predictors: BMI, BP, S5, SEX, S4, S6, and S3.

Hasterberg, Choi, Meier and Fraley (2008) claimed that BMI and S5 appear to be most important whereas AGE, S1 and S2 enter the model relatively late. The predictors S1 and S2 have strong positive correlation, so these terms largely cancel out. The BVSCV performs the result of AGE, S1, and S2 similar to the result of Hasterberg, Choi, Meier and Fraley (2008).

The BVSCV performs group selection and variable selection efficiency better than BEN proposed by Li and Lin (2010). Li and Lin (2010) showed that the excluded predictors using their Bayesian elastic net (BEN) are as follows.

- Using the 95% credible interval criterion, the predictors AGE and S2 are excluded from the model.
- Using the scaled neighborhood criterion with probability threshold 0.5, the predictors AGE, S2 and S4 are excluded from the model.

The result of Li and Lin (2010) did not perform the group selection performance. The predictors S1 and S2 are highly correlated, so both S1 and S2 should be excluded from the model together.

For BVSBF, the predictor SEX is excluded from the optimal model. This differs from BVSCV. Maximum posterior probability of the optimal models derived from BVSCV and BVSBF are 1.46379E-246 and 2.563666E-263, respectively. The optimal model derived from BVSCV has posterior probability higher than the BVSBF does. The result reveals that the prior of the BVSCV is appropriate for the BVS method of diabetes data.

4.10.2 Prostate cancer data ($p = 8, n = 97$)

The response variable (y) is the logarithm of prostate specific antigen (lpsa). The predictor variables are eight clinical measures:

X1 - the logarithm of cancer volume (lcavol),

X2 - the logarithm of prostate weight (lweight),

X3 - age,

X4 - the logarithm of the amount of benign prostatic hyperplasia (lbph),

X5 - seminal vesicle invasion (svi),

X6 - the logarithm of capsular penetration (lcp),

X7 - the Gleason score (gleason), and

X8 - the percentage Gleason score 4 or 5 (pgg45).

Appendix L shows the values of the Pearson correlation and the partial correlation coefficients for prostate cancer data. Table 4.53 shows variable selection results for prostate cancer data using BVSCV and BVSBF.

Table 4.53 Variable selection results for prostate cancer data using BVSCV and BVSBF

Prostate cancer data									Prediction error
Predictor variables [a]	X1	X5	X6	X2	X8	X7	X4	X3	
λ_1 and λ_2 are estimated by the 10-fold cross-validation method (λ_1 = 0.0224 and λ_2 = 0.0224)									
BVSCV	X1	X5	X6	X2	X8	X7	X4	X3	0.4475
ENETCV $\alpha = 0.5$	X1	X5	-	X2	X8	X7	X4	X3	0.4555
λ_2 is based on Bayes factor (λ_1 = 0.0286 and λ_2BF = 0.0286)									
BVSBF	X1	X5	X6	X2	X8	X7	X4	X3	0.4475
ENETBF $\alpha = 0.5$	X1	X5	-	X2	X8	X7	X4	X3	0.4583

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots, 8$.

The result reveals that BVSCV and BVSBF perform variable selection efficiency differs from elastic net method. For BVSCV and BVSBF, all predictors are in the optimal model. This is the same as the result of naïve elastic net method of Zou and Hastie (2005). Li and Lin (2010) showed that the selected predictors using their Bayesian elastic net (BEN) is X1, X2, X4, X7, and X8. The variable selection efficiency of BVSCV and BVSBF differ from the BEN proposed by Li and Lin (2010). For prostate cancer data, we note that highly correlated predictors have close correlation coefficients to each other. This may be affected to all predictors are included in the optimal model of BVS. Maximum posterior probability of the optimal models derived from BVSCV and BVSBF are 2.792154E-99 and 8.20063E-99, respectively.

Chapter 5

Conclusions and Recommendations

In Chapter 5, there are four sections as follows.

5.1 The method for estimating the value of the penalty parameter λ_2 based on Bayes factor

5.2 Comparative study between two adaptive elastic net methods

5.3 Bayesian variable selection for elastic net linear regression model

5.4 Recommendations and Suggestions

5.1 The method for estimating the value of the penalty parameter λ_2 based on Bayes factor

Conclusions and Discussions

5.1.1 Prediction accuracy

The method for estimating the value of the penalty parameter λ_2 based on Bayes factor, λ_2BF, improves the prediction accuracy of the elastic net method. The method of λ_2BF has the prediction performance better than the value of λ_2 obtained from 10-fold cross-validation method (λ_2CV) when α is not close to zero. This is expected according to the L_2 part stabilizes the solution parts and improves the prediction. Using λ_2BF, the result reveals that the elastic net model is significance model as defined by Bayes factor ($\text{BF}_{elastic\ net} > 1$) when $\alpha \in (0,1)$.

As expect according to research hypothesis (3):

(3) Using the penalty parameter (λ_2) associated with Bayes factor, the derived elastic net model is significance model as defined by Bayes factor and this model has prediction error less than the elastic net model derived by using the value of λ_2 obtained from 10-fold cross-validation method does.

Using λ_2BF to fit elastic net model, the penalty parameter λ_1 is derived from the relationship between the shrinkage parameters, i.e., $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$. This may cause the value of λ_1 associated with λ_2BF becomes higher than the value of λ_1

derived from CV method when α is close to zero. It affected the prediction error derived from the λ_2BF is larger than the λ_2CV when α is small.

For adaptive elastic net methods, the λ_2BF has the prediction performance better than the λ_2CV does. The result reveals that the λ_2BF improves the prediction accuracy of the adaptive elastic net AENET2009 and AENET2011. This is expected according to the L_2 part of the adaptive elastic net offers the same kind of contribution as the naïve elastic net estimate. When α is small (α = 0.01, 0.05), the λ_2CV has the prediction performance better than the λ_2BF does. This prediction performance is the same as the result of the elastic net estimate. Thus, the method of λ_2BF can be applied for adaptive elastic net estimate.

5.1.2 Estimation accuracy

For elastic net and adaptive elastic net estimates, the result reveals that λ_2CV performs the estimation accuracy better than λ_2BF does. For some small value of α, the λ_2BF has the estimation accuracy better than the λ_2CV does. For example:

- For simulation dataset 1, the λ_2BF has the estimation accuracy better than the λ_2CV when $0.08 < \alpha < 0.6$.

- For simulation dataset 2, the λ_2BF has the estimation accuracy better than the λ_2CV when α = 0.02, 0.03.

The research hypothesis (1):

(1) Using the penalty parameter (λ_2) associated with Bayes factor, the elastic net and adaptive elastic net estimators have mean squared error less than those estimators derived by using the value of λ_2 obtained from 10-fold cross-validation method do.

The simulation result reveals that the λ_2BF has the estimation accuracy better than the λ_2CV for some small value of α. Hence, the research hypothesis (1) is true for some small value of α.

5.1.3 Variable selection performance

For elastic net and adaptive elastic net estimates, the simulation result reveals that the λ_2CV has the variable selection performance better than the λ_2BF does. For small α value (α is close to zero), the λ_2BF has the value C tends to the true value of C better than the λ_2CV does.

Research hypothesis (2):

(2) Using the penalty parameter (λ_2) associated with Bayes factor, the elastic net and adaptive elastic net linear regression models have the variable selection performance – the number of zero coefficients that are correctly estimated by zero (C) is maximum, whereas the number of nonzero coefficients that are incorrectly estimated by zero (IC) is minimum – better than the elastic net and adaptive elastic net models derived by using the value of λ_2 obtained from 10-fold cross-validation method do.

Hence, this research hypothesis is true when α is close to zero.

Elastic net does both parameter estimation and variable selection. The elastic net is based on a combination of the ridge (L_2) and the lasso (L_1) penalties. The L_1 part of the elastic net performs automatic variable selection, while the L_2 part stabilizes the solution parts and, hence, improves the prediction.

The λ_2BF is the penalty parameter of the L_2 part of the elastic net method; nevertheless, the λ_2BF performs the variable selection performance better than the λ_2CV when α is close to zero. For some small value of α, the λ_2BF has the estimation accuracy better less than λ_2CV does.

In this thesis, we propose the λ_2BF which is the penalty parameter of the L_2 part of the elastic net. The method of λ_2BF has the prediction performance better than λ_2CV does. But the λ_2BF performs the estimation accuracy and the variable selection performance worse than λ_2CV does. This due to the elastic net is based on a combination of the L_2 and the L_1 penalties. The appropriate combination of λ_1 and λ_2

can make the elastic net perform best in the prediction performance, the estimation accuracy and the variable selection performance. For example:

For simulation dataset 1, the λ_2BF has the prediction performance and the estimation accuracy better than λ_2CV when $0.2 < \alpha < 0.6$. The λ_2BF has the variable selection performance close to λ_2CV when $\alpha = 0.2$.

5.2 Comparative study between two adaptive elastic net methods

Conclusions and Discussions

Zou and Zhang (2009) proposed the adaptive elastic net using the elastic net estimator to construct the adaptive weight (AENET2009). Ghosh (2011) proposed the adaptive elastic net using the least squares estimator to construct the adaptive weight (AENET2011). The comparative study between the AENET2009 and AENET2011 is summarized as follows.

The adaptive elastic net method is suitable for high dimensional data analysis. The AENET2009 and AENET2011 incorporate the adaptive weight in the L_1 penalty of the naïve elastic net estimate. The L_1 penalty is responsible for the sparsity of the estimator. The AENET2009 performs parameter estimation and variable selection efficiency better than the elastic net and AENET2011 do. The AENET2009 has the best performance when the sample size is large (n = 400). Using the adaptive weights RENETCV09 or RENETBF09, the AENET2009 has the prediction accuracy as well as the AENET2011 does. Hence, the AENET2009 where the adaptive weight is constructed using rescaled elastic net with α = 0.9 performs the estimation accuracy, the variable selection performance, and the prediction performance best.

5.3 Bayesian variable selection for elastic net linear regression model (BVS)

In this thesis, we study the BVS with different priors as follows:

(1) The Bayesian variable selection for elastic net linear regression model where the penalty parameters λ_1 and λ_2 are estimated by the 10-fold cross-validation method (BVSCV).

(2) The Bayesian variable selection for elastic net linear regression model where the penalty parameter λ_2 is based on Bayes factor (BVSBF). For BVSBF, the penalty parameter λ_1 is derived from the relationship between λ_1 and λ_2: $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$.

For the process of Bayesian variable selection for elastic net linear regression model studied in this thesis, we set $\alpha = 0.5$, i.e., $\lambda_1 = \lambda_2$.

Conclusions and Discussions

Process of BVS differs from ordinary elastic net. Elastic net is a penalized regression method which does both parameter estimation and variable selection. The elastic net performs variable selection when some parameter estimation is shrunk to zero. The BVS method performs variable selection where the optimal model is the model with highest posterior probability. For BVS, the predictors which are highly correlated with the response variable are included into the optimal model. The result reveals that the BVS has the number of predictors in the optimal model less than elastic net does. The BVS performs group selection where the pair of predictors which are highly correlated with the response variable is included into the optimal model. For elastic net, the prior distribution of $\boldsymbol{\beta}$ is a compromise between normal and double exponential distribution. The prior distribution of $\boldsymbol{\beta}$ which is used in BVS is transformed into normal distribution. Hence, the variable selection result of BVS differs from elastic net.

The variable selection result of BVSCV differs from BVSBF. The BVSCV and BVSBF have different priors for the penalty parameters. The BVSCV uses the penalty parameters λ_1 and λ_2 estimated by the 10-fold CV method whereas the BVSBF uses the penalty parameter λ_2 based on Bayes factor (λ_2BF). For BVS method, the prior for the penalty parameters λ_1 and λ_2 estimated by the 10-fold CV method is the best since the optimal model derived from BVSCV has posterior probability more than the BVSBF does.

Research hypothesis (4):

(4) For linear regression analysis of microarray classification and gene selection where the prior distribution of $\boldsymbol{\beta}$ is a compromise between normal and double exponential distribution, the combination of Bayesian variable selection and elastic net model performs the variable selection efficiency – the optimal model exhibits group selection – better than the classical elastic net model does.

This research hypothesis is true. The BVS performs both variable selection and group selection where the pair of predictors which are highly correlated with the response variable is included into the optimal model whereas some pair of predictors which are highly correlated with the response variable is excluded from the elastic net model. For BVS, the variable selection result derived from the optimal model with highest posterior probability whereas the elastic net performs variable selection when some parameter estimation is shrunk to zero. This may cause the elastic net and BVS perform different variable selection result.

5.4 Recommendations and Suggestions

The method for estimating the value of the penalty parameter λ_2 based on Bayes factor, λ_2BF, improves the prediction accuracy of the elastic net estimate. In this thesis, we study the performance of λ_2BF with variety values of α. Using λ_2BF, the result reveals that the elastic net model is significance model as defined by Bayes factor ($\mathrm{BF}_{elastic\ net} > 1$) when $\alpha \in (0,1)$. The λ_2BF has the prediction performance better than λ_2CV when α is not close to zero. The λ_2BF can be applied to different dataset where the number of parameters (p) less than the sample size (n), e.g. small p (p = 15) or the cases where the number of parameters diverges with the sample size. The method of λ_2BF can be used for adaptive elastic net estimate where the adaptive weight is included in the L_1 penalty. The λ_2BF improves the prediction accuracy of the adaptive elastic net AENET2009 and AENET2011 where theirs adaptive weights are different. In this thesis, the method of λ_2BF is based on the linear regression model. Friedman, Hastie and Tibshirani (2010) developed the cyclical coordinate descent algorithms for fitting generalized linear models with elastic net penalties. The extensions of the λ_2BF approach for generalized linear models (e.g. logistic regression, multinomial regression) are interesting for future research. Another research direction is analysis the performance of the λ_2BF for fitting the elastic net estimator for censored data proposed by Johnson (2009).

This thesis provides Bayesian variable selection for elastic net linear regression model as an alternative method for variable selection in linear regression analysis of microarray classification and genes selection when the prior distribution of $\boldsymbol{\beta}$ is a compromise between normal and double exponential distribution. For Bayesian variable selection for elastic net linear regression model (BVS), the prior for the penalty parameters λ_1 and λ_2 estimated by the 10-fold CV method is the best. The BVS is more parsimonious than the elastic net. Thus, the prior for the penalty parameters λ_1 and λ_2 with $\alpha \to 0$ ($\lambda_1 > \lambda_2$) affects the optimal model derived from BVS has the number of predictors close to the lasso. In this thesis, we set α = 0.5 for prior of the penalty parameters λ_1 and λ_2 used in the BVS. The other values of α are interesting for analysis the performance of the BVS method. For the model space prior, the dilution prior proposed by George (2010) is interesting for analysis the performance of the BVS method.

In this thesis, the prior for σ^2 is inverse gamma distribution. The other choice is a noninformative prior $p(\sigma^2) \propto 1/\sigma^2$, and Gibbs sampling method can be used to search for the model having highest posterior probability rather than compute the entire posterior probability.

References

Books

Ando, T. (2010). *Bayesian model selection and statistical modeling*. New York: CRC Press.

Box, G.E.P., & Tiao, G.C. (1973). *Bayesian inference in statistical analysis*. California: Addison-Wesley.

Birkes, D., & Dodge, Y. (1993). *Alternative method of regression*. New York: John Wiley & Sons.

Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (3rd ed.). New York: John Wiley & Sons.

Gamerman, D., & Lopes, H.F. (2006). *Markov chain Monte Carlo: Stochastic simulation for Bayesian inference* (2nd ed.). New York: CRC Press.

Gelman, A., Carlin, J.B., Stern, H.S., & Rubin, D.B. (2003). *Bayesian data analysis* (2nd ed.). New York: CRC Press.

Gruber, M. H. J. (1990). *Regression estimators: A comparative study*. Boston: Academic Press.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). New York: Springer.

Jeffreys, H. (1983). *Theory of probability* (3rd ed.). New York: Oxford University Press.

Lehmann, E. L. (1983). *Theory of point estimation*. Pacific Grove, CA: Wadsworth and Brooks/Cole.

Marin, J.-M. (2007). *Bayesian core: A practical approach to computational Bayesian statistics*. New York: Springer.

Miller, A. (2002). *Subset selection in regression* (2nd ed.). New York: CRC Press.

Ntzoufras, I. (2009). *Bayesian modeling using WinBUGS*. New Jersey: John Wiley & Sons.

Robert, C. P. (2001). *The Bayesian choice* (2nd ed.). New York: Springer.

Whittaker, J. (1990). *Graphical models in applied multivariate statistics*. Chichester, United Kingdom: John Wiley & Sons.

Articles

Andrews, D. F., & Mallows, C. L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society. Series B, 36*, 99-102.

Barbieri, M., & Berger, J. O. (2004). Optimal predictive variable selection. *The annals of Statistics, 32*, 870-897.

Bartlett, M. S. (1957). A comment on D.V. Lindley’s statistical paradox. *Biometrika, 44*, 533-534.

Berger, J. O. (1999). Bayes factors. In the *Encyclopedia of Statistical Sciences*, (Update Vol. 3, pp.20-29). New York: Wiley.

Berger, J. O., Ghosh, J. K., & Mukhopadhyay, N. (2003). Approximations and consistency of Bayes factors as model dimension grows. *Journal of Statistical Planning and Inference, 112*, 241-258.

Berger, J. O., & Pericchi, L. R. (2001). Objective Bayesian methods for model selection: Introduction and comparison (with discussion). In *Model Selection*, ed. P. Lahiri, Beachwood, OH: IMS Lecture Notes-Monograph Series, *38*, 135-207.

Brown, P. J., Vannucci, M., & Fearn, T. (1998). Multivariate Bayesian variable selection and prediction. *Journal of the Royal Statistical Society. Series B*, *60*, 627-641.

Brown, P. J., Vannucci, M., & Fearn, T. (2002). Bayesian model averaging with selection of regressors. *Journal of the Royal Statistical Society. Series B, 64*, 519-536.

Casella, G. (2001). Empirical Bayes Gibbs sampling. *Biostatistics*, *2*, 485-500.

Celeux, G., Anbari, M. E., Marin, J.-M., & Robert, C. P. (2012). Regularization in regression: Comparing Bayesian and frequentist methods in a poorly informative situation. *Bayesian Analysis*, *7*(2), 477-502.

Chipman, H. A., George, E. I., & McCulloch, R. E. (2001). The practical implementation of Bayesian model selection (with discussion). In *Model Selection*, ed. P. Lahiri, Beachwood, OH: IMS Lecture Notes-Monograph Series, *38*, 65–134.

Clyde, M. (2001). Discussion of “The practical implementation of Bayesian model selection" by H.A. Chipman, E.I. George, and R.E. McCulloch. *Model Selection*, ed. P. Lahiri, Beachwood, OH: IMS Lecture Notes-Monograph Series, *38*, 117-124.

Clyde, M., & George, E. I. (2004a). Flexible empirical Bayes estimation for wavelets. *Journal of the Royal Statistical Society. Series B, 62,* 681-698.

Clyde, M., & George, E. I. (2004b). Model uncertainty. *Statistical Science, 19*, 81-94.

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

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

Fernandez, C., Ley, E., & Steel, M. F. J. (2001). Benchmark priors for Bayesian model averaging. *Journal of Econometrics, 100,* 381-427.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), 1-22.

Fu, W. J. (1998). Penalized regressions: The bridge versus the lasso. *Journal of Computational and Graphical Statistics, 7,* 397-416.

George, E. I. (1999). Bayesian model selection. In *Encyclopedia of Statistical Sciences* (Update Vol. 3, pp.39-46). New York: Wiley.

George, E. I. (2000). The variable selection problem. *Journal of the American Statistical Association, 95,* 1304-1308.

George, E. I. (2010). Dilution priors: Compensating for model space redundancy. *IMS Collections Borrowing Strength: Theory Powering Applications – A Festschrift for Lawrence D. Brown, 6,* 158-165.
DOI: 10.1214/10-IMSCOLL611

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

George, E. I., & McCulloch, R. E. (1997). Approaches for Bayesian variable selection. *Statistica Sinica, 7,* 339-373.

Ghosh, S. (2011). On the grouped selection and model complexity of the adaptive elastic net. *Statistics and Computing, 21*, 451-462.

Good, I. J. (1958). Significance tests in parallel and in series. *Journal of the American Statistical Association, 53*, 799-813.

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

Hans, C. (2010). Model uncertainty and variable selection in Bayesian lasso regression. *Statistics and Computing, 20*, 221-229.

Hans, C. (2011). Elastic net regression modeling with the Orthant normal prior. *Journal of the American Statistical Association, 106*, 1383-1393.

Hesterberg, T., Choi, N. H., Meier, L., & Fraley, C. (2008). Least angle and ℓ_1 penalized regression: A review. *Statistics Surveys, 2*, 61-93.

Hoerl, A. E., & Kennard, R. W. (1970a). Ridge regression: Applications to nonorthogonal problems. *Technometrics, 12*(1), 69-82.

Hoerl, A. E., & Kennard, R. W. (1970b). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*(1), 55-67.

Hoerl, A. E., Kennard, R. W., & Baldwin, K. F. (1975). Ridge regression, some simulations. *Communications in Statistics, A4*, 105-123.

Johnson, B.A. (2009). On lasso for censored data. *Electronic Journal of Statistics, 3*, 485-506.

Kass, R. E. (1991). About Theory of Probability. *Chance, 4*, 13.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*, 773-795.

Kyung, M., Gill, J., Ghosh, M., & Casella, G. (2010). Penalized regression, standard errors, and Bayesian lassos. *Bayesian Analysis, 5*(2), 369-412.

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

Lindley, D. V. (1957). A statistical paradox. *Biometrika, 44*, 187-192.

Lykou, A., & Ntzoufras, I. (2012). On Bayesian lasso variable selection and the specification of the shrinkage parameter. *Statistics and Computing*, (published online: 2012, February 14).

O'Hara, R. B., & Sillanpää, M. J. (2009). A review of Bayesian variable selection methods: What, how and which. *Bayesian Analysis, 4*(1), 85-118.

Osborne, M. R., Presnell, B., & Turlach, B. A. (2000). On the lasso and its dual. *Journal of Computational and Graphical Statistics, 9*, 319-337.

Park, T., & Casella, G. (2008). The Bayesian lasso. *Journal of the American*

Statistical Association, 103, 681-686.

Pauler, D. K. (1998). The Schwarz criterion and related methods for normal linear models. *Biometrika, 85*, 13-27.

Pauler, D. K., Wakefield, J. C., & Kass, R.E. (1999). Bayes factors and approximations for variance component models. *Journal of the American Statistical Association, 94,* 1242-1253.

Raftery, A. E. (1999). Bayes factors and BIC: Comment on “A critique of the Bayesian Information Criterion for model selection”. *Sociological Methods & Research*, *27*(3), 411-427.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*, 461-464.

Shafer, G. (1982). Lindley’s paradox. *Journal of the American Statistical Association, 77*, 325-351.

Shao, J. (1997). An asymptotic theory for linear model selection. *Statistica Sinica, 7,* 221-264.

Shibata, R. (1981). An optimal selection of regression variables. *Biometrika*, 68, 45-54.

Smith, M. & Kohn, R. (1996). Nonparametric regression using Bayesian variable selection. *Journal of Econometrics, 75,* 317-344.

Stamey, T. A., Kabalin, J. N., Mcneal, J. E., Johnstone, I. M., Freiha, F., Redwine, E. A., & Yang, N. (1989). Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate: II. Radical prostatectomy treated patients, *Journal of Urology, 141*(5), 1076-1083.

Stone, M. (1979). Comments on model selection criteria of Akaike and Schwarz. *Journal of the Royal Statistical Society. Series B, 41*, 267-278.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B, 58,* 267-288.

Yang, Y., & Zou, H. (2012). An efficient algorithm for computing the HHSVM and its generalizations, *Journal of Computational and Graphical Statistics*, 22, 396-415.

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

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B, 67,* 301-320.

Zou, H., & Zhang, H. H. (2009). On the adaptive elastic-net with a diverging number of parameters. *The Annals of Statistics, 37*, 1733-1751.

Zou, H., Hastie, T., & Tibshirani, R. (2007). On the degrees of freedom of the lasso. *The Annals of Statistics, 35,* 2173-2192.

Other material

Barbieri, M., & Berger, J. O. (2001). *Optimal predictive variable selection*. ISDS Discussion Paper, Duke University.

Berger, J. O., Ghosh, J. K., & Mukhopadhyay, N. (1999). *Approximations and consistency of Bayes factors as model dimension grows*. Technical report, Purdue University, Department of Statistics.

Bornn, L., Gottardo, R., & Doucet, A. (2010). *Grouping priors and the Bayesian elastic net*. Technical report, University of British Columbia, Department of Statistics.

Hyndman, R. J. (2010, October 4). Why every statistician should know about cross-validation [Web blog message]. Retrieved February 21, 2013, from http://robjhydman.com/hyndsight/crossvalidation/.

MATLAB & Simulink., Lasso and Elastic Net. Retrieved February 13, 2013, from http://www.mathworks.com/help/stats/lasso-and-elastic-net.html

MATLAB & Simulink., Regularized least-squares regression using lasso or elastic net algorithm. Retrieved February 13, 2013, from http://www.mathworks.com/help/stats/lasso.html

R Development Core Team. (2013). R: *A language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.r-project.org/

Yang, Y., and Zou, H. (2013). gcdnet: A generalized coordinate descent (GCD) algorithm for computing the solution path of the hybrid Huberized support vector machine (HHSVM) and its generalization. R package version 1.0.3, URL: http://code.google.com/p/gcdnet/

Appendix A

Oracle properties of the adaptive elastic net

In the language of Fan and Li (2001), let $\mathcal{A} = \{j : \beta_j \neq 0\}$ denote the true model set and $|\mathcal{A}| = p_0 < p$. Let $\widehat{\boldsymbol{\beta}}(\delta)$ is the coefficient of estimator produced by fitting a procedure δ. The procedure δ is called an oracle procedure if $\widehat{\boldsymbol{\beta}}(\delta)$ has following properties asymptotically (i.e. as $n \to \infty$):

- Identifies the correct subset model, $\{j : \hat{\beta}_j(\delta) \neq 0\} = \mathcal{A}$
- Has the optimal estimation rate, $\sqrt{n}\left[\widehat{\boldsymbol{\beta}}(\delta)_{\mathcal{A}} - \boldsymbol{\beta}_{\mathcal{A}}\right] \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is the covariance matrix knowing the true covariates subset.

As pointed out by Fan and Li (2001), the oracle properties are closely related to the super-efficiency phenomenon (Lehman, 1983, p.405).

A.1 Adaptive elastic net (Zou & Zhang, 2009)

$$\widehat{\boldsymbol{\beta}}_{\text{AdaEnet2009}} = (1+\lambda_2)\left\{\arg\min_{\boldsymbol{\beta}} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda_2\|\boldsymbol{\beta}\|_2^2 + \lambda_1^* \textstyle\sum_{j=1}^{p} \hat{w}_j |\beta_j|\right\}.$$

The adaptive weight $\hat{w}_j = \left(|\hat{\beta}_j(\text{elastic net})|\right)^{-\gamma}, j = 1, \ldots, p,$ where γ is a positive constant.

The adaptive elastic net has the oracle property; that is, the estimator $\widehat{\boldsymbol{\beta}}_{\text{AdaEnet2009}}$ must satisfy:

1. *Consistency in selection*: $\Pr\left(\left\{j : \widehat{\boldsymbol{\beta}}_{(\text{AdaEnet2009})_j} \neq 0\right\} = \mathcal{A}\right) \to 1$,
   where $\mathcal{A} = \{j : \beta_j \neq 0, j = 1,2, \ldots, p\}$

2. *Asymptotic normality*: $\boldsymbol{\alpha}^T \frac{\mathbf{I} + \lambda_2 \boldsymbol{\Sigma}_{\mathcal{A}}^{-1}}{1+\lambda_2/n}\left(\widehat{\boldsymbol{\beta}}_{(\text{AdaEnet2009})_{\mathcal{A}}} - \boldsymbol{\beta}_{\mathcal{A}}\right) \xrightarrow{d} N(0, \sigma^2)$,
   where $\boldsymbol{\Sigma}_{\mathcal{A}} = \mathbf{X}_{\mathcal{A}}^T \mathbf{X}_{\mathcal{A}}$ and $\boldsymbol{\alpha}$ is a vector of norm 1.

A.2 Adaptive elastic net (Ghosh, 2011)

$\widehat{\boldsymbol{\beta}}_{\text{AdaEnet2011}}$

$$= (1+\lambda_2)\left\{\arg\min_{\boldsymbol{\beta}} \|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda_2 \sum_{j=1}^{p}\left(|\beta_j| + \frac{\lambda_1}{2\lambda_2}\widehat{w}_j\right)^2 - \frac{(\lambda_1)^2}{4\lambda_2}\sum_{j=1}^{p}\widehat{w}_j^2\right\}.$$

The adaptive weight vector $\widehat{\mathbf{w}} = 1/\left|\widehat{\boldsymbol{\beta}}_{least\ square}\right|^{\gamma}$, for some $\gamma > 0$.

Assume that $\mathcal{A} = \{j :\ \beta_j \neq 0, j = 1, \dots, p\}$, and $|\mathcal{A}| = p_0 < p$.

Suppose that $\lambda_1/\sqrt{n} \to 0$, $\lambda_1 n^{(\gamma-1)/2} \to \infty$ and $\lambda_2/n \to 0$. Then the adaptive elastic net estimate $\widehat{\boldsymbol{\beta}}_{\text{AdaEnet2011}}$ must satisfy the following:

1. *Consistency in selection*: $\lim_{n\to\infty} P(\mathcal{A}^* = \mathcal{A}) = 1$ where $\mathcal{A}^* = \{j :\ \widehat{\boldsymbol{\beta}}_j \neq 0\}$.

2. *Asymptotic normality*: $\sqrt{n}\left(\widehat{\boldsymbol{\beta}}_{\mathcal{A}} - \boldsymbol{\beta}_{\mathcal{A}}\right) \xrightarrow{d} N(\mathbf{0}, \mathbf{C}_{11}^{-1})$ where $\mathbf{C}_{11}$ is a $p_0 \times p_0$ matrix.

Appendix B

Example of Hypotheses for Bayes Factor

The Bayes factor is used in the process for specification the penalty parameter (λ_2). This appendix shows an example of the hypotheses used for computing Bayes factor when $p = 3$. There are 12 hypotheses and the 12 Bayes factors are computed under their corresponding hypotheses.

Hypotheses (1):

H_0: M_R is a constant model where $\mathbf{y}|\boldsymbol{\beta}, \sigma^2, M_R \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$.

H_1: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}$

Hypotheses (2):

H_0: M_R is a constant model where $\mathbf{y}|\boldsymbol{\beta}, \sigma^2, M_R \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$.

H_1: $\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$

Hypotheses (3):

H_0: M_R is a constant model where $\mathbf{y}|\boldsymbol{\beta}, \sigma^2, M_R \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$.

H_1: $\mathbf{y} = \mathbf{X}_3\boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}$

Hypotheses (4):

H_0: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}$

H_1: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$

Hypotheses (5):

H_0: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}$

H_1: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_3\boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}$

Hypotheses (6):

H_0: $\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$

H_1: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$

Hypotheses (7):

H_0: $\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$

H_1: $\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{X}_3\boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}$

Hypotheses (8):

H_0: $\mathbf{y} = \mathbf{X}_3\boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}$

H_1: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_3\boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}$

Hypotheses (9):

H_0: $\mathbf{y} = \mathbf{X}_3\boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}$

H_1: $\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{X}_3\boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}$

Hypotheses (10):

H_0: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$

H_1: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{X}_3\boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}$

Hypotheses (11):

H_0: $\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{X}_3\boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}$

H_1: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{X}_3\boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}$

Hypotheses (12):

H_0: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_3\boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}$

H_1: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{X}_3\boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}$

Appendix C

Mixture of normal distribution

The Laplace (double exponential) distribution is a scale mixture of a normal distribution with an exponential mixing density (Andrews & Mallows, 1974; cited by Kyung, Gill, Ghosh, & Casella, 2010), that is

$$\frac{a}{2}exp(-a|z|) = \int_0^\infty \frac{1}{\sqrt{2\pi s}} exp\left(-\frac{z^2}{2s}\right)\frac{a^2}{2} exp\left(-\frac{a^2}{2}s\right) ds. \qquad \text{(C.1)}$$

This idea is to introduce the appropriate latent parameters (Kyung, Gill, Ghosh, & Casella, 2010).

Proof.

Andrews & Mallows (1974, p.100) provided the identity

$$\int_0^\infty exp\left[-\frac{1}{2}(a^2u^2 + b^2u^{-2})\right] du = (\pi/2a^2)^{1/2} exp(-|ab|).$$

Then,

$$exp(-|ab|) = (\pi/2a^2)^{-1/2}\int_0^\infty exp\left[-\frac{1}{2}(a^2u^2 + b^2u^{-2})\right] du$$

$$exp(-|ab|) = \left(\frac{2a^2}{\pi}\right)^{1/2}\int_0^\infty exp\left[-\frac{1}{2}(a^2u^2 + b^2u^{-2})\right] du$$

$$\frac{a}{2}exp(-|ab|) = \frac{a}{2}\left(\frac{2a^2}{\pi}\right)^{1/2}\int_0^\infty exp\left[-\frac{1}{2}(a^2u^2 + b^2u^{-2})\right] du$$

$$\frac{a}{2}exp(-|ab|) = \frac{a^2}{\sqrt{2\pi}}\int_0^\infty exp\left[-\frac{1}{2}(a^2u^2 + b^2u^{-2})\right] du$$

$$\frac{a}{2}exp(-|ab|) = \frac{a^2}{\sqrt{2\pi}}\int_0^\infty exp\left(-\frac{a^2u^2}{2}\right)exp\left(-\frac{b^2u^{-2}}{2}\right)du.$$

Let $u = \sqrt{s}$ and $b = z$, so

$$\frac{a}{2}exp(-|az|) = \frac{a^2}{\sqrt{2\pi}}\int_0^\infty exp\left(-\frac{a^2s}{2}\right)exp\left(-\frac{z^2}{2s}\right)d(\sqrt{s})$$

$$= \frac{a^2}{\sqrt{2\pi}}\int_0^\infty \frac{1}{2\sqrt{s}}exp\left(-\frac{a^2s}{2}\right)exp\left(-\frac{z^2}{2s}\right)ds$$

$$= \frac{a^2}{2}\int_0^\infty \frac{1}{\sqrt{2\pi s}}exp\left(-\frac{a^2s}{2}\right)exp\left(-\frac{z^2}{2s}\right)ds$$

$$= \int_0^\infty \frac{1}{\sqrt{2\pi s}}exp\left(-\frac{z^2}{2s}\right)\frac{a^2}{2}exp\left(-\frac{a^2}{2}s\right)ds \;;\; a > 0. \qquad \blacksquare$$

Appendix D

Prior distribution for $\boldsymbol{\beta}$ of the elastic net linear regression model

Kyung, Gill, Ghosh, and Casella (2010) proposed hierarchical model prior for elastic net. By introducing latent variables $\tau_1^2, \dots, \tau_p^2$, using the basic identity (C.1) in Appendix C, this prior can be written as a normal mixture of gammas,

$$f\left(\boldsymbol{\beta}\middle|\sigma^2, \tau_1^2, \dots, \tau_p^2\right) \propto$$

$$\prod_{j=1}^{p} \frac{\sqrt{\lambda_2}}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{\lambda_2}{2\sigma^2}\beta_j^2\right] \times \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}} exp\left[-\frac{\beta_j^2}{2\sigma^2\tau_j^2}\right]\frac{\lambda_1^2}{2} exp\left[-\frac{\lambda_1^2\tau_j^2}{2}\right] d\tau_j^2.$$

Thus $\beta_j \sim N\left(0, \sigma^2\left(\tau_j^{-2} + \lambda_2\right)^{-1}\right), j = 1, \dots, p.$ (D.1)

Proof.

Zou and Hastie (2005) pointed out that, solving the elastic net problem is equivalent to find the marginal posterior mode of $\boldsymbol{\beta}|\mathbf{y}$ when the prior distribution of $\boldsymbol{\beta}$ is given by

$$f(\boldsymbol{\beta}|\sigma^2) \propto exp\left\{-\frac{\lambda_1}{\sigma}\sum_{j=1}^{p}\left|\beta_j\right| - \frac{\lambda_2}{2\sigma^2}\sum_{j=1}^{p}\beta_j^2\right\}, \tag{D.2}$$

a compromise between the Gaussian (normal) and Laplace (double exponential) priors.

The equation (D.2) can be written as

$$f(\boldsymbol{\beta}|\sigma^2)$$

$$= (2\pi\sigma^2)^{-p/2}(\lambda_2)^{p/2} exp\left(-\frac{\lambda_2}{2\sigma^2}\sum_{j=1}^{p}\beta_j^2\right) \times \left(2\sqrt{\sigma^2}\right)^{-p}(\lambda_1)^p exp\left(-\frac{\lambda_1}{\sigma}\sum_{j=1}^{p}\left|\beta_j\right|\right)$$

$f(\boldsymbol{\beta}|\sigma^2)$

$$= \prod_{j=1}^{p}\left[\frac{\sqrt{\lambda_2}}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{\lambda_2}{2\sigma^2}\sum_{j=1}^{p}\beta_j^2\right) \times \frac{\lambda_1}{2\sigma} exp\left(-\frac{\lambda_1}{\sigma}\sum_{j=1}^{p}|\beta_j|\right)\right] \quad \text{(D.3)}$$

Thus,

$$f(\beta_j|\sigma^2) = \frac{\sqrt{\lambda_2}}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{\lambda_2}{2\sigma^2}\beta_j^2\right) \times \frac{\lambda_1}{2\sigma} exp\left(-\frac{\lambda_1}{\sigma}|\beta_j|\right) \quad \text{(D.4)}$$

Let $a = \frac{\lambda_1}{\sigma}$, $z = \beta_j$, $s = \sigma^2\tau_j^2$ and using the basic identity (C.1) in Appendix C, the second term in (D.4) can be written as

$$\begin{aligned}
\frac{\lambda_1}{2\sigma} exp\left(-\frac{\lambda_1}{\sigma}|\beta_j|\right) &= \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}} exp\left[-\frac{\beta_j^2}{2\sigma^2\tau_j^2}\right] \frac{\lambda_1^2}{2\sigma^2} exp\left[-\frac{\lambda_1^2}{2\sigma^2}\sigma^2\tau_j^2\right] d(\sigma^2\tau_j^2) \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}} exp\left[-\frac{\beta_j^2}{2\sigma^2\tau_j^2}\right] \frac{\lambda_1^2}{2\sigma^2} exp\left[-\frac{\lambda_1^2\tau_j^2}{2}\right] \sigma^2 d\tau_j^2 \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}} exp\left[-\frac{\beta_j^2}{2\sigma^2\tau_j^2}\right] \frac{\lambda_1^2}{2} exp\left[-\frac{\lambda_1^2\tau_j^2}{2}\right] d\tau_j^2 \qquad \text{(D.5)} \\
&= \int_0^\infty N(\beta_j|0,\sigma^2\tau_j^2)\ \frac{\lambda_1^2}{2} exp\left[-\frac{\lambda_1^2\tau_j^2}{2}\right] d\tau_j^2
\end{aligned}$$

The equation (D.5) shows that the double exponential distribution is a scale mixture of a normal distribution with an exponential mixing density. This concept was used by Park and Casella (2008), and Kyung, Gill, Ghosh, and casella (2010).

By substituting (D.5) into (D.4), we can write equation (D.4) as

$$\begin{aligned}
f(\beta_j|\sigma^2) &= \frac{\sqrt{\lambda_2}}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{\lambda_2}{2\sigma^2}\beta_j^2\right) \times \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}} exp\left[-\frac{\beta_j^2}{2\sigma^2\tau_j^2}\right] \frac{\lambda_1^2}{2} exp\left[-\frac{\lambda_1^2\tau_j^2}{2}\right] d\tau_j^2 \\
&= \frac{1}{\sqrt{2\pi\sigma^2\lambda_2^{-1}}} \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}} exp\left[-\frac{1}{2\sigma^2}\left(\lambda_2 + \frac{1}{\tau_j^2}\right)\beta_j^2\right] \frac{\lambda_1^2}{2} exp\left[-\frac{\lambda_1^2\tau_j^2}{2}\right] d\tau_j^2 \\
&= \frac{1}{\sqrt{2\pi\sigma^2\lambda_2^{-1}}} \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}} exp\left[-\frac{\beta_j^2}{2\sigma^2\left(\lambda_2+\tau_j^{-2}\right)^{-1}}\right] \frac{\lambda_1^2}{2} exp\left[-\frac{\lambda_1^2\tau_j^2}{2}\right] d\tau_j^2
\end{aligned}$$

$$f\left(\beta_j|\sigma^2\right)=\frac{1}{\sqrt{2\pi\sigma^2\lambda_2^{-1}}}\int_0^\infty \frac{1}{\sqrt{\tau_j^2}}\frac{\left(\lambda_2+\tau_j^{-2}\right)^{-1/2}}{\sqrt{2\pi\sigma^2\left(\lambda_2+\tau_j^{-2}\right)^{-1}}}exp\left[-\frac{\beta_j^2}{2\sigma^2\left(\lambda_2+\tau_j^{-2}\right)^{-1}}\right]\frac{\lambda_1^2}{2}exp\left[-\frac{\lambda_1^2\tau_j^2}{2}\right]d\tau_j^2$$

$$=\frac{1}{\sqrt{2\pi\sigma^2\lambda_2^{-1}}}\int_0^\infty \frac{\left(\lambda_2+\tau_j^{-2}\right)^{-1/2}}{\sqrt{\tau_j^2}}N\left(\beta_j|0,\sigma^2\left(\lambda_2+\tau_j^{-2}\right)^{-1}\right)\frac{\lambda_1^2}{2}exp\left[-\frac{\lambda_1^2\tau_j^2}{2}\right]d\tau_j^2$$

(D.6)

Hence, Kyung, Gill, Ghosh, and casella (2010) suggested the prior of β_j conditional on $\tau_j^2,\ \beta_j \sim N\left(0,\sigma^2\left(\tau_j^{-2}+\lambda_2\right)^{-1}\right), j=1,\ldots,p.$ ■

Kyung, Gill, Ghosh, and casella (2010) gave the hierarchical prior of $\boldsymbol{\beta}$ as

$\boldsymbol{\beta}|\sigma^2,\mathbf{D}_\tau \sim N_p(\mathbf{0},\sigma^2\mathbf{D}_\tau)$,

where $\mathbf{D}_\tau$ is a diagonal matrix with diagonal elements $\left(\tau_j^{-2}+\lambda_2\right)^{-1}$,

$\tau_1^2,\ldots,\tau_p^2 \sim \prod_{j=1}^{p}\frac{\lambda_1^2}{2}e^{-\lambda_1^2\tau_j^2/2}d\tau_j^2,\ \tau_1^2,\ldots,\tau_p^2>0.$

From (D.6), we see that

$$f\left(\tau_j^2\right)=\frac{1}{\Gamma(1)}\left(\frac{\lambda_1^2}{2}\right)\left(\tau_j^2\right)^{1-1}exp\left(-\frac{\lambda_1^2\tau_j^2}{2}\right)=\frac{\lambda_1^2}{2}exp\left[-\frac{\lambda_1^2\tau_j^2}{2}\right].$$

Hence, the prior distribution of τ_j^2 is $\tau_j^2\sim Gamma\left(1,\frac{\lambda_1^2}{2}\right)$ or $\tau_j^2\sim Exponential\left(\frac{\lambda_1^2}{2}\right)$.

Appendix E

Standardization and correlation transformation

E.1 Standardization of the response and the predictor variables

The standardization of the response variable is

$$\frac{y_i - \bar{y}}{s_y} \tag{E.1}$$

where y_i is the ith observation of the response variable, $i = 1,2,3, \dots, n$,

$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $s_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}}$.

The standardization of predictor variables are

$$\frac{x_{ij} - \bar{x}_j}{s_{x_j}} \quad , j = 1,2,3, \dots, p \tag{E.2}$$

where x_{ij} is the ith observation of the jth predictor variable, $i = 1,2,3, \dots, n$, and $j = 1,2,3, \dots, p$,

$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}$ and $s_{x_j} = \sqrt{\frac{\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2}{n-1}}$.

E.2 Correlation transformation of the response and the predictor variables

The correlation transformation of the response variable is

$$y_i^* = \frac{1}{\sqrt{n-1}}\left(\frac{y_i-\bar{y}}{s_y}\right) \tag{E.3}$$

where y_i is the ith observation of the response variable, $i = 1,2,3,\ldots,n$,

$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $s_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1}}$.

The correlation transformation of the predictor variables are

$$x_{ij}^* = \frac{1}{\sqrt{n-1}}\left(\frac{x_{ij}-\bar{x}_j}{s_{x_j}}\right) \quad , \; j = 1,2,3,\ldots,p \tag{E.4}$$

where x_{ij} is the ith observation of the jth predictor variable, $i = 1,2,3,\ldots,n$, and $j = 1,2,3,\ldots,p$,

$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}$ and $s_{x_j} = \sqrt{\frac{\sum_{i=1}^{n}(x_{ij}-\bar{x}_j)^2}{n-1}}$.

Appendix F

Posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$ when $\boldsymbol{\lambda_1} = \mathbf{0}$

In Chapter 3, we use the hierarchical prior models prior (3.5) as follows.

$$\mathbf{y}|\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma} \sim N_n(\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2\mathbf{I}),$$

$$\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^2, \boldsymbol{\gamma} \sim N_{q_{\boldsymbol{\gamma}}}(\mathbf{0}, \frac{\sigma^2}{\lambda_2}\mathbf{I}),$$

$$\sigma^2|\boldsymbol{\gamma} = \sigma^2 \sim \text{inverse gamma}\left(\frac{\nu}{2}, \frac{\nu\xi}{2}\right),$$

$$p(\boldsymbol{\gamma}) = \left(\frac{1}{2}\right)^p.$$

Let $f(\boldsymbol{\gamma}|\mathbf{y})$ be the posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$.

We obtain the posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$.

$$f(\boldsymbol{\gamma}|\mathbf{y}) = p(\mathbf{y}|\boldsymbol{\gamma})p(\boldsymbol{\gamma}) .$$

where $$p(\mathbf{y}|\boldsymbol{\gamma}) \propto \left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right|^{-\frac{1}{2}} (\lambda_2)^{\frac{q_{\boldsymbol{\gamma}}}{2}} \left(\nu\xi + S_{\boldsymbol{\gamma}}^2\right)^{-\frac{(n+\nu)}{2}} ,$$

and $$S_{\boldsymbol{\gamma}}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_{\boldsymbol{\gamma}}\left[\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right]^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}.$$

Proof

Suppose a linear regression model

$\mathbf{y}|\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma} \sim N_n(\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2\mathbf{I})$,

$$p(\mathbf{y}|\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}) = (2\pi\sigma^2)^{-n/2} exp\left[-\frac{1}{2\sigma^2}\left(\mathbf{y}-\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right)^T\left(\mathbf{y}-\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right)\right]. \qquad \text{(F.1)}$$

Prior distribution for $\boldsymbol{\beta}_{\boldsymbol{\gamma}}$ is

$\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^2, \boldsymbol{\gamma} \sim N_{q_{\boldsymbol{\gamma}}}(\mathbf{0}, \frac{\sigma^2}{\lambda_2}\mathbf{I})$,

$$p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^2, \boldsymbol{\gamma}) = \left(2\pi\frac{\sigma^2}{\lambda_2}\right)^{-q_{\boldsymbol{\gamma}}/2}\left|\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right|^{-1/2} exp\left[-\frac{1}{2\sigma^2}\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\left(\frac{1}{\lambda_2}\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right].$$

$$p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^2, \boldsymbol{\gamma}) = [2\pi\sigma^2(\lambda_2)^{-1}]^{-q_{\boldsymbol{\gamma}}/2} exp\left[-\frac{1}{2\sigma^2}\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\left(\frac{1}{\lambda_2}\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right]. \qquad \text{(F.2)}$$

Prior distribution for σ^2 is

$\sigma^2|\boldsymbol{\gamma} = \sigma^2 \sim$ inverse gamma $\left(\frac{\nu}{2}, \frac{\nu\xi}{2}\right)$,

$$p(\sigma^2|\boldsymbol{\gamma}) = \left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-(\nu/2)-1} exp\left[-\frac{1}{\sigma^2}\left(\frac{\nu\xi}{2}\right)\right]. \qquad \text{(F.3)}$$

Combining the likelihood from (F.1) with the priors (F.2) and (F.3) yields the joint posterior probability $p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}|\mathbf{y})$ is

$$p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}|\mathbf{y}) = p(\mathbf{y}|\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma})p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^2, \boldsymbol{\gamma})p(\sigma^2|\boldsymbol{\gamma})p(\boldsymbol{\gamma}) .$$

$$p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}|\mathbf{y}) = (2\pi\sigma^2)^{-(n+q_{\boldsymbol{\gamma}})/2}(\lambda_2)^{q_{\boldsymbol{\gamma}}/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-(\nu/2)-1}$$
$$\times\, exp\left\{-\frac{1}{2\sigma^2}\left[(\mathbf{y}-\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}})^T(\mathbf{y}-\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}) + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\left(\frac{1}{\lambda_2}\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \nu\xi\right]\right\}p(\boldsymbol{\gamma}).$$

$$p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}|\mathbf{y}) = p(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma})p(\boldsymbol{\gamma}),$$

where $p(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma}) = p(\mathbf{y}|\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma})p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^2, \boldsymbol{\gamma})p(\sigma^2|\boldsymbol{\gamma})$.

$$p(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma}) = (2\pi\sigma^2)^{-(n+q_{\boldsymbol{\gamma}})/2}(\lambda_2)^{q_{\boldsymbol{\gamma}}/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-(\nu/2)-1}$$
$$\times\, exp\left\{-\frac{1}{2\sigma^2}\left[(\mathbf{y}-\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}})^T(\mathbf{y}-\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}) + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\left(\frac{1}{\lambda_2}\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \nu\xi\right]\right\}.$$

(F.4)

Since

$$(\mathbf{y}-\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}})^T(\mathbf{y}-\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}) + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\left(\frac{1}{\lambda_2}\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\boldsymbol{\beta}_{\boldsymbol{\gamma}}$$

$$= \mathbf{y}^T\mathbf{y} - 2\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y} + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\left(\frac{1}{\lambda_2}\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\boldsymbol{\beta}_{\boldsymbol{\gamma}}.$$

$$= \mathbf{y}^T\mathbf{y} - 2\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y} + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)\boldsymbol{\beta}_{\boldsymbol{\gamma}} \,. \qquad \text{(F.5)}$$

By considering the term $-2\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y} + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)\boldsymbol{\beta}_{\boldsymbol{\gamma}}$ in equation (F.5),

$$-2\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y} + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)\boldsymbol{\beta}_{\boldsymbol{\gamma}}$$

$$= \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)\left[-2\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y} + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right].$$

$$= \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\boldsymbol{\beta}_{\boldsymbol{\gamma}} - 2\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y} + (\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y})^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-2}(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}) - (\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y})^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-2}(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y})\right].$$

$$= \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)\left\{\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]^T\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right] - (\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y})^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-2}(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y})\right\}.$$

$$= \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right] - (\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y})^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}).$$

(F.6)

Replacing equation (F.6) into (F.5), so

$$\left(\mathbf{y} - \mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right)^T\left(\mathbf{y} - \mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right) + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\left(\frac{1}{\lambda_2}\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\boldsymbol{\beta}_{\boldsymbol{\gamma}}$$

$$= \mathbf{y}^T\mathbf{y} - (\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y})^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y})$$
$$+ \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right].$$

(F.7)

Hence, $p(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma})$ in equation (F.4) becomes

$$p(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma})$$
$$= (2\pi\sigma^2)^{-(n+q_{\gamma})/2}(\lambda_2)^{q_{\gamma}/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-(\nu/2)-1}$$
$$\times\, exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi + \mathbf{y}^T\mathbf{y} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right)^{-1}\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)\right]\right\}$$
$$\times\, exp\left\{-\frac{1}{2\sigma^2}\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right)\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]\right\}.$$

(F.8)

Integrating out $\boldsymbol{\beta}_{\boldsymbol{\gamma}}$ from $p(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma})$ yields

$$p(\mathbf{y}, \sigma^2|\boldsymbol{\gamma}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} p(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma})\, d\beta_1 d\beta_2 \dots d\beta_{q_{\gamma}}$$

$$= (2\pi\sigma^2)^{-(n+q_{\gamma})/2}(\lambda_2)^{q_{\gamma}/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-(\nu/2)-1}$$
$$\times\, exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi + \mathbf{y}^T\mathbf{y} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right)^{-1}\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)\right]\right\}$$
$$\times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} exp\left\{-\frac{1}{2\sigma^2}\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right)\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]\right\} d\beta_1 d\beta_2 \dots d\beta_{q_{\gamma}}$$

(F.9)

In equation (F.9), the integral term can be expressed as

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} exp\left\{-\frac{1}{2\sigma^2}\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right)\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]\right\} d\beta_1 d\beta_2 \dots d\beta_{q_{\gamma}}$$

$$= (2\pi\sigma^2)^{q_{\gamma}/2}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right|^{-\frac{1}{2}}$$
$$\times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty}(2\pi\sigma^2)^{-q_{\gamma}/2}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right|^{\frac{1}{2}} exp\left\{-\frac{1}{2\sigma^2}\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right)\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right.\right.$$
$$\left.\left. - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]\right\} d\beta_1 d\beta_2 \dots d\beta_{q_{\gamma}}$$

(F.10)

$$= (2\pi\sigma^2)^{q_{\gamma}/2}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\gamma}}\right|^{-\frac{1}{2}} \quad \text{(F.11)}$$

Replacing equation (F.11) into (F.9), so

$$p(\mathbf{y}, \sigma^2|\boldsymbol{\gamma}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p\big(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma}\big)\, d\beta_1 d\beta_2 \dots d\beta_{q_{\boldsymbol{\gamma}}}$$

$$= (2\pi\sigma^2)^{-(n+q_{\boldsymbol{\gamma}})/2}(\lambda_2)^{q_{\boldsymbol{\gamma}}/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-(\nu/2)-1}$$
$$\times\, exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi + \mathbf{y}^T\mathbf{y} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)\right]\right\}$$
$$\times\, (2\pi\sigma^2)^{q_{\boldsymbol{\gamma}}/2}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right|^{-\frac{1}{2}}.$$

$$= (\lambda_2)^{q_{\boldsymbol{\gamma}}/2}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right|^{-\frac{1}{2}}(2\pi\sigma^2)^{-n/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-(\nu/2)-1}$$
$$\times\, exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi + \mathbf{y}^T\mathbf{y} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)\right]\right\}.$$

(F.12)

Let $S_{\boldsymbol{\gamma}}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_{\boldsymbol{\gamma}}\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}.$ (F.13)

Thus, equation (F.12) can be expressed as

$p(\mathbf{y}, \sigma^2|\boldsymbol{\gamma})$

$$= (\lambda_2)^{q_{\boldsymbol{\gamma}}/2}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right|^{-\frac{1}{2}}(2\pi\sigma^2)^{-n/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-(\nu/2)-1}\, exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi + S_{\boldsymbol{\gamma}}^2\right]\right\}.$$

(F.14)

Integrating out σ^2 from $p(\mathbf{y},\sigma^2|\boldsymbol{\gamma})$ yields

$$p(\mathbf{y}|\boldsymbol{\gamma}) = \int_0^\infty p(\mathbf{y},\sigma^2|\boldsymbol{\gamma})d\sigma^2$$

$$= (\lambda_2)^{q_\gamma/2}\left|\mathbf{X}_\gamma^T\mathbf{X}_\gamma+\lambda_2\mathbf{I}_{q_\gamma}\right|^{-\frac{1}{2}}(2\pi)^{-n/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}$$
$$\times\int_0^\infty(\sigma^2)^{-[(n+\nu)/2]-1}exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi+S_\gamma^2\right]\right\}d\sigma^2.$$

$$= (\lambda_2)^{q_\gamma/2}\left|\mathbf{X}_\gamma^T\mathbf{X}_\gamma+\lambda_2\mathbf{I}_{q_\gamma}\right|^{-\frac{1}{2}}(2\pi)^{-n/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}\left[\frac{1}{2}\left(\nu\xi+S_\gamma^2\right)\right]^{-(n+\nu)/2}\left[\Gamma\left(\frac{n}{2}+\frac{\nu}{2}\right)\right]$$
$$\times\int_0^\infty\left[\frac{1}{2}\left(\nu\xi+S_\gamma^2\right)\right]^{(n+\nu)/2}\left[\Gamma\left(\frac{n}{2}+\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-[(n+\nu)/2]-1}exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi+S_\gamma^2\right]\right\}d\sigma^2.$$

Since

$$\int_0^\infty\left[\frac{1}{2}\left(\nu\xi+S_\gamma^2\right)\right]^{(n+\nu)/2}\left[\Gamma\left(\frac{n}{2}+\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-[(n+\nu)/2]-1}exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi+S_\gamma^2\right]\right\}d\sigma^2 = 1 \tag{F.15}$$

Thus,

$$p(\mathbf{y}|\boldsymbol{\gamma})$$
$$= (2\pi)^{-n/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}\left[\Gamma\left(\frac{n}{2}+\frac{\nu}{2}\right)\right]\left(\frac{1}{2}\right)^{-(n+\nu)/2}(\lambda_2)^{q_\gamma/2}\left|\mathbf{X}_\gamma^T\mathbf{X}_\gamma+\lambda_2\mathbf{I}_{q_\gamma}\right|^{-\frac{1}{2}}\left(\nu\xi+S_\gamma^2\right)^{-\frac{(n+\nu)}{2}}, \tag{F.16}$$

where $S_\gamma^2 = \mathbf{y}^T\mathbf{y}-\mathbf{y}^T\mathbf{X}_\gamma\left(\mathbf{X}_\gamma^T\mathbf{X}_\gamma+\lambda_2\mathbf{I}_{q_\gamma}\right)^{-1}\mathbf{X}_\gamma^T\mathbf{y}$.

Posterior probability of model $\boldsymbol{\gamma}$ given $\mathbf{y}$ is

$$f(\boldsymbol{\gamma}|\mathbf{y}) = p(\mathbf{y}|\boldsymbol{\gamma})p(\boldsymbol{\gamma}). \tag{F.17}$$

In this thesis, we use the uniform prior for $\boldsymbol{\gamma}$, $p(\boldsymbol{\gamma}) = \left(\frac{1}{2}\right)^{p}$.

Thus,

$$f(\boldsymbol{\gamma}|\mathbf{y})$$
$$= (2\pi)^{-n/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}\left[\Gamma\left(\frac{n}{2}+\frac{\nu}{2}\right)\right]\left(\frac{1}{2}\right)^{-\frac{(n+\nu)}{2}+p}(\lambda_2)^{q_\gamma/2}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}+\lambda_2\mathbf{I}_{q_\gamma}\right|^{-\frac{1}{2}}\left(\nu\xi+S_{\boldsymbol{\gamma}}^2\right)^{-\frac{(n+\nu)}{2}}. \tag{F.18}$$

The posterior probability of model $\boldsymbol{\gamma}$ given $\mathbf{y}$ in equation (F.18) can be expressed as

$$f(\boldsymbol{\gamma}|\mathbf{y}) \propto \left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}+\lambda_2\mathbf{I}_{q_\gamma}\right|^{-\frac{1}{2}}(\lambda_2)^{\frac{q_\gamma}{2}}\left(\nu\xi+S_{\boldsymbol{\gamma}}^2\right)^{-\frac{(n+\nu)}{2}}, \tag{F.19}$$

and the posterior probability $p(\mathbf{y}|\boldsymbol{\gamma})$ in equation (F.16) can be expressed as

$$p(\mathbf{y}|\boldsymbol{\gamma}) \propto \left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}+\lambda_2\mathbf{I}_{q_\gamma}\right|^{-\frac{1}{2}}(\lambda_2)^{\frac{q_\gamma}{2}}\left(\nu\xi+S_{\boldsymbol{\gamma}}^2\right)^{-\frac{(n+\nu)}{2}}, \tag{F.20}$$

where $$S_{\boldsymbol{\gamma}}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_{\boldsymbol{\gamma}}\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}+\lambda_2\mathbf{I}_{q_\gamma}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}.$$ ■

From equation (F.10), we see that the posterior distribution of $\boldsymbol{\beta}_{\boldsymbol{\gamma}}$ is

$$Normal\left(\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\,, \sigma^2\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \lambda_2\mathbf{I}_{q_{\boldsymbol{\gamma}}}\right)^{-1}\right)$$

which is the same distribution as prior distribution for $\boldsymbol{\beta}_{\boldsymbol{\gamma}}$. ■

From equation (F.15), we see that the posterior distribution of σ^2 is

$$inverse\ gamma\ \left(\frac{n}{2} + \frac{\nu}{2},\ \frac{1}{2}\left(\nu\xi + S_{\boldsymbol{\gamma}}^2\right)\right)$$

which is the same distribution as prior distribution for σ^2. ■

Appendix G

Posterior model probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$ for elastic net linear regression model

In Chapter 3, we use the hierarchical model prior for elastic net linear regression models as follows.

$$\mathbf{y}|\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma} \sim N_n(\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2\mathbf{I}),$$

$$\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^2, \mathbf{D}_\tau, \boldsymbol{\gamma} \sim N_{q_{\boldsymbol{\gamma}}}(\mathbf{0}, \sigma^2\mathbf{D}_\tau),$$

where $\mathbf{D}_\tau$ is a diagonal matrix with diagonal elements $\left(\tau_j^{-2} + \lambda_2\right)^{-1}$,

$$\tau_j^2 \sim Exponential\left(\frac{\lambda_1^2}{2}\right), j = 1,2,3,\ldots,p,$$

$$\sigma^2|\boldsymbol{\gamma} = \sigma^2 \sim \text{inverse gamma}\left(\frac{\nu}{2}, \frac{\nu\xi}{2}\right),$$

$$p(\boldsymbol{\gamma}) = \left(\frac{1}{2}\right)^p.$$

Let $g(\boldsymbol{\gamma}|\mathbf{y})$ be the posterior probability of $\boldsymbol{\gamma}$ given $\mathbf{y}$.

We obtain the posterior probability $g(\boldsymbol{\gamma}|\mathbf{y})$.

$$g(\boldsymbol{\gamma}|\mathbf{y}) \propto |\mathbf{D}_\tau|^{-\frac{1}{2}}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_\tau^{-1}\right|^{-\frac{1}{2}}\left(\nu\xi + S_{\boldsymbol{\gamma}\mathbf{D}}^2\right)^{-\frac{(n+\nu)}{2}},$$

where $S_{\boldsymbol{\gamma}\mathbf{D}}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_{\boldsymbol{\gamma}}\left[\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_\tau^{-1}\right]^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}$.

Proof

Suppose a linear regression model

$\mathbf{y}|\boldsymbol{\beta}_\gamma, \sigma^2, \boldsymbol{\gamma} \sim N_n(\mathbf{X}_\gamma\boldsymbol{\beta}_\gamma, \sigma^2\mathbf{I})$,

$$p(\mathbf{y}|\boldsymbol{\beta}_\gamma, \sigma^2, \boldsymbol{\gamma}) = (2\pi\sigma^2)^{-n/2} exp\left[-\frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{X}_\gamma\boldsymbol{\beta}_\gamma)^T(\mathbf{y}-\mathbf{X}_\gamma\boldsymbol{\beta}_\gamma)\right]. \quad \text{(G.1)}$$

Prior distribution for $\boldsymbol{\beta}_\gamma$ is

$\boldsymbol{\beta}_\gamma|\sigma^2, \mathbf{D}_\tau, \boldsymbol{\gamma} \sim N_{q_\gamma}(\mathbf{0}, \sigma^2\mathbf{D}_\tau)$ where $\mathbf{D}_\tau$ is a diagonal matrix with diagonal elements $\left(\tau_j^{-2}+\lambda_2\right)^{-1}$,

and $\tau_j^2 \sim Exponential\left(\frac{\lambda_1^2}{2}\right)$, $j = 1,2,3,\dots,p$,

$p(\tau_j^2) = \frac{\lambda_1^2}{2} exp\left[-\frac{\lambda_1^2\tau_j^2}{2}\right]$; $\tau_j^2 > 0$.

$$p(\boldsymbol{\beta}_\gamma|\sigma^2, \mathbf{D}_\tau, \boldsymbol{\gamma}) = (2\pi\sigma^2)^{-q_\gamma/2}|\mathbf{D}_\tau|^{-1/2} exp\left[-\frac{1}{2\sigma^2}\boldsymbol{\beta}_\gamma^T\mathbf{D}_\tau^{-1}\boldsymbol{\beta}_\gamma\right]. \quad \text{(G.2)}$$

Prior distribution for σ^2 is

$\sigma^2|\boldsymbol{\gamma} = \sigma^2 \sim$ inverse gamma $\left(\frac{\nu}{2}, \frac{\nu\xi}{2}\right)$,

$$p(\sigma^2|\boldsymbol{\gamma}) = \left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-(\nu/2)-1} exp\left[-\frac{1}{\sigma^2}\left(\frac{\nu\xi}{2}\right)\right]. \quad \text{(G.3)}$$

Combining the likelihood from (G.1) with the priors (G.2) and (G.3) yields the joint posterior probability $p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}|\mathbf{y})$ is

$$p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}|\mathbf{y}) = p(\mathbf{y}|\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma})p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^2, \boldsymbol{\gamma})p(\sigma^2|\boldsymbol{\gamma})p(\boldsymbol{\gamma}).$$

$$\begin{aligned} p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}|\mathbf{y}) &= (2\pi\sigma^2)^{-(n+q_{\gamma})/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-(\nu/2)-1}|\mathbf{D}_{\tau}|^{-1/2} \\ &\times exp\left\{-\frac{1}{2\sigma^2}\left[(\mathbf{y}-\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}})^T(\mathbf{y}-\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}) + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\mathbf{D}_{\tau}^{-1}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \nu\xi\right]\right\}p(\boldsymbol{\gamma}). \end{aligned}$$

$$p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}|\mathbf{y}) = p(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma})p(\boldsymbol{\gamma}),$$

where $p(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma}) = p(\mathbf{y}|\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma})p(\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^2, \boldsymbol{\gamma})p(\sigma^2|\boldsymbol{\gamma})$.

$$\begin{aligned} p(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma}) &= (2\pi\sigma^2)^{-(n+q_{\gamma})/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-(\nu/2)-1}|\mathbf{D}_{\tau}|^{-1/2} \\ &\times exp\left\{-\frac{1}{2\sigma^2}\left[(\mathbf{y}-\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}})^T(\mathbf{y}-\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}) + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\mathbf{D}_{\tau}^{-1}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \nu\xi\right]\right\}. \end{aligned} \tag{G.4}$$

Since

$(\mathbf{y}-\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}})^T(\mathbf{y}-\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}) + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\mathbf{D}_{\tau}^{-1}\boldsymbol{\beta}_{\boldsymbol{\gamma}}$

$= \mathbf{y}^T\mathbf{y} - 2\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y} + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}} + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\mathbf{D}_{\tau}^{-1}\boldsymbol{\beta}_{\boldsymbol{\gamma}}$.

$$= \mathbf{y}^T\mathbf{y} - 2\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y} + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1})\boldsymbol{\beta}_{\boldsymbol{\gamma}} . \tag{G.5}$$

By considering the term $-2\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}+\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}}+\mathbf{D}_\tau^{-1}\right)\boldsymbol{\beta}_{\boldsymbol{\gamma}}$ in equation (G.5),

$$-2\boldsymbol{\beta}_{\gamma}^T\mathbf{X}_{\gamma}^T\mathbf{y}+\boldsymbol{\beta}_{\gamma}^T(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1})\boldsymbol{\beta}_{\gamma}$$

$$=\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)\left[-2\boldsymbol{\beta}_{\gamma}^T\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)^{-1}\mathbf{X}_{\gamma}^T\mathbf{y}+\boldsymbol{\beta}_{\gamma}^T\boldsymbol{\beta}_{\gamma}\right].$$

$$=\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)\left[\boldsymbol{\beta}_{\gamma}^T\boldsymbol{\beta}_{\gamma}-2\boldsymbol{\beta}_{\gamma}^T\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)^{-1}\mathbf{X}_{\gamma}^T\mathbf{y}+\left(\mathbf{X}_{\gamma}^T\mathbf{y}\right)^T\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)^{-2}\mathbf{X}_{\gamma}^T\mathbf{y}-\left(\mathbf{X}_{\gamma}^T\mathbf{y}\right)^T\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)^{-2}\mathbf{X}_{\gamma}^T\mathbf{y}\right].$$

$$=\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)\left\{\left[\boldsymbol{\beta}_{\gamma}-\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)^{-1}\mathbf{X}_{\gamma}^T\mathbf{y}\right]^T\left[\boldsymbol{\beta}_{\gamma}-\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)^{-1}\mathbf{X}_{\gamma}^T\mathbf{y}\right]-\left(\mathbf{X}_{\gamma}^T\mathbf{y}\right)^T\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)^{-2}\mathbf{X}_{\gamma}^T\mathbf{y}\right\}.$$

$$=\left[\boldsymbol{\beta}_{\gamma}-\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)^{-1}\mathbf{X}_{\gamma}^T\mathbf{y}\right]^T\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)\left[\boldsymbol{\beta}_{\gamma}-\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)^{-1}\mathbf{X}_{\gamma}^T\mathbf{y}\right]-(\mathbf{X}_{\gamma}^T\mathbf{y})^T\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)^{-1}(\mathbf{X}_{\gamma}^T\mathbf{y}).$$

(G.6)

Replacing equation (G.6) into (G.5), therefore,

$$\left(\mathbf{y}-\mathbf{X}_{\gamma}\boldsymbol{\beta}_{\gamma}\right)^T\left(\mathbf{y}-\mathbf{X}_{\gamma}\boldsymbol{\beta}_{\gamma}\right)+\boldsymbol{\beta}_{\gamma}^T\mathbf{D}_\tau^{-1}\boldsymbol{\beta}_{\gamma}$$

$$=\mathbf{y}^T\mathbf{y}+\left[\boldsymbol{\beta}_{\gamma}-\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)^{-1}\mathbf{X}_{\gamma}^T\mathbf{y}\right]^T\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)\left[\boldsymbol{\beta}_{\gamma}-\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)^{-1}\mathbf{X}_{\gamma}^T\mathbf{y}\right]$$
$$-(\mathbf{X}_{\gamma}^T\mathbf{y})^T\left(\mathbf{X}_{\gamma}^T\mathbf{X}_{\gamma}+\mathbf{D}_\tau^{-1}\right)^{-1}(\mathbf{X}_{\gamma}^T\mathbf{y}).$$

(G.7)

Hence $p(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma})$ in equation (G.4) becomes

$$
\begin{aligned}
&p(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma}) \\
&= (2\pi\sigma^2)^{-(n+q_{\gamma})/2} \left(\frac{\nu\xi}{2}\right)^{\nu/2} \left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1} (\sigma^2)^{-(\nu/2)-1} |\mathbf{D}_{\tau}|^{-1/2} \\
&\quad \times exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi + \mathbf{y}^T\mathbf{y} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)^{-1}\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)\right]\right\} \\
&\quad \times exp\left\{-\frac{1}{2\sigma^2}\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]^T \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]\right\}.
\end{aligned}
\tag{G.8}
$$

Integrating out $\boldsymbol{\beta}_{\boldsymbol{\gamma}}$ from $p(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma})$ yields

$$
\begin{aligned}
&p(\mathbf{y}, \sigma^2|\boldsymbol{\gamma}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma})\, d\beta_1 d\beta_2 \dots d\beta_{q_{\boldsymbol{\gamma}}} \\
&= (2\pi\sigma^2)^{-(n+q_{\gamma})/2} \left(\frac{\nu\xi}{2}\right)^{\nu/2} \left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1} (\sigma^2)^{-(\nu/2)-1} |\mathbf{D}_{\tau}|^{-1/2} \\
&\quad \times exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi + \mathbf{y}^T\mathbf{y} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)^{-1}\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)\right]\right\} \\
&\quad \times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} exp\left\{-\frac{1}{2\sigma^2}\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]^T \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]\right\} \\
&\qquad d\beta_1 d\beta_2 \dots d\beta_{q_{\boldsymbol{\gamma}}}
\end{aligned}
\tag{G.9}
$$

In equation (G.9), the integral term can be expressed as

$$
\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} exp\left\{-\frac{1}{2\sigma^2}\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]^T \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]\right\} \\
&\qquad d\beta_1 d\beta_2 \dots d\beta_{q_{\boldsymbol{\gamma}}} \\
&= (2\pi\sigma^2)^{q_{\gamma}/2}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right|^{-1/2} \times \\
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (2\pi\sigma^2)^{-q_{\gamma}/2}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right|^{1/2} exp\left\{-\frac{1}{2\sigma^2}\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]^T \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right]\right\} \\
&\qquad d\beta_1 d\beta_2 \dots d\beta_{q_{\boldsymbol{\gamma}}}
\end{aligned}
\tag{G.10}
$$

$$
= (2\pi\sigma^2)^{q_{\gamma}/2}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right|^{-1/2}. \tag{G.11}
$$

Replacing equation (G.11) into (G.9), therefore

$$p(\mathbf{y}, \sigma^2|\boldsymbol{\gamma}) \; = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p\left(\mathbf{y}, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2|\boldsymbol{\gamma}\right) d\beta_1 d\beta_2 \dots d\beta_{q_{\boldsymbol{\gamma}}}$$

$$= (2\pi\sigma^2)^{-(n+q_{\boldsymbol{\gamma}})/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-(\nu/2)-1}|\mathbf{D}_{\tau}|^{-1/2}$$
$$\times\, exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi + \mathbf{y}^T\mathbf{y} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)^{-1}\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)\right]\right\}$$
$$\times\, (2\pi\sigma^2)^{q_{\boldsymbol{\gamma}}/2}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right|^{-1/2}.$$

$$= (2\pi\sigma^2)^{-n/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-(\nu/2)-1}|\mathbf{D}_{\tau}|^{-1/2}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right|^{-1/2}$$
$$\times\, exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi + \mathbf{y}^T\mathbf{y} - \left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)^T\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right)^{-1}\left(\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}\right)\right]\right\}. \tag{G.12}$$

Let $\quad S_{\boldsymbol{\gamma}\mathbf{D}}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_{\boldsymbol{\gamma}}\left[\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right]^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{y}.$ (G.13)

Thus, equation (G.12) can be expressed as

$$p(\mathbf{y}, \sigma^2|\boldsymbol{\gamma})$$

$$= (2\pi\sigma^2)^{-n/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-(\nu/2)-1}|\mathbf{D}_{\tau}|^{-1/2}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right|^{-1/2} exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi + S_{\boldsymbol{\gamma}\mathbf{D}}^2\right]\right\}. \tag{G.14}$$

Integrating out σ^2 from $p(\mathbf{y}, \sigma^2|\boldsymbol{\gamma})$ yields

$$p(\mathbf{y}|\boldsymbol{\gamma}) \; = \int_0^{\infty} p(\mathbf{y}, \sigma^2|\boldsymbol{\gamma})\, d\sigma^2$$

$$= |\mathbf{D}_{\tau}|^{-1/2}\left|\mathbf{X}_{\boldsymbol{\gamma}}^T\mathbf{X}_{\boldsymbol{\gamma}} + \mathbf{D}_{\tau}^{-1}\right|^{-1/2}(2\pi)^{-n/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}$$
$$\times \int_0^{\infty}(\sigma^2)^{-[(n+\nu)/2]-1} exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi + S_{\boldsymbol{\gamma}\mathbf{D}}^2\right]\right\} d\sigma^2.$$

$p(\mathbf{y}|\boldsymbol{\gamma})$

$$= |\mathbf{D}_\tau|^{-1/2}\left|\mathbf{X}_\gamma^T\mathbf{X}_\gamma + \mathbf{D}_\tau^{-1}\right|^{-1/2}(2\pi)^{-n/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}\left[\Gamma\left(\frac{n}{2}+\frac{\nu}{2}\right)\right]\left[\frac{1}{2}\left(\nu\xi + S_{\gamma\mathbf{D}}^2\right)\right]^{-(n+\nu)/2}$$
$$\times \int_0^\infty \left[\frac{1}{2}\left(\nu\xi + S_{\gamma\mathbf{D}}^2\right)\right]^{(n+\nu)/2}\left[\Gamma\left(\frac{n}{2}+\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-[(n+\nu)/2]-1} exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi + S_{\gamma\mathbf{D}}^2\right]\right\} d\sigma^2.$$

Since

$$\int_0^\infty \left[\frac{1}{2}\left(\nu\xi + S_{\gamma\mathbf{D}}^2\right)\right]^{(n+\nu)/2}\left[\Gamma\left(\frac{n}{2}+\frac{\nu}{2}\right)\right]^{-1}(\sigma^2)^{-[(n+\nu)/2]-1} exp\left\{-\frac{1}{2\sigma^2}\left[\nu\xi + S_{\gamma\mathbf{D}}^2\right]\right\} d\sigma^2 = 1. \tag{G.15}$$

Hence,

$p(\mathbf{y}|\boldsymbol{\gamma})$

$$= (2\pi)^{-n/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}\left[\Gamma\left(\frac{n}{2}+\frac{\nu}{2}\right)\right]\left(\frac{1}{2}\right)^{-(n+\nu)/2}|\mathbf{D}_\tau|^{-\frac{1}{2}}\left|\mathbf{X}_\gamma^T\mathbf{X}_\gamma + \mathbf{D}_\tau^{-1}\right|^{-\frac{1}{2}}\left(\nu\xi + S_{\gamma\mathbf{D}}^2\right)^{-\frac{(n+\nu)}{2}}, \tag{G.16}$$

where $S_{\gamma\mathbf{D}}^2 = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{X}_\gamma\left[\mathbf{X}_\gamma^T\mathbf{X}_\gamma + \mathbf{D}_\tau^{-1}\right]^{-1}\mathbf{X}_\gamma^T\mathbf{y}$.

Posterior probability of model $\boldsymbol{\gamma}$ given $\mathbf{y}$ is

$$g(\boldsymbol{\gamma}|\mathbf{y}) = p(\mathbf{y}|\boldsymbol{\gamma})p(\boldsymbol{\gamma}). \tag{G.17}$$

In this thesis, we use the uniform prior for $\boldsymbol{\gamma}$, $p(\boldsymbol{\gamma}) = \left(\frac{1}{2}\right)^p$.

Hence,

$$g(\boldsymbol{\gamma}|\mathbf{y})$$

$$= (2\pi)^{-n/2}\left(\frac{\nu\xi}{2}\right)^{\nu/2}\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}\left[\Gamma\left(\frac{n}{2}+\frac{\nu}{2}\right)\right]\left(\frac{1}{2}\right)^{-\frac{(n+\nu)}{2}+p}|\mathbf{D}_{\tau}|^{-\frac{1}{2}}\left|\mathbf{X}_{\boldsymbol{\gamma}}^{T}\mathbf{X}_{\boldsymbol{\gamma}}+\mathbf{D}_{\tau}^{-1}\right|^{-\frac{1}{2}}\left(\nu\xi+S_{\boldsymbol{\gamma}\mathbf{D}}^{2}\right)^{-\frac{(n+\nu)}{2}}.$$

(G.18)

The posterior probability of model $\boldsymbol{\gamma}$ given $\mathbf{y}$ in equation (G.18) can be expressed as

$$g(\boldsymbol{\gamma}|\mathbf{y}) \propto |\mathbf{D}_{\tau}|^{-\frac{1}{2}}\left|\mathbf{X}_{\boldsymbol{\gamma}}^{T}\mathbf{X}_{\boldsymbol{\gamma}}+\mathbf{D}_{\tau}^{-1}\right|^{-\frac{1}{2}}\left(\nu\xi+S_{\boldsymbol{\gamma}\mathbf{D}}^{2}\right)^{-\frac{(n+\nu)}{2}}, \qquad \text{(G.19)}$$

where $S_{\boldsymbol{\gamma}\mathbf{D}}^{2} = \mathbf{y}^{T}\mathbf{y}-\mathbf{y}^{T}\mathbf{X}_{\boldsymbol{\gamma}}\left[\mathbf{X}_{\boldsymbol{\gamma}}^{T}\mathbf{X}_{\boldsymbol{\gamma}}+\mathbf{D}_{\tau}^{-1}\right]^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^{T}\mathbf{y}$. ■

From equation (G.10), we see that the posterior distribution of $\boldsymbol{\beta}_{\boldsymbol{\gamma}}$ is

$$Normal\left(\left(\mathbf{X}_{\boldsymbol{\gamma}}^{T}\mathbf{X}_{\boldsymbol{\gamma}}+\mathbf{D}_{\tau}^{-1}\right)^{-1}\mathbf{X}_{\boldsymbol{\gamma}}^{T}\mathbf{y}\,,\sigma^{2}\left(\mathbf{X}_{\boldsymbol{\gamma}}^{T}\mathbf{X}_{\boldsymbol{\gamma}}+\mathbf{D}_{\tau}^{-1}\right)^{-1}\right)$$

which is the same distribution as prior distribution for $\boldsymbol{\beta}_{\boldsymbol{\gamma}}$. ■

From equation (G.15), we see that the posterior distribution of σ^2 is

$$inverse\ gamma\ \left(\frac{n}{2}+\frac{\nu}{2},\ \frac{1}{2}\left(\nu\xi+S_{\boldsymbol{\gamma}\mathbf{D}}^{2}\right)\right)$$

which is the same distribution as prior distribution for σ^2. ■

Appendix H

Significance and Important Predictors

Appendix H presents results of the set of significance and important predictors for 100 datasets. The 100 datasets are generated using the simulation design proposed by Lykou and Ntzoufras (2012), which consists of 15 predictor variables of 50 observations each. This dataset has different correlated between predictor variables. The last five predictors are highly correlated, whereas, there are small to moderate correlations between x_j , $j = 1, \ldots ,5$ and $x_{11},\ldots,x_{15}$.

By the method for specification the penalty parameter λ_2 associated with Bayes factor and the Flowchart 3.1 described in Chapter 3, we can identify the set of predictors in the following.

Set of significance predictors is $\{X_j | X_j \in H_1 \text{ and } \mathrm{BF}_{10} > 1\}$.

Set of important predictors is $\{X_j | X_j \in H_1 \text{ and } \mathrm{BF}_{10} > 3\}$.

Table H shows the results of the set of significance and important predictors for 100 datasets.

Table H Set of significance and important predictors for 100 datasets

Data no. 1															
Predictor variables [a]	x_1	x_{13}	x_{11}	x_{12}	x_{15}	x_{14}	x_4	x_6	x_8	x_2	x_7	x_3	x_5	x_9	x_{10}
Significance predictors [b]	x_1, x_{13}, x_{11}, x_{12}, x_4, x_6, x_{15}, x_{14}, x_8, x_2, x_7, x_5, x_{10}														
Important predictors [c]	x_1 , x_{11}, x_{13}, x_{12}, x_4, x_{15}, x_{14}														

Data no. 2															
Predictor variables [a]	x_{13}	x_7	x_1	x_{14}	x_3	x_{11}	x_{12}	x_8	x_6	x_4	x_{15}	x_9	x_2	x_{10}	x_5
Significance predictors [b]	x_7, x_1, x_3, x_{13}, x_{14}, x_{11}, x_5, x_{12}, x_{15}, x_4, x_6, x_8, x_9, x_2														
Important predictors [c]	x_1, x_3, x_7, x_{13}, x_{11}, x_{14}														

Data no. 3															
Predictor variables [a]	x_1	x_{11}	x_{15}	x_{13}	x_4	x_{12}	x_{14}	x_3	x_7	x_{10}	x_6	x_9	x_2	x_5	x_8
Significance predictors [b]	x_1, x_{11}, x_{13}, x_{15}, x_4, x_7, x_3, x_{12}, x_{14}, x_{10}, x_5, x_9, x_2, x_6														
Important predictors [c]	x_1, x_{11}, x_4, x_{13}, x_7, x_{15}														

Data no. 4															
Predictor variables [a]	x_{11}	x_{13}	x_1	x_4	x_{12}	x_{14}	x_{15}	x_7	x_{10}	x_2	x_6	x_8	x_3	x_5	x_9
Significance predictors [b]	x_1, x_{13}, x_4, x_{11}, x_7, x_{14}, x_{12}, x_{15}, x_6, x_8, x_5, x_2, x_{10}														
Important predictors [c]	x_1, x_{11}, x_{13}, x_4, x_{12}, x_{14}														

Data no. 5															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{12}	x_{15}	x_7	x_{14}	x_5	x_2	x_3	x_8	x_6	x_4	x_{10}	x_9
Significance predictors [b]	x_1, x_7, x_{13}, x_{11}, x_{12}, x_{15}, x_{14}, x_5, x_8, x_2, x_3, x_4, x_6														
Important predictors [c]	x_1, x_{11}, x_7, x_{13}, x_{12}, x_{15}														

Data no. 6															
Predictor variables [a]	x_7	x_1	x_{11}	x_2	x_{12}	x_{14}	x_8	x_{10}	x_{15}	x_5	x_{13}	x_4	x_9	x_6	x_3
Significance predictors [b]	x_2, x_7, x_{12}, x_1, x_{14}, x_{11}, x_8, x_{15}, x_{13}, x_{10}, x_4, x_5, x_6, x_3														
Important predictors [c]	x_1, x_7, x_{11}, x_2, x_{12}														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\dots,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 7															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{12}	x_{14}	x_{15}	x_7	x_5	x_2	x_3	x_4	x_{10}	x_9	x_8	x_6
Significance predictors [b]	$x_1, x_{11}, x_{13}, x_{12}, x_{14}, x_7, x_{15}, x_5, x_4, x_2, x_3, x_9$														
Important predictors [c]	$x_1, x_{11}, x_{13}, x_{12}, x_{14}$														

Data no. 8															
Predictor variables [a]	x_{11}	x_1	x_{14}	x_7	x_{13}	x_{15}	x_{12}	x_5	x_4	x_8	x_6	x_{10}	x_9	x_3	x_2
Significance predictors [b]	$x_1, x_7, x_{11}, x_{13}, x_{15}, x_{14}, x_{12}, x_5, x_8, x_4, x_6, x_9, x_3$														
Important predictors [c]	$x_1, x_{11}, x_7, x_{14}, x_{13}$														

Data no. 9															
Predictor variables [a]	x_1	x_{13}	x_{14}	x_{15}	x_{12}	x_3	x_{11}	x_8	x_2	x_7	x_4	x_{10}	x_6	x_5	x_9
Significance predictors [b]	$x_1, x_7, x_3, x_{13}, x_{15}, x_{14}, x_{12}, x_{11}, x_2, x_8, x_5, x_4$														
Important predictors [c]	$x_1, x_{13}, x_3, x_7, x_{14}, x_{12}, x_{15}$														

Data no. 10															
Predictor variables [a]	x_{11}	x_{14}	x_{13}	x_{15}	x_{12}	x_2	x_1	x_4	x_9	x_5	x_3	x_{10}	x_6	x_8	x_7
Significance predictors [b]	$x_{13}, x_{15}, x_{12}, x_2, x_{14}, x_1, x_{11}, x_4, x_9, x_3, x_{10}, x_5, x_7$														
Important predictors [c]	x_1, x_{11}														

Data no. 11															
Predictor variables [a]	x_{11}	x_1	x_{15}	x_{14}	x_{13}	x_7	x_4	x_{12}	x_6	x_5	x_9	x_3	x_8	x_2	x_{10}
Significance predictors [b]	$x_1, x_7, x_{11}, x_{14}, x_{13}, x_{15}, x_4, x_{12}, x_6, x_5, x_3, x_2$														
Important predictors [c]	$x_1, x_7, x_{11}, x_{14}, x_{13}, x_{15}, x_4$														

Data no. 12															
Predictor variables [a]	x_{11}	x_{13}	x_{14}	x_7	x_4	x_1	x_8	x_{15}	x_{10}	x_{12}	x_2	x_6	x_3	x_9	x_5
Significance predictors [b]	$x_{13}, x_7, x_{11}, x_{14}, x_8, x_4, x_1, x_2, x_{15}, x_{12}, x_{10}, x_3$														
Important predictors [c]	$x_{13}, x_{11}, x_1, x_{14}$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 13															
Predictor variables [a]	x_{11}	x_1	x_{15}	x_{13}	x_{14}	x_4	x_{12}	x_8	x_3	x_2	x_5	x_9	x_{10}	x_6	x_7
Significance predictors [b]	$x_1, x_4, x_{15}, x_{13}, x_{14}, x_{12}, x_{11}, x_3, x_5, x_7, x_8, x_2, x_6, x_9$														
Important predictors [c]	x_1, x_4, x_{11}, x_{13}														

Data no. 14															
Predictor variables [a]	x_1	x_{11}	x_{15}	x_{14}	x_{13}	x_{12}	x_4	x_2	x_3	x_7	x_5	x_9	x_8	x_{10}	x_6
Significance predictors [b]	$x_1, x_{11}, x_{15}, x_{14}, x_{13}, x_4, x_{12}, x_7, x_3, x_8, x_2, x_5, x_6$														
Important predictors [c]	$x_1, x_{11}, x_{15}, x_{14}, x_4, x_{13}, x_7$														

Data no. 15															
Predictor variables [a]	x_1	x_{13}	x_{11}	x_{12}	x_{15}	x_{14}	x_7	x_8	x_4	x_3	x_2	x_5	x_6	x_{10}	x_9
Significance predictors [b]	$x_1, x_7, x_{13}, x_{12}, x_{15}, x_{11}, x_8, x_{14}, x_3, x_2, x_4, x_{10}$														
Important predictors [c]	x_1, x_{13}, x_7, x_{11}														

Data no. 16															
Predictor variables [a]	x_7	x_{13}	x_{11}	x_1	x_{15}	x_{14}	x_{12}	x_5	x_6	x_4	x_3	x_9	x_2	x_8	x_{10}
Significance predictors [b]	$x_7, x_1, x_{13}, x_{11}, x_{15}, x_{14}, x_4, x_{12}, x_6, x_5, x_8, x_2, x_3$														
Important predictors [c]	$x_7, x_1, x_{13}, x_{11}, x_{14}, x_{15}, x_{12}$														

Data no. 17															
Predictor variables [a]	x_{11}	x_{14}	x_{15}	x_{13}	x_1	x_{12}	x_7	x_4	x_3	x_2	x_5	x_6	x_8	x_{10}	x_9
Significance predictors [b]	$x_1, x_{11}, x_7, x_4, x_{14}, x_{13}, x_{15}, x_3, x_{12}, x_5, x_2, x_6$														
Important predictors [c]	$x_1, x_{11}, x_{14}, x_4, x_{15}, x_{13}, x_{12}, x_3$														

Data no. 18															
Predictor variables [a]	x_1	x_{11}	x_{14}	x_{13}	x_{15}	x_{12}	x_4	x_7	x_3	x_2	x_8	x_6	x_5	x_{10}	x_9
Significance predictors [b]	$x_1, x_7, x_{11}, x_{14}, x_{13}, x_{15}, x_4, x_{12}, x_5, x_3, x_2, x_8$														
Important predictors [c]	$x_1, x_{11}, x_{14}, x_{13}, x_7, x_{15}, x_{12}, x_4$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 19															
Predictor variables [a]	x_{11}	x_1	x_{14}	x_{15}	x_{13}	x_{12}	x_{10}	x_2	x_3	x_5	x_6	x_4	x_7	x_8	x_9
Significance predictors [b]	$x_1, x_{14}, x_{15}, x_{13}, x_{11}, x_{12}, x_3, x_{10}, x_2, x_6, x_5, x_4, x_7$														
Important predictors [c]	$x_1, x_{11}, x_{15}, x_{14}, x_{13}, x_{12}$														

Data no. 20															
Predictor variables [a]	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_1	x_3	x_6	x_4	x_5	x_7	x_8	x_9	x_{10}	x_2
Significance predictors [b]	$x_1, x_{12}, x_{14}, x_{11}, x_{13}, x_{15}, x_6, x_3, x_4, x_8, x_9, x_5$														
Important predictors [c]	$x_{11}, x_1, x_{13}, x_{14}, x_{15}, x_{12}$														

Data no. 21															
Predictor variables [a]	x_{11}	x_{13}	x_{14}	x_{12}	x_1	x_{15}	x_5	x_3	x_7	x_8	x_2	x_4	x_9	x_{10}	x_6
Significance predictors [b]	$x_1, x_{13}, x_{14}, x_{11}, x_{12}, x_8, x_3, x_7, x_{15}, x_5, x_{10}, x_2, x_9$														
Important predictors [c]	$x_1, x_{11}, x_{13}, x_{14}, x_{12}, x_3, x_{15}$														

Data no. 22															
Predictor variables [a]	x_1	x_{13}	x_{15}	x_{14}	x_{12}	x_{11}	x_6	x_3	x_2	x_4	x_7	x_5	x_8	x_9	x_{10}
Significance predictors [b]	$x_1, x_{15}, x_7, x_{13}, x_6, x_{14}, x_{12}, x_{11}, x_2, x_3, x_4, x_8$														
Important predictors [c]	$x_1, x_7, x_{13}, x_{15}, x_{14}, x_{12}, x_{11}, x_6, x_3$														

Data no. 23															
Predictor variables [a]	x_{11}	x_{13}	x_{15}	x_1	x_7	x_{14}	x_{12}	x_3	x_4	x_5	x_9	x_8	x_{10}	x_2	x_6
Significance predictors [b]	$x_1, x_7, x_{13}, x_{15}, x_{11}, x_{14}, x_3, x_{12}, x_4, x_5, x_9, x_8, x_6, x_{10}$														
Important predictors [c]	$x_1, x_{11}, x_{13}, x_7, x_{15}, x_3, x_{14}$														

Data no. 24															
Predictor variables [a]	x_1	x_{11}	x_{14}	x_{13}	x_{15}	x_4	x_3	x_{12}	x_7	x_5	x_2	x_{10}	x_8	x_6	x_9
Significance predictors [b]	$x_1, x_{11}, x_{14}, x_4, x_{13}, x_{15}, x_7, x_5, x_3, x_{12}, x_{10}, x_2, x_9$														
Important predictors [c]	$x_1, x_{11}, x_4, x_{14}, x_{13}, x_3, x_5, x_{15}$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \dots ,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 25															
Predictor variables [a]	x_1	x_{11}	x_{15}	x_4	x_{12}	x_{13}	x_7	x_{14}	x_{10}	x_6	x_3	x_5	x_9	x_8	x_2
Significance predictors [b]	$x_1, x_7, x_4, x_{11}, x_{15}, x_{10}, x_{13}, x_{12}, x_{14}, x_6, x_5, x_9$														
Important predictors [c]	$x_1, x_7, x_{11}, x_4, x_{15}, x_{13}, x_{12}, x_{14}$														

Data no. 26															
Predictor variables [a]	x_{11}	x_{13}	x_{15}	x_1	x_{12}	x_{14}	x_2	x_4	x_7	x_5	x_3	x_8	x_9	x_{10}	x_6
Significance predictors [b]	$x_1, x_7, x_{15}, x_{13}, x_{11}, x_{12}, x_2, x_8, x_4, x_{14}, x_3, x_5, x_9$														
Important predictors [c]	$x_1, x_7, x_{11}, x_{15}, x_{13}, x_{12}, x_{14}, x_2, x_8, x_4$														

Data no. 27															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_2	x_4	x_9	x_5	x_3	x_{10}	x_6	x_7	x_8
Significance predictors [b]	$x_1, x_{11}, x_{13}, x_{15}, x_{14}, x_6, x_2, x_{12}, x_4, x_{10}, x_9, x_5, x_3, x_7$														
Important predictors [c]	$x_1, x_{11}, x_{13}, x_{15}, x_{14}, x_{12}, x_2, x_4$														

Data no. 28															
Predictor variables [a]	x_{11}	x_{13}	x_{12}	x_1	x_{14}	x_{15}	x_7	x_4	x_5	x_{10}	x_2	x_6	x_3	x_9	x_8
Significance predictors [b]	$x_1, x_7, x_{12}, x_{13}, x_{11}, x_{14}, x_{15}, x_4, x_5, x_2, x_{10}, x_6, x_3, x_9$														
Important predictors [c]	$x_1, x_{11}, x_{12}, x_{13}, x_{14}, x_7, x_{15}, x_4, x_5$														

Data no. 29															
Predictor variables [a]	x_{11}	x_{14}	x_{15}	x_{12}	x_{13}	x_1	x_3	x_2	x_4	x_7	x_5	x_9	x_6	x_{10}	x_8
Significance predictors [b]	$x_{14}, x_{12}, x_{15}, x_{13}, x_{11}, x_1, x_3, x_7, x_2, x_4, x_8, x_5, x_6, x_{10}$														
Important predictors [c]	$x_1, x_{11}, x_{15}, x_7, x_{14}, x_{12}, x_{13}$														

Data no. 30															
Predictor variables [a]	x_1	x_{11}	x_{14}	x_{13}	x_{15}	x_2	x_7	x_{12}	x_9	x_4	x_{10}	x_3	x_8	x_5	x_6
Significance predictors [b]	$x_1, x_7, x_{11}, x_2, x_{14}, x_{13}, x_{15}, x_{12}, x_5, x_9, x_4, x_{10}, x_3$														
Important predictors [c]	$x_1, x_7, x_{11}, x_2, x_{14}, x_{13}, x_{15}, x_{12}$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 31															
Predictor variables [a]	x_1	x_{13}	x_{11}	x_{14}	x_7	x_{15}	x_{12}	x_2	x_3	x_5	x_9	x_{10}	x_6	x_4	x_8
Significance predictors [b]	$x_7, x_1, x_2, x_{11}, x_{14}, x_{13}, x_{15}, x_{12}, x_5, x_6, x_3, x_9$														
Important predictors [c]	$x_1, x_7, x_{13}, x_2, x_{11}, x_{14}, x_{15}, x_{12}$														

Data no. 32															
Predictor variables [a]	x_1	x_{13}	x_{11}	x_{14}	x_{12}	x_{15}	x_4	x_6	x_9	x_8	x_3	x_5	x_{10}	x_7	x_2
Significance predictors [b]	$x_1, x_{13}, x_{14}, x_{11}, x_{12}, x_{15}, x_4, x_3, x_7$														
Important predictors [c]	$x_1, x_{13}, x_{14}, x_{11}, x_4, x_{12}, x_{15}$														

Data no. 33															
Predictor variables [a]	x_1	x_{11}	x_{12}	x_{14}	x_{13}	x_{15}	x_7	x_3	x_4	x_5	x_8	x_9	x_{10}	x_6	x_2
Significance predictors [b]	$x_1, x_7, x_{11}, x_{12}, x_{14}, x_{13}, x_{15}, x_3, x_4, x_8, x_6, x_5, x_{10}$														
Important predictors [c]	$x_1, x_7, x_{11}, x_{12}, x_{14}, x_{13}, x_{15}, x_4$														

Data no. 34															
Predictor variables [a]	x_{14}	x_{11}	x_1	x_{15}	x_{13}	x_4	x_{12}	x_2	x_6	x_5	x_{10}	x_9	x_7	x_3	x_8
Significance predictors [b]	$x_1, x_4, x_{14}, x_{15}, x_{11}, x_{13}, x_{12}, x_6, x_2, x_7, x_5, x_3$														
Important predictors [c]	$x_1, x_4, x_{14}, x_{11}, x_{15}$														

Data no. 35															
Predictor variables [a]	x_7	x_{11}	x_{13}	x_{12}	x_1	x_{14}	x_3	x_{15}	x_9	x_4	x_{10}	x_2	x_6	x_8	x_5
Significance predictors [b]	$x_7, x_1, x_{11}, x_{13}, x_{12}, x_3, x_{14}, x_9, x_{15}, x_5, x_4, x_6, x_{10}, x_2$														
Important predictors [c]	$x_7, x_{11}, x_1, x_{13}, x_3, x_{14}, x_5$														

Data no. 36															
Predictor variables [a]	x_1	x_{11}	x_{14}	x_{15}	x_{12}	x_{13}	x_4	x_{10}	x_2	x_9	x_7	x_5	x_8	x_3	x_6
Significance predictors [b]	$x_1, x_{11}, x_{14}, x_{15}, x_{12}, x_{10}, x_4, x_{13}, x_6, x_9, x_2, x_3, x_8$														
Important predictors [c]	$x_1, x_{11}, x_{15}, x_{14}, x_{12}, x_{13}$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 37															
Predictor variables [a]	x_1	x_{12}	x_{13}	x_{14}	x_{15}	x_{11}	x_4	x_3	x_7	x_9	x_5	x_2	x_{10}	x_8	x_6
Significance predictors [b]	$x_1, x_{13}, x_{12}, x_7, x_{14}, x_{11}, x_4, x_{15}, x_3, x_8, x_5, x_6, x_2$														
Important predictors [c]	$x_1, x_3, x_{12}, x_{13}, x_{11}, x_7, x_{14}, x_{15}$														

Data no. 38															
Predictor variables [a]	x_{13}	x_{11}	x_{15}	x_{14}	x_1	x_{12}	x_5	x_7	x_3	x_4	x_{10}	x_2	x_9	x_8	x_6
Significance predictors [b]	$x_1, x_{11}, x_7, x_{15}, x_{13}, x_{14}, x_3, x_{12}, x_5, x_4, x_9$														
Important predictors [c]	$x_{11}, x_{13}, x_1, x_{15}, x_7, x_{14}$														

Data no. 39															
Predictor variables [a]	x_{14}	x_{13}	x_{11}	x_{12}	x_1	x_{15}	x_7	x_{10}	x_4	x_5	x_3	x_8	x_9	x_2	x_6
Significance predictors [b]	$x_1, x_{14}, x_{11}, x_7, x_{13}, x_{12}, x_{15}, x_{10}, x_4, x_3, x_8, x_5$														
Important predictors [c]	$x_1, x_{14}, x_{13}, x_{11}, x_{12}, x_7, x_{15}$														

Data no. 40															
Predictor variables [a]	x_1	x_3	x_{11}	x_{13}	x_{14}	x_{15}	x_7	x_{10}	x_{12}	x_4	x_6	x_5	x_9	x_2	x_8
Significance predictors [b]	$x_1, x_3, x_{13}, x_7, x_{11}, x_{14}, x_{15}, x_{10}, x_5, x_{12}, x_4, x_6$														
Important predictors [c]	x_1, x_3, x_{11}, x_7														

Data no. 41															
Predictor variables [a]	x_{11}	x_{13}	x_1	x_{12}	x_{14}	x_{15}	x_7	x_3	x_5	x_4	x_9	x_{10}	x_6	x_2	x_8
Significance predictors [b]	$x_1, x_7, x_{11}, x_{13}, x_{12}, x_{14}, x_3, x_{15}, x_5, x_9, x_4, x_6, x_{10}$														
Important predictors [c]	$x_1, x_7, x_{11}, x_{13}, x_{12}, x_{15}, x_{14}, x_3$														

Data no. 42															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	x_{12}	x_8	x_9	x_6	x_2	x_5	x_3	x_{10}
Significance predictors [b]	$x_1, x_{11}, x_{14}, x_4, x_{13}, x_{15}, x_{12}, x_7, x_9, x_8, x_6, x_{10}, x_3, x_5, x_2$														
Important predictors [c]	$x_1, x_{11}, x_{13}, x_4, x_{14}$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \dots ,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 43															
Predictor variables [a]	x_{11}	x_{15}	x_1	x_{13}	x_{12}	x_4	x_{14}	x_7	x_3	x_8	x_9	x_2	x_{10}	x_5	x_6
Significance predictors [b]	$x_1, x_{11}, x_7, x_{15}, x_{13}, x_4, x_{12}, x_3, x_{14}, x_5, x_8, x_6, x_{10}, x_9$														
Important predictors [c]	$x_1, x_{11}, x_{15}, x_{13}, x_4, x_{12}, x_7$														

Data no. 44															
Predictor variables [a]	x_{11}	x_{13}	x_1	x_{12}	x_{15}	x_{14}	x_5	x_3	x_7	x_4	x_{10}	x_6	x_8	x_2	x_9
Significance predictors [b]	$x_1, x_{11}, x_{13}, x_{12}, x_{15}, x_{14}, x_3, x_7, x_5, x_4, x_{10}, x_8, x_2$														
Important predictors [c]	$x_1, x_{11}, x_{13}, x_{12}, x_{15}, x_{14}$														

Data no. 45															
Predictor variables [a]	x_{11}	x_{15}	x_{12}	x_{14}	x_1	x_3	x_{13}	x_{10}	x_2	x_7	x_9	x_5	x_4	x_6	x_8
Significance predictors [b]	$x_1, x_{11}, x_3, x_{12}, x_{15}, x_{14}, x_{10}, x_{13}, x_4, x_2, x_9, x_5, x_7$														
Important predictors [c]	$x_1, x_3, x_{11}, x_{12}, x_{15}, x_{14}$														

Data no. 46															
Predictor variables [a]	x_{11}	x_1	x_{14}	x_{15}	x_{12}	x_{13}	x_3	x_4	x_5	x_6	x_7	x_9	x_{10}	x_2	x_8
Significance predictors [b]	$x_1, x_{11}, x_{14}, x_{15}, x_3, x_{12}, x_{13}, x_4, x_5, x_2, x_7$														
Important predictors [c]	$x_1, x_{11}, x_{14}, x_{15}, x_3, x_{12}$														

Data no. 47															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{15}	x_4	x_{12}	x_{14}	x_8	x_{10}	x_2	x_3	x_7	x_9	x_5	x_6
Significance predictors [b]	$x_1, x_{11}, x_{13}, x_4, x_{15}, x_8, x_7, x_{12}, x_{14}, x_5, x_2, x_{10}, x_3$														
Important predictors [c]	$x_1, x_{11}, x_{13}, x_{15}, x_4, x_7, x_{14}$														

Data no. 48															
Predictor variables [a]	x_1	x_{14}	x_{13}	x_{11}	x_3	x_{15}	x_{12}	x_5	x_7	x_2	x_8	x_9	x_4	x_{10}	x_6
Significance predictors [b]	$x_1, x_{14}, x_3, x_{13}, x_{11}, x_7, x_{15}, x_{12}, x_2, x_5, x_8, x_6, x_{10}$														
Important predictors [c]	$x_1, x_3, x_{14}, x_{13}, x_{11}$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 49															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_{12}	x_8	x_5	x_6	x_4	x_7	x_9	x_{10}	x_3	x_2
Significance predictors [b]	$x_1, x_{11}, x_{13}, x_6, x_7, x_8, x_{15}, x_{14}, x_9, x_4, x_{10}, x_5, x_{12}, x_3$														
Important predictors [c]	$x_1, x_{11}, x_{13}, x_{15}, x_{14}$														

Data no. 50															
Predictor variables [a]	x_{15}	x_{13}	x_{11}	x_{12}	x_1	x_4	x_{14}	x_7	x_6	x_5	x_3	x_8	x_9	x_2	x_{10}
Significance predictors [b]	$x_1, x_{13}, x_{15}, x_7, x_{11}, x_{12}, x_4, x_6, x_{14}, x_8, x_3, x_5, x_{10}$														
Important predictors [c]	$x_1, x_{15}, x_{11}, x_{13}, x_4, x_{12}, x_7, x_{14}$														

Data no. 51															
Predictor variables [a]	x_{11}	x_1	x_{13}	x_{14}	x_7	x_3	x_{15}	x_{12}	x_{10}	x_8	x_4	x_9	x_6	x_5	x_2
Significance predictors [b]	$x_7, x_1, x_{11}, x_3, x_{14}, x_{13}, x_{15}, x_{12}, x_5, x_{10}, x_8, x_4$														
Important predictors [c]	$x_1, x_7, x_{11}, x_3, x_{13}, x_{14}, x_{15}$														

Data no. 52															
Predictor variables [a]	x_{11}	x_{15}	x_{13}	x_{14}	x_1	x_4	x_{12}	x_3	x_7	x_2	x_5	x_{10}	x_9	x_8	x_6
Significance predictors [b]	$x_1, x_{15}, x_4, x_{13}, x_7, x_{11}, x_{14}, x_{12}, x_3, x_2, x_5$														
Important predictors [c]	$x_1, x_{11}, x_{15}, x_4, x_{13}, x_{14}, x_{12}$														

Data no. 53															
Predictor variables [a]	x_1	x_7	x_{13}	x_{15}	x_{14}	x_{11}	x_{12}	x_9	x_{10}	x_3	x_5	x_8	x_4	x_6	x_2
Significance predictors [b]	$x_1, x_{13}, x_7, x_{14}, x_{15}, x_{11}, x_{12}, x_{10}, x_3, x_9, x_5, x_8, x_2$														
Important predictors [c]	x_1, x_7														

Data no. 54															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_7	x_{15}	x_2	x_{14}	x_{12}	x_8	x_3	x_4	x_6	x_{10}	x_5	x_9
Significance predictors [b]	$x_1, x_7, x_{13}, x_2, x_{11}, x_{15}, x_{14}, x_{12}, x_5, x_8, x_3, x_6, x_9$														
Important predictors [c]	$x_1, x_7, x_{11}, x_2, x_{13}$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 55															
Predictor variables [a]	x_{13}	x_{11}	x_{12}	x_1	x_{14}	x_{15}	x_7	x_5	x_4	x_3	x_2	x_9	x_{10}	x_8	x_6
Significance predictors [b]	$x_1, x_7, x_{13}, x_{12}, x_{11}, x_{14}, x_{15}, x_4, x_5, x_2, x_9$														
Important predictors [c]	$x_1, x_{11}, x_{13}, x_{12}, x_7, x_{14}, x_{15}$														

Data no. 56															
Predictor variables [a]	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_4	x_1	x_7	x_3	x_2	x_5	x_6	x_{10}	x_9	x_8
Significance predictors [b]	$x_7, x_1, x_4, x_{12}, x_{13}, x_{14}, x_{15}, x_{11}, x_2, x_3, x_5, x_6$														
Important predictors [c]	$x_1, x_4, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_7$														

Data no. 57															
Predictor variables [a]	x_{11}	x_3	x_1	x_7	x_{13}	x_{12}	x_{14}	x_2	x_{15}	x_4	x_5	x_6	x_{10}	x_9	x_8
Significance predictors [b]	$x_1, x_7, x_{11}, x_3, x_5, x_{13}, x_{12}, x_2, x_{14}, x_{15}$														
Important predictors [c]	$x_1, x_7, x_3, x_{11}, x_{13}, x_{12}, x_5$														

Data no. 58															
Predictor variables [a]	x_1	x_{11}	x_{14}	x_{15}	x_{13}	x_{12}	x_4	x_7	x_{10}	x_5	x_3	x_8	x_6	x_9	x_2
Significance predictors [b]	$x_1, x_{11}, x_{15}, x_{14}, x_{13}, x_{12}, x_7, x_4, x_3, x_{10}, x_8, x_5$														
Important predictors [c]	$x_1, x_{11}, x_{15}, x_{14}, x_{13}, x_{12}, x_4$														

Data no. 59															
Predictor variables [a]	x_{11}	x_{15}	x_{13}	x_{12}	x_{14}	x_1	x_3	x_7	x_4	x_5	x_{10}	x_6	x_2	x_8	x_9
Significance predictors [b]	$x_7, x_1, x_{15}, x_{13}, x_{11}, x_{14}, x_{12}, x_4, x_3, x_5, x_{10}, x_2$														
Important predictors [c]	$x_7, x_1, x_{11}, x_{15}, x_{13}, x_{12}, x_{14}, x_3, x_4$														

Data no. 60															
Predictor variables [a]	x_1	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_7	x_5	x_9	x_4	x_3	x_{10}	x_6	x_8	x_2
Significance predictors [b]	$x_1, x_{12}, x_{11}, x_{13}, x_7, x_{14}, x_9, x_{15}, x_4, x_5, x_6, x_3, x_2, x_{10}$														
Important predictors [c]	$x_1, x_{11}, x_{12}, x_{13}, x_{14}, x_7, x_{15}$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \dots ,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 61															
Predictor variables [a]	x_{11}	x_{13}	x_{1}	x_{15}	x_{14}	x_{9}	x_{5}	x_{12}	x_{4}	x_{7}	x_{2}	x_{3}	x_{10}	x_{6}	x_{8}
Significance predictors [b]	$x_1, x_7, x_{11}, x_{15}, x_9, x_{13}, x_{14}, x_4, x_{12}, x_2, x_5, x_{10}, x_6, x_3$														
Important predictors [c]	$x_{11}, x_1, x_7, x_{13}, x_9$														

Data no. 62															
Predictor variables [a]	x_{11}	x_{1}	x_{13}	x_{15}	x_{14}	x_{9}	x_{12}	x_{5}	x_{8}	x_{7}	x_{4}	x_{6}	x_{2}	x_{3}	x_{10}
Significance predictors [b]	$x_1, x_{11}, x_{15}, x_{13}, x_{14}, x_7, x_9, x_5, x_{12}, x_4, x_8, x_6, x_3, x_2$														
Important predictors [c]	x_1, x_{11}, x_{13}														

Data no. 63															
Predictor variables [a]	x_{13}	x_{11}	x_{1}	x_{15}	x_{14}	x_{7}	x_{5}	x_{12}	x_{8}	x_{6}	x_{3}	x_{4}	x_{9}	x_{2}	x_{10}
Significance predictors [b]	$x_1, x_{13}, x_{11}, x_7, x_8, x_{14}, x_{15}, x_5, x_{12}, x_3, x_6, x_2$														
Important predictors [c]	$x_1, x_{13}, x_{11}, x_{15}, x_7, x_{14}$														

Data no. 64															
Predictor variables [a]	x_{1}	x_{11}	x_{15}	x_{14}	x_{12}	x_{13}	x_{4}	x_{7}	x_{6}	x_{3}	x_{10}	x_{8}	x_{5}	x_{2}	x_{9}
Significance predictors [b]	$x_1, x_7, x_{11}, x_4, x_{15}, x_{14}, x_{13}, x_{12}, x_6, x_2, x_5, x_{10}, x_3$														
Important predictors [c]	$x_1, x_7, x_{11}, x_4, x_{15}, x_{12}, x_{14}, x_{13}$														

Data no. 65															
Predictor variables [a]	x_{11}	x_{1}	x_{13}	x_{14}	x_{12}	x_{15}	x_{4}	x_{8}	x_{2}	x_{7}	x_{9}	x_{5}	x_{3}	x_{10}	x_{6}
Significance predictors [b]	$x_1, x_{13}, x_{14}, x_{12}, x_{11}, x_7, x_4, x_{15}, x_8, x_9, x_2, x_5, x_3, x_{10}, x_6$														
Important predictors [c]	$x_1, x_{11}, x_{13}, x_{14}, x_{12}$														

Data no. 66															
Predictor variables [a]	x_{1}	x_{14}	x_{11}	x_{13}	x_{12}	x_{15}	x_{7}	x_{2}	x_{9}	x_{5}	x_{3}	x_{4}	x_{6}	x_{8}	x_{10}
Significance predictors [b]	$x_1, x_{14}, x_{11}, x_{13}, x_7, x_{12}, x_{15}, x_5, x_2, x_9, x_3, x_6, x_4, x_{10}$														
Important predictors [c]	$x_1, x_{14}, x_{11}, x_{13}$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 67															
Predictor variables [a]	x_1	x_{11}	x_{15}	x_{13}	x_{12}	x_7	x_{14}	x_5	x_6	x_2	x_4	x_9	x_3	x_{10}	x_8
Significance predictors [b]	$x_1, x_7, x_{11}, x_{15}, x_{13}, x_{12}, x_{14}, x_5, x_6, x_4, x_2, x_9, x_{10}$														
Important predictors [c]	$x_1, x_7, x_{11}, x_{15}, x_{13}, x_{12}, x_{14}, x_5$														

Data no. 68															
Predictor variables [a]	x_1	x_{11}	x_7	x_3	x_{15}	x_{13}	x_8	x_{14}	x_{12}	x_6	x_4	x_5	x_{10}	x_9	x_2
Significance predictors [b]	$x_1, x_7, x_3, x_{11}, x_8, x_{15}, x_{13}, x_{14}, x_{12}, x_4, x_5, x_9, x_{10}, x_2$														
Important predictors [c]	x_1, x_3, x_{11}, x_7														

Data no. 69															
Predictor variables [a]	x_1	x_{11}	x_7	x_{14}	x_{12}	x_{13}	x_4	x_{15}	x_2	x_6	x_{10}	x_8	x_3	x_5	x_9
Significance predictors [b]	$x_1, x_7, x_{11}, x_4, x_{12}, x_{14}, x_{13}, x_2, x_{15}, x_5, x_{10}, x_6, x_8, x_9$														
Important predictors [c]	$x_1, x_7, x_4, x_{11}, x_{14}, x_2, x_{12}, x_{13}$														

Data no. 70															
Predictor variables [a]	x_7	x_{12}	x_{13}	x_{11}	x_1	x_{14}	x_3	x_6	x_{15}	x_4	x_8	x_2	x_5	x_{10}	x_9
Significance predictors [b]	$x_1, x_7, x_6, x_{13}, x_{11}, x_{12}, x_3, x_{14}, x_5, x_4, x_{15}, x_2, x_8$														
Important predictors [c]	$x_7, x_1, x_{13}, x_{11}, x_{12}, x_6, x_3$														

Data no. 71															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_7	x_3	x_4	x_2	x_8	x_6	x_9	x_{10}	x_5
Significance predictors [b]	$x_1, x_7, x_{11}, x_{13}, x_{15}, x_{14}, x_{12}, x_4, x_3, x_5, x_2, x_8, x_6$														
Important predictors [c]	$x_1, x_{11}, x_7, x_{13}, x_{15}, x_{14}, x_{12}, x_4$														

Data no. 72															
Predictor variables [a]	x_{14}	x_{12}	x_{11}	x_{13}	x_1	x_3	x_{15}	x_4	x_{10}	x_5	x_7	x_2	x_9	x_8	x_6
Significance predictors [b]	$x_1, x_3, x_{14}, x_{12}, x_{11}, x_{13}, x_4, x_{15}, x_7, x_2, x_{10}$														
Important predictors [c]	$x_1, x_{14}, x_3, x_{11}, x_{12}, x_{13}, x_{15}, x_4$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 73															
Predictor variables [a]	x_{11}	x_{15}	x_{14}	x_{12}	x_1	x_{13}	x_4	x_3	x_7	x_2	x_5	x_8	x_{10}	x_9	x_6
Significance predictors [b]	$x_1, x_{11}, x_{15}, x_{14}, x_{12}, x_7, x_4, x_{13}, x_{10}, x_3, x_2, x_8, x_5$														
Important predictors [c]	$x_1, x_{11}, x_{15}, x_{14}, x_{12}, x_{13}, x_4, x_{10}, x_2$														

Data no. 74															
Predictor variables [a]	x_{15}	x_1	x_{13}	x_{11}	x_7	x_{14}	x_{12}	x_2	x_4	x_3	x_9	x_8	x_6	x_5	x_{10}
Significance predictors [b]	$x_7, x_{15}, x_1, x_{13}, x_{11}, x_{14}, x_3, x_5, x_{12}, x_2, x_4, x_9, x_8$														
Important predictors [c]	$x_1, x_7, x_{15}, x_3, x_{11}, x_{13}$														

Data no. 75															
Predictor variables [a]	x_1	x_7	x_3	x_{11}	x_{14}	x_{13}	x_{15}	x_{12}	x_8	x_4	x_2	x_{10}	x_9	x_5	x_6
Significance predictors [b]	$x_1, x_7, x_3, x_{11}, x_5, x_{14}, x_8, x_{13}, x_{15}, x_4, x_{12}, x_{10}, x_2$														
Important predictors [c]	$x_1, x_7, x_3, x_{11}, x_{14}, x_5, x_{13}, x_{15}, x_{12}$														

Data no. 76															
Predictor variables [a]	x_7	x_1	x_{11}	x_{14}	x_3	x_{15}	x_{13}	x_{12}	x_8	x_6	x_2	x_5	x_{10}	x_9	x_4
Significance predictors [b]	$x_1, x_7, x_3, x_{14}, x_{11}, x_8, x_{15}, x_{13}, x_{12}, x_2, x_6, x_5, x_{10}$														
Important predictors [c]	x_1, x_7														

Data no. 77															
Predictor variables [a]	x_{11}	x_{13}	x_1	x_{15}	x_{12}	x_{14}	x_3	x_2	x_7	x_5	x_4	x_6	x_{10}	x_9	x_8
Significance predictors [b]	$x_1, x_{11}, x_{15}, x_{13}, x_3, x_{14}, x_{12}, x_5, x_7, x_2, x_8, x_4, x_{10}, x_9$														
Important predictors [c]	x_1, x_{11}, x_{13}, x_3														

Data no. 78															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_7	x_{15}	x_2	x_{12}	x_{14}	x_3	x_4	x_9	x_6	x_{10}	x_8	x_5
Significance predictors [b]	$x_1, x_7, x_{11}, x_{13}, x_2, x_3, x_{15}, x_{12}, x_4, x_{14}, x_5, x_6, x_8, x_9$														
Important predictors [c]	$x_1, x_7, x_{11}, x_{13}, x_2, x_{15}, x_{12}, x_4, x_{14}, x_3, x_5$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 79															
Predictor variables [a]	x_{11}	x_{13}	x_{1}	x_{15}	x_{12}	x_{14}	x_{3}	x_{4}	x_{5}	x_{2}	x_{10}	x_{7}	x_{9}	x_{8}	x_{6}
Significance predictors [b]	$x_1, x_{11}, x_{13}, x_{15}, x_{12}, x_{14}, x_3, x_4, x_5, x_2, x_{10}, x_8$														
Important predictors [c]	$x_1, x_{11}, x_{13}, x_{15}, x_{12}$														

Data no. 80															
Predictor variables [a]	x_{15}	x_{12}	x_{13}	x_{11}	x_{1}	x_{14}	x_{7}	x_{4}	x_{8}	x_{5}	x_{2}	x_{10}	x_{3}	x_{6}	x_{9}
Significance predictors [b]	$x_1, x_{11}, x_7, x_{15}, x_{12}, x_{13}, x_4, x_{10}, x_{14}, x_8, x_3, x_2, x_5, x_6$														
Important predictors [c]	$x_1, x_{15}, x_{13}, x_{12}, x_{11}, x_7, x_4$														

Data no. 81															
Predictor variables [a]	x_{11}	x_{15}	x_{13}	x_{14}	x_{1}	x_{3}	x_{12}	x_{10}	x_{9}	x_{4}	x_{8}	x_{2}	x_{5}	x_{7}	x_{6}
Significance predictors [b]	$x_1, x_3, x_{11}, x_{13}, x_{15}, x_7, x_{14}, x_{10}, x_{12}, x_9, x_4, x_2, x_8, x_5$														
Important predictors [c]	$x_1, x_{11}, x_3, x_{13}, x_{15}, x_{14}, x_{12}, x_{10}, x_4, x_9$														

Data no. 82															
Predictor variables [a]	x_{1}	x_{13}	x_{11}	x_{15}	x_{14}	x_{12}	x_{7}	x_{3}	x_{4}	x_{2}	x_{8}	x_{6}	x_{10}	x_{5}	x_{9}
Significance predictors [b]	$x_1, x_{13}, x_{11}, x_7, x_{15}, x_{12}, x_{14}, x_3, x_8, x_2, x_4, x_6, x_5$														
Important predictors [c]	x_1, x_{13}, x_{11}, x_7														

Data no. 83															
Predictor variables [a]	x_{1}	x_{13}	x_{14}	x_{11}	x_{15}	x_{4}	x_{12}	x_{3}	x_{9}	x_{7}	x_{8}	x_{6}	x_{5}	x_{10}	x_{2}
Significance predictors [b]	$x_1, x_{13}, x_{11}, x_{14}, x_{15}, x_4, x_7, x_3, x_{12}, x_9, x_5, x_8$														
Important predictors [c]	$x_1, x_{13}, x_{14}, x_{11}, x_4, x_{15}, x_7, x_3, x_{12}$														

Data no.84															
Predictor variables [a]	x_{1}	x_{7}	x_{11}	x_{13}	x_{6}	x_{12}	x_{4}	x_{3}	x_{14}	x_{15}	x_{2}	x_{9}	x_{8}	x_{10}	x_{5}
Significance predictors [b]	$x_1, x_7, x_3, x_{11}, x_6, x_{12}, x_{13}, x_4, x_5, x_{14}, x_{15}, x_2, x_{10}$														
Important predictors [c]	$x_1, x_7, x_3, x_{11}, x_{13}, x_6, x_{12}$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 85															
Predictor variables [a]	x_{11}	x_4	x_{13}	x_{15}	x_{14}	x_{12}	x_1	x_2	x_9	x_5	x_{10}	x_7	x_3	x_6	x_8
Significance predictors [b]	$x_4, x_1, x_{15}, x_{11}, x_{13}, x_{12}, x_{14}, x_9, x_5, x_2, x_7, x_{10}$														
Important predictors [c]	$x_4, x_1, x_{11}, x_{13}, x_{15}, x_{12}, x_{14}$														

Data no. 86															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{12}	x_{15}	x_{14}	x_4	x_8	x_5	x_9	x_7	x_2	x_6	x_{10}	x_3
Significance predictors [b]	$x_1, x_{11}, x_{13}, x_{12}, x_4, x_{15}, x_8, x_{14}, x_7, x_9, x_2, x_3, x_5, x_6, x_{10}$														
Important predictors [c]	$x_1, x_{11}, x_4, x_{12}, x_{13}, x_{15}, x_{14}$														

Data no. 87															
Predictor variables [a]	x_{11}	x_1	x_{13}	x_7	x_3	x_{15}	x_4	x_{12}	x_{10}	x_5	x_8	x_9	x_{14}	x_6	x_2
Significance predictors [b]	$x_1, x_7, x_{13}, x_{11}, x_3, x_5, x_4, x_{15}, x_{12}, x_{14}, x_{10}, x_8$														
Important predictors [c]	$x_1, x_{11}, x_{13}, x_3, x_4, x_{14}$														

Data no. 88															
Predictor variables [a]	x_{11}	x_{13}	x_1	x_{15}	x_{12}	x_7	x_{14}	x_3	x_4	x_5	x_9	x_8	x_{10}	x_6	x_2
Significance predictors [b]	$x_7, x_1, x_{13}, x_{11}, x_{15}, x_3, x_{12}, x_{14}, x_4, x_5, x_2$														
Important predictors [c]	$x_1, x_7, x_{11}, x_{13}, x_{15}, x_{12}, x_3, x_{14}$														

Data no. 89															
Predictor variables [a]	x_1	x_{11}	x_{12}	x_{13}	x_{15}	x_{14}	x_4	x_3	x_7	x_8	x_{10}	x_6	x_5	x_9	x_2
Significance predictors [b]	$x_1, x_{11}, x_{12}, x_{15}, x_{13}, x_7, x_4, x_{14}, x_3, x_9, x_{10}, x_5, x_6, x_8$														
Important predictors [c]	$x_1, x_{11}, x_{12}, x_{15}, x_{13}, x_{14}$														

Data no. 90															
Predictor variables [a]	x_1	x_{13}	x_{11}	x_{12}	x_7	x_{14}	x_{15}	x_5	x_2	x_8	x_9	x_3	x_{10}	x_4	x_6
Significance predictors [b]	$x_1, x_7, x_{13}, x_{11}, x_{12}, x_{14}, x_{15}, x_5, x_2, x_8, x_3, x_9, x_4, x_{10}$														
Important predictors [c]	$x_1, x_7, x_{13}, x_{11}, x_{12}, x_{14}, x_{15}$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 91															
Predictor variables [a]	x_7	x_1	x_{13}	x_{11}	x_{12}	x_{15}	x_{14}	x_6	x_5	x_4	x_2	x_8	x_9	x_3	x_{10}
Significance predictors [b]	$x_1, x_7, x_{11}, x_{13}, x_{12}, x_{15}, x_{14}, x_2, x_5, x_6, x_4, x_9$														
Important predictors [c]	$x_7, x_1, x_{11}, x_{13}, x_{15}, x_{12}$														

Data no. 92															
Predictor variables [a]	x_1	x_7	x_{13}	x_{11}	x_3	x_{14}	x_{15}	x_{12}	x_4	x_2	x_5	x_8	x_6	x_{10}	x_9
Significance predictors [b]	$x_1, x_7, x_{13}, x_3, x_{11}, x_{14}, x_{15}, x_{12}, x_4, x_5, x_2, x_{10}, x_8$														
Important predictors [c]	$x_1, x_7, x_3, x_{13}, x_{11}, x_{14}, x_{15}$														

Data no. 93															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_{12}	x_2	x_3	x_{10}	x_4	x_6	x_5	x_8	x_9	x_7
Significance predictors [b]	$x_1, x_{11}, x_{13}, x_{10}, x_{14}, x_{15}, x_2, x_3, x_{12}, x_4, x_6, x_5, x_7$														
Important predictors [c]	$x_1, x_{11}, x_{13}, x_{14}, x_{15}, x_{10}$														

Data no. 94															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_4	x_{14}	x_{12}	x_{15}	x_7	x_9	x_3	x_{10}	x_6	x_8	x_5	x_2
Significance predictors [b]	$x_1, x_7, x_{11}, x_{13}, x_4, x_{14}, x_{12}, x_3, x_{15}, x_9, x_5, x_6, x_{10}$														
Important predictors [c]	x_1, x_{11}, x_7, x_4														

Data no. 95															
Predictor variables [a]	x_1	x_{12}	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	x_9	x_5	x_3	x_8	x_6	x_2	x_{10}
Significance predictors [b]	$x_1, x_{12}, x_{13}, x_{11}, x_{14}, x_7, x_4, x_{15}, x_8, x_5, x_3, x_9, x_{10}$														
Important predictors [c]	$x_1, x_{12}, x_{11}, x_{13}, x_4, x_{14}$														

Data no. 96															
Predictor variables [a]	x_1	x_7	x_{11}	x_{13}	x_2	x_{12}	x_{14}	x_6	x_4	x_9	x_{15}	x_{10}	x_8	x_3	x_5
Significance predictors [b]	$x_7, x_1, x_2, x_{13}, x_{11}, x_6, x_{12}, x_4, x_{14}, x_9, x_{15}, x_5, x_{10}, x_3$														
Important predictors [c]	$x_1, x_7, x_2, x_{11}, x_{13}$														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\dots,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Table H Set of significance and important predictors for 100 datasets (Cont.)

Data no. 97															
Predictor variables [a]	x_1	x_{11}	x_{12}	x_{13}	x_{15}	x_7	x_4	x_{14}	x_2	x_6	x_3	x_5	x_9	x_{10}	x_8
Significance predictors [b]	x_1, x_7, x_{11}, x_{13}, x_{15}, x_{12}, x_4, x_2, x_{14}, x_5, x_6, x_9														
Important predictors [c]	x_1, x_{11}, x_4, x_{12}, x_{13}, x_7, x_{15}														

Data no. 98															
Predictor variables [a]	x_{11}	x_{14}	x_1	x_{13}	x_{15}	x_3	x_8	x_{12}	x_4	x_7	x_{10}	x_6	x_9	x_5	x_2
Significance predictors [b]	x_{14}, x_1, x_3, x_{13}, x_{11}, x_{15}, x_8, x_{12}, x_7, x_9, x_4, x_5, x_6														
Important predictors [c]	x_1, x_3														

Data no. 99															
Predictor variables [a]	x_1	x_{12}	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	x_5	x_6	x_3	x_{10}	x_9	x_8	x_2
Significance predictors [b]	x_1, x_{12}, x_{11}, x_7, x_{13}, x_{14}, x_{15}, x_4, x_5, x_3, x_6, x_{10}														
Important predictors [c]	x_1, x_{12}, x_{11}, x_{13}, x_{14}, x_{15}, x_7														

Data no. 100															
Predictor variables [a]	x_1	x_{11}	x_{14}	x_{12}	x_{13}	x_{15}	x_7	x_5	x_{10}	x_4	x_6	x_9	x_8	x_3	x_2
Significance predictors [b]	x_1, x_{11}, x_7, x_{12}, x_{13}, x_{15}, x_{14}, x_5, x_{10}, x_9, x_4, x_3, x_6														
Important predictors [c]	x_1, x_{11}, x_7, x_5														

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.
[b] Significance predictors are in descending order according to the frequency of each significance predictor.
[c] Important predictors are in descending order according to the frequency of each important predictor.

Appendix I

Naïve elastic net estimates using different α values

This appendix presents the results of the naïve elastic net estimates ($\hat{\beta}$) using different α values where the values of λ_2 are estimated by 10-fold cross validation method (λ_2CV) and Bayes factor method (λ_2BF).

The 100 simulation datasets using simulation method for dataset 1 described in Chapter 3 are used for analysis the performance of the method for estimating the value of the penalty parameter λ_2 based on Bayes factor. Each dataset consists of 15 predictor variables of 50 observations. The dataset was fitted by the elastic net method using nine α values: 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9. Table I shows the results of the naïve elastic net estimates ($\hat{\beta}$) where the values of λ_2 are estimated by λ_2CV and λ_2BF for 100 datasets.

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF

Data no.1: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0468	0.0661	0.1134	0.0479	0.2001	0.2492	0.3218	0.1646	0.2553
λ_1		0.4212	0.2645	0.2645	0.0719	0.2001	0.1661	0.1379	0.0411	0.0284
Predictors	x_1	1.6885	1.7457	1.6570	1.8290	1.5508	1.4994	1.4255	1.6507	1.5370
	x_2	0	0.0810	0.0838	0.0170	0.1145	0.1261	0.1355	0.0853	0.1204
	x_3	-0	-0	-0	-0.2265	-0	-0.0175	-0.0443	-0.2016	-0.1853
	x_4	0.6968	0.7453	0.7167	0.5255	0.6665	0.6426	0.6180	0.5648	0.5783
	x_5	-0	-0.0475	-0.0646	-0.6055	-0.1774	-0.2289	-0.2599	-0.5492	-0.5135
	x_6	-0.3914	-0.5016	-0.4978	-0.6040	-0.5289	-0.5388	-0.5368	-0.6065	-0.5913
	x_7	0.4027	0.5096	0.4924	0.6275	0.5081	0.5090	0.4955	0.5863	0.5571
	x_8	0.4966	0.5926	0.5724	0.7171	0.5740	0.5762	0.5693	0.6862	0.6634
	x_9	-0	-0.0796	-0.0556	-0.2637	-0.0777	-0.0881	-0.0876	-0.2063	-0.1736
	x_{10}	0	0	0	0	0	0	0	0.0334	0.0634
	x_{11}	0.7197	0.7350	0.7070	0.9350	0.6827	0.6735	0.6566	0.8189	0.7470
	x_{12}	0.0905	0.1184	0.1542	0.2789	0.2208	0.2520	0.2829	0.3230	0.3282
	x_{13}	0.3933	0.4177	0.4258	0.5855	0.4502	0.4614	0.4685	0.5619	0.5366
	x_{14}	-0	-0	-0	-0.2327	-0	-0	-0	-0.1431	-0.0879
	x_{15}	0	0	0	-0	0	0	0	0	0.0294
df		8	11	11	13	11	12	12	14	15
PE		6.0551	5.5166	5.6517	4.6814	5.6603	5.6801	5.7944	4.9156	5.1600

Data no.1: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02799551.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0280	0.0280	0.0280	0.0280	0.0280	0.0280	0.0280	0.0280	0.0280
λ_1		0.2520	0.1120	0.0653	0.0420	0.0280	0.0187	0.0120	0.0070	0.0031
Predictors	x_1	1.8335	1.8740	1.8632	1.8628	1.8605	1.8542	1.8438	1.8359	1.8298
	x_2	0.0839	0.0502	0	0	-0	-0.0194	-0.0593	-0.0884	-0.1108
	x_3	-0	-0.1188	-0.2613	-0.3129	-0.3444	-0.3806	-0.4270	-0.4590	-0.4832
	x_4	0.7670	0.6144	0.4885	0.4415	0.4142	0.3845	0.3474	0.3213	0.3016
	x_5	-0.0520	-0.4437	-0.6663	-0.7521	-0.8072	-0.8652	-0.9342	-0.9804	-1.0151
	x_6	-0.5125	-0.5917	-0.6075	-0.6205	-0.6274	-0.6276	-0.6221	-0.6195	-0.6178
	x_7	0.5344	0.6340	0.6398	0.6501	0.6545	0.6510	0.6404	0.6322	0.6260
	x_8	0.6178	0.6976	0.7307	0.7435	0.7530	0.7633	0.7752	0.7829	0.7886
	x_9	-0.1163	-0.2467	-0.2831	-0.2981	-0.3028	-0.3063	-0.3105	-0.3127	-0.3143
	x_{10}	0	0	0	0	0.0145	0.0310	0.0482	0.0587	0.0665
	x_{11}	0.7660	0.9041	0.9804	1.0148	1.0348	1.0509	1.0661	1.0789	1.0891
	x_{12}	0.0868	0.2082	0.2848	0.3144	0.3326	0.3516	0.3742	0.3940	0.4097
	x_{13}	0.4120	0.5382	0.6050	0.6314	0.6484	0.6656	0.6854	0.7009	0.7130
	x_{14}	-0	-0.1972	-0.2745	-0.3159	-0.3399	-0.3534	-0.3600	-0.3646	-0.3683
	x_{15}	0	-0	-0	-0	-0	-0	-0	-0.0079	-0.0153
df		11	13	12	12	13	14	14	15	15
PE		5.3724	4.7701	4.6166	4.5660	4.5422	4.5258	4.5114	4.5018	4.4953

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.2: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0149	0.0368	0.0435	0.0512	0.1015	0.1049	0.1235	0.1006	0.1559
λ_1		0.1342	0.1472	0.1015	0.0768	0.1015	0.0699	0.0529	0.0251	0.0173
Predictors	x_1	1.9131	1.8596	1.8841	1.8823	1.7714	1.7847	1.763	1.8132	1.7308
	x_2	-0	-0	-0	-0	-0	-0	-0	-0.0306	-0.0140
	x_3	0.9807	0.9473	1.0365	1.0791	0.9847	1.0375	1.049	1.1097	1.0709
	x_4	0.0318	0.0553	0.0408	0.0290	0.0763	0.0632	0.0662	0.0321	0.0726
	x_5	-0.4896	-0.4530	-0.4947	-0.5198	-0.4493	-0.4829	-0.4885	-0.5505	-0.5046
	x_6	-0.0157	-0.0187	-0.0687	-0.1025	-0.1010	-0.1360	-0.1605	-0.1858	-0.2072
	x_7	1.5216	1.4828	1.4920	1.4890	1.4064	1.4148	1.3979	1.4429	1.3769
	x_8	0.2134	0.1983	0.2564	0.2942	0.2557	0.2983	0.3181	0.3650	0.3583
	x_9	0.2994	0.2793	0.3533	0.3975	0.3377	0.3885	0.4088	0.4612	0.4471
	x_{10}	0	0	0.0446	0.0911	0.0279	0.0822	0.1045	0.1646	0.1503
	x_{11}	0.2828	0.2558	0.3212	0.3494	0.2812	0.3141	0.3200	0.3781	0.3407
	x_{12}	-0	-0	-0	-0	-0	-0	-0	-0.0189	-0.0163
	x_{13}	0.8383	0.7907	0.8098	0.8043	0.7200	0.7244	0.7059	0.7510	0.6839
	x_{14}	0	0	0	0.0263	0.0472	0.0755	0.0996	0.1148	0.1421
	x_{15}	-0.1170	-0.0712	-0.1556	-0.2039	-0.1216	-0.1802	-0.1996	-0.2623	-0.2339
df		11	11	12	13	13	13	13	15	15
PE		3.9994	4.1240	3.9322	3.8519	4.1108	3.9834	3.9759	3.8084	3.9467

Data no.2: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.0358675.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0359	0.0359	0.0359	0.0359	0.0359	0.0359	0.0359	0.0359	0.0359
λ_1		0.3228	0.1435	0.0837	0.0538	0.0359	0.0239	0.0154	0.0090	0.0040
Predictors	x_1	1.6498	1.8648	1.9099	1.9212	1.9284	1.9340	1.9379	1.9301	1.9131
	x_2	0	-0	-0	-0.0199	-0.0486	-0.0665	-0.0793	-0.1031	-0.1361
	x_3	0.7995	0.9547	1.0810	1.1364	1.1641	1.1814	1.1936	1.1898	1.1734
	x_4	0.1771	0.0537	0.0198	0	0	0	-0	-0.0314	-0.0876
	x_5	-0.1002	-0.4575	-0.5264	-0.5718	-0.5939	-0.6072	-0.6166	-0.6545	-0.7154
	x_6	-0	-0.0215	-0.0846	-0.1209	-0.1472	-0.1673	-0.1816	-0.1931	-0.2028
	x_7	1.3543	1.4857	1.5109	1.5221	1.5339	1.5431	1.5496	1.5541	1.5572
	x_8	0.0626	0.2025	0.2842	0.3340	0.3666	0.3894	0.4056	0.4204	0.4344
	x_9	0.0445	0.2845	0.3905	0.4448	0.4740	0.4941	0.5085	0.5206	0.5314
	x_{10}	-0	0	0.0840	0.1431	0.1801	0.2065	0.2252	0.2409	0.2547
	x_{11}	0.0155	0.2612	0.3555	0.3956	0.4298	0.4587	0.4792	0.5022	0.5276
	x_{12}	0	-0	-0	-0	-0.0300	-0.0624	-0.0856	-0.0937	-0.0908
	x_{13}	0.7305	0.7950	0.8308	0.8425	0.8551	0.8643	0.8710	0.8793	0.8889
	x_{14}	0	0	0.0025	0.0358	0.0574	0.0717	0.0821	0.0936	0.1066
	x_{15}	-0	-0.0795	-0.1995	-0.2689	-0.3010	-0.3184	-0.3307	-0.3387	-0.3435
df		9	11	13	13	14	14	14	15	15
PE		5.0008	4.1030	3.8380	3.7348	3.6784	3.6461	3.6276	3.6130	3.6004

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.3: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0072	0.0194	0.0303	0.0623	0.0935	0.1061	0.0861	0.1116	0.1900
λ_1		0.0645	0.0776	0.0707	0.0935	0.0935	0.0707	0.0369	0.0279	0.0211
Predictors	x_1	1.6401	1.6643	1.6625	1.6682	1.6467	1.6242	1.6222	1.6014	1.5346
	x_2	-0.2354	-0.2012	-0.1827	-0.13208	-0.1137	-0.1233	-0.1513	-0.1470	-0.1381
	x_3	0.1471	0.1619	0.1951	0.2111	0.2329	0.2619	0.2904	0.3096	0.3243
	x_4	0	0.0475	0.0995	0.1798	0.2287	0.2508	0.2412	0.2742	0.3239
	x_5	-1.2127	-1.0777	-1.0224	-0.8068	-0.7127	-0.7408	-0.8794	-0.8315	-0.6986
	x_6	-0.3162	-0.2649	-0.2688	-0.1863	-0.1630	-0.1994	-0.2842	-0.2782	-0.2318
	x_7	1.0963	1.0550	1.0454	0.9711	0.9347	0.9477	1.0098	0.9903	0.9204
	x_8	0.4518	0.4275	0.4305	0.3801	0.3599	0.3791	0.4346	0.4270	0.3824
	x_9	-0	-0	-0	-0	-0	-0	-0	-0	-0
	x_{10}	0.4292	0.4150	0.4152	0.3863	0.3767	0.3909	0.4234	0.4232	0.4058
	x_{11}	1.2463	1.1618	1.1341	0.9882	0.9186	0.9320	1.0337	0.9900	0.8634
	x_{12}	-0	-0	-0	-0	-0	-0	-0	-0	-0
	x_{13}	0.5948	0.5551	0.5448	0.4734	0.4458	0.4592	0.5131	0.4980	0.4477
	x_{14}	-0.2446	-0.2035	-0.2196	-0.1541	-0.1363	-0.1739	-0.2572	-0.2524	-0.2025
	x_{15}	0.03810	0.0110	0.0200	0	0.0036	0.0366	0.0736	0.0892	0.1120
df		12	13	13	13	13	13	13	13	13
PE		3.9186	4.0062	4.0135	4.2283	4.3405	4.2703	4.0642	4.1128	4.3363

Data no.3: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.04932405.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0493	0.0493	0.0493	0.0493	0.0493	0.0493	0.0493	0.0493	0.0493
λ_1		0.4439	0.1973	0.1151	0.0740	0.0493	0.0329	0.0211	0.0123	0.0055
Predictors	x_1	1.5453	1.6796	1.6745	1.6649	1.6442	1.6302	1.6203	1.6110	1.6036
	x_2	-0	-0.0869	-0.1407	-0.1535	-0.1708	-0.1825	-0.1909	-0.1947	-0.1978
	x_3	0	0.0645	0.1611	0.2221	0.2472	0.2638	0.2756	0.2929	0.3064
	x_4	0.2279	0.1058	0.1273	0.1609	0.1659	0.1689	0.1711	0.1756	0.1788
	x_5	-0	-0.6075	-0.8186	-0.9090	-0.9855	-1.0369	-1.0735	-1.1050	-1.1298
	x_6	0	-0	-0.1506	-0.2408	-0.2965	-0.3337	-0.3602	-0.3821	-0.3988
	x_7	0.6277	0.8744	0.9650	1.0118	1.0440	1.0655	1.0808	1.0874	1.0923
	x_8	0	0.2488	0.3581	0.4143	0.4455	0.4663	0.4812	0.4933	0.5027
	x_9	-0	-0	-0	-0	-0	0	0	0.0217	0.0387
	x_{10}	0.0179	0.2952	0.3736	0.4059	0.4247	0.4372	0.4462	0.4595	0.4700
	x_{11}	0.6151	0.8271	0.9803	1.0629	1.1141	1.1482	1.1726	1.1911	1.2057
	x_{12}	0	-0	-0	-0	-0	-0	-0	-0	-0.0010
	x_{13}	0.1749	0.3476	0.4573	0.5131	0.5427	0.5625	0.5766	0.5884	0.5975
	x_{14}	-0	-0	-0.1148	-0.2035	-0.2605	-0.2984	-0.3255	-0.3464	-0.3622
	x_{15}	0	0	0	0.0152	0.0529	0.0780	0.0960	0.1082	0.1181
df		6	10	12	13	13	13	13	14	15
PE		6.1804	4.7990	4.2918	4.0910	3.9932	3.9405	3.9090	3.8869	3.8721

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.4: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0036	0.0354	0.0607	0.0540	0.1176	0.1464	0.1723	0.1690	0.1135
λ_1		0.0320	0.1416	0.1416	0.0810	0.1176	0.0976	0.0738	0.0423	0.0126
Predictors	x_1	2.1100	1.8066	1.7645	1.8175	1.6888	1.6574	1.6323	1.6534	1.7817
	x_2	0.0948	0	0	0	0	0	0	0	0.0332
	x_3	0.4343	0	0	0.0124	0	0	0	0.0124	0.1282
	x_4	1.2299	0.7843	0.7842	0.8204	0.7831	0.7834	0.7843	0.8039	0.9056
	x_5	-0.0178	-0.3597	-0.3335	-0.3973	-0.3197	-0.3243	-0.3352	-0.3599	-0.3232
	x_6	-0.5235	-0.4722	-0.4589	-0.5061	-0.4531	-0.4563	-0.4618	-0.4830	-0.5221
	x_7	0.9376	0.8853	0.8831	0.9155	0.8803	0.8820	0.8902	0.9146	0.9559
	x_8	0	-0.0679	-0.0992	-0.0774	-0.1460	-0.1638	-0.1782	-0.1765	-0.1285
	x_9	-0.0591	-0	-0	-0.0251	-0	-0.0059	-0.0273	-0.0562	-0.0845
	x_{10}	-0.1152	-0.0355	-0.0344	-0.0699	-0.0470	-0.0572	-0.0672	-0.0798	-0.1016
	x_{11}	1.1878	0.8917	0.8355	0.9306	0.7633	0.7419	0.7300	0.7623	0.8913
	x_{12}	-0.3982	-0	-0	-0.0756	-0	-0	-0	-0.0223	-0.1515
	x_{13}	0.3313	0.2730	0.2888	0.3374	0.3253	0.3433	0.3616	0.3827	0.4024
	x_{14}	-0.1506	-0	-0	-0	-0	0	0	0	-0.0214
	x_{15}	-0.4900	-0.2252	-0.1908	-0.2641	-0.1592	-0.1558	-0.1595	-0.1931	-0.2854
df		14	10	10	13	10	11	11	13	15
PE		3.4779	4.1008	4.1907	3.9231	4.3024	4.3278	4.3280	4.1991	3.8581

Data no.4: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02724963.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0272	0.0272	0.0272	0.0272	0.0272	0.0272	0.0272	0.0272	0.0272
λ_1		0.2452	0.1090	0.0636	0.0409	0.0272	0.0182	0.0117	0.0068	0.0030
Predictors	x_1	1.7816	1.8448	1.9131	1.9636	2.0018	2.0272	2.0451	2.0590	2.0696
	x_2	0	0	0	0.0055	0.0483	0.0768	0.0971	0.1127	0.1246
	x_3	0	0	0.1246	0.2222	0.3004	0.3518	0.3886	0.4167	0.4384
	x_4	0.7631	0.7994	0.9110	1.0070	1.0834	1.1342	1.1704	1.1984	1.2196
	x_5	-0.2099	-0.4083	-0.3371	-0.2596	-0.1831	-0.1321	-0.0957	-0.0675	-0.0463
	x_6	-0.3850	-0.5029	-0.5309	-0.5358	-0.5381	-0.5387	-0.5389	-0.5389	-0.5396
	x_7	0.8445	0.8967	0.9235	0.9388	0.9491	0.9564	0.9615	0.9653	0.9682
	x_8	-0.0855	-0.0455	-0.0212	-0.0088	-0.0016	-0	-0	-0	0.0021
	x_9	-0	-0	-0.0354	-0.0611	-0.0736	-0.0820	-0.0881	-0.0926	-0.0961
	x_{10}	-0	-0.0582	-0.0967	-0.1040	-0.1117	-0.1161	-0.1189	-0.1211	-0.1233
	x_{11}	0.7838	0.9613	1.0431	1.0890	1.1136	1.1290	1.1399	1.1482	1.1550
	x_{12}	-0	-0.0428	-0.1879	-0.2616	-0.3117	-0.3451	-0.3688	-0.3867	-0.4007
	x_{13}	0.1997	0.2954	0.3369	0.3598	0.3656	0.3699	0.3731	0.3752	0.3768
	x_{14}	-0	-0	-0.0148	-0.0674	-0.1031	-0.1269	-0.1438	-0.1569	-0.1668
	x_{15}	-0.0828	-0.2721	-0.3479	-0.3887	-0.4195	-0.4390	-0.4531	-0.4638	-0.4723
df		9	11	14	15	15	14	14	14	15
PE		4.5344	3.9438	3.7162	3.6100	3.5529	3.5314	3.5019	3.4887	3.4793

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.5: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0351	0.0719	0.1123	0.0074	0.0161	0.0066	0.0452	0.0085	0.1319
λ_1		0.3157	0.2877	0.2621	0.0111	0.0161	0.0044	0.0194	0.0021	0.0147
Predictors	x_1	2.2975	2.2302	2.1567	2.6876	2.6946	2.6650	2.6278	2.5068	2.3730
	x_2	-0	-0	-0	-0.8289	-0.7029	-0.9051	-0.5151	-0.3783	-0.3054
	x_3	0	0	0	-0.3654	-0.2164	-0.4743	-0.0665	-0.0124	-0.0279
	x_4	0	0	0	-0.9032	-0.6849	-1.0492	-0.4059	-0.2405	-0.1893
	x_5	-0	-0	-0	-1.3536	-1.0844	-1.5226	-0.7128	-0.4693	-0.3600
	x_6	0	0	0	0.0491	0.0604	0.0459	0.0863	0.1085	0.1253
	x_7	0.3734	0.4174	0.4532	0.5121	0.5029	0.5234	0.5051	0.5173	0.5357
	x_8	0	0	0.0232	0.1248	0.1280	0.1322	0.1549	0.1877	0.2242
	x_9	0	0	0	0.3433	0.3379	0.3471	0.3195	0.2936	0.2755
	x_{10}	-0.3167	-0.3270	-0.3283	-1.0243	-0.9732	-1.0501	-0.8698	-0.7701	-0.6996
	x_{11}	0.9529	0.9143	0.8683	1.7625	1.6631	1.8072	1.4613	1.2735	1.1357
	x_{12}	0	0	0.0198	0.3113	0.2530	0.3558	0.1906	0.1612	0.1637
	x_{13}	0.1185	0.1488	0.1661	0.4584	0.4145	0.4883	0.3541	0.3127	0.2976
	x_{14}	-0	-0	-0	-0.1334	-0.1525	-0.1156	-0.1517	-0.1265	-0.0959
	x_{15}	-0	-0	-0	-0.2311	-0.2314	-0.2273	-0.2054	-0.1625	-0.1251
df		5	5	7	15	15	15	15	15	15
PE		5.6168	5.6420	5.6853	3.9345	3.9919	3.9136	4.1578	4.3858	4.5974

Data no.5: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02417284.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242
λ_1		0.2176	0.0967	0.0564	0.0363	0.0242	0.0161	0.0104	0.0060	0.0027
Predictors	x_1	2.4060	2.6313	2.6889	2.7059	2.6945	2.6792	2.6683	2.6600	2.6536
	x_2	-0.0198	-0.2742	-0.4167	-0.5060	-0.5843	-0.6454	-0.6890	-0.7218	-0.7473
	x_3	-0	0	0	-0	-0.0821	-0.1668	-0.2270	-0.2724	-0.3078
	x_4	0	-0	-0.1856	-0.3346	-0.4803	-0.5977	-0.6814	-0.7444	-0.7934
	x_5	-0	-0.2590	-0.5076	-0.6713	-0.8352	-0.9682	-1.0631	-1.1344	-1.1899
	x_6	0.0351	0.0901	0.0841	0.0740	0.0702	0.0686	0.0673	0.0664	0.0658
	x_7	0.3959	0.4247	0.4591	0.4789	0.4932	0.5033	0.5108	0.5163	0.5205
	x_8	0.0248	0.0758	0.0981	0.1154	0.1278	0.1367	0.1431	0.1479	0.1516
	x_9	0.0638	0.2148	0.2761	0.3135	0.3278	0.3343	0.3390	0.3424	0.3452
	x_{10}	-0.4646	-0.7126	-0.8221	-0.8843	-0.9203	-0.9438	-0.9607	-0.9733	-0.9831
	x_{11}	1.0244	1.2998	1.4359	1.5147	1.5670	1.6036	1.6299	1.6495	1.6648
	x_{12}	0	0.0307	0.1036	0.1506	0.1966	0.2339	0.2603	0.2802	0.2957
	x_{13}	0.1252	0.2319	0.2961	0.3366	0.3706	0.3966	0.4153	0.4293	0.4401
	x_{14}	-0	-0.1421	-0.1699	-0.1764	-0.1662	-0.1546	-0.1461	-0.1398	-0.1349
	x_{15}	-0	-0.0865	-0.1678	-0.2145	-0.2253	-0.2260	-0.2268	-0.2273	-0.2275
df		9	13	14	14	15	15	15	15	15
PE		5.3240	4.5413	4.2659	4.1455	4.0750	4.0311	4.0038	3.9857	3.9729

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.6: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0267	0.0602	0.1032	0.1008	0.1512	0.2489	0.2929	0.3153	0.3699
λ_1		0.2407	0.2407	0.2407	0.1512	0.1512	0.1659	0.1255	0.0788	0.0411
Predictors	x_1	1.1886	1.1386	1.0802	1.2236	1.1513	1.0167	1.0081	1.0254	0.9934
	x_2	0.4986	0.4858	0.4700	0.5348	0.5157	0.4722	0.4822	0.5005	0.4994
	x_3	-0	-0	-0	-0	-0	-0	-0	-0	-0.0501
	x_4	-0	-0	-0	-0.0032	-0.0080	-0.0005	-0.0471	-0.0875	-0.1320
	x_5	0	0	0	0	0	0	0	0	0
	x_6	-0	-0	-0	-0	-0	-0	-0	-0	-0.0100
	x_7	1.1927	1.1488	1.0971	1.1918	1.1325	1.0218	1.0132	1.0242	1.0019
	x_8	0.0096	0.0180	0.0272	0.0864	0.0961	0.1008	0.1293	0.1562	0.1832
	x_9	0	0	0	0	0.0013	0	0.0389	0.0937	0.1402
	x_{10}	0.5613	0.5321	0.4984	0.6030	0.5653	0.4890	0.5118	0.5477	0.5596
	x_{11}	0.4866	0.4646	0.4391	0.4991	0.4596	0.3904	0.3800	0.3885	0.3819
	x_{12}	0	0	0	0	0	0	0.0079	0.0450	0.0866
	x_{13}	-0	-0	-0	-0.0062	-0	-0	-0	-0.0541	-0.0821
	x_{14}	0	0	0	0	0.0109	0.0330	0.0644	0.1033	0.1333
	x_{15}	-0	-0	-0	-0	-0	-0	-0	-0	-0.0150
df		6	6	6	8	9	8	10	11	14
PE		5.6142	5.7046	5.8227	5.5202	5.6456	5.9357	5.8863	5.7416	5.7080

Data no.6: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.09135656.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0914	0.0914	0.0914	0.0914	0.0914	0.0914	0.0914	0.0914	0.0914
λ_1		0.8222	0.3654	0.2132	0.1370	0.0914	0.0609	0.0392	0.0228	0.0102
Predictors	x_1	0.2031	0.9031	1.1396	1.2565	1.3077	1.3427	1.3676	1.3853	1.3992
	x_2	0	0.3698	0.4943	0.5477	0.5881	0.6161	0.6273	0.6459	0.6620
	x_3	0	0	-0	-0	-0	-0	-0	-0	-0
	x_4	-0	-0	-0	-0.0084	-0.0197	-0.0258	-0.0321	-0.0295	-0.0263
	x_5	0	0	0	0	0	0	0	0.0339	0.0643
	x_6	-0	-0	-0	-0	-0	-0	-0.0127	-0.0225	-0.0298
	x_7	0.5082	0.9749	1.1396	1.2171	1.2744	1.3136	1.3390	1.3558	1.3689
	x_8	0	0	0.0439	0.0897	0.1105	0.1235	0.1374	0.1480	0.1561
	x_9	0	0	0	0.0137	0.0893	0.1401	0.1784	0.2115	0.2375
	x_{10}	0	0.3616	0.5393	0.6281	0.7009	0.7496	0.7842	0.8118	0.8336
	x_{11}	0.0714	0.3662	0.4626	0.5293	0.5853	0.6243	0.6502	0.6591	0.6652
	x_{12}	0	0	0	0	0	0	0.0124	0.0370	0.0564
	x_{13}	0	-0	-0	-0.0294	-0.0969	-0.1400	-0.1739	-0.2092	-0.2372
	x_{14}	0	0	0	0.0020	0.0896	0.1492	0.1878	0.2115	0.2297
	x_{15}	0	0	-0	-0	-0.0622	-0.1102	-0.1463	-0.1785	-0.2040
df		3	5	6	10	11	11	13	14	14
PE		9.1343	6.3090	5.6931	5.4333	5.1943	5.0635	4.9853	4.9344	4.9011

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.7: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0475	0.0671	0.0722	0.1123	0.0878	0.1201	0.1702	0.1832	0.1626
λ_1		0.4271	0.2683	0.1685	0.1685	0.0878	0.0800	0.0729	0.0458	0.0181
Predictors	x_1	2.3310	2.3745	2.4167	2.3341	2.4300	2.3687	2.2775	2.2686	2.3182
	x_2	0	0	0	0	0	0	0	0	0
	x_3	0	0	0	0	0	0	0	0	-0
	x_4	0	0	0	0	0	0	0	-0	-0.0302
	x_5	-0	-0.3242	-0.6022	-0.5383	-0.8181	-0.7716	-0.6966	-0.7364	-0.8533
	x_6	0	0	0	0	0.0240	0.0394	0.0546	0.0863	0.1256
	x_7	0.8661	1.0970	1.2614	1.1858	1.3600	1.3048	1.2238	1.2365	1.3005
	x_8	0	0	0.0080	0	0.0511	0.0520	0.0506	0.0703	0.0978
	x_9	-0	-0	-0	-0	0	0	0	0	0.0349
	x_{10}	0	0	0.0123	0.0354	0.1116	0.1361	0.1590	0.1931	0.2269
	x_{11}	0.9843	1.0799	1.1484	1.0511	1.1362	1.0659	0.9752	0.9646	1.0102
	x_{12}	0	0	0.0064	0.0451	0.0569	0.0862	0.1192	0.1345	0.1446
	x_{13}	0.0793	0.1752	0.2291	0.2299	0.2549	0.2577	0.2600	0.2685	0.2881
	x_{14}	0	0	0.0194	0.0410	0.0786	0.0954	0.1141	0.1311	0.1405
	x_{15}	0	0	0	0	0	0	0	-0	-0.0079
df		4	5	9	9	10	10	10	10	13
PE		7.3278	6.5927	6.1581	6.3732	5.9279	6.0578	6.2889	6.2604	6.0618

Data no.7: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02873431.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0287	0.0287	0.0287	0.0287	0.0287	0.0287	0.0287	0.0287	0.0287
λ_1		0.2586	0.1149	0.0670	0.0431	0.0287	0.0192	0.0123	0.0072	0.0032
Predictors	x_1	2.4625	2.5432	2.5705	2.5808	2.5342	2.4918	2.4617	2.4382	2.4198
	x_2	0	-0	-0	-0.0103	-0.1250	-0.2144	-0.2780	-0.3296	-0.3699
	x_3	0	0	0	-0	-0.1385	-0.2659	-0.3566	-0.4387	-0.5028
	x_4	0	0	-0	-0.0150	-0.1913	-0.3325	-0.4330	-0.5247	-0.5963
	x_5	-0.4344	-0.8747	-1.0283	-1.1105	-1.3258	-1.4987	-1.6215	-1.7267	-1.8086
	x_6	0	0	0.0345	0.0656	0.0922	0.1098	0.1223	0.1350	0.1450
	x_7	1.1877	1.4568	1.5400	1.5773	1.5927	1.6047	1.6132	1.6226	1.6300
	x_8	0	0.0436	0.0770	0.1003	0.1288	0.1466	0.1592	0.1691	0.1768
	x_9	-0	0	0	0.0010	0.0432	0.0729	0.0940	0.1074	0.1177
	x_{10}	0	0.0439	0.1004	0.1330	0.1616	0.1818	0.1962	0.2066	0.2147
	x_{11}	1.1651	1.2911	1.3181	1.3335	1.3673	1.3938	1.4129	1.4274	1.4388
	x_{12}	0	0	0	0.0062	0.0848	0.1490	0.1947	0.2296	0.2567
	x_{13}	0.1564	0.2412	0.2614	0.2741	0.3241	0.3649	0.3941	0.4135	0.4286
	x_{14}	0	0.0220	0.0624	0.0831	0.1176	0.1452	0.1646	0.1779	0.1881
	x_{15}	0	0	-0	-0	-0	0	0	0.0205	0.0370
df		5	8	9	13	14	14	14	15	15
PE		6.3179	5.7622	5.6367	5.5855	5.4873	5.4225	5.3830	5.3534	5.3329

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.8: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0179	0.0533	0.0691	0.1075	0.1220	0.0870	0.1353	0.1209	0.1709
λ_1		0.1613	0.2133	0.1613	0.1613	0.1220	0.0580	0.0580	0.0302	0.0190
Predictors	x_1	2.1578	1.9916	2.0120	1.9228	1.9314	2.0649	1.9615	2.0054	1.9177
	x_2	-0	-0	-0	-0	-0	-0.0562	-0.0340	-0.1100	-0.1005
	x_3	0.1415	0.0747	0.1058	0.1006	0.1031	0.0898	0.0978	0.0734	0.0851
	x_4	-0.0856	-0	-0.0394	-0	-0.0859	-0.3412	-0.2515	-0.3838	-0.3197
	x_5	0	0	0.0210	0.0815	0.0872	0	0.0660	0	0.0832
	x_6	0.8054	0.7196	0.7508	0.7095	0.7321	0.8344	0.7807	0.8272	0.7812
	x_7	1.5157	1.4117	1.4232	1.3596	1.3697	1.4850	1.4039	1.4545	1.3848
	x_8	0.0803	0.0917	0.1157	0.1358	0.1507	0.1439	0.1701	0.1679	0.1889
	x_9	0.1880	0.1357	0.2065	0.2159	0.2655	0.3374	0.3402	0.3686	0.3823
	x_{10}	-0	-0	-0	-0	-0	-0	-0	-0.0024	-0.0288
	x_{11}	1.0300	0.9074	0.8832	0.7924	0.7849	0.9017	0.7978	0.8474	0.7557
	x_{12}	0.0299	0.0269	0.0352	0.0321	0.0407	0.0558	0.0571	0.0666	0.0674
	x_{13}	0.0512	0.0798	0.0951	0.1117	0.1285	0.1431	0.1546	0.1667	0.1776
	x_{14}	0	0.0342	0.0612	0.0938	0.1223	0.1420	0.1643	0.1808	0.2025
	x_{15}	0.1581	0.1315	0.1516	0.1484	0.1702	0.2297	0.2127	0.2490	0.2386
df		11	11	13	12	13	13	14	14	15
PE		5.7955	6.1206	5.9875	6.1509	6.0377	5.6566	5.8461	5.6917	5.8559

Data no.8: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03228914.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0323	0.0323	0.0323	0.0323	0.0323	0.0323	0.0323	0.0323	0.0323
λ_1		0.2906	0.1292	0.0753	0.0484	0.0323	0.0215	0.0138	0.0081	0.0036
Predictors	x_1	1.9511	2.1494	2.2042	2.2194	2.1988	2.204	2.1884	2.1706	2.1564
	x_2	-0	-0	-0.0158	-0.0894	-0.2125	-0.2571	-0.3323	-0.4015	-0.4559
	x_3	0.0129	0.1288	0.1255	0.1095	0	0	-0.0739	-0.1534	-0.2159
	x_4	-0	-0.1760	-0.3559	-0.4732	-0.6667	-0.7322	-0.8500	-0.9601	-1.0464
	x_5	0	0	0	-0	-0.2143	-0.2895	-0.4325	-0.5656	-0.6699
	x_6	0.6660	0.8216	0.8757	0.9041	0.9518	0.9757	1.0046	1.0295	1.0489
	x_7	1.3930	1.5178	1.5674	1.5952	1.6305	1.6513	1.6707	1.6866	1.6991
	x_8	0.0521	0.0963	0.1032	0.1061	0.1033	0.1022	0.1033	0.1035	0.1035
	x_9	0.0290	0.2373	0.3094	0.3412	0.3491	0.3575	0.3573	0.3561	0.3552
	x_{10}	-0	-0	-0	-0	0	0	0.0156	0.0253	0.0329
	x_{11}	0.9422	1.0049	1.0355	1.0485	1.1022	1.1266	1.1627	1.1945	1.2192
	x_{12}	0.0062	0.0371	0.0419	0.0399	0.0749	0.0862	0.1097	0.1323	0.1499
	x_{13}	0.0649	0.0785	0.0993	0.1129	0.1434	0.1471	0.1651	0.1841	0.1992
	x_{14}	0.0103	0.0376	0.0765	0.1053	0.1406	0.1481	0.1682	0.1882	0.2040
	x_{15}	0.1103	0.1809	0.2258	0.2662	0.3152	0.3364	0.3657	0.3913	0.4114
df		11	12	13	13	13	13	15	15	15
PE		6.3384	5.7116	5.5400	5.4672	5.3736	5.3382	5.2916	5.2537	5.2268

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.9: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0246	0.0608	0.0949	0.1117	0.1527	0.1733	0.2456	0.0655	0.1474
λ_1		0.2215	0.2431	0.2215	0.1676	0.1527	0.1155	0.1052	0.0164	0.0164
Predictors	x_1	1.9216	1.8327	1.7862	1.7919	1.7300	1.7225	1.6195	1.9768	1.8320
	x_2	0.4837	0.4608	0.4453	0.4481	0.4373	0.4391	0.4219	0.4086	0.4323
	x_3	0.1728	0.2048	0.2338	0.2491	0.2729	0.2860	0.3168	0	0.1724
	x_4	-0	0	0	-0	-0	-0	-0	-0.4652	-0.2707
	x_5	-0.1608	-0.1221	-0.1634	-0.2638	-0.2882	-0.3513	-0.3518	-0.8818	-0.6877
	x_6	-0	-0	-0	-0	-0	-0	-0	-0.1358	-0.0847
	x_7	1.2268	1.1487	1.124	1.1568	1.1199	1.1373	1.0649	1.4205	1.2926
	x_8	-0.1258	-0.1309	-0.1550	-0.1894	-0.2077	-0.2281	-0.2432	-0.2673	-0.2735
	x_9	-0	-0	-0	-0	-0	-0.0218	-0.0127	-0.1729	-0.1378
	x_{10}	-0.3693	-0.3333	-0.3318	-0.3587	-0.3492	-0.3612	-0.3425	-0.5033	-0.4467
	x_{11}	0.0391	0.0376	0.0548	0.0769	0.0833	0.0958	0.0993	0.2468	0.1910
	x_{12}	0.0879	0.0947	0.1097	0.1193	0.1249	0.1339	0.1397	0.2306	0.1956
	x_{13}	0.7808	0.7298	0.7033	0.7016	0.6671	0.6593	0.6096	1.0325	0.8194
	x_{14}	-0	0	0	-0	-0	-0	0	-0.2490	-0.1237
	x_{15}	0	0	0.0012	0.0246	0.0458	0.0625	0.0841	0.1798	0.1474
df		10	10	11	11	11	12	12	14	15
PE		4.4546	4.6141	4.6266	4.4967	4.5592	4.5023	4.6653	3.6850	3.9871

Data no.9: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03611225.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0361	0.0361	0.0361	0.0361	0.0361	0.0361	0.0361	0.0361	0.0361
λ_1		0.3250	0.1444	0.0843	0.0542	0.0361	0.0241	0.0155	0.0090	0.0040
Predictors	x_1	1.8133	1.9621	1.9909	2.0082	2.0227	2.0314	2.0272	2.0242	2.0219
	x_2	0.4588	0.4791	0.4249	0.4117	0.4180	0.4191	0.3845	0.3588	0.3389
	x_3	0.1660	0.1835	0.0535	0	0	-0.0066	-0.0899	-0.1523	-0.2007
	x_4	0	-0	-0.2224	-0.3765	-0.4481	-0.5004	-0.5890	-0.6552	-0.7065
	x_5	-0	-0.3207	-0.6219	-0.7810	-0.8501	-0.9019	-1.0039	-1.0802	-1.1393
	x_6	0	-0	-0.0500	-0.0975	-0.1216	-0.1386	-0.1617	-0.1790	-0.1925
	x_7	1.0956	1.3078	1.4009	1.4270	1.4338	1.4401	1.4664	1.4861	1.5014
	x_8	-0.0827	-0.1694	-0.2082	-0.2339	-0.2485	-0.2582	-0.2675	-0.2745	-0.2799
	x_9	-0	-0.0421	-0.1208	-0.1482	-0.1621	-0.1719	-0.1850	-0.1948	-0.2024
	x_{10}	-0.2814	-0.4087	-0.4528	-0.4896	-0.5149	-0.5312	-0.5354	-0.5385	-0.5409
	x_{11}	0	0.0799	0.1558	0.2082	0.2352	0.2540	0.2781	0.2962	0.3103
	x_{12}	0.0670	0.1187	0.1801	0.2103	0.2222	0.2316	0.2537	0.2703	0.2831
	x_{13}	0.7269	0.7959	0.9224	1.0222	1.0747	1.1114	1.1557	1.1889	1.2147
	x_{14}	0	-0	-0.0816	-0.1902	-0.2609	-0.3067	-0.3256	-0.3399	-0.3512
	x_{15}	0	0	0.0648	0.1254	0.1582	0.1807	0.2058	0.2246	0.2392
df		8	11	15	14	14	15	15	15	15
PE		4.8638	4.2521	3.9285	3.7453	3.6668	3.6226	3.5773	3.5475	3.5269

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.10: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0133	0.0227	0.0469	0.0664	0.0626	0.0446	0.0693	0.0747	0.0780
λ_1		0.1200	0.0908	0.1093	0.0996	0.0626	0.0297	0.0297	0.0187	0.0087
Predictors	x_1	1.5024	1.5535	1.4665	1.4519	1.5415	1.6692	1.6217	1.6493	1.6744
	x_2	0.2543	0.2808	0.3012	0.3223	0.3368	0.3649	0.3883	0.4171	0.4458
	x_3	-0	-0	0	0	0	0	0.0312	0.0766	0.1184
	x_4	0.4256	0.4839	0.4745	0.4981	0.5567	0.6875	0.6863	0.7493	0.8072
	x_5	-0.6196	-0.6594	-0.5346	-0.5048	-0.5956	-0.6583	-0.5817	-0.5453	-0.5105
	x_6	-0.1302	-0.1792	-0.1625	-0.1814	-0.2335	-0.2803	-0.2820	-0.2988	-0.3119
	x_7	1.7113	1.7338	1.5768	1.5169	1.6069	1.7485	1.6353	1.6288	1.6208
	x_8	0	0	0	0	0.0308	0.0748	0.0723	0.0832	0.0972
	x_9	-0	-0	-0	-0.0010	-0	0	-0	-0.0102	-0.0175
	x_{10}	-0.5704	-0.6112	-0.5685	-0.5714	-0.6279	-0.6512	-0.6470	-0.6516	-0.6558
	x_{11}	1.4423	1.4271	1.2976	1.2324	1.2865	1.4290	1.3086	1.2954	1.2831
	x_{12}	-0	-0.0085	-0	-0	-0.0210	-0.1130	-0.0830	-0.1062	-0.1229
	x_{13}	0	-0	0	0	-0	-0	-0	-0.0043	-0.0205
	x_{14}	0.1490	0.1881	0.1785	0.1963	0.2374	0.3247	0.3077	0.3260	0.3454
	x_{15}	-0	-0	-0	-0	-0	-0.0849	-0.0686	-0.0933	-0.1111
df		9	10	9	11	11	12	13	15	15
PE		5.2138	5.1379	5.3620	5.4392	5.2394	4.9590	5.1056	5.0893	5.0774

Data no.10: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.04530483.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0453	0.0453	0.0453	0.0453	0.0453	0.0453	0.0453	0.0453	0.0453
λ_1		0.4077	0.1812	0.1057	0.0680	0.0453	0.0302	0.0194	0.0113	0.0050
Predictors	x_1	0.8197	1.3066	1.4775	1.5618	1.6236	1.6663	1.7024	1.7391	1.7678
	x_2	0.0926	0.2669	0.3013	0.3187	0.3454	0.3649	0.3871	0.4162	0.4389
	x_3	0	0	0	-0	0	0	0.0149	0.0530	0.0823
	x_4	0.0071	0.3697	0.4788	0.5457	0.6276	0.6852	0.7407	0.8040	0.8528
	x_5	-0	-0.3600	-0.5480	-0.6311	-0.6475	-0.6560	-0.6481	-0.6218	-0.6023
	x_6	-0	-0.0617	-0.1670	-0.2238	-0.2576	-0.2797	-0.2953	-0.3078	-0.3175
	x_7	0.8517	1.4208	1.5916	1.6734	1.7160	1.7441	1.7629	1.7786	1.7916
	x_8	0	0	0	0.0218	0.0533	0.0741	0.0914	0.1092	0.1240
	x_9	-0	-0	-0	-0	-0	0	0	0.0175	0.0357
	x_{10}	-0.1823	-0.4553	-0.5752	-0.6285	-0.6425	-0.6509	-0.6567	-0.6634	-0.6693
	x_{11}	1.0271	1.2366	1.3071	1.3536	1.3960	1.4246	1.4416	1.4502	1.4573
	x_{12}	0	-0	-0	-0.0315	-0.0795	-0.1107	-0.1346	-0.1571	-0.1747
	x_{13}	0	0	0	-0	-0	-0	-0.0101	-0.0298	-0.0454
	x_{14}	0	0.1073	0.1813	0.2292	0.2846	0.3228	0.3515	0.3755	0.3946
	x_{15}	0	-0	-0	-0	-0.0469	-0.0831	-0.1092	-0.1291	-0.1439
df		6	9	9	11	12	12	14	15	15
PE		7.6124	5.7678	5.3355	5.1572	5.0342	4.9650	4.9215	4.8895	4.8672

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.11: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0266	0.0452	0.1124	0.1206	0.1136	0.1174	0.1259	0.1125	0.0474
λ_1		0.2390	0.1808	0.2623	0.1808	0.1136	0.0783	0.0540	0.0281	0.0053
Predictors	x_1	1.9627	1.9512	1.7901	1.8152	1.8203	1.8125	1.7991	1.818	1.8949
	x_2	-0	-0.0446	-0	-0.0662	-0.1604	-0.1992	-0.2246	-0.2570	-0.2631
	x_3	0.3050	0.3392	0.2247	0.2845	0.2977	0.3053	0.3084	0.3165	0.3618
	x_4	0.3128	0.3291	0.2960	0.3271	0.2806	0.2683	0.2690	0.2569	0.2272
	x_5	-0	-0.0678	-0	-0.0426	-0.2217	-0.3096	-0.3634	-0.4637	-0.7219
	x_6	0	0	0	0	0	0	0.0138	0.0256	0.0067
	x_7	1.7260	1.7709	1.577	1.6635	1.7489	1.7867	1.8029	1.8536	2.0335
	x_8	0	0	-0	0	0	0	0.0084	0.0450	0.1394
	x_9	-0	-0	-0	-0	-0	-0.0342	-0.0656	-0.1156	-0.2174
	x_{10}	-0	-0	-0	-0	-0	-0.0186	-0.0543	-0.0929	-0.1382
	x_{11}	0.9402	0.9485	0.8555	0.8584	0.9131	0.9366	0.9448	0.9899	1.1909
	x_{12}	-0	-0	-0	-0	-0.0424	-0.0882	-0.1217	-0.1628	-0.1734
	x_{13}	0.1088	0.1484	0.1552	0.1751	0.2141	0.2395	0.2579	0.2770	0.3622
	x_{14}	0	0	0.0324	0.0687	0.1436	0.1932	0.2287	0.2566	0.2657
	x_{15}	0	0	0	0	0	0	0	-0	-0.1807
df		6	8	7	9	10	12	14	14	15
PE		3.7908	3.6690	4.1167	3.8914	3.6124	3.4863	3.4205	3.2896	3.0172

Data no.11: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.02020411.

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0202	0.0202	0.0202	0.0202	0.0202	0.0202	0.0202	0.0202	0.0202
λ_1		0.1818	0.0808	0.0471	0.0303	0.0202	0.0135	0.0087	0.0051	0.0022
Predictors	x_1	1.9903	1.9372	1.9159	1.9136	1.9121	1.9111	1.9103	1.9097	1.9089
	x_2	-0.0484	-0.2250	-0.2691	-0.2682	-0.2677	-0.2674	-0.2672	-0.2668	-0.2668
	x_3	0.3460	0.3017	0.3001	0.3236	0.3377	0.3471	0.3537	0.3590	0.3626
	x_4	0.3088	0.1423	0.1215	0.1434	0.1565	0.1653	0.1715	0.1770	0.1807
	x_5	-0.0979	-0.5530	-0.7177	-0.7784	-0.8150	-0.8393	-0.8573	-0.8712	-0.8827
	x_6	0	0	-0	-0	-0	-0	-0.0017	-0.0065	-0.0104
	x_7	1.8092	1.9405	2.0031	2.0435	2.0678	2.0839	2.0960	2.1063	2.1143
	x_8	0	0.0376	0.0901	0.1240	0.1444	0.1580	0.1678	0.1755	0.1814
	x_9	-0	-0.1160	-0.1799	-0.2113	-0.2301	-0.2426	-0.2515	-0.2580	-0.2630
	x_{10}	-0	-0.0183	-0.0776	-0.1052	-0.1218	-0.1328	-0.1408	-0.1469	-0.1517
	x_{11}	0.9836	1.1351	1.2007	1.2401	1.2638	1.2795	1.2910	1.2999	1.3069
	x_{12}	-0	-0.0940	-0.1291	-0.1385	-0.1441	-0.1479	-0.1501	-0.1509	-0.1514
	x_{13}	0.1351	0.2450	0.3035	0.3450	0.3698	0.3864	0.3983	0.4074	0.4146
	x_{14}	0	0.1370	0.1996	0.2228	0.2368	0.2461	0.2525	0.2563	0.2596
	x_{15}	0	-0	-0.0638	-0.1458	-0.1950	-0.2278	-0.2511	-0.2686	-0.2819
df		8	13	14	14	14	14	15	15	15
PE		3.6043	3.2005	3.0704	3.0126	2.9862	2.9721	2.9636	2.9579	2.9541

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.12: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0443	0.0996	0.1556	0.1150	0.1571	0.1957	0.3044	0.2986	0.3844
λ_1		0.3983	0.3983	0.3630	0.1724	0.1571	0.1304	0.1304	0.0746	0.0427
Predictors	x_1	0.9518	0.8906	0.8701	1.1239	1.0886	1.0749	0.9684	1.0241	0.9732
	x_2	0	0	0.0364	0.2576	0.2770	0.3071	0.2933	0.3501	0.3596
	x_3	-0	-0	-0	-0	-0	-0	-0	-0.0249	-0.0491
	x_4	0.0013	0.0307	0.0788	0.2620	0.2728	0.2874	0.2698	0.3001	0.3025
	x_5	-0	-0	-0	-0.0115	-0.0240	-0.0424	-0.0566	-0.0937	-0.1224
	x_6	-0	-0	-0	-0	-0.0053	-0.02907	-0.0508	-0.0762	-0.0999
	x_7	0.8530	0.8122	0.7980	0.8769	0.8646	0.8492	0.8034	0.8082	0.7812
	x_8	-0	-0.0424	-0.1113	-0.3084	-0.3506	-0.4050	-0.4279	-0.5052	-0.5446
	x_9	-0	-0	-0	-0	-0	-0	-0	-0.0133	-0.0364
	x_{10}	0	0	0	0	0	0.0329	0.0534	0.1260	0.1683
	x_{11}	0.8279	0.7663	0.7168	0.8126	0.7602	0.7231	0.6309	0.6528	0.6122
	x_{12}	-0	-0	-0	-0.1999	-0.1940	-0.2026	-0.1590	-0.2096	-0.2028
	x_{13}	0.3307	0.3379	0.3484	0.3865	0.3838	0.3866	0.3686	0.3911	0.3892
	x_{14}	0	0	0	0.1477	0.1716	0.1987	0.1901	0.2353	0.2444
	x_{15}	-0	-0	-0	-0.1279	-0.1266	-0.1399	-0.0903	-0.1473	-0.1382
df		5	6	7	11	12	13	13	15	15
PE		7.1184	7.2316	7.2179	6.0829	6.1180	6.0814	6.3601	6.0719	6.1316

Data no.12: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.0451382.

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0451	0.0451	0.0451	0.0451	0.0451	0.0451	0.0451	0.0451	0.0451
λ_1		0.4062	0.1806	0.1053	0.0677	0.0451	0.0301	0.0193	0.0113	0.0050
Predictors	x_1	0.9418	1.2175	1.3224	1.3680	1.3843	1.3952	1.4030	1.4089	1.4137
	x_2	0	0.2285	0.3534	0.4193	0.4485	0.4679	0.4816	0.4919	0.4999
	x_3	-0	-0	-0	-0.0433	-0.0869	-0.1157	-0.1359	-0.1511	-0.1626
	x_4	0	0.2569	0.3693	0.4044	0.4107	0.4145	0.4169	0.4187	0.4204
	x_5	-0	-0	-0	-0.0180	-0.0518	-0.0745	-0.0910	-0.1032	-0.1124
	x_6	-0	-0	-0	-0	-0	-0.0028	-0.0066	-0.0095	-0.0117
	x_7	0.8455	0.8967	0.9026	0.8896	0.8822	0.8771	0.8734	0.8706	0.8684
	x_8	-0	-0.2299	-0.3544	-0.4453	-0.5025	-0.5405	-0.5676	-0.5880	-0.6035
	x_9	-0	-0	-0	-0.0017	-0.0207	-0.0333	-0.0423	-0.0490	-0.0543
	x_{10}	0	0	0.0145	0.0815	0.1239	0.1528	0.1737	0.1895	0.2017
	x_{11}	0.8241	0.9485	0.9673	0.9713	0.9782	0.9822	0.9848	0.9867	0.9885
	x_{12}	-0	-0.2335	-0.3387	-0.3995	-0.4343	-0.4573	-0.4735	-0.4858	-0.4952
	x_{13}	0.3285	0.3937	0.3993	0.4095	0.4217	0.4299	0.4360	0.4405	0.4440
	x_{14}	0	0.1123	0.2408	0.3078	0.3488	0.3761	0.3955	0.4101	0.4212
	x_{15}	-0	-0.1597	-0.2959	-0.3510	-0.3775	-0.3949	-0.4071	-0.4162	-0.4236
df		4	10	11	14	14	15	15	15	15
PE		7.1416	5.9328	5.5628	5.4208	5.3543	5.3198	5.3001	5.2879	5.2799

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.13: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0203	0.0417	0.0651	0.0439	0.0600	0.1305	0.0964	0.1991	0.1767
λ_1		0.1831	0.1668	0.1520	0.0658	0.0600	0.0870	0.0413	0.0498	0.0196
Predictors	x_1	1.4462	1.4371	1.4241	1.5698	1.5521	1.402	1.5143	1.3537	1.4219
	x_2	-0	-0	-0	-0.0087	-0.0096	-0	-0.0428	-0.0105	-0.0637
	x_3	0.2478	0.2741	0.2973	0.3532	0.3573	0.3264	0.3560	0.3383	0.3530
	x_4	0.6075	0.6235	0.6365	0.6928	0.6774	0.5873	0.6363	0.5539	0.5806
	x_5	-0.2410	-0.2309	-0.2177	-0.4196	-0.4125	-0.3213	-0.4231	-0.3429	-0.3956
	x_6	-0	-0	-0	-0	-0	-0	-0	-0	0
	x_7	0.6819	0.6802	0.6733	0.8761	0.8639	0.7232	0.8503	0.7203	0.7915
	x_8	0.4208	0.4355	0.4474	0.5601	0.5632	0.5012	0.5698	0.5068	0.5618
	x_9	-0	-0	-0	-0	-0	-0	-0.0067	-0.0209	-0.0568
	x_{10}	0.0518	0.0707	0.0866	0.19122	0.1976	0.1627	0.2121	0.1933	0.2207
	x_{11}	0.8715	0.8517	0.8298	0.9965	0.9625	0.7787	0.9051	0.7117	0.7770
	x_{12}	-0	-0	-0	-0.0353	-0.0333	-0	-0.0449	-0	-0.0453
	x_{13}	0	0	0	0.1211	0.1346	0.0926	0.1644	0.1330	0.1817
	x_{14}	-0	-0	-0	-0.1333	-0.1221	-0	-0.1161	-0.0142	-0.0799
	x_{15}	0	0	0	0	0	0.0114	0.0464	0.0634	0.0991
df		8	8	8	12	12	11	14	13	14
PE		4.4965	4.4958	4.5076	4.0534	4.0823	4.4494	4.1178	4.5058	4.2876

Data no.13: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.04261893.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0426	0.0426	0.0426	0.0426	0.0426	0.0426	0.0426	0.0426	0.0426
λ_1		0.3836	0.1705	0.0994	0.0639	0.0426	0.0284	0.0183	0.0107	0.0047
Predictors	x_1	1.2051	1.4333	1.5045	1.5758	1.6139	1.6337	1.6493	1.6610	1.6697
	x_2	0	-0	-0	-0.0118	-0.0519	-0.0947	-0.1253	-0.1482	-0.1662
	x_3	0.0955	0.2724	0.3171	0.3544	0.3611	0.3543	0.3481	0.3435	0.3404
	x_4	0.4410	0.6216	0.6567	0.6955	0.6992	0.6876	0.6790	0.6726	0.6688
	x_5	-0	-0.2199	-0.3870	-0.4234	-0.4638	-0.5057	-0.5339	-0.5549	-0.5713
	x_6	-0	-0	0	-0	-0	-0	0	0	0.0026
	x_7	0.2666	0.6725	0.8074	0.8817	0.9249	0.9558	0.9782	0.9950	1.0075
	x_8	0.1889	0.4312	0.5105	0.5644	0.6020	0.6241	0.6431	0.6573	0.6686
	x_9	-0	-0	-0	-0	-0	-0.0085	-0.0276	-0.0418	-0.0539
	x_{10}	0	0.0662	0.1537	0.1931	0.2137	0.2256	0.2337	0.2398	0.2453
	x_{11}	0.6786	0.8454	0.9385	1.0042	1.0502	1.0799	1.0987	1.1128	1.1237
	x_{12}	0	-0	-0	-0.0388	-0.0722	-0.0983	-0.1182	-0.1330	-0.1448
	x_{13}	0	0	0.0406	0.1264	0.1832	0.2205	0.2507	0.2733	0.2906
	x_{14}	0	-0	-0.0413	-0.1404	-0.1971	-0.2361	-0.2645	-0.2859	-0.3022
	x_{15}	0	0	0	0	0.0177	0.0540	0.0787	0.0971	0.1113
df		6	8	10	12	13	14	14	14	15
PE		5.4870	4.5133	4.2297	4.0395	3.9414	3.8851	3.8512	3.8297	3.8153

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.14: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0185	0.0346	0.0541	0.0208	0.1047	0.0323	0.1398	0.1506	0.1109
λ_1		0.1668	0.1385	0.1262	0.0312	0.1047	0.0215	0.0599	0.0376	0.0123
Predictors	x_1	2.5418	2.5270	2.4866	2.6904	2.3768	2.6632	2.3332	2.3277	2.4481
	x_2	0	0	0	0	0	0	-0	-0	-0.0135
	x_3	0.3125	0.3268	0.3240	0.3757	0.3035	0.3499	0.2944	0.2989	0.3036
	x_4	0.8960	0.8843	0.8542	1.0372	0.7826	1.0064	0.7464	0.7439	0.8248
	x_5	-0	-0.0336	-0.0600	-0.0783	-0.1128	-0.1185	-0.2023	-0.2317	-0.2330
	x_6	-0.0926	-0.1227	-0.1325	-0.2278	-0.1467	-0.2316	-0.1840	-0.2010	-0.2265
	x_7	1.7165	1.7199	1.6894	1.9759	1.5994	1.9518	1.5850	1.5950	1.7330
	x_8	0.4855	0.5275	0.5410	0.6135	0.5530	0.6125	0.5961	0.6180	0.6316
	x_9	-0	-0.0287	-0.0297	-0.2431	-0.0229	-0.2482	-0.0756	-0.1053	-0.1792
	x_{10}	-0	-0	-0	-0	-0	-0	-0	-0	-0
	x_{11}	0.6181	0.6177	0.6167	0.7649	0.6017	0.7664	0.5959	0.6004	0.6722
	x_{12}	-0	-0	-0	-0.1207	-0	-0.1265	-0	-0.0126	-0.0688
	x_{13}	-0	-0	-0	-0.1561	-0	-0.1701	-0	-0	-0.0833
	x_{14}	0.0751	0.1001	0.1237	0.1799	0.1590	0.1929	0.1778	0.1865	0.2086
	x_{15}	0	0	0	0.0496	0.0277	0.0839	0.0745	0.0925	0.1230
df		8	10	10	13	11	13	11	12	14
PE		4.2823	4.2338	4.2497	3.8497	4.3474	3.8509	4.3345	4.3053	4.0599

Data no.14: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.02800986.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0280	0.0280	0.0280	0.0280	0.0280	0.0280	0.0280	0.0280	0.0280
λ_1		0.2521	0.1120	0.0654	0.0420	0.0280	0.0187	0.0120	0.0070	0.0031
Predictors	x_1	2.4439	2.5745	2.6287	2.6565	2.6704	2.6797	2.6864	2.6913	2.6960
	x_2	0	0	0	0	0	0	0	0	0.0034
	x_3	0.2131	0.3418	0.3732	0.3788	0.3621	0.3507	0.3425	0.3366	0.3343
	x_4	0.7991	0.9023	0.9489	0.9950	1.0137	1.0257	1.0346	1.0410	1.0492
	x_5	-0	-0.08084	-0.1112	-0.0948	-0.1026	-0.1084	-0.1122	-0.1151	-0.1135
	x_6	-0.0015	-0.1511	-0.1978	-0.2179	-0.2280	-0.2347	-0.2394	-0.2430	-0.2457
	x_7	1.5722	1.7826	1.8743	1.9273	1.9549	1.9733	1.9864	1.9962	2.0037
	x_8	0.3556	0.5690	0.6209	0.6224	0.6141	0.6087	0.6046	0.6018	0.5992
	x_9	-0	-0.0871	-0.1717	-0.2129	-0.2407	-0.2592	-0.2724	-0.2823	-0.2906
	x_{10}	-0	-0	-0	-0	-0	-0	-0	-0	-0.0005
	x_{11}	0.6221	0.6331	0.6812	0.7288	0.7598	0.7805	0.7952	0.8062	0.8139
	x_{12}	-0	-0	-0.0358	-0.0854	-0.1179	-0.1394	-0.1550	-0.1666	-0.1760
	x_{13}	0	-0	-0.0403	-0.1056	-0.1554	-0.1885	-0.2124	-0.2301	-0.2444
	x_{14}	0.0717	0.1046	0.1413	0.1706	0.1853	0.1951	0.2022	0.2075	0.2108
	x_{15}	0	0	0	0.0240	0.0639	0.0906	0.1098	0.1241	0.1340
df		8	10	12	13	13	13	13	13	15
PE		4.5647	4.1322	3.9829	3.8991	3.8572	3.8349	3.8216	3.8132	3.8074

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.15: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0134	0.0302	0.0357	0.0609	0.0523	0.1249	0.1220	0.1090	0.1691
λ_1		0.1208	0.1208	0.0832	0.0914	0.0523	0.0832	0.0523	0.0273	0.0188
Predictors	x_1	1.9220	1.8881	1.9216	1.8614	1.9165	1.7523	1.7863	1.8303	1.7350
	x_2	0.4099	0.4023	0.4367	0.4183	0.4712	0.4021	0.4350	0.4774	0.4521
	x_3	0.1596	0.1593	0.1836	0.1749	0.1938	0.1711	0.1860	0.1976	0.2042
	x_4	0	0	0	0	0	0	0	0	0.0070
	x_5	0	0	0	0	0	0	-0	-0	-0
	x_6	0	0	0.0094	0	0.0614	0.0086	0.0575	0.0948	0.0930
	x_7	1.0773	1.0576	1.0911	1.0506	1.0894	0.9831	1.0100	1.0358	0.9794
	x_8	0.4263	0.4209	0.4545	0.4355	0.4886	0.4212	0.4579	0.4942	0.4760
	x_9	-0	-0	-0	-0	-0.0262	-0	-0.0111	-0.0411	-0.0350
	x_{10}	0.5583	0.5433	0.5906	0.5604	0.6356	0.5313	0.5839	0.6340	0.6021
	x_{11}	0.2954	0.2979	0.3075	0.3044	0.3275	0.3001	0.3090	0.3384	0.3240
	x_{12}	0	0	0.00036	0.0176	0.0297	0.05489	0.0623	0.0748	0.0886
	x_{13}	0.3958	0.3904	0.3940	0.3764	0.3899	0.3435	0.3470	0.3675	0.3383
	x_{14}	-0	-0	-0	-0	-0.0328	-0	-0.0075	-0.0600	-0.0558
	x_{15}	0	0	0	0	0	0	0	0	0.0204
df		8	8	10	9	12	11	12	12	14
PE		2.7166	2.7411	2.6834	2.7344	2.6337	2.8268	2.7500	2.6595	2.7473

Data no.15: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03957326.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0396	0.0396	0.0396	0.0396	0.0396	0.0396	0.0396	0.0396	0.0396
λ_1		0.3562	0.1583	0.0923	0.0594	0.0396	0.0264	0.0170	0.0099	0.0044
Predictors	x_1	1.5889	1.8250	1.9035	1.9366	1.9531	1.9641	1.9725	2.0028	2.0315
	x_2	0.1694	0.3617	0.4259	0.4670	0.4995	0.5212	0.5382	0.5915	0.6417
	x_3	0.0023	0.1341	0.1778	0.1930	0.1999	0.2044	0.2127	0.2557	0.2957
	x_4	0	0	0	0	0	0	0.0115	0.0932	0.1700
	x_5	0	0	0	0	0	0	0	0.0804	0.1601
	x_6	0	0	0	0.0504	0.0827	0.1043	0.1211	0.1432	0.1626
	x_7	0.7921	1.0062	1.0774	1.1024	1.1104	1.1158	1.1176	1.1130	1.1084
	x_8	0.2096	0.3847	0.4430	0.4852	0.5123	0.5304	0.5428	0.5520	0.5593
	x_9	0	-0	-0	-0.0211	-0.0447	-0.0605	-0.0695	-0.0711	-0.0717
	x_{10}	0.2244	0.4856	0.5731	0.6325	0.6684	0.6923	0.7102	0.7257	0.7382
	x_{11}	0.2423	0.29	0.3057	0.3219	0.3484	0.3661	0.3779	0.3785	0.3777
	x_{12}	0	0	0.0013	0.0155	0.0329	0.0445	0.0517	0.0561	0.0601
	x_{13}	0.3514	0.3818	0.3912	0.3960	0.4087	0.4173	0.4231	0.4196	0.4155
	x_{14}	0	-0	-0	-0.0215	-0.0695	-0.1015	-0.1270	-0.1593	-0.1866
	x_{15}	0	0	0	0	0	-0	-0	-0.0368	-0.0743
df		8	8	9	12	12	12	13	14	15
PE		3.5761	2.8391	2.7044	2.6350	2.5864	2.5615	2.5467	2.5245	2.5073

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.16: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0133	0.0249	0.0354	0.0502	0.0753	0.1361	0.1211	0.1571	0.1530
λ_1		0.1199	0.0996	0.0827	0.0753	0.0753	0.0907	0.0519	0.0393	0.0170
Predictors	x_1	1.3670	1.3725	1.3761	1.3668	1.3396	1.2592	1.3067	1.2778	1.3004
	x_2	-0	-0	-0	-0	-0	-0	-0	-0.0082	-0.0630
	x_3	0	0	0.00031	0.0053	0.0126	0.0149	0.0194	0.0233	0.0164
	x_4	0.7539	0.7560	0.7551	0.7517	0.7413	0.6953	0.7184	0.6982	0.6911
	x_5	-0.2540	-0.2915	-0.3214	-0.3155	-0.2774	-0.1890	-0.2890	-0.2878	-0.3562
	x_6	-0.2677	-0.2869	-0.3015	-0.3090	-0.3115	-0.3061	-0.3363	-0.3498	-0.3738
	x_7	1.6939	1.6901	1.6867	1.6713	1.6369	1.5443	1.5897	1.5520	1.5764
	x_8	0.4927	0.5118	0.5265	0.5263	0.5104	0.4550	0.5104	0.5034	0.5276
	x_9	-0.1515	-0.1638	-0.1735	-0.1774	-0.1766	-0.1625	-0.1833	-0.1861	-0.2028
	x_{10}	-0.2331	-0.2619	-0.2834	-0.2857	-0.2690	-0.2065	-0.2753	-0.2732	-0.3177
	x_{11}	0.4890	0.5054	0.5131	0.5057	0.4853	0.4367	0.4670	0.4521	0.4716
	x_{12}	-0	-0	-0	-0	-0	0	-0	-0	-0.0030
	x_{13}	0.5634	0.5680	0.5677	0.5565	0.5347	0.4884	0.5095	0.4902	0.4933
	x_{14}	0	0	0.0099	0.0220	0.0362	0.0550	0.0562	0.0675	0.0742
	x_{15}	0	0	0	0	0	0.0170	0.0280	0.0486	0.0668
df		10	10	12	12	12	13	13	14	15
PE		3.7398	3.7013	3.6759	3.6876	3.7445	3.9560	3.7917	3.8486	3.7687

Data no.16: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.03513.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0351	0.0351	0.0351	0.0351	0.0351	0.0351	0.0351	0.0351	0.0351
λ_1		0.3162	0.1405	0.0820	0.0527	0.0351	0.0234	0.0151	0.0088	0.0039
Predictors	x_1	1.1553	1.3238	1.3770	1.4057	1.4237	1.4360	1.4447	1.4467	1.4463
	x_2	0	0	-0	-0.0311	-0.0798	-0.1091	-0.1300	-0.1569	-0.1828
	x_3	0	0.0031	0	0	0	0	-0	-0.0304	-0.0673
	x_4	0.5844	0.7426	0.7552	0.7512	0.7367	0.7258	0.7181	0.6914	0.6613
	x_5	-0	-0.1619	-0.3241	-0.4123	-0.4782	-0.5261	-0.5603	-0.6128	-0.6656
	x_6	-0.0960	-0.2585	-0.3019	-0.3261	-0.3469	-0.3630	-0.3745	-0.3839	-0.3914
	x_7	1.5146	1.6467	1.6876	1.7129	1.7302	1.7396	1.7463	1.7490	1.7501
	x_8	0.2327	0.4517	0.5276	0.5633	0.5837	0.5992	0.6103	0.6116	0.6095
	x_9	-0.0131	-0.1410	-0.1738	-0.1942	-0.2077	-0.2137	-0.2180	-0.2224	-0.2263
	x_{10}	-0	-0.1848	-0.2847	-0.3406	-0.3812	-0.4095	-0.4297	-0.4397	-0.4452
	x_{11}	0.3716	0.4562	0.5139	0.5460	0.5693	0.5835	0.5936	0.6081	0.6226
	x_{12}	0	0	-0	-0	-0.0145	-0.0441	-0.0652	-0.0767	-0.0835
	x_{13}	0.5086	0.5339	0.5685	0.5777	0.5803	0.5875	0.5925	0.6001	0.6076
	x_{14}	0	0	0.0101	0.0240	0.0339	0.0403	0.0448	0.0564	0.0688
	x_{15}	0	0	0	0	0.0156	0.0395	0.0566	0.0727	0.0867
df		8	11	11	12	14	14	14	15	15
PE		4.5894	3.8690	3.6734	3.5986	3.5540	3.5257	3.5095	3.4967	3.4870

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.17: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0290	0.0134	0.0083	0.1094	0.1495	0.1862	0.0451	0.1021	0.2297
λ_1		0.2612	0.0537	0.0193	0.1641	0.1495	0.1241	0.0193	0.0255	0.0255
Predictors	x_1	1.7876	1.6834	1.6064	1.7055	1.6620	1.6284	1.7020	1.7189	1.6059
	x_2	-0	-0.6851	-1.0635	-0	-0	-0	-0.7096	-0.4338	-0.2474
	x_3	0.4314	0	-0.2453	0.3861	0.3921	0.3976	0.0367	0.2718	0.4017
	x_4	1.0088	0.3234	-0	0.9309	0.9120	0.8965	0.4173	0.7130	0.8176
	x_5	-0.0151	-1.3982	-1.9586	-0.2920	-0.3059	-0.3396	-1.3575	-0.8918	-0.5863
	x_6	-0.4178	-0.7332	-0.8052	-0.5086	-0.5087	-0.5199	-0.7455	-0.6824	-0.6059
	x_7	0.9922	1.1451	1.1853	0.9786	0.9595	0.9494	1.1439	1.0779	0.9745
	x_8	-0	-0.0545	-0.0683	-0	-0	-0	-0.0541	-0.0490	-0.0665
	x_9	0	0.0496	0.0625	0	0	0	0.0673	0.0484	0.0083
	x_{10}	0	0.1767	0.3041	0	0	0	0.2131	0.1330	0.0631
	x_{11}	0.9869	1.2619	1.3730	0.8230	0.7611	0.7185	1.1202	0.9024	0.6990
	x_{12}	0	0.2630	0.4563	0	0	0	0.2748	0.1547	0.1085
	x_{13}	0.0309	0.3974	0.4724	0.1955	0.2132	0.23445	0.4017	0.3404	0.2983
	x_{14}	0	0	0	0.1152	0.1478	0.1731	0.0600	0.1180	0.1889
	x_{15}	0	0.2010	0.3020	0.0979	0.1313	0.1591	0.2532	0.2224	0.2302
df		8	13	13	10	10	10	15	15	15
PE		5.7655	4.7157	4.5295	5.6346	5.7184	5.7601	4.7455	5.0494	5.5184

Data no.17: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.03274476.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0327	0.0327	0.0327	0.0327	0.0327	0.0327	0.0327	0.0327	0.0327
λ_1		0.2947	0.1310	0.0764	0.0491	0.0327	0.0218	0.0140	0.0082	0.0036
Predictors	x_1	1.7563	1.7878	1.7379	1.7076	1.6932	1.6875	1.6834	1.6705	1.6606
	x_2	0	-0.0476	-0.3824	-0.5885	-0.7056	-0.7766	-0.8274	-0.8833	-0.9262
	x_3	0.4059	0.2988	0.1242	0.0317	0	0	-0	-0.0555	-0.1004
	x_4	0.9737	0.8752	0.5969	0.4450	0.3696	0.3347	0.3096	0.2515	0.2060
	x_5	-0	-0.5026	-0.9690	-1.2436	-1.3891	-1.4673	-1.5233	-1.6135	-1.6838
	x_6	-0.3798	-0.5803	-0.6686	-0.7137	-0.7409	-0.7587	-0.7715	-0.7815	-0.7890
	x_7	0.9715	1.0537	1.0991	1.1274	1.1442	1.1555	1.1635	1.1695	1.1743
	x_8	-0	-0	-0.0213	-0.0380	-0.0504	-0.0605	-0.0677	-0.0679	-0.0679
	x_9	0	0	0.0314	0.0480	0.0590	0.0675	0.0735	0.0751	0.0764
	x_{10}	0	0	0.0704	0.1492	0.1978	0.2313	0.2552	0.2701	0.2816
	x_{11}	0.9882	1.0398	1.1121	1.1532	1.1737	1.1832	1.1898	1.2052	1.2173
	x_{12}	0	0	0.1172	0.2169	0.2734	0.3072	0.3314	0.3588	0.3798
	x_{13}	0.0161	0.2319	0.3317	0.3758	0.4013	0.4166	0.4276	0.4398	0.4485
	x_{14}	0	0.0432	0.0655	0.05707	0.0454	0.0320	0.0225	0.0305	0.0373
	x_{15}	0	0.0405	0.1431	0.2037	0.2372	0.2571	0.2715	0.2888	0.3025
df		7	11	15	15	14	14	14	15	15
PE		5.8728	5.2857	4.9566	4.7997	4.7264	4.6894	4.6662	4.6375	4.6166

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.18: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV.

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0318	0.0593	0.0401	0.0569	0.0708	0.0968	0.1372	0.0927	0.0080
λ_1		0.2859	0.2373	0.0936	0.0853	0.0708	0.0645	0.0588	0.0232	0.0009
Predictors	x_1	2.5246	2.4980	2.5455	2.5019	2.4655	2.4116	2.3409	2.4013	2.2169
	x_2	0.1695	0.2116	0.2752	0.2684	0.2657	0.2618	0.2672	0.2731	-0.4460
	x_3	-0	-0	-0.2917	-0.3047	-0.3376	-0.3278	-0.2868	-0.4616	-1.6426
	x_4	0	0.0313	0	0	0	0	0.0351	-0.0183	-1.4325
	x_5	-0.0199	-0.0981	-0.5868	-0.6178	-0.6693	-0.6740	-0.6444	-0.8335	-2.2643
	x_6	-0	-0	-0.0238	-0.0183	-0.0185	-0.0126	-0.0026	-0.0362	-0.1436
	x_7	1.2065	1.2383	1.5058	1.4926	1.4942	1.4622	1.4077	1.5408	1.7758
	x_8	0	0	0.0614	0.0786	0.0998	0.1128	0.1317	0.1487	-0.0043
	x_9	-0	-0.0071	-0.2298	-0.2320	-0.2458	-0.2396	-0.2238	-0.3067	-0.5344
	x_{10}	0	0.0039	0.2281	0.2195	0.2236	0.2070	0.1840	0.2654	0.2268
	x_{11}	1.0395	1.0026	1.1010	1.0568	1.0284	0.9744	0.8987	1.0090	1.5460
	x_{12}	0	0	-0	-0	-0	-0	-0	-0.0175	0.2697
	x_{13}	0.1462	0.1965	0.3124	0.3129	0.3185	0.3139	0.2975	0.3521	0.7668
	x_{14}	0	0.0061	0.0430	0.0829	0.1169	0.1552	0.1917	0.1854	0.2848
	x_{15}	0	0.0010	0.1115	0.1345	0.1594	0.1759	0.1905	0.2207	0.2984
df		6	11	13	13	13	13	14	15	15
PE		4.9480	4.8643	4.0638	4.0850	4.0764	4.1485	4.2963	3.9841	3.4033

Data no.18: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.0200494.

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200
λ_1		0.1804	0.0802	0.0468	0.0301	0.0200	0.0134	0.0086	0.0050	0.0022
Predictors	x_1	2.6222	2.5866	2.5053	2.4421	2.4028	2.3786	2.3627	2.3392	2.3208
	x_2	0.2721	0.2839	0.1894	0.1068	0.0467	0	-0.0281	-0.0939	-0.1449
	x_3	-0.0173	-0.3615	-0.6234	-0.8076	-0.9290	-1.0260	-1.0917	-1.1738	-1.2373
	x_4	0	-0	-0.2247	-0.4283	-0.5653	-0.6805	-0.7584	-0.8573	-0.9335
	x_5	-0.2345	-0.6524	-0.9744	-1.1999	-1.3505	-1.4740	-1.5571	-1.6657	-1.7499
	x_6	-0	-0.0418	-0.0672	-0.0860	-0.0977	-0.1071	-0.1139	-0.1179	-0.1210
	x_7	1.3691	1.5656	1.6454	1.6919	1.7176	1.7291	1.7375	1.7422	1.7458
	x_8	0	0.0681	0.0733	0.0599	0.0518	0.0470	0.0436	0.0403	0.0377
	x_9	-0.1029	-0.2671	-0.3454	-0.3935	-0.4229	-0.4441	-0.4594	-0.4707	-0.4795
	x_{10}	0.1165	0.2726	0.2999	0.3068	0.3077	0.3030	0.3007	0.2896	0.2810
	x_{11}	1.0858	1.1755	1.2412	1.2948	1.3291	1.3534	1.3702	1.3890	1.4033
	x_{12}	0	-0	-0	0	0.0108	0.0443	0.0664	0.0986	0.1233
	x_{13}	0.2188	0.3313	0.4266	0.5015	0.5477	0.5761	0.5954	0.6194	0.6378
	x_{14}	0	0.0063	0.0803	0.1340	0.1669	0.1883	0.2030	0.2202	0.2338
	x_{15}	0	0.1121	0.1762	0.2133	0.2348	0.2433	0.2485	0.2596	0.2686
df		9	13	14	14	15	14	15	15	15
PE		4.4771	3.9514	3.7561	3.6521	3.5979	3.5612	3.5386	3.5133	3.4953

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.19: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0162	0.1016	0.0570	0.2974	0.2326	0.1041	0.1619	0.0909	0.2045
λ_1		0.1461	0.4065	0.1331	0.4461	0.2326	0.0694	0.0694	0.0227	0.0227
Predictors	x_1	1.3067	1.1062	1.2478	0.9027	1.0096	1.1563	1.0799	1.1300	1.0203
	x_2	0.1444	0	0.1344	0	0.0276	0.1539	0.1368	0.1474	0.1455
	x_3	0.4408	0.1864	0.4502	0.1352	0.3228	0.4722	0.4519	0.4411	0.4451
	x_4	-0	0	0	0	0	-0	-0	-0.1396	-0.0464
	x_5	-0.3509	-0	-0.2969	-0	-0	-0.3971	-0.3173	-0.6302	-0.3989
	x_6	-0	-0	-0	-0	-0.0083	-0.0334	-0.0560	-0.0549	-0.0893
	x_7	0.3478	0.0813	0.3413	0.0345	0.2187	0.3641	0.3412	0.3936	0.3523
	x_8	0.3122	0	0.2818	0	0.0451	0.3222	0.2702	0.4428	0.3028
	x_9	-0	-0	-0	-0	0	-0	-0	-0.0347	-0.0003
	x_{10}	0.0327	0	0.0823	0.0023	0.1213	0.1901	0.2216	0.2231	0.2696
	x_{11}	1.2138	0.7341	1.0800	0.5268	0.6329	0.9982	0.8592	1.1398	0.8305
	x_{12}	0	0	0	0	0	0	0.0103	0.0598	0.0421
	x_{13}	-0	0	0	0	0	-0	0	-0.0693	-0
	x_{14}	0	0.0467	0.0108	0.1204	0.1331	0.0774	0.1183	0.0938	0.1457
	x_{15}	0	0.1056	0.0839	0.1691	0.1887	0.1594	0.1998	0.1921	0.2369
df		8	6	10	7	10	11	12	15	14
PE		4.7494	5.9476	4.8362	6.4874	5.6945	4.7956	4.9621	4.5808	4.9472

Data no.19: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03606571.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0361	0.0361	0.0361	0.0361	0.0361	0.0361	0.0361	0.0361	0.0361
λ_1		0.3246	0.1443	0.0842	0.0541	0.0361	0.0240	0.0155	0.0090	0.0040
Predictors	x_1	1.2387	1.2809	1.2756	1.2658	1.2252	1.1966	1.1765	1.1611	1.1492
	x_2	0	0.1378	0.1837	0.1979	0.1713	0.1522	0.1389	0.1286	0.1207
	x_3	0.2784	0.4448	0.4854	0.4801	0.4334	0.4008	0.3780	0.3604	0.3468
	x_4	0	0	-0	-0.0465	-0.1600	-0.2382	-0.2934	-0.3354	-0.3680
	x_5	-0	-0.3054	-0.4974	-0.6236	-0.7618	-0.8561	-0.9228	-0.9734	-1.0127
	x_6	-0	-0	-0	-0.0098	-0.0266	-0.0371	-0.0446	-0.0502	-0.0546
	x_7	0.1764	0.3407	0.3847	0.4003	0.4075	0.4127	0.4163	0.4191	0.4212
	x_8	0.0107	0.2893	0.3968	0.4775	0.5402	0.5817	0.6114	0.6336	0.6509
	x_9	-0	-0	-0	-0.0413	-0.0669	-0.0842	-0.0966	-0.1058	-0.1130
	x_{10}	0	0.0512	0.1066	0.1319	0.1503	0.1634	0.1727	0.1798	0.1853
	x_{11}	0.9070	1.1416	1.2416	1.3315	1.4003	1.4455	1.4778	1.5019	1.5207
	x_{12}	0	0	0	0.0197	0.0729	0.1081	0.1331	0.1520	0.1667
	x_{13}	0	-0	-0	-0.0457	-0.1021	-0.1393	-0.1661	-0.1860	-0.2015
	x_{14}	0	0	0	0	0.0181	0.0327	0.0429	0.0507	0.0568
	x_{15}	0.0510	0.0404	0.0516	0.0685	0.1033	0.1266	0.1430	0.1556	0.1653
df		6	9	9	14	15	15	15	15	15
PE		5.5593	4.8068	4.6150	4.4959	4.4073	4.3587	4.3293	4.3101	4.2968

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.20: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0365	0.0682	0.0971	0.1658	0.1297	0.2135	0.2512	0.3575	0.3173
λ_1		0.3288	0.2730	0.2266	0.2487	0.1297	0.1423	0.1077	0.0894	0.0353
Predictors	x_1	1.0261	1.0969	1.1486	1.0704	1.2461	1.1654	1.1609	1.0862	1.1527
	x_2	-0	-0	-0	-0	-0.0238	-0	-0.0304	-0.0353	-0.1093
	x_3	0	0	0	0	0.0250	0.0513	0.0635	0.0956	0.0973
	x_4	0	0	0	0	0	0	0.0020	0.0431	0.0351
	x_5	-0	-0	-0	-0	-0.0991	-0.0512	-0.1008	-0.0807	-0.1775
	x_6	-0.4414	-0.4951	-0.5383	-0.4904	-0.6058	-0.5664	-0.5692	-0.5325	-0.5666
	x_7	0.1983	0.2662	0.3168	0.2671	0.4361	0.3783	0.4072	0.3887	0.4728
	x_8	0	0	0	0	0	0	0	0.0327	0.0611
	x_9	0	0.0249	0.0756	0.0552	0.1648	0.1479	0.1799	0.1990	0.2557
	x_{10}	0	0	0	0	0	0	0	0	0.0170
	x_{11}	1.2213	1.1482	1.0916	0.9969	1.0407	0.9376	0.8988	0.8024	0.8429
	x_{12}	0.0554	0.0838	0.1080	0.1191	0.1436	0.1498	0.1608	0.1639	0.1756
	x_{13}	0.0068	0.0643	0.1056	0.1453	0.1803	0.2039	0.2251	0.2303	0.2506
	x_{14}	0	0	0	0	0	0	0	0.0059	0.0035
	x_{15}	0	0	0	0	0	0.0044	0.0318	0.0668	0.0759
df		6	7	7	7	10	10	12	14	15
PE		5.7005	5.5833	5.4986	5.7146	5.2803	5.5063	5.4907	5.6934	5.4547

Data no.20: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.0188469.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0188	0.0188	0.0188	0.0188	0.0188	0.0188	0.0188	0.0188	0.0188
λ_1		0.1696	0.0754	0.0440	0.0283	0.0188	0.0126	0.0081	0.0047	0.0021
Predictors	x_1	1.2864	1.3793	1.3787	1.3852	1.3757	1.3665	1.3592	1.3448	1.3337
	x_2	-0.0059	-0.1813	-0.2494	-0.2942	-0.3390	-0.3722	-0.3969	-0.4263	-0.4492
	x_3	0	0	0	-0	-0.0667	-0.1243	-0.1667	-0.2123	-0.2478
	x_4	0	-0	-0.0858	-0.1469	-0.2384	-0.3101	-0.3625	-0.4145	-0.4549
	x_5	-0.0347	-0.2935	-0.4517	-0.5532	-0.6760	-0.7698	-0.8383	-0.9066	-0.9596
	x_6	-0.6302	-0.6433	-0.6239	-0.6179	-0.6222	-0.6265	-0.6296	-0.6312	-0.6325
	x_7	0.4288	0.6090	0.6764	0.7099	0.7231	0.7306	0.7358	0.73858	0.7407
	x_8	0	-0	-0	-0.0468	-0.0876	-0.1164	-0.1369	-0.1526	-0.1648
	x_9	0.1324	0.2268	0.2403	0.2326	0.2202	0.2108	0.2044	0.2028	0.2015
	x_{10}	0	0.0391	0.0768	0.0992	0.1126	0.1215	0.1275	0.1289	0.1299
	x_{11}	1.2358	1.2733	1.3306	1.3633	1.3934	1.4153	1.4309	1.4412	1.4492
	x_{12}	0.1007	0.1335	0.1588	0.1705	0.1930	0.2110	0.2241	0.2369	0.2468
	x_{13}	0.0493	0.1650	0.2703	0.3333	0.3824	0.4171	0.4418	0.4594	0.4731
	x_{14}	-0	-0.0393	-0.1299	-0.1774	-0.1877	-0.1907	-0.1927	-0.1918	-0.1911
	x_{15}	0	-0	-0	-0	-0	0	0.0013	0.0156	0.0266
df		9	11	12	13	14	14	15	15	15
PE		5.2111	4.9127	4.7906	4.7385	4.7008	4.6782	4.6647	4.6546	4.6479

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.21: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0055	0.0179	0.0232	0.0330	0.0374	0.2135	0.0601	0.1031	0.0396
λ_1		0.0494	0.0717	0.0542	0.0494	0.0374	0.1423	0.0258	0.0258	0.0044
Predictors	x_1	2.1569	2.0711	2.0875	2.0649	2.0705	2.0691	2.0166	1.8967	2.099
	x_2	-0	-0	-0	-0	-0	-0	-0	0	-0.1271
	x_3	0.8323	0.7943	0.8031	0.7936	0.7967	0.7965	0.7732	0.7201	0.7513
	x_4	0	0	0	0	0	0	-0	-0	-0.1174
	x_5	-0.5761	-0.4434	-0.4826	-0.4634	-0.4944	-0.5124	-0.4740	-0.3848	-0.7151
	x_6	-0.1165	-0.0690	-0.0962	-0.0979	-0.1198	-0.1352	-0.1329	-0.1189	-0.1914
	x_7	1.3106	1.2803	1.2826	1.2719	1.2698	1.2657	1.2401	1.1800	1.2825
	x_8	-1.0208	-0.9858	-0.9953	-0.9888	-0.9902	-0.9892	-0.9691	-0.9255	-1.0096
	x_9	-0.3082	-0.2938	-0.3039	-0.3053	-0.3125	-0.3178	-0.3176	-0.3151	-0.3303
	x_{10}	-0.3663	-0.3474	-0.3604	-0.3620	-0.3774	-0.3887	-0.3938	-0.3951	-0.3992
	x_{11}	1.1253	1.0207	1.0349	1.0056	1.0076	1.0024	0.9413	0.8193	1.0693
	x_{12}	0.6154	0.5431	0.5596	0.5431	0.5485	0.5483	0.5109	0.4383	0.6133
	x_{13}	0	0.0227	0.0363	0.0585	0.0611	0.0660	0.0883	0.1266	0.0787
	x_{14}	0	0	0	0.0005	0.0208	0.0377	0.0613	0.0974	0.0842
	x_{15}	-0.6469	-0.5424	-0.5705	-0.5526	-0.5695	-0.5770	-0.5343	-0.4371	-0.6274
df		11	12	12	13	13	13	13	13	15
PE		4.1654	4.3052	4.2614	4.2923	4.2697	4.2630	4.3433	4.5728	4.1691

Data no.21: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.05082661.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0508	0.0508	0.0508	0.0508	0.0508	0.0508	0.0508	0.0508	0.0508
λ_1		0.4574	0.2033	0.1186	0.0762	0.0508	0.0339	0.0218	0.0127	0.0056
Predictors	x_1	1.4253	1.7446	1.8869	1.9616	2.0041	2.0324	2.0525	2.0631	2.0633
	x_2	0	0	0	0	0	-0	-0	-0.0418	-0.0904
	x_3	0.4061	0.6275	0.7083	0.7452	0.7660	0.7797	0.7895	0.7779	0.7518
	x_4	0	0	0	0	0	0	-0	-0.0302	-0.0929
	x_5	-0	-0	-0.1980	-0.3265	-0.4125	-0.4704	-0.5116	-0.5724	-0.6442
	x_6	-0	-0	-0	-0.0422	-0.0896	-0.1213	-0.1439	-0.1659	-0.1788
	x_7	0.8870	1.1136	1.2001	1.2324	1.2433	1.2505	1.2556	1.2602	1.2646
	x_8	-0.6853	-0.8478	-0.9143	-0.9484	-0.9645	-0.9752	-0.9827	-0.9887	-0.9943
	x_9	-0.0461	-0.2267	-0.2646	-0.2866	-0.3027	-0.3133	-0.3210	-0.3265	-0.3289
	x_{10}	-0.0372	-0.2706	-0.3100	-0.3363	-0.3646	-0.3836	-0.3973	-0.4021	-0.4046
	x_{11}	0.5865	0.6836	0.8193	0.8944	0.9348	0.9616	0.9807	0.9994	1.0170
	x_{12}	0.1400	0.2902	0.4012	0.4645	0.4982	0.5209	0.5370	0.5562	0.5753
	x_{13}	0.1238	0.0486	0.0781	0.0876	0.0824	0.0787	0.0761	0.0799	0.0909
	x_{14}	0	0	0	0	0.0250	0.0423	0.0548	0.0705	0.0898
	x_{15}	-0	-0.1340	-0.3412	-0.4398	-0.4999	-0.5402	-0.5689	-0.5852	-0.5953
df		9	10	11	12	13	13	13	15	15
PE		6.7025	5.4240	4.7761	4.5204	4.3950	4.3251	4.2820	4.2487	4.2232

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.22: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0194	0.0397	0.0470	0.0381	0.0998	0.0940	0.0837	0.1191	0.1274
λ_1		0.1745	0.1590	0.1096	0.0571	0.0998	0.0627	0.0359	0.0298	0.0142
Predictors	x_1	2.0611	2.0372	2.0663	2.1285	1.9739	2.0138	2.0596	1.9934	1.9898
	x_2	0.5776	0.5813	0.6104	0.6557	0.5946	0.6196	0.6510	0.6325	0.6397
	x_3	0.2842	0.2924	0.3033	0.2474	0.2898	0.2650	0.2441	0.2350	0.2219
	x_4	0.3138	0.3105	0.3488	0.3448	0.2882	0.2923	0.3040	0.2724	0.2662
	x_5	-0	-0	-0.0387	-0.1825	-0.0656	-0.1508	-0.2138	-0.2248	-0.2582
	x_6	-0.8254	-0.8231	-0.8441	-0.8915	-0.8319	-0.8632	-0.8857	-0.8747	-0.8813
	x_7	1.6196	1.5969	1.6512	1.7312	1.5523	1.6036	1.6560	1.5836	1.5816
	x_8	-0	-0	-0	0	-0	-0	-0	-0.0140	-0.0244
	x_9	0	0	0	0.0599	0	0.0047	0.0669	0.0589	0.0855
	x_{10}	-0	-0	-0.0099	-0.0944	-0.0095	-0.0591	-0.1057	-0.0985	-0.1180
	x_{11}	0.2372	0.2371	0.2476	0.2855	0.2487	0.2674	0.2869	0.2754	0.2793
	x_{12}	0.0527	0.0709	0.0768	0.0961	0.1242	0.1367	0.1407	0.1500	0.1542
	x_{13}	0.4987	0.4829	0.4877	0.5236	0.4603	0.4796	0.4982	0.4692	0.4666
	x_{14}	0	0	0	0	0	0	0	0.0282	0.0402
	x_{15}	0	0	0	-0	0	0	0	0.0133	0.0200
df		9	9	11	12	11	12	12	15	15
PE		4.1308	4.1511	4.0619	3.9178	4.1737	4.0609	3.9691	4.0554	4.0482

Data no.22: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02459862.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0246	0.0246	0.0246	0.0246	0.0246	0.0246	0.0246	0.0246	0.0246
λ_1		0.2214	0.0984	0.0574	0.0369	0.0246	0.0164	0.0105	0.0061	0.0027
Predictors	x_1	2.0042	2.1208	2.1567	2.1783	2.2157	2.2408	2.2589	2.2721	2.2820
	x_2	0.5389	0.6294	0.6654	0.6899	0.7372	0.7689	0.7918	0.8084	0.8210
	x_3	0.2510	0.2967	0.2443	0.2219	0.2364	0.2463	0.2536	0.2632	0.2716
	x_4	0.2328	0.3813	0.3627	0.3583	0.3894	0.4104	0.4257	0.4378	0.4469
	x_5	-0	-0.0617	-0.1842	-0.2411	-0.2512	-0.2578	-0.2622	-0.2602	-0.2577
	x_6	-0.8076	-0.8622	-0.8970	-0.9122	-0.9114	-0.9108	-0.9102	-0.9097	-0.9093
	x_7	1.5341	1.7161	1.7637	1.7887	1.8132	1.8297	1.8414	1.8530	1.8629
	x_8	-0	-0	0	0	0.0253	0.0423	0.0545	0.0630	0.0692
	x_9	0	0	0.0720	0.1187	0.1570	0.1826	0.2010	0.2112	0.2182
	x_{10}	-0	-0.0313	-0.1010	-0.1368	-0.1552	-0.1674	-0.1761	-0.1826	-0.1876
	x_{11}	0.2325	0.2555	0.2893	0.3108	0.3476	0.3723	0.3903	0.4026	0.4122
	x_{12}	0.0721	0.0566	0.0809	0.0923	0.0958	0.0981	0.0997	0.1054	0.1115
	x_{13}	0.4912	0.5102	0.5357	0.5505	0.5747	0.5908	0.6023	0.6123	0.6210
	x_{14}	0	0	0	-0	-0	-0	-0	-0.0101	-0.0208
	x_{15}	0	0	-0	-0.0095	-0.0805	-0.1281	-0.1626	-0.1863	-0.2047
df		9	11	12	13	14	14	14	15	15
PE		4.2610	3.9874	3.8946	3.8567	3.8094	3.7838	3.7684	3.7578	3.7504

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.23: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0121	0.0359	0.0424	0.0660	0.0749	0.1023	0.0911	0.1076	0.0413
λ_1		0.1086	0.1436	0.0990	0.0990	0.0749	0.0682	0.0390	0.0269	0.0046
Predictors	x_1	1.9930	1.8871	1.9428	1.8963	1.9174	1.8778	1.9427	1.9314	2.1215
	x_2	0.1020	0.0164	0.0795	0.0577	0.0862	0.0761	0.1277	0.1366	0.3198
	x_3	0.2216	0.2176	0.2457	0.2587	0.2754	0.2910	0.3028	0.3183	0.3191
	x_4	-0	0	-0	0	0	0	0	0	-0.1348
	x_5	-0.9093	-0.7092	-0.8047	-0.7171	-0.7428	-0.6684	-0.7656	-0.7332	-1.0040
	x_6	-0.3553	-0.3431	-0.3728	-0.3775	-0.3934	-0.3985	-0.4167	-0.4245	-0.4435
	x_7	0.9675	0.9241	0.9612	0.9488	0.9630	0.9521	0.9812	0.9798	1.0611
	x_8	-0.3133	-0.1773	-0.2928	-0.2594	-0.3113	-0.2920	-0.3820	-0.3900	-0.5914
	x_9	0.1013	0.0243	0.0813	0.0598	0.0847	0.0721	0.1193	0.1230	0.3198
	x_{10}	-0	-0	-0	-0.0047	-0.0195	-0.0374	-0.0431	-0.0570	-0.0031
	x_{11}	1.2871	1.1661	1.1972	1.1318	1.133	1.0752	1.1285	1.1037	1.4046
	x_{12}	-0	0	0	0	0	0	0	0	0.0391
	x_{13}	0.3122	0.3251	0.3407	0.3534	0.3626	0.3720	0.3781	0.3876	0.4425
	x_{14}	-0	-0	-0	-0	-0	-0	-0.0053	-0.0136	-0.0978
	x_{15}	-0	0	-0	0	-0	0	-0	-0	-0.1309
df		10	10	10	11	11	11	12	12	15
PE		5.5036	5.7858	5.6105	5.7330	5.6766	5.7906	5.6208	5.6536	5.1696

Data no.23: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02356816.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0236	0.0236	0.0236	0.0236	0.0236	0.0236	0.0236	0.0236	0.0236
λ_1		0.2121	0.0943	0.0550	0.0354	0.0236	0.0157	0.0101	0.0059	0.0026
Predictors	x_1	1.8267	1.9904	2.0482	2.0968	2.1259	2.1432	2.1531	2.1583	2.1616
	x_2	0	0.1094	0.1698	0.2501	0.2983	0.3248	0.3373	0.3459	0.3514
	x_3	0.1814	0.2366	0.2261	0.2596	0.2798	0.2871	0.2856	0.2848	0.2831
	x_4	0	-0	-0.1101	-0.1372	-0.1534	-0.1739	-0.1993	-0.2199	-0.2374
	x_5	-0.5616	-0.8970	-1.0343	-1.0412	-1.0453	-1.0633	-1.0930	-1.1194	-1.1422
	x_6	-0.2844	-0.3694	-0.4047	-0.4179	-0.4258	-0.4325	-0.4387	-0.4439	-0.4482
	x_7	0.8639	0.9749	1.0138	1.0375	1.0517	1.0621	1.0703	1.0785	1.0852
	x_8	-0.0045	-0.3358	-0.4690	-0.5370	-0.5777	-0.6052	-0.6250	-0.6397	-0.6513
	x_9	0	0.1084	0.2085	0.2826	0.3271	0.3532	0.3679	0.3793	0.3879
	x_{10}	-0	-0	-0	-0	-0	0	0	0.0096	0.0187
	x_{11}	1.1378	1.2632	1.3494	1.4128	1.4509	1.4748	1.4901	1.5028	1.5125
	x_{12}	0	-0	0	0	0	0.0116	0.0321	0.0485	0.0615
	x_{13}	0.2746	0.3270	0.3738	0.4035	0.4213	0.4347	0.4459	0.4555	0.4634
	x_{14}	0	-0	-0.0171	-0.0522	-0.0732	-0.0872	-0.0968	-0.1036	-0.1085
	x_{15}	0	-0	-0.0180	-0.0752	-0.1095	-0.1319	-0.1469	-0.1592	-0.1681
df		8	10	13	13	13	14	14	15	15
PE		6.0772	5.5019	5.3075	5.2038	5.1547	5.1259	5.1070	5.0935	5.0842

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.24: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0108	0.0561	0.0457	0.1243	0.0349	0.1103	0.1424	0.1273	0.1639
λ_1		0.0972	0.2245	0.1067	0.1864	0.0349	0.0735	0.0610	0.0318	0.0182
Predictors	x_1	1.8106	1.5711	1.7077	1.5001	1.8393	1.5948	1.5424	1.5792	1.5184
	x_2	-0	-0	-0.0156	-0	-0.0342	-0.0862	-0.1076	-0.1565	-0.1793
	x_3	0.7108	0.4611	0.6117	0.4698	0.8065	0.5439	0.5123	0.5328	0.4993
	x_4	1.2269	0.8407	1.0755	0.8319	1.3540	0.9672	0.9152	0.9567	0.8994
	x_5	-0.1352	-0.2961	-0.2639	-0.3269	-0.1094	-0.4149	-0.4680	-0.4917	-0.5429
	x_6	-0	-0	-0	-0	-0.0120	-0	-0	-0.0454	-0.0718
	x_7	1.3277	1.0498	1.2433	1.0262	1.4363	1.2118	1.1913	1.2667	1.2446
	x_8	-0	-0	-0	-0	-0.0086	-0	-0	-0.0491	-0.0764
	x_9	0.1490	0.0082	0.1286	0.0301	0.1965	0.1439	0.1489	0.1812	0.1844
	x_{10}	0	0	0	0	0.0552	0.0504	0.0721	0.0970	0.1240
	x_{11}	0.6062	0.5433	0.6032	0.5414	0.6706	0.6041	0.5895	0.6112	0.5898
	x_{12}	-0.2637	-0	-0.1619	-0	-0.2860	-0.1117	-0.1023	-0.1552	-0.1446
	x_{13}	0	0	0	0	0	0.0488	0.0821	0.0902	0.1148
	x_{14}	-0.0118	0	-0	0	-0.1256	-0	0	-0	-0
	x_{15}	-0	0	0	0	0	0	0.0170	0.0519	0.0933
df		9	7	9	7	13	11	12	14	14
PE		3.2221	3.8660	3.3740	3.9227	3.0801	3.4582	3.5149	3.3790	3.4455

Data no.24: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03091316.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0309	0.0309	0.0309	0.0309	0.0309	0.0309	0.0309	0.0309	0.0309
λ_1		0.2782	0.1237	0.0721	0.0464	0.0309	0.0206	0.0132	0.0077	0.0034
Predictors	x_1	1.5801	1.7216	1.7918	1.8373	1.8634	1.8796	1.8892	1.8943	1.898
	x_2	-0	-0	-0.0218	-0.0203	-0.0211	-0.0247	-0.0306	-0.0382	-0.0443
	x_3	0.4348	0.6120	0.7117	0.7936	0.8415	0.8696	0.8831	0.8891	0.8935
	x_4	0.8128	1.0780	1.2241	1.3365	1.4017	1.4400	1.4591	1.4684	1.4752
	x_5	-0.2385	-0.2364	-0.1767	-0.1041	-0.0631	-0.0412	-0.0331	-0.0312	-0.0301
	x_6	-0	-0	-0	-0.0018	-0.0119	-0.0219	-0.0288	-0.0377	-0.0447
	x_7	1.0161	1.2391	1.3408	1.4140	1.4585	1.4882	1.5092	1.5267	1.5402
	x_8	-0	-0	-0	-0	-0.0103	-0.0278	-0.0405	-0.0525	-0.0619
	x_9	0	0.1162	0.1663	0.1879	0.2001	0.2077	0.2141	0.2198	0.2243
	x_{10}	0	0	0.0087	0.0390	0.0578	0.0708	0.0807	0.0907	0.0985
	x_{11}	0.5292	0.5927	0.6306	0.6578	0.6741	0.6857	0.6927	0.6967	0.6997
	x_{12}	-0	-0.1604	-0.2527	-0.2853	-0.3033	-0.3137	-0.3228	-0.3314	-0.3380
	x_{13}	0	0	0	0	0	0.0037	0.0127	0.0198	0.0253
	x_{14}	0	-0	-0.0337	-0.1081	-0.1512	-0.1798	-0.2013	-0.2196	-0.2336
	x_{15}	0	-0	-0	0	0	0	0.0062	0.0227	0.0358
df		6	8	11	12	13	14	15	15	15
PE		3.9528	3.3866	3.1995	3.1006	3.0533	3.0272	3.0122	3.0029	2.9970

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.25: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0318	0.0594	0.0846	0.1199	0.1360	0.2041	0.1252	0.1119	0.1735
λ_1		0.2864	0.2377	0.1974	0.1798	0.1360	0.1360	0.0537	0.0280	0.0193
Predictors	x_1	1.7582	1.7415	1.7295	1.6902	1.6867	1.5979	1.694	1.698	1.6325
	x_2	-0	-0	-0	-0	-0.0747	-0.0786	-0.2602	-0.3393	-0.2970
	x_3	-0	-0.0455	-0.1069	-0.1196	-0.1900	-0.1550	-0.4196	-0.5220	-0.4607
	x_4	0.8307	0.8348	0.8285	0.8090	0.7802	0.7487	0.6478	0.5768	0.6044
	x_5	0	0	-0	-0	-0	-0	-0.2515	-0.3910	-0.3337
	x_6	0	0	0	0	0.0473	0.0600	0.1281	0.1482	0.1611
	x_7	1.2612	1.2902	1.3185	1.3244	1.3527	1.3212	1.4238	1.4491	1.4363
	x_8	0	0	0	0	0.0032	0	0.0761	0.1021	0.1059
	x_9	-0	-0	-0	-0	-0	-0	-0	-0	-0.0216
	x_{10}	-0.5244	-0.5545	-0.5773	-0.5827	-0.6146	-0.6018	-0.6815	-0.7082	-0.6882
	x_{11}	0.4172	0.4126	0.4063	0.3940	0.4022	0.3830	0.4928	0.5421	0.4927
	x_{12}	0	0	0	0	0	0	0	-0	0
	x_{13}	0	0	0.0202	0.0433	0.0652	0.0837	0.1327	0.1617	0.1615
	x_{14}	0	0	0	0	0	0	0	0	0.0256
	x_{15}	0.2244	0.2537	0.2678	0.2687	0.2821	0.2756	0.3471	0.3801	0.3518
df		6	7	8	8	11	10	12	12	14
PE		3.6955	3.6296	3.5694	3.5964	3.4910	3.6327	3.1583	3.0443	3.1559

Data no.25: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03653586.

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0365	0.0365	0.0365	0.0365	0.0365	0.0365	0.0365	0.0365	0.0365
λ_1		0.3288	0.1461	0.0853	0.0548	0.0365	0.0244	0.0157	0.0091	0.0041
Predictors	x_1	1.7290	1.8343	1.8169	1.7812	1.7599	1.7477	1.7407	1.7355	1.7314
	x_2	-0	-0.0395	-0.2262	-0.3593	-0.4392	-0.4996	-0.5415	-0.5730	-0.5974
	x_3	-0	-0.2338	-0.4352	-0.5809	-0.6682	-0.7301	-0.7699	-0.7999	-0.8231
	x_4	0.8109	0.8547	0.6889	0.5379	0.4475	0.3827	0.3438	0.3147	0.2921
	x_5	0	-0	-0.2170	-0.4503	-0.5902	-0.6810	-0.7371	-0.7792	-0.8119
	x_6	0	0.0085	0.0803	0.1047	0.1193	0.1305	0.1391	0.1455	0.1505
	x_7	1.2195	1.3737	1.4248	1.4466	1.4597	1.4663	1.4709	1.4744	1.4771
	x_8	0	0.0101	0.0636	0.0926	0.1099	0.1196	0.1237	0.1268	0.1293
	x_9	-0	-0	-0	0	0	0.0166	0.0337	0.0464	0.0564
	x_{10}	-0.4960	-0.6228	-0.6826	-0.7193	-0.7413	-0.7573	-0.7705	-0.7805	-0.7882
	x_{11}	0.4125	0.4386	0.5232	0.6008	0.6474	0.6791	0.7024	0.7199	0.7335
	x_{12}	0	-0	-0	-0	-0	-0.0007	-0.0125	-0.0214	-0.0282
	x_{13}	0	0.0086	0.0768	0.1327	0.1663	0.1885	0.2056	0.2185	0.2285
	x_{14}	0	0	0	0	0	0	-0	-0	-0
	x_{15}	0.2175	0.2972	0.3574	0.4109	0.4429	0.4676	0.4877	0.5028	0.5145
df		6	11	12	12	12	14	14	14	14
PE		3.7845	3.3521	3.1033	2.9666	2.9020	2.8637	2.8383	2.8219	2.8106

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.26: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0062	0.0106	0.0165	0.0213	0.0889	0.0633	0.0514	0.2952	0.1033
λ_1		0.0558	0.0422	0.0385	0.0319	0.0889	0.0422	0.0220	0.0738	0.0115
Predictors	x_1	1.7858	1.7888	1.7897	1.7949	1.6702	1.7484	1.7843	1.4492	1.7241
	x_2	0	0	0	0.0081	0.1584	0.1515	0.1425	0.2887	0.2318
	x_3	0	0.0149	0.0473	0.0813	0.1167	0.1826	0.2073	0.1314	0.2552
	x_4	0.0046	0.0117	0.0353	0.0622	0.2294	0.2355	0.2306	0.3033	0.3168
	x_5	-1.0588	-1.0676	-1.0290	-0.9973	-0.5223	-0.6633	-0.7391	-0.2483	-0.5679
	x_6	0.6264	0.6558	0.6348	0.6330	0.2155	0.4151	0.5217	0.0726	0.4011
	x_7	1.5731	1.5927	1.5956	1.6047	1.4155	1.5342	1.5895	1.1918	1.5265
	x_8	-0.4090	-0.4237	-0.4281	-0.4365	-0.4023	-0.4441	-0.4635	-0.4164	-0.4764
	x_9	0.2081	0.2116	0.2231	0.2301	0.2692	0.2642	0.2565	0.2733	0.2840
	x_{10}	0.1456	0.1616	0.1638	0.1695	0.0911	0.1467	0.1725	0.0662	0.1645
	x_{11}	0.8656	0.8648	0.8342	0.8150	0.5023	0.6052	0.6655	0.3340	0.5364
	x_{12}	0.2414	0.2632	0.2664	0.2732	0.1917	0.2421	0.2685	0.1889	0.2571
	x_{13}	0.6495	0.6535	0.6331	0.6199	0.3955	0.4693	0.5115	0.2903	0.4292
	x_{14}	-0.3319	-0.3736	-0.3672	-0.3794	-0.0666	-0.2535	-0.3517	-0	-0.2781
	x_{15}	0.1401	0.1549	0.1673	0.1760	0.1976	0.2029	0.2017	0.2040	0.2226
df		13	14	14	15	15	15	15	14	15
PE		3.6046	3.5706	3.5830	3.5817	4.2256	3.8322	3.6900	4.8699	3.8687

Data no.26: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02675825.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0268	0.0268	0.0268	0.0268	0.0268	0.0268	0.0268	0.0268	0.0268
λ_1		0.2408	0.1070	0.0624	0.0401	0.0268	0.0178	0.0115	0.0067	0.0030
Predictors	x_1	1.7029	1.7225	1.7570	1.7849	1.8014	1.8121	1.8201	1.8260	1.8304
	x_2	0.1118	0	0.0038	0.0322	0.0488	0.0595	0.0676	0.0735	0.0777
	x_3	0	0	0.0393	0.0946	0.1271	0.1483	0.1642	0.1758	0.1843
	x_4	0.1907	0.0494	0.0550	0.0943	0.1169	0.1313	0.1426	0.1507	0.1564
	x_5	-0.2356	-0.8214	-0.9188	-0.9226	-0.9260	-0.9291	-0.9301	-0.9313	-0.9331
	x_6	0.0033	0.3162	0.4862	0.5699	0.6200	0.6533	0.6772	0.6951	0.7093
	x_7	1.2419	1.4576	1.5382	1.5841	1.6115	1.6297	1.6428	1.6526	1.6601
	x_8	-0.2834	-0.3568	-0.4032	-0.4313	-0.4481	-0.4592	-0.4672	-0.4732	-0.4778
	x_9	0.1738	0.2510	0.2426	0.2376	0.2348	0.2330	0.2315	0.2305	0.2297
	x_{10}	0	0.0816	0.1335	0.1592	0.1746	0.1848	0.1922	0.1977	0.2019
	x_{11}	0.5343	0.6717	0.7425	0.7673	0.7822	0.7921	0.7992	0.8045	0.8090
	x_{12}	0.1013	0.1703	0.2320	0.2597	0.2763	0.2874	0.2953	0.3013	0.3059
	x_{13}	0.3468	0.5109	0.5682	0.5839	0.5936	0.6003	0.6048	0.6082	0.6113
	x_{14}	-0	-0.0576	-0.2346	-0.3333	-0.3920	-0.4309	-0.4591	-0.4802	-0.4963
	x_{15}	0.1393	0.1633	0.1776	0.1799	0.1816	0.1830	0.1836	0.1842	0.1845
df		12	13	15	15	15	15	15	15	15
PE		4.8210	4.0351	3.7432	3.6397	3.5914	3.5649	3.5487	3.5382	3.5307

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.27: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0106	0.0316	0.0310	0.0401	0.0149	0.0566	0.0966	0.0787	0.0767
λ_1		0.0957	0.1265	0.0724	0.0601	0.0149	0.0377	0.0414	0.0197	0.0085
Predictors	x_1	2.6734	2.5634	2.6126	2.5733	2.7196	2.4961	2.3681	2.4205	2.4338
	x_2	0.5492	0.4998	0.5638	0.5512	0.5895	0.5195	0.5012	0.5087	0.5168
	x_3	-0.0422	-0.0205	-0.0790	-0.1242	-0.2380	-0.2184	-0.2087	-0.2696	-0.2959
	x_4	0.5744	0.5076	0.5883	0.5608	0.5811	0.4935	0.4689	0.4638	0.4664
	x_5	-0	-0	-0	-0.0560	-0.1626	-0.1962	-0.1895	-0.2709	-0.2995
	x_6	-0.2379	-0.2463	-0.2687	-0.2837	-0.2683	-0.3096	-0.3348	-0.3340	-0.3362
	x_7	0.7948	0.7190	0.8043	0.8133	0.9453	0.8321	0.7726	0.8351	0.8606
	x_8	0.0514	0	0.0601	0.0700	0.1736	0.0906	0.0600	0.0988	0.1152
	x_9	-0.3196	-0.2967	-0.3407	-0.3525	-0.4055	-0.3753	-0.3681	-0.3920	-0.4056
	x_{10}	0.3335	0.3224	0.3617	0.3788	0.4581	0.4131	0.4196	0.4411	0.4590
	x_{11}	0.4433	0.4181	0.4422	0.4540	0.5513	0.4865	0.4640	0.5023	0.5223
	x_{12}	-0.4259	-0.2425	-0.4303	-0.4297	-0.6742	-0.4204	-0.3007	-0.3913	-0.4208
	x_{13}	0.4102	0.3433	0.4205	0.4401	0.6304	0.4803	0.4352	0.4933	0.5208
	x_{14}	-0	0	-0	-0	-0.2283	-0.0076	0	-0.0190	-0.0570
	x_{15}	0.3617	0.3173	0.3737	0.3872	0.5460	0.4151	0.3869	0.4267	0.4505
df		13	12	13	14	15	15	14	15	15
PE		3.6256	3.9172	3.6026	3.5814	3.2468	3.5441	3.7378	3.5575	3.4918

Data no.27: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.01654181.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165	0.0165
λ_1		0.1489	0.0662	0.0386	0.0248	0.0165	0.0110	0.0071	0.0041	0.0018
Predictors	x_1	2.5998	2.6771	2.7096	2.7090	2.7086	2.7083	2.7081	2.708	2.7079
	x_2	0.4815	0.5803	0.6046	0.5931	0.5864	0.5819	0.5787	0.5762	0.5743
	x_3	-0	-0.0776	-0.1394	-0.1994	-0.2353	-0.2593	-0.2764	-0.2892	-0.2992
	x_4	0.4825	0.6121	0.6274	0.5960	0.5774	0.5650	0.5561	0.5495	0.5442
	x_5	-0	-0	-0.0419	-0.1175	-0.1627	-0.1928	-0.2144	-0.2306	-0.2433
	x_6	-0.2209	-0.2561	-0.2623	-0.2672	-0.2702	-0.2721	-0.2735	-0.2746	-0.2754
	x_7	0.7071	0.8330	0.8915	0.9212	0.9391	0.9511	0.9596	0.9660	0.9710
	x_8	0	0.0838	0.1279	0.1531	0.1682	0.1783	0.1855	0.1909	0.1951
	x_9	-0.2761	-0.3430	-0.3762	-0.3934	-0.4037	-0.4105	-0.4154	-0.4191	-0.4220
	x_{10}	0.2966	0.3586	0.4099	0.4379	0.4546	0.4657	0.4736	0.4796	0.4842
	x_{11}	0.4163	0.4524	0.4971	0.5284	0.5473	0.5599	0.56888	0.5756	0.5809
	x_{12}	-0.2054	-0.5128	-0.6014	-0.6375	-0.6591	-0.6734	-0.6837	-0.6914	-0.6974
	x_{13}	0.3282	0.4515	0.5344	0.5879	0.6199	0.6414	0.6567	0.6681	0.6771
	x_{14}	0	-0.0024	-0.1258	-0.1808	-0.2138	-0.2358	-0.2515	-0.2633	-0.2725
	x_{15}	0.2951	0.3916	0.4672	0.5108	0.5368	0.5541	0.5664	0.5757	0.5829
df		11	14	15	15	15	15	15	15	15
PE		3.9959	3.5115	3.3452	3.2852	3.2572	3.2419	3.2327	3.2266	3.2224

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.28: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0225	0.0348	0.0597	0.0641	0.1394	0.1313	0.1545	0.0720	0.1476
λ_1		0.2022	0.1394	0.1394	0.0961	0.1394	0.0875	0.0662	0.0180	0.0164
Predictors	x_1	1.5670	1.5897	1.5519	1.5588	1.4386	1.4716	1.4446	1.5551	1.4600
	x_2	0	0	0	0	0	0.0029	0.0076	-0.0250	0
	x_3	-0	-0	-0.0046	-0.1006	-0.0397	-0.1269	-0.1717	-0.3004	-0.2775
	x_4	0.3259	0.3910	0.3721	0.3594	0.3044	0.3188	0.2975	0.2782	0.2727
	x_5	-0	-0	-0	-0.0682	-0	-0.0463	-0.1030	-0.3821	-0.2734
	x_6	-0.3368	-0.4029	-0.3998	-0.4625	-0.3921	-0.4606	-0.4840	-0.6044	-0.5629
	x_7	0.9648	1.0077	0.9785	0.9851	0.9013	0.9267	0.9192	1.0012	0.9481
	x_8	-0.0570	-0.1173	-0.1069	-0.1536	-0.0640	-0.1223	-0.1276	-0.2251	-0.1729
	x_9	0	0	0	0.0261	0	0.0254	0.0478	0.1338	0.1131
	x_{10}	0	0.0295	0.0436	0.0727	0.0672	0.0946	0.1133	0.1290	0.1446
	x_{11}	0.7759	0.7527	0.7069	0.7125	0.5896	0.6094	0.5945	0.7791	0.6385
	x_{12}	0.1015	0.1209	0.1450	0.1731	0.1904	0.2103	0.2248	0.2374	0.2496
	x_{13}	0.1282	0.1411	0.1645	0.1985	0.1970	0.2203	0.2367	0.3033	0.2790
	x_{14}	0	0	0	0	0.0539	0.0453	0.0679	0.0305	0.0838
	x_{15}	0	-0	-0	-0	0	-0	-0	-0.0827	-0.0270
df		8	9	10	12	11	14	14	15	14
PE		3.1820	3.0682	3.1072	2.9873	3.2644	3.0997	3.0789	2.7554	2.9185

Data no.28: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02375604.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0238	0.0238	0.0238	0.0238	0.0238	0.0238	0.0238	0.0238	0.0238
λ_1		0.2138	0.0950	0.0554	0.0356	0.0238	0.0158	0.0102	0.0059	0.0026
Predictors	x_1	1.5566	1.6066	1.6115	1.6105	1.5995	1.5922	1.5870	1.5834	1.5805
	x_2	0	0	-0	-0.0315	-0.0857	-0.1218	-0.1476	-0.1666	-0.1814
	x_3	-0	-0.0991	-0.2088	-0.2742	-0.3323	-0.3710	-0.3987	-0.4190	-0.4347
	x_4	0.3112	0.3674	0.3210	0.2839	0.2385	0.2082	0.1866	0.1710	0.1589
	x_5	-0	-0.1497	-0.3270	-0.4380	-0.5375	-0.6037	-0.6511	-0.6858	-0.7129
	x_6	-0.3240	-0.4670	-0.5488	-0.5921	-0.6162	-0.6323	-0.6438	-0.6524	-0.6591
	x_7	0.9524	1.0358	1.0374	1.0403	1.0454	1.0488	1.0512	1.0531	1.0546
	x_8	-0.0437	-0.1717	-0.2174	-0.2444	-0.2646	-0.2780	-0.2877	-0.2948	-0.3004
	x_9	0	0.0301	0.0890	0.1188	0.1381	0.1509	0.1601	0.1670	0.1723
	x_{10}	0	0.0548	0.0871	0.1001	0.1064	0.1106	0.1135	0.1158	0.1175
	x_{11}	0.7720	0.8141	0.8664	0.9033	0.9358	0.9575	0.9730	0.9845	0.9935
	x_{12}	0.1022	0.1308	0.1659	0.1915	0.2130	0.2274	0.2377	0.2453	0.2512
	x_{13}	0.1297	0.1730	0.2436	0.2856	0.3150	0.3345	0.3485	0.3587	0.3667
	x_{14}	0	0	0	0	0	0	0	0	0
	x_{15}	0	-0	-0.0583	-0.0971	-0.1111	-0.1205	-0.1272	-0.1322	-0.1363
df		8	12	13	14	14	14	14	14	14
PE		3.2129	2.9084	2.7643	2.7028	2.6682	2.6497	2.6388	2.6319	2.6273

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.29: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0175	0.0327	0.0387	0.0499	0.0622	0.0933	0.1000	0.1297	0.1386
λ_1		0.1576	0.1309	0.0902	0.0749	0.0622	0.0622	0.0429	0.0324	0.0154
Predictors	x_1	1.3132	1.3183	1.3604	1.3726	1.3822	1.3518	1.3682	1.3393	1.3479
	x_2	0	0	0	0	0.0340	0.0969	0.1020	0.1300	0.1312
	x_3	0	0	0.0029	0.0484	0.1066	0.1986	0.2156	0.2559	0.2685
	x_4	0	0	0.0218	0.0732	0.1317	0.2041	0.2307	0.2624	0.2794
	x_5	-0.4687	-0.4966	-0.5584	-0.5236	-0.4572	-0.3006	-0.3028	-0.2381	-0.2420
	x_6	-0.3582	-0.3801	-0.4169	-0.4299	-0.4394	-0.4304	-0.4450	-0.4442	-0.4522
	x_7	1.3421	1.3472	1.4008	1.4045	1.3983	1.3275	1.3443	1.3025	1.3129
	x_8	0	0	0	0	0	0.02902	0.0465	0.07405	0.0928
	x_9	-0.3281	-0.3500	-0.4010	-0.4227	-0.4374	-0.4231	-0.4465	-0.4419	-0.4598
	x_{10}	0	0	0.0370	0.0596	0.0789	0.0752	0.1009	0.1083	0.1286
	x_{11}	0.9462	0.9436	0.9610	0.9371	0.9049	0.8216	0.8114	0.7533	0.7417
	x_{12}	0	0	0	0	0	0	0	0	0
	x_{13}	0	0	0	0	0	0	0	0	0.0097
	x_{14}	0	0	0	0	0	0.0138	0.0251	0.0593	0.0685
	x_{15}	0.6008	0.6158	0.6377	0.6284	0.6037	0.5470	0.5465	0.5165	0.5129
df		7	7	10	10	11	13	13	13	14
PE		5.9048	5.8718	5.7645	5.7596	5.7750	5.8988	5.8727	5.9572	5.9483

Data no.29: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.04299953.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0430	0.0430	0.0430	0.0430	0.0430	0.0430	0.0430	0.0430	0.0430
λ_1		0.3870	0.1720	0.1003	0.0645	0.0430	0.0287	0.0184	0.0107	0.0048
Predictors	x_1	1.0381	1.2618	1.3424	1.3937	1.4245	1.4462	1.4671	1.4839	1.5038
	x_2	0.0484	0.0138	0	0	0	0	0	0.0044	0.0395
	x_3	0.0825	0.0378	0.0163	0.0273	0.0339	0.0408	0.0563	0.0716	0.1066
	x_4	0	0	0.0255	0.0634	0.0862	0.1035	0.1258	0.1496	0.1890
	x_5	-0	-0.3570	-0.5214	-0.5720	-0.6024	-0.6204	-0.6237	-0.6224	-0.5941
	x_6	-0.1316	-0.3399	-0.4069	-0.4406	-0.4608	-0.4743	-0.4856	-0.5007	-0.5171
	x_7	0.9764	1.2683	1.3736	1.4377	1.4763	1.5014	1.5162	1.5228	1.5254
	x_8	0	0	0	0	0	0	0	-0.0034	-0.0353
	x_9	-0.0994	-0.3069	-0.3851	-0.4398	-0.4726	-0.4947	-0.5114	-0.5238	-0.5331
	x_{10}	-0	0	0.0220	0.0755	0.1076	0.1286	0.1412	0.1483	0.1527
	x_{11}	0.7403	0.8880	0.9436	0.9610	0.9714	0.9799	0.9940	1.0084	1.0276
	x_{12}	0	0	0	-0	-0	-0.0067	-0.0385	-0.0586	-0.0840
	x_{13}	0	0	0	0	-0	-0	-0.0037	-0.0253	-0.0514
	x_{14}	0	0	0	0	0	-0	-0	-0.0022	-0.0180
	x_{15}	0.4334	0.5760	0.6265	0.6429	0.6528	0.6615	0.6778	0.6954	0.7034
df		8	9	10	10	10	11	12	15	15
PE		7.1463	6.0627	5.8109	5.7104	5.6641	5.6367	5.6099	5.5885	5.5692

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.30: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0162	0.0068	0.0155	0.0290	0.0300	0.0374	0.0530	0.0687	0.0884
λ_1		0.1459	0.0273	0.0361	0.0435	0.0300	0.0249	0.0227	0.0172	0.0098
Predictors	x_1	1.6834	1.6765	1.6650	1.6490	1.6451	1.6346	1.6138	1.5932	1.5685
	x_2	0.5770	0.6996	0.6790	0.6677	0.6858	0.6901	0.6811	0.6780	0.6759
	x_3	0.2086	0.5131	0.4650	0.4232	0.4585	0.4591	0.4315	0.4157	0.4001
	x_4	0	0.2101	0.1691	0.1405	0.1702	0.1742	0.1634	0.1616	0.1615
	x_5	-0.3872	-0.3611	-0.3785	-0.3786	-0.3701	-0.3675	-0.3780	-0.3830	-0.3858
	x_6	0	0.2081	0.1708	0.1329	0.1691	0.1734	0.1569	0.1508	0.1459
	x_7	1.5473	1.8709	1.8241	1.7725	1.8061	1.8016	1.7661	1.7408	1.7121
	x_8	0.1073	0.2360	0.2253	0.2168	0.2349	0.2418	0.2420	0.2463	0.2522
	x_9	-0.0415	-0.0494	-0.0551	-0.0610	-0.0637	-0.0680	-0.0721	-0.0770	-0.0833
	x_{10}	0.1958	0.6584	0.5937	0.5276	0.5768	0.5752	0.5351	0.5107	0.4857
	x_{11}	0.9555	0.9960	0.9804	0.9513	0.9512	0.9382	0.9114	0.8866	0.8578
	x_{12}	-0.1227	-0.6064	-0.5358	-0.4682	-0.5192	-0.5190	-0.4917	-0.4782	-0.4642
	x_{13}	0	0	0	0	0	0.0021	0.0235	0.0415	0.0601
	x_{14}	0.2128	0.3267	0.3187	0.3095	0.3201	0.3216	0.3118	0.3057	0.2999
	x_{15}	0	0.2960	0.2633	0.2343	0.2655	0.2709	0.2627	0.2617	0.2614
df		11	14	14	14	14	15	15	15	15
PE		2.8063	2.2142	2.2541	2.3132	2.2711	2.2749	2.3119	2.3414	2.3794

Data no.30: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02421479.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242	0.0242
λ_1		0.2179	0.0969	0.0565	0.0363	0.0242	0.0161	0.0104	0.0061	0.0027
Predictors	x_1	1.6534	1.6626	1.6573	1.6537	1.6516	1.6503	1.6492	1.6483	1.6477
	x_2	0.5970	0.5897	0.6483	0.6776	0.6952	0.7072	0.7156	0.7215	0.7264
	x_3	0.1222	0.2775	0.3923	0.4500	0.4846	0.5079	0.5244	0.5364	0.5460
	x_4	0	0.0114	0.1112	0.1604	0.1899	0.2099	0.2239	0.2340	0.2422
	x_5	-0.2874	-0.4270	-0.3926	-0.3763	-0.3666	-0.3596	-0.3550	-0.3521	-0.3494
	x_6	-0	0	0.1024	0.1589	0.1927	0.2153	0.2314	0.2435	0.2529
	x_7	1.4029	1.6412	1.7484	1.8029	1.8355	1.8573	1.8728	1.8845	1.8936
	x_8	0.0776	0.1472	0.1989	0.2260	0.2422	0.2530	0.2608	0.2666	0.2712
	x_9	-0.0117	-0.0532	-0.0571	-0.0598	-0.0615	-0.0625	-0.0633	-0.0639	-0.0644
	x_{10}	0.0200	0.3329	0.4890	0.5682	0.6156	0.6473	0.6699	0.6869	0.7001
	x_{11}	0.8605	0.9595	0.9608	0.9622	0.9631	0.9636	0.9640	0.9644	0.9647
	x_{12}	-0	-0.2606	-0.4266	-0.5097	-0.5594	-0.5927	-0.6165	-0.6341	-0.6480
	x_{13}	0	0	0	0	0	0	0	0	-0
	x_{14}	0.1236	0.2689	0.3005	0.3165	0.3261	0.3324	0.3370	0.3406	0.3433
	x_{15}	0	0.1069	0.2059	0.2552	0.2845	0.3041	0.3182	0.3288	0.3370
df		10	13	14	14	14	14	14	14	14
PE		3.2502	2.5609	2.3504	2.2758	2.2415	2.2230	2.2119	2.2047	2.1998

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.31: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .

	α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0192	0.0359	0.0561	0.0602	0.1437	0.1024	0.2105	0.0676	0.2205
λ_1		0.1731	0.1437	0.1309	0.0902	0.1437	0.0683	0.0902	0.0169	0.0245
Predictors	x_1	1.7402	1.7269	1.6981	1.7193	1.5490	1.6630	1.4958	1.8591	1.5243
	x_2	0.8355	0.8084	0.8098	0.7637	0.8467	0.7804	0.8359	0.7621	0.8019
	x_3	-0	-0.0807	-0.0934	-0.2328	-0	-0.2211	-0.0285	-0.4021	-0.1699
	x_4	0.1236	0.1096	0.1091	0.0942	0.1008	0.1059	0.1104	0.3006	0.1166
	x_5	-0.7908	-0.8595	-0.8386	-0.9918	-0.5790	-0.9115	-0.5853	-1.0067	-0.7192
	x_6	-0.2785	-0.3065	-0.3190	-0.3563	-0.2943	-0.3833	-0.3502	-0.4112	-0.4284
	x_7	1.9800	1.9858	1.9662	2.0058	1.8177	1.9614	1.7794	2.1347	1.8357
	x_8	-0	-0	-0	-0	-0	-0.0164	-0.0238	-0.0423	-0.0844
	x_9	-0.1775	-0.2179	-0.2424	-0.2820	-0.2669	-0.3200	-0.3225	-0.4068	-0.3797
	x_{10}	0.0212	0.0583	0.0746	0.1210	0.0656	0.1491	0.1242	0.2336	0.2018
	x_{11}	0.4818	0.5055	0.5057	0.5467	0.4606	0.5317	0.4593	0.6456	0.4903
	x_{12}	-0	-0	-0	-0	0	-0	0	-0.1089	-0
	x_{13}	0.6703	0.6636	0.6397	0.6653	0.5347	0.6194	0.5122	0.7320	0.5341
	x_{14}	0.1085	0.1545	0.1807	0.2237	0.2211	0.2563	0.2644	0.2864	0.3072
	x_{15}	-0	-0	-0	-0	0	-0	0	-0.2018	-0
df		11	12	12	12	12	13	13	15	13
PE		5.8448	5.7512	5.7648	5.5981	6.1485	5.6656	6.1625	5.2670	5.9230

Data no.31: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.04228049.

	α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0423	0.0423	0.0423	0.0423	0.0423	0.0423	0.0423	0.0423	0.0423
λ_1		0.3805	0.1691	0.0987	0.0634	0.0423	0.0282	0.0181	0.0106	0.0047
Predictors	x_1	1.5676	1.6968	1.7456	1.8064	1.8715	1.9134	1.9436	1.9659	1.9835
	x_2	0.8390	0.8409	0.7510	0.7310	0.7438	0.7497	0.7543	0.7573	0.7598
	x_3	0	-0	-0.2434	-0.3612	-0.4021	-0.4295	-0.4485	-0.4634	-0.4747
	x_4	0	0.1169	0.0856	0.1481	0.2523	0.3180	0.3658	0.4006	0.4281
	x_5	-0.1576	-0.7343	-1.0357	-1.1099	-1.0872	-1.0700	-1.0569	-1.0481	-1.0408
	x_6	-0	-0.2797	-0.3463	-0.3671	-0.3755	-0.3847	-0.3914	-0.3963	-0.4001
	x_7	1.7333	1.9461	2.0275	2.0900	2.1374	2.1704	2.1941	2.2117	2.2255
	x_8	-0	-0	-0	-0	-0.0008	-0.0116	-0.0193	-0.0250	-0.0295
	x_9	-0.0413	-0.1995	-0.2645	-0.3207	-0.3593	-0.3863	-0.4056	-0.4201	-0.4313
	x_{10}	0	0.0297	0.1114	0.1493	0.1854	0.2128	0.2325	0.2472	0.2586
	x_{11}	0.4249	0.4782	0.5547	0.6053	0.6438	0.6683	0.6859	0.6991	0.7093
	x_{12}	0	-0	-0	-0.0216	-0.0798	-0.1174	-0.1444	-0.1645	-0.1802
	x_{13}	0.4732	0.6324	0.6899	0.7318	0.7592	0.7759	0.7878	0.7968	0.8038
	x_{14}	0.1124	0.1382	0.2087	0.2421	0.2542	0.2637	0.2702	0.2754	0.2793
	x_{15}	0	-0	-0	-0.0890	-0.1693	-0.2226	-0.2609	-0.2893	-0.3116
df		8	11	12	14	15	15	15	15	15
PE		6.9857	5.9024	5.5666	5.3881	5.2772	5.2171	5.1809	5.1576	5.1416

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.32: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0350	0.0596	0.0403	0.1448	0.1643	0.1864	0.1659	0.1786	0.2096
λ_1		0.3151	0.2383	0.0940	0.2172	0.1643	0.1243	0.0711	0.0447	0.0233
Predictors	x_1	2.4058	2.4064	2.5578	2.2263	2.2213	2.1998	2.2713	2.2557	2.1826
	x_2	-0	-0	-0.1358	-0	-0.0268	-0.0784	-0.1504	-0.1894	-0.2433
	x_3	-0	-0	-0.2048	-0	-0	-0.0431	-0.1402	-0.1814	-0.2311
	x_4	0	0.0084	0.0264	0.0521	0.0865	0.1022	0.1007	0.1037	0.0922
	x_5	0	0	0	0	0	0	-0	-0.0170	-0.1057
	x_6	0	0	0	0	0	0	0.0077	0.0249	0.0381
	x_7	0.2724	0.3577	0.5239	0.3481	0.4068	0.4361	0.4933	0.5109	0.5176
	x_8	-0	-0	-0	-0	-0	-0	-0	-0	-0.0260
	x_9	-0	-0	-0.0940	-0	-0	-0.0321	-0.0967	-0.1268	-0.1462
	x_{10}	-0.0639	-0.1234	-0.2474	-0.1254	-0.1570	-0.1835	-0.2317	-0.2516	-0.2622
	x_{11}	0.0167	0.0562	0.0246	0.0959	0.1083	0.1135	0.1012	0.1062	0.1250
	x_{12}	0	0	-0	0	0	0	0	0	0.0047
	x_{13}	0.4708	0.4721	0.5074	0.4453	0.4450	0.4456	0.4613	0.4628	0.4665
	x_{14}	0.5142	0.5183	0.6386	0.4663	0.4701	0.4812	0.5312	0.5428	0.5442
	x_{15}	0	0	0.0377	0.0251	0.0343	0.0502	0.0632	0.0735	0.0935
df		6	7	11	8	9	11	12	13	15
PE		4.0808	3.9463	3.5343	4.1296	4.0530	3.9966	3.7899	3.7637	3.7994

Data no.32: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.01591737.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0159	0.0159	0.0159	0.0159	0.0159	0.0159	0.0159	0.0159	0.0159
λ_1		0.1433	0.0637	0.0371	0.0239	0.0159	0.0106	0.0068	0.0040	0.0018
Predictors	x_1	2.5941	2.6450	2.6612	2.6692	2.6605	2.6544	2.6496	2.6293	2.6131
	x_2	-0.0599	-0.1900	-0.2323	-0.2533	-0.2834	-0.3016	-0.3153	-0.3494	-0.3760
	x_3	-0.1225	-0.3036	-0.3657	-0.3960	-0.4364	-0.4567	-0.4725	-0.5244	-0.5657
	x_4	0.0117	0	0.0059	0.0114	0	0	-0.0010	-0.0506	-0.0899
	x_5	0	-0	-0	-0	-0.0407	-0.0661	-0.0854	-0.1526	-0.2058
	x_6	0	0	0	0	0	0	-0	-0.0032	-0.0087
	x_7	0.4899	0.5578	0.5592	0.5596	0.5608	0.5651	0.5678	0.5604	0.5545
	x_8	0	-0	-0	-0	-0	-0	-0	0	0.0022
	x_9	-0.0253	-0.1502	-0.1885	-0.2075	-0.2197	-0.2267	-0.2320	-0.2367	-0.2401
	x_{10}	-0.2076	-0.2882	-0.3478	-0.3790	-0.3957	-0.4069	-0.4150	-0.4219	-0.4274
	x_{11}	0.0171	0	0	0	0.0016	0.0064	0.0096	0.0192	0.0271
	x_{12}	0	-0	-0.0949	-0.1471	-0.1694	-0.1861	-0.1980	-0.1943	-0.1924
	x_{13}	0.5135	0.5191	0.5523	0.5703	0.5844	0.5926	0.5985	0.6101	0.6194
	x_{14}	0.6172	0.7065	0.7714	0.8054	0.8330	0.8485	0.8603	0.8798	0.8962
	x_{15}	0.0086	0.0489	0.0745	0.0875	0.0994	0.1036	0.1073	0.1206	0.1315
df		11	9	11	11	12	12	13	14	15
PE		3.6134	3.4420	3.3738	3.3491	3.3364	3.3302	3.3266	3.3223	3.3195

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.33: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0163	0.0191	0.0298	0.0976	0.0578	0.1662	0.0846	0.1205	0.1551
λ_1		0.1464	0.0764	0.0696	0.1464	0.0578	0.1108	0.0363	0.0301	0.0172
Predictors	x_1	2.3995	2.4625	2.4408	2.2188	2.3703	2.0914	2.3169	2.2399	2.1743
	x_2	0.1072	0.1431	0.1383	0.0572	0.1371	0.0635	0.1481	0.1462	0.1469
	x_3	0.0360	0	0	0.0977	0.0119	0.1074	0.0199	0.0510	0.0673
	x_4	0.4324	0.4331	0.4159	0.3289	0.3846	0.2741	0.3650	0.3426	0.3185
	x_5	-0.2520	-0.4702	-0.4767	-0.1651	-0.4599	-0.1925	-0.4741	-0.4231	-0.4062
	x_6	-0	-0	-0.0095	-0.0276	-0.0592	-0.0996	-0.1017	-0.1341	-0.1618
	x_7	1.7632	1.9026	1.8852	1.5733	1.8284	1.4992	1.8071	1.7323	1.6795
	x_8	0	0	0	0	0.0442	0.0699	0.1022	0.1360	0.1680
	x_9	-0.1006	-0.2327	-0.2358	-0.0499	-0.2332	-0.0745	-0.2537	-0.2353	-0.2325
	x_{10}	-0.4541	-0.5022	-0.4993	-0.4110	-0.4852	-0.4012	-0.4859	-0.4710	-0.4651
	x_{11}	0.8978	0.9321	0.9074	0.7263	0.8391	0.6392	0.7847	0.7222	0.6742
	x_{12}	0.0599	0.0951	0.1138	0.1347	0.1574	0.1811	0.1862	0.2044	0.2162
	x_{13}	0.0371	0.0851	0.1003	0.1072	0.1265	0.1406	0.1445	0.1563	0.1647
	x_{14}	0	0	0	0	0	0.0424	0	0.0114	0.0328
	x_{15}	0	0	0	0	0	0	0.0211	0.0203	0.0262
df		11	10	11	12	13	14	14	15	15
PE		4.9345	4.7300	4.7393	5.2247	4.7843	5.3397	4.8106	4.9103	4.9966

Data no.33: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03103979.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0310	0.0310	0.0310	0.0310	0.0310	0.0310	0.0310	0.0310	0.0310
λ_1		0.2794	0.1242	0.0724	0.0466	0.0310	0.0207	0.0133	0.0078	0.0034
Predictors	x_1	2.2253	2.3855	2.4344	2.4578	2.4716	2.4843	2.4819	2.4804	2.4793
	x_2	0	0.1030	0.1347	0.1594	0.1741	0.1925	0.1936	0.1950	0.1963
	x_3	0.0681	0.0316	0	0	-0	-0	-0.0471	-0.0827	-0.1102
	x_4	0.3225	0.4039	0.4123	0.4211	0.4231	0.4456	0.4465	0.4485	0.4505
	x_5	-0	-0.3117	-0.4686	-0.5361	-0.5821	-0.6046	-0.6656	-0.7109	-0.7456
	x_6	-0	-0	-0.0096	-0.0238	-0.0291	-0.0367	-0.0473	-0.0553	-0.0617
	x_7	1.4887	1.7683	1.8761	1.9307	1.9673	2.0091	2.0587	2.0968	2.1262
	x_8	0	0	0	0.0225	0.0450	0.0685	0.0950	0.1149	0.1304
	x_9	-0	-0.1320	-0.2294	-0.2818	-0.3165	-0.3420	-0.3657	-0.3836	-0.3974
	x_{10}	-0.3641	-0.4632	-0.4966	-0.5163	-0.5324	-0.5396	-0.5425	-0.5447	-0.5462
	x_{11}	0.8013	0.8730	0.9032	0.9065	0.9050	0.9060	0.9170	0.9257	0.9325
	x_{12}	0.0672	0.0899	0.1146	0.1210	0.1208	0.1227	0.1362	0.1462	0.1541
	x_{13}	0.0194	0.0708	0.1005	0.1083	0.1104	0.1106	0.1167	0.1207	0.1238
	x_{14}	0	0	0	0	-0	-0.0265	-0.0601	-0.0859	-0.1062
	x_{15}	0	0	0	0.0149	0.0367	0.0618	0.0999	0.1289	0.1510
df		8	11	11	13	13	11	15	15	15
PE		5.4745	4.8984	4.7485	4.6952	4.6694	4.6468	4.6207	4.6041	4.5931

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.34: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0237	0.0641	0.1099	0.1558	0.2129	0.1382	0.1482	0.2541	0.3590
λ_1		0.2129	0.2564	0.2564	0.2336	0.2129	0.0922	0.0635	0.0635	0.0399
Predictors	x_1	1.8427	1.7082	1.6305	1.5808	1.5151	1.7713	1.7734	1.5979	1.4785
	x_2	0.0667	0.1021	0.1273	0.1506	0.1717	0.1672	0.1695	0.2078	0.2312
	x_3	-0	-0	-0	-0	-0	-0.1278	-0.2226	-0.1877	-0.2117
	x_4	0.7647	0.7679	0.7573	0.7435	0.7209	0.7614	0.7359	0.6960	0.6551
	x_5	-0.4774	-0.2872	-0.2288	-0.2292	-0.2212	-0.4812	-0.5346	-0.4349	-0.3968
	x_6	-0.5201	-0.4883	-0.4755	-0.4810	-0.4831	-0.6100	-0.6477	-0.6092	-0.5982
	x_7	0.5055	0.4305	0.4061	0.4083	0.4047	0.5482	0.5578	0.5051	0.4743
	x_8	0.0278	0	0.0021	0.0225	0.0391	0.1617	0.1989	0.1878	0.1995
	x_9	0	0	0	0.0062	0.0166	0.1351	0.1732	0.1501	0.1593
	x_{10}	-0.0065	-0	-0	-0.0012	-0.0113	-0.1017	-0.1235	-0.1038	-0.1082
	x_{11}	0.6646	0.5832	0.5433	0.5166	0.4827	0.6156	0.6285	0.5277	0.4765
	x_{12}	-0	-0	-0	-0	-0	-0.1045	-0.1443	-0.0925	-0.0819
	x_{13}	0	0	0	0	0	0.0285	0.0589	0.0715	0.0938
	x_{14}	0.2810	0.2574	0.2645	0.2726	0.2727	0.3409	0.3535	0.3174	0.3031
	x_{15}	0	0	0	0.0168	0.0454	0.0512	0.0742	0.0993	0.1209
df		10	8	9	12	12	15	15	15	15
PE		5.9191	6.2270	6.3580	6.3780	6.4344	5.6251	5.5099	5.8142	5.9845

Data no.34: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.0535475.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0535	0.0535	0.0535	0.0535	0.0535	0.0535	0.0535	0.0535	0.0535
λ_1		0.4819	0.2142	0.1249	0.0803	0.0535	0.0357	0.0229	0.0134	0.0059
Predictors	x_1	1.4699	1.7824	1.9213	1.9686	1.9805	1.9883	1.994	1.9982	2.0014
	x_2	0.0158	0.0960	0.1153	0.0991	0.0793	0.0660	0.0566	0.0494	0.0437
	x_3	0	-0	-0.0345	-0.1973	-0.3322	-0.4223	-0.4866	-0.5349	-0.5727
	x_4	0.6699	0.7702	0.8082	0.7736	0.7159	0.6771	0.6496	0.6287	0.6122
	x_5	-0	-0.4066	-0.5246	-0.6348	-0.7387	-0.8083	-0.8579	-0.8953	-0.9248
	x_6	-0.2976	-0.5141	-0.5839	-0.6474	-0.6923	-0.7221	-0.7434	-0.7593	-0.7717
	x_7	0.1705	0.4851	0.5822	0.6015	0.6090	0.6140	0.6176	0.6203	0.6223
	x_8	0	0.0324	0.1237	0.1773	0.2170	0.2435	0.2624	0.2766	0.2877
	x_9	0	0.0058	0.1017	0.1825	0.2294	0.2607	0.2830	0.2998	0.3128
	x_{10}	-0	-0.0087	-0.0844	-0.1393	-0.1633	-0.1792	-0.1906	-0.1992	-0.2058
	x_{11}	0.5025	0.6225	0.6976	0.7686	0.8050	0.8292	0.8464	0.8594	0.8695
	x_{12}	0	-0	-0.1017	-0.1782	-0.2281	-0.2614	-0.2852	-0.3030	-0.3168
	x_{13}	0	0	0	0.02501	0.0682	0.0972	0.1177	0.1333	0.1455
	x_{14}	0.1941	0.2806	0.3494	0.4005	0.4242	0.4401	0.4514	0.4600	0.4668
	x_{15}	0	0	0	0.0063	0.0322	0.0495	0.0619	0.07125	0.0784
df		7	11	13	15	15	15	15	15	15
PE		7.2445	6.0067	5.5520	5.3126	5.1920	5.1277	5.0897	5.0655	5.0491

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.35: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0094	0.0371	0.0275	0.0822	0.0486	0.0879	0.0491	0.1113	0.0565
λ_1		0.0850	0.1485	0.0643	0.1233	0.0486	0.0586	0.0210	0.0278	0.0063
Predictors	x_1	1.4502	1.4475	1.4887	1.4285	1.5174	1.4882	1.551	1.4993	1.5685
	x_2	0.0847	0.1521	0.1355	0.1648	0.1811	0.2044	0.2096	0.2259	0.2320
	x_3	0.0761	0.2399	0.1578	0.2930	0.2522	0.3165	0.2835	0.3654	0.3197
	x_4	-0.3475	-0	-0.2707	-0	-0.1796	-0.0538	-0.2013	-0.0549	-0.1835
	x_5	-1.2760	-0.7548	-1.1288	-0.6673	-0.9650	-0.7564	-0.9928	-0.7229	-0.9607
	x_6	0.1484	0.1021	0.1676	0.1220	0.1811	0.1759	0.1928	0.1993	0.1992
	x_7	1.9821	1.9409	1.9812	1.8919	1.9743	1.9196	1.9916	1.9071	1.9941
	x_8	0	0	0	0	0	0	0	0	0.0022
	x_9	0.6208	0.5619	0.6393	0.5805	0.6495	0.6302	0.6656	0.6508	0.6718
	x_{10}	0	0	0	-0	0	-0	0.0156	-0	0.0323
	x_{11}	1.4001	1.0921	1.3171	0.9984	1.2328	1.0711	1.2949	1.0552	1.2957
	x_{12}	-0	-0	-0.0024	-0	-0.0449	-0.0143	-0.1012	-0.0572	-0.1292
	x_{13}	0.2565	0.2015	0.2750	0.2429	0.2961	0.2877	0.3306	0.3189	0.3488
	x_{14}	-0	-0	-0	-0	-0	-0	-0.0348	-0	-0.0534
	x_{15}	-0.0124	-0.0149	-0.0656	-0.0340	-0.0991	-0.1048	-0.1150	-0.1235	-0.1271
df		11	10	12	11	12	12	14	12	15
PE		2.8651	3.2652	2.8852	3.3445	2.9221	3.1200	2.8397	3.1027	2.8278

Data no.35: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02464791.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0246	0.0246	0.0246	0.0246	0.0246	0.0246	0.0246	0.0246	0.0246
λ_1		0.2218	0.0986	0.0575	0.0370	0.0246	0.0164	0.0106	0.0062	0.0027
Predictors	x_1	1.3881	1.4727	1.4911	1.5162	1.5330	1.5443	1.5525	1.5587	1.5635
	x_2	0.1119	0.1364	0.1298	0.1509	0.1662	0.1764	0.1847	0.1911	0.1959
	x_3	0.2042	0.1731	0.1471	0.1699	0.1866	0.1978	0.2064	0.2129	0.2178
	x_4	-0	-0.1802	-0.3039	-0.3214	-0.3260	-0.3290	-0.3306	-0.3316	-0.3326
	x_5	-0.6313	-1.0195	-1.1768	-1.2039	-1.2174	-1.2264	-1.2319	-1.2360	-1.2395
	x_6	0.0398	0.1413	0.1723	0.1802	0.1835	0.1856	0.1866	0.1872	0.1876
	x_7	1.9178	1.9718	1.9868	2.0004	2.0098	2.0160	2.0203	2.0234	2.0258
	x_8	0	0	0	0	0	0	0.0036	0.0068	0.0093
	x_9	0.5230	0.6077	0.6456	0.6557	0.6589	0.6610	0.6622	0.6630	0.6636
	x_{10}	-0	0	0	0.0174	0.0402	0.0554	0.0663	0.0744	0.0807
	x_{11}	1.0502	1.2568	1.3503	1.4048	1.4417	1.4663	1.4844	1.4981	1.5088
	x_{12}	-0	-0	-0.0143	-0.0609	-0.0905	-0.1103	-0.1238	-0.1336	-0.1412
	x_{13}	0.1689	0.2409	0.2834	0.3111	0.3281	0.3394	0.3469	0.3525	0.3569
	x_{14}	-0	-0	-0	-0.0300	-0.0557	-0.0728	-0.0855	-0.0951	-0.1025
	x_{15}	-0	-0.0415	-0.0644	-0.0749	-0.0804	-0.0841	-0.0868	-0.0889	-0.0905
df		9	11	12	14	14	14	15	15	15
PE		3.4860	2.9865	2.8522	2.7897	2.7572	2.7395	2.7286	2.7216	2.7167

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.36: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0105	0.0544	0.0443	0.0756	0.0858	0.0809	0.1146	0.1486	0.0998
λ_1		0.0942	0.2176	0.1034	0.1135	0.0858	0.0539	0.0491	0.0372	0.0111
Predictors	x_1	1.9968	1.8628	1.9089	1.8335	1.8130	1.8244	1.7569	1.6965	1.7780
	x_2	0	0	0	0	0	-0	-0	-0	-0.0595
	x_3	0	0	0	0	0	0	0	0	-0.0021
	x_4	0	0.0487	0.0369	0.0732	0.0858	0.0834	0.1175	0.1467	0.0999
	x_5	-0.3389	-0.0501	-0.2760	-0.2277	-0.2851	-0.3729	-0.3501	-0.3469	-0.4831
	x_6	-0.1236	-0	-0.1207	-0.1132	-0.1473	-0.1856	-0.1910	-0.2029	-0.2476
	x_7	1.1668	0.8833	1.0675	0.9868	1.0020	1.0510	0.9941	0.9526	1.0761
	x_8	-0.0529	-0	-0.0180	-0	-0.0475	-0.1300	-0.1225	-0.1319	-0.2232
	x_9	0.1574	0.0487	0.1399	0.1249	0.1552	0.1974	0.1950	0.2018	0.2578
	x_{10}	-0.3191	-0.2545	-0.3437	-0.3596	-0.3838	-0.3988	-0.4198	-0.4402	-0.4328
	x_{11}	0.9106	0.7674	0.8396	0.7775	0.7774	0.8061	0.7570	0.7199	0.8077
	x_{12}	0	0	0.0179	0.0496	0.0595	0.0525	0.0826	0.1076	0.0758
	x_{13}	0	0	0	0	0	0	0	0	-0.0046
	x_{14}	0.0460	0.0998	0.0957	0.1311	0.1554	0.1713	0.2029	0.2324	0.2221
	x_{15}	0.6197	0.5116	0.5887	0.5567	0.5654	0.5881	0.5656	0.5500	0.6079
df		10	9	12	11	12	12	12	12	15
PE		4.2457	4.7221	4.3377	4.4434	4.3879	4.2912	4.3663	4.4288	4.2403

Data no.36: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.02881948.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0288	0.0288	0.0288	0.0288	0.0288	0.0288	0.0288	0.0288	0.0288
λ_1		0.2594	0.1153	0.0672	0.0432	0.0288	0.0192	0.0124	0.0072	0.0032
Predictors	x_1	1.9017	1.9489	1.9455	1.9393	1.9187	1.8983	1.8728	1.8467	1.8264
	x_2	0	0	-0	-0.0234	-0.0774	-0.1180	-0.1648	-0.2110	-0.2469
	x_3	0	0	-0	-0	-0.0376	-0.0787	-0.1423	-0.2117	-0.2656
	x_4	0.0196	0.0151	0.0097	0.0026	0	-0	-0.0538	-0.1279	-0.1856
	x_5	-0	-0.2604	-0.4031	-0.4869	-0.5727	-0.6391	-0.7356	-0.8387	-0.9189
	x_6	-0	-0.1024	-0.1645	-0.2008	-0.2325	-0.2590	-0.2735	-0.2816	-0.2880
	x_7	0.8734	1.0888	1.1503	1.1844	1.2146	1.2350	1.2511	1.2641	1.2742
	x_8	0	-0	-0.1271	-0.1970	-0.2392	-0.2704	-0.2897	-0.3026	-0.3125
	x_9	0.0347	0.1285	0.1956	0.2357	0.2585	0.2734	0.2838	0.2915	0.2975
	x_{10}	-0.2007	-0.3215	-0.3513	-0.3654	-0.3734	-0.3776	-0.3811	-0.3841	-0.3865
	x_{11}	0.7825	0.8642	0.9002	0.9218	0.9427	0.9630	0.9921	1.0231	1.0473
	x_{12}	0	0	0	0	0	0	0.0113	0.0262	0.0377
	x_{13}	0	0	0	-0	-0.0013	-0.0201	-0.0243	-0.0216	-0.0196
	x_{14}	0.0670	0.0667	0.0997	0.1203	0.1441	0.1677	0.1874	0.2041	0.2171
	x_{15}	0.4991	0.5953	0.6256	0.6448	0.6688	0.6956	0.7195	0.7405	0.7569
df		8	10	11	12	13	13	15	15	15
PE		4.7857	4.3389	4.2069	4.1545	4.1167	4.0899	4.0655	4.0445	4.0300

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.37: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0189	0.0352	0.0604	0.0939	0.1065	0.1456	0.1714	0.2439	0.3140
λ_1		0.1697	0.1408	0.1408	0.1408	0.1065	0.0971	0.0734	0.0610	0.0349
Predictors	x_1	1.7767	1.7659	1.7250	1.6757	1.6777	1.631	1.6108	1.5278	1.4616
	x_2	0	0.0194	0.0009	0	0.0258	0.0225	0.0421	0.0327	0.0365
	x_3	0.4324	0.4631	0.4647	0.4682	0.4997	0.5079	0.5280	0.5253	0.5221
	x_4	0.7868	0.7845	0.7695	0.7547	0.7591	0.7451	0.7425	0.7080	0.6731
	x_5	-0	-0	-0	-0	-0	-0	-0	-0.0210	-0.0645
	x_6	0.4896	0.5295	0.4961	0.4639	0.5155	0.4949	0.5129	0.4743	0.4669
	x_7	1.2187	1.2317	1.1941	1.155	1.1796	1.1453	1.146	1.0909	1.0581
	x_8	0.6332	0.6605	0.6464	0.6244	0.6579	0.6445	0.6511	0.6152	0.6007
	x_9	0	0	0	0	0	0	0	0	0.0070
	x_{10}	-0.0252	-0.0552	-0.0591	-0.0626	-0.0933	-0.1030	-0.1214	-0.1288	-0.1445
	x_{11}	0.4397	0.4382	0.4103	0.3823	0.3902	0.3679	0.3565	0.3172	0.2997
	x_{12}	0.4420	0.4397	0.4256	0.4096	0.4124	0.3994	0.3923	0.3693	0.3560
	x_{13}	0.0725	0.0853	0.1213	0.1509	0.1518	0.1757	0.1791	0.1995	0.2123
	x_{14}	0	0	0	0	0	0.0030	0.0215	0.0669	0.0985
	x_{15}	0	0	0	0	0	0	0	0	0.0059
df		10	11	11	10	11	12	12	13	15
PE		5.2486	5.1954	5.2612	5.3431	5.2654	5.3345	5.3330	5.4864	5.5843

Data no.37: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.03745729.

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0375	0.0375	0.0375	0.0375	0.0375	0.0375	0.0375	0.0375	0.0375
λ_1		0.3371	0.1498	0.0874	0.0562	0.0375	0.0250	0.0161	0.0094	0.0042
Predictors	x_1	1.6550	1.7565	1.7952	1.8059	1.8102	1.8230	1.8269	1.8297	1.8320
	x_2	0	0.0026	0.1060	0.1739	0.2227	0.2586	0.2861	0.3074	0.3240
	x_3	0.3121	0.4538	0.5195	0.5436	0.5557	0.5700	0.5815	0.5910	0.5984
	x_4	0.7289	0.7794	0.8049	0.8270	0.8465	0.8674	0.8850	0.8995	0.9107
	x_5	0	-0	-0	-0.0452	-0.0885	-0.0992	-0.1095	-0.1161	-0.1212
	x_6	0.1821	0.5054	0.6499	0.7232	0.7680	0.8015	0.8256	0.8437	0.8577
	x_7	1.0101	1.2146	1.3085	1.3669	1.4062	1.4279	1.4489	1.4649	1.4774
	x_8	0.3795	0.6490	0.7207	0.7503	0.7665	0.7875	0.7979	0.8050	0.8106
	x_9	0	0	0	0	-0	-0	-0.0179	-0.0331	-0.0449
	x_{10}	-0	-0.0476	-0.1036	-0.1469	-0.1799	-0.2045	-0.2224	-0.2355	-0.2457
	x_{11}	0.3264	0.4292	0.4723	0.5119	0.5457	0.5848	0.6101	0.6286	0.6430
	x_{12}	0.3646	0.43605	0.4523	0.4823	0.5099	0.5310	0.5496	0.5635	0.5744
	x_{13}	0.1258	0.0952	0.0528	0.0715	0.1005	0.1238	0.1439	0.1584	0.1698
	x_{14}	0	0	0	-0	-0.0055	-0.0511	-0.0813	-0.1037	-0.1212
	x_{15}	0	0	-0	-0.0742	-0.1514	-0.2015	-0.2407	-0.2702	-0.2933
df		9	11	11	13	14	14	15	15	15
PE		5.9707	5.2325	5.0476	4.9224	4.8397	4.7803	4.7431	4.7193	4.7031

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.38: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0223	0.0503	0.0652	0.1113	0.1151	0.1433	0.2447	0.2401	0.4086
λ_1		0.2011	0.2011	0.1522	0.1670	0.1151	0.0956	0.1049	0.0600	0.0454
Predictors	x_1	1.3423	1.2975	1.3311	1.2494	1.3017	1.2868	1.1590	1.2028	1.0321
	x_2	0	0	0	0	0	0	0	0	0.0087
	x_3	0.1882	0.2021	0.2386	0.2445	0.2715	0.2855	0.3021	0.3125	0.3308
	x_4	-0	-0	-0	0	-0	-0	0	0	0
	x_5	0	0	0	0	0	0	0.0602	0.1184	0.1846
	x_6	-0	-0	-0	-0	-0	-0	-0	-0.0395	-0.0557
	x_7	1.1160	1.0757	1.0984	1.0236	1.0586	1.0407	0.9352	0.9747	0.8473
	x_8	-0	-0	-0	-0	-0	-0.0007	-0.0121	-0.0463	-0.0712
	x_9	0	0	0.0146	0.0224	0.0596	0.0825	0.0993	0.1217	0.1520
	x_{10}	-0	-0	-0	-0	-0	-0	-0.0129	-0.0554	-0.0842
	x_{11}	0.3669	0.3794	0.3893	0.4010	0.4021	0.4031	0.3906	0.3913	0.3621
	x_{12}	-0	-0	-0	0	-0	-0	-0	-0	-0
	x_{13}	0.6418	0.6207	0.6193	0.5869	0.5827	0.5608	0.4878	0.4784	0.3938
	x_{14}	0	0	0	0.0057	0.0305	0.0575	0.0934	0.1138	0.1305
	x_{15}	0	0	0	0	0	0	0.0200	0.0179	0.0976
df		5	5	6	7	7	8	11	12	13
PE		4.3268	4.3709	4.3206	4.4150	4.3423	4.3599	4.5534	4.4684	4.7549

Data no.38: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03222113.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322
λ_1		0.2900	0.1289	0.0752	0.0483	0.0322	0.0215	0.0138	0.0081	0.0036
Predictors	x_1	1.2173	1.4129	1.4642	1.5154	1.5572	1.5853	1.6051	1.6199	1.6331
	x_2	0	0	0	-0	-0	-0	-0	-0	-0
	x_3	0.1352	0.2392	0.2773	0.2751	0.3023	0.3213	0.3342	0.3436	0.3546
	x_4	0	-0	-0	-0	-0	-0	-0	0	0.0107
	x_5	0	0	0	0	0.1185	0.2007	0.2597	0.3038	0.3450
	x_6	-0	-0	-0	-0.0422	-0.0850	-0.1139	-0.1344	-0.1497	-0.1627
	x_7	1.0078	1.1706	1.2004	1.2273	1.2533	1.2708	1.2835	1.2930	1.3020
	x_8	-0	-0	-0	-0	-0	-0	-0.0018	-0.0038	-0.0051
	x_9	0	0.0144	0.0459	0.0467	0.0393	0.0343	0.0307	0.0282	0.0254
	x_{10}	-0	-0	-0	-0.0100	-0.0319	-0.0468	-0.0580	-0.0667	-0.0756
	x_{11}	0.3699	0.3760	0.4030	0.4385	0.4503	0.4580	0.4635	0.4676	0.4699
	x_{12}	0	-0	-0.0409	-0.1056	-0.1689	-0.2118	-0.2424	-0.2652	-0.2836
	x_{13}	0.6161	0.6469	0.6766	0.7180	0.7237	0.7270	0.7289	0.7302	0.7299
	x_{14}	0	0	0	0.0673	0.1299	0.1721	0.2025	0.2252	0.2412
	x_{15}	0	-0	-0.0104	-0.0971	-0.1583	-0.1990	-0.2280	-0.2496	-0.2689
df		5	6	8	11	12	12	13	13	14
PE		4.4979	4.2487	4.1677	4.0652	4.0037	3.9713	3.9526	3.9411	3.9331

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.39: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		2.8308E-05	0.0014	0.0028	0.0048	0.0003	0.0047	0.0186	0.0010	0.0058
λ_1		0.0003	0.0055	0.0066	0.0073	0.0003	0.0031	0.0080	0.0003	0.0006
Predictors	x_1	1.5319	1.5749	1.5987	1.6246	1.5357	1.6069	1.733	1.5465	1.6096
	x_2	-1.9255	-1.7899	-1.7192	-1.6424	-1.915	-1.7057	-1.2948	-1.8851	-1.7048
	x_3	-2.6993	-2.4587	-2.3327	-2.1960	-2.6806	-2.3073	-1.5821	-2.6268	-2.3049
	x_4	-2.7391	-2.4697	-2.3265	-2.1702	-2.7176	-2.2929	-1.466	-2.6559	-2.2869
	x_5	-4.7580	-4.4323	-4.2576	-4.0659	-4.7318	-4.2131	-3.1917	-4.6562	-4.2036
	x_6	0.6238	0.6200	0.6187	0.6172	0.6238	0.6203	0.6051	0.6237	0.6217
	x_7	1.5856	1.5833	1.5825	1.5814	1.5855	1.5836	1.5676	1.5856	1.5844
	x_8	0.5208	0.5149	0.5139	0.5135	0.5208	0.5183	0.5134	0.5210	0.5214
	x_9	-0.7853	-0.7526	-0.7373	-0.7211	-0.7832	-0.7384	-0.6482	-0.7773	-0.7410
	x_{10}	-0.7238	-0.6435	-0.6041	-0.5621	-0.7183	-0.6023	-0.3741	-0.7024	-0.6058
	x_{11}	0.7789	0.7578	0.7480	0.7375	0.7775	0.7486	0.6860	0.7737	0.75013
	x_{12}	1.6559	1.5588	1.5046	1.4444	1.6476	1.4863	1.1664	1.6236	1.4803
	x_{13}	1.5981	1.5164	1.4708	1.4201	1.5912	1.4555	1.1847	1.5711	1.4506
	x_{14}	0.5214	0.4750	0.4526	0.4293	0.5183	0.4520	0.3339	0.5090	0.4542
	x_{15}	0.3031	0.2697	0.2535	0.2364	0.3008	0.2530	0.1646	0.2942	0.2545
df		15	15	15	15	15	15	15	15	15
PE		4.9195	4.9334	4.9510	4.9784	4.9197	4.9561	5.2120	4.9211	4.9571

Data no.39: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.08699396.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0870	0.0870	0.0870	0.0870	0.0870	0.0870	0.0870	0.0870	0.0870
λ_1		0.7829	0.3480	0.2030	0.1305	0.0870	0.0580	0.0373	0.0217	0.0097
Predictors	x_1	1.2002	1.5921	1.7186	1.7836	1.8150	1.8315	1.8372	1.8233	1.8145
	x_2	-0	-0	-0.1757	-0.3155	-0.4103	-0.4768	-0.5339	-0.6075	-0.6688
	x_3	0	0	0	-0	-0.0694	-0.1859	-0.2881	-0.4243	-0.5331
	x_4	0	0.1602	0.0949	0.0416	0	0	-0.0324	-0.1643	-0.2732
	x_5	-0	-0.0924	-0.5893	-0.8638	-1.0559	-1.1946	-1.3184	-1.4877	-1.6240
	x_6	0	0.0150	0.2322	0.3378	0.4052	0.4562	0.4893	0.5063	0.5195
	x_7	0.7039	1.0436	1.2178	1.3011	1.3558	1.4019	1.4312	1.4422	1.4521
	x_8	0	0.1679	0.3075	0.3776	0.4190	0.4504	0.4736	0.4883	0.5004
	x_9	-0	-0	-0.2048	-0.3115	-0.3737	-0.4155	-0.4467	-0.4704	-0.4926
	x_{10}	0	0	0	0	0	0	0	0	-0.0203
	x_{11}	0.2582	0.3817	0.4230	0.4554	0.4802	0.4953	0.5077	0.5301	0.5474
	x_{12}	0.1094	0.1979	0.3748	0.4699	0.5272	0.5569	0.5855	0.6315	0.6677
	x_{13}	0.3872	0.4548	0.5658	0.6208	0.6562	0.6733	0.6899	0.7206	0.7477
	x_{14}	0.2298	0.1886	0.1511	0.1394	0.1492	0.1671	0.1815	0.2044	0.2239
	x_{15}	0	0	0	0	0	0.0138	0.0332	0.0553	0.0732
df		6	10	12	12	12	13	14	14	15
PE		10.4902	8.7074	7.4606	6.9568	6.6828	6.5087	6.3836	6.2390	6.1275

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.40: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0341	0.0439	0.0686	0.1698	0.1328	0.1654	0.1774	0.3662	0.4715
λ_1		0.3069	0.1756	0.1600	0.2548	0.1328	0.1103	0.0760	0.0916	0.0524
Predictors	x_1	1.6207	1.7491	1.7208	1.4623	1.6339	1.5913	1.5917	1.3459	1.2634
	x_2	-0	-0	-0	-0	-0.0248	-0.0820	-0.1451	-0.1423	-0.1883
	x_3	0.8689	0.8100	0.7863	0.7637	0.7382	0.7042	0.6757	0.6399	0.6097
	x_4	0.1095	0.0866	0.0801	0.0913	0.0734	0.0646	0.0539	0.0919	0.1032
	x_5	-0	-0.1738	-0.2014	-0.0777	-0.2451	-0.2872	-0.3449	-0.2689	-0.2851
	x_6	0	0	0	0	0	0	0	0	0.0201
	x_7	0.8759	1.0148	1.0073	0.8229	0.9741	0.9669	0.9895	0.8337	0.8029
	x_8	-0	-0.1823	-0.1917	-0.0125	-0.1904	-0.1958	-0.2266	-0.1258	-0.1288
	x_9	0	0.0134	0.0283	0	0.0538	0.0744	0.1049	0.0867	0.1104
	x_{10}	0.2486	0.3069	0.3136	0.2609	0.3215	0.3289	0.3447	0.3130	0.3202
	x_{11}	0.3695	0.4324	0.4288	0.3584	0.4176	0.4194	0.4313	0.3722	0.3599
	x_{12}	0	-0	-0	0	-0	-0	-0	0	0
	x_{13}	0.1488	0.2081	0.2198	0.1877	0.2403	0.2584	0.2809	0.2607	0.2645
	x_{14}	0	0.0560	0.0706	0.0444	0.0977	0.1198	0.1454	0.1389	0.1525
	x_{15}	0	-0	-0	0	0	-0	-0	0	0.0085
df		7	11	11	10	12	12	12	12	14
PE		5.7300	5.2631	5.2578	5.8605	5.2982	5.2880	5.2066	5.6780	5.7898

Data no.40: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.04150941.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0415	0.0415	0.0415	0.0415	0.0415	0.0415	0.0415	0.0415	0.0415
λ_1		0.3736	0.1660	0.0969	0.0623	0.0415	0.0277	0.0178	0.0104	0.0046
Predictors	x_1	1.5449	1.7647	1.8247	1.8328	1.8429	1.8517	1.8580	1.8627	1.8688
	x_2	-0	-0	-0.0529	-0.1677	-0.2337	-0.2728	-0.3007	-0.3215	-0.3363
	x_3	0.8346	0.8053	0.7323	0.6873	0.6785	0.6791	0.6795	0.6799	0.6793
	x_4	0.0682	0.0829	0.0301	0.0163	0.0367	0.0639	0.0834	0.0981	0.1106
	x_5	-0	-0.1896	-0.3433	-0.4485	-0.4879	-0.4961	-0.5020	-0.5063	-0.5085
	x_6	0	0	-0	-0	-0	-0	-0	-0	-0.0087
	x_7	0.8022	1.0287	1.1095	1.1319	1.1421	1.1531	1.1610	1.1669	1.1725
	x_8	-0	-0.1995	-0.3054	-0.3621	-0.4042	-0.4363	-0.4591	-0.4763	-0.4895
	x_9	0	0.0222	0.0841	0.1012	0.1067	0.1084	0.1096	0.1105	0.1126
	x_{10}	0.2080	0.3127	0.3520	0.3620	0.3637	0.3625	0.3616	0.3609	0.3588
	x_{11}	0.3507	0.4367	0.4743	0.5386	0.5858	0.6170	0.6392	0.6558	0.6662
	x_{12}	0	-0	-0	-0	-0.0124	-0.0406	-0.0607	-0.0758	-0.0890
	x_{13}	0.1447	0.2131	0.2639	0.3141	0.3458	0.3671	0.3823	0.3937	0.4046
	x_{14}	0	0.0618	0.1199	0.1759	0.2142	0.2442	0.2656	0.2817	0.2954
	x_{15}	0	-0	-0	-0.0751	-0.1469	-0.1941	-0.2279	-0.2532	-0.2712
df		7	11	12	13	14	14	14	14	15
PE		5.9533	5.2259	5.0226	4.8928	4.8187	4.7762	4.7516	4.7363	4.7261

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.41: The naïve elastic net estimates $(\hat{\beta})$ using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0094	0.0233	0.0332	0.0429	0.0643	0.0879	0.0859	0.1616	0.1434
λ_1		0.0850	0.0933	0.0775	0.0643	0.0643	0.0586	0.0368	0.0404	0.0159
	x_1	2.0860	2.0443	2.0402	2.0307	1.9873	1.9451	1.9598	1.8224	1.8652
	x_2	0.0971	0.0856	0.0969	0.1022	0.0994	0.0979	0.1018	0.0831	0.0909
	x_3	0.1508	0.1571	0.1820	0.1987	0.2199	0.2412	0.2409	0.2907	0.2852
	x_4	0.0783	0.0742	0.0968	0.1067	0.1195	0.1290	0.1160	0.1397	0.1262
	x_5	-0	-0	-0	-0.0142	-0	-0	-0.0687	-0.0021	-0.0840
	x_6	0.2986	0.2935	0.3048	0.3153	0.3134	0.3160	0.3369	0.3195	0.3442
	x_7	1.7516	1.7037	1.7048	1.7028	1.6531	1.6107	1.6496	1.4923	1.5613
Predictors	x_8	-0.2913	-0.2695	-0.2814	-0.2888	-0.2752	-0.2669	-0.2866	-0.2549	-0.2822
	x_9	-0.7792	-0.7583	-0.7607	-0.7619	-0.7407	-0.7236	-0.7442	-0.6761	-0.7094
	x_{10}	-0.5736	-0.5502	-0.5593	-0.5646	-0.5472	-0.5351	-0.5554	-0.5161	-0.5455
	x_{11}	0.3799	0.3880	0.3860	0.3870	0.3903	0.3931	0.3971	0.3910	0.3971
	x_{12}	0.2871	0.2802	0.2794	0.2801	0.2732	0.2695	0.2796	0.2591	0.2718
	x_{13}	0.4190	0.4084	0.4007	0.3966	0.3816	0.3699	0.3824	0.3345	0.3536
	x_{14}	0	0	0	0	0	0	0	0.0412	0.0381
	x_{15}	0.0817	0.0875	0.0959	0.1055	0.1113	0.1201	0.1355	0.1340	0.1482
df		13	13	13	14	13	13	14	15	15
PE		2.9102	2.9496	2.9365	2.9285	2.9735	3.0153	2.9597	3.1727	3.0596

Data no.41: The naïve elastic net estimates $(\hat{\beta})$ using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02579272.

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0258	0.0258	0.0258	0.0258	0.0258	0.0258	0.0258	0.0258	0.0258
λ_1		0.2321	0.1032	0.0602	0.0387	0.0258	0.0172	0.0111	0.0064	0.0029
	x_1	1.8935	2.0275	2.0633	2.0647	2.0655	2.0700	2.0752	2.0794	2.0826
	x_2	0	0.0762	0.0980	0.0889	0.0834	0.0843	0.0919	0.0982	0.1029
	x_3	0.0611	0.1500	0.1692	0.1511	0.1402	0.1414	0.1416	0.1417	0.1418
	x_4	0	0.0635	0.0736	0.0278	0.0002	0	0	0	-0
	x_5	-0	-0	-0.0672	-0.1830	-0.2526	-0.2833	-0.3067	-0.3243	-0.3381
	x_6	0.2028	0.2868	0.3233	0.3496	0.3654	0.3752	0.3831	0.3891	0.3938
	x_7	1.4864	1.6822	1.7534	1.7933	1.8172	1.8327	1.8453	1.8547	1.8621
Predictors	x_8	-0.0836	-0.2554	-0.3019	-0.3172	-0.3264	-0.3332	-0.3333	-0.3329	-0.3327
	x_9	-0.6316	-0.7482	-0.7854	-0.8087	-0.8227	-0.8298	-0.8361	-0.8410	-0.8448
	x_{10}	-0.3727	-0.5364	-0.5823	-0.6004	-0.6112	-0.6181	-0.6164	-0.6145	-0.6132
	x_{11}	0.4332	0.3923	0.3872	0.3961	0.4015	0.4031	0.4071	0.4107	0.4132
	x_{12}	0.2407	0.2777	0.2919	0.3076	0.3173	0.3210	0.3256	0.3293	0.3321
	x_{13}	0.3941	0.4068	0.4185	0.4405	0.4538	0.4579	0.4678	0.4755	0.4817
	x_{14}	0	0	0	0	-0	-0	-0.0227	-0.0417	-0.0562
	x_{15}	0.0533	0.0866	0.1089	0.1359	0.1518	0.1590	0.1692	0.1771	0.1833
df		11	13	14	14	14	13	14	14	14
PE		3.3962	2.9759	2.8867	2.8473	2.8292	2.8203	2.8120	2.8065	2.8030

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.42: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0234	0.0578	0.0823	0.1280	0.1594	0.1809	0.2129	0.2292	0.2450
λ_1		0.2108	0.2313	0.1920	0.1920	0.1594	0.1206	0.0912	0.0573	0.0272
Predictors	x_1	2.3049	2.2176	2.2130	2.1306	2.1012	2.0867	2.0501	2.0426	2.0322
	x_2	-0	-0	-0	-0	-0.0181	-0.0725	-0.1021	-0.1335	-0.1502
	x_3	0.1514	0.1135	0.1202	0.0946	0.0915	0.0904	0.0861	0.0904	0.0965
	x_4	0	0	0	0	0	0	0	0	0
	x_5	-0.6036	-0.5299	-0.5798	-0.5366	-0.5683	-0.6231	-0.6412	-0.6743	-0.6911
	x_6	0	0	0.0041	0.0064	0.0313	0.0560	0.0796	0.1032	0.1202
	x_7	0.6695	0.6118	0.6398	0.6047	0.6188	0.6467	0.6548	0.6764	0.6899
	x_8	-0	-0	-0	-0	-0	-0	-0	-0.0066	-0.0415
	x_9	-0	-0	-0	-0	-0	-0	-0.0244	-0.0556	-0.0923
	x_{10}	0.0518	0.0392	0.0627	0.0622	0.0804	0.1074	0.1298	0.1557	0.1794
	x_{11}	0.9217	0.8394	0.8148	0.7500	0.7318	0.7355	0.7237	0.7334	0.7459
	x_{12}	-0	-0	-0	-0	-0	-0	-0	-0.0100	-0.0335
	x_{13}	0.0224	0.0385	0.0713	0.0899	0.1156	0.1419	0.1603	0.1813	0.2022
	x_{14}	0.0793	0.0915	0.1096	0.1220	0.1353	0.1510	0.1630	0.1789	0.2008
	x_{15}	-0	-0	-0	-0	-0	-0	-0	-0.0103	-0.0353
df		8	8	9	9	10	10	11	14	14
PE		6.1429	6.3370	6.2658	6.4276	6.3988	6.3009	6.2917	6.2098	6.1389

Data no.42: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03774683.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0377	0.0377	0.0377	0.0377	0.0377	0.0377	0.0377	0.0377	0.0377
λ_1		0.3397	0.1510	0.0881	0.0566	0.0377	0.0252	0.0162	0.0094	0.0042
Predictors	x_1	2.1383	2.3155	2.3515	2.3434	2.322	2.3098	2.3148	2.3156	2.2954
	x_2	-0	-0.0615	-0.1773	-0.2576	-0.3229	-0.3649	-0.3823	-0.3971	-0.4258
	x_3	0.0567	0.1525	0.1552	0.1002	0.0377	0	0	-0.0055	-0.0502
	x_4	0	0	-0	-0.1061	-0.2205	-0.2939	-0.3252	-0.3527	-0.4048
	x_5	-0.3386	-0.7299	-0.8721	-1.0174	-1.1515	-1.2374	-1.2719	-1.3031	-1.3662
	x_6	0	0.0355	0.0792	0.1015	0.1122	0.1191	0.1218	0.1237	0.1277
	x_7	0.4938	0.7360	0.8207	0.8577	0.8788	0.8929	0.9028	0.9102	0.9159
	x_8	-0	-0	-0	-0	-0	-0	-0	-0	-0.0020
	x_9	-0	-0	-0	-0.0252	-0.0381	-0.0469	-0.0551	-0.0608	-0.0632
	x_{10}	0	0.0921	0.1484	0.1823	0.2007	0.2129	0.2195	0.2241	0.2303
	x_{11}	0.8131	0.9336	0.9958	1.0463	1.0838	1.1081	1.1211	1.1315	1.1462
	x_{12}	-0	-0	-0	-0	-0	-0	-0.0045	-0.0090	-0.0093
	x_{13}	0	0.0746	0.1337	0.1911	0.2366	0.2662	0.2830	0.2969	0.3144
	x_{14}	0.0488	0.1023	0.1332	0.1741	0.2081	0.2299	0.2403	0.2496	0.2656
	x_{15}	-0	-0	-0.0480	-0.0799	-0.0861	-0.0910	-0.1000	-0.1050	-0.0999
df		6	10	11	13	13	12	13	14	15
PE		6.7698	5.9441	5.7708	5.5647	5.4868	5.4445	5.4233	5.4087	5.3919

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.43: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0118	0.0242	0.0378	0.0489	0.0554	0.0629	0.0386	0.0726	0.1236
λ_1		0.1063	0.0969	0.0883	0.0733	0.0554	0.0419	0.0165	0.0182	0.0137
Predictors	x_1	1.9579	1.9510	1.9407	1.9423	1.9544	1.9594	2.0433	1.9799	1.8981
	x_2	-0	-0	-0	-0	-0	0	0.0412	0.0197	0.0063
	x_3	0.0933	0.1331	0.1687	0.1960	0.2141	0.2312	0.2019	0.2575	0.3101
	x_4	0.2566	0.2993	0.3379	0.3825	0.4256	0.4594	0.5135	0.5304	0.5568
	x_5	-1.0124	-0.9723	-0.9333	-0.9102	-0.9028	-0.8917	-0.9453	-0.8698	-0.7836
	x_6	-0.0970	-0.0988	-0.1002	-0.1085	-0.1209	-0.1294	-0.1618	-0.1432	-0.1296
	x_7	1.3197	1.298	1.2761	1.2584	1.2474	1.2364	1.2929	1.2362	1.1693
	x_8	0	0	0	0	0	0	0.0882	0.0412	0.0040
	x_9	0.2329	0.2399	0.2457	0.2615	0.2831	0.2982	0.3531	0.3336	0.3173
	x_{10}	-0	-0	-0	-0	-0	-0	-0.0157	-0.0209	-0.0312
	x_{11}	1.0779	1.0454	1.0124	0.9970	0.9979	0.9928	1.0776	0.9896	0.8919
	x_{12}	-0	-0	-0	-0.0292	-0.0760	-0.1072	-0.2075	-0.1595	-0.1209
	x_{13}	0.4401	0.4432	0.4447	0.4644	0.4950	0.5145	0.5920	0.5487	0.5102
	x_{14}	-0.3514	-0.3608	-0.3653	-0.3751	-0.3888	-0.3966	-0.4761	-0.4240	-0.3709
	x_{15}	0.3323	0.3381	0.3434	0.3607	0.3845	0.4009	0.4489	0.4249	0.4105
df		11	11	11	12	12	12	15	15	15
PE		3.9132	3.9203	3.9362	3.9127	3.8646	3.8415	3.6903	3.7908	3.9459

Data no.43: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.03106005.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0311	0.0311	0.0311	0.0311	0.0311	0.0311	0.0311	0.0311	0.0311
λ_1		0.2795	0.1242	0.0725	0.0466	0.0311	0.0207	0.0133	0.0078	0.0035
Predictors	x_1	1.7977	1.9230	1.9667	1.9998	2.0217	2.0432	2.0648	2.0811	2.0938
	x_2	-0	-0	-0	0	0	0.0222	0.0586	0.0859	0.1072
	x_3	0.0542	0.1391	0.1595	0.1655	0.1602	0.1711	0.1920	0.2074	0.2199
	x_4	0.3164	0.3162	0.3437	0.3961	0.4301	0.4742	0.5250	0.5631	0.5929
	x_5	-0.5611	-0.8863	-0.9767	-0.9934	-1.0019	-0.9866	-0.9558	-0.9327	-0.9146
	x_6	-0	-0.0623	-0.1192	-0.1423	-0.1586	-0.1658	-0.1682	-0.1699	-0.1713
	x_7	1.2039	1.2778	1.2873	1.2857	1.2920	1.3005	1.3100	1.3170	1.3226
	x_8	-0	-0	0	0.0099	0.0487	0.0817	0.1108	0.1325	0.1496
	x_9	0.0142	0.2071	0.2684	0.3054	0.3318	0.3499	0.3624	0.3717	0.3790
	x_{10}	-0	-0	-0	-0	-0	-0.0079	-0.0177	-0.0248	-0.0307
	x_{11}	0.8549	0.9958	1.0491	1.0762	1.0930	1.0997	1.1006	1.1014	1.1019
	x_{12}	-0	-0	-0.0364	-0.1184	-0.1676	-0.2027	-0.2304	-0.2515	-0.2675
	x_{13}	0.2344	0.4040	0.4803	0.5389	0.5766	0.5975	0.6081	0.6161	0.6223
	x_{14}	-0	-0.2884	-0.3968	-0.4281	-0.4518	-0.4743	-0.4967	-0.5134	-0.5265
	x_{15}	0.1633	0.3048	0.3683	0.4094	0.4357	0.4505	0.4597	0.4669	0.4719
df		9	11	12	13	13	15	15	15	15
PE		4.9126	4.0524	3.8511	3.7596	3.7118	3.6838	3.6656	3.6541	3.6465

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.44: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0221	0.0723	0.1029	0.1103	0.1039	0.1710	0.1834	0.1240	0.2316
λ_1		0.1993	0.2891	0.2400	0.1654	0.1039	0.1140	0.0786	0.0310	0.0257
Predictors	x_1	2.1101	1.8747	1.8829	1.9681	2.0861	1.9179	1.9562	2.2216	1.9518
	x_2	0.0556	0	0.0026	0.0406	0.1056	0.0511	0.0898	0.2057	0.1073
	x_3	0.2151	0.2282	0.2666	0.3022	0.3488	0.3542	0.3897	0.4532	0.4363
	x_4	0.4071	0.4213	0.4187	0.4099	0.4405	0.4117	0.4444	0.5745	0.4941
	x_5	-0.3580	-0	-0.0898	-0.2816	-0.4174	-0.3048	-0.3330	-0.4392	-0.3481
	x_6	-0.0927	-0.0592	-0.0633	-0.0822	-0.0940	-0.0946	-0.0993	-0.1082	-0.1208
	x_7	1.0844	0.9034	0.9339	1.0170	1.1036	1.0062	1.0331	1.1617	1.0297
	x_8	-0	-0	-0	-0	-0	-0.0183	-0.0436	-0.0761	-0.1215
	x_9	0	0	0	0	0	0	0	0.0849	0.0724
	x_{10}	0	0	0	0.0320	0.0557	0.0837	0.1069	0.1045	0.1532
	x_{11}	1.1552	1.0047	0.9798	1.0056	1.0469	0.9443	0.9454	1.0465	0.9097
	x_{12}	0	0	0	0	0	0	0	-0	0.0405
	x_{13}	0.0967	0.1132	0.1441	0.1654	0.1934	0.1990	0.2214	0.2627	0.2497
	x_{14}	-0	-0	-0	-0	-0.0592	-0	-0.0474	-0.2277	-0.1153
	x_{15}	0.1637	0.1045	0.1438	0.1948	0.2375	0.2283	0.2515	0.2836	0.2743
df		10	8	10	11	12	12	13	14	15
PE		5.0245	5.6326	5.5126	5.2025	4.9058	5.2230	5.0989	4.6261	5.0405

Data no.44: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.02985552.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0299	0.0299	0.0299	0.0299	0.0299	0.0299	0.0299	0.0299	0.0299
λ_1		0.2687	0.1194	0.0697	0.0448	0.0299	0.0199	0.0128	0.0075	0.0033
Predictors	x_1	1.9906	2.2511	2.4237	2.5421	2.6322	2.6923	2.7352	2.7674	2.7924
	x_2	0.0194	0.1454	0.2816	0.3715	0.4442	0.4927	0.5273	0.5533	0.5735
	x_3	0.2157	0.2844	0.4079	0.4880	0.5472	0.5868	0.6149	0.6361	0.6526
	x_4	0.4457	0.4425	0.5932	0.7008	0.7834	0.8385	0.8777	0.9072	0.9302
	x_5	-0.0807	-0.5473	-0.5279	-0.4824	-0.4304	-0.3956	-0.3710	-0.3524	-0.3379
	x_6	-0.0902	-0.0836	-0.0813	-0.0824	-0.0803	-0.0788	-0.0778	-0.0770	-0.0764
	x_7	0.9731	1.1958	1.2743	1.3125	1.3365	1.3525	1.3639	1.3725	1.3792
	x_8	-0	-0	-0	-0	-0	-0	-0	-0	0
	x_9	0	0	0.0385	0.0816	0.1078	0.1253	0.1377	0.1471	0.1544
	x_{10}	0	0.0001	0.0123	0.0182	0.0219	0.0243	0.02603	0.0273	0.0283
	x_{11}	1.0918	1.1849	1.2045	1.2120	1.2157	1.2181	1.2198	1.2211	1.2221
	x_{12}	0	-0	-0	-0.0403	-0.0950	-0.1314	-0.1575	-0.1770	-0.1922
	x_{13}	0.0709	0.1621	0.2336	0.2795	0.3146	0.3379	0.3546	0.3671	0.3769
	x_{14}	-0	-0.0917	-0.2978	-0.4052	-0.4711	-0.5152	-0.5465	-0.5701	-0.5884
	x_{15}	0.1014	0.2317	0.2622	0.2742	0.2816	0.2866	0.2901	0.2928	0.2948
df		10	12	13	14	14	14	14	14	14
PE		5.3846	4.6941	4.4305	4.3147	4.2487	4.2121	4.1898	4.1750	4.1647

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.45: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0141	0.0263	0.0410	0.0401	0.0601	0.0682	0.0732	0.1254	0.1615
λ_1		0.1266	0.1051	0.0957	0.0601	0.0601	0.0455	0.0314	0.0314	0.0179
Predictors	x_1	1.4851	1.4860	1.4689	1.5057	1.4769	1.4812	1.5045	1.4105	1.3727
	x_2	0	0	0	-0	0	0	0	0.0588	0.0993
	x_3	0.4635	0.4584	0.4612	0.4562	0.4900	0.5048	0.5172	0.5660	0.5882
	x_4	0	0	0	0.0117	0.0573	0.0852	0.1075	0.1805	0.2204
	x_5	-0.8635	-0.9306	-0.9263	-1.0846	-0.9893	-1.0085	-1.0308	-0.8414	-0.7788
	x_6	0.0678	0.0872	0.0919	0.1396	0.1331	0.1498	0.1604	0.1426	0.1427
	x_7	1.2434	1.2585	1.2443	1.3322	1.2900	1.3084	1.3325	1.2258	1.1894
	x_8	-0	-0	-0	-0.0537	-0.0572	-0.0823	-0.0983	-0.0950	-0.1042
	x_9	0.1424	0.1623	0.1736	0.1972	0.2008	0.2117	0.2262	0.2370	0.2545
	x_{10}	0.6397	0.6489	0.6479	0.6705	0.6633	0.6699	0.6827	0.6457	0.6322
	x_{11}	0.7630	0.7380	0.7140	0.7088	0.6771	0.6640	0.6698	0.5818	0.5422
	x_{12}	0.3524	0.3549	0.3515	0.3714	0.3561	0.3577	0.3676	0.3286	0.3165
	x_{13}	0.1496	0.1800	0.1861	0.2310	0.2086	0.2158	0.2362	0.1841	0.1760
	x_{14}	0.0410	0.0716	0.0894	0.1181	0.1186	0.1282	0.1378	0.1412	0.1518
	x_{15}	-0	-0	-0	-0	-0	-0.0020	-0.0413	-0	-0
df		11	11	11	13	13	14	14	14	14
PE		3.5385	3.4871	3.4966	3.3675	3.4244	3.4009	3.3535	3.5317	3.5966

Data no.45: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.04505885.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0451	0.0451	0.0451	0.0451	0.0451	0.0451	0.0451	0.0451	0.0451
λ_1		0.4055	0.1802	0.1051	0.0676	0.0451	0.0300	0.0193	0.0113	0.0050
Predictors	x_1	1.1419	1.3689	1.4513	1.4911	1.5326	1.572	1.6001	1.6188	1.6333
	x_2	0	0	0	0	-0	-0	-0	-0.0162	-0.0297
	x_3	0.4856	0.4985	0.4671	0.4654	0.4711	0.4764	0.4802	0.4751	0.4711
	x_4	0	0.00783	0.0014	0.0197	0.0379	0.0506	0.0596	0.0547	0.0506
	x_5	-0	-0.5371	-0.8719	-1.0279	-1.1094	-1.1579	-1.1925	-1.2327	-1.2643
	x_6	0	0	0.0797	0.1281	0.1512	0.1617	0.1692	0.1760	0.1815
	x_7	0.7254	1.0774	1.2181	1.3031	1.3599	1.3985	1.4261	1.4483	1.4657
	x_8	-0	-0	-0	-0.0422	-0.0717	-0.0854	-0.0952	-0.1037	-0.1105
	x_9	0	0.1172	0.1685	0.1931	0.2132	0.2301	0.2421	0.2489	0.2539
	x_{10}	0.4573	0.5839	0.6387	0.6642	0.6851	0.7030	0.7159	0.7264	0.7348
	x_{11}	0.6167	0.7117	0.7084	0.7024	0.7149	0.7345	0.74841	0.7620	0.7729
	x_{12}	0.1564	0.3172	0.3461	0.3634	0.3826	0.4001	0.4127	0.4224	0.4301
	x_{13}	0	0.0572	0.1699	0.2161	0.2544	0.2887	0.3131	0.3335	0.3492
	x_{14}	0	0.0222	0.0839	0.1121	0.1279	0.1382	0.1456	0.1538	0.1600
	x_{15}	0	0	-0	-0	-0.0438	-0.1012	-0.1421	-0.1679	-0.1878
df		6	11	12	13	14	14	14	15	15
PE		5.1356	3.9194	3.5457	3.4053	3.3139	3.2501	3.2117	3.1841	3.1650

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.46: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV.

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0472	0.0881	0.0717	0.1344	0.2213	0.2756	0.1541	0.2898	0.3732
λ_1		0.4244	0.3523	0.1674	0.2016	0.2213	0.1837	0.0660	0.0725	0.0415
Predictors	x_1	1.6428	1.6328	1.7870	1.6497	1.5073	1.4592	1.7068	1.5112	1.4340
	x_2	-0	-0.0182	-0.2985	-0.2300	-0.1886	-0.2281	-0.4090	-0.3377	-0.3326
	x_3	0.3217	0.3738	0.3737	0.4152	0.4445	0.4626	0.4431	0.5148	0.5510
	x_4	0	0	0	0	0	0	0	0.0386	0.0775
	x_5	0	0	0	0	0	0	0.0594	0.0941	0.1344
	x_6	-0	-0	-0	-0	-0	-0	-0	-0	-0.0184
	x_7	0.4947	0.5431	0.8365	0.6863	0.5625	0.5562	0.8376	0.6637	0.6217
	x_8	0	0	0.1913	0.1124	0.07013	0.1159	0.3244	0.2599	0.2733
	x_9	0	0.0340	0.2678	0.2037	0.1638	0.1943	0.3664	0.3070	0.3131
	x_{10}	-0	-0	-0	-0	-0	-0	0	0	0
	x_{11}	0.6480	0.6167	0.6269	0.5918	0.5509	0.5306	0.5755	0.5195	0.4905
	x_{12}	0	0	0.1680	0.1341	0.1327	0.1615	0.2467	0.2165	0.2212
	x_{13}	0	0	0	0	0	0.0233	0.0353	0.0499	0.0639
	x_{14}	0.4119	0.4303	0.5184	0.4718	0.4409	0.4343	0.4993	0.4428	0.4242
	x_{15}	0.0916	0.1323	0.0514	0.1266	0.1767	0.1837	0.0897	0.1596	0.1722
df		6	8	10	10	10	11	12	13	14
PE		7.0862	6.9658	6.1463	6.4437	6.7642	6.7595	6.0592	6.4237	6.5534

Data no.46: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.03253366.

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0325	0.0325	0.0325	0.0325	0.0325	0.0325	0.0325	0.0325	0.0325
λ_1		0.2928	0.1301	0.0759	0.0488	0.0325	0.0217	0.0139	0.0081	0.0036
Predictors	x_1	1.7865	1.8966	1.9333	1.9570	1.9714	1.9724	1.9756	1.9807	1.9847
	x_2	-0.0891	-0.37461	-0.4673	-0.5039	-0.5258	-0.5710	-0.6015	-0.6179	-0.6304
	x_3	0.3626	0.32907	0.3187	0.3215	0.3232	0.2931	0.2759	0.2697	0.2642
	x_4	0	0	0	-0	-0	-0.0508	-0.0777	-0.0897	-0.1004
	x_5	0	0	0.0070	0.0357	0.0530	0.0189	0	0	-0
	x_6	-0	-0	0	0	0	0.0130	0.0223	0.0301	0.0371
	x_7	0.7140	0.97321	1.0534	1.0949	1.1198	1.1417	1.1625	1.1794	1.1934
	x_8	0	0.2795	0.3777	0.4293	0.4603	0.4944	0.5238	0.5456	0.5630
	x_9	0.1085	0.33872	0.4157	0.4590	0.4850	0.5053	0.5206	0.5330	0.5431
	x_{10}	-0	-0	-0	-0	-0	-0	-0	-0.0035	-0.0117
	x_{11}	0.6578	0.64192	0.6301	0.6178	0.6103	0.6199	0.6292	0.6333	0.6367
	x_{12}	0.0044	0.21351	0.2749	0.3005	0.3157	0.3489	0.3778	0.3960	0.4088
	x_{13}	0	0	0	0.0065	0.0105	0.0241	0.0367	0.0442	0.0501
	x_{14}	0.4722	0.55929	0.5774	0.5804	0.5822	0.5992	0.6163	0.6279	0.6372
	x_{15}	0.0770	0	0	0	-0	-0.0179	-0.0532	-0.0810	-0.1018
df		9	9	10	11	11	14	13	14	14
PE		6.5794	5.9418	5.8258	5.7849	5.7667	5.7429	5.7241	5.7128	5.7053

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.47: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0097	0.0200	0.0343	0.0642	0.0457	0.0907	0.1067	0.1149	0.1348
λ_1		0.0877	0.0799	0.0799	0.0963	0.0457	0.0605	0.0457	0.0287	0.0150
Predictors	x_1	2.2033	2.2098	2.1870	2.0941	2.2268	2.0882	2.0702	2.0724	2.0385
	x_2	0.3457	0.39383	0.4217	0.4178	0.4985	0.4770	0.4942	0.5199	0.5435
	x_3	0	0.0794	0.1316	0.1427	0.2561	0.2456	0.2783	0.3122	0.3350
	x_4	0.0010	0.1084	0.2011	0.2942	0.3151	0.4055	0.4482	0.4795	0.5189
	x_5	-1.6123	-1.4855	-1.3561	-1.1766	-1.2718	-1.0809	-1.0389	-1.0228	-0.9760
	x_6	-0.1802	-0.1972	-0.1978	-0.1664	-0.2596	-0.2239	-0.2444	-0.2695	-0.2840
	x_7	1.4228	1.3963	1.3624	1.2969	1.3492	1.2684	1.2518	1.2474	1.2232
	x_8	-0.4080	-0.4069	-0.4099	-0.4212	-0.4134	-0.4367	-0.4466	-0.4546	-0.4682
	x_9	0.2504	0.2581	0.2448	0.1837	0.3105	0.2388	0.2562	0.2844	0.2966
	x_{10}	-0.4359	-0.4378	-0.4289	-0.3964	-0.4625	-0.4256	-0.4351	-0.4480	-0.4505
	x_{11}	0.7950	0.7692	0.7425	0.7025	0.7203	0.6762	0.6646	0.6623	0.6554
	x_{12}	-0	-0	-0	-0	-0	-0	-0	-0.0107	-0.0371
	x_{13}	0.8313	0.8043	0.7617	0.6772	0.7797	0.6799	0.6751	0.6849	0.6790
	x_{14}	-0.0383	-0.0711	-0.0784	-0.0396	-0.1616	-0.1050	-0.1238	-0.1484	-0.1548
	x_{15}	0.2390	0.2256	0.2146	0.2061	0.2114	0.2116	0.2164	0.2238	0.2374
df		13	14	14	14	14	14	14	15	15
PE		4.9469	4.9455	4.9917	5.1806	4.9109	5.1281	5.1332	5.0996	5.1221

Data no.47: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03445204.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0345	0.0345	0.0345	0.0345	0.0345	0.0345	0.0345	0.0345	0.0345
λ_1		0.3101	0.1378	0.0804	0.0517	0.0345	0.0230	0.0148	0.0086	0.0038
Predictors	x_1	1.9846	2.1011	2.1859	2.2430	2.2759	2.2933	2.3057	2.3149	2.3221
	x_2	0.2571	0.3342	0.4212	0.4804	0.5210	0.5727	0.6095	0.6371	0.6586
	x_3	0	0.0125	0.1309	0.2218	0.2771	0.3211	0.3524	0.3758	0.3940
	x_4	0.2765	0.1438	0.2015	0.2588	0.2948	0.3295	0.3542	0.3726	0.3869
	x_5	-0.8239	-1.3045	-1.3545	-1.3418	-1.3316	-1.3087	-1.2925	-1.2805	-1.2711
	x_6	-0	-0.0908	-0.1970	-0.2505	-0.2823	-0.3023	-0.3166	-0.3273	-0.3356
	x_7	1.1149	1.3211	1.3618	1.3687	1.3709	1.3635	1.3583	1.3543	1.3513
	x_8	-0.2698	-0.3927	-0.4099	-0.4073	-0.4065	-0.4086	-0.4101	-0.4112	-0.4121
	x_9	0	0.1015	0.2436	0.3085	0.3478	0.3756	0.3954	0.4103	0.4219
	x_{10}	-0.1882	-0.3561	-0.4282	-0.4622	-0.4812	-0.4872	-0.4914	-0.4946	-0.4971
	x_{11}	0.7720	0.7618	0.7422	0.7354	0.7321	0.7340	0.7354	0.7365	0.7373
	x_{12}	0	0	-0	-0	-0.0084	-0.0506	-0.0808	-0.1034	-0.1210
	x_{13}	0.4490	0.6795	0.7606	0.7998	0.8245	0.8453	0.8601	0.8713	0.8799
	x_{14}	-0	-0	-0.0774	-0.1510	-0.1942	-0.2199	-0.2382	-0.2520	-0.2626
	x_{15}	0.1719	0.2148	0.2145	0.2138	0.2160	0.2273	0.2355	0.2416	0.2464
df		10	13	14	14	15	15	15	15	15
PE		6.0640	5.2446	4.9941	4.8936	4.8462	4.8149	4.7968	4.7855	4.7781

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.48: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0424	0.0599	0.1236	0.1101	0.1504	0.2256	0.2009	0.2605	0.2784
λ_1		0.3814	0.2395	0.2885	0.1651	0.1504	0.1504	0.0861	0.0651	0.0309
Predictors	x_1	1.7847	1.8835	1.7280	1.8739	1.8159	1.6955	1.7911	1.7173	1.7164
	x_2	0.1719	0.3188	0.2534	0.4041	0.4041	0.3725	0.4571	0.4481	0.4655
	x_3	0.5240	0.7190	0.6254	0.8086	0.7943	0.7424	0.8350	0.8113	0.8217
	x_4	0	0.0232	0	0.1056	0.0956	0.0549	0.1375	0.1228	0.1347
	x_5	0	-0	-0	-0	-0	-0	-0	-0.0149	-0.0682
	x_6	0	0	0	0.0794	0.0931	0.0841	0.1611	0.1724	0.2065
	x_7	0.6945	0.8383	0.7238	0.8873	0.8601	0.7882	0.8811	0.8482	0.8705
	x_8	0	0.1270	0.0815	0.1645	0.1677	0.1601	0.1952	0.1985	0.2122
	x_9	0	0	0	0	0.0165	0.0261	0.0921	0.1134	0.1421
	x_{10}	-0	-0	-0	-0	-0.0029	-0.0023	-0.0808	-0.1003	-0.1398
	x_{11}	0.2718	0.2789	0.2568	0.2742	0.2660	0.2527	0.2685	0.2653	0.2789
	x_{12}	0	0	0	0	0	0	-0	-0	-0.0021
	x_{13}	0.1869	0.2031	0.2040	0.2090	0.2153	0.2204	0.2266	0.2341	0.2486
	x_{14}	0.2386	0.2126	0.2345	0.2014	0.2106	0.2277	0.2079	0.2202	0.2268
	x_{15}	0	0	0	0	0	0	0	0	0
df		7	9	8	10	12	12	12	13	14
PE		6.2767	5.7925	6.1122	5.6135	5.6576	5.8440	5.5362	5.6012	5.5196

Data no.48: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.04271525.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0427	0.0427	0.0427	0.0427	0.0427	0.0427	0.0427	0.0427	0.0427
λ_1		0.3844	0.1709	0.0997	0.0641	0.0427	0.0285	0.0183	0.0107	0.0047
Predictors	x_1	1.7810	2.0064	2.0903	2.1437	2.1867	2.2154	2.2358	2.2512	2.2699
	x_2	0.1690	0.4353	0.5465	0.6143	0.6561	0.6840	0.7039	0.7189	0.7414
	x_3	0.5208	0.8714	1.0064	1.0615	1.0873	1.1045	1.1168	1.1261	1.1484
	x_4	0	0.1594	0.2772	0.3637	0.4322	0.4778	0.5104	0.5348	0.5721
	x_5	0	-0	-0	-0	-0	-0	-0	0	0.0310
	x_6	0	0.0846	0.1788	0.2259	0.2529	0.2708	0.2837	0.2933	0.2983
	x_7	0.6909	0.9687	1.0766	1.1033	1.1126	1.1188	1.1233	1.1266	1.1262
	x_8	0	0.1682	0.2086	0.2281	0.2405	0.2488	0.2547	0.2591	0.2645
	x_9	0	0	0.0507	0.0999	0.1298	0.1497	0.1640	0.1746	0.1846
	x_{10}	-0	-0	-0.0783	-0.1714	-0.2377	-0.2819	-0.3135	-0.3373	-0.3569
	x_{11}	0.2711	0.2976	0.3165	0.3764	0.4250	0.4573	0.4804	0.4978	0.5084
	x_{12}	0	0	-0	-0.0806	-0.1439	-0.1861	-0.2163	-0.2389	-0.2613
	x_{13}	0.1866	0.1984	0.2013	0.2241	0.2468	0.2619	0.2727	0.2808	0.2819
	x_{14}	0.2391	0.1679	0.1281	0.1315	0.1420	0.1489	0.1539	0.1576	0.1556
	x_{15}	0	0	-0	-0.0286	-0.0693	-0.0964	-0.1157	-0.1302	-0.1474
df		7	10	12	14	14	14	14	14	15
PE		6.2877	5.4618	5.2200	5.0450	4.9399	4.8841	4.8512	4.8303	4.8125

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.49: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0218	0.0446	0.0697	0.0038	0.0083	0.0086	0.0121	0.0579	0.0986
λ_1		0.1959	0.1785	0.1626	0.0057	0.0083	0.0057	0.0052	0.0145	0.0110
Predictors	x_1	1.8229	1.7810	1.7372	2.1174	2.0779	2.0834	2.0623	1.8642	1.7679
	x_2	0	0	0	0.9896	0.8475	0.8856	0.8301	0.3839	0.3119
	x_3	0	0	0	1.1102	0.9484	0.9898	0.9248	0.4154	0.3322
	x_4	0.1466	0.1483	0.1531	1.7196	1.5032	1.556	1.4667	0.7657	0.6320
	x_5	-0	-0	0	1.5243	1.2858	1.3457	1.2492	0.5013	0.3790
	x_6	-0.5475	-0.5623	-0.5746	-0.6561	-0.6677	-0.6666	-0.6726	-0.6969	-0.6948
	x_7	0.8174	0.8052	0.7903	0.9564	0.9539	0.9547	0.9532	0.9106	0.8742
	x_8	0.2708	0.2987	0.3210	0.3918	0.4002	0.4014	0.4076	0.4367	0.4480
	x_9	0.0777	0.1173	0.1520	0.2486	0.2441	0.2488	0.2507	0.2673	0.2953
	x_{10}	-0	-0	-0.0105	-0.1225	-0.1142	-0.1188	-0.1183	-0.1221	-0.1465
	x_{11}	0.8154	0.7785	0.7437	0.8685	0.8886	0.8813	0.8856	0.8567	0.7890
	x_{12}	-0	-0	-0	-0.4906	-0.4365	-0.4533	-0.4338	-0.2564	-0.2207
	x_{13}	0.0130	0.0589	0.0990	0.1647	0.1933	0.1913	0.2067	0.2827	0.2967
	x_{14}	-0	-0	-0	-0.3109	-0.2852	-0.2927	-0.2824	-0.1728	-0.1379
	x_{15}	-0	-0	-0	-0.7052	-0.6304	-0.6470	-0.6133	-0.3117	-0.2233
df		8	8	9	15	15	15	15	15	15
PE		3.7325	3.7335	3.7464	2.6430	2.6731	2.6641	2.6795	2.9482	3.0905

Data no.49: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.02638164.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0264	0.0264	0.0264	0.0264	0.0264	0.0264	0.0264	0.0264	0.0264
λ_1		0.2374	0.1055	0.0616	0.0396	0.0264	0.0176	0.0113	0.0066	0.0029
Predictors	x_1	1.7913	1.8600	1.8832	1.9192	1.9461	1.9641	1.9769	1.9865	1.9940
	x_2	0	0	0.0259	0.2413	0.3990	0.5041	0.5791	0.6354	0.6792
	x_3	0	0	0.0404	0.2737	0.4444	0.5581	0.6393	0.7002	0.7476
	x_4	0.1215	0.2359	0.3260	0.6151	0.8327	0.9777	1.0813	1.1590	1.2194
	x_5	0	-0	0	0.3040	0.5493	0.7128	0.8296	0.9171	0.9852
	x_6	-0.514	-0.6376	-0.6887	-0.6919	-0.6906	-0.6897	-0.6890	-0.6885	-0.6882
	x_7	0.7818	0.8753	0.9111	0.9262	0.9330	0.9375	0.9407	0.9432	0.9451
	x_8	0.2443	0.3484	0.3905	0.4054	0.4128	0.4178	0.4214	0.4241	0.4262
	x_9	0.0410	0.1555	0.1810	0.2076	0.2269	0.2397	0.2488	0.2557	0.2610
	x_{10}	-0	-0.0076	-0.0374	-0.0704	-0.0889	-0.1013	-0.1101	-0.1167	-0.1219
	x_{11}	0.8080	0.8832	0.9619	0.9440	0.9226	0.9084	0.8982	0.8906	0.8846
	x_{12}	-0	-0.0378	-0.0889	-0.1852	-0.2556	-0.3025	-0.3361	-0.3612	-0.3808
	x_{13}	0.0096	0.1432	0.2491	0.2555	0.2530	0.2514	0.2502	0.24933	0.2486
	x_{14}	-0	-0.0383	-0.1006	-0.1539	-0.1890	-0.2124	-0.2291	-0.2416	-0.2514
	x_{15}	-0	-0.1067	-0.1995	-0.3031	-0.3783	-0.4284	-0.4641	-0.4910	-0.5118
df		8	12	14	15	15	15	15	15	15
PE		3.8329	3.3945	3.1814	2.9927	2.8840	2.8227	2.7844	2.7587	2.7404

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.50: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0221	0.0545	0.0707	0.1100	0.1650	0.2475	0.2654	0.3777	0.4431
λ_1		0.1988	0.2181	0.1650	0.1650	0.1650	0.1650	0.1137	0.0944	0.0492
Predictors	x_1	1.7711	1.6754	1.6768	1.5961	1.5004	1.3836	1.3907	1.2766	1.2353
	x_2	-0	-0	-0	-0	-0	-0	-0	-0	-0.0365
	x_3	0.1615	0.1345	0.1925	0.1859	0.1786	0.1704	0.2101	0.2088	0.2260
	x_4	0.2947	0.2913	0.3293	0.3370	0.3455	0.3544	0.3795	0.3878	0.3987
	x_5	-0	-0	-0	0	0	0	0	0	0.0070
	x_6	-0.2555	-0.2694	-0.2867	-0.3034	-0.3218	-0.3408	-0.3568	-0.3758	-0.3926
	x_7	1.0344	0.9770	1.0128	0.9725	0.9233	0.8611	0.8888	0.8312	0.8245
	x_8	-0.5448	-0.5151	-0.5541	-0.5366	-0.5131	-0.4807	-0.5066	-0.4726	-0.4727
	x_9	0	0	0	0	0	0	0.0141	0.0478	0.0723
	x_{10}	0.1917	0.1663	0.2465	0.2425	0.2335	0.2173	0.2699	0.2551	0.2859
	x_{11}	0.7016	0.6533	0.6321	0.5902	0.5445	0.4941	0.4865	0.4423	0.4260
	x_{12}	0.0092	0.0262	0.0449	0.0636	0.0849	0.1088	0.1187	0.1405	0.1558
	x_{13}	0.2979	0.3036	0.2916	0.2926	0.2924	0.2905	0.2861	0.2831	0.2823
	x_{14}	-0	-0	-0	-0	-0	-0	-0	-0	-0
	x_{15}	0.1489	0.1744	0.1878	0.2082	0.2283	0.2471	0.2535	0.2681	0.2785
df		11	11	11	11	11	11	12	12	14
PE		5.8439	5.9598	5.8503	5.9331	6.0565	6.2470	6.1482	6.3424	6.3532

Data no.50: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03124969.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0312	0.0312	0.0312	0.0312	0.0312	0.0312	0.0312	0.0312	0.0312
λ_1		0.2812	0.1250	0.0729	0.0469	0.0312	0.0208	0.0134	0.0078	0.0035
Predictors	x_1	1.6859	1.8015	1.8227	1.8159	1.8145	1.8159	1.8169	1.8177	1.8183
	x_2	-0	-0	-0.0349	-0.0653	-0.0849	-0.0991	-0.1093	-0.1169	-0.1228
	x_3	0.0603	0.2487	0.2987	0.3185	0.3319	0.3422	0.3496	0.3551	0.3594
	x_4	0.2419	0.3466	0.3647	0.3654	0.3681	0.3721	0.3749	0.3770	0.3786
	x_5	-0	-0	-0	-0	-0	-0	-0	-0	-0
	x_6	-0.2473	-0.2726	-0.2866	-0.3016	-0.3098	-0.3144	-0.3177	-0.3202	-0.3221
	x_7	0.9364	1.1008	1.1448	1.1491	1.1521	1.1547	1.1565	1.1578	1.1589
	x_8	-0.4674	-0.6079	-0.6455	-0.6510	-0.6586	-0.6676	-0.6740	-0.6788	-0.6826
	x_9	0	0	-0	-0	-0.0161	-0.0413	-0.0592	-0.0727	-0.0832
	x_{10}	0.0630	0.3101	0.3976	0.4379	0.4667	0.4899	0.5065	0.5190	0.5287
	x_{11}	0.6883	0.6852	0.6942	0.7125	0.7232	0.7301	0.7350	0.7387	0.7416
	x_{12}	0	0.0306	0.0550	0.0832	0.1016	0.1151	0.1247	0.1319	0.1374
	x_{13}	0.3205	0.2779	0.2693	0.2711	0.2716	0.2716	0.2715	0.2715	0.2714
	x_{14}	0	-0	-0.0299	-0.0937	-0.1315	-0.1562	-0.1738	-0.1870	-0.1973
	x_{15}	0.1523	0.1613	0.1815	0.2109	0.2269	0.2362	0.2429	0.2479	0.2517
df		10	11	13	13	14	14	14	14	14
PE		6.1206	5.6956	5.5982	5.5445	5.5186	5.5035	5.4951	5.4900	5.4868

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.51: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0059	0.0121	0.0188	0.0322	0.0365	0.0260	0.0232	0.0397	0.0466
λ_1		0.0530	0.0483	0.0440	0.0483	0.0365	0.0173	0.0099	0.0099	0.0052
Predictors	x_1	1.8916	1.8875	1.8817	1.8506	1.8599	1.8759	1.8772	1.8609	1.8574
	x_2	-0.9533	-0.9370	-0.9175	-0.8415	-0.8577	-0.9873	-1.0363	-0.9450	-0.9247
	x_3	0	0	0	0.0421	0.0497	0	-0.0260	0.0142	0.0377
	x_4	-0	-0	-0	-0	-0	-0.0935	-0.1414	-0.0986	-0.0942
	x_5	-1.5916	-1.5746	-1.5541	-1.4579	-1.4736	-1.6611	-1.7375	-1.6119	-1.5800
	x_6	-0.4522	-0.4515	-0.4493	-0.4247	-0.4381	-0.4905	-0.5114	-0.4842	-0.4829
	x_7	1.3040	1.2997	1.2948	1.2784	1.2777	1.2772	1.2729	1.2661	1.2605
	x_8	-0.2777	-0.2807	-0.2832	-0.2700	-0.2782	-0.2903	-0.2947	-0.2924	-0.2939
	x_9	0.2559	0.2505	0.2429	0.1942	0.2124	0.2933	0.3243	0.2739	0.2664
	x_{10}	0.0793	0.0891	0.0984	0.1006	0.1177	0.1435	0.1568	0.1602	0.1713
	x_{11}	1.5016	1.4830	1.4619	1.3901	1.3910	1.4783	1.5123	1.4348	1.4131
	x_{12}	-0.9676	-0.9528	-0.9335	-0.8466	-0.8765	-0.9960	-1.0353	-0.9498	-0.9346
	x_{13}	0.9160	0.9020	0.8857	0.8287	0.8388	0.9322	0.9684	0.8992	0.8852
	x_{14}	0.1975	0.2065	0.2144	0.2067	0.2255	0.3003	0.3389	0.3142	0.3224
	x_{15}	-0	-0	-0	-0	-0	-0.0348	-0.0600	-0.0443	-0.0549
df		12	12	12	13	13	14	15	15	15
PE		3.7272	3.7413	3.7615	3.8708	3.8412	3.6819	3.6368	3.7263	3.7486

Data no.51: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.0434764.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0435	0.0435	0.0435	0.0435	0.0435	0.0435	0.0435	0.0435	0.0435
λ_1		0.3913	0.1739	0.1014	0.0652	0.0435	0.0290	0.0186	0.0109	0.0048
Predictors	x_1	1.5003	1.6693	1.7549	1.8065	1.8375	1.8581	1.8572	1.8586	1.8596
	x_2	-0	-0.4410	-0.6406	-0.7452	-0.8080	-0.8499	-0.8915	-0.9204	-0.9428
	x_3	0.1282	0.1132	0.0975	0.0821	0.0728	0.0666	0.0456	0.0331	0.0235
	x_4	0	0	0	-0	-0	-0	-0.0479	-0.0807	-0.1062
	x_5	-0.2695	-0.9688	-1.2159	-1.3380	-1.4113	-1.4602	-1.5279	-1.5731	-1.6082
	x_6	-0	-0.2129	-0.3284	-0.3848	-0.4186	-0.4412	-0.4610	-0.4764	-0.4884
	x_7	1.0327	1.2024	1.2447	1.2588	1.2673	1.2730	1.2685	1.2648	1.2619
	x_8	-0	-0.1175	-0.2100	-0.2453	-0.2665	-0.2806	-0.2857	-0.2909	-0.2949
	x_9	-0	0	0.0377	0.1254	0.1779	0.2130	0.2400	0.2604	0.2763
	x_{10}	0	0	0.0417	0.0860	0.1125	0.1302	0.1459	0.1599	0.1709
	x_{11}	0.9110	1.1524	1.2501	1.3116	1.3486	1.3732	1.3960	1.4144	1.4287
	x_{12}	-0	-0.3393	-0.5788	-0.7270	-0.8159	-0.8752	-0.9070	-0.9306	-0.9491
	x_{13}	0.1660	0.5363	0.6872	0.7593	0.8025	0.8314	0.8600	0.8815	0.8982
	x_{14}	0	0.0378	0.1235	0.1800	0.2139	0.2365	0.2738	0.3046	0.3284
	x_{15}	0	0	0	0	-0	-0	-0.0177	-0.0401	-0.0573
df		6	11	13	13	13	13	15	15	15
PE		6.9279	4.9595	4.3278	4.0570	3.9261	3.8520	3.7945	3.7551	3.7274

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.52: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0287	0.0535	0.1007	0.1185	0.1620	0.1838	0.3138	0.1106	0.0616
λ_1		0.2579	0.2141	0.2350	0.1778	0.1620	0.1225	0.1345	0.0277	0.0068
Predictors	x_1	1.5170	1.5095	1.4077	1.4231	1.3647	1.3498	1.1893	1.5160	1.6004
	x_2	0	0	0	0	0	0	0.0378	0.0094	0
	x_3	0	0	0	0	0	0	0	-0.0059	-0.1048
	x_4	1.0456	1.0576	0.9880	1.0190	0.9886	0.9668	0.8655	1.0399	1.0337
	x_5	-0	-0	-0	-0	-0	-0.0754	-0.0386	-0.2577	-0.3938
	x_6	0	0	0	0	0	0	0	-0	-0.0838
	x_7	0.8292	0.8567	0.8149	0.8524	0.8403	0.8488	0.7736	0.9161	0.9685
	x_8	-0	-0	-0	-0	-0	-0.0241	-0.0057	-0.1437	-0.1648
	x_9	-0	-0	-0	-0	-0	-0	-0	-0.0633	-0.1162
	x_{10}	-0	-0	-0	-0	-0	-0.0052	-0	-0.1112	-0.1387
	x_{11}	0.8249	0.7903	0.7113	0.6963	0.6439	0.6431	0.5291	0.8790	1.1097
	x_{12}	-0	-0	-0	-0	-0	-0.0097	-0	-0.2103	-0.3025
	x_{13}	0.1230	0.1625	0.1955	0.2068	0.2171	0.2398	0.2357	0.3396	0.4270
	x_{14}	0	0	0	0	0	0	0.0123	-0.0615	-0.2061
	x_{15}	0	0	0.0263	0.0447	0.0842	0.1129	0.1640	0.1087	0.1049
df		5	5	6	6	6	10	10	14	14
PE		4.9244	4.9071	5.0946	5.0336	5.1396	5.0737	5.5135	4.3684	4.0417

Data no.52: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03429558.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0343	0.0343	0.0343	0.0343	0.0343	0.0343	0.0343	0.0343	0.0343
λ_1		0.3087	0.1372	0.0800	0.0514	0.0343	0.0229	0.0147	0.0086	0.0038
Predictors	x_1	1.4613	1.6063	1.6324	1.6607	1.6708	1.6627	1.6558	1.6476	1.6401
	x_2	0	0	0	0	0	-0	-0	-0.0155	-0.0335
	x_3	0	0	-0	-0	-0	-0.0537	-0.0962	-0.1350	-0.1680
	x_4	0.9909	1.1448	1.1205	1.1085	1.1135	1.0776	1.0486	1.0172	0.9888
	x_5	-0	-0.0071	-0.1477	-0.2286	-0.2718	-0.3445	-0.4002	-0.4530	-0.4988
	x_6	0	0	-0	-0	-0.0315	-0.0667	-0.0922	-0.1101	-0.1235
	x_7	0.7824	0.9086	0.9205	0.9314	0.9452	0.9611	0.9728	0.9826	0.9906
	x_8	-0	-0	-0.0492	-0.0955	-0.1194	-0.1350	-0.1462	-0.1524	-0.1564
	x_9	-0	-0	-0	-0.0314	-0.0567	-0.0855	-0.1070	-0.1233	-0.1360
	x_{10}	-0	-0.0760	-0.1400	-0.1688	-0.1819	-0.1740	-0.1669	-0.1623	-0.1589
	x_{11}	0.8028	0.9051	1.0332	1.1194	1.1704	1.2091	1.2369	1.2615	1.2820
	x_{12}	-0	-0.1152	-0.2442	-0.2952	-0.3296	-0.3422	-0.3503	-0.3545	-0.3569
	x_{13}	0.1274	0.1766	0.2730	0.3389	0.3844	0.4194	0.4444	0.4661	0.4842
	x_{14}	0	-0	-0.0385	-0.1193	-0.1747	-0.2191	-0.2513	-0.2720	-0.2865
	x_{15}	0	0	0	0	0	0.0293	0.0531	0.0713	0.0854
df		5	8	10	11	12	14	14	15	15
PE		5.0509	4.5653	4.2695	4.1136	4.0323	3.9797	3.9474	3.9249	3.9087

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.53: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0355	0.0799	0.1504	0.2339	0.2913	0.4369	0.5141	0.5535	0.5916
λ_1		0.3197	0.3197	0.3508	0.3508	0.2913	0.2913	0.2203	0.1384	0.0657
Predictors	x_1	1.2056	1.1556	1.0568	0.9884	0.9968	0.8999	0.9044	0.9368	0.9654
	x_2	0	0	0	0	0	0	0	0	0
	x_3	0.0650	0.0577	0.0250	0.0185	0.0544	0.0470	0.0870	0.1265	0.1521
	x_4	0	0	0	-0	-0	-0	-0	-0	-0.0260
	x_5	-0	-0	-0	-0	-0	-0	-0	-0.0548	-0.1721
	x_6	-0	-0	-0	-0	-0	-0	-0	-0	-0
	x_7	1.2377	1.1884	1.0677	0.9946	1.0275	0.9147	0.9351	0.9901	1.0240
	x_8	-0	-0	-0	-0	-0	-0	-0	-0	-0.0657
	x_9	-0	-0	-0	-0	-0	-0.0197	-0.0595	-0.1058	-0.1635
	x_{10}	0	0.0059	0	0.0132	0.0537	0.0660	0.1070	0.1480	0.1815
	x_{11}	0.2130	0.1906	0.1537	0.1330	0.1388	0.1131	0.1159	0.1297	0.1484
	x_{12}	0.0269	0.0311	0.0263	0.0310	0.0496	0.0453	0.0537	0.0707	0.0939
	x_{13}	0.2898	0.2762	0.2613	0.2420	0.2293	0.2064	0.1973	0.1981	0.2107
	x_{14}	0.0049	0.0191	0.0257	0.0342	0.0475	0.0488	0.0574	0.0711	0.0868
	x_{15}	0	0	0	0	0	0.0129	0.0291	0.0461	0.0674
df		7	8	7	8	8	10	10	11	13
PE		7.1449	7.2395	7.5131	7.6899	7.5752	7.8518	7.7335	7.5031	7.2611

Data no.53: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.08436095.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0844	0.0844	0.0844	0.0844	0.0844	0.0844	0.0844	0.0844	0.0844
λ_1		0.7592	0.3374	0.1968	0.1265	0.0844	0.0562	0.0362	0.0211	0.0094
Predictors	x_1	0.7221	1.1327	1.2739	1.3467	1.3866	1.3747	1.3660	1.3593	1.3477
	x_2	0	0	0	0	0	0	0	-0	-0.0184
	x_3	0	0.0422	0.1368	0.1548	0.1568	0.1050	0.0676	0.0395	0.0072
	x_4	0	0	0	-0	-0.0301	-0.1652	-0.2623	-0.3352	-0.4003
	x_5	0	-0	-0.0885	-0.3597	-0.5415	-0.7454	-0.8916	-1.0015	-1.0959
	x_6	0	-0	-0	-0	-0.0002	-0.0304	-0.0521	-0.0684	-0.0810
	x_7	0.5791	1.1539	1.3984	1.5111	1.5654	1.5884	1.6048	1.6171	1.6269
	x_8	0	-0	-0	-0.0553	-0.1237	-0.1794	-0.2192	-0.2491	-0.2718
	x_9	-0	-0	-0.0009	-0.0970	-0.1651	-0.2086	-0.2396	-0.2629	-0.2803
	x_{10}	0	0	0.0847	0.1132	0.1291	0.1456	0.1575	0.1664	0.1757
	x_{11}	0	0.1817	0.2448	0.2966	0.3301	0.3508	0.3654	0.3763	0.3886
	x_{12}	0	0.0257	0.0924	0.1544	0.1959	0.2377	0.2676	0.2900	0.3101
	x_{13}	0.1802	0.2774	0.2605	0.2802	0.3075	0.3632	0.4031	0.4332	0.4570
	x_{14}	0	0.0159	0.0549	0.0930	0.1129	0.1285	0.1395	0.1479	0.1560
	x_{15}	0	0	0	0	0	0.0243	0.0424	0.0559	0.0677
df		3	7	10	11	13	14	14	14	15
PE		9.0891	7.3160	6.7882	6.4193	6.2266	6.0692	5.9729	5.9095	5.8617

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.54: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0089	0.0201	0.0021	0.0091	0.0137	0.0056	0.0614	0.0500	0.0132
λ_1		0.0803	0.0803	0.0049	0.0137	0.0137	0.0037	0.0263	0.0125	0.0015
Predictors	x_1	1.2908	1.3056	0.9771	1.1452	1.1952	1.0302	1.354	1.3432	1.1215
	x_2	-0	0	-0.8537	-0.5204	-0.4218	-0.7586	0.0122	-0.0360	-0.5897
	x_3	-0.4460	-0.3790	-1.5417	-1.1550	-1.0356	-1.4248	-0.4131	-0.5460	-1.2174
	x_4	-0.5744	-0.5093	-1.7182	-1.2748	-1.1387	-1.5818	-0.4923	-0.6143	-1.3423
	x_5	-1.9141	-1.8013	-3.6043	-2.9531	-2.7518	-3.4091	-1.7293	-1.9345	-3.0604
	x_6	-0.1053	-0.0731	-0.5623	-0.4177	-0.3740	-0.5226	-0.1154	-0.1805	-0.4500
	x_7	1.1244	1.1156	1.1369	1.1321	1.1301	1.1352	1.1055	1.1171	1.1321
	x_8	-0.6418	-0.6360	-0.9909	-0.8531	-0.8170	-0.9596	-0.6542	-0.6887	-0.9025
	x_9	0.5263	0.5272	0.6247	0.6398	0.6466	0.6360	0.6059	0.6369	0.6537
	x_{10}	0.2246	0.1941	0.4561	0.3897	0.3704	0.4398	0.2303	0.2819	0.4092
	x_{11}	1.5471	1.4554	2.2597	2.0054	1.9210	2.1840	1.3914	1.5196	2.0417
	x_{12}	-0	-0	0.3921	0.1758	0.1212	0.3384	-0.0163	-0.0013	0.2459
	x_{13}	0.5133	0.5047	0.9261	0.7969	0.7613	0.8949	0.5640	0.6124	0.8378
	x_{14}	-0.1908	-0.1704	-0.1928	-0.2173	-0.2267	-0.2077	-0.2016	-0.2372	-0.2300
	x_{15}	-0	-0	-0.0697	-0.0137	-0.0028	-0.0634	0	0	-0.0499
df		12	12	15	15	15	15	14	14	15
PE		4.5431	4.6529	3.8551	3.9779	4.0403	3.8771	4.6493	4.4592	3.9439

Data no.54: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.0567022.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0567	0.0567	0.0567	0.0567	0.0567	0.0567	0.0567	0.0567	0.0567
λ_1		0.5103	0.2268	0.1323	0.0851	0.0567	0.0378	0.0243	0.0142	0.0063
Predictors	x_1	1.0677	1.2279	1.3268	1.3350	1.3400	1.3443	1.3500	1.3553	1.3556
	x_2	0.3305	0.2206	0.1554	0.0966	0.0475	0.0094	0	-0	-0.0284
	x_3	0	-0	-0	-0.1674	-0.2954	-0.3891	-0.4468	-0.4868	-0.5348
	x_4	0	-0	-0.1390	-0.2972	-0.4049	-0.4809	-0.5248	-0.5541	-0.5956
	x_5	-0.0182	-0.7955	-1.1712	-1.4273	-1.5969	-1.7161	-1.7862	-1.8330	-1.8962
	x_6	0	0	-0	-0	-0.0464	-0.0977	-0.1311	-0.1550	-0.1780
	x_7	0.8626	1.0323	1.0505	1.0865	1.1001	1.1049	1.1092	1.1128	1.1146
	x_8	-0	-0.3508	-0.5261	-0.5864	-0.6235	-0.6494	-0.6631	-0.6714	-0.6838
	x_9	0	0.1734	0.4014	0.4954	0.5506	0.5871	0.6121	0.6302	0.6467
	x_{10}	0	0	0.0538	0.1266	0.1778	0.2151	0.2432	0.2649	0.2795
	x_{11}	0.6231	0.9111	1.0634	1.2033	1.3043	1.3780	1.4276	1.4634	1.4936
	x_{12}	-0	-0	-0	-0.0020	-0.0048	-0.0065	-0.0110	-0.0156	-0.0161
	x_{13}	0	0.2096	0.3277	0.4387	0.5051	0.5486	0.5749	0.5928	0.6131
	x_{14}	0	-0	-0.0583	-0.1285	-0.1665	-0.1903	-0.2111	-0.2279	-0.2355
	x_{15}	0	0	-0	-0	-0	0	0	0	0.0086
df		5	8	11	13	14	14	13	13	15
PE		8.0987	6.2047	5.5136	5.0751	4.8322	4.6827	4.5943	4.5372	4.4845

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.55: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0364	0.0620	0.1541	0.1652	0.2058	0.2335	0.3310	0.2039	0.2179
λ_1		0.3277	0.2479	0.3596	0.2479	0.2058	0.1557	0.1418	0.0510	0.0242
Predictors	x_1	1.8717	1.8780	1.6352	1.6922	1.6595	1.6530	1.5390	1.8188	1.8219
	x_2	0	0	0	0	0	0	0	0.0927	0.1109
	x_3	0	0	0	0	0	0	0.0167	0.0516	0.0605
	x_4	0.0030	0.0893	0.0399	0.1270	0.1724	0.2217	0.2333	0.3405	0.3591
	x_5	-0	-0	-0	-0	-0	-0	-0	-0.0832	-0.1249
	x_6	-0.0919	-0.1583	-0.0313	-0.1245	-0.1541	-0.1944	-0.1800	-0.2688	-0.2779
	x_7	1.2178	1.2779	1.0967	1.1884	1.1963	1.2237	1.1503	1.3568	1.3691
	x_8	0	0	0	0	0.0279	0.0727	0.0692	0.1875	0.2086
	x_9	0	0	0	0	0	0	0.0050	0.1187	0.1421
	x_{10}	-0	-0	-0	-0	-0	-0	-0	-0.0543	-0.0849
	x_{11}	0.6858	0.6577	0.5492	0.5516	0.5287	0.5175	0.4749	0.5547	0.5575
	x_{12}	0	0.0276	0.1474	0.1534	0.1815	0.1983	0.2311	0.1879	0.1996
	x_{13}	0.3387	0.3519	0.3177	0.3330	0.3322	0.3315	0.3275	0.2955	0.2905
	x_{14}	0.3974	0.4018	0.3642	0.3779	0.3770	0.3792	0.3686	0.4041	0.4126
	x_{15}	0	0	0	0	0	0	0.0085	-0	-0
df		7	8	8	8	9	9	12	14	14
PE		6.4247	6.2756	6.9202	6.5930	6.5786	6.4992	6.7559	6.0208	5.9753

Data no.55: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03128163.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0313	0.0313	0.0313	0.0313	0.0313	0.0313	0.0313	0.0313	0.0313
λ_1		0.2815	0.1251	0.0730	0.0469	0.0313	0.0209	0.0134	0.0078	0.0035
Predictors	x_1	1.9149	2.0977	2.1721	2.2319	2.2681	2.2922	2.3094	2.3223	2.3324
	x_2	0	0.1020	0.1590	0.2047	0.2324	0.2509	0.2641	0.2740	0.2817
	x_3	0	0	0	0.0368	0.0597	0.0748	0.0856	0.0938	0.1002
	x_4	0.0414	0.2289	0.2583	0.3338	0.3798	0.4102	0.4320	0.4484	0.4612
	x_5	-0	-0.0847	-0.2438	-0.2778	-0.2975	-0.3108	-0.3203	-0.3273	-0.3327
	x_6	-0.1366	-0.2823	-0.2977	-0.2885	-0.2829	-0.2792	-0.2765	-0.2745	-0.2729
	x_7	1.2713	1.4680	1.5320	1.5630	1.5816	1.5940	1.6029	1.6095	1.6147
	x_8	0	0.1588	0.2228	0.2618	0.2852	0.3008	0.3120	0.3203	0.3269
	x_9	0	0.0818	0.1499	0.1851	0.2063	0.2203	0.2304	0.2380	0.2438
	x_{10}	-0	-0.0355	-0.1437	-0.2034	-0.2392	-0.2631	-0.2801	-0.2929	-0.3028
	x_{11}	0.6935	0.7266	0.7910	0.8277	0.8496	0.8642	0.8746	0.8824	0.8885
	x_{12}	0	0	0.0168	0.0449	0.0616	0.0725	0.0804	0.0863	0.0909
	x_{13}	0.3426	0.2967	0.2559	0.2229	0.2028	0.1896	0.1802	0.1730	0.1674
	x_{14}	0.4028	0.4434	0.5066	0.5610	0.5936	0.6153	0.6308	0.6424	0.6515
	x_{15}	0	-0	-0.0255	-0.1176	-0.1730	-0.2098	-0.2361	-0.2559	-0.2713
df		7	12	14	15	15	15	15	15	15
PE		6.2812	5.7546	5.5466	5.4303	5.3769	5.3482	5.3310	5.3200	5.3125

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.56: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0355	0.0664	0.0785	0.1112	0.2420	0.1432	0.1274	0.1991	0.2335
λ_1		0.3199	0.2656	0.1831	0.1668	0.2420	0.0954	0.0546	0.0498	0.0259
Predictors	x_1	1.4307	1.4588	1.5581	1.5410	1.2911	1.5889	1.6617	1.5604	1.5355
	x_2	0.2464	0.2999	0.3765	0.3894	0.2744	0.4314	0.4620	0.4239	0.4149
	x_3	0	0	0.0720	0.1144	0.0542	0.1803	0.2023	0.1953	0.1950
	x_4	0.8489	0.8798	0.9292	0.9322	0.7654	0.9114	0.9054	0.8309	0.7828
	x_5	-0	-0	-0.0503	-0.0344	-0	-0.1641	-0.2873	-0.2593	-0.3063
	x_6	-0	-0	-0	-0	-0	-0.0495	-0.0749	-0.1017	-0.1304
	x_7	1.3322	1.3485	1.4129	1.3959	1.2062	1.4277	1.4800	1.4116	1.3953
	x_8	0	0	0.0656	0.0804	0.0046	0.1544	0.2076	0.2033	0.2295
	x_9	0	0.0512	0.1496	0.1679	0.0896	0.2532	0.2991	0.3034	0.3215
	x_{10}	-0	-0	-0	-0	-0	-0.0599	-0.1372	-0.1119	-0.1311
	x_{11}	0.6435	0.6300	0.6209	0.5944	0.4971	0.5887	0.6177	0.5431	0.5166
	x_{12}	0	0	0	0	0.0523	0	0	0.0326	0.0463
	x_{13}	0	0	0	0	0.0282	0	0	0.0045	0.0188
	x_{14}	0	0	0	0	0	0	0	0.0220	0.0421
	x_{15}	0	0	0	0	0.0387	0.0439	0.0664	0.0951	0.1110
df		5	6	9	9	11	12	12	15	15
PE		6.3828	6.2965	6.0335	6.0605	6.7455	5.8793	5.7099	5.9104	5.9589

Data no.56: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.04722241.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0472	0.0472	0.0472	0.0472	0.0472	0.0472	0.0472	0.0472	0.0472
λ_1		0.4250	0.1889	0.1102	0.0708	0.0472	0.0315	0.0202	0.0118	0.0052
Predictors	x_1	1.3016	1.5852	1.6981	1.7545	1.7868	1.8307	1.8651	1.8907	1.9110
	x_2	0.1569	0.3697	0.4429	0.4814	0.5003	0.5446	0.5823	0.6102	0.6326
	x_3	0	0.0282	0.0970	0.1368	0.1564	0.2240	0.2876	0.3346	0.3725
	x_4	0.7594	0.9211	0.9341	0.9358	0.9285	0.9924	1.0581	1.1064	1.1457
	x_5	-0	-0.1003	-0.2908	-0.3894	-0.4566	-0.4364	-0.3995	-0.3729	-0.3503
	x_6	-0	-0	-0	-0.0213	-0.0416	-0.0472	-0.0566	-0.0637	-0.0692
	x_7	1.2276	1.4371	1.5076	1.5421	1.5636	1.5912	1.6127	1.6287	1.6413
	x_8	0	0.0563	0.1400	0.1909	0.2223	0.2299	0.2394	0.2465	0.2519
	x_9	0	0.1399	0.2197	0.2588	0.2858	0.3211	0.3455	0.3637	0.3781
	x_{10}	-0	-0	-0.0767	-0.1480	-0.1929	-0.2337	-0.2636	-0.2860	-0.3034
	x_{11}	0.6340	0.6603	0.7067	0.7325	0.7468	0.7571	0.7684	0.7770	0.7835
	x_{12}	0	0	0	-0	-0	-0.0067	-0.0372	-0.0599	-0.0778
	x_{13}	0	0	0	-0	-0.0003	-0.0481	-0.0766	-0.0977	-0.1147
	x_{14}	0	0	-0	-0	-0	-0.0281	-0.0526	-0.0708	-0.0854
	x_{15}	0	0	0	0	0.0095	0.0381	0.0591	0.0748	0.0869
df		5	9	10	11	13	15	15	15	15
PE		6.7194	5.9710	5.6814	5.5664	5.5135	5.4420	5.3895	5.3563	5.3335

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.57: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0060	0.0149	0.0281	0.0363	0.0311	0.0293	0.0603	0.0713	0.0918
λ_1		0.0544	0.0597	0.0655	0.0544	0.0311	0.0196	0.0259	0.0178	0.0102
Predictors	x_1	1.8913	1.8746	1.8637	1.8671	1.8847	1.8819	1.8453	1.8298	1.8007
	x_2	0	0.0006	0.0595	0.0879	0.0687	0.0563	0.1531	0.1779	0.2169
	x_3	0.4209	0.4521	0.5244	0.5472	0.5030	0.4716	0.5824	0.5968	0.6251
	x_4	-0.2152	-0.1875	-0.1174	-0.1086	-0.1718	-0.2195	-0.1277	-0.1338	-0.1308
	x_5	-1.2222	-1.1838	-1.0849	-1.0709	-1.1585	-1.2198	-1.0750	-1.0714	-1.0476
	x_6	-0.1593	-0.1434	-0.1302	-0.1441	-0.1845	-0.2075	-0.1822	-0.1913	-0.1948
	x_7	2.0646	2.0339	1.9911	1.9790	2.0164	2.0332	1.9423	1.9224	1.8805
	x_8	0.2255	0.1981	0.1699	0.1888	0.2473	0.2750	0.2289	0.2352	0.2334
	x_9	0.2409	0.2267	0.2073	0.2220	0.2606	0.2771	0.2498	0.2541	0.2544
	x_{10}	0	0	0	0	0	0	0	0	0
	x_{11}	1.1939	1.1593	1.1004	1.0878	1.1329	1.1496	1.0429	1.0193	0.9751
	x_{12}	0	0	0	0	0.0144	0.0390	0.0431	0.0630	0.0842
	x_{13}	0.6634	0.6240	0.5614	0.5682	0.6463	0.6850	0.5766	0.5689	0.5426
	x_{14}	-0.1723	-0.1468	-0.1205	-0.1295	-0.1757	-0.1979	-0.1469	-0.1470	-0.1369
	x_{15}	-0.2839	-0.2729	-0.2724	-0.2805	-0.3014	-0.3100	-0.2938	-0.2943	-0.2873
df		12	13	13	13	14	14	14	14	14
PE		3.5635	3.6136	3.6952	3.6844	3.5796	3.5400	3.6762	3.6937	3.7533

Data no.57: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02849428.

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0285	0.0285	0.0285	0.0285	0.0285	0.0285	0.0285	0.0285	0.0285
λ_1		0.2564	0.1140	0.0665	0.0427	0.0285	0.0190	0.0122	0.0071	0.0032
Predictors	x_1	1.7010	1.8188	1.8625	1.8844	1.8862	1.8822	1.8794	1.8772	1.877
	x_2	0.0329	0.0609	0.0610	0.0611	0.0570	0.0524	0.0491	0.0466	0.0449
	x_3	0.4980	0.5639	0.5268	0.5083	0.4861	0.4663	0.4521	0.4414	0.4329
	x_4	-0	-0.0421	-0.1145	-0.1508	-0.1904	-0.2250	-0.2498	-0.2684	-0.2829
	x_5	-0.8330	-0.9761	-1.0807	-1.1330	-1.1843	-1.2276	-1.2585	-1.2818	-1.2998
	x_6	-0	-0.0552	-0.1287	-0.1654	-0.1906	-0.2089	-0.2220	-0.2318	-0.2404
	x_7	1.8076	1.9424	1.9890	2.0123	2.0265	2.0361	2.0429	2.0481	2.0509
	x_8	0	0.0542	0.1673	0.2239	0.2563	0.2773	0.2922	0.3035	0.3130
	x_9	0	0.1257	0.2055	0.2453	0.2662	0.2786	0.2875	0.2941	0.2999
	x_{10}	0	0	0	0	0	0	0	-0	-0.0066
	x_{11}	0.9102	1.0332	1.0977	1.1299	1.1451	1.1533	1.1592	1.1635	1.1655
	x_{12}	0	0	0	0	0.0188	0.0399	0.0552	0.0665	0.07657
	x_{13}	0.1916	0.4261	0.5575	0.6233	0.6631	0.6898	0.7088	0.7232	0.7353
	x_{14}	-0	-0.0364	-0.1184	-0.1594	-0.1839	-0.2001	-0.2117	-0.2204	-0.2265
	x_{15}	-0.0005	-0.2302	-0.2716	-0.2923	-0.3034	-0.3102	-0.3152	-0.3188	-0.3232
df		8	13	13	13	14	14	14	14	15
PE		4.7936	3.9627	3.7012	3.6079	3.5619	3.5359	3.5201	3.5098	3.5025

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.58: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0364	0.0620	0.1063	0.1372	0.0891	0.2336	0.1726	0.2239	0.2180
λ_1		0.3278	0.2480	0.2480	0.2059	0.0891	0.1557	0.0740	0.0560	0.0242
Predictors	x_1	1.7272	1.7582	1.6870	1.6747	1.8428	1.5658	1.7161	1.6476	1.6784
	x_2	-0	-0	-0	-0	-0.0872	-0	-0.0942	-0.1076	-0.1624
	x_3	0.2562	0.3090	0.2904	0.3078	0.3889	0.2954	0.3558	0.3379	0.3518
	x_4	0	0	0	0	0	0.0282	0	0	-0
	x_5	-0	-0	-0	-0	-0.3465	-0.0601	-0.3277	-0.3452	-0.4464
	x_6	0	0	0	0	0.0918	0.0013	0.0870	0.0972	0.1261
	x_7	0.8848	0.9209	0.8831	0.8881	1.0009	0.8495	0.9398	0.9128	0.9322
	x_8	0	0	0	0	-0	0	0	0	0
	x_9	0	0	0	0	0.1053	0	0.0921	0.0994	0.1478
	x_{10}	0.0897	0.1243	0.1270	0.1454	0.1861	0.1651	0.1964	0.2023	0.2144
	x_{11}	0.5669	0.5380	0.5179	0.5025	0.5280	0.4612	0.5020	0.4812	0.4902
	x_{12}	0.0477	0.0623	0.0807	0.0918	0.1057	0.1094	0.1278	0.1305	0.1382
	x_{13}	0.1535	0.1764	0.1835	0.1950	0.2421	0.2102	0.2493	0.2537	0.2689
	x_{14}	0	0	0	0	0	0.0445	0.0106	0.0478	0.0440
	x_{15}	0.2568	0.2599	0.2545	0.2554	0.3055	0.2489	0.2934	0.2864	0.3011
df		8	8	8	8	12	12	13	13	13
PE		5.2140	5.0903	5.1848	5.1649	4.6622	5.2649	4.7969	4.8725	4.7601

Data no.58: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03476409.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0348	0.0348	0.0348	0.0348	0.0348	0.0348	0.0348	0.0348	0.0348
λ_1		0.3129	0.1391	0.0811	0.0521	0.0348	0.0232	0.0149	0.0087	0.0039
Predictors	x_1	1.7439	1.9051	1.9491	1.9780	1.9925	2.0148	2.0307	2.0427	2.0520
	x_2	-0	-0	-0.1246	-0.2849	-0.4248	-0.5161	-0.5811	-0.6298	-0.6678
	x_3	0.2689	0.4052	0.4195	0.3675	0.3029	0.2752	0.2556	0.2409	0.2295
	x_4	0	0	-0	-0.1683	-0.3610	-0.4671	-0.5427	-0.5992	-0.6435
	x_5	-0	-0.1799	-0.4407	-0.7210	-0.9723	-1.1173	-1.2205	-1.2979	-1.3582
	x_6	0	0.0592	0.1181	0.1713	0.2140	0.2344	0.2490	0.2597	0.2684
	x_7	0.8975	1.0284	1.0558	1.0662	1.0708	1.0710	1.0711	1.0712	1.0713
	x_8	0	-0	-0	-0.0702	-0.1232	-0.1657	-0.1960	-0.2185	-0.2364
	x_9	0	0.0405	0.1543	0.2243	0.2660	0.2963	0.3179	0.3340	0.3467
	x_{10}	0.0958	0.1641	0.1847	0.2019	0.2123	0.2234	0.2314	0.2374	0.2421
	x_{11}	0.5653	0.5407	0.5510	0.5783	0.6133	0.6437	0.6656	0.6820	0.6946
	x_{12}	0.0470	0.0649	0.0868	0.1128	0.1443	0.1726	0.1929	0.2084	0.2199
	x_{13}	0.1564	0.2071	0.2466	0.3036	0.3566	0.3880	0.4103	0.4272	0.4401
	x_{14}	0	0	-0	-0	-0.0065	-0.0472	-0.0764	-0.0985	-0.1154
	x_{15}	0.2585	0.2911	0.3286	0.3906	0.4448	0.4859	0.5152	0.5368	0.5542
df		8	11	12	14	15	15	15	15	15
PE		5.1748	4.7478	4.5356	4.3605	4.2340	4.1561	4.1087	4.0776	4.0560

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.59: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0269	0.0316	0.0594	0.0439	0.0658	0.0988	0.1059	0.1507	0.1940
λ_1		0.2422	0.1263	0.1386	0.0658	0.0658	0.0658	0.0454	0.0377	0.0216
Predictors	x_1	1.4544	1.5561	1.5164	1.5825	1.5551	1.5055	1.5055	1.4481	1.4024
	x_2	-0	-0.0342	-0	-0.1308	-0.0712	-0.0197	-0.0531	-0.0243	-0.0233
	x_3	0.1934	0.1663	0.2615	0.1544	0.2296	0.2930	0.2834	0.3311	0.3562
	x_4	0.2341	0.2864	0.3932	0.2930	0.3838	0.4486	0.4338	0.4712	0.4820
	x_5	-0.5872	-0.8857	-0.7108	-1.0234	-0.9055	-0.7945	-0.8483	-0.7615	-0.7241
	x_6	-0.4956	-0.6210	-0.5483	-0.6642	-0.6078	-0.5394	-0.5450	-0.4888	-0.4560
	x_7	1.5659	1.6781	1.6485	1.7268	1.7106	1.6760	1.6849	1.6389	1.6012
	x_8	0.0862	0.2126	0.1912	0.2637	0.2586	0.2495	0.2680	0.2616	0.2628
	x_9	0	0.0235	0.0281	0.0634	0.0861	0.1098	0.1231	0.1439	0.1603
	x_{10}	-0.0415	-0.1086	-0.1525	-0.1808	-0.2158	-0.2551	-0.2836	-0.3225	-0.3560
	x_{11}	1.2946	1.3190	1.1761	1.2983	1.1794	1.0474	1.0331	0.9160	0.8360
	x_{12}	0	-0	0	-0	-0	0	-0	0	0.0043
	x_{13}	0	0.0906	0.1040	0.1747	0.1697	0.1644	0.1767	0.1777	0.1837
	x_{14}	0	0	0	0.0042	0.0053	0.0137	0.0363	0.0526	0.0719
	x_{15}	0	0	0	0	0.0270	0.0745	0.0891	0.1308	0.1596
df		9	12	11	13	14	14	14	14	15
PE		7.7208	7.2347	7.4486	7.0856	7.2274	7.4334	7.4069	7.6367	7.8147

Data no.59: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.0331199.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0331	0.0331	0.0331	0.0331	0.0331	0.0331	0.0331	0.0331	0.0331
λ_1		0.2981	0.1325	0.0773	0.0497	0.0331	0.0221	0.0142	0.0083	0.0037
Predictors	x_1	1.3916	1.5517	1.5824	1.5952	1.6057	1.6138	1.6196	1.6239	1.6275
	x_2	-0	-0.0137	-0.1455	-0.2264	-0.2678	-0.2920	-0.3093	-0.3226	-0.3322
	x_3	0.2105	0.1817	0.1133	0.0609	0.03561	0.0220	0.0121	0.0044	0
	x_4	0.2209	0.3046	0.2353	0.1719	0.1377	0.1181	0.1038	0.0928	0.0854
	x_5	-0.4309	-0.8506	-1.0616	-1.1984	-1.2573	-1.2866	-1.3079	-1.3244	-1.3357
	x_6	-0.4157	-0.6091	-0.6788	-0.7129	-0.7488	-0.7780	-0.7989	-0.8144	-0.8266
	x_7	1.5083	1.6723	1.7205	1.7415	1.7499	1.7543	1.7574	1.7597	1.7615
	x_8	0.0258	0.2074	0.2539	0.2852	0.3053	0.3186	0.3281	0.3354	0.3407
	x_9	0	0.0222	0.0476	0.0508	0.0633	0.0759	0.0848	0.0913	0.0968
	x_{10}	-0.0306	-0.1067	-0.1507	-0.1794	-0.2010	-0.2164	-0.2274	-0.2358	-0.2421
	x_{11}	1.2328	1.3024	1.3581	1.3866	1.4129	1.4335	1.4482	1.4590	1.4675
	x_{12}	0	0	-0	-0	-0.0397	-0.0797	-0.1082	-0.1296	-0.1463
	x_{13}	0	0.0821	0.1613	0.1934	0.2136	0.2279	0.2382	0.2461	0.2518
	x_{14}	0	0	0	0.0355	0.0686	0.0924	0.1094	0.1226	0.1320
	x_{15}	0	0	0	0	0	0	0	0	0
df		9	12	12	13	14	14	14	14	13
PE		8.0758	7.2743	7.0535	6.9631	6.9042	6.8672	6.8445	6.8294	6.8191

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.60: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0221	0.0497	0.0776	0.0758	0.1370	0.1872	0.2203	0.1969	0.2104
λ_1		0.1987	0.1987	0.1810	0.1137	0.1370	0.1248	0.0944	0.0492	0.0234
Predictors	x_1	1.8107	1.7645	1.7292	1.7606	1.6655	1.6056	1.5801	1.6157	1.6000
	x_2	-0.1825	-0.1804	-0.1913	-0.2519	-0.2146	-0.2130	-0.2244	-0.2838	-0.3137
	x_3	0	0	0	0	0	0	0.0195	0.0172	0.0194
	x_4	0	0	0	0.0032	0.0216	0.0447	0.0758	0.0643	0.0615
	x_5	-0	-0	-0	-0.0350	-0	-0	-0	-0.0941	-0.1503
	x_6	0.5884	0.5561	0.5590	0.6809	0.5744	0.5460	0.5625	0.6456	0.6698
	x_7	1.1257	1.0849	1.0683	1.1522	1.0449	1.0026	1.0019	1.0758	1.0912
	x_8	0.3620	0.3247	0.3206	0.4384	0.3285	0.3032	0.3180	0.4111	0.4393
	x_9	0.2771	0.2754	0.2917	0.3607	0.3329	0.3418	0.3665	0.4063	0.4224
	x_{10}	0	-0	-0	0	-0	-0	-0	-0	-0.0005
	x_{11}	0.7845	0.7525	0.7336	0.7741	0.7026	0.6724	0.6616	0.7064	0.7111
	x_{12}	0.3645	0.3681	0.3766	0.3999	0.3908	0.3931	0.3964	0.4168	0.4255
	x_{13}	0.2149	0.2383	0.2552	0.2502	0.2810	0.2975	0.3028	0.3053	0.3097
	x_{14}	0	0	0	0	0	0	0.0009	0.0080	0.0262
	x_{15}	0	0	0	0	0	0	0	0	0
df		9	9	9	11	10	10	12	13	14
PE		5.2622	5.3501	5.3733	5.1246	5.4042	5.5255	5.5266	5.2762	5.2362

Data no.60: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03517935.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0352	0.0352	0.0352	0.0352	0.0352	0.0352	0.0352	0.0352	0.0352
λ_1		0.3166	0.1407	0.0821	0.0528	0.0352	0.0235	0.0151	0.0088	0.0039
Predictors	x_1	1.7275	1.8184	1.8176	1.8160	1.8180	1.8096	1.7903	1.7758	1.7629
	x_2	-0.0885	-0.2274	-0.3139	-0.3574	-0.3758	-0.4143	-0.4692	-0.5104	-0.5543
	x_3	0	0	0	-0	-0	-0.0485	-0.1259	-0.1838	-0.2433
	x_4	0	0	0	0	-0	-0.0493	-0.1325	-0.1948	-0.2615
	x_5	0	-0	-0.1724	-0.2687	-0.3248	-0.4021	-0.5016	-0.5761	-0.6498
	x_6	0.3401	0.6868	0.7851	0.8364	0.8764	0.8945	0.9039	0.9109	0.9119
	x_7	0.9566	1.1792	1.2619	1.3055	1.3370	1.3485	1.3517	1.3542	1.3546
	x_8	0.1376	0.4447	0.5789	0.6471	0.6853	0.7204	0.7531	0.7777	0.8019
	x_9	0.1487	0.3390	0.3997	0.4288	0.4433	0.4506	0.4500	0.4496	0.4520
	x_{10}	-0	0	0	0.0102	0.0489	0.0611	0.0603	0.0597	0.0512
	x_{11}	0.7049	0.7999	0.8681	0.9046	0.9283	0.9580	0.9880	1.0105	1.0315
	x_{12}	0.3274	0.3854	0.4204	0.4382	0.4471	0.4690	0.4985	0.5206	0.5388
	x_{13}	0.2479	0.2163	0.2258	0.2328	0.2417	0.2645	0.2892	0.3078	0.3169
	x_{14}	0	0	-0	-0	-0.0037	-0.0134	-0.0102	-0.0078	-0.0098
	x_{15}	0	0	-0	-0	-0	-0	-0	0	0.0206
df		9	9	10	11	12	14	14	14	15
PE		5.8712	5.1074	4.8878	4.8062	4.7633	4.7239	4.6902	4.6686	4.6508

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.61: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0131	0.0324	0.0461	0.0654	0.0561	0.0580	0.0622	0.0735	0.0786
λ_1		0.1181	0.1296	0.1076	0.0980	0.0561	0.0387	0.0267	0.0184	0.0087
Predictors	x_1	1.6016	1.5344	1.5572	1.5396	1.6495	1.6883	1.7085	1.7023	1.7135
	x_2	0.1846	0.1828	0.2162	0.2335	0.2846	0.3167	0.3390	0.3550	0.3716
	x_3	0	0	0	0	0	0.0542	0.0933	0.1122	0.1395
	x_4	0.9846	0.8836	0.9175	0.8932	1.0433	1.0786	1.0953	1.0821	1.0903
	x_5	-0	-0	-0	-0	-0	-0	-0	-0	-0
	x_6	-0	0	0	0	0	0.0248	0.0521	0.0795	0.1012
	x_7	1.5939	1.538	1.5321	1.5010	1.5641	1.5779	1.5818	1.5658	1.5655
	x_8	-0.1782	-0.1295	-0.1488	-0.1401	-0.2162	-0.2415	-0.2532	-0.2501	-0.2574
	x_9	-0.3105	-0.3452	-0.3598	-0.3794	-0.3645	-0.3742	-0.3858	-0.4044	-0.4152
	x_{10}	-0.1906	-0.1947	-0.2121	-0.2225	-0.2441	-0.2745	-0.2997	-0.3173	-0.3360
	x_{11}	1.2239	1.1382	1.1105	1.0537	1.1127	1.1195	1.1104	1.0757	1.0643
	x_{12}	-0.3632	-0.2832	-0.3037	-0.2917	-0.4041	-0.4395	-0.4601	-0.4514	-0.4621
	x_{13}	0	0	0.0289	0.0610	0.1227	0.1624	0.1881	0.2048	0.2222
	x_{14}	0	0	0	0	0	0.0171	0.0403	0.0511	0.0669
	x_{15}	-0	-0	-0	-0	-0.0261	-0.0823	-0.1210	-0.1356	-0.1610
df		9	9	10	10	11	14	14	14	14
PE		4.3902	4.5822	4.5429	4.6113	4.3326	4.2393	4.1975	4.2234	4.2075

Data no.61: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.0268111.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0268	0.0268	0.0268	0.0268	0.0268	0.0268	0.0268	0.0268	0.0268
λ_1		0.2413	0.1072	0.0626	0.0402	0.0268	0.0179	0.0115	0.0067	0.0030
Predictors	x_1	1.2934	1.5982	1.7025	1.7602	1.7947	1.8247	1.8589	1.8850	1.9050
	x_2	0.0396	0.2076	0.2676	0.3042	0.3243	0.3516	0.3944	0.4270	0.4520
	x_3	-0	0	0.0040	0.0802	0.1282	0.1773	0.2421	0.2914	0.3292
	x_4	0.5077	0.9789	1.1240	1.1796	1.2148	1.2615	1.3454	1.4101	1.4594
	x_5	-0	-0	-0	0	0	0.0290	0.1148	0.1809	0.2312
	x_6	0	-0	0	0	0	0.0145	0.0212	0.0258	0.0296
	x_7	1.4267	1.5759	1.6215	1.6487	1.6669	1.6783	1.6889	1.6970	1.7032
	x_8	-0	-0.1769	-0.2510	-0.2878	-0.3075	-0.3195	-0.3252	-0.3295	-0.3328
	x_9	-0.3353	-0.3335	-0.3210	-0.3225	-0.3228	-0.3232	-0.3179	-0.3137	-0.3106
	x_{10}	-0.1165	-0.2050	-0.2294	-0.2662	-0.2910	-0.3108	-0.3249	-0.3352	-0.3434
	x_{11}	1.0131	1.1842	1.2189	1.2466	1.2614	1.2624	1.2527	1.2453	1.2396
	x_{12}	-0.0822	-0.3453	-0.4652	-0.5316	-0.5781	-0.6129	-0.6492	-0.6769	-0.6982
	x_{13}	0	0.0060	0.0953	0.1508	0.1815	0.1994	0.2006	0.2006	0.2011
	x_{14}	-0	0	0	0.0326	0.0677	0.0866	0.0899	0.0921	0.0940
	x_{15}	-0	-0	-0.0310	-0.1167	-0.1719	-0.2139	-0.2517	-0.2796	-0.3015
df		8	10	12	13	13	15	15	15	15
PE		5.3491	4.4216	4.1947	4.0649	3.9990	3.9587	3.9253	3.9030	3.8875

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.62: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0150	0.0176	0.0331	0.0469	0.0641	0.0876	0.0780	0.0922	0.1081
λ_1		0.1350	0.0704	0.0772	0.0704	0.0641	0.0584	0.0334	0.0230	0.0120
Predictors	x_1	1.9834	2.0672	2.0159	1.9903	1.9574	1.9119	1.9670	1.9492	1.9329
	x_2	0	0.0514	0.0351	0.0438	0.0505	0.0551	0.09202	0.1032	0.1092
	x_3	-0.1769	-0.2639	-0.2603	-0.2739	-0.2860	-0.2964	-0.3331	-0.3480	-0.3655
	x_4	-0	-0.0072	-0	-0	-0	-0	-0.0452	-0.0559	-0.0752
	x_5	-0	-0	-0	-0	-0	-0	-0	0	0
	x_6	-0.1283	-0.1588	-0.1704	-0.1836	-0.1970	-0.2109	-0.2184	-0.2295	-0.2380
	x_7	1.2032	1.1796	1.142	1.1046	1.0635	1.0151	1.0159	0.9853	0.9559
	x_8	-0.1380	-0.2432	-0.2426	-0.2636	-0.2833	-0.3023	-0.3346	-0.3536	-0.3740
	x_9	-0.1916	-0.2538	-0.2489	-0.2577	-0.2659	-0.2738	-0.2986	-0.3094	-0.3188
	x_{10}	0.2182	0.3186	0.3070	0.3170	0.3241	0.3273	0.3761	0.3886	0.3974
	x_{11}	1.2307	1.3492	1.2807	1.2500	1.2112	1.1594	1.2300	1.2108	1.1837
	x_{12}	-0.0783	-0.1777	-0.1568	-0.1576	-0.1548	-0.1470	-0.1874	-0.1905	-0.200
	x_{13}	0.3699	0.4959	0.4684	0.4682	0.4643	0.4552	0.5082	0.5131	0.5120
	x_{14}	-0.3062	-0.4255	-0.3662	-0.3441	-0.3162	-0.2794	-0.3377	-0.3259	-0.3151
	x_{15}	0	-0	-0	0	0	0	0	0.0003	0.0294
df		11	13	12	12	12	12	13	14	14
PE		4.1172	3.8221	3.9096	3.9362	3.9808	4.0557	3.9074	3.9243	3.9482

Data no.62: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02941342.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0294	0.0294	0.0294	0.0294	0.0294	0.0294	0.0294	0.0294	0.0294
λ_1		0.2647	0.1177	0.0686	0.0441	0.0294	0.0196	0.0126	0.0074	0.0033
Predictors	x_1	1.7473	1.9695	2.0376	2.0725	2.0935	2.1037	2.1064	2.1082	2.1088
	x_2	0	0	0.0512	0.0883	0.1105	0.1199	0.1203	0.1200	0.1128
	x_3	-0.0149	-0.2066	-0.2707	-0.3128	-0.3380	-0.3635	-0.3919	-0.4138	-0.4388
	x_4	0	-0	-0.0028	-0.0611	-0.0961	-0.1277	-0.1601	-0.1850	-0.2158
	x_5	-0	-0	-0	-0	-0	-0.0181	-0.0523	-0.0790	-0.1130
	x_6	-0.0898	-0.1494	-0.1707	-0.1815	-0.1879	-0.1917	-0.1940	-0.1958	-0.1955
	x_7	1.1330	1.1605	1.1471	1.1222	1.1073	1.0983	1.0928	1.0886	1.0868
	x_8	-0	-0.1762	-0.2554	-0.2947	-0.3182	-0.3327	-0.3417	-0.3484	-0.3535
	x_9	-0.0614	-0.2104	-0.2570	-0.2852	-0.3021	-0.3116	-0.3162	-0.3194	-0.3197
	x_{10}	0.0636	0.2465	0.3210	0.3740	0.4058	0.4276	0.4441	0.4566	0.4655
	x_{11}	0.9690	1.2163	1.3103	1.3631	1.3949	1.4193	1.4408	1.4569	1.4685
	x_{12}	-0	-0.0989	-0.1720	-0.2114	-0.2349	-0.2466	-0.2504	-0.2532	-0.2570
	x_{13}	0.1481	0.3967	0.4868	0.5417	0.5746	0.5986	0.6178	0.6325	0.6427
	x_{14}	-0	-0.2997	-0.3937	-0.4394	-0.4669	-0.4824	-0.4899	-0.4953	-0.4978
	x_{15}	0	0	-0	-0	-0	-0	0	0	0.0156
df		8	11	13	13	13	14	14	14	15
PE		5.1253	4.0876	3.8593	3.7639	3.7197	3.6944	3.6778	3.6671	3.6595

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.63: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0222	0.0958	0.0711	0.1213	0.1377	0.2065	0.2927	0.1243	0.4057
λ_1		0.1997	0.3830	0.1658	0.1820	0.1377	0.1377	0.1254	0.0311	0.0451
Predictors	x_1	2.2117	1.7625	2.1359	1.9864	2.0194	1.8807	1.7549	2.2378	1.6836
	x_2	0.3401	0	0.3310	0.2359	0.2953	0.2299	0.1977	0.5675	0.2487
	x_3	0.5898	0.3031	0.6011	0.5396	0.5887	0.5445	0.5213	0.8396	0.5632
	x_4	-0	0	-0	-0	-0	-0	-0	-0	-0
	x_5	0	0	0	0	0	0	0.0380	0.1451	0.1367
	x_6	0.3094	0.2434	0.3294	0.3325	0.3465	0.3515	0.3586	0.4095	0.3838
	x_7	0.8954	0.9532	0.8807	0.8934	0.8712	0.8677	0.8513	0.8918	0.8217
	x_8	0.3812	0.2882	0.4153	0.4126	0.4421	0.4463	0.4520	0.4794	0.4846
	x_9	0.0783	0	0.1106	0.1005	0.1352	0.1355	0.1460	0.1973	0.2015
	x_{10}	-0.2379	-0	-0.2507	-0.1852	-0.2413	-0.1990	-0.1760	-0.4452	-0.2204
	x_{11}	0.5350	0.5686	0.5276	0.5300	0.5212	0.5161	0.4976	0.5802	0.4656
	x_{12}	-0	0	-0	-0	-0	-0	0	-0.1645	-0
	x_{13}	0.5730	0.5063	0.5410	0.5132	0.5095	0.4845	0.4558	0.5286	0.4231
	x_{14}	0.0012	0	0.0496	0.0720	0.0930	0.1182	0.1390	0.2045	0.1636
	x_{15}	-0	0	-0	-0	-0	-0	0	-0.1639	0
df		11	7	11	11	11	11	12	14	12
PE		6.8272	7.8222	6.8319	7.0392	6.9198	7.1150	7.3142	6.2892	7.3603

Data no.63: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03552071.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0355	0.0355	0.0355	0.0355	0.0355	0.0355	0.0355	0.0355	0.0355
λ_1		0.3197	0.1421	0.0829	0.0533	0.0355	0.0237	0.0152	0.0089	0.0039
Predictors	x_1	1.9289	2.2869	2.4117	2.4877	2.5435	2.5952	2.6322	2.6599	2.6815
	x_2	0.0383	0.4480	0.6090	0.7265	0.8163	0.9038	0.9663	1.0132	1.0496
	x_3	0.3802	0.6750	0.8364	0.9552	1.0496	1.1454	1.2138	1.2650	1.3049
	x_4	0	-0	-0	0	0.0547	0.1697	0.2518	0.3133	0.3612
	x_5	0	0	0.0170	0.1249	0.2331	0.3671	0.4628	0.5345	0.5904
	x_6	0.2850	0.3283	0.3653	0.3963	0.4145	0.4260	0.4342	0.4403	0.4451
	x_7	0.9969	0.8492	0.8713	0.8927	0.9046	0.9113	0.9161	0.9195	0.9223
	x_8	0.3085	0.4247	0.4433	0.4413	0.4431	0.4484	0.4522	0.4551	0.4574
	x_9	0	0.1273	0.1478	0.1604	0.1669	0.1697	0.1716	0.1731	0.1743
	x_{10}	-0	-0.3327	-0.4701	-0.5475	-0.5897	-0.6118	-0.6276	-0.6395	-0.6487
	x_{11}	0.5675	0.5135	0.6009	0.6328	0.6366	0.6173	0.6036	0.5931	0.5851
	x_{12}	0	-0	-0.1441	-0.2329	-0.2941	-0.3462	-0.3835	-0.4114	-0.4333
	x_{13}	0.5346	0.5696	0.5856	0.5786	0.5725	0.5657	0.5609	0.5573	0.5545
	x_{14}	0	0.0347	0.1498	0.2129	0.2429	0.2515	0.2577	0.2623	0.2659
	x_{15}	0	-0	-0.1389	-0.2363	-0.3037	-0.3614	-0.4027	-0.4333	-0.4574
df		8	11	14	14	15	15	15	15	15
PE		7.4725	6.6523	6.2370	6.0444	5.9412	5.8658	5.8189	5.7876	5.7655

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.64: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .

	α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0061	0.0503	0.0177	0.0251	0.0454	0.0390	0.1059	0.0044	0.1769
λ_1		0.0547	0.2012	0.0414	0.0377	0.0454	0.0260	0.0454	0.0011	0.0197
Predictors	x_1	2.0663	1.9717	2.0570	2.0560	2.0425	2.0627	1.9622	2.0553	1.8667
	x_2	0.1210	0.2249	0.1312	0.1529	0.2045	0.1986	0.2463	-0.0096	0.2469
	x_3	-0.9154	-0	-0.9235	-0.8755	-0.6651	-0.8180	-0.4112	-1.6883	-0.3598
	x_4	0	0.6325	0	0.0543	0.2376	0.1483	0.4476	-0.5691	0.5046
	x_5	-1.3362	-0.0715	-1.3362	-1.2632	-0.9729	-1.1655	-0.6325	-2.2771	-0.5533
	x_6	0.6540	0.3999	0.6596	0.6571	0.6251	0.6667	0.5624	0.8101	0.5342
	x_7	0.9301	0.9047	0.9356	0.9377	0.9425	0.9402	0.9516	0.8406	0.9457
	x_8	-0.3822	-0.0129	-0.3902	-0.3826	-0.3293	-0.3851	-0.2617	-0.5582	-0.2540
	x_9	0	0	0.0057	0.0115	0.0126	0.0229	0.0209	0.0316	0.0403
	x_{10}	0	-0	0	0	-0	0.0054	-0.0443	0.3032	-0.0829
	x_{11}	0.8397	0.4895	0.8319	0.8070	0.7181	0.7758	0.5944	1.1055	0.5375
	x_{12}	0.6128	0.2827	0.6079	0.5884	0.5116	0.5698	0.4109	0.8827	0.3734
	x_{13}	0.3066	0.0520	0.3209	0.3119	0.2575	0.3129	0.2216	0.5342	0.2329
	x_{14}	-0	0	-0	-0.0023	-0	-0.0376	-0	-0.1477	0.0001
	x_{15}	0	0	0	0	0	-0	0	0.0870	0.0230
df		10	10	11	13	12	14	13	15	15
PE		3.6193	4.6407	3.6158	3.6400	3.7769	3.6582	4.0174	3.3695	4.1393

Data no.64: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.03791819.

	α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0379	0.0379	0.0379	0.0379	0.0379	0.0379	0.0379	0.0379	0.0379
λ_1		0.3413	0.1517	0.0885	0.0569	0.0379	0.0253	0.0163	0.0095	0.0042
Predictors	x_1	1.8406	2.0468	2.0497	2.0486	2.0507	2.0649	2.0793	2.0901	2.1005
	x_2	0.0804	0.2549	0.2178	0.1969	0.1884	0.1974	0.2081	0.2162	0.2269
	x_3	0	-0	-0.4208	-0.6365	-0.7623	-0.8330	-0.8995	-0.9493	-0.9853
	x_4	0.5734	0.5955	0.3567	0.2341	0.1649	0.1383	0.1172	0.1016	0.0957
	x_5	-0	-0.2033	-0.6974	-0.9502	-1.1012	-1.1835	-1.2504	-1.3005	-1.3333
	x_6	0.2568	0.4699	0.5695	0.6169	0.6437	0.6707	0.6981	0.7188	0.7355
	x_7	0.8191	0.9505	0.9448	0.9433	0.9407	0.9385	0.9286	0.9212	0.9157
	x_8	-0	-0.0957	-0.2421	-0.3167	-0.3577	-0.3887	-0.4103	-0.4264	-0.4375
	x_9	0	0	0	0.0005	0.0160	0.0234	0.0316	0.0377	0.0437
	x_{10}	-0	-0.0229	-0	-0	0	0.0115	0.0531	0.0844	0.1074
	x_{11}	0.4999	0.5221	0.6537	0.7199	0.7556	0.7821	0.8069	0.8256	0.8403
	x_{12}	0.2386	0.3183	0.4445	0.5079	0.5459	0.5758	0.6035	0.6242	0.6399
	x_{13}	0.0325	0.0575	0.1785	0.2426	0.2849	0.3170	0.3413	0.3595	0.3743
	x_{14}	0	0	-0	-0	-0.0041	-0.0410	-0.0751	-0.1007	-0.1192
	x_{15}	0	0	0	0	-0	-0	-0	-0	-0.0073
df		8	11	11	12	13	14	14	14	15
PE		5.0676	4.4187	3.9694	3.7951	3.7072	3.6483	3.6021	3.5715	3.5505

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.65: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0472	0.0968	0.1660	0.1953	0.1676	0.2088	0.0733	0.5072	0.3405
λ_1		0.4250	0.3872	0.3872	0.2929	0.1676	0.1392	0.0314	0.1268	0.0378
Predictors	x_1	1.7756	1.7422	1.6593	1.6946	1.7918	1.7600	2.0495	1.4782	1.6754
	x_2	0	0	0	0.0384	0.1177	0.1438	0.3980	0.1692	0.2409
	x_3	0	0	0	0.0728	0.1655	0.1860	0.3348	0.1848	0.2598
	x_4	0	0	0	0.0182	0.0489	0.0894	0.1727	0.2098	0.2078
	x_5	-0	-0	-0	-0	-0.1509	-0.1551	-0.3654	-0.0143	-0.1789
	x_6	0	0	0	0.0031	0.0975	0.1149	0.2002	0.1085	0.1669
	x_7	0.7677	0.7720	0.7285	0.8063	0.9512	0.9487	1.1252	0.7742	0.9403
	x_8	-0.0892	-0.1239	-0.1390	-0.1912	-0.2343	-0.2499	-0.3594	-0.2662	-0.3055
	x_9	-0.1632	-0.1856	-0.1890	-0.2465	-0.3286	-0.3424	-0.3922	-0.3152	-0.3783
	x_{10}	0	0	0	0	0	0	-0.1260	0	0
	x_{11}	0.9751	0.9064	0.8222	0.7940	0.8489	0.8078	1.1596	0.6104	0.7305
	x_{12}	0	0	0	0	-0	-0	-0.3299	0	-0.0272
	x_{13}	0.1768	0.2223	0.2341	0.2353	0.2301	0.2353	0.2789	0.2442	0.2591
	x_{14}	0	0.0143	0.0646	0.1050	0.1366	0.1554	0.2727	0.1890	0.2145
	x_{15}	0	0	0	0	-0	-0	-0.2112	0	-0.0426
df		6	7	7	11	12	12	15	12	14
PE		7.9661	7.9767	8.1712	7.9075	7.3292	7.3524	6.3420	8.1458	7.3616

Data no.65: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.04194738.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0419	0.0419	0.0419	0.0419	0.0419	0.0419	0.0419	0.0419	0.0419
λ_1		0.3775	0.1678	0.0979	0.0629	0.0419	0.0280	0.0180	0.0105	0.0047
Predictors	x_1	1.8167	1.9345	1.9910	2.0362	2.0795	2.1084	2.1290	2.1446	2.1566
	x_2	0	0.0977	0.2213	0.3097	0.3928	0.4481	0.4876	0.5177	0.5408
	x_3	0.0130	0.1653	0.2154	0.2546	0.3114	0.3492	0.3762	0.3971	0.4129
	x_4	0	0	0	0.0236	0.1140	0.1741	0.2171	0.2502	0.2753
	x_5	-0	-0.2831	-0.4057	-0.4599	-0.4282	-0.4073	-0.3924	-0.3801	-0.3714
	x_6	0	0.1208	0.1543	0.1826	0.1998	0.2113	0.2195	0.2257	0.2305
	x_7	0.8234	1.0576	1.1091	1.1241	1.1369	1.1455	1.1516	1.1562	1.1598
	x_8	-0.1146	-0.1959	-0.2649	-0.3160	-0.3533	-0.3782	-0.3959	-0.4093	-0.4197
	x_9	-0.1959	-0.3422	-0.3710	-0.3811	-0.3856	-0.3887	-0.3908	-0.3924	-0.3937
	x_{10}	0	-0	-0	-0.0802	-0.1346	-0.1709	-0.1968	-0.2162	-0.2313
	x_{11}	0.9893	1.0575	1.1525	1.2249	1.2545	1.2742	1.2883	1.2989	1.3071
	x_{12}	0	-0	-0.1637	-0.2813	-0.3674	-0.4248	-0.4658	-0.4966	-0.5205
	x_{13}	0.1841	0.1895	0.2248	0.2566	0.2701	0.2791	0.2856	0.2902	0.2940
	x_{14}	0	0.0809	0.1939	0.2598	0.2835	0.2993	0.3106	0.3188	0.3253
	x_{15}	0	-0	-0.0776	-0.1612	-0.2206	-0.2602	-0.2885	-0.3099	-0.3265
df		7	11	13	15	15	15	15	15	15
PE		7.7937	6.9916	6.5968	6.3781	6.2704	6.2124	6.1778	6.1553	6.1402

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.66: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0385	0.0789	0.1233	0.1592	0.2176	0.2250	0.1663	0.2156	0.1447
λ_1		0.3465	0.3157	0.2877	0.2388	0.2176	0.1500	0.0713	0.0539	0.0161
Predictors	x_1	2.0308	1.9772	1.9191	1.8922	1.8161	1.8446	1.9632	1.8958	2.0021
	x_2	0	0.0078	0.0381	0.0827	0.0983	0.1511	0.1991	0.1985	0.2033
	x_3	0	-0	-0	-0	-0	-0	-0.0596	-0.0829	-0.2350
	x_4	0	0	0	0	0	0	0	0	-0
	x_5	-0	-0	-0	-0	-0	-0.0174	-0.2179	-0.2389	-0.3942
	x_6	0	0	0	0	-0	-0	-0	-0	-0.0041
	x_7	0.9959	0.9962	0.9955	1.0409	1.0109	1.1149	1.3222	1.2784	1.4713
	x_8	0	0	0	0	0	0.0006	0.1054	0.1131	0.1819
	x_9	0.0956	0.1362	0.1742	0.2280	0.2484	0.3134	0.4172	0.4235	0.4941
	x_{10}	0.0352	0.0578	0.0730	0.1130	0.1143	0.1792	0.2957	0.2889	0.3804
	x_{11}	0.4705	0.4373	0.4082	0.3947	0.3724	0.3809	0.4471	0.4313	0.5352
	x_{12}	0	0	0	0	0	-0	-0	-0	-0.1054
	x_{13}	0.1166	0.1375	0.1548	0.1720	0.1801	0.1978	0.2501	0.2522	0.3240
	x_{14}	0.2033	0.2196	0.2280	0.2343	0.2403	0.2469	0.2688	0.2733	0.2832
	x_{15}	0	0	0	0	0	0	0	0.0170	0.0705
df		7	8	8	8	8	10	11	12	14
PE		6.4287	6.4298	6.4505	6.3600	6.4812	6.2358	5.6608	5.7492	5.2942

Data no.66: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.03208399.

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0321	0.0321	0.0321	0.0321	0.0321	0.0321	0.0321	0.0321	0.0321
λ_1		0.2888	0.1283	0.0749	0.0481	0.0321	0.0214	0.0138	0.0080	0.0036
Predictors	x_1	2.0853	2.1883	2.1760	2.1674	2.1623	2.1489	2.1309	2.1173	2.1064
	x_2	0.0312	0.1965	0.1928	0.1750	0.1645	0.1431	0.1156	0.0949	0.0781
	x_3	-0	-0	-0.1558	-0.2994	-0.3854	-0.4565	-0.5191	-0.5661	-0.6026
	x_4	0	0	0	-0	-0	-0.0381	-0.0978	-0.1426	-0.1774
	x_5	-0	-0.1359	-0.3153	-0.4463	-0.5246	-0.6051	-0.6870	-0.7484	-0.7973
	x_6	0	0	0	0	0	-0	-0	-0	-0.0034
	x_7	1.1169	1.4638	1.5772	1.6557	1.7028	1.7370	1.7638	1.7838	1.7994
	x_8	0	0.0625	0.1441	0.1813	0.2036	0.2131	0.2153	0.2170	0.2185
	x_9	0.1570	0.3637	0.4514	0.4885	0.5108	0.5254	0.5355	0.5432	0.5488
	x_{10}	0.1107	0.3104	0.3904	0.4338	0.4598	0.4780	0.4918	0.5021	0.5110
	x_{11}	0.4907	0.5234	0.6073	0.6780	0.7203	0.7514	0.7759	0.7944	0.8088
	x_{12}	0	-0	-0.0815	-0.1650	-0.2152	-0.2396	-0.2490	-0.2561	-0.2615
	x_{13}	0.1314	0.2087	0.3121	0.3701	0.4049	0.4307	0.4517	0.4675	0.4799
	x_{14}	0.2029	0.2105	0.2384	0.2345	0.2323	0.2367	0.2447	0.2508	0.2552
	x_{15}	0	0	0	0.0641	0.1026	0.1354	0.1648	0.1869	0.2047
df		8	10	12	13	13	14	14	14	15
PE		6.1453	5.5062	5.2119	5.0500	4.9751	4.9301	4.8993	4.8790	4.8649

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.67: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0276	0.0116	0.0151	0.0283	0.0292	0.1339	0.0749	0.1169	0.0449
λ_1		0.2484	0.0465	0.0352	0.0424	0.0292	0.0893	0.0321	0.0292	0.0050
Predictors	x_1	2.0239	2.1458	2.1433	2.1255	2.1269	1.9629	2.0848	2.0285	2.1185
	x_2	0	-0.1267	-0.1596	-0.0964	-0.1407	0	-0	0	-0.1738
	x_3	0	0	0	0	0	0	0.0453	0.0391	0.0216
	x_4	0.1256	0	0	0	0	0.1255	0.1190	0.1452	0.0073
	x_5	-0	-0.7773	-0.8063	-0.7054	-0.7534	-0.1933	-0.4601	-0.3560	-0.7584
	x_6	-0.0770	-0.1853	-0.1988	-0.2024	-0.2152	-0.2149	-0.2422	-0.2622	-0.2452
	x_7	1.5544	1.7305	1.7295	1.7063	1.7102	1.5388	1.6456	1.5884	1.6984
	x_8	0.1230	0.4490	0.4751	0.4484	0.4791	0.3283	0.4360	0.4196	0.5228
	x_9	0	0.0558	0.0673	0.0717	0.0824	0.0796	0.1014	0.1149	0.1087
	x_{10}	-0.5276	-0.9228	-0.9429	-0.9041	-0.9315	-0.6941	-0.8656	-0.8175	-0.9555
	x_{11}	0.7057	0.9671	0.9819	0.9051	0.9346	0.6083	0.7671	0.6876	0.9289
	x_{12}	0	-0.1022	-0.1240	-0.0865	-0.1145	-0	-0.0627	-0.0320	-0.1454
	x_{13}	0	0.3562	0.3957	0.3538	0.3993	0.2157	0.3249	0.3116	0.4594
	x_{14}	0	-0.2625	-0.2912	-0.2384	-0.2762	-0	-0.2003	-0.1542	-0.3133
	x_{15}	0.3021	0.4760	0.4928	0.4716	0.4921	0.3434	0.4312	0.4073	0.5127
df		8	13	13	13	13	11	14	14	15
PE		4.4087	3.3296	3.2927	3.3820	3.3243	4.1135	3.5535	3.7230	3.2959

Data no.67: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.02402008.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0240	0.0240	0.0240	0.0240	0.0240	0.0240	0.0240	0.0240	0.0240
λ_1		0.2162	0.0961	0.0560	0.0360	0.0240	0.0160	0.0103	0.0060	0.0027
Predictors	x_1	2.0459	2.1183	2.1285	2.1321	2.1346	2.1379	2.1347	2.1335	2.1215
	x_2	0	-0	-0.0568	-0.1300	-0.1736	-0.2008	-0.2379	-0.2657	-0.3078
	x_3	0	0	0	0	0.0009	0.0077	0	-0	-0.0350
	x_4	0.1420	0.0302	0.0022	0	0	-0	-0.0361	-0.0671	-0.1201
	x_5	-0.0327	-0.5141	-0.6690	-0.7523	-0.8017	-0.8334	-0.8908	-0.9366	-1.0034
	x_6	-0.0956	-0.1466	-0.1868	-0.2052	-0.2162	-0.2227	-0.2255	-0.2265	-0.2290
	x_7	1.5878	1.6897	1.7073	1.7154	1.7203	1.7239	1.7258	1.7276	1.7268
	x_8	0.1628	0.3325	0.4179	0.4665	0.4956	0.5153	0.5319	0.5448	0.5558
	x_9	0	0.0230	0.0583	0.0737	0.0827	0.0880	0.0917	0.0936	0.0984
	x_{10}	-0.5793	-0.7910	-0.8808	-0.9253	-0.9521	-0.9694	-0.9785	-0.9843	-0.9886
	x_{11}	0.7192	0.8084	0.8868	0.9391	0.9703	0.9907	1.0103	1.0256	1.0430
	x_{12}	-0	-0	-0.0613	-0.1076	-0.1353	-0.1542	-0.1645	-0.1721	-0.1738
	x_{13}	0.0094	0.1877	0.3080	0.3810	0.4248	0.4543	0.4843	0.5080	0.5317
	x_{14}	-0	-0.0768	-0.2055	-0.2672	-0.3045	-0.3300	-0.3438	-0.3547	-0.3570
	x_{15}	0.3068	0.3784	0.4519	0.4851	0.5045	0.5161	0.5319	0.5429	0.5629
df		10	12	14	13	14	14	14	14	15
PE		4.2744	3.6862	3.4384	3.3326	3.2833	3.2562	3.2343	3.2192	3.2045

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.68: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0216	0.0534	0.0524	0.1077	0.1015	0.2209	0.2368	0.2550	0.2726
λ_1		0.1947	0.2137	0.1223	0.1616	0.1015	0.1473	0.1015	0.0637	0.0303
Predictors	x_1	1.7133	1.6311	1.6845	1.5707	1.6143	1.4128	1.4246	1.4265	1.4226
	x_2	-0	-0	-0.1332	-0	-0.1494	-0	-0.0771	-0.1426	-0.2012
	x_3	0.4724	0.4774	0.4238	0.4834	0.4418	0.4847	0.4792	0.4757	0.4828
	x_4	0.2863	0.2537	0.2041	0.2108	0.1833	0.1547	0.1541	0.1526	0.1601
	x_5	-0.3226	-0.2551	-0.5554	-0.3226	-0.5314	-0.2730	-0.3597	-0.4232	-0.4570
	x_6	-0	-0	-0	-0	-0	-0	-0	-0	-0.0105
	x_7	1.3220	1.2550	1.3574	1.2401	1.3112	1.1253	1.1513	1.1680	1.1790
	x_8	0.1773	0.1852	0.2307	0.2221	0.2585	0.2492	0.2765	0.2955	0.3046
	x_9	-0	-0	-0	-0	-0	-0	-0	-0.0080	-0.0449
	x_{10}	-0.0947	-0.0618	-0.1659	-0.1039	-0.1682	-0.0875	-0.1250	-0.1540	-0.1755
	x_{11}	0.5425	0.5036	0.6079	0.5075	0.5765	0.4578	0.4810	0.4964	0.5126
	x_{12}	0	0	0	0	0	-0	-0	-0	-0
	x_{13}	0	0.0029	0.1104	0.0610	0.1364	0.0950	0.1316	0.1575	0.1827
	x_{14}	-0	-0	-0	-0	-0	-0	-0	-0.0079	-0.0505
	x_{15}	0	0	0	0	0	0	0	0.0146	0.0459
df		8	9	10	9	10	9	10	13	14
PE		5.6036	5.7682	5.3720	5.7217	5.4311	5.9999	5.8404	5.7368	5.6475

Data no.68: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.0443203.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0443	0.0443	0.0443	0.0443	0.0443	0.0443	0.0443	0.0443	0.0443
λ_1		0.3989	0.1773	0.1034	0.0665	0.0443	0.0295	0.0190	0.0111	0.0049
Predictors	x_1	1.4606	1.6794	1.7001	1.7053	1.7122	1.7179	1.7221	1.7252	1.7276
	x_2	0	-0	-0.2137	-0.3514	-0.4548	-0.5415	-0.6032	-0.6495	-0.6856
	x_3	0.4286	0.4766	0.3846	0.3274	0.2929	0.2575	0.2323	0.2135	0.1988
	x_4	0.1103	0.2660	0.1716	0.1136	0.0819	0.0572	0.0397	0.0266	0.0164
	x_5	-0	-0.3399	-0.6766	-0.8652	-0.9829	-1.0780	-1.1456	-1.1963	-1.2357
	x_6	-0	-0	0	0	0.0060	0.0407	0.0654	0.0840	0.0985
	x_7	1.0766	1.3083	1.3899	1.4304	1.4648	1.4920	1.5114	1.5260	1.5373
	x_8	0.1075	0.1936	0.2422	0.2694	0.2838	0.2902	0.2949	0.2983	0.3010
	x_9	-0	-0	-0	-0	-0.0128	-0.0345	-0.0499	-0.0616	-0.0706
	x_{10}	-0	-0.1070	-0.1904	-0.2320	-0.2660	-0.2912	-0.3092	-0.3227	-0.3332
	x_{11}	0.3775	0.5408	0.6515	0.7139	0.7550	0.7870	0.8099	0.8270	0.8403
	x_{12}	0	0	0	0	0.0454	0.0837	0.1111	0.1316	0.1475
	x_{13}	0	0.0199	0.1525	0.2319	0.2724	0.2932	0.3080	0.3190	0.3277
	x_{14}	-0	-0	-0	-0.0048	-0.0535	-0.0935	-0.1221	-0.1435	-0.1602
	x_{15}	0	0	0	0	0.0114	0.0367	0.0546	0.0681	0.0785
df		6	9	10	11	15	15	15	15	15
PE		6.5787	5.6018	5.2630	5.1227	5.0233	4.9554	4.9144	4.8877	4.8694

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.69: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0275	0.0427	0.0667	0.0861	0.1177	0.1010	0.1893	0.1541	0.2178
λ_1		0.2477	0.1707	0.1556	0.1292	0.1177	0.0673	0.0811	0.0385	0.0242
Predictors	x_1	2.1420	2.1611	2.1088	2.0724	1.9917	2.0202	1.8449	1.9066	1.7982
	x_2	0.2023	0.2375	0.2676	0.2987	0.3179	0.3199	0.3553	0.3522	0.3708
	x_3	0	0	0	0	-0	-0	-0	-0.0108	-0.0435
	x_4	0.8408	0.8806	0.8748	0.8803	0.8401	0.8298	0.7795	0.7772	0.7274
	x_5	-0	-0	-0	-0.0009	-0.0687	-0.1825	-0.1848	-0.2727	-0.3189
	x_6	0	0.0152	0.0394	0.0670	0.0803	0.0908	0.1164	0.1202	0.1457
	x_7	1.7694	1.7989	1.7451	1.7095	1.6429	1.6963	1.5259	1.6022	1.5006
	x_8	0.3145	0.3903	0.3837	0.3936	0.3789	0.4379	0.3688	0.4252	0.4006
	x_9	-0	-0	-0	-0	-0	-0	-0	-0.0094	-0.0104
	x_{10}	0	0	0.0264	0.0664	0.1100	0.1574	0.1943	0.2226	0.2573
	x_{11}	0.3340	0.3353	0.3420	0.3478	0.3603	0.3747	0.3803	0.3938	0.4003
	x_{12}	0.1757	0.2034	0.1965	0.1994	0.1986	0.2291	0.2127	0.2384	0.2394
	x_{13}	0	0	0	0	0	0	0	0	0.0159
	x_{14}	0.0566	0.0520	0.0673	0.0744	0.1146	0.1378	0.1721	0.1842	0.2010
	x_{15}	0	0	0	0	0	0	0	0	0.0143
df		8	9	10	11	11	11	11	13	15
PE		4.8107	4.6946	4.7156	4.7094	4.7476	4.6216	4.8338	4.6820	4.8015

Data no.69: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03140259.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0314	0.0314	0.0314	0.0314	0.0314	0.0314	0.0314	0.0314	0.0314
λ_1		0.2826	0.1256	0.0733	0.0471	0.0314	0.0209	0.0135	0.0079	0.0035
Predictors	x_1	2.1079	2.2118	2.2146	2.2050	2.1987	2.194	2.1905	2.1879	2.1859
	x_2	0.1957	0.2433	0.2630	0.2628	0.2625	0.2687	0.2728	0.2762	0.2788
	x_3	0	0	-0	-0	-0.0198	-0.0547	-0.0771	-0.0939	-0.1067
	x_4	0.8144	0.9139	0.9160	0.8646	0.8236	0.7974	0.7828	0.7720	0.7639
	x_5	-0	-0	-0.0669	-0.1654	-0.2294	-0.2621	-0.2815	-0.2958	-0.3066
	x_6	0	0.0419	0.0699	0.0754	0.0804	0.0875	0.0930	0.0972	0.1005
	x_7	1.7291	1.8590	1.8806	1.8955	1.9049	1.9132	1.9201	1.9252	1.9292
	x_8	0.2701	0.4555	0.5099	0.5345	0.5454	0.5461	0.5469	0.5475	0.5481
	x_9	-0	-0	-0	-0.0350	-0.0613	-0.0871	-0.1055	-0.1192	-0.1299
	x_{10}	0	0.0138	0.0707	0.1066	0.1249	0.1345	0.1423	0.1482	0.1528
	x_{11}	0.3360	0.3341	0.3451	0.3593	0.3758	0.4005	0.4174	0.4302	0.4402
	x_{12}	0.1561	0.2404	0.2697	0.2895	0.3104	0.3419	0.3662	0.3843	0.3986
	x_{13}	0	0	-0	-0	-0.0110	-0.0463	-0.0710	-0.0897	-0.1043
	x_{14}	0.0686	0.0240	0.0354	0.0684	0.0933	0.1216	0.1420	0.1575	0.1693
	x_{15}	0	0	-0	-0	-0	-0.0138	-0.0309	-0.0437	-0.0537
df		8	10	11	12	14	15	15	15	15
PE		4.9020	4.5993	4.5268	4.4914	4.4711	4.4491	4.4362	4.4286	4.4240

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.70: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0157	0.0514	0.0553	0.0651	0.0889	0.1009	0.1303	0.1061	0.1367
λ_1		0.1416	0.2055	0.1290	0.0976	0.0889	0.0673	0.0559	0.0265	0.0152
Predictors	x_1	1.3737	1.2989	1.3588	1.3672	1.3534	1.3557	1.3368	1.3784	1.3579
	x_2	-0	-0	-0	-0	-0	-0	-0	-0	-0
	x_3	0.2376	0.2722	0.3271	0.3556	0.3881	0.4100	0.4372	0.4372	0.4643
	x_4	0.2818	0.2857	0.3346	0.3579	0.3735	0.3892	0.3996	0.4212	0.4298
	x_5	-0.8295	-0.5893	-0.6969	-0.6963	-0.6460	-0.6363	-0.5925	-0.6411	-0.5995
	x_6	-0.7847	-0.7346	-0.7801	-0.7971	-0.7937	-0.8047	-0.7999	-0.8307	-0.8254
	x_7	1.6301	1.4809	1.5607	1.5800	1.5454	1.5492	1.5119	1.5927	1.5507
	x_8	0	0	0	0.0108	0.0132	0.0318	0.0353	0.0767	0.0773
	x_9	-0.1315	-0.0707	-0.1622	-0.2008	-0.2121	-0.2366	-0.2462	-0.2778	-0.2867
	x_{10}	-0.1649	-0.0895	-0.1750	-0.1981	-0.2026	-0.2166	-0.2211	-0.2442	-0.2485
	x_{11}	0.6357	0.5410	0.5734	0.5862	0.5612	0.5644	0.5414	0.5988	0.5713
	x_{12}	-0	0	0	-0	0	0	0	-0	0
	x_{13}	0.4586	0.4197	0.4395	0.4649	0.4583	0.4715	0.4659	0.5088	0.4992
	x_{14}	-0	-0	-0	-0	-0	-0	-0	-0.0323	-0.0248
	x_{15}	-0	-0	-0	-0.0432	-0.0427	-0.0665	-0.0657	-0.1125	-0.1083
df		10	10	10	12	12	12	12	13	13
PE		3.3068	3.6882	3.3929	3.2964	3.3448	3.3022	3.3595	3.1728	3.2342

Data no.70: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.03555224.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0356	0.0356	0.0356	0.0356	0.0356	0.0356	0.0356	0.0356	0.0356
λ_1		0.3200	0.1422	0.0830	0.0533	0.0356	0.0237	0.0152	0.0089	0.0040
Predictors	x_1	1.2211	1.3625	1.3906	1.4158	1.4513	1.4749	1.4918	1.5045	1.5144
	x_2	-0	-0	-0.0156	-0.0233	-0.0162	-0.0115	-0.0081	-0.0056	-0.0037
	x_3	0.1985	0.2847	0.2999	0.3174	0.3679	0.4016	0.4256	0.4437	0.4577
	x_4	0.2053	0.3095	0.3270	0.3531	0.4003	0.4317	0.4541	0.4710	0.4840
	x_5	-0.4236	-0.7432	-0.8116	-0.8083	-0.7425	-0.6986	-0.6673	-0.6438	-0.6255
	x_6	-0.6603	-0.7786	-0.8175	-0.8357	-0.8480	-0.8561	-0.8619	-0.8663	-0.8697
	x_7	1.3817	1.5853	1.6619	1.7146	1.7612	1.7922	1.8143	1.8309	1.8438
	x_8	0	0	0.0342	0.0728	0.0845	0.0922	0.0977	0.1019	0.1051
	x_9	-0	-0.1408	-0.2130	-0.2404	-0.2536	-0.2624	-0.2687	-0.2734	-0.2771
	x_{10}	-0	-0.1626	-0.2096	-0.2327	-0.2569	-0.2730	-0.2845	-0.2932	-0.2999
	x_{11}	0.4914	0.5974	0.6558	0.7037	0.7432	0.7696	0.7884	0.8025	0.8135
	x_{12}	0	-0	-0	-0.0162	-0.0946	-0.1468	-0.1841	-0.2121	-0.2339
	x_{13}	0.3971	0.4452	0.5096	0.5528	0.5896	0.6141	0.6316	0.6447	0.6549
	x_{14}	0	-0	-0	-0.0510	-0.0822	-0.1030	-0.1179	-0.1291	-0.1377
	x_{15}	-0	-0	-0.0837	-0.1198	-0.1393	-0.1523	-0.1616	-0.1686	-0.1740
df		8	10	13	15	15	15	15	15	15
PE		4.2242	3.3718	3.1430	3.0283	2.9471	2.9024	2.8750	2.8569	2.8444

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.71: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0243	0.0499	0.0710	0.1007	0.1254	0.1423	0.1837	0.1978	0.1755
λ_1		0.2191	0.1996	0.1657	0.1510	0.1254	0.0948	0.0787	0.0494	0.0195
Predictors	x_1	2.0044	1.9795	1.9609	1.9258	1.8971	1.8827	1.8323	1.8231	1.8631
	x_2	0.1121	0.1506	0.1918	0.2176	0.2407	0.2823	0.2968	0.3335	0.3961
	x_3	0	0.0267	0.0637	0.1036	0.1286	0.1448	0.1718	0.1835	0.2006
	x_4	0.2286	0.2822	0.3268	0.3755	0.4050	0.4264	0.4556	0.4707	0.4946
	x_5	-0.9088	-0.8678	-0.8771	-0.8274	-0.8278	-0.8480	-0.8079	-0.8317	-0.8974
	x_6	0	0	0.0234	0.0443	0.0751	0.1139	0.1351	0.1699	0.1933
	x_7	1.0207	1.0278	1.0439	1.0394	1.0439	1.0691	1.0556	1.0795	1.1304
	x_8	-0	-0	-0	-0	-0	-0.0574	-0.0647	-0.1226	-0.2017
	x_9	-0.0135	-0.0303	-0.0644	-0.0742	-0.0987	-0.1375	-0.1460	-0.1814	-0.2342
	x_{10}	-0	-0	-0	-0	-0	-0	-0	-0	-0
	x_{11}	0.9460	0.8873	0.8441	0.7893	0.7552	0.7387	0.6925	0.6843	0.7243
	x_{12}	0	0	0	0	-0	-0	-0	-0	-0.0561
	x_{13}	0.3768	0.3742	0.3803	0.3748	0.3768	0.3825	0.3758	0.3817	0.4167
	x_{14}	0.0673	0.0844	0.0955	0.1025	0.1128	0.1290	0.1397	0.1534	0.1703
	x_{15}	0	0.0128	0.0383	0.0568	0.0780	0.0864	0.1056	0.1112	0.1011
df		9	11	12	12	12	13	13	13	14
PE		5.6887	5.7025	5.6515	5.6999	5.6979	5.6362	5.7214	5.6662	5.4684

Data no.71: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.03218101.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322	0.0322
λ_1		0.2896	0.1287	0.0751	0.0483	0.0322	0.0215	0.0138	0.0080	0.0036
Predictors	x_1	1.9734	2.0049	2.0137	2.0111	2.0231	2.0316	2.0377	2.0411	2.0426
	x_2	0.0678	0.1917	0.2917	0.3455	0.3963	0.4309	0.4557	0.4729	0.4859
	x_3	0	0	0	0	0.0401	0.0688	0.0892	0.1028	0.1110
	x_4	0.2403	0.2260	0.2347	0.2507	0.2985	0.3320	0.3559	0.3716	0.3836
	x_5	-0.7165	-1.1198	-1.2322	-1.3008	-1.3006	-1.2986	-1.2973	-1.2993	-1.3035
	x_6	0	0.0650	0.1367	0.1556	0.1620	0.1660	0.1689	0.1712	0.1726
	x_7	0.9664	1.1014	1.1867	1.2197	1.2405	1.2544	1.2643	1.2712	1.2770
	x_8	-0	-0.0571	-0.2297	-0.2956	-0.3345	-0.3604	-0.3790	-0.3916	-0.4006
	x_9	-0	-0.1365	-0.2377	-0.2953	-0.3247	-0.3441	-0.3579	-0.3683	-0.3767
	x_{10}	-0	-0	0	0	0	0	0	0	0.0047
	x_{11}	0.9000	0.9540	0.9806	1.0167	1.0272	1.0335	1.0381	1.0418	1.0449
	x_{12}	0	-0	-0.0280	-0.1155	-0.1789	-0.2216	-0.2521	-0.2746	-0.2918
	x_{13}	0.3421	0.4205	0.4549	0.5028	0.5342	0.5553	0.5703	0.5815	0.5896
	x_{14}	0.0366	0.1117	0.1573	0.1913	0.2022	0.2091	0.2140	0.2175	0.2212
	x_{15}	0	0.0325	0.0219	0.0227	0.0130	0.0060	0.0011	0	0
df		8	12	13	13	14	14	14	13	14
PE		5.9863	5.3863	5.1844	5.0732	5.0229	4.9962	4.9805	4.9707	4.9640

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.72: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0112	0.0773	0.1207	0.1877	0.1012	0.3849	0.1628	0.2542	0.2256
λ_1		0.1012	0.3090	0.2816	0.2816	0.1012	0.2566	0.0698	0.0636	0.0251
Predictors	x_1	1.7271	1.4792	1.4322	1.3463	1.5408	1.1643	1.4524	1.3426	1.3834
	x_2	0	0	0	0	0	0	0	0	0
	x_3	0.9999	0.9881	0.9764	0.9317	0.9830	0.8356	0.9477	0.9165	0.9104
	x_4	-0	0	0	0	-0	0	-0	-0	-0.0264
	x_5	-0.1817	-0	-0	-0	-0.1792	-0	-0.2501	-0.2397	-0.3385
	x_6	-0.2766	-0	-0.0215	-0.0342	-0.2627	-0.0672	-0.2916	-0.2757	-0.3292
	x_7	1.5410	1.0380	1.0125	0.9367	1.3151	0.8003	1.2643	1.1409	1.2429
	x_8	-0.0877	-0	-0	-0	-0.0294	-0	-0.0469	-0.0267	-0.0791
	x_9	0.0078	0	0	0	0	0	0	0	0
	x_{10}	0	0	0	0	0.0413	0.0393	0.0763	0.1074	0.1101
	x_{11}	0.6734	0.4329	0.4056	0.3693	0.5228	0.3225	0.4889	0.4352	0.4758
	x_{12}	0.1023	0.0996	0.1293	0.1592	0.1855	0.2002	0.2168	0.2337	0.2454
	x_{13}	0.3909	0.2483	0.2458	0.2419	0.3267	0.2341	0.3249	0.3009	0.3334
	x_{14}	0.4171	0.2975	0.3057	0.3108	0.3942	0.3096	0.4003	0.3803	0.4114
	x_{15}	-0.3889	-0	-0	-0	-0.2348	-0	-0.2116	-0.1402	-0.2057
df		12	7	8	8	12	9	12	12	13
PE		6.0859	7.3847	7.4341	7.6239	6.4478	8.0748	6.5284	6.8330	6.5471

Data no.72: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.07373199.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0737	0.0737	0.0737	0.0737	0.0737	0.0737	0.0737	0.0737	0.0737
λ_1		0.6636	0.2949	0.1720	0.1106	0.0737	0.0492	0.0316	0.0184	0.0082
Predictors	x_1	1.2247	1.4952	1.5766	1.5893	1.5938	1.5923	1.5828	1.5580	1.5388
	x_2	0	0	0	0	0	-0	-0.0146	-0.0784	-0.1278
	x_3	0.7336	1.0007	1.0442	0.9988	0.9682	0.9296	0.8794	0.8160	0.7667
	x_4	0	0	0	-0	-0	-0.0628	-0.1616	-0.2594	-0.3352
	x_5	-0	-0	-0	-0.1516	-0.2640	-0.3788	-0.5017	-0.6286	-0.7269
	x_6	-0	-0	-0.1537	-0.2536	-0.3120	-0.3475	-0.3725	-0.3957	-0.4137
	x_7	0.5933	1.0608	1.2380	1.3524	1.4316	1.4850	1.5271	1.5710	1.6052
	x_8	-0	-0	-0	-0.0288	-0.0856	-0.1140	-0.1289	-0.1392	-0.1471
	x_9	0	0	0	0	0	0.0300	0.0725	0.1177	0.1528
	x_{10}	0	0	0.0120	0.0249	0.0308	0.0316	0.0309	0.0343	0.0370
	x_{11}	0.3139	0.4413	0.4997	0.5485	0.5799	0.6064	0.6307	0.6509	0.6670
	x_{12}	0.1029	0.0967	0.1335	0.1670	0.1778	0.1949	0.2166	0.2351	0.2498
	x_{13}	0.2308	0.2494	0.2825	0.3322	0.3677	0.4028	0.4375	0.4695	0.4943
	x_{14}	0.3075	0.2963	0.3446	0.3932	0.4267	0.4512	0.4743	0.5062	0.5307
	x_{15}	0	-0	-0.1389	-0.2564	-0.3240	-0.3599	-0.3767	-0.3842	-0.3903
df		7	7	10	12	12	14	15	15	15
PE		8.6810	7.3381	6.7486	6.3867	6.2023	6.0810	5.9891	5.9143	5.8610

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.73: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0095	0.0282	0.0303	0.0430	0.0777	0.0882	0.0652	0.0639	0.2087
λ_1		0.0853	0.1127	0.0708	0.0645	0.0777	0.0588	0.0279	0.0160	0.0232
Predictors	x_1	1.6186	1.5719	1.5979	1.5848	1.5280	1.5273	1.5797	1.5773	1.4138
	x_2	0	0	0	0.0327	0.0831	0.1141	0.1071	0.0926	0.2429
	x_3	-0	0	-0	-0	0	0	-0.0024	-0.0604	0.0562
	x_4	0.0476	0.1305	0.0836	0.1209	0.2195	0.2313	0.1610	0.1102	0.3870
	x_5	-1.2317	-1.0576	-1.1876	-1.1404	-0.9604	-0.9691	-1.1349	-1.2195	-0.7162
	x_6	0	0	0.0121	0.0229	0.0326	0.0464	0.0535	0.0647	0.0852
	x_7	0.9481	0.8795	0.9521	0.9477	0.8887	0.9069	0.9803	0.9999	0.8469
	x_8	0.3503	0.3377	0.3572	0.3538	0.3381	0.3415	0.3574	0.3637	0.3329
	x_9	0.2718	0.2065	0.2814	0.2788	0.2255	0.2467	0.3163	0.3357	0.2173
	x_{10}	0.1112	0.1073	0.1349	0.1636	0.1981	0.2289	0.2355	0.2445	0.3261
	x_{11}	1.0692	1.0245	1.0118	0.9766	0.9023	0.8784	0.9165	0.9242	0.7111
	x_{12}	0.2254	0.2085	0.2397	0.2397	0.2259	0.2336	0.2533	0.2685	0.2327
	x_{13}	0.0732	0.0503	0.0889	0.0978	0.0925	0.1099	0.1345	0.1572	0.1321
	x_{14}	0.0874	0.0977	0.1023	0.1103	0.1269	0.1326	0.1249	0.1445	0.1573
	x_{15}	0.3881	0.3586	0.3922	0.3870	0.3598	0.3652	0.3938	0.4036	0.3446
df		12	12	13	14	14	14	15	15	15
PE		3.4472	3.5605	3.4574	3.4733	3.5963	3.5789	3.4599	3.4245	3.8341

Data no.73: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02407508.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0241	0.0241	0.0241	0.0241	0.0241	0.0241	0.0241	0.0241	0.0241
λ_1		0.2167	0.0963	0.0562	0.0361	0.0241	0.0161	0.0103	0.0060	0.0027
Predictors	x_1	1.5101	1.5888	1.6175	1.6147	1.6131	1.6102	1.6027	1.5924	1.5845
	x_2	0	0	0	0	0	0	-0.0165	-0.0628	-0.0987
	x_3	0	0	-0	-0.0783	-0.1238	-0.1629	-0.2144	-0.2708	-0.3144
	x_4	0.2315	0.1014	0.0488	0	-0	-0.0257	-0.1034	-0.1904	-0.2578
	x_5	-0.7371	-1.1294	-1.2634	-1.3759	-1.4253	-1.4774	-1.5629	-1.6584	-1.7322
	x_6	0	0	0.0160	0.0336	0.0421	0.0502	0.0602	0.0674	0.0729
	x_7	0.7115	0.9125	0.9855	1.0174	1.0333	1.0461	1.0598	1.0717	1.0810
	x_8	0.2718	0.3458	0.3638	0.3710	0.3788	0.3810	0.3799	0.3843	0.3879
	x_9	0.0128	0.2402	0.3132	0.3444	0.3641	0.3760	0.3834	0.3910	0.3971
	x_{10}	0.0494	0.1149	0.1398	0.1674	0.1854	0.1978	0.1999	0.1864	0.1759
	x_{11}	1.0440	1.0325	1.0245	1.0269	1.0285	1.0299	1.0355	1.0463	1.0546
	x_{12}	0.1183	0.2202	0.2498	0.2711	0.2842	0.2934	0.3045	0.3203	0.3326
	x_{13}	0	0.0642	0.1023	0.1389	0.1569	0.1738	0.1957	0.2146	0.2291
	x_{14}	0.0856	0.0961	0.0993	0.1267	0.1377	0.1520	0.1770	0.2025	0.2222
	x_{15}	0.2672	0.3734	0.4068	0.4175	0.4196	0.4239	0.4346	0.4480	0.4586
df		11	12	13	13	13	14	15	15	15
PE		3.9627	3.5043	3.4164	3.3649	3.3441	3.3289	3.3094	3.2896	3.2761

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.74: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0144	0.0427	0.0732	0.0862	0.0614	0.0697	0.0820	0.1406	0.1369
λ_1		0.1293	0.1709	0.1709	0.1293	0.0614	0.0465	0.0351	0.0351	0.0152
Predictors	x_1	2.0278	1.8835	1.8126	1.8270	1.9851	1.9837	1.9639	1.8246	1.8521
	x_2	0	0	0	0	-0	-0	-0.0079	0	-0
	x_3	0.7693	0.6682	0.6345	0.6723	0.7151	0.7104	0.6989	0.6736	0.6817
	x_4	-0	-0	-0	-0	-0.2423	-0.2888	-0.3204	-0.2207	-0.3031
	x_5	-0.8809	-0.8128	-0.7943	-0.8672	-1.0633	-1.0967	-1.1251	-1.0480	-1.1170
	x_6	0	0	0	0	0.0463	0.0610	0.0712	0.0426	0.0822
	x_7	1.5650	1.4846	1.4206	1.4254	1.4686	1.4581	1.4423	1.3622	1.3735
	x_8	0.3055	0.2231	0.2236	0.2936	0.4461	0.4743	0.4879	0.4472	0.4854
	x_9	-0.4054	-0.3612	-0.3511	-0.3787	-0.4481	-0.4578	-0.4570	-0.4218	-0.4299
	x_{10}	-0.2896	-0.2103	-0.1998	-0.2617	-0.3806	-0.4005	-0.4131	-0.3776	-0.4129
	x_{11}	0.8217	0.7387	0.6927	0.6881	0.8300	0.8370	0.8259	0.7047	0.7287
	x_{12}	-0.0754	-0	-0	-0	-0.1991	-0.2243	-0.2342	-0.1599	-0.2147
	x_{13}	0.3512	0.3142	0.3286	0.3574	0.5147	0.5433	0.5612	0.5114	0.5523
	x_{14}	-0	0	0	0	0	0	0.0137	0.0450	0.0759
	x_{15}	0.1416	0.1593	0.1897	0.2076	0.2624	0.2819	0.2975	0.3063	0.3194
df		11	10	10	10	13	13	15	14	14
PE		5.4533	5.7473	5.8557	5.7331	5.2073	5.1640	5.1564	5.4130	5.2955

Data no.74: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.04018307.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0402	0.0402	0.0402	0.0402	0.0402	0.0402	0.0402	0.0402	0.0402
λ_1		0.3616	0.1607	0.0938	0.0603	0.0402	0.0268	0.0172	0.0100	0.0045
Predictors	x_1	1.6831	1.9006	1.9997	2.0470	2.0632	2.0693	2.0692	2.0691	2.0692
	x_2	0	0	-0	-0.0023	-0.0869	-0.1463	-0.1915	-0.2255	-0.2518
	x_3	0.4428	0.6845	0.7342	0.7254	0.6463	0.5905	0.5477	0.5155	0.4908
	x_4	0	-0	-0.1339	-0.2938	-0.4706	-0.6001	-0.7035	-0.7811	-0.8409
	x_5	-0.4358	-0.8355	-0.9840	-1.0982	-1.2532	-1.3716	-1.4703	-1.5444	-1.6015
	x_6	0	0	0.0173	0.0680	0.0906	0.1126	0.1349	0.1516	0.1646
	x_7	1.3335	1.4978	1.5034	1.5040	1.5081	1.5096	1.5098	1.5099	1.5101
	x_8	0	0.2415	0.3826	0.4637	0.5007	0.5234	0.5377	0.5485	0.5568
	x_9	-0.2093	-0.3699	-0.4277	-0.4601	-0.4766	-0.4805	-0.4765	-0.4734	-0.4712
	x_{10}	-0	-0.2285	-0.3377	-0.3955	-0.4365	-0.4660	-0.4889	-0.5061	-0.5194
	x_{11}	0.6767	0.7460	0.8216	0.8882	0.9557	0.9944	1.0161	1.0324	1.0451
	x_{12}	0	-0	-0.1447	-0.2446	-0.2903	-0.3262	-0.3567	-0.3795	-0.3973
	x_{13}	0.1979	0.3192	0.4461	0.5385	0.6245	0.6862	0.7344	0.7707	0.7986
	x_{14}	0	0	0	0	0	0.0196	0.0520	0.0763	0.0949
	x_{15}	0.1399	0.1581	0.2125	0.2529	0.2911	0.3157	0.3321	0.3444	0.3538
df		8	10	13	14	14	15	15	15	15
PE		6.7297	5.6977	5.3102	5.1081	4.9752	4.8979	4.8476	4.8142	4.7908

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.75: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0283	0.0190	0.0247	0.1286	0.0917	0.0310	0.1224	0.1094	0.1066
λ_1		0.2551	0.0761	0.0576	0.1930	0.0917	0.0207	0.0525	0.0274	0.0118
Predictors	x_1	2.5981	2.5071	2.4898	2.4012	2.4585	2.4806	2.3984	2.4068	2.4035
	x_2	-0	-0.2059	-0.2346	-0	-0.0568	-0.2966	-0.1127	-0.1738	-0.2075
	x_3	0.5441	0.5484	0.5764	0.5886	0.6349	0.6174	0.6717	0.6895	0.7006
	x_4	0.3877	0.3638	0.3884	0.4496	0.5179	0.4254	0.5475	0.5545	0.5601
	x_5	-1.1778	-1.8696	-1.8912	-1.0763	-1.3707	-1.9595	-1.3758	-1.4842	-1.5364
	x_6	0	0	0	-0	-0	0	-0	-0	-0.0033
	x_7	0.9667	1.0235	1.0280	0.9517	0.9895	1.0451	0.9958	1.0104	1.0178
	x_8	0.1239	0.2365	0.2557	0.2763	0.3282	0.2856	0.3885	0.3984	0.4085
	x_9	-0.0175	-0.2638	-0.2902	-0.0888	-0.2054	-0.3442	-0.2448	-0.2809	-0.3019
	x_{10}	-0	-0	-0	-0.0881	-0.0801	0	-0.1182	-0.1121	-0.1130
	x_{11}	0.4931	0.9247	0.9541	0.4552	0.6400	1.0290	0.6534	0.7235	0.7583
	x_{12}	-0	-0.4498	-0.5003	-0	-0.2297	-0.5970	-0.2746	-0.3569	-0.4005
	x_{13}	0.3120	0.4752	0.4814	0.2882	0.3646	0.5132	0.3717	0.3987	0.4123
	x_{14}	0	-0	-0	0	0	-0.0552	0	0	0
	x_{15}	0.2366	0.4690	0.4856	0.2244	0.3266	0.5422	0.3403	0.3786	0.3983
df		10	12	12	12	13	13	13	13	14
PE		5.8559	4.6772	4.6198	5.9652	5.1960	4.5035	5.1304	4.9420	4.8615

Data no.75: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02523561.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0252	0.0252	0.0252	0.0252	0.0252	0.0252	0.0252	0.0252	0.0252
λ_1		0.2271	0.1009	0.0589	0.0379	0.0252	0.0168	0.0108	0.0063	0.0028
Predictors	x_1	2.6056	2.5356	2.4914	2.4734	2.4777	2.4853	2.4937	2.5002	2.5052
	x_2	-0	-0.1241	-0.2296	-0.2808	-0.3054	-0.3197	-0.3286	-0.3351	-0.3401
	x_3	0.5153	0.5555	0.5773	0.5885	0.5966	0.6037	0.6090	0.6133	0.6165
	x_4	0.3716	0.4002	0.3912	0.3887	0.3953	0.4002	0.4031	0.4056	0.4074
	x_5	-1.2834	-1.7003	-1.8801	-1.9666	-2.0051	-2.0328	-2.0554	-2.0720	-2.0851
	x_6	0	0	0	0	0	0.0082	0.0187	0.0266	0.0327
	x_7	0.9747	1.0094	1.0271	1.0368	1.0456	1.0510	1.0547	1.0574	1.0595
	x_8	0.1478	0.2311	0.2558	0.2679	0.2741	0.2724	0.2666	0.2623	0.2589
	x_9	-0.0465	-0.2210	-0.2878	-0.3211	-0.3412	-0.3568	-0.3698	-0.3796	-0.3872
	x_{10}	-0	-0	-0	-0	0	0.0158	0.0394	0.0571	0.0709
	x_{11}	0.5191	0.8150	0.9468	1.0122	1.0494	1.0780	1.1010	1.1184	1.1319
	x_{12}	-0	-0.3428	-0.4937	-0.5672	-0.6044	-0.6317	-0.6526	-0.6684	-0.6807
	x_{13}	0.3336	0.4349	0.4789	0.5023	0.5212	0.5341	0.5440	0.5513	0.5570
	x_{14}	0	-0	-0	-0.0091	-0.0478	-0.0780	-0.1027	-0.1214	-0.1359
	x_{15}	0.2541	0.4118	0.4820	0.5194	0.5498	0.5732	0.5924	0.6068	0.6181
df		10	12	12	13	13	15	15	15	15
PE		5.7283	4.8643	4.6291	4.5406	4.4918	4.4603	4.4380	4.4231	4.4126

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.76: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0146	0.0396	0.0678	0.1270	0.0994	0.0936	0.1209	0.2072	0.1676
λ_1		0.1314	0.1582	0.1582	0.1906	0.0994	0.0624	0.0518	0.0518	0.0186
Predictors	x_1	1.3266	1.2975	1.2761	1.2187	1.2827	1.3213	1.3109	1.2457	1.3087
	x_2	-0.6380	-0.5752	-0.5403	-0.4404	-0.5669	-0.6212	-0.6071	-0.5435	-0.6139
	x_3	0.3183	0.3322	0.3406	0.3552	0.3435	0.3214	0.3243	0.3326	0.3212
	x_4	-0	0	0	0	0	0	0	0	0
	x_5	-0.6808	-0.5265	-0.4550	-0.2567	-0.5353	-0.6688	-0.6381	-0.4887	-0.6410
	x_6	0.1681	0.1781	0.2056	0.2058	0.2842	0.3176	0.3424	0.3603	0.3907
	x_7	1.7315	1.6307	1.5716	1.4236	1.5910	1.6461	1.6077	1.4672	1.5677
	x_8	-0.4552	-0.4485	-0.4542	-0.4244	-0.5009	-0.5359	-0.5451	-0.5310	-0.5721
	x_9	-0	-0	-0	-0	-0	-0	-0.0068	-0	-0.0344
	x_{10}	0.1767	0.1347	0.1305	0.0895	0.1947	0.2385	0.2420	0.2191	0.2554
	x_{11}	0.9618	0.8192	0.7434	0.6062	0.7545	0.8036	0.7531	0.6159	0.6867
	x_{12}	-0.0869	-0.0063	-0	-0	-0.0871	-0.1647	-0.1615	-0.1076	-0.2023
	x_{13}	0	0	0	0	0.0139	0.0719	0.0861	0.0813	0.1233
	x_{14}	0.0864	0.0518	0.0770	0.0849	0.1908	0.2503	0.2637	0.2483	0.2887
	x_{15}	0	0	0	0	0	0	0	0.0073	0.0487
df		11	11	10	10	12	12	13	13	14
PE		4.8011	5.0786	5.1923	5.5844	4.9723	4.7549	4.8073	5.1210	4.8221

Data no.76: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.06184907.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0618	0.0618	0.0618	0.0618	0.0618	0.0618	0.0618	0.0618	0.0618
λ_1		0.5566	0.2474	0.1443	0.0928	0.0618	0.0412	0.0265	0.0155	0.0069
Predictors	x_1	1.0074	1.252	1.2865	1.3142	1.3450	1.3668	1.3839	1.3803	1.3772
	x_2	-0.0440	-0.4173	-0.5613	-0.6131	-0.6572	-0.6836	-0.7133	-0.7513	-0.7812
	x_3	0.2318	0.3659	0.3387	0.3331	0.3098	0.2974	0.2825	0.2522	0.2283
	x_4	0	0	0	0	-0	-0	-0.0064	-0.0556	-0.0942
	x_5	-0	-0.1978	-0.5048	-0.6432	-0.7570	-0.8309	-0.8936	-0.9759	-1.0405
	x_6	0	0.1128	0.2141	0.2639	0.2960	0.3208	0.3378	0.3503	0.3600
	x_7	1.0961	1.4508	1.6034	1.6754	1.7144	1.7433	1.7628	1.7779	1.7895
	x_8	-0.0922	-0.3759	-0.4637	-0.5023	-0.5335	-0.5552	-0.5735	-0.5874	-0.5983
	x_9	0	-0	-0	-0	-0.0065	-0.0328	-0.0505	-0.0649	-0.0761
	x_{10}	0	0.0265	0.1485	0.2127	0.2504	0.2759	0.2867	0.2931	0.2981
	x_{11}	0.4191	0.6538	0.7791	0.8589	0.8913	0.9119	0.9176	0.9273	0.9347
	x_{12}	0	-0	-0.0238	-0.1299	-0.1994	-0.2470	-0.2987	-0.3294	-0.3531
	x_{13}	0	0	0	0.0163	0.0726	0.1089	0.1383	0.1645	0.1850
	x_{14}	0	0	0.0967	0.1953	0.2524	0.2903	0.3110	0.3301	0.3451
	x_{15}	0	0	0	0	0	0.0003	0.0391	0.0768	0.1061
df		6	9	11	12	13	14	15	15	15
PE		7.4070	5.6756	5.0822	4.7863	4.6309	4.5432	4.4834	4.4373	4.4056

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.77: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0113	0.0404	0.0575	0.0353	0.0439	0.0956	0.0934	0.0835	0.0892
λ_1		0.1015	0.1616	0.1342	0.0529	0.0439	0.0637	0.0400	0.0209	0.0099
Predictors	x_1	2.1066	1.9224	1.9223	2.1053	2.0873	1.9309	1.9615	2.0031	1.9984
	x_2	0.1266	0.1342	0.1616	0.1684	0.1680	0.2073	0.2041	0.1866	0.1838
	x_3	0.4127	0.4451	0.4686	0.5063	0.5111	0.5359	0.5517	0.5380	0.5366
	x_4	-0.1779	-0	-0	-0.2567	-0.2739	-0.0705	-0.1524	-0.2578	-0.2894
	x_5	-1.2350	-0.8570	-0.8859	-1.2478	-1.2492	-0.9630	-1.0501	-1.1712	-1.1957
	x_6	-0.6058	-0.5255	-0.5348	-0.6230	-0.6223	-0.5608	-0.5822	-0.6069	-0.6116
	x_7	1.2320	1.0630	1.0615	1.2557	1.2485	1.0866	1.1267	1.1860	1.1921
	x_8	0.1818	0	0.0008	0.2089	0.2068	0.0631	0.1084	0.1669	0.1784
	x_9	0.2010	0.1231	0.1579	0.2788	0.2918	0.2543	0.2893	0.3179	0.3316
	x_{10}	0.4811	0.3586	0.3889	0.5762	0.5886	0.4929	0.5476	0.5984	0.6157
	x_{11}	1.1041	0.9019	0.8785	1.1011	1.0827	0.8740	0.9177	0.9801	0.9806
	x_{12}	-0.0752	-0	-0	-0.2371	-0.2566	-0.1115	-0.1847	-0.2641	-0.2895
	x_{13}	0.3078	0.2461	0.2732	0.4357	0.4489	0.3672	0.4127	0.4559	0.4701
	x_{14}	0	0	0	0	0.0136	0	0.0005	0.0363	0.0528
	x_{15}	0	0	0	0.0296	0.0518	0.0492	0.0847	0.1078	0.1257
df		13	10	11	14	15	14	15	15	15
PE		4.7312	5.3287	5.2833	4.5423	4.5387	5.0329	4.8472	4.6564	4.6306

Data no.77: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.03977744.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0398	0.0398	0.0398	0.0398	0.0398	0.0398	0.0398	0.0398	0.0398
λ_1		0.3580	0.1591	0.0928	0.0597	0.0398	0.0265	0.0170	0.0099	0.0044
Predictors	x_1	1.6176	1.9279	2.0441	2.0867	2.0999	2.1065	2.1111	2.1146	2.1174
	x_2	0	0.1364	0.1766	0.1782	0.1544	0.1330	0.1174	0.1058	0.0970
	x_3	0.2550	0.4475	0.4815	0.5092	0.4961	0.4762	0.4613	0.4504	0.4424
	x_4	0	-0	-0.0905	-0.2099	-0.3185	-0.4010	-0.4604	-0.5047	-0.5388
	x_5	-0.3210	-0.8655	-1.0765	-1.1925	-1.3021	-1.3857	-1.4463	-1.4914	-1.5259
	x_6	-0.3019	-0.5287	-0.5840	-0.6128	-0.6301	-0.6418	-0.6502	-0.6565	-0.6614
	x_7	0.8264	1.0668	1.1695	1.2310	1.2708	1.3003	1.3215	1.3373	1.3496
	x_8	0	0	0.1185	0.1834	0.2290	0.2625	0.2864	0.3044	0.3184
	x_9	0	0.1258	0.2158	0.2669	0.2987	0.3191	0.3335	0.3444	0.3529
	x_{10}	0.1307	0.3618	0.4749	0.5538	0.6032	0.6352	0.6579	0.6750	0.6884
	x_{11}	0.8020	0.9044	0.9952	1.0670	1.1110	1.1417	1.1636	1.1800	1.1928
	x_{12}	0	-0	-0.0837	-0.2037	-0.2819	-0.3359	-0.3741	-0.4028	-0.4255
	x_{13}	0.1577	0.2468	0.3237	0.4132	0.4647	0.4986	0.5232	0.5414	0.5555
	x_{14}	0	0	0	0	0.0279	0.0604	0.0839	0.1014	0.1149
	x_{15}	0	0	0	0.0185	0.0578	0.0836	0.1018	0.1155	0.1263
df		8	10	13	14	15	15	15	15	15
PE		6.6924	5.3133	4.8691	4.6169	4.4840	4.4061	4.3578	4.3257	4.3031

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.78: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0190	0.0323	0.0504	0.0715	0.0890	0.1216	0.1304	0.1404	0.1501
λ_1		0.1707	0.1291	0.1176	0.1072	0.0890	0.0811	0.0559	0.0351	0.0167
Predictors	x_1	2.5205	2.4974	2.4589	2.4157	2.3851	2.3234	2.3173	2.3057	2.2943
	x_2	0.6325	0.6459	0.6569	0.6640	0.6717	0.6799	0.6840	0.6857	0.6875
	x_3	0.8524	0.8465	0.8408	0.8361	0.8362	0.8301	0.8290	0.8314	0.8345
	x_4	0.7325	0.7118	0.6985	0.6847	0.6758	0.6612	0.6514	0.6515	0.6533
	x_5	-0.1945	-0.2797	-0.2942	-0.3054	-0.3287	-0.3250	-0.3706	-0.4018	-0.4269
	x_6	-0.1934	-0.2335	-0.2481	-0.2658	-0.2887	-0.3002	-0.3274	-0.3566	-0.3827
	x_7	1.0139	1.0246	1.0210	1.0143	1.0116	0.9982	0.9995	0.9961	0.9925
	x_8	0.2387	0.2604	0.2661	0.2727	0.2841	0.2885	0.3018	0.3131	0.3232
	x_9	0	0	0.0031	0.0354	0.0708	0.0967	0.1285	0.1610	0.1898
	x_{10}	-0	-0	-0	-0	-0	-0	-0	-0.0320	-0.0644
	x_{11}	0.7081	0.7125	0.6939	0.6670	0.6490	0.6194	0.6120	0.6072	0.6031
	x_{12}	0	0	0	0	0	0	0.0191	0.0355	0.0496
	x_{13}	0.1604	0.2130	0.2398	0.2702	0.2991	0.3205	0.3409	0.3544	0.3652
	x_{14}	0	-0	-0	0	0	0	0	0	-0
	x_{15}	0	0	0	0	0	0	0	0	0
df		10	10	11	11	11	11	12	13	13
PE		4.2059	4.1301	4.1383	4.1530	4.1478	4.1999	4.1700	4.1495	4.1365

Data no.78: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.0294623.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0295	0.0295	0.0295	0.0295	0.0295	0.0295	0.0295	0.0295	0.0295
λ_1		0.2652	0.1178	0.0687	0.0442	0.0295	0.0196	0.0126	0.0074	0.0033
Predictors	x_1	2.4698	2.5067	2.5167	2.5191	2.5220	2.5236	2.5248	2.5256	2.5263
	x_2	0.6154	0.6468	0.6489	0.6463	0.6567	0.6657	0.6721	0.6769	0.6806
	x_3	0.8383	0.8483	0.8525	0.8649	0.9111	0.9427	0.9651	0.9821	0.9952
	x_4	0.7367	0.7117	0.7019	0.7169	0.7727	0.8119	0.8398	0.8608	0.8771
	x_5	-0.0033	-0.3041	-0.4092	-0.4513	-0.4186	-0.3947	-0.3777	-0.3648	-0.3549
	x_6	-0.1144	-0.2426	-0.2975	-0.3405	-0.3615	-0.3746	-0.3841	-0.3911	-0.3966
	x_7	0.9731	1.0302	1.0466	1.0483	1.0600	1.0678	1.0734	1.0776	1.0809
	x_8	0.1867	0.2666	0.2961	0.3129	0.3437	0.3640	0.3785	0.3894	0.3978
	x_9	0	0	0.0557	0.1048	0.1378	0.1586	0.1735	0.1846	0.1933
	x_{10}	-0	-0	-0.0082	-0.0664	-0.0945	-0.1118	-0.1241	-0.1334	-0.1406
	x_{11}	0.6447	0.7224	0.7342	0.7438	0.7585	0.7685	0.7757	0.7811	0.7852
	x_{12}	0	0	0	0	0	0	0	0	0
	x_{13}	0.0955	0.2196	0.2758	0.2984	0.3157	0.3278	0.3364	0.3429	0.3479
	x_{14}	0	-0	-0	-0	-0.0154	-0.0303	-0.0408	-0.0488	-0.0550
	x_{15}	0	0	-0	-0.0018	-0.0599	-0.0978	-0.1248	-0.1452	-0.1609
df		10	10	12	13	14	14	14	14	14
PE		4.5392	4.1024	4.0138	3.9720	3.9326	3.9114	3.8991	3.8914	3.8863

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.79: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0220	0.0494	0.0772	0.0908	0.1495	0.2043	0.0864	0.0641	0.1091
λ_1		0.1976	0.1976	0.1800	0.1362	0.1495	0.1362	0.0370	0.0160	0.0121
Predictors	x_1	1.8047	1.7659	1.7392	1.7414	1.6632	1.6060	1.7503	1.7416	1.7274
	x_2	0	0	0	0.0462	0.0338	0.0570	0.1079	0.0513	0.1291
	x_3	-0	-0	-0	-0	-0	-0	-0.1747	-0.3489	-0.2366
	x_4	-0	-0	-0	-0	-0	-0	-0.1726	-0.3475	-0.2122
	x_5	-0.0032	-0	-0	-0.0846	-0.0179	-0.0146	-0.4598	-0.7052	-0.5145
	x_6	-0	-0	-0	-0	-0	-0	-0.0245	-0.0612	-0.0363
	x_7	0.3020	0.3031	0.3149	0.3564	0.3256	0.3201	0.4657	0.4740	0.4818
	x_8	0	0	0	0	0	0	0.0190	0.0902	0.0583
	x_9	0	0	0	0	0	0	0.1128	0.1515	0.1230
	x_{10}	0.1788	0.1879	0.2053	0.2393	0.2387	0.2529	0.3137	0.3100	0.3263
	x_{11}	0.8101	0.7625	0.7376	0.7618	0.6826	0.6457	0.8995	0.9973	0.8817
	x_{12}	-0	-0	-0	-0	-0	-0	-0.0455	-0.0654	-0.0588
	x_{13}	0.5441	0.5470	0.5570	0.5826	0.5626	0.5559	0.7214	0.8074	0.7370
	x_{14}	-0.2173	-0.1723	-0.1585	-0.1943	-0.1167	-0.0864	-0.2398	-0.2370	-0.2350
	x_{15}	0	0	0	0	0	0.0010	0.0251	0.0501	0.0817
df		7	6	6	8	8	9	15	15	15
PE		4.5119	4.6294	4.6741	4.5163	4.8278	4.9858	3.9739	3.3741	3.9586

Data no.79: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.01906026.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0191	0.0191	0.0191	0.0191	0.0191	0.0191	0.0191	0.0191	0.0191
λ_1		0.1715	0.0762	0.0445	0.0286	0.0191	0.0127	0.0082	0.0048	0.0021
Predictors	x_1	1.8155	1.7819	1.7416	1.7211	1.6894	1.6660	1.6493	1.6365	1.6267
	x_2	0	0	0	-0	-0.0735	-0.1271	-0.1652	-0.1942	-0.21654
	x_3	-0	-0.1928	-0.3993	-0.50374	-0.6081	-0.6873	-0.7438	-0.7868	-0.8199
	x_4	-0	-0.2428	-0.4510	-0.5566	-0.6798	-0.7669	-0.8290	-0.8764	-0.9128
	x_5	-0.0680	-0.5431	-0.8543	-1.0122	-1.1563	-1.2548	-1.3251	-1.3788	-1.4200
	x_6	-0	-0.0185	-0.0702	-0.0965	-0.1101	-0.1166	-0.1212	-0.1248	-0.1275
	x_7	0.3284	0.4386	0.4285	0.4218	0.4160	0.4173	0.4181	0.41845	0.4189
	x_8	0	0	0.0889	0.1383	0.1739	0.1992	0.2173	0.2309	0.2414
	x_9	0	0.1002	0.1640	0.1965	0.2167	0.2249	0.2307	0.2352	0.2386
	x_{10}	0.1927	0.2667	0.2889	0.2993	0.2991	0.2965	0.2946	0.2933	0.2922
	x_{11}	0.8512	1.0250	1.1134	1.1583	1.1970	1.2179	1.2329	1.2443	1.2531
	x_{12}	-0	-0.0339	-0.0339	-0.0340	-0.0268	-0.0254	-0.0242	-0.0229	-0.0222
	x_{13}	0.5616	0.7250	0.8221	0.8709	0.9197	0.9529	0.9764	0.9943	1.0080
	x_{14}	-0.2597	-0.2830	-0.2466	-0.2278	-0.1956	-0.1770	-0.1639	-0.1539	-0.1463
	x_{15}	0	0	0	0	0	0.0184	0.0315	0.0413	0.0490
df		7	12	13	13	14	15	15	15	15
PE		4.3750	3.8475	3.6538	3.5814	3.5281	3.4959	3.4755	3.4615	3.4515

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.80: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0142	0.0183	0.0377	0.0852	0.0881	0.1746	0.1176	0.1837	0.3431
λ_1		0.1278	0.0731	0.0881	0.1278	0.0881	0.1164	0.0504	0.0459	0.0381
Predictors	x_1	1.9709	2.0013	1.9418	1.8056	1.8323	1.6546	1.8062	1.6965	1.4847
	x_2	0	-0	-0	0	-0	0	-0	0	0.0201
	x_3	0.5288	0.5681	0.5222	0.4178	0.4569	0.3480	0.4675	0.4171	0.3286
	x_4	0.5579	0.5160	0.5106	0.4922	0.4860	0.4581	0.4762	0.4661	0.4472
	x_5	-0.1662	-0.3238	-0.2726	-0.1589	-0.2451	-0.1477	-0.3047	-0.2672	-0.1961
	x_6	0.4100	0.5050	0.4595	0.3541	0.4180	0.3182	0.4570	0.4176	0.3407
	x_7	1.2874	1.3248	1.2810	1.1822	1.2082	1.0853	1.1976	1.1236	0.9823
	x_8	0.2520	0.2838	0.2845	0.2776	0.3003	0.2982	0.3270	0.3372	0.3405
	x_9	-0.2170	-0.2680	-0.2475	-0.1985	-0.2319	-0.1861	-0.2548	-0.2391	-0.2102
	x_{10}	-0.4195	-0.4622	-0.4441	-0.4004	-0.4305	-0.3903	-0.4520	-0.4396	-0.4217
	x_{11}	0.2298	0.2507	0.2494	0.2410	0.2537	0.2449	0.2656	0.2645	0.2565
	x_{12}	0.2890	0.2998	0.2884	0.2669	0.2755	0.2576	0.2790	0.2715	0.2515
	x_{13}	0.3934	0.4465	0.4177	0.3623	0.3885	0.3400	0.3982	0.3726	0.3286
	x_{14}	0	-0	-0	0	0	0	-0	0	0.0217
	x_{15}	0.1428	0.1668	0.1799	0.2006	0.2106	0.2314	0.2294	0.2453	0.2525
df		13	13	13	13	13	13	13	13	15
PE		4.7365	4.5856	4.6593	4.9127	4.7738	5.1203	4.7308	4.8807	5.2973

Data no.80: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03418234.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0342	0.0342	0.0342	0.0342	0.0342	0.0342	0.0342	0.0342	0.0342
λ_1		0.3076	0.1367	0.0798	0.0513	0.0342	0.0228	0.0146	0.0085	0.0038
Predictors	x_1	1.7151	1.9121	1.9570	1.9788	1.9851	1.9902	1.9938	1.9967	1.9988
	x_2	0	0	-0	-0.0060	-0.0844	-0.1296	-0.1622	-0.1862	-0.2049
	x_3	0.1832	0.4813	0.5359	0.5601	0.5388	0.5394	0.5395	0.5402	0.5405
	x_4	0.3760	0.5364	0.5096	0.4917	0.4287	0.3976	0.3749	0.3586	0.3458
	x_5	-0	-0.1447	-0.2960	-0.3765	-0.4810	-0.5338	-0.5721	-0.5999	-0.6218
	x_6	0.0630	0.3765	0.4774	0.5279	0.5582	0.5799	0.5955	0.6072	0.6163
	x_7	1.1149	1.2480	1.2933	1.3159	1.3273	1.3338	1.3385	1.3420	1.3447
	x_8	0.1463	0.2563	0.2877	0.3035	0.3171	0.3270	0.3340	0.3394	0.3435
	x_9	-0.0449	-0.2036	-0.2563	-0.2819	-0.2861	-0.2892	-0.2913	-0.2930	-0.2943
	x_{10}	-0.2698	-0.4062	-0.4518	-0.4732	-0.4662	-0.4551	-0.4473	-0.4413	-0.4367
	x_{11}	0.2051	0.2323	0.2521	0.2621	0.2686	0.2635	0.2599	0.2572	0.2552
	x_{12}	0.2493	0.2782	0.2916	0.3000	0.3276	0.3533	0.3718	0.3855	0.3962
	x_{13}	0.3081	0.3756	0.4272	0.4541	0.4840	0.5105	0.5296	0.5437	0.5547
	x_{14}	0	0	-0	-0	-0.0058	-0.0342	-0.0541	-0.0695	-0.0812
	x_{15}	0.1854	0.1622	0.1792	0.1893	0.2178	0.2413	0.2580	0.2705	0.2802
df		12	13	13	14	15	15	15	15	15
PE		5.6883	4.8160	4.6298	4.5647	4.5183	4.4886	4.4709	4.4598	4.4522

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.81: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0138	0.0373	0.0640	0.0753	0.0778	0.1281	0.1374	0.1782	0.1581
λ_1		0.1239	0.1493	0.1493	0.1129	0.0778	0.0854	0.0589	0.0445	0.0176
Predictors	x_1	1.2452	1.2285	1.2101	1.2057	1.2075	1.1722	1.1691	1.1439	1.1522
	x_2	0	0	0	0	0	0	0	0.0257	0.0447
	x_3	0.7491	0.7494	0.7505	0.7566	0.7626	0.7582	0.7621	0.7589	0.7851
	x_4	0	0	0	0	0	0	0	0	-0
	x_5	-0.4945	-0.4013	-0.3700	-0.4374	-0.5110	-0.4338	-0.4766	-0.4532	-0.5081
	x_6	0	0	0	0	0	0	0	0.0074	0.0071
	x_7	1.2268	1.1489	1.1179	1.1707	1.2310	1.1522	1.1853	1.1618	1.2326
	x_8	-0.4453	-0.4019	-0.3847	-0.4122	-0.4433	-0.4076	-0.4254	-0.4203	-0.4597
	x_9	0.0145	0.0183	0.0394	0.0713	0.0959	0.1185	0.1389	0.1688	0.1836
	x_{10}	0.5990	0.5804	0.5708	0.5785	0.5883	0.5791	0.5863	0.5901	0.6142
	x_{11}	0.9238	0.8609	0.8069	0.7928	0.7939	0.7240	0.7167	0.6682	0.6983
	x_{12}	-0	-0	-0	-0	-0	-0	-0	-0	-0.0523
	x_{13}	0.2616	0.2655	0.2769	0.2935	0.3065	0.3130	0.3213	0.3187	0.3306
	x_{14}	0	0	0	0	0	0	0.0048	0.0313	0.0268
	x_{15}	0.1430	0.1596	0.1856	0.2116	0.2298	0.2511	0.2642	0.2726	0.2980
df		10	10	10	10	10	10	11	13	14
PE		3.3175	3.4279	3.4910	3.4198	3.3466	3.4783	3.4407	3.5143	3.3821

Data no.81: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.02585734.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259	0.0259
λ_1		0.2327	0.1034	0.0603	0.0388	0.0259	0.0172	0.0111	0.0065	0.0029
Predictors	x_1	1.2100	1.2394	1.2242	1.2134	1.2064	1.1900	1.1730	1.1604	1.1503
	x_2	0	0	-0	-0	-0	-0.0061	-0.0304	-0.0483	-0.0628
	x_3	0.7423	0.7544	0.7868	0.8078	0.8172	0.7990	0.7747	0.7570	0.7425
	x_4	0	0	-0	-0	-0.0069	-0.0529	-0.1011	-0.1366	-0.1652
	x_5	-0.2149	-0.5264	-0.6148	-0.6562	-0.6860	-0.7378	-0.7943	-0.8359	-0.8694
	x_6	0	0	-0	-0	-0.0133	-0.0230	-0.0284	-0.0325	-0.0357
	x_7	1.0038	1.2524	1.3414	1.3866	1.4129	1.4323	1.4450	1.4546	1.4621
	x_8	-0.3163	-0.4562	-0.5033	-0.5274	-0.5468	-0.5613	-0.5672	-0.5717	-0.5751
	x_9	0	0.0401	0.0714	0.0873	0.0966	0.1082	0.1138	0.1179	0.1212
	x_{10}	0.5769	0.5972	0.6203	0.6341	0.6453	0.6542	0.6563	0.6580	0.6591
	x_{11}	0.8715	0.8941	0.9114	0.9221	0.9363	0.9487	0.9612	0.9704	0.9776
	x_{12}	0	-0	-0.0774	-0.1300	-0.1605	-0.1736	-0.1802	-0.1853	-0.1891
	x_{13}	0.2290	0.2770	0.2995	0.3118	0.3267	0.3443	0.3613	0.3738	0.3838
	x_{14}	0	0	-0	-0	-0.0226	-0.0334	-0.0365	-0.0391	-0.0408
	x_{15}	0.0945	0.1710	0.2336	0.2720	0.3010	0.3293	0.3539	0.3724	0.3870
df		9	10	11	11	14	15	15	15	15
PE		3.6776	3.2949	3.1756	3.1276	3.0959	3.0752	3.0610	3.0519	3.0458

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.82: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0262	0.0489	0.1465	0.1083	0.1348	0.1843	0.1800	0.3086	0.2072
λ_1		0.2356	0.1956	0.3418	0.1624	0.1348	0.1228	0.0771	0.0771	0.0230
Predictors	x_1	2.1226	2.1028	1.8290	2.0017	1.9651	1.8867	1.9147	1.7291	1.8717
	x_2	0.3534	0.3940	0.1810	0.4001	0.4023	0.3808	0.4177	0.3547	0.4245
	x_3	0.1489	0.1454	0.1913	0.1471	0.1240	0.1236	0.0937	0.1314	0.0216
	x_4	0	0	0	0	-0	-0	-0	-0	-0.1589
	x_5	-0.0831	-0.1372	-0	-0.1963	-0.2701	-0.2997	-0.3920	-0.3596	-0.5541
	x_6	-0.3216	-0.3596	-0.1622	-0.3775	-0.4001	-0.3932	-0.4448	-0.3915	-0.4980
	x_7	1.0592	1.0684	0.9112	1.0442	1.0364	1.0059	1.0308	0.9476	1.0468
	x_8	0.0955	0.1469	0.0016	0.2100	0.2533	0.2734	0.3294	0.3217	0.4056
	x_9	0.1265	0.1688	0	0.2001	0.2369	0.2424	0.3046	0.2660	0.3711
	x_{10}	0	0	-0	-0	-0	-0	-0	-0	-0.0106
	x_{11}	0.3106	0.3317	0.3190	0.3250	0.3107	0.2981	0.2929	0.2782	0.3046
	x_{12}	0	0	0	0.0152	0.0389	0.0579	0.0764	0.0951	0.1209
	x_{13}	0.5481	0.5411	0.4963	0.5021	0.4795	0.4517	0.4477	0.4026	0.4477
	x_{14}	0	0	0	0.0074	0.0386	0.0629	0.0829	0.1038	0.1192
	x_{15}	0	0	0	0.0578	0.0891	0.1141	0.1373	0.1598	0.1867
df		10	10	8	13	13	13	13	13	15
PE		6.6301	6.5301	7.4082	6.5447	6.5149	6.6020	6.4598	6.7984	6.3501

Data no.82: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03350562.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0335	0.0335	0.0335	0.0335	0.0335	0.0335	0.0335	0.0335	0.0335
λ_1		0.3016	0.1340	0.0782	0.0503	0.0335	0.0223	0.0144	0.0084	0.0037
Predictors	x_1	2.0611	2.1716	2.1791	2.1679	2.1468	2.1044	2.0710	2.0447	2.0242
	x_2	0.2605	0.4810	0.5176	0.5250	0.5056	0.4468	0.4055	0.3747	0.3508
	x_3	0.1752	0.1066	0.0069	-0	-0.0678	-0.2400	-0.3651	-0.4599	-0.5334
	x_4	0	-0	-0.0653	-0.1957	-0.3199	-0.4936	-0.6213	-0.7188	-0.7946
	x_5	-0	-0.2411	-0.4382	-0.5475	-0.6842	-0.9144	-1.0795	-1.2040	-1.3005
	x_6	-0.2331	-0.4480	-0.5340	-0.5669	-0.5914	-0.6181	-0.6390	-0.6556	-0.6683
	x_7	1.0106	1.1192	1.1486	1.1689	1.1792	1.1830	1.1878	1.1922	1.1956
	x_8	0.0149	0.2263	0.3264	0.3819	0.4176	0.4467	0.4694	0.4874	0.5012
	x_9	0.0339	0.2609	0.3617	0.3980	0.4295	0.4699	0.5005	0.5243	0.5426
	x_{10}	0	0	0	-0	-0	-0	-0.0120	-0.0259	-0.0368
	x_{11}	0.2945	0.3413	0.3335	0.3473	0.3630	0.3892	0.4077	0.4215	0.4325
	x_{12}	0	0	0.0470	0.0766	0.1077	0.1552	0.1897	0.2157	0.2361
	x_{13}	0.5380	0.5537	0.5509	0.5520	0.5637	0.5928	0.6137	0.6295	0.6420
	x_{14}	0	0	0.0303	0.0387	0.0658	0.1258	0.1682	0.1999	0.2242
	x_{15}	0	0.0181	0.0824	0.1211	0.1498	0.1796	0.2014	0.2180	0.2305
df		9	11	14	14	14	14	15	15	15
PE		6.9439	6.3086	6.1533	6.0801	6.0278	5.9624	5.9211	5.8935	5.8742

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.83: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0066	0.0136	0.0255	0.0300	0.0494	0.0615	0.0660	0.1362	0.1003
λ_1		0.0595	0.0542	0.0595	0.0450	0.0494	0.0410	0.0283	0.0340	0.0111
Predictors	x_1	2.5872	2.5598	2.4967	2.4994	2.4191	2.3892	2.3912	2.1917	2.3086
	x_2	-0	-0	-0	-0	-0	-0	-0	-0	-0.0055
	x_3	0.6178	0.6007	0.5609	0.5705	0.5330	0.5283	0.5375	0.4864	0.5235
	x_4	0.6722	0.6408	0.5941	0.5885	0.5469	0.5330	0.5313	0.4873	0.5050
	x_5	-0.2891	-0.3270	-0.3664	-0.3881	-0.4186	-0.4378	-0.4530	-0.4544	-0.4841
	x_6	0.2417	0.2453	0.2243	0.2478	0.2248	0.2309	0.2491	0.2005	0.2529
	x_7	1.3671	1.3687	1.3457	1.3639	1.3311	1.3282	1.3421	1.2431	1.3216
	x_8	0.2764	0.2822	0.2825	0.2921	0.2927	0.2989	0.3068	0.3044	0.3175
	x_9	-0.4359	-0.4365	-0.4041	-0.4322	-0.3968	-0.4017	-0.4239	-0.3566	-0.4235
	x_{10}	-0.1452	-0.1529	-0.1418	-0.1632	-0.1495	-0.1574	-0.1741	-0.1396	-0.1827
	x_{11}	0.5165	0.5136	0.5072	0.5049	0.4919	0.4844	0.4829	0.4377	0.4646
	x_{12}	-0.6979	-0.6781	-0.5978	-0.6301	-0.5376	-0.5235	-0.5483	-0.3606	-0.4970
	x_{13}	0.3517	0.3584	0.3611	0.3687	0.3693	0.3731	0.3788	0.3649	0.3831
	x_{14}	0.2440	0.2525	0.2484	0.2644	0.2583	0.2649	0.2763	0.2588	0.2833
	x_{15}	0.6456	0.6435	0.6092	0.6308	0.5878	0.5837	0.5991	0.5001	0.5767
df		14	14	14	14	14	14	14	14	15
PE		3.8566	3.8629	3.9312	3.8908	3.9861	3.9996	3.9656	4.2782	4.0322

Data no.83: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.02583583.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0258	0.0258	0.0258	0.0258	0.0258	0.0258	0.0258	0.0258	0.0258
λ_1		0.2325	0.1033	0.0603	0.0388	0.0258	0.0172	0.0111	0.0065	0.0029
Predictors	x_1	2.2420	2.4288	2.4940	2.5267	2.5462	2.5576	2.5650	2.5705	2.5747
	x_2	-0	-0	-0	-0	-0	-0.0078	-0.0171	-0.0239	-0.0296
	x_3	0.3590	0.5004	0.5589	0.5883	0.6059	0.6109	0.6116	0.6120	0.6122
	x_4	0.4825	0.5706	0.5926	0.6037	0.6104	0.6078	0.6031	0.5995	0.5965
	x_5	-0.2540	-0.3388	-0.3670	-0.3810	-0.3894	-0.4029	-0.4161	-0.4260	-0.4342
	x_6	0	0.1411	0.2226	0.2634	0.2878	0.3033	0.3141	0.3222	0.3286
	x_7	1.0801	1.2724	1.3440	1.3797	1.4012	1.4167	1.4283	1.4370	1.4439
	x_8	0.1685	0.2568	0.2821	0.2948	0.3024	0.3066	0.3090	0.3108	0.3122
	x_9	-0.0137	-0.2976	-0.4017	-0.4538	-0.4850	-0.5045	-0.5178	-0.5278	-0.5356
	x_{10}	-0	-0.0726	-0.1405	-0.1744	-0.1948	-0.2080	-0.2172	-0.2241	-0.2295
	x_{11}	0.4694	0.5061	0.5070	0.5074	0.5077	0.5105	0.5136	0.5159	0.5177
	x_{12}	-0	-0.4349	-0.5934	-0.6726	-0.7202	-0.7499	-0.7704	-0.7857	-0.7975
	x_{13}	0.2913	0.3422	0.3609	0.3703	0.3759	0.3815	0.3862	0.3898	0.3925
	x_{14}	0.0907	0.2001	0.2476	0.2713	0.2855	0.2950	0.3017	0.3066	0.3105
	x_{15}	0.2842	0.5170	0.6070	0.6520	0.6789	0.6984	0.7132	0.7244	0.7336
df		11	14	14	14	14	15	15	15	15
PE		5.2112	4.1682	3.9360	3.8534	3.8145	3.7931	3.7799	3.7713	3.7654

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.84: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0047	0.0050	0.0072	0.0178	0.0153	0.0209	0.0567	0.0241	0.0947
λ_1		0.0425	0.0202	0.0167	0.0267	0.0153	0.0139	0.0243	0.0060	0.0105
Predictors	x_1	3.0563	3.0679	3.0563	2.9925	3.0114	2.9848	2.8283	2.9728	2.7213
	x_2	0.8668	0.9371	0.9284	0.8198	0.8698	0.8390	0.6384	0.8400	0.5828
	x_3	1.8224	1.8973	1.8780	1.7213	1.7831	1.7321	1.4320	1.7225	1.3204
	x_4	1.3969	1.5014	1.4752	1.2651	1.3479	1.2817	0.90554	1.2700	0.7978
	x_5	-0.0461	-0.0229	-0.0586	-0.2171	-0.1791	-0.2413	-0.5374	-0.2767	-0.6363
	x_6	0.4284	0.4686	0.4755	0.4601	0.4797	0.4824	0.4577	0.4955	0.4697
	x_7	1.2120	1.2417	1.2422	1.2097	1.2291	1.2209	1.1477	1.2250	1.1123
	x_8	-0.0597	-0.0706	-0.0744	-0.0769	-0.0814	-0.0852	-0.0817	-0.0908	-0.0848
	x_9	-0.0947	-0.1197	-0.1209	-0.0996	-0.1142	-0.1111	-0.0813	-0.1170	-0.0860
	x_{10}	0.3338	0.4036	0.4105	0.3621	0.4008	0.3954	0.2921	0.4117	0.2746
	x_{11}	0.3629	0.3814	0.3949	0.4239	0.4272	0.4427	0.4544	0.4567	0.4501
	x_{12}	0	-0	-0.0002	0	0	0	0.0892	0.0062	0.1340
	x_{13}	0.3022	0.3407	0.3534	0.3599	0.3750	0.3861	0.3840	0.4036	0.3916
	x_{14}	-0.5134	-0.5802	-0.5792	-0.5021	-0.5451	-0.5260	-0.3904	-0.5349	-0.3502
	x_{15}	-0.2701	-0.3359	-0.3381	-0.2766	-0.3157	-0.3040	-0.2070	-0.3166	-0.1946
df		14	14	15	14	14	14	15	15	15
PE		2.7896	2.7371	2.7359	2.7836	2.7514	2.7625	2.9080	2.7568	2.9866

Data no.84: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02490737.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0249	0.0249	0.0249	0.0249	0.0249	0.0249	0.0249	0.0249	0.0249
λ_1		0.2242	0.0996	0.0581	0.0374	0.0249	0.0166	0.0107	0.0062	0.0028
Predictors	x_1	2.8558	2.9166	2.9396	2.9511	2.9580	2.9626	2.9659	2.9684	2.9703
	x_2	0.3867	0.5653	0.6845	0.7441	0.7799	0.8037	0.8208	0.8336	0.8435
	x_3	1.2544	1.4271	1.5541	1.6177	1.6558	1.6812	1.6994	1.7130	1.7236
	x_4	0.8135	0.8715	1.0430	1.1288	1.1803	1.2145	1.2391	1.2574	1.2717
	x_5	-0.1551	-0.3456	-0.3198	-0.3069	-0.2992	-0.2941	-0.2903	-0.2876	-0.2854
	x_6	0.1844	0.3355	0.4064	0.4418	0.4631	0.4773	0.4874	0.4949	0.5009
	x_7	0.9584	1.1031	1.1565	1.1831	1.1992	1.2098	1.2175	1.2232	1.2276
	x_8	-0	-0.0394	-0.0623	-0.0737	-0.0805	-0.0851	-0.0884	-0.0908	-0.0927
	x_9	-0	-0.0156	-0.0603	-0.0826	-0.0960	-0.1049	-0.1113	-0.1161	-0.1198
	x_{10}	0	0.1307	0.2543	0.3161	0.3532	0.3779	0.3956	0.4089	0.4192
	x_{11}	0.2042	0.3688	0.4084	0.4282	0.4401	0.4480	0.4537	0.4579	0.4612
	x_{12}	0	0.02068	0.0157	0.0132	0.0117	0.0107	0.0100	0.0095	0.0091
	x_{13}	0.0863	0.2466	0.3166	0.3516	0.3726	0.3866	0.3966	0.4041	0.4099
	x_{14}	-0	-0.2737	-0.3881	-0.4453	-0.4796	-0.5025	-0.5188	-0.5311	-0.5406
	x_{15}	-0	-0.0611	-0.1735	-0.2296	-0.2633	-0.2858	-0.3018	-0.3139	-0.3232
df		9	15	15	15	15	15	15	15	15
PE		3.9249	3.1695	2.9255	2.8406	2.8014	2.7803	2.7676	2.7594	2.7538

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.85: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0155	0.0265	0.0163	0.0706	0.1058	0.0754	0.1069	0.1833	0.2360
λ_1		0.1399	0.1058	0.0380	0.1058	0.1058	0.0503	0.0458	0.0458	0.0262
Predictors	x_1	1.3590	1.3787	1.4523	1.3426	1.3020	1.3925	1.3577	1.2638	1.2239
	x_2	0.1086	0.1245	0.1135	0.1858	0.2007	0.2061	0.2157	0.2118	0.2145
	x_3	0.1314	0.1762	0.2872	0.1931	0.1798	0.2747	0.2548	0.1970	0.1934
	x_4	0.7540	0.7842	0.7751	0.8827	0.8965	0.9208	0.9271	0.8921	0.8764
	x_5	-0.7657	-0.8033	-1.0140	-0.6200	-0.5389	-0.7031	-0.6421	-0.5424	-0.5156
	x_6	0.1955	0.2564	0.3609	0.2389	0.2238	0.3228	0.3163	0.2907	0.2994
	x_7	0.7964	0.8219	0.9848	0.7576	0.7138	0.8402	0.7993	0.7042	0.6803
	x_8	0	0	0.1479	0	0	0.0616	0.0519	0.0234	0.0383
	x_9	-0.1470	-0.1508	-0.1652	-0.1583	-0.1657	-0.1701	-0.1799	-0.1996	-0.2161
	x_{10}	-0.1057	-0.1628	-0.3517	-0.1528	-0.14005	-0.2660	-0.2524	-0.2103	-0.2248
	x_{11}	0.8161	0.8266	0.9875	0.6970	0.6294	0.7542	0.6876	0.5668	0.5250
	x_{12}	0	0	-0	0	0	0	0.0134	0.0531	0.0808
	x_{13}	0.3620	0.3672	0.4824	0.3299	0.3130	0.3743	0.3526	0.3062	0.2971
	x_{14}	-0	-0	-0.1780	-0	-0	-0.0814	-0.0692	-0.0247	-0.0370
	x_{15}	0	0	0	0.0226	0.0543	0.0301	0.0597	0.1053	0.1295
df		11	11	13	12	12	14	15	15	15
PE		4.7891	4.7031	4.3772	4.8576	4.9771	4.6282	4.7220	4.9827	5.0551

Data no.85: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.0834155.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0834	0.0834	0.0834	0.0834	0.0834	0.0834	0.0834	0.0834	0.0834
λ_1		0.7507	0.3337	0.1946	0.1251	0.0834	0.0556	0.0357	0.0209	0.0093
Predictors	x_1	0.6141	1.1320	1.2541	1.3121	1.3468	1.3768	1.3993	1.4152	1.4277
	x_2	0	0.1839	0.1840	0.1908	0.1949	0.2075	0.2172	0.2192	0.2210
	x_3	0	0.0087	0.0890	0.1668	0.2135	0.2602	0.2943	0.3121	0.3265
	x_4	0.7337	0.9025	0.8703	0.8863	0.8959	0.9198	0.9383	0.9451	0.9505
	x_5	-0	-0.0829	-0.3931	-0.5447	-0.6357	-0.6753	-0.7002	-0.7265	-0.7466
	x_6	0	0	0.0789	0.1997	0.2722	0.3117	0.3398	0.3621	0.3795
	x_7	0	0.4670	0.6398	0.7189	0.7664	0.8166	0.8553	0.8848	0.9078
	x_8	-0	-0	0	0	0	0.0467	0.0833	0.1110	0.1326
	x_9	-0	-0.1156	-0.1528	-0.1592	-0.1631	-0.1699	-0.1766	-0.1843	-0.1904
	x_{10}	-0	-0	-0.0085	-0.1176	-0.1831	-0.2467	-0.2939	-0.3297	-0.3576
	x_{11}	0.2889	0.4114	0.5692	0.6476	0.6947	0.7285	0.7522	0.7655	0.7762
	x_{12}	0	0	0	0	0	0	0.0027	0.0172	0.0284
	x_{13}	0.1765	0.2220	0.2859	0.3148	0.3321	0.3608	0.3838	0.3998	0.4122
	x_{14}	0	0	0	-0	-0	-0.0607	-0.1143	-0.1556	-0.1878
	x_{15}	0.0176	0.0523	0.0531	0.0390	0.0305	0.0359	0.0419	0.0489	0.0542
df		5	10	12	12	12	14	15	15	15
PE		8.2933	6.0984	5.3218	4.9799	4.8211	4.6827	4.5940	4.5396	4.5033

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.86: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2CV .

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0409	0.0527	0.0683	0.0883	0.1750	0.0783	0.1012	0.1580	0.1061
λ_1		0.3684	0.2108	0.1595	0.1324	0.1750	0.0522	0.0434	0.0395	0.0118
Predictors	x_1	2.0941	2.1319	2.1556	2.1498	1.9826	2.2233	2.1886	2.0781	2.1953
	x_2	0	-0	-0	-0	0	-0.1622	-0.1504	-0.1001	-0.2432
	x_3	0	0.0349	0.1227	0.1813	0.1543	0.2907	0.3121	0.3020	0.3386
	x_4	0.0413	0.0365	0.0897	0.1458	0.2415	0.1116	0.1583	0.2266	0.1270
	x_5	-0.0341	-0.4130	-0.4579	-0.4435	-0.2121	-0.7320	-0.6804	-0.5634	-0.8065
	x_6	0.0321	0.1279	0.1894	0.2290	0.2320	0.3619	0.3787	0.3761	0.4324
	x_7	1.1488	1.3542	1.3828	1.3703	1.1621	1.4703	1.4287	1.3265	1.4415
	x_8	-0.3081	-0.3976	-0.4296	-0.4480	-0.4230	-0.4527	-0.4656	-0.4814	-0.4720
	x_9	-0	-0	-0	-0.0175	-0.0574	-0.2054	-0.2345	-0.2574	-0.3098
	x_{10}	-0	-0	-0	-0	-0	-0	-0.0095	-0.0294	-0.0338
	x_{11}	0.8930	1.0605	1.0490	1.0106	0.8135	1.0926	1.0217	0.8825	1.0197
	x_{12}	0	0	0	0.0046	0.0568	0.1461	0.1640	0.1691	0.2185
	x_{13}	0	0	0	0	0	0	0	0.0297	0.0383
	x_{14}	-0	-0	-0	-0	-0	-0.2044	-0.2012	-0.1593	-0.2677
	x_{15}	0	0	0	0	0.0200	0.1077	0.1198	0.1238	0.1667
df		7	8	8	10	11	13	14	15	15
PE		7.8830	7.0872	6.9612	6.9635	7.5748	6.3920	6.4821	6.7965	6.3645

Data no.86: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, and λ_2BF = 0.03433671.

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0343	0.0343	0.0343	0.0343	0.0343	0.0343	0.0343	0.0343	0.0343
λ_1		0.3090	0.1373	0.0801	0.0515	0.0343	0.0229	0.0147	0.0086	0.0038
Predictors	x_1	2.1193	2.1994	2.2640	2.2850	2.2976	2.3060	2.2974	2.2852	2.2752
	x_2	0	-0	-0.1571	-0.2509	-0.3073	-0.3449	-0.4061	-0.4690	-0.5185
	x_3	0	0.0882	0.1956	0.2421	0.2700	0.2886	0.2563	0.2133	0.1788
	x_4	0.0070	0	0	0	0	-0	-0.0757	-0.1642	-0.2342
	x_5	-0.2201	-0.6507	-0.8409	-0.9629	-1.0361	-1.0849	-1.1929	-1.3052	-1.3940
	x_6	0.0577	0.1697	0.2824	0.3477	0.3869	0.4130	0.4294	0.4420	0.4517
	x_7	1.2547	1.5064	1.5530	1.5711	1.5820	1.5892	1.6032	1.6166	1.6274
	x_8	-0.3398	-0.4243	-0.4209	-0.4200	-0.4195	-0.4192	-0.4201	-0.4196	-0.4191
	x_9	-0	-0	-0.0972	-0.1887	-0.2436	-0.2802	-0.3033	-0.3226	-0.3376
	x_{10}	-0	-0	-0	-0	-0	-0	-0.0094	-0.0122	-0.0146
	x_{11}	0.9890	1.2096	1.2821	1.2980	1.3075	1.3139	1.3271	1.3387	1.3480
	x_{12}	0	0	0.0721	0.1428	0.1852	0.2135	0.2483	0.2775	0.3008
	x_{13}	0	0	0	0	0	0	0.0153	0.0418	0.0627
	x_{14}	-0	-0.0380	-0.1736	-0.2621	-0.3152	-0.3506	-0.3638	-0.3704	-0.3754
	x_{15}	0	0	0.0568	0.1108	0.1433	0.1649	0.1939	0.2213	0.2425
df		7	8	12	12	12	12	15	15	15
PE		7.4712	6.6649	6.2915	6.1335	6.0607	6.0215	5.9819	5.9492	5.9255

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.87: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0134	0.0331	0.0296	0.0667	0.0690	0.0650	0.1218	0.1579	0.0802
λ_1		0.1206	0.1323	0.0690	0.1001	0.0690	0.0433	0.0522	0.0395	0.0089
Predictors	x_1	1.4818	1.4687	1.6027	1.4907	1.5531	1.6131	1.5107	1.4822	1.6562
	x_2	-0.3695	-0.2961	-0.3787	-0.2724	-0.3046	-0.3385	-0.2687	-0.2561	-0.3549
	x_3	0.6492	0.6889	0.7622	0.7725	0.8093	0.8373	0.8330	0.8376	0.8821
	x_4	0	0.0781	0.1352	0.1986	0.2358	0.2570	0.3071	0.3368	0.3108
	x_5	-0.9950	-0.8800	-0.8559	-0.7722	-0.7495	-0.7406	-0.6840	-0.6589	-0.7078
	x_6	0.0060	0.0154	0.0374	0.0564	0.0705	0.0779	0.1066	0.1264	0.1028
	x_7	1.5213	1.4694	1.5465	1.4337	1.4616	1.4967	1.3832	1.3380	1.5029
	x_8	0.2760	0.2258	0.3897	0.2443	0.3167	0.3863	0.2795	0.2624	0.4394
	x_9	0	0	0.0038	0	0	0.0039	0	0	0.0247
	x_{10}	0.1846	0.1961	0.2249	0.2421	0.2586	0.2682	0.2928	0.3103	0.2945
	x_{11}	1.1989	1.1048	1.2163	1.0433	1.0815	1.1308	0.9741	0.9194	1.1371
	x_{12}	-0	-0	-0.1628	-0.0321	-0.1161	-0.1924	-0.1030	-0.0976	-0.2570
	x_{13}	0.8333	0.7721	0.8722	0.7502	0.7896	0.8325	0.7323	0.7035	0.8519
	x_{14}	-0.4683	-0.4353	-0.5100	-0.4509	-0.4773	-0.5049	-0.4405	-0.4179	-0.5218
	x_{15}	0	0	0.0605	0.0295	0.0649	0.0925	0.0931	0.1111	0.1294
df		11	12	15	14	14	15	14	14	15
PE		5.0816	5.2365	4.8757	5.2253	5.0581	4.9189	5.2261	5.3325	4.8529

Data no.87: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.05737284.

α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0574	0.0574	0.0574	0.0574	0.0574	0.0574	0.0574	0.0574	0.0574
λ_1		0.5164	0.2295	0.1339	0.0861	0.0574	0.0382	0.0246	0.0143	0.0064
Predictors	x_1	1.2312	1.3870	1.4557	1.5313	1.5946	1.6362	1.6659	1.6881	1.7054
	x_2	-0	-0.1063	-0.2393	-0.3029	-0.3349	-0.3542	-0.3674	-0.3774	-0.3851
	x_3	0.4056	0.6290	0.7339	0.7809	0.8153	0.8398	0.8578	0.8712	0.8817
	x_4	0	0.1206	0.1630	0.1948	0.2279	0.2494	0.2648	0.2762	0.2852
	x_5	-0.4609	-0.7059	-0.7793	-0.7843	-0.7631	-0.7508	-0.7421	-0.7358	-0.7307
	x_6	0	0	0.0323	0.0556	0.0667	0.0738	0.0788	0.0826	0.0855
	x_7	1.0757	1.3294	1.4199	1.4668	1.4974	1.5189	1.5344	1.5461	1.5552
	x_8	0	0.0366	0.1906	0.2932	0.3656	0.4134	0.4475	0.4729	0.4929
	x_9	-0	0	0	0	0	0.0119	0.0229	0.0313	0.0377
	x_{10}	0	0.1480	0.2152	0.2424	0.2559	0.2649	0.2711	0.2759	0.2795
	x_{11}	0.6125	0.8851	1.0244	1.0914	1.1336	1.1632	1.1847	1.2008	1.2134
	x_{12}	-0	-0	-0	-0.0795	-0.1636	-0.2179	-0.2565	-0.2854	-0.3079
	x_{13}	0.3806	0.6069	0.7233	0.7866	0.8275	0.8547	0.8741	0.8886	0.8999
	x_{14}	-0	-0.2641	-0.4149	-0.4723	-0.4989	-0.5171	-0.5303	-0.5401	-0.5478
	x_{15}	0	0	0	0.0436	0.0763	0.0962	0.1101	0.1205	0.1286
df		6	11	12	14	14	15	15	15	15
PE		7.9622	6.0223	5.3734	5.0937	4.9468	4.8651	4.8153	4.7829	4.7605

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.88: The naïve elastic net estimates $(\hat{\beta})$ using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0170	0.0317	0.0411	0.0530	0.0873	0.0903	0.0804	0.1660	0.0925
λ_1		0.1526	0.1267	0.0958	0.0796	0.0873	0.0602	0.0344	0.0415	0.0103
Predictors	x_1	2.3414	2.3159	2.3126	2.2895	2.1912	2.2036	2.2410	2.0427	2.2102
	x_2	-0.2893	-0.3221	-0.3642	-0.3864	-0.3593	-0.4009	-0.4734	-0.3987	-0.5456
	x_3	0.9155	0.9112	0.9143	0.9174	0.8953	0.9100	0.9148	0.8690	0.8862
	x_4	0	0	0	-0	-0	-0	-0.0576	-0.0197	-0.1543
	x_5	-0.0613	-0.1365	-0.2223	-0.2671	-0.2314	-0.3043	-0.4182	-0.3275	-0.5426
	x_6	-0	-0	-0.0066	-0.0296	-0.0192	-0.0546	-0.1030	-0.0716	-0.1483
	x_7	1.1973	1.1990	1.2023	1.2064	1.1992	1.2072	1.2226	1.1914	1.2377
	x_8	-0	-0	-0	-0	-0	-0	-0	-0	-0.0029
	x_9	-0	-0	-0	-0	-0	-0	-0.0348	0	-0.0581
	x_{10}	0.1632	0.1864	0.2100	0.2271	0.2342	0.2545	0.2741	0.2792	0.3050
	x_{11}	0.5558	0.5695	0.5943	0.5950	0.5435	0.5661	0.6144	0.5181	0.6374
	x_{12}	0.3323	0.3573	0.3942	0.4017	0.3537	0.3844	0.4278	0.3472	0.4473
	x_{13}	0.6072	0.6052	0.6084	0.6055	0.5856	0.5913	0.6107	0.5636	0.6230
	x_{14}	-0.5285	-0.5600	-0.6192	-0.6222	-0.5029	-0.5582	-0.6349	-0.4276	-0.6292
	x_{15}	0.1489	0.1840	0.2164	0.2366	0.2478	0.2684	0.2925	0.3022	0.3271
df		11	11	12	12	12	12	14	13	15
PE		4.9315	4.8470	4.7277	4.7029	4.9028	4.7823	4.6099	5.0280	4.5543

Data no.88: The naïve elastic net estimates $(\hat{\beta})$ using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.05526184.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0553	0.0553	0.0553	0.0553	0.0553	0.0553	0.0553	0.0553	0.0553
λ_1		0.4974	0.2210	0.1289	0.0829	0.0553	0.0368	0.0237	0.0138	0.0061
Predictors	x_1	1.8466	2.1379	2.2443	2.2807	2.3075	2.3165	2.3111	2.3071	2.3040
	x_2	-0	-0.1840	-0.3106	-0.3800	-0.4282	-0.4779	-0.5286	-0.5666	-0.5971
	x_3	0.5945	0.8454	0.8974	0.9148	0.9398	0.9399	0.9188	0.9026	0.8896
	x_4	0	0	0	-0	-0	-0.0404	-0.1071	-0.1573	-0.1973
	x_5	-0	-0	-0.1273	-0.2559	-0.3307	-0.4118	-0.5030	-0.5717	-0.6237
	x_6	-0	-0	-0	-0.0251	-0.0683	-0.1046	-0.1347	-0.1571	-0.1742
	x_7	1.0645	1.1836	1.1970	1.2054	1.2125	1.2218	1.2329	1.2411	1.2477
	x_8	-0	-0	-0	-0	-0	-0	-0	-0.0003	-0.0089
	x_9	0	0	-0	-0	-0.0276	-0.0576	-0.0759	-0.0897	-0.0996
	x_{10}	0	0.1311	0.1963	0.2259	0.2421	0.2571	0.2727	0.2843	0.2933
	x_{11}	0.4186	0.4476	0.5355	0.5879	0.6208	0.6504	0.6785	0.6998	0.7168
	x_{12}	0.0552	0.2016	0.3270	0.3941	0.4302	0.4553	0.4760	0.4918	0.5040
	x_{13}	0.5434	0.5705	0.5904	0.6035	0.6108	0.6208	0.6321	0.6407	0.6466
	x_{14}	-0	-0.2286	-0.4811	-0.6061	-0.6833	-0.7236	-0.7401	-0.7526	-0.7627
	x_{15}	0.0714	0.1419	0.1992	0.2353	0.2578	0.2792	0.2997	0.3148	0.3265
df		7	10	11	12	13	14	14	15	15
PE		6.8903	5.6025	4.9794	4.7292	4.5963	4.5124	4.4568	4.4208	4.3962

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.89: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0257	0.0480	0.0622	0.1166	0.1323	0.1502	0.2336	0.2088	0.2449
λ_1		0.2313	0.1920	0.1452	0.1749	0.1323	0.1001	0.1001	0.0522	0.0272
Predictors	x_1	1.7442	1.6987	1.6695	1.5680	1.5462	1.5224	1.4121	1.4458	1.4064
	x_2	0	0	0	-0	-0	-0	-0	-0.0082	-0.0485
	x_3	0.5610	0.5144	0.5163	0.4522	0.4613	0.4657	0.4272	0.4475	0.4384
	x_4	0.6483	0.6209	0.6361	0.5498	0.5679	0.5770	0.5158	0.5473	0.5227
	x_5	-0.1159	-0.2227	-0.2965	-0.2718	-0.3248	-0.3620	-0.3464	-0.4234	-0.4606
	x_6	0	0	0	0	0	0	0	0	0.0318
	x_7	0.8265	0.8573	0.9081	0.8231	0.8622	0.8854	0.8208	0.8957	0.8906
	x_8	-0.2177	-0.2544	-0.2952	-0.2598	-0.2947	-0.3195	-0.3064	-0.3514	-0.3648
	x_9	-0	-0.0142	-0.0918	-0.0380	-0.1028	-0.1490	-0.1423	-0.2159	-0.2505
	x_{10}	0.1209	0.1648	0.2307	0.1954	0.2490	0.2876	0.2829	0.3413	0.3660
	x_{11}	0.6403	0.6298	0.6273	0.5943	0.5932	0.5902	0.5564	0.57116	0.5586
	x_{12}	0.0628	0.0975	0.1164	0.1576	0.1712	0.1842	0.2147	0.2096	0.2298
	x_{13}	-0	-0	-0	0	-0	-0	0	-0	-0
	x_{14}	0	0	0	0	0	0	0.0146	0.0368	0.0586
	x_{15}	0.0470	0.1036	0.1399	0.1532	0.1776	0.1960	0.2072	0.2194	0.2293
df		10	11	11	11	11	11	12	13	14
PE		4.7583	4.6634	4.5291	4.7577	4.6466	4.5911	4.7933	4.5985	4.6125

Data no.89: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03784048.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0378	0.0378	0.0378	0.0378	0.0378	0.0378	0.0378	0.0378	0.0378
λ_1		0.3406	0.1514	0.0883	0.0568	0.0378	0.0252	0.0162	0.0095	0.0042
Predictors	x_1	1.6686	1.7221	1.7377	1.7788	1.8034	1.8210	1.8336	1.8459	1.8563
	x_2	0	0	0	0	0	0	0	0.0112	0.0229
	x_3	0.5294	0.5494	0.5900	0.6356	0.6569	0.6651	0.6708	0.6792	0.6869
	x_4	0.5518	0.6724	0.7407	0.8102	0.8444	0.8599	0.8708	0.8869	0.9017
	x_5	-0	-0.2689	-0.3381	-0.3327	-0.3374	-0.3487	-0.3570	-0.3546	-0.3501
	x_6	-0	0	0	0	0.0054	0.0234	0.0363	0.0454	0.0523
	x_7	0.6947	0.9226	1.0109	1.0559	1.0850	1.1058	1.1207	1.1326	1.1421
	x_8	-0.0915	-0.2942	-0.3506	-0.3728	-0.3864	-0.3955	-0.4019	-0.4069	-0.4108
	x_9	-0	-0.0881	-0.1863	-0.2061	-0.2204	-0.2334	-0.2428	-0.2497	-0.2550
	x_{10}	0.0255	0.2206	0.3043	0.3342	0.3554	0.3752	0.3894	0.4017	0.4118
	x_{11}	0.6266	0.6414	0.6537	0.6692	0.6755	0.6730	0.6711	0.6669	0.6629
	x_{12}	0.0703	0.0882	0.1073	0.1491	0.1737	0.1909	0.2031	0.2093	0.2133
	x_{13}	0	-0	-0.0450	-0.1653	-0.2387	-0.2882	-0.3235	-0.3512	-0.3730
	x_{14}	0	0	0	0	0.0117	0.0331	0.0486	0.0599	0.0687
	x_{15}	0.0058	0.1168	0.1601	0.1984	0.2218	0.2373	0.2486	0.2561	0.2616
df		9	11	12	12	14	14	14	15	15
PE		5.2160	4.4996	4.3044	4.1763	4.1158	4.0815	4.0614	4.0476	4.0380

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.90: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0292	0.0312	0.0588	0.0914	0.1139	0.1292	0.0870	0.1637	0.1001
λ_1		0.2631	0.1250	0.1372	0.1372	0.1139	0.0861	0.0373	0.0409	0.0111
Predictors	x_1	1.9095	2.0097	1.9455	1.8848	1.8580	1.8482	1.9592	1.8223	1.9381
	x_2	0.0435	0.1062	0.1137	0.1250	0.1486	0.1699	0.2056	0.2119	0.2069
	x_3	-0	-0.0109	-0	-0	-0	-0.0064	-0.0931	-0.0478	-0.1711
	x_4	-0.1653	-0.3999	-0.3555	-0.3314	-0.3543	-0.3891	-0.4551	-0.4305	-0.5337
	x_5	0	0	0	0	0	0	-0	0	-0.1241
	x_6	-0	-0	-0	-0	-0	-0.0029	-0.0361	-0.0460	-0.0602
	x_7	1.5142	1.6110	1.5610	1.5151	1.5080	1.5127	1.6640	1.5181	1.6858
	x_8	-0.5245	-0.6312	-0.6052	-0.5866	-0.5902	-0.6000	-0.6859	-0.6132	-0.7041
	x_9	0	0.0664	0.0747	0.0932	0.1326	0.1701	0.2183	0.2407	0.2732
	x_{10}	0	0	0	0	0	0	-0	0	-0.0213
	x_{11}	0.4553	0.5037	0.4881	0.4762	0.4652	0.4587	0.4708	0.4444	0.4825
	x_{12}	0	-0	-0	-0	-0	-0	-0.1816	-0.0536	-0.2298
	x_{13}	0.5265	0.5352	0.5227	0.5105	0.4962	0.4866	0.5510	0.4889	0.5774
	x_{14}	0.0774	0.1240	0.1307	0.1428	0.1558	0.1678	0.2526	0.2105	0.3056
	x_{15}	0	0	0	0	0.0286	0.0550	0.1247	0.1082	0.1775
df		8	10	9	9	10	12	13	13	15
PE		5.0960	4.7224	4.8156	4.9058	4.9114	4.8880	4.5171	4.8287	4.4385

Data no.90: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02669419.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267	0.0267
λ_1		0.2402	0.1068	0.0623	0.0400	0.0267	0.0178	0.0114	0.0067	0.0030
Predictors	x_1	1.9324	2.0371	2.0619	2.0752	2.0653	2.0581	2.0530	2.0491	2.0461
	x_2	0.0526	0.1050	0.1647	0.2040	0.1806	0.16364	0.1514	0.1419	0.1348
	x_3	-0	-0.0501	-0.1003	-0.1243	-0.2175	-0.2819	-0.3280	-0.3629	-0.3899
	x_4	-0.2069	-0.4386	-0.4435	-0.4393	-0.5299	-0.5929	-0.6382	-0.6727	-0.6992
	x_5	0	0	-0	-0	-0.1547	-0.2622	-0.3392	-0.3975	-0.4425
	x_6	-0	-0	-0	-0.0132	-0.0250	-0.0329	-0.0385	-0.0428	-0.0461
	x_7	1.5343	1.6338	1.7504	1.8158	1.8496	1.8720	1.8880	1.8998	1.9091
	x_8	-0.5439	-0.6491	-0.7259	-0.7648	-0.7872	-0.80216	-0.8127	-0.8206	-0.8268
	x_9	0	0.0752	0.1462	0.1946	0.2315	0.2563	0.2739	0.2871	0.2974
	x_{10}	0	-0	-0	-0.0222	-0.0534	-0.0742	-0.0889	-0.1001	-0.1087
	x_{11}	0.4636	0.5086	0.4981	0.4820	0.4994	0.5118	0.5207	0.5275	0.5327
	x_{12}	0	-0	-0.2291	-0.3543	-0.3985	-0.4272	-0.4475	-0.4625	-0.4743
	x_{13}	0.5304	0.5399	0.5975	0.6272	0.6667	0.6936	0.7128	0.7274	0.7386
	x_{14}	0.0834	0.1353	0.2519	0.3161	0.3724	0.4105	0.4376	0.4580	0.4739
	x_{15}	0	0	0.0927	0.1535	0.2003	0.2319	0.2545	0.2714	0.2846
df		8	10	12	14	15	15	15	15	15
PE		5.0126	4.6682	4.4143	4.3123	4.2386	4.1983	4.1741	4.1586	4.1481

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.91: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0343	0.0771	0.1000	0.0890	0.1335	0.3420	0.1957	0.0757	0.1704
λ_1		0.3085	0.3085	0.2334	0.1335	0.1335	0.2280	0.0839	0.0189	0.0189
Predictors	x_1	1.3699	1.3279	1.3571	1.4290	1.3853	1.5114	1.3590	1.4905	1.4107
	x_2	0	0	0.0534	0.1547	0.1496	0.0422	0.1859	0.1002	0.1555
	x_3	-0	-0	-0	-0.1593	-0.1423	-0.7300	-0.1985	-0.6089	-0.4393
	x_4	0	0	0	-0	-0	-0.4920	-0	-0.3653	-0.2066
	x_5	0	0	0	-0	-0	-0.8850	-0.0068	-0.6658	-0.3873
	x_6	-0	-0	-0	-0.0528	-0.0739	-0.1788	-0.1328	-0.1833	-0.1973
	x_7	1.7335	1.6561	1.7024	1.8444	1.7629	2.0110	1.7135	1.9383	1.7813
	x_8	0	0	0	0	0	0.1167	0	0.0913	0.0373
	x_9	0.2268	0.2013	0.2926	0.4334	0.4090	0.6082	0.4371	0.5767	0.5125
	x_{10}	-0	-0	-0	-0	-0	-0.1126	-0.0153	-0.1002	-0.0740
	x_{11}	0.7729	0.7280	0.7245	0.7447	0.6960	0.8968	0.6433	0.8232	0.7010
	x_{12}	0	0	0	0.0104	0.0245	0.1072	0.0468	0.0960	0.0907
	x_{13}	0.2108	0.2453	0.2354	0.1885	0.2105	0.2279	0.2204	0.2299	0.2483
	x_{14}	0	0	0	0	0.0113	0.0714	0.0318	0.0728	0.0746
	x_{15}	0	0	0.0477	0.1283	0.1222	0.4479	0.1471	0.3670	0.2635
df		5	5	7	10	11	15	13	15	15
PE		4.2743	4.3945	4.2376	3.9124	4.0141	3.3503	4.0367	3.4512	3.7095

Data no.91: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02782389.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0278	0.0278	0.0278	0.0278	0.0278	0.0278	0.0278	0.0278	0.0278
λ_1		0.2504	0.1113	0.0649	0.0417	0.0278	0.0185	0.0119	0.0070	0.0031
Predictors	x_1	1.4196	1.5100	1.5572	1.5313	1.5168	1.5072	1.5060	1.5024	1.4943
	x_2	0.0269	0.1849	0.2018	0.1071	0.0468	0.0067	0	-0.0153	-0.0478
	x_3	-0	-0.2287	-0.3500	-0.5755	-0.7138	-0.8059	-0.8619	-0.9084	-0.9540
	x_4	0	-0	-0.0426	-0.3066	-0.4705	-0.5798	-0.6410	-0.6949	-0.7524
	x_5	0	-0	-0.2105	-0.6124	-0.8581	-1.0218	-1.1143	-1.1947	-1.2807
	x_6	-0	-0.0252	-0.0785	-0.1351	-0.1688	-0.1913	-0.2070	-0.2189	-0.2285
	x_7	1.8075	2.0116	2.0556	2.0373	2.0252	2.0171	2.0149	2.0115	2.0056
	x_8	0	0.0028	0.0859	0.1053	0.1168	0.1246	0.1321	0.1367	0.1385
	x_9	0.3083	0.5115	0.5777	0.5971	0.6094	0.6176	0.6219	0.6258	0.6303
	x_{10}	-0	-0.0260	-0.0796	-0.1003	-0.1114	-0.1187	-0.1245	-0.1285	-0.1313
	x_{11}	0.8072	0.8256	0.8441	0.8814	0.9034	0.9181	0.9246	0.9313	0.9403
	x_{12}	0	0	0.0036	0.0631	0.0994	0.1237	0.1356	0.1472	0.1611
	x_{13}	0.1968	0.1228	0.1121	0.1786	0.2178	0.2439	0.2557	0.2679	0.2837
	x_{14}	0	0	0	0.0328	0.0626	0.0824	0.0911	0.1004	0.1125
	x_{15}	0.0085	0.1531	0.2706	0.3826	0.4456	0.4876	0.5179	0.5401	0.5577
df		7	11	14	15	15	15	14	15	15
PE		4.1143	3.7397	3.5814	3.4277	3.3542	3.3137	3.2925	3.2771	3.2630

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.92: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0465	0.0546	0.0444	0.0691	0.0715	0.1291	0.1830	0.1490	0.1919
λ_1		0.4186	0.2182	0.1037	0.1037	0.0715	0.0861	0.0784	0.0373	0.0213
Predictors	x_1	1.9801	2.1245	2.1583	2.1204	2.1146	2.0194	1.9402	1.9939	1.9388
	x_2	0	0	0	0	0	0	0	0	0.0088
	x_3	0.4661	0.6413	0.6536	0.6726	0.6747	0.6978	0.7108	0.7125	0.7275
	x_4	0.0735	0.2780	0.3090	0.3188	0.3192	0.3185	0.3170	0.3275	0.3375
	x_5	-0	-0	-0.2396	-0.2187	-0.2935	-0.2425	-0.2415	-0.3302	-0.3336
	x_6	0	0	0.0406	0.0422	0.0708	0.0520	0.0557	0.0920	0.1028
	x_7	1.2409	1.3592	1.4211	1.3984	1.4096	1.3503	1.3077	1.3533	1.3220
	x_8	-0	-0	-0	-0	0	-0	-0	0	0
	x_9	0	0.1451	0.2818	0.2705	0.3037	0.2573	0.2412	0.2964	0.2906
	x_{10}	0	0.2013	0.3405	0.3331	0.3775	0.3477	0.3465	0.4067	0.4144
	x_{11}	0.3132	0.3360	0.4515	0.4402	0.4680	0.4245	0.4073	0.4486	0.4373
	x_{12}	0	0	-0	-0	-0	0	0	-0	0
	x_{13}	0.4872	0.4918	0.5655	0.5489	0.5597	0.5086	0.4799	0.5114	0.4918
	x_{14}	0	0	-0	0	-0	0	0	0	0
	x_{15}	0	0	0	0	0.0186	0.0519	0.0820	0.0865	0.1082
df		6	8	10	10	11	11	11	11	12
PE		5.5312	4.8856	4.4947	4.5363	4.4651	4.6062	4.7015	4.5352	4.5886

Data no.92: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03225788.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0323	0.0323	0.0323	0.0323	0.0323	0.0323	0.0323	0.0323	0.0323
λ_1		0.2903	0.1290	0.0753	0.0484	0.0323	0.0215	0.0138	0.0081	0.0036
Predictors	x_1	2.1093	2.1767	2.1775	2.1900	2.2009	2.1897	2.1818	2.1758	2.1711
	x_2	0	0	-0	-0	-0.0030	-0.0505	-0.0843	-0.1096	-0.1293
	x_3	0.5675	0.6429	0.6412	0.6592	0.6833	0.6813	0.6801	0.6791	0.6784
	x_4	0.2027	0.3012	0.3042	0.3224	0.3381	0.3054	0.2826	0.2655	0.2521
	x_5	-0	-0.1936	-0.3167	-0.3545	-0.3705	-0.4192	-0.4532	-0.4788	-0.4987
	x_6	0	0.0133	0.0696	0.0898	0.0898	0.0926	0.0954	0.0976	0.0992
	x_7	1.3328	1.4209	1.4455	1.4639	1.4804	1.5053	1.5235	1.5371	1.5477
	x_8	-0	-0	0	0	0	0.0129	0.0252	0.0343	0.0415
	x_9	0.0670	0.2568	0.3222	0.3679	0.4051	0.4351	0.4566	0.4727	0.4853
	x_{10}	0.1137	0.3106	0.3818	0.4196	0.4488	0.4727	0.4902	0.5034	0.5136
	x_{11}	0.3243	0.4284	0.4911	0.5236	0.5424	0.5661	0.5830	0.5957	0.6056
	x_{12}	0	-0	-0	-0	-0	-0	-0	-0	-0
	x_{13}	0.4991	0.5581	0.5938	0.6214	0.6409	0.6562	0.6671	0.6752	0.6815
	x_{14}	0	-0	-0	-0.0346	-0.0882	-0.1161	-0.1362	-0.1512	-0.1629
	x_{15}	0	0	0	0.0025	0.0266	0.0574	0.0792	0.0955	0.1083
df		8	10	10	12	13	14	14	14	14
PE		5.0697	4.5430	4.4147	4.3536	4.3109	4.2817	4.2646	4.2539	4.2469

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.93: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0105	0.0237	0.2802	0.0525	0.0495	0.0424	0.1051	0.1242	0.1328
λ_1		0.0948	0.0948	0.6537	0.0787	0.0495	0.0283	0.0451	0.0311	0.0148
Predictors	x_1	2.6492	2.6028	2.6036	2.5202	2.5335	2.5696	2.3813	2.3353	2.3207
	x_2	0.6162	0.6158	0.6281	0.6199	0.6313	0.6977	0.6157	0.6181	0.6395
	x_3	0.2998	0.2978	0.3144	0.3038	0.3219	0.3885	0.3203	0.3280	0.3507
	x_4	0	0	0	0	0	0.1023	0	0	0.0323
	x_5	-0.1067	-0.1208	-0.1356	-0.1485	-0.1708	-0.1185	-0.2023	-0.2214	-0.2235
	x_6	-0	-0	-0	-0	-0	-0	-0	-0	-0
	x_7	0.2367	0.2405	0.2652	0.2597	0.2813	0.2816	0.2928	0.3035	0.3114
	x_8	0.5152	0.5080	0.5351	0.5087	0.5347	0.5536	0.5072	0.5079	0.5160
	x_9	0	0	0.0113	0	0.0374	0.1083	0.0155	0.0353	0.0682
	x_{10}	-0.4628	-0.4672	-0.5028	-0.4921	-0.5223	-0.5636	-0.5210	-0.5342	-0.5547
	x_{11}	0.9300	0.9041	0.9342	0.8732	0.9163	0.9473	0.8321	0.8224	0.8257
	x_{12}	-0.4082	-0.3913	-0.4240	-0.3761	-0.4247	-0.4853	-0.3707	-0.3716	-0.3890
	x_{13}	0.4670	0.4632	0.5118	0.4804	0.5245	0.5451	0.4991	0.5075	0.5186
	x_{14}	0	0	0.0048	0	0.0347	0.0375	0.0588	0.0793	0.0924
	x_{15}	-0.2902	-0.2564	-0.3044	-0.2229	-0.2905	-0.3619	-0.2000	-0.1971	-0.2154
df		11	11	13	11	13	14	13	13	14
PE		3.6853	3.7288	3.6427	3.7724	3.6522	3.5525	3.8369	3.8559	3.8255

Data no.93: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.02356857.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0236	0.0236	0.0236	0.0236	0.0236	0.0236	0.0236	0.0236	0.0236
λ_1		0.2121	0.0943	0.0550	0.0354	0.0236	0.0157	0.0101	0.0059	0.0026
Predictors	x_1	2.5270	2.6038	2.6177	2.6358	2.6468	2.6534	2.6534	2.6534	2.6582
	x_2	0.5722	0.6161	0.6417	0.7143	0.7579	0.7871	0.8001	0.8096	0.8259
	x_3	0.2312	0.2983	0.3280	0.4041	0.4498	0.4826	0.4946	0.5034	0.5235
	x_4	0	0	0.0117	0.1288	0.1992	0.2480	0.2687	0.2840	0.3120
	x_5	-0.0921	-0.1205	-0.1341	-0.0687	-0.0294	-0	-0	0	0.0235
	x_6	0	-0	-0	-0	-0	-0.0069	-0.0114	-0.0148	-0.0198
	x_7	0.1423	0.2409	0.2687	0.2679	0.2675	0.2672	0.2682	0.2690	0.2683
	x_8	0.3849	0.5087	0.5447	0.5582	0.5664	0.5720	0.5758	0.5787	0.5811
	x_9	-0	0	0.0427	0.1099	0.1504	0.1749	0.1918	0.2044	0.2144
	x_{10}	-0.3350	-0.4678	-0.5185	-0.5604	-0.5855	-0.6025	-0.6127	-0.6204	-0.6286
	x_{11}	0.7567	0.9051	0.9585	0.9731	0.9819	0.9871	0.9940	0.9992	0.9998
	x_{12}	-0.2475	-0.3923	-0.4505	-0.5055	-0.5386	-0.5618	-0.5746	-0.5842	-0.5959
	x_{13}	0.2636	0.4641	0.5306	0.5428	0.5500	0.5565	0.5646	0.5708	0.5724
	x_{14}	0	0	0.0149	0.0154	0.0157	0.0149	0.0183	0.02099	0.0182
	x_{15}	-0.0255	-0.2578	-0.3411	-0.3983	-0.4326	-0.4539	-0.4668	-0.4765	-0.4865
df		11	11	14	14	14	14	14	14	15
PE		4.3645	3.7262	3.5908	3.5185	3.4850	3.4665	3.4568	3.4506	3.4447

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.94: he naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0148	0.0638	0.0687	0.1068	0.1105	0.1510	0.1776	0.2099	0.3572
λ_1		0.1330	0.2552	0.1602	0.1602	0.1105	0.1006	0.0761	0.0525	0.0397
Predictors	x_1	1.9545	1.8146	1.8423	1.7765	1.7797	1.7162	1.6782	1.6343	1.4693
	x_2	0.1225	0	0.0526	0.0302	0.0735	0.0647	0.0760	0.0807	0.0506
	x_3	0.8051	0.5757	0.6968	0.6557	0.7112	0.6851	0.6886	0.6804	0.5947
	x_4	0.8596	0.6840	0.7479	0.6930	0.6991	0.6572	0.6357	0.6052	0.5205
	x_5	-0.0288	-0	-0.0407	-0.0688	-0.1475	-0.1777	-0.2226	-0.2738	-0.3018
	x_6	0.3652	0.2522	0.3381	0.3366	0.3813	0.3818	0.3935	0.3991	0.3685
	x_7	1.5906	1.3158	1.4534	1.3958	1.4629	1.4200	1.4192	1.4100	1.2629
	x_8	-0.2813	-0.1206	-0.2222	-0.2077	-0.2559	-0.2488	-0.2615	-0.2750	-0.2487
	x_9	-0.3955	-0.2637	-0.3635	-0.3629	-0.4203	-0.4294	-0.4556	-0.4773	-0.4599
	x_{10}	-0	-0	-0	-0	-0	-0.0232	-0.0531	-0.0853	-0.1266
	x_{11}	0.6416	0.6095	0.6139	0.5908	0.5989	0.5758	0.5650	0.5487	0.4721
	x_{12}	0	0	0	0	0	0	0.0079	0.0364	0.0778
	x_{13}	0	0.0321	0.0565	0.0855	0.1023	0.1208	0.1325	0.1413	0.1565
	x_{14}	0	0	0	0.0199	0.0400	0.0671	0.0870	0.1042	0.1305
	x_{15}	-0	0	0	0	0	0	0	0	0.0187
df		10	9	11	12	12	13	14	14	15
PE		5.1579	5.7508	5.3838	5.4966	5.3217	5.4058	5.4058	5.4342	5.8573

Data no.94: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03452689.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0345	0.0345	0.0345	0.0345	0.0345	0.0345	0.0345	0.0345	0.0345
λ_1		0.3107	0.1381	0.0806	0.0518	0.0345	0.0230	0.0148	0.0086	0.0038
Predictors	x_1	1.8388	1.9117	1.9209	1.9193	1.9302	1.944	1.9549	1.9631	1.9695
	x_2	0	0.0991	0.1648	0.1997	0.2469	0.2898	0.3225	0.3469	0.3660
	x_3	0.5264	0.7685	0.8512	0.8927	0.9410	0.9894	1.0268	1.0548	1.0765
	x_4	0.6611	0.8144	0.8328	0.8395	0.8632	0.9007	0.9311	0.9539	0.9717
	x_5	-0	-0.0446	-0.1289	-0.1723	-0.1453	-0.1022	-0.0670	-0.0408	-0.0204
	x_6	0.1999	0.3601	0.4167	0.4419	0.4625	0.4809	0.4949	0.5054	0.5135
	x_7	1.2706	1.5456	1.6374	1.6853	1.7060	1.7203	1.7307	1.7384	1.7445
	x_8	-0.0641	-0.2632	-0.3237	-0.3513	-0.3665	-0.3719	-0.3752	-0.3776	-0.3794
	x_9	-0.2074	-0.3899	-0.4662	-0.5079	-0.5449	-0.5700	-0.5881	-0.6015	-0.6120
	x_{10}	-0	-0	-0	-0.0228	-0.0426	-0.0493	-0.0530	-0.0558	-0.0579
	x_{11}	0.6311	0.6352	0.6525	0.6640	0.6750	0.6802	0.6836	0.6862	0.6881
	x_{12}	0	0	-0	-0	-0	-0.0255	-0.0479	-0.0649	-0.0779
	x_{13}	0	0.0248	0.0488	0.0578	0.0619	0.0649	0.0669	0.0686	0.0697
	x_{14}	0	0	0	0	0	0	0	0	-0
	x_{15}	0	-0	-0	-0	-0.0392	-0.0634	-0.0803	-0.0929	-0.1028
df		8	11	11	12	13	14	14	14	14
PE		5.9039	5.2158	5.0584	5.0034	4.9627	4.9345	4.9167	4.9052	4.8975

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.95: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0097	0.0181	0.0374	0.0400	0.0547	0.0390	0.0607	0.0595	0.0766
λ_1		0.0872	0.0724	0.0872	0.0601	0.0547	0.0260	0.0260	0.0149	0.0085
Predictors	x_1	2.0257	2.0177	1.9458	1.9709	1.9382	2.0254	1.9595	1.9722	1.9323
	x_2	0	0.0025	0	0.0148	0.0157	0.1010	0.0741	0.0913	0.0851
	x_3	-0	-0	-0	-0	-0	-0.0282	-0.0223	-0.0567	-0.0645
	x_4	0.0088	0.0251	0.0156	0.0444	0.0526	0.0842	0.0826	0.0758	0.0788
	x_5	-0.6139	-0.6219	-0.5757	-0.6052	-0.5937	-0.6361	-0.6207	-0.6526	-0.6486
	x_6	0	0	0	0	0	0.0241	0	0.0169	0.0084
	x_7	1.0854	1.0984	1.0471	1.0899	1.0789	1.1534	1.1281	1.1445	1.1336
	x_8	0.4751	0.4902	0.5016	0.5147	0.5276	0.5263	0.5408	0.5504	0.5638
	x_9	-0	-0	-0	-0	-0	-0.0113	-0.0039	-0.0217	-0.0247
	x_{10}	-0.2413	-0.2638	-0.2537	-0.2898	-0.2992	-0.3626	-0.3540	-0.3837	-0.3908
	x_{11}	1.1721	1.1613	1.0299	1.0769	1.0243	1.1575	1.0551	1.0865	1.0298
	x_{12}	0.3724	0.3786	0.3763	0.3849	0.3868	0.4075	0.4013	0.4164	0.4177
	x_{13}	0.2652	0.2707	0.2536	0.2682	0.2668	0.2694	0.2737	0.2728	0.2747
	x_{14}	0.0222	0.0445	0.0295	0.0654	0.0738	0.1283	0.1204	0.1449	0.1508
	x_{15}	-0.6357	-0.6563	-0.4959	-0.5990	-0.5559	-0.7718	-0.6508	-0.6992	-0.6431
df		11	12	11	12	12	15	14	15	15
PE		3.1477	3.1250	3.2942	3.1797	3.2303	3.0261	3.1334	3.0809	3.1380

Data no.95: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.0256012.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0256	0.0256	0.0256	0.0256	0.0256	0.0256	0.0256	0.0256	0.0256
λ_1		0.2304	0.1024	0.0597	0.0384	0.0256	0.0171	0.0110	0.0064	0.0028
Predictors	x_1	1.8176	1.9610	2.0159	2.0567	2.0752	2.0833	2.0890	2.0932	2.0968
	x_2	-0	0	0.0320	0.0981	0.1273	0.1406	0.1500	0.1571	0.1629
	x_3	0	0	-0	-0	-0.0277	-0.0588	-0.0811	-0.0982	-0.1108
	x_4	0	0	0.0484	0.0909	0.0939	0.0859	0.0801	0.0753	0.0724
	x_5	-0.3159	-0.5694	-0.6160	-0.6135	-0.6341	-0.6597	-0.6779	-0.6919	-0.7020
	x_6	-0	0	0	0.0191	0.0506	0.0710	0.0858	0.0969	0.1056
	x_7	0.8255	1.0381	1.1138	1.1529	1.1673	1.1766	1.1833	1.1882	1.1920
	x_8	0.4356	0.4854	0.4993	0.5039	0.5143	0.5228	0.5287	0.5332	0.5366
	x_9	0	-0	-0	-0	-0.0149	-0.0287	-0.0385	-0.0459	-0.0516
	x_{10}	-0.1268	-0.2321	-0.2898	-0.3337	-0.3699	-0.3944	-0.4127	-0.4266	-0.4371
	x_{11}	0.7205	1.0486	1.1490	1.1987	1.2349	1.2604	1.2787	1.2926	1.3035
	x_{12}	0.3214	0.3703	0.3848	0.3951	0.4122	0.4271	0.4379	0.4462	0.4524
	x_{13}	0.0667	0.2477	0.2699	0.2682	0.2621	0.2580	0.2551	0.2528	0.2511
	x_{14}	0	0.0064	0.0670	0.1058	0.1362	0.1589	0.1753	0.1879	0.1974
	x_{15}	-0.0022	-0.4826	-0.6811	-0.8038	-0.8732	-0.9145	-0.9442	-0.9668	-0.9846
df		9	10	12	13	15	15	15	15	15
PE		4.2514	3.3103	3.1029	3.0116	2.9611	2.9325	2.9149	2.9032	2.8951

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.96: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0144	0.0324	0.0383	0.0596	0.0894	0.0439	0.1194	0.1548	0.1816
λ_1		0.1297	0.1297	0.0894	0.0894	0.0894	0.0293	0.0512	0.0387	0.0202
Predictors	x_1	2.3144	2.2540	2.3050	2.2371	2.1515	2.4263	2.1282	2.0587	2.0190
	x_2	0.3475	0.3638	0.3855	0.4009	0.4162	0.4690	0.4476	0.4633	0.4756
	x_3	0	0	0	0	-0	0.0905	-0	-0	-0.0097
	x_4	0.5044	0.4885	0.5256	0.5069	0.4829	0.6997	0.4974	0.4847	0.4811
	x_5	-0.1064	-0.0825	-0.1519	-0.1261	-0.1011	-0.1276	-0.1348	-0.1304	-0.1432
	x_6	-0.3448	-0.3537	-0.3651	-0.3750	-0.3851	-0.3634	-0.4069	-0.4192	-0.4308
	x_7	1.1480	1.1455	1.1256	1.1228	1.1144	1.1235	1.0965	1.0821	1.0705
	x_8	0.1147	0.1111	0.1461	0.1427	0.1372	0.1842	0.1674	0.1717	0.1815
	x_9	0.5364	0.5296	0.5794	0.5699	0.5571	0.6372	0.5893	0.5875	0.5968
	x_{10}	-0.4871	-0.4868	-0.5233	-0.5201	-0.5127	-0.5854	-0.5386	-0.5369	-0.5439
	x_{11}	0.8065	0.7572	0.8277	0.7751	0.7170	0.9477	0.7277	0.6944	0.6857
	x_{12}	-0	-0	-0	-0	-0	-0.1063	-0	-0	-0.0004
	x_{13}	0.2903	0.2873	0.3351	0.3284	0.3199	0.3866	0.3491	0.3489	0.3576
	x_{14}	-0.2646	-0.2267	-0.3173	-0.2734	-0.2227	-0.4502	-0.2547	-0.2297	-0.2269
	x_{15}	-0	-0	-0	-0	-0	-0.0146	-0	0	0
df		12	12	12	12	12	15	12	12	14
PE		4.6728	4.7479	4.5730	4.6541	4.7682	4.3358	4.7123	4.7884	4.8112

Data no.96: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.03697153.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0370	0.0370	0.0370	0.0370	0.0370	0.0370	0.0370	0.0370	0.0370
λ_1		0.3327	0.1479	0.0863	0.0555	0.0370	0.0246	0.0158	0.0092	0.0041
Predictors	x_1	1.9748	2.2080	2.3149	2.3676	2.4316	2.4816	2.5173	2.5441	2.5646
	x_2	0.2077	0.3597	0.3857	0.4101	0.4554	0.4847	0.5057	0.5215	0.5333
	x_3	-0	0	0	0	0.0762	0.1374	0.1811	0.2139	0.2388
	x_4	0.3047	0.4657	0.5301	0.5921	0.6858	0.7629	0.8181	0.8596	0.8909
	x_5	-0	-0.0430	-0.1597	-0.1914	-0.1423	-0.0884	-0.0499	-0.0209	-0
	x_6	-0.2465	-0.3516	-0.3651	-0.3564	-0.3569	-0.3601	-0.3624	-0.3641	-0.3653
	x_7	1.1146	1.1529	1.1243	1.1243	1.1259	1.1227	1.1204	1.1186	1.1172
	x_8	0	0.0938	0.1492	0.1572	0.1749	0.1942	0.2079	0.2182	0.2260
	x_9	0.2966	0.5044	0.5841	0.6167	0.6314	0.6442	0.6534	0.6603	0.6655
	x_{10}	-0.2387	-0.4697	-0.5262	-0.5637	-0.5802	-0.5875	-0.5927	-0.5966	-0.5997
	x_{11}	0.6104	0.7078	0.8379	0.9195	0.9527	0.9783	0.9966	1.0104	1.0216
	x_{12}	0	-0	-0	-0.0608	-0.1014	-0.1253	-0.1423	-0.1551	-0.1651
	x_{13}	0.1419	0.2644	0.3393	0.3769	0.3824	0.3859	0.3884	0.3903	0.3919
	x_{14}	-0	-0.1719	-0.3281	-0.3913	-0.4490	-0.4878	-0.5156	-0.5364	-0.5524
	x_{15}	0	-0	-0	-0	-0.0059	-0.0321	-0.0508	-0.0649	-0.0755
df		9	12	12	13	15	15	15	15	14
PE		5.7774	4.8738	4.5548	4.4208	4.3411	4.2909	4.2600	4.2394	4.2251

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.97: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0056	0.0614	0.0960	0.0232	0.0883	0.1207	0.1559	0.1842	0.1794
λ_1		0.0505	0.2458	0.2240	0.0348	0.0883	0.0805	0.0668	0.0461	0.0199
Predictors	x_1	3.1808	2.4698	2.4084	3.0756	2.6029	2.5166	2.4404	2.3971	2.4462
	x_2	1.3412	0.9161	0.8891	1.2411	0.9429	0.9071	0.8802	0.8682	0.8913
	x_3	1.2335	0.4037	0.3985	1.1228	0.6060	0.5667	0.5430	0.5441	0.6003
	x_4	1.3977	0.6206	0.6146	1.2359	0.7377	0.7045	0.6881	0.6929	0.7382
	x_5	-0.2004	-0.2965	-0.3335	-0.4067	-0.6475	-0.6545	-0.6604	-0.6771	-0.7175
	x_6	0.4940	0.3103	0.3160	0.4804	0.4015	0.3990	0.3989	0.4036	0.4179
	x_7	1.1349	1.0382	1.0340	1.1386	1.0846	1.0734	1.0625	1.0574	1.0711
	x_8	0.0355	0	0	0.0411	0	0	0	-0	-0
	x_9	0.71305	0.3771	0.3929	0.7141	0.5510	0.5499	0.5522	0.5620	0.5914
	x_{10}	0.2042	0	0	0.2158	0.1365	0.1389	0.1491	0.1699	0.2042
	x_{11}	0.7837	0.5908	0.5644	0.7697	0.6231	0.5899	0.5612	0.5447	0.5598
	x_{12}	0.0526	0	0	0.1178	0.0482	0.0797	0.1096	0.1361	0.1597
	x_{13}	0.0440	0.0542	0.1025	0.1155	0.1874	0.1999	0.2119	0.2244	0.2383
	x_{14}	-0.5163	-0	-0	-0.4835	-0.1419	-0.1252	-0.1212	-0.1369	-0.1912
	x_{15}	-0.3086	0	0	-0.2854	-0	-0	-0	-0	-0.0163
df		15	10	10	15	13	13	13	13	14
PE		5.9021	7.3285	7.3521	5.9289	6.6105	6.7088	6.7893	6.8059	6.6535

Data no.97: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.04859737.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0486	0.0486	0.0486	0.0486	0.0486	0.0486	0.0486	0.0486	0.0486
λ_1		0.4374	0.1944	0.1134	0.0729	0.0486	0.0324	0.0208	0.0121	0.0054
Predictors	x_1	2.2475	2.5620	2.6954	2.7854	2.8561	2.9078	2.9468	2.9761	2.9991
	x_2	0.8877	0.9254	0.9836	1.0424	1.0926	1.1234	1.1436	1.1588	1.1710
	x_3	0.1644	0.4641	0.6303	0.7629	0.8744	0.9483	1.0004	1.0396	1.0704
	x_4	0.5416	0.6130	0.7557	0.8870	0.9821	1.0435	1.0857	1.1175	1.1427
	x_5	-0	-0.4584	-0.6024	-0.6198	-0.5990	-0.5835	-0.5725	-0.5639	-0.5569
	x_6	0.1327	0.3517	0.3987	0.4205	0.4393	0.4551	0.4679	0.4776	0.4851
	x_7	0.9680	1.0630	1.0904	1.1098	1.1247	1.1320	1.1359	1.1389	1.1411
	x_8	-0	0	0	0	0	0.0173	0.0379	0.0533	0.0653
	x_9	0.1073	0.4459	0.5359	0.5859	0.6336	0.6730	0.7048	0.7287	0.7473
	x_{10}	0	0	0.1085	0.1709	0.2006	0.2168	0.2265	0.2338	0.2395
	x_{11}	0.5501	0.6316	0.6650	0.6861	0.7067	0.7232	0.7364	0.7463	0.7540
	x_{12}	0	0	0	0.0378	0.0949	0.1326	0.1598	0.1801	0.1957
	x_{13}	0	0.0810	0.1475	0.1714	0.1757	0.1767	0.1767	0.1766	0.1764
	x_{14}	0	-0	-0.1297	-0.2529	-0.3394	-0.3939	-0.4316	-0.4598	-0.4818
	x_{15}	0	0	-0	-0.0462	-0.1332	-0.1949	-0.2408	-0.2753	-0.3021
df		8	10	12	14	14	15	15	15	15
PE		8.3883	7.0411	6.5796	6.3267	6.1591	6.0635	6.0036	5.9638	5.9359

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.98: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0174	0.0224	0.0508	0.0867	0.1185	0.1952	0.1091	0.2053	0.2643
λ_1		0.1567	0.0897	0.1185	0.1301	0.1185	0.1301	0.0468	0.0513	0.0294
Predictors	x_1	1.6258	1.6720	1.6090	1.5457	1.5070	1.3958	1.5880	1.4615	1.4139
	x_2	0	0	0	0	0	0.0038	0.0484	0.0649	0.0840
	x_3	0.0424	0.0134	0.0878	0.1438	0.1751	0.2289	0.1709	0.2718	0.3084
	x_4	0	-0	0	0	0	0.0006	0	0.0684	0.0988
	x_5	-1.2141	-1.4916	-1.1847	-0.9680	-0.8850	-0.6466	-1.1735	-0.8239	-0.7644
	x_6	-0.1926	-0.2494	-0.2135	-0.1988	-0.2064	-0.1927	-0.2767	-0.2454	-0.2525
	x_7	0.9340	1.0149	0.9314	0.8579	0.8214	0.7140	0.9051	0.7919	0.7595
	x_8	-0.0365	-0.0423	-0.0644	-0.0927	-0.1188	-0.1540	-0.1344	-0.1782	-0.2045
	x_9	0	0	0	0.0211	0.0526	0.0878	0.0849	0.1378	0.1726
	x_{10}	0	0.0035	0	0	0	0	0.0827	0.0409	0.0671
	x_{11}	0.8238	0.9118	0.7913	0.7035	0.6647	0.5657	0.7450	0.6040	0.5574
	x_{12}	0	0	0	0	0	0	0.0374	0.0227	0.0379
	x_{13}	0.5061	0.5721	0.4907	0.4300	0.4036	0.3378	0.4514	0.3582	0.3263
	x_{14}	0.1990	0.2009	0.2282	0.2503	0.2646	0.2796	0.2673	0.2723	0.2727
	x_{15}	0	0	0	0	0	0	0.0021	0.0163	0.0358
df		9	10	9	10	10	12	14	15	15
PE		5.4217	5.2204	5.4297	5.6292	5.7141	6.0448	5.4087	5.7701	5.8759

Data no.98: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.05588364.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0559	0.0559	0.0559	0.0559	0.0559	0.0559	0.0559	0.0559	0.0559
λ_1		0.5030	0.2235	0.1304	0.0838	0.0559	0.0373	0.0240	0.0140	0.0062
Predictors	x_1	1.2271	1.5164	1.5922	1.6291	1.6558	1.6754	1.6785	1.6740	1.6751
	x_2	0	0	0	0	0.0144	0.0299	0.0275	0.0157	0.0157
	x_3	0.1559	0.1378	0.1012	0.0825	0.0753	0.0737	0.0463	0.0082	-0
	x_4	0	0	0	0	-0	-0	-0.0612	-0.1455	-0.1947
	x_5	-0	-0.7347	-1.1100	-1.2964	-1.4316	-1.5252	-1.6276	-1.7295	-1.7916
	x_6	-0	-0.1269	-0.2024	-0.2422	-0.2785	-0.3031	-0.3265	-0.3465	-0.3604
	x_7	0.3977	0.7880	0.9083	0.9647	0.9884	1.0062	1.0141	1.0194	1.0244
	x_8	-0.0213	-0.0615	-0.0674	-0.0732	-0.0856	-0.0930	-0.1012	-0.1098	-0.1132
	x_9	0	0	0	0.0071	0.0256	0.0386	0.0475	0.0548	0.0602
	x_{10}	-0	-0	0	0.0062	0.0769	0.1256	0.1639	0.1958	0.2194
	x_{11}	0.3991	0.6433	0.7641	0.8244	0.8551	0.8737	0.8901	0.9027	0.9097
	x_{12}	0	0	0	0	0.0307	0.0589	0.0871	0.1119	0.1292
	x_{13}	0.1599	0.3802	0.4715	0.5154	0.52879	0.5362	0.5466	0.5560	0.5601
	x_{14}	0.2250	0.2306	0.2319	0.2343	0.2410	0.2424	0.2529	0.2658	0.2724
	x_{15}	0	0	0	0	0	0	0.0109	0.0335	0.0469
df		7	9	9	11	13	13	15	15	15
PE		7.6275	5.9304	5.4961	5.3350	5.2356	5.1794	5.1342	5.0975	5.0767

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.99: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0210	0.0393	0.0614	0.1150	0.1084	0.1481	0.1318	0.1178	0.2006
λ_1		0.1894	0.1572	0.1433	0.1726	0.1084	0.0987	0.0565	0.0295	0.0223
Predictors	x_1	2.6130	2.5916	2.5467	2.3973	2.4601	2.3786	2.4455	2.4953	2.3196
	x_2	0	-0	-0	-0	-0	-0	-0	-0.0189	-0.0437
	x_3	-0	-0	-0	-0	-0.0141	-0.0107	-0.0842	-0.1369	-0.1047
	x_4	0.1079	0.1288	0.1427	0.1467	0.1657	0.1804	0.1676	0.1517	0.1828
	x_5	0	0	0	0	0	0	0	0	0
	x_6	-0.5526	-0.5796	-0.5776	-0.4980	-0.5814	-0.5634	-0.6200	-0.6584	-0.5986
	x_7	1.1073	1.1131	1.0977	1.0173	1.0867	1.0602	1.1222	1.1655	1.0875
	x_8	0.1743	0.2107	0.2146	0.1369	0.2348	0.2241	0.3052	0.3667	0.3257
	x_9	-0.2102	-0.2485	-0.2656	-0.2380	-0.3092	-0.3206	-0.3665	-0.3953	-0.3953
	x_{10}	0	0	0	0	0	0	0.0364	0.0655	0.0839
	x_{11}	0.1840	0.1970	0.2099	0.2252	0.2208	0.2265	0.2301	0.2348	0.2503
	x_{12}	0.2783	0.2854	0.2984	0.3238	0.3093	0.3154	0.3071	0.2985	0.3111
	x_{13}	0.6561	0.6473	0.6208	0.5303	0.5719	0.5294	0.5666	0.5982	0.5114
	x_{14}	0	0	0	0.0142	0.0345	0.0573	0.0701	0.0786	0.1091
	x_{15}	0	0	0	0	0	0	0	-0.0001	0
df		9	9	9	10	11	11	12	14	13
PE		4.2364	4.2009	4.2307	4.4995	4.2861	4.3983	4.2193	4.1131	4.3672

Data no.99: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.01981771.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0198	0.0198	0.0198	0.0198	0.0198	0.0198	0.0198	0.0198	0.0198
λ_1		0.1784	0.0793	0.0462	0.0297	0.0198	0.0132	0.0085	0.0050	0.0022
Predictors	x_1	2.6258	2.7163	2.7537	2.7766	2.7903	2.7992	2.8059	2.8105	2.8159
	x_2	0	-0	-0	-0	-0	0	0	0	0.0037
	x_3	-0	-0.1102	-0.1550	-0.1727	-0.1836	-0.1913	-0.1961	-0.2004	-0.2015
	x_4	0.1123	0.1105	0.1162	0.1315	0.1404	0.1457	0.1506	0.1532	0.1585
	x_5	0	0.0110	0.0350	0.0554	0.0673	0.0745	0.0808	0.0844	0.0913
	x_6	-0.5699	-0.7171	-0.7637	-0.7836	-0.7957	-0.8039	-0.8094	-0.8139	-0.8173
	x_7	1.1199	1.2310	1.2654	1.2787	1.2866	1.2921	1.2958	1.2987	1.3009
	x_8	0.1934	0.3622	0.4291	0.4774	0.5064	0.5258	0.5396	0.5500	0.5576
	x_9	-0.2221	-0.3261	-0.3610	-0.3814	-0.3936	-0.4018	-0.4076	-0.4119	-0.4157
	x_{10}	0	0	0.0216	0.0659	0.0925	0.1102	0.1229	0.1323	0.1398
	x_{11}	0.1836	0.1886	0.1995	0.2227	0.2366	0.2457	0.2525	0.2573	0.2599
	x_{12}	0.2738	0.2645	0.2606	0.2572	0.2553	0.2543	0.2532	0.2528	0.2516
	x_{13}	0.6679	0.7566	0.7866	0.8030	0.8130	0.8199	0.8244	0.8282	0.8302
	x_{14}	0	0	0.0041	0.0246	0.0369	0.0455	0.0511	0.0558	0.0584
	x_{15}	0	-0	-0.0248	-0.0853	-0.1217	-0.1459	-0.1632	-0.1762	-0.1864
df		9	11	14	14	14	14	14	14	15
PE		4.2015	3.9581	3.8949	3.8524	3.8335	3.8237	3.8180	3.8145	3.8122

Table I Naïve elastic net estimates where λ_2 are estimated by λ_2CV and λ_2BF (Cont.)

Data no.100: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2CV .										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0315	0.0536	0.0837	0.1081	0.1779	0.1840	0.2165	0.1935	0.2270
λ_1		0.2833	0.2143	0.1953	0.1621	0.1779	0.1226	0.0928	0.0484	0.0252
Predictors	x_1	1.8695	1.9066	1.8795	1.8746	1.7485	1.7910	1.7690	1.8599	1.8303
	x_2	0	0	0	0	0	0	0	0.0785	0.1042
	x_3	-0	-0	-0	-0	-0	-0	-0.0085	-0.0515	-0.0877
	x_4	0.3956	0.4402	0.4508	0.4731	0.4149	0.4628	0.4673	0.5167	0.5134
	x_5	0	0	0.0436	0.0910	0.1035	0.1489	0.1668	0.146	0.1545
	x_6	0	0	0	0	0	0	0.0198	0.1335	0.1521
	x_7	1.4292	1.4664	1.4313	1.4238	1.3092	1.3496	1.3315	1.3930	1.3638
	x_8	-0	-0	-0	-0	-0	-0	-0	-0.0259	-0.0415
	x_9	-0	-0	-0	-0	0	0	0	0	0
	x_{10}	0.0239	0.0770	0.1049	0.1337	0.1517	0.1832	0.2111	0.2582	0.2857
	x_{11}	0.7310	0.7108	0.6741	0.6417	0.5801	0.5715	0.5519	0.5782	0.5617
	x_{12}	0	0	0	0	0.0277	0.0296	0.0432	0.0488	0.0695
	x_{13}	0.1113	0.1575	0.1675	0.1754	0.1578	0.1714	0.1758	0.1942	0.1987
	x_{14}	0	0	0.0069	0.0227	0.0469	0.0506	0.0617	0.0715	0.0942
	x_{15}	0	0	0	0	0	0	0	-0	-0.0175
df		6	6	8	8	9	9	11	13	14
PE		5.8837	5.7677	5.7994	5.7890	6.0434	5.9246	5.9465	5.7236	5.7516

Data no.100: The naïve elastic net estimates ($\hat{\beta}$) using different α values; $\alpha = \lambda_2/(\lambda_1+\lambda_2)$, and λ_2BF = 0.0339025.										
α		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
λ_2		0.0339	0.0339	0.0339	0.0339	0.0339	0.0339	0.0339	0.0339	0.0339
λ_1		0.3051	0.1356	0.0791	0.0509	0.0339	0.0226	0.0145	0.0085	0.0038
Predictors	x_1	1.8423	2.0326	2.1422	2.2185	2.2651	2.3149	2.3551	2.4002	2.4388
	x_2	0	0	0.1326	0.2347	0.3042	0.3810	0.4432	0.5087	0.5641
	x_3	-0	-0	-0	0	0	0.0569	0.1112	0.1833	0.2466
	x_4	0.3762	0.5285	0.6186	0.7265	0.7924	0.8716	0.9367	1.0132	1.0791
	x_5	0	0	0.0069	0.0602	0.0952	0.1640	0.2244	0.3080	0.3819
	x_6	0	0.1146	0.2907	0.3875	0.4446	0.4946	0.5329	0.5606	0.5818
	x_7	1.4007	1.5805	1.6355	1.6535	1.6679	1.6756	1.6806	1.6873	1.6932
	x_8	-0	-0	-0.0352	-0.0435	-0.0471	-0.0472	-0.0467	-0.0486	-0.0507
	x_9	-0	-0	-0.0278	-0.0873	-0.1259	-0.1481	-0.1631	-0.1734	-0.1812
	x_{10}	0.0136	0.1213	0.2000	0.2477	0.2735	0.2941	0.3097	0.3188	0.3253
	x_{11}	0.7284	0.7446	0.7589	0.7927	0.8059	0.8052	0.8024	0.7919	0.7819
	x_{12}	0	0	0	0	0	0	-0	-0.0289	-0.0581
	x_{13}	0.1013	0.1953	0.2114	0.2385	0.2508	0.2495	0.2463	0.2439	0.2419
	x_{14}	0	0	0	0.0006	0.0276	0.0275	0.0231	0.0187	0.0151
	x_{15}	0	-0	-0	-0.0983	-0.1766	-0.2379	-0.2839	-0.3168	-0.3419
df		6	7	11	13	13	14	14	15	15
PE		5.9533	5.5154	5.3366	5.2010	5.1318	5.0835	5.0520	5.0235	5.0020

Appendix J

Bayesian variable selection for elastic net linear regression model

This section shows the result of the Bayesian variable selection for elastic net linear regression model with different priors.

(1) The Bayesian variable selection for elastic net linear regression model where the penalty parameters λ_1 and λ_2 are estimated by the 10-fold cross-validation method (BVSCV).

(2) The Bayesian variable selection for elastic net linear regression model where the penalty parameter λ_2 is based on Bayes factor (BVSBF). For BVSBF, the penalty parameter λ_1 is derived from the relationship between λ_1 and λ_2: $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$.

For the process of Bayesian variable selection for elastic net linear regression model studied in this thesis, we set $\alpha = 0.5$, i.e. $\lambda_1 = \lambda_2$.

The 100 simulation datasets using simulation method for dataset 1 are used for analysis the performance of the Bayesian variable selection for elastic net linear regression model. The dataset consists of 15 predictor variables of 50 observations each ($p = 15, n = 50$). The result of Bayesian variable selection for elastic net linear regression model is the optimal model which has highest posterior model probability.

BVSCV is compared to elastic net method where the penalty parameter λ_1 and λ_2 are estimated by the 10-fold cross-validation method (ENETCV). BVSBF is compared to elastic net method where the penalty parameter λ_2 is based on Bayes factor (ENETBF). Variable selection results of ENETCV and ENETBF using $\alpha = 0.5$ are derived from Appendix I.

Table J shows variable selection results of BVSCV, BVSBF, ENETCV, and ENETBF using $\alpha = 0.5$ for 100 simulation datasets.

Table J Variable selection results of BVSCV, BVSBF, and elastic net

Data no. 1															
Predictor variables [a]	x_1	x_{13}	x_{11}	x_{12}	x_{15}	x_{14}	x_4	x_6	x_8	x_2	x_7	x_3	x_5	x_9	x_{10}
Correlation coefficient between predictors	$corr(x_{12}, x_{15}) = 0.7530$ $corr(x_{14}, x_{15}) = 0.6100$ $corr(x_{12}, x_{14}) = 0.5425$														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{13}	x_{11}	x_{12}	-	-	x_4	x_6	x_8	-	x_7	-	-	-	-
ENETCV	x_1	x_{13}	x_{11}	x_{12}	-	-	x_4	x_6	x_8	x_2	x_7	-	x_5	x_9	-
λ_2 based on Bayes factor															
BVSBF	x_1	x_{13}	x_{11}	x_{12}	x_{15}	x_{14}	x_4	x_6	x_8	x_2	x_7	-	-	-	-
ENETBF	x_1	x_{13}	x_{11}	x_{12}	-	x_{14}	x_4	x_6	x_8	-	x_7	x_3	x_5	x_9	x_{10}

Data no. 2															
Predictor variables [a]	x_{13}	x_7	x_1	x_{14}	x_3	x_{11}	x_{12}	x_8	x_6	x_4	x_{15}	x_9	x_2	x_{10}	x_5
Correlation coefficient between predictors	$corr(x_{14}, x_{15}) = 0.7337$ $corr(x_{12}, x_{15}) = 0.7280$ $corr(x_{12}, x_{14}) = 0.6746$														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{13}	x_7	x_1	x_{14}	x_3	x_{11}	x_{12}	x_8	x_6	x_4	-	x_9	-	-	-
ENETCV	x_{13}	x_7	x_1	x_{14}	x_3	x_{11}	-	x_8	x_6	x_4	x_{15}	x_9	-	x_{10}	x_5
λ_2 based on Bayes factor															
BVSBF	x_{13}	x_7	x_1	x_{14}	x_3	x_{11}	x_{12}	x_8	x_6	x_4	x_{15}	x_9	-	-	-
ENETBF	x_{13}	x_7	x_1	x_{14}	x_3	x_{11}	x_{12}	x_8	x_6	-	x_{15}	x_9	x_2	x_{10}	x_5

Data no. 3															
Predictor variables [a]	x_1	x_{11}	x_{15}	x_{13}	x_4	x_{12}	x_{14}	x_3	x_7	x_{10}	x_6	x_9	x_2	x_5	x_8
Correlation coefficient between predictors	$corr(x_{12}, x_{14}) = 0.7544$ $corr(x_{12}, x_{15}) = 0.7191$ $corr(x_{14}, x_{15}) = 0.7188$														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{15}	x_{13}	x_4	x_{12}	-	x_3	x_7	-	-	-	-	-	-
ENETCV	x_1	x_{11}	x_{15}	x_{13}	x_4	-	x_{14}	x_3	x_7	x_{10}	x_6	-	x_2	x_5	x_8
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{15}	x_{13}	x_4	x_{12}	x_{14}	x_3	x_7	-	-	-	-	-	-
ENETBF	x_1	x_{11}	x_{15}	x_{13}	x_4	-	x_{14}	x_3	x_7	x_{10}	x_6	-	x_2	x_5	x_8

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 4															
Predictor variables [a]	x_{11}	x_{13}	x_1	x_4	x_{12}	x_{14}	x_{15}	x_7	x_{10}	x_2	x_6	x_8	x_3	x_5	x_9
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7884 $corr(x_{14}, x_{15})$ =0.7574 $corr(x_{12}, x_{14})$ =0.7507														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{13}	x_1	x_4	x_{12}	x_{14}	-	x_7	-	-	x_6	x_8	-	-	-
ENETCV	x_{11}	x_{13}	x_1	x_4	-	-	x_{15}	x_7	x_{10}	-	x_6	x_8	-	x_5	-
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{13}	x_1	x_4	x_{12}	x_{14}	x_{15}	x_7	x_{10}	x_2	x_6	x_8	-	-	-
ENETBF	x_{11}	x_{13}	x_1	x_4	x_{12}	x_{14}	x_{15}	x_7	x_{10}	x_2	x_6	x_8	x_3	x_5	x_9

Data no. 5															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{12}	x_{15}	x_7	x_{14}	x_5	x_2	x_3	x_8	x_6	x_4	x_{10}	x_9
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7296 $corr(x_{12}, x_{14})$ =0.7119 $corr(x_{12}, x_{15})$ =0.6909														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{13}	x_{12}	x_{15}	x_7	x_{14}	x_5	x_2	x_3	x_8	x_6	-	-	-
ENETCV	x_1	x_{11}	x_{13}	x_{12}	x_{15}	x_7	x_{14}	x_5	x_2	x_3	x_8	x_6	x_4	x_{10}	x_9
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{13}	x_{12}	x_{15}	x_7	x_{14}	x_5	x_2	x_3	x_8	x_6	-	-	-
ENETBF	x_1	x_{11}	x_{13}	x_{12}	x_{15}	x_7	x_{14}	x_5	x_2	x_3	x_8	x_6	x_4	x_{10}	x_9

Data no. 6															
Predictor variables [a]	x_7	x_1	x_{11}	x_2	x_{12}	x_{14}	x_8	x_{10}	x_{15}	x_5	x_{13}	x_4	x_9	x_6	x_3
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7376 $corr(x_{12}, x_{14})$ =0.7117 $corr(x_{14}, x_{15})$ =0.6714														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_7	x_1	x_{11}	x_2	x_{12}	-	-	x_{10}	-	-	-	-	-	-	-
ENETCV	x_7	x_1	x_{11}	x_2	-	x_{14}	x_8	x_{10}	-	-	-	x_4	x_9	-	-
λ_2 based on Bayes factor															
BVSBF	x_7	x_1	x_{11}	x_2	x_{12}	x_{14}	-	x_{10}	-	-	-	-	-	-	-
ENETBF	x_7	x_1	x_{11}	x_2	-	x_{14}	x_8	x_{10}	x_{15}	-	x_{13}	x_4	x_9	-	-

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 7															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{12}	x_{14}	x_{15}	x_7	x_5	x_2	x_3	x_4	x_{10}	x_9	x_8	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7225 $corr(x_{14}, x_{15})$ =0.6714 $corr(x_{12}, x_{15})$ =0.6629														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{13}	x_{12}	x_{14}	x_{15}	x_7	-	-	-	-	-	-	-	-
ENETCV	x_1	x_{11}	x_{13}	x_{12}	x_{14}	-	x_7	x_5	-	-	-	x_{10}	-	x_8	x_6
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{13}	x_{12}	x_{14}	x_{15}	x_7	x_5	x_2	-	-	-	-	-	-
ENETBF	x_1	x_{11}	x_{13}	x_{12}	x_{14}	-	x_7	x_5	x_2	x_3	x_4	x_{10}	x_9	x_8	x_6

Data no. 8															
Predictor variables [a]	x_{11}	x_1	x_{14}	x_7	x_{13}	x_{15}	x_{12}	x_5	x_4	x_8	x_6	x_{10}	x_9	x_3	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7169 $corr(x_{14}, x_{15})$ =0.7065 $corr(x_{12}, x_{14})$ =0.6721														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_1	x_{14}	x_7	x_{13}	x_{15}	x_{12}	x_5	-	x_8	x_6	-	-	-	-
ENETCV	x_{11}	x_1	x_{14}	x_7	x_{13}	x_{15}	x_{12}	x_5	x_4	x_8	x_6	-	x_9	x_3	-
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_1	x_{14}	x_7	x_{13}	x_{15}	x_{12}	x_5	x_4	x_8	x_6	-	-	-	-
ENETBF	x_{11}	x_1	x_{14}	x_7	x_{13}	x_{15}	x_{12}	x_5	x_4	x_8	x_6	-	x_9	-	x_2

Data no. 9															
Predictor variables [a]	x_1	x_{13}	x_{14}	x_{15}	x_{12}	x_3	x_{11}	x_8	x_2	x_7	x_4	x_{10}	x_6	x_5	x_9
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7562 $corr(x_{12}, x_{15})$ =0.6433 $corr(x_{12}, x_{14})$ =0.6390														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{13}	x_{14}	x_{15}	x_{12}	x_3	x_{11}	x_8	x_2	x_7	-	x_{10}	-	x_5	-
ENETCV	x_1	x_{13}	-	x_{15}	x_{12}	x_3	x_{11}	x_8	x_2	x_7	-	x_{10}	-	x_5	-
λ_2 based on Bayes factor															
BVSBF	x_1	x_{13}	x_{14}	x_{15}	x_{12}	x_3	x_{11}	x_8	x_2	x_7	x_4	-	-	-	-
ENETBF	x_1	x_{13}	x_{14}	x_{15}	x_{12}	-	x_{11}	x_8	x_2	x_7	x_4	x_{10}	x_6	x_5	x_9

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 10															
Predictor variables [a]	x_{11}	x_{14}	x_{13}	x_{15}	x_{12}	x_2	x_1	x_4	x_9	x_5	x_3	x_{10}	x_6	x_8	x_7
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.8132 $corr(x_{12}, x_{15})$ =0.7618 $corr(x_{12}, x_{14})$ =0.7560														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{14}	x_{13}	x_{15}	x_{12}	x_2	x_1	x_4	x_9	-	x_3	x_{10}	-	-	-
ENETCV	x_{11}	x_{14}	-	-	x_{12}	x_2	x_1	x_4	-	x_5	-	x_{10}	x_6	x_8	x_7
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{14}	x_{13}	x_{15}	x_{12}	x_2	x_1	x_4	x_9	x_5	x_3	x_{10}	-	-	-
ENETBF	x_{11}	x_{14}	-	x_{15}	x_{12}	x_2	x_1	x_4	-	x_5	-	x_{10}	x_6	x_8	x_7

Data no. 11															
Predictor variables [a]	x_{11}	x_1	x_{15}	x_{14}	x_{13}	x_7	x_4	x_{12}	x_6	x_5	x_9	x_3	x_8	x_2	x_{10}
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7153 $corr(x_{12}, x_{14})$ =0.6291 $corr(x_{12}, x_{15})$ =0.6567														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_1	x_{15}	x_{14}	x_{13}	x_7	x_4	-	x_6	-	-	-	-	-	-
ENETCV	x_{11}	x_1	-	x_{14}	x_{13}	x_7	x_4	x_{12}	-	x_5	-	x_3	-	x_2	-
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_1	x_{15}	x_{14}	x_{13}	x_7	x_4	x_{12}	x_6	x_5	-	-	-	-	-
ENETBF	x_{11}	x_1	x_{15}	x_{14}	x_{13}	x_7	x_4	x_{12}	-	x_5	x_9	x_3	x_8	x_2	x_{10}

Data no. 12															
Predictor variables [a]	x_{11}	x_{13}	x_{14}	x_7	x_4	x_1	x_8	x_{15}	x_{10}	x_{12}	x_2	x_6	x_3	x_9	x_5
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.6894 $corr(x_{12}, x_{15})$ =0.6888 $corr(x_{12}, x_{14})$ =0.6846														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{13}	x_{14}	x_7	x_4	x_1	x_8	-	-	-	x_2	-	-	-	-
ENETCV	x_{11}	x_{13}	x_{14}	x_7	x_4	x_1	x_8	x_{15}	-	x_{12}	x_2	x_6	-	-	x_5
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{13}	x_{14}	x_7	x_4	x_1	x_8	x_{15}	x_{10}	-	x_2	-	-	-	-
ENETBF	x_{11}	x_{13}	x_{14}	x_7	x_4	x_1	x_8	x_{15}	x_{10}	x_{12}	x_2	-	x_3	x_9	x_5

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 13															
Predictor variables [a]	x_{11}	x_1	x_{15}	x_{13}	x_{14}	x_4	x_{12}	x_8	x_3	x_2	x_5	x_9	x_{10}	x_6	x_7
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.8168 $corr(x_{12}, x_{14})$ =0.7439 $corr(x_{14}, x_{15})$ =0.7572														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_1	x_{15}	x_{13}	x_{14}	x_4	x_{12}	x_8	x_3	-	-	-	-	-	-
ENETCV	x_{11}	x_1	-	x_{13}	x_{14}	x_4	x_{12}	x_8	x_3	x_2	x_5	-	x_{10}	-	x_7
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_1	x_{15}	x_{13}	x_{14}	x_4	x_{12}	x_8	x_3	x_2	-	-	-	-	-
ENETBF	x_{11}	x_1	x_{15}	x_{13}	x_{14}	x_4	x_{12}	x_8	x_3	x_2	x_5	-	x_{10}	-	x_7

Data no. 14															
Predictor variables [a]	x_1	x_{11}	x_{15}	x_{14}	x_{13}	x_{12}	x_4	x_2	x_3	x_7	x_5	x_9	x_8	x_{10}	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.8035 $corr(x_{14}, x_{15})$ =0.7931 $corr(x_{12}, x_{14})$ =0.7486														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{15}	x_{14}	x_{13}	x_{12}	x_4	x_2	x_3	x_7	-	-	x_8	-	-
ENETCV	x_1	x_{11}	x_{15}	x_{14}	-	-	x_4	-	x_3	x_7	x_5	x_9	x_8	-	x_6
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{15}	x_{14}	x_{13}	x_{12}	x_4	x_2	x_3	x_7	x_5	-	-	-	-
ENETBF	x_1	x_{11}	x_{15}	x_{14}	x_{13}	x_{12}	x_4	-	x_3	x_7	x_5	x_9	x_8	-	x_6

Data no. 15															
Predictor variables [a]	x_1	x_{13}	x_{11}	x_{12}	x_{15}	x_{14}	x_7	x_8	x_4	x_3	x_2	x_5	x_6	x_{10}	x_9
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.8005 $corr(x_{14}, x_{15})$ =0.7542 $corr(x_{12}, x_{14})$ =0.6973														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{13}	x_{11}	x_{12}	x_{15}	x_{14}	x_7	x_8	x_4	x_3	x_2	-	-	-	-
ENETCV	x_1	x_{13}	x_{11}	x_{12}	-	x_{14}	x_7	x_8	-	x_3	x_2	-	x_6	x_{10}	x_9
λ_2 based on Bayes factor															
BVSBF	x_1	x_{13}	x_{11}	x_{12}	x_{15}	x_{14}	x_7	x_8	x_4	x_3	x_2	-	-	-	-
ENETBF	x_1	x_{13}	x_{11}	x_{12}	-	x_{14}	x_7	x_8	-	x_3	x_2	-	x_6	x_{10}	x_9

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 16															
Predictor variables [a]	x_7	x_{13}	x_{11}	x_1	x_{15}	x_{14}	x_{12}	x_5	x_6	x_4	x_3	x_9	x_2	x_8	x_{10}
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7595 $corr(x_{12}, x_{15})$ =0.7442 $corr(x_{12}, x_{14})$ =0.6979														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_7	x_{13}	x_{11}	x_1	x_{15}	x_{14}	x_{12}	x_5	x_6	x_4	-	-	-	-	-
ENETCV	x_7	x_{13}	x_{11}	x_1	-	x_{14}	-	x_5	x_6	x_4	x_3	x_9	-	x_8	x_{10}
λ_2 based on Bayes factor															
BVSBF	x_7	x_{13}	x_{11}	x_1	x_{15}	x_{14}	x_{12}	x_5	x_6	x_4	x_3	-	-	-	-
ENETBF	x_7	x_{13}	x_{11}	x_1	x_{15}	x_{14}	x_{12}	x_5	x_6	x_4	-	x_9	x_2	x_8	x_{10}

Data no. 17															
Predictor variables [a]	x_{11}	x_{14}	x_{15}	x_{13}	x_1	x_{12}	x_7	x_4	x_3	x_2	x_5	x_6	x_8	x_{10}	x_9
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7760 $corr(x_{14}, x_{15})$ =0.7302 $corr(x_{12}, x_{15})$ =0.7028														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{14}	x_{15}	x_{13}	x_1	x_{12}	x_7	x_4	x_3	-	-	x_6	-	-	-
ENETCV	x_{11}	x_{14}	x_{15}	x_{13}	x_1	-	x_7	x_4	x_3	-	x_5	x_6	-	-	-
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{14}	x_{15}	x_{13}	x_1	x_{12}	x_7	x_4	x_3	x_2	x_5	x_6	x_8	-	-
ENETBF	x_{11}	x_{14}	x_{15}	x_{13}	x_1	x_{12}	x_7	x_4	-	x_2	x_5	x_6	x_8	x_{10}	x_9

Data no. 18															
Predictor variables [a]	x_1	x_{11}	x_{14}	x_{13}	x_{15}	x_{12}	x_4	x_7	x_3	x_2	x_8	x_6	x_5	x_{10}	x_9
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7072 $corr(x_{12}, x_{15})$ =0.6864 $corr(x_{14}, x_{15})$ =0.6581														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{14}	x_{13}	x_{15}	x_{12}	x_4	x_7	x_3	x_2	-	-	-	-	-
ENETCV	x_1	x_{11}	x_{14}	x_{13}	x_{15}	-	-	x_7	x_3	x_2	x_8	x_6	x_5	x_{10}	x_9
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{14}	x_{13}	x_{15}	x_{12}	x_4	x_7	x_3	x_2	-	-	-	-	-
ENETBF	x_1	x_{11}	x_{14}	x_{13}	x_{15}	x_{12}	x_4	x_7	x_3	x_2	x_8	x_6	x_5	x_{10}	x_9

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 19															
Predictor variables [a]	x_{11}	x_1	x_{14}	x_{15}	x_{13}	x_{12}	x_{10}	x_2	x_3	x_5	x_6	x_4	x_7	x_8	x_9
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7004 $corr(x_{12}, x_{14})$ =0.6292 $corr(x_{12}, x_{15})$ =0.5435														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_1	x_{14}	x_{15}	-	-	x_{10}	-	x_3	-	-	-	x_7	-	-
ENETCV	x_{11}	x_1	x_{14}	x_{15}	-	-	x_{10}	x_2	x_3	-	x_6	-	x_7	x_8	-
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_1	x_{14}	x_{15}	x_{13}	x_{12}	x_{10}	x_2	x_3	x_5	x_6	-	-	-	-
ENETBF	x_{11}	x_1	x_{14}	x_{15}	x_{13}	x_{12}	x_{10}	x_2	x_3	x_5	x_6	x_4	x_7	x_8	x_9

Data no. 20															
Predictor variables [a]	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_1	x_3	x_6	x_4	x_5	x_7	x_8	x_9	x_{10}	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.6477 $corr(x_{14}, x_{15})$ =0.5955 $corr(x_{12}, x_{15})$ =0.4832														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_1	x_3	x_6	-	-	-	-	-	-	-
ENETCV	x_{11}	x_{13}	-	-	x_{12}	x_1	x_3	x_6	-	x_5	x_7	-	x_9	-	x_2
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_1	x_3	x_6	x_4	-	-	-	-	-	-
ENETBF	x_{11}	x_{13}	-	x_{14}	x_{12}	x_1	x_3	x_6	x_4	x_5	x_7	x_8	x_9	x_{10}	x_2

Data no. 21															
Predictor variables [a]	x_{11}	x_{13}	x_{14}	x_{12}	x_1	x_{15}	x_5	x_3	x_7	x_8	x_2	x_4	x_9	x_{10}	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.8185 $corr(x_{14}, x_{15})$ =0.7658 $corr(x_{12}, x_{15})$ =0.7175														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{13}	x_{14}	x_{12}	x_1	x_{15}	x_5	x_3	x_7	x_8	x_2	x_4	x_9	-	-
ENETCV	x_{11}	x_{13}	x_{14}	x_{12}	x_1	x_{15}	x_5	x_3	x_7	x_8	-	-	x_9	x_{10}	x_6
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{13}	x_{14}	x_{12}	x_1	x_{15}	x_5	x_3	x_7	x_8	x_2	x_4	x_9	-	-
ENETBF	x_{11}	x_{13}	x_{14}	x_{12}	x_1	x_{15}	x_5	x_3	x_7	x_8	-	-	x_9	x_{10}	x_6

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 22															
Predictor variables [a]	x_1	x_{13}	x_{15}	x_{14}	x_{12}	x_{11}	x_6	x_3	x_2	x_4	x_7	x_5	x_8	x_9	x_{10}
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.8428 $corr(x_{14}, x_{15})$ =0.7905 $corr(x_{12}, x_{15})$ =0.7681														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{13}	x_{15}	x_{14}	x_{12}	x_{11}	x_6	x_3	x_2	-	x_7	-	x_8	-	-
ENETCV	x_1	x_{13}	-	-	x_{12}	x_{11}	x_6	x_3	x_2	x_4	x_7	x_5	-	-	x_{10}
λ_2 based on Bayes factor															
BVSBF	x_1	x_{13}	x_{15}	x_{14}	x_{12}	x_{11}	x_6	x_3	x_2	x_4	x_7	x_5	x_8	x_9	-
ENETBF	x_1	x_{13}	x_{15}	-	x_{12}	x_{11}	x_6	x_3	x_2	x_4	x_7	x_5	x_8	x_9	x_{10}

Data no. 23															
Predictor variables [a]	x_{11}	x_{13}	x_{15}	x_1	x_7	x_{14}	x_{12}	x_3	x_4	x_5	x_9	x_8	x_{10}	x_2	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.8038 $corr(x_{14}, x_{15})$ =0.7977 $corr(x_{12}, x_{14})$ =0.7729														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{13}	x_{15}	x_1	x_7	x_{14}	x_{12}	x_3	x_4	-	-	-	-	-	-
ENETCV	x_{11}	x_{13}	-	x_1	x_7	-	-	x_3	-	x_5	x_9	x_8	x_{10}	x_2	x_6
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{13}	x_{15}	x_1	x_7	x_{14}	x_{12}	x_3	x_4	x_5	x_9	-	-	-	-
ENETBF	x_{11}	x_{13}	x_{15}	x_1	x_7	x_{14}	-	x_3	x_4	x_5	x_9	x_8	-	x_2	x_6

Data no. 24															
Predictor variables [a]	x_1	x_{11}	x_{14}	x_{13}	x_{15}	x_4	x_3	x_{12}	x_7	x_5	x_2	x_{10}	x_8	x_6	x_9
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7738 $corr(x_{12}, x_{14})$ =0.7706 $corr(x_{12}, x_{15})$ =0.7550														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{14}	x_{13}	x_{15}	x_4	x_3	x_{12}	x_7	x_5	-	-	-	-	-
ENETCV	x_1	x_{11}	x_{14}	-	-	x_4	x_3	x_{12}	x_7	x_5	x_2	x_{10}	x_8	x_6	x_9
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{14}	x_{13}	x_{15}	x_4	x_3	x_{12}	x_7	x_5	-	-	-	-	-
ENETBF	x_1	x_{11}	x_{14}	-	-	x_4	x_3	x_{12}	x_7	x_5	x_2	x_{10}	x_8	x_6	x_9

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 25															
Predictor variables [a]	x_1	x_{11}	x_{15}	x_4	x_{12}	x_{13}	x_7	x_{14}	x_{10}	x_6	x_3	x_5	x_9	x_8	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7877 $corr(x_{12}, x_{14})$ =0.7788 $corr(x_{14}, x_{15})$ =0.7505														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{15}	x_4	x_{12}	x_{13}	x_7	x_{14}	x_{10}	-	-	-	-	-	-
ENETCV	x_1	x_{11}	x_{15}	x_4	-	x_{13}	x_7	-	x_{10}	x_6	x_3	-	-	x_8	x_2
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{15}	x_4	x_{12}	x_{13}	x_7	x_{14}	x_{10}	x_6	x_3	-	x_9	-	-
ENETBF	x_1	x_{11}	x_{15}	x_4	-	x_{13}	x_7	-	x_{10}	x_6	x_3	x_5	-	x_8	x_2

Data no. 26															
Predictor variables [a]	x_{11}	x_{13}	x_{15}	x_1	x_{12}	x_{14}	x_2	x_4	x_7	x_5	x_3	x_8	x_9	x_{10}	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.8215 $corr(x_{14}, x_{15})$ =0.8207 $corr(x_{12}, x_{15})$ =0.7394														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{13}	x_{15}	x_1	x_{12}	x_{14}	x_2	x_4	x_7	-	-	x_8	x_9	-	-
ENETCV	x_{11}	x_{13}	x_{15}	x_1	x_{12}	x_{14}	x_2	x_4	x_7	x_5	x_3	x_8	x_9	x_{10}	x_6
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{13}	x_{15}	x_1	x_{12}	x_{14}	x_2	x_4	x_7	x_5	x_3	x_8	x_9	-	-
ENETBF	x_{11}	x_{13}	x_{15}	x_1	x_{12}	x_{14}	x_2	x_4	x_7	x_5	x_3	x_8	x_9	x_{10}	x_6

Data no. 27															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_2	x_4	x_9	x_5	x_3	x_{10}	x_6	x_7	x_8
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.8585 $corr(x_{14}, x_{15})$ =0.8440 $corr(x_{12}, x_{14})$ =0.7746														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_2	x_4	x_9	x_5	x_3	x_{10}	-	-	-
ENETCV	x_1	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_2	x_4	x_9	x_5	x_3	x_{10}	x_6	x_7	x_8
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_2	x_4	x_9	x_5	x_3	x_{10}	-	-	-
ENETBF	x_1	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_2	x_4	x_9	x_5	x_3	x_{10}	x_6	x_7	x_8

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 28															
Predictor variables [a]	x_{11}	x_{13}	x_{12}	x_1	x_{14}	x_{15}	x_7	x_4	x_5	x_{10}	x_2	x_6	x_3	x_9	x_8
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7142 $corr(x_{14}, x_{15})$ =0.6944 $corr(x_{12}, x_{15})$ =0.6854														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{13}	x_{12}	x_1	x_{14}	x_{15}	x_7	x_4	-	x_{10}	-	x_6	-	-	-
ENETCV	x_{11}	x_{13}	x_{12}	x_1	x_{14}	-	x_7	x_4	-	x_{10}	-	x_6	x_3	-	x_8
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{13}	x_{12}	x_1	x_{14}	x_{15}	x_7	x_4	x_5	x_{10}	x_2	-	-	-	-
ENETBF	x_{11}	x_{13}	x_{12}	x_1	-	x_{15}	x_7	x_4	x_5	x_{10}	x_2	x_6	x_3	x_9	x_8

Data no. 29															
Predictor variables [a]	x_{11}	x_{14}	x_{15}	x_{12}	x_{13}	x_1	x_3	x_2	x_4	x_7	x_5	x_9	x_6	x_{10}	x_8
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7813 $corr(x_{12}, x_{15})$ =0.7812 $corr(x_{14}, x_{15})$ =0.7141														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{14}	x_{15}	x_{12}	x_{13}	x_1	x_3	x_2	x_4	x_7	-	-	-	-	-
ENETCV	x_{11}	-	x_{15}	-	-	x_1	x_3	x_2	x_4	x_7	x_5	x_9	x_6	x_{10}	-
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{14}	x_{15}	x_{12}	x_{13}	x_1	x_3	x_2	x_4	x_7	x_5	-	-	-	-
ENETBF	x_{11}	-	x_{15}	-	-	x_1	x_3	-	x_4	x_7	x_5	x_9	x_6	x_{10}	-

Data no. 30															
Predictor variables [a]	x_1	x_{11}	x_{14}	x_{13}	x_{15}	x_2	x_7	x_{12}	x_9	x_4	x_{10}	x_3	x_8	x_5	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7804 $corr(x_{14}, x_{15})$ =0.7352 $corr(x_{12}, x_{14})$ =0.6543														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{14}	x_{13}	x_{15}	x_2	x_7	x_{12}	x_9	x_4	-	-	-	-	-
ENETCV	x_1	x_{11}	x_{14}	-	x_{15}	x_2	x_7	x_{12}	x_9	x_4	x_{10}	x_3	x_8	x_5	x_6
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{14}	x_{13}	x_{15}	x_2	x_7	x_{12}	x_9	x_4	-	-	-	-	-
ENETBF	x_1	x_{11}	x_{14}	-	x_{15}	x_2	x_7	x_{12}	x_9	x_4	x_{10}	x_3	x_8	x_5	x_6

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 31															
Predictor variables [a]	x_1	x_{13}	x_{11}	x_{14}	x_7	x_{15}	x_{12}	x_2	x_3	x_5	x_9	x_{10}	x_6	x_4	x_8
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.5837 $corr(x_{12}, x_{15})$ =0.5790 $corr(x_{14}, x_{15})$ =0.5511														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{13}	x_{11}	x_{14}	x_7	x_{15}	x_{12}	x_2	-	-	-	-	-	-	-
ENETCV	x_1	x_{13}	x_{11}	x_{14}	x_7	-	-	x_2	-	x_5	x_9	x_{10}	x_6	x_4	-
λ_2 based on Bayes factor															
BVSBF	x_1	x_{13}	x_{11}	x_{14}	x_7	x_{15}	x_{12}	x_2	x_3	x_5	x_9	x_{10}	-	-	-
ENETBF	x_1	x_{13}	x_{11}	x_{14}	x_7	x_{15}	x_{12}	x_2	x_3	x_5	x_9	x_{10}	x_6	x_4	x_8

Data no. 32															
Predictor variables [a]	x_1	x_{13}	x_{11}	x_{14}	x_{12}	x_{15}	x_4	x_6	x_9	x_8	x_3	x_5	x_{10}	x_7	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7463 $corr(x_{12}, x_{15})$ =0.6610 $corr(x_{14}, x_{15})$ =0.5800														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{13}	x_{11}	x_{14}	x_{12}	x_{15}	x_4	-	-	-	-	-	-	-	-
ENETCV	x_1	x_{13}	x_{11}	x_{14}	-	x_{15}	x_4	-	-	-	-	-	x_{10}	x_7	x_2
λ_2 based on Bayes factor															
BVSBF	x_1	x_{13}	x_{11}	x_{14}	x_{12}	x_{15}	x_4	-	-	-	-	-	-	-	-
ENETBF	x_1	x_{13}	x_{11}	x_{14}	x_{12}	x_{15}	-	-	x_9	-	x_3	x_5	x_{10}	x_7	x_2

Data no. 33															
Predictor variables [a]	x_1	x_{11}	x_{12}	x_{14}	x_{13}	x_{15}	x_7	x_3	x_4	x_5	x_8	x_9	x_{10}	x_6	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.8258 $corr(x_{14}, x_{15})$ =0.8239 $corr(x_{12}, x_{14})$ =0.8048														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{12}	x_{14}	x_{13}	x_{15}	x_7	x_3	x_4	x_5	x_8	-	-	-	-
ENETCV	x_1	x_{11}	x_{12}	-	x_{13}	-	x_7	x_3	x_4	x_5	x_8	x_9	x_{10}	x_6	x_2
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{12}	x_{14}	x_{13}	x_{15}	x_7	x_3	x_4	x_5	x_8	x_9	-	-	-
ENETBF	x_1	x_{11}	x_{12}	-	x_{13}	x_{15}	x_7	-	x_4	x_5	x_8	x_9	x_{10}	x_6	x_2

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 34															
Predictor variables [a]	x_{14}	x_{11}	x_1	x_{15}	x_{13}	x_4	x_{12}	x_2	x_6	x_5	x_{10}	x_9	x_7	x_3	x_8
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7746 $corr(x_{12}, x_{14})$ =0.7581 $corr(x_{12}, x_{15})$ =0.7352														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{14}	x_{11}	x_1	x_{15}	-	x_4	-	x_2	x_6	-	-	-	x_7	-	-
ENETCV	x_{14}	x_{11}	x_1	x_{15}	-	x_4	-	x_2	x_6	x_5	x_{10}	x_9	x_7	-	x_8
λ_2 based on Bayes factor															
BVSBF	x_{14}	x_{11}	x_1	x_{15}	x_{13}	x_4	x_{12}	x_2	x_6	-	-	-	-	-	-
ENETBF	x_{14}	x_{11}	x_1	x_{15}	x_{13}	x_4	x_{12}	x_2	x_6	x_5	x_{10}	x_9	x_7	x_3	x_8

Data no. 35															
Predictor variables [a]	x_7	x_{11}	x_{13}	x_{12}	x_1	x_{14}	x_3	x_{15}	x_9	x_4	x_{10}	x_2	x_6	x_8	x_5
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7046 $corr(x_{12}, x_{15})$ =0.6840 $corr(x_{14}, x_{15})$ =0.6423														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_7	x_{11}	x_{13}	x_{12}	x_1	x_{14}	x_3	x_{15}	x_9	x_4	x_{10}	-	-	-	-
ENETCV	x_7	x_{11}	x_{13}	x_{12}	x_1	-	x_3	x_{15}	x_9	x_4	-	x_2	x_6	-	x_5
λ_2 based on Bayes factor															
BVSBF	x_7	x_{11}	x_{13}	x_{12}	x_1	x_{14}	x_3	x_{15}	x_9	x_4	x_{10}	-	-	-	-
ENETBF	x_7	x_{11}	x_{13}	x_{12}	x_1	x_{14}	x_3	x_{15}	x_9	x_4	x_{10}	x_2	x_6	-	x_5

Data no. 36															
Predictor variables [a]	x_1	x_{11}	x_{14}	x_{15}	x_{12}	x_{13}	x_4	x_{10}	x_2	x_9	x_7	x_5	x_8	x_3	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.6461 $corr(x_{11}, x_{14})$ =0.5862 $corr(x_{12}, x_{14})$ =0.5361														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{14}	x_{15}	x_{12}	x_{13}	x_4	x_{10}	-	-	-	-	-	-	-
ENETCV	x_1	x_{11}	x_{14}	x_{15}	x_{12}	-	x_4	x_{10}	-	x_9	x_7	x_5	x_8	-	x_6
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{14}	x_{15}	x_{12}	x_{13}	x_4	x_{10}	x_2	x_9	-	-	-	-	-
ENETBF	x_1	x_{11}	x_{14}	x_{15}	-	x_{13}	-	x_{10}	x_2	x_9	x_7	x_5	x_8	x_3	x_6

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 37															
Predictor variables [a]	x_1	x_{12}	x_{13}	x_{14}	x_{15}	x_{11}	x_4	x_3	x_7	x_9	x_5	x_2	x_{10}	x_8	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7676 $corr(x_{12}, x_{14})$ =0.7479 $corr(x_{14}, x_{15})$ =0.7372														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{12}	x_{13}	x_{14}	x_{15}	x_{11}	x_4	x_3	x_7	-	-	-	-	x_8	-
ENETCV	x_1	x_{12}	x_{13}	-	-	x_{11}	x_4	x_3	x_7	-	-	x_2	x_{10}	x_8	x_6
λ_2 based on Bayes factor															
BVSBF	x_1	x_{12}	x_{13}	x_{14}	x_{15}	x_{11}	x_4	x_3	x_7	x_9	-	-	-	-	-
ENETBF	x_1	x_{12}	x_{13}	x_{14}	x_{15}	x_{11}	x_4	x_3	x_7	-	x_5	x_2	x_{10}	x_8	x_6

Data no. 38															
Predictor variables [a]	x_{13}	x_{11}	x_{15}	x_{14}	x_1	x_{12}	x_5	x_7	x_3	x_4	x_{10}	x_2	x_9	x_8	x_6
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7866 $corr(x_{12}, x_{14})$ =0.7806 $corr(x_{12}, x_{15})$ =0.7342														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{13}	x_{11}	x_{15}	x_{14}	x_1	x_{12}	x_5	x_7	x_3	-	-	-	-	-	-
ENETCV	x_{13}	x_{11}	-	x_{14}	x_1	-	-	x_7	x_3	-	-	-	x_9	-	-
λ_2 based on Bayes factor															
BVSBF	x_{13}	x_{11}	x_{15}	x_{14}	x_1	x_{12}	x_5	x_7	x_3	x_4	x_{10}	-	-	-	-
ENETBF	x_{13}	x_{11}	x_{15}	x_{14}	x_1	x_{12}	x_5	x_7	x_3	-	x_{10}	-	x_9	-	x_6

Data no. 39															
Predictor variables [a]	x_{14}	x_{13}	x_{11}	x_{12}	x_1	x_{15}	x_7	x_{10}	x_4	x_5	x_3	x_8	x_9	x_2	x_6
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7914 $corr(x_{12}, x_{14})$ =0.7695 $corr(x_{12}, x_{15})$ =0.6798														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{14}	x_{13}	x_{11}	x_{12}	x_1	x_{15}	x_7	x_{10}	x_4	x_5	x_3	-	-	-	-
ENETCV	x_{14}	x_{13}	x_{11}	x_{12}	x_1	x_{15}	x_7	x_{10}	x_4	x_5	x_3	x_8	x_9	x_2	x_6
λ_2 based on Bayes factor															
BVSBF	x_{14}	x_{13}	x_{11}	x_{12}	x_1	x_{15}	x_7	x_{10}	x_4	-	-	-	-	-	-
ENETBF	x_{14}	x_{13}	x_{11}	x_{12}	x_1	-	x_7	-	-	x_5	x_3	x_8	x_9	x_2	x_6

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 40															
Predictor variables [a]	x_1	x_3	x_{11}	x_{13}	x_{14}	x_{15}	x_7	x_{10}	x_{12}	x_4	x_6	x_5	x_9	x_2	x_8
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7204 $corr(x_{14}, x_{15})$ =0.6847 $corr(x_{12}, x_{15})$ =0.6504														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_3	x_{11}	x_{13}	x_{14}	x_{15}	x_7	x_{10}	-	-	-	x_5	-	-	-
ENETCV	x_1	x_3	x_{11}	x_{13}	x_{14}	-	x_7	x_{10}	-	x_4	-	x_5	x_9	x_2	x_8
λ_2 based on Bayes factor															
BVSBF	x_1	x_3	x_{11}	x_{13}	x_{14}	x_{15}	x_7	x_{10}	x_{12}	x_4	-	-	-	-	-
ENETBF	x_1	x_3	x_{11}	x_{13}	x_{14}	x_{15}	x_7	x_{10}	x_{12}	x_4	-	x_5	x_9	x_2	x_8

Data no. 41															
Predictor variables [a]	x_{11}	x_{13}	x_1	x_{12}	x_{14}	x_{15}	x_7	x_3	x_5	x_4	x_9	x_{10}	x_6	x_2	x_8
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.8019 $corr(x_{14}, x_{15})$ =0.7431 $corr(x_{12}, x_{14})$ =0.7176														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{13}	x_1	x_{12}	x_{14}	x_{15}	x_7	x_3	x_5	x_4	x_9	x_{10}	x_6	-	-
ENETCV	x_{11}	x_{13}	x_1	x_{12}	-	x_{15}	x_7	x_3	-	x_4	x_9	x_{10}	x_6	x_2	x_8
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{13}	x_1	x_{12}	x_{14}	x_{15}	x_7	x_3	x_5	x_4	x_9	x_{10}	x_6	-	-
ENETBF	x_{11}	x_{13}	x_1	x_{12}	-	x_{15}	x_7	x_3	x_5	x_4	x_9	x_{10}	x_6	x_2	x_8

Data no. 42															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	x_{12}	x_8	x_9	x_6	x_2	x_5	x_3	x_{10}
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7349 $corr(x_{12}, x_{15})$ =0.7230 $corr(x_{14}, x_{15})$ =0.6437														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{13}	x_{14}	-	-	x_7	-	-	-	-	-	-	-	-
ENETCV	x_1	x_{11}	x_{13}	x_{14}	-	-	x_7	-	-	-	x_6	x_2	x_5	x_3	x_{10}
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	x_{12}	-	-	-	-	-	-	-
ENETBF	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	-	-	x_9	x_6	x_2	x_5	x_3	x_{10}

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 43															
Predictor variables [a]	x_{11}	x_{15}	x_{1}	x_{13}	x_{12}	x_{4}	x_{14}	x_{7}	x_{3}	x_{8}	x_{9}	x_{2}	x_{10}	x_{5}	x_{6}
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.8320 $corr(x_{12}, x_{15})$ =0.7500 $corr(x_{14}, x_{15})$ =0.7334														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{15}	x_{1}	x_{13}	x_{12}	x_{4}	x_{14}	x_{7}	x_{3}	-	-	-	-	-	-
ENETCV	x_{11}	x_{15}	x_{1}	x_{13}	x_{12}	x_{4}	x_{14}	x_{7}	x_{3}	-	x_{9}	-	-	x_{5}	x_{6}
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{15}	x_{1}	x_{13}	x_{12}	x_{4}	x_{14}	x_{7}	x_{3}	-	-	-	-	-	-
ENETBF	x_{11}	x_{15}	x_{1}	x_{13}	x_{12}	x_{4}	x_{14}	x_{7}	x_{3}	x_{8}	x_{9}	-	-	x_{5}	x_{6}

Data no. 44															
Predictor variables [a]	x_{11}	x_{13}	x_{1}	x_{12}	x_{15}	x_{14}	x_{5}	x_{3}	x_{7}	x_{4}	x_{10}	x_{6}	x_{8}	x_{2}	x_{9}
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7436 $corr(x_{14}, x_{15})$ =0.6647 $corr(x_{12}, x_{15})$ =0.6381														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{13}	x_{1}	x_{12}	x_{15}	x_{14}	-	x_{3}	x_{7}	x_{4}	-	-	-	-	-
ENETCV	x_{11}	x_{13}	x_{1}	-	x_{15}	x_{14}	x_{5}	x_{3}	x_{7}	x_{4}	x_{10}	x_{6}	-	x_{2}	-
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{13}	x_{1}	x_{12}	x_{15}	x_{14}	x_{5}	x_{3}	x_{7}	x_{4}	-	-	-	-	-
ENETBF	x_{11}	x_{13}	x_{1}	x_{12}	x_{15}	x_{14}	x_{5}	x_{3}	x_{7}	x_{4}	x_{10}	x_{6}	-	x_{2}	x_{9}

Data no. 45															
Predictor variables [a]	x_{11}	x_{15}	x_{12}	x_{14}	x_{1}	x_{3}	x_{13}	x_{10}	x_{2}	x_{7}	x_{9}	x_{5}	x_{4}	x_{6}	x_{8}
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.8063 $corr(x_{12}, x_{14})$ =0.7714 $corr(x_{14}, x_{15})$ =0.7484														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{15}	x_{12}	x_{14}	x_{1}	x_{3}	x_{13}	x_{10}	x_{2}	x_{7}	x_{9}	-	-	-	-
ENETCV	x_{11}	-	x_{12}	x_{14}	x_{1}	x_{3}	x_{13}	x_{10}	-	x_{7}	x_{9}	x_{5}	x_{4}	x_{6}	x_{8}
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{15}	x_{12}	x_{14}	x_{1}	x_{3}	x_{13}	x_{10}	x_{2}	x_{7}	x_{9}	x_{5}	-	-	-
ENETBF	x_{11}	x_{15}	x_{12}	x_{14}	x_{1}	x_{3}	x_{13}	x_{10}	-	x_{7}	x_{9}	x_{5}	x_{4}	x_{6}	x_{8}

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 46															
Predictor variables [a]	x_{11}	x_1	x_{14}	x_{15}	x_{12}	x_{13}	x_3	x_4	x_5	x_6	x_7	x_9	x_{10}	x_2	x_8
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7941 $corr(x_{12}, x_{15})$ =0.7596 $corr(x_{12}, x_{14})$ =0.6998														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_1	x_{14}	x_{15}	x_{12}	-	x_3	-	-	-	x_7	-	-	-	-
ENETCV	x_{11}	x_1	x_{14}	x_{15}	x_{12}	-	x_3	-	-	-	x_7	x_9	-	x_2	x_8
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_1	x_{14}	x_{15}	x_{12}	x_{13}	x_3	x_4	x_5	-	-	-	-	-	-
ENETBF	x_{11}	x_1	x_{14}	-	x_{12}	x_{13}	x_3	-	x_5	-	x_7	x_9	-	x_2	x_8

Data no. 47															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{15}	x_4	x_{12}	x_{14}	x_8	x_{10}	x_2	x_3	x_7	x_9	x_5	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7379 $corr(x_{12}, x_{14})$ =0.6767 $corr(x_{14}, x_{15})$ =0.6453														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{13}	x_{15}	x_4	x_{12}	x_{14}	x_8	x_{10}	x_2	x_3	x_7	-	-	-
ENETCV	x_1	x_{11}	x_{13}	x_{15}	x_4	-	x_{14}	x_8	x_{10}	x_2	x_3	x_7	x_9	x_5	x_6
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{13}	x_{15}	x_4	x_{12}	x_{14}	x_8	x_{10}	x_2	x_3	x_7	-	-	-
ENETBF	x_1	x_{11}	x_{13}	x_{15}	x_4	x_{12}	x_{14}	x_8	x_{10}	x_2	x_3	x_7	x_9	x_5	x_6

Data no. 48															
Predictor variables [a]	x_1	x_{14}	x_{13}	x_{11}	x_3	x_{15}	x_{12}	x_5	x_7	x_2	x_8	x_9	x_4	x_{10}	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7687 $corr(x_{12}, x_{14})$ =0.7118 $corr(x_{14}, x_{15})$ =0.6514														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{14}	x_{13}	x_{11}	x_3	x_{15}	-	-	x_7	x_2	-	-	-	-	-
ENETCV	x_1	x_{14}	x_{13}	x_{11}	x_3	-	-	-	x_7	x_2	x_8	x_9	x_4	x_{10}	x_6
λ_2 based on Bayes factor															
BVSBF	x_1	x_{14}	x_{13}	x_{11}	x_3	x_{15}	x_{12}	x_5	x_7	x_2	-	-	-	-	-
ENETBF	x_1	x_{14}	x_{13}	x_{11}	x_3	x_{15}	x_{12}	-	x_7	x_2	x_8	x_9	x_4	x_{10}	x_6

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 49															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_{12}	x_8	x_5	x_6	x_4	x_7	x_9	x_{10}	x_3	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.8197 $corr(x_{12}, x_{14})$ =0.7736 $corr(x_{14}, x_{15})$ =0.7701														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_{12}	x_8	x_5	x_6	x_4	x_7	x_9	x_{10}	x_3	-
ENETCV	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_{12}	x_8	x_5	x_6	x_4	x_7	x_9	x_{10}	x_3	x_2
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_{12}	x_8	x_5	x_6	x_4	x_7	x_9	x_{10}	x_3	-
ENETBF	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_{12}	x_8	x_5	x_6	x_4	x_7	x_9	x_{10}	x_3	x_2

Data no. 50															
Predictor variables [a]	x_{15}	x_{13}	x_{11}	x_{12}	x_1	x_4	x_{14}	x_7	x_6	x_5	x_3	x_8	x_9	x_2	x_{10}
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7471 $corr(x_{12}, x_{15})$ =0.7320 $corr(x_{14}, x_{15})$ =0.7207														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{15}	x_{13}	x_{11}	x_{12}	x_1	x_4	-	x_7	x_6	-	-	x_8	-	-	-
ENETCV	x_{15}	x_{13}	x_{11}	x_{12}	x_1	x_4	-	x_7	x_6	-	x_3	x_8	-	-	x_{10}
λ_2 based on Bayes factor															
BVSBF	x_{15}	x_{13}	x_{11}	x_{12}	x_1	x_4	x_{14}	x_7	x_6	x_5	x_3	x_8	x_9	-	-
ENETBF	x_{15}	x_{13}	x_{11}	x_{12}	x_1	x_4	x_{14}	x_7	x_6	-	x_3	x_8	x_9	x_2	x_{10}

Data no. 51															
Predictor variables [a]	x_{11}	x_1	x_{13}	x_{14}	x_7	x_3	x_{15}	x_{12}	x_{10}	x_8	x_4	x_9	x_6	x_5	x_2
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7901 $corr(x_{12}, x_{14})$ =0.7249 $corr(x_{12}, x_{15})$ =0.7215														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_1	x_{13}	x_{14}	x_7	x_3	x_{15}	x_{12}	x_{10}	-	-	-	-	-	-
ENETCV	x_{11}	x_1	x_{13}	x_{14}	x_7	x_3	-	x_{12}	x_{10}	x_8	-	x_9	x_6	x_5	x_2
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_1	x_{13}	x_{14}	x_7	x_3	x_{15}	x_{12}	x_{10}	-	-	-	-	-	-
ENETBF	x_{11}	x_1	x_{13}	x_{14}	x_7	x_3	-	x_{12}	x_{10}	x_8	-	x_9	x_6	x_5	x_2

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\dots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 52															
Predictor variables [a]	x_{11}	x_{15}	x_{13}	x_{14}	x_1	x_4	x_{12}	x_3	x_7	x_2	x_5	x_{10}	x_9	x_8	x_6
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7879 $corr(x_{12}, x_{15})$ =0.7757 $corr(x_{12}, x_{14})$ =0.7315														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{15}	x_{13}	x_{14}	x_1	x_4	-	-	x_7	-	-	-	-	-	-
ENETCV	x_{11}	x_{15}	x_{13}	-	x_1	x_4	-	-	x_7	-	-	-	-	-	-
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{15}	x_{13}	x_{14}	x_1	x_4	x_{12}	x_3	x_7	x_2	x_5	-	-	-	-
ENETBF	x_{11}	-	x_{13}	x_{14}	x_1	x_4	x_{12}	-	x_7	-	x_5	x_{10}	x_9	x_8	x_6

Data no. 53															
Predictor variables [a]	x_1	x_7	x_{13}	x_{15}	x_{14}	x_{11}	x_{12}	x_9	x_{10}	x_3	x_5	x_8	x_4	x_6	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7663 $corr(x_{14}, x_{15})$ =0.7282 $corr(x_{12}, x_{15})$ =0.7041														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_7	x_{13}	-	-	x_{11}	-	-	-	-	-	-	-	-	-
ENETCV	x_1	x_7	x_{13}	-	x_{14}	x_{11}	x_{12}	-	x_{10}	x_3	-	-	-	-	-
λ_2 based on Bayes factor															
BVSBF	x_1	x_7	x_{13}	x_{15}	x_{14}	x_{11}	-	x_9	x_{10}	-	-	-	-	-	-
ENETBF	x_1	x_7	x_{13}	-	x_{14}	x_{11}	x_{12}	x_9	x_{10}	x_3	x_5	x_8	x_4	x_6	-

Data no. 54															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_7	x_{15}	x_2	x_{14}	x_{12}	x_8	x_3	x_4	x_6	x_{10}	x_5	x_9
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7988 $corr(x_{12}, x_{14})$ =0.7627 $corr(x_{14}, x_{15})$ =0.7475														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{13}	x_7	x_{15}	x_2	x_{14}	x_{12}	-	-	-	-	-	-	-
ENETCV	x_1	x_{11}	x_{13}	x_7	x_{15}	x_2	x_{14}	x_{12}	x_8	x_3	x_4	x_6	x_{10}	x_5	x_9
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{13}	x_7	x_{15}	x_2	x_{14}	x_{12}	-	-	-	-	-	-	-
ENETBF	x_1	x_{11}	x_{13}	x_7	-	x_2	x_{14}	x_{12}	x_8	x_3	x_4	x_6	x_{10}	x_5	x_9

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 55															
Predictor variables [a]	x_{13}	x_{11}	x_{12}	x_1	x_{14}	x_{15}	x_7	x_5	x_4	x_3	x_2	x_9	x_{10}	x_8	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7874 $corr(x_{14}, x_{15})$ =0.7149 $corr(x_{12}, x_{14})$ =0.7078														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{13}	x_{11}	x_{12}	x_1	x_{14}	x_{15}	x_7	-	-	-	-	-	-	-	-
ENETCV	x_{13}	x_{11}	x_{12}	x_1	x_{14}	-	x_7	-	x_4	-	-	-	-	x_8	x_6
λ_2 based on Bayes factor															
BVSBF	x_{13}	x_{11}	x_{12}	x_1	x_{14}	x_{15}	x_7	x_5	x_4	x_3	-	-	-	-	-
ENETBF	x_{13}	x_{11}	x_{12}	x_1	x_{14}	x_{15}	x_7	x_5	x_4	x_3	x_2	x_9	x_{10}	x_8	x_6

Data no. 56															
Predictor variables [a]	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_4	x_1	x_7	x_3	x_2	x_5	x_6	x_{10}	x_9	x_8
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.8237 $corr(x_{14}, x_{15})$ =0.7913 $corr(x_{12}, x_{15})$ =0.7909														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{12}	x_{13}	-	x_{15}	x_4	x_1	x_7	-	-	-	-	-	-	-
ENETCV	x_{11}	x_{12}	x_{13}	-	x_{15}	x_4	x_1	x_7	x_3	x_2	-	-	-	x_9	x_8
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_4	x_1	x_7	x_3	x_2	x_5	x_6	-	-	-
ENETBF	x_{11}	-	x_{13}	-	x_{15}	x_4	x_1	x_7	x_3	x_2	x_5	x_6	x_{10}	x_9	x_8

Data no. 57															
Predictor variables [a]	x_{11}	x_3	x_1	x_7	x_{13}	x_{12}	x_{14}	x_2	x_{15}	x_4	x_5	x_6	x_{10}	x_9	x_8
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7906 $corr(x_{12}, x_{14})$ =0.7641 $corr(x_{12}, x_{15})$ =0.7473														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_3	x_1	x_7	x_{13}	x_{12}	x_{14}	x_2	x_{15}	-	-	-	-	-	-
ENETCV	x_{11}	x_3	x_1	x_7	x_{13}	x_{12}	x_{14}	x_2	x_{15}	x_4	x_5	x_6	-	x_9	x_8
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_3	x_1	x_7	x_{13}	x_{12}	x_{14}	x_2	x_{15}	-	-	-	-	-	-
ENETBF	x_{11}	x_3	x_1	x_7	x_{13}	x_{12}	x_{14}	x_2	x_{15}	x_4	x_5	x_6	-	x_9	x_8

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 58															
Predictor variables [a]	x_1	x_{11}	x_{14}	x_{15}	x_{13}	x_{12}	x_4	x_7	x_{10}	x_5	x_3	x_8	x_6	x_9	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.8220 $corr(x_{12}, x_{15})$ =0.7781 $corr(x_{14}, x_{15})$ =0.7615														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{14}	x_{15}	x_{13}	x_{12}	x_4	x_7	x_{10}	-	x_3	x_8	-	-	-
ENETCV	x_1	x_{11}	-	x_{15}	x_{13}	x_{12}	-	x_7	x_{10}	x_5	x_3	-	x_6	x_9	x_2
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{14}	x_{15}	x_{13}	x_{12}	x_4	x_7	x_{10}	x_5	x_3	x_8	-	-	-
ENETBF	x_1	x_{11}	x_{14}	x_{15}	x_{13}	x_{12}	x_4	x_7	x_{10}	x_5	x_3	x_8	x_6	x_9	x_2

Data no. 59															
Predictor variables [a]	x_{11}	x_{15}	x_{13}	x_{12}	x_{14}	x_1	x_3	x_7	x_4	x_5	x_{10}	x_6	x_2	x_8	x_9
Correlation coefficient between predictors	$orr(x_{12}, x_{14})$ =0.8129 $corr(x_{14}, x_{15})$ =0.7908 $corr(x_{12}, x_{15})$ =0.7899														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{15}	x_{13}	x_{12}	x_{14}	x_1	x_3	x_7	x_4	-	x_{10}	x_6	-	-	-
ENETCV	x_{11}	x_{15}	x_{13}	-	x_{14}	x_1	x_3	x_7	x_4	x_5	x_{10}	x_6	x_2	x_8	x_9
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{15}	x_{13}	x_{12}	x_{14}	x_1	x_3	x_7	x_4	x_5	x_{10}	x_6	x_2	-	-
ENETBF	x_{11}	-	x_{13}	x_{12}	x_{14}	x_1	x_3	x_7	x_4	x_5	x_{10}	x_6	x_2	x_8	x_9

Data no. 60															
Predictor variables [a]	x_1	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_7	x_5	x_9	x_4	x_3	x_{10}	x_6	x_8	x_2
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7398 $corr(x_{12}, x_{15})$ =0.6710 $corr(x_{12}, x_{14})$ =0.6614														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_7	-	x_9	-	-	-	-	-	-
ENETCV	x_1	x_{11}	x_{12}	x_{13}	-	-	x_7	-	x_9	x_4	-	-	x_6	x_8	x_2
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_7	x_5	x_9	x_4	x_3	-	-	-	-
ENETBF	x_1	x_{11}	x_{12}	x_{13}	x_{14}	-	x_7	x_5	x_9	-	-	x_{10}	x_6	x_8	x_2

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 61															
Predictor variables [a]	x_{11}	x_{13}	x_1	x_{15}	x_{14}	x_9	x_5	x_{12}	x_4	x_7	x_2	x_3	x_{10}	x_6	x_8
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.8445 $corr(x_{14}, x_{15})$ =0.7982 $corr(x_{12}, x_{15})$ =0.7967														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{13}	x_1	x_{15}	x_{14}	x_9	x_5	x_{12}	x_4	x_7	x_2	-	-	-	-
ENETCV	x_{11}	x_{13}	x_1	x_{15}	-	x_9	-	x_{12}	x_4	x_7	x_2	-	x_{10}	-	x_8
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{13}	x_1	x_{15}	x_{14}	x_9	x_5	x_{12}	x_4	x_7	x_2	-	-	-	-
ENETBF	x_{11}	x_{13}	x_1	x_{15}	x_{14}	x_9	-	x_{12}	x_4	x_7	x_2	x_3	x_{10}	-	x_8

Data no. 62															
Predictor variables [a]	x_{11}	x_1	x_{13}	x_{15}	x_{14}	x_9	x_{12}	x_5	x_8	x_7	x_4	x_6	x_2	x_3	x_{10}
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7555 $corr(x_{12}, x_{14})$ =0.7262 $corr(x_{14}, x_{15})$ =0.6841														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_1	x_{13}	x_{15}	x_{14}	x_9	x_{12}	x_5	x_8	x_7	-	x_6	-	-	-
ENETCV	x_{11}	x_1	x_{13}	-	x_{14}	x_9	x_{12}	-	x_8	x_7	-	x_6	x_2	x_3	x_{10}
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_1	x_{13}	x_{15}	x_{14}	x_9	x_{12}	x_5	x_8	x_7	x_4	x_6	-	-	-
ENETBF	x_{11}	x_1	x_{13}	-	x_{14}	x_9	x_{12}	-	x_8	x_7	x_4	x_6	x_2	x_3	x_{10}

Data no. 63															
Predictor variables [a]	x_{13}	x_{11}	x_1	x_{15}	x_{14}	x_7	x_5	x_{12}	x_8	x_6	x_3	x_4	x_9	x_2	x_{10}
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7402 $corr(x_{12}, x_{14})$ =0.6880 $corr(x_{12}, x_{15})$ =0.6489														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{13}	x_{11}	x_1	x_{15}	x_{14}	x_7	x_5	x_{12}	x_8	x_6	x_3	-	-	-	-
ENETCV	x_{13}	x_{11}	x_1	-	x_{14}	x_7	-	-	x_8	x_6	x_3	-	x_9	x_2	x_{10}
λ_2 based on Bayes factor															
BVSBF	x_{13}	x_{11}	x_1	x_{15}	x_{14}	x_7	x_5	x_{12}	x_8	x_6	x_3	-	-	-	-
ENETBF	x_{13}	x_{11}	x_1	x_{15}	x_{14}	x_7	x_5	x_{12}	x_8	x_6	x_3	x_4	x_9	x_2	x_{10}

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 64															
Predictor variables [a]	x_1	x_{11}	x_{15}	x_{14}	x_{12}	x_{13}	x_4	x_7	x_6	x_3	x_{10}	x_8	x_5	x_2	x_9
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7483 $corr(x_{12}, x_{14})$ =0.7182 $corr(x_{12}, x_{15})$ =0.6956														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{15}	x_{14}	x_{12}	x_{13}	x_4	x_7	x_6	-	-	-	-	-	-
ENETCV	x_1	x_{11}	-	-	x_{12}	x_{13}	x_4	x_7	x_6	x_3	-	x_8	x_5	x_2	x_9
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{15}	x_{14}	x_{12}	x_{13}	x_4	x_7	x_6	-	-	-	-	-	-
ENETBF	x_1	x_{11}	-	x_{14}	x_{12}	x_{13}	x_4	x_7	x_6	x_3	-	x_8	x_5	x_2	x_9

Data no. 65															
Predictor variables [a]	x_{11}	x_1	x_{13}	x_{14}	x_{12}	x_{15}	x_4	x_8	x_2	x_7	x_9	x_5	x_3	x_{10}	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7777 $corr(x_{12}, x_{15})$ =0.7447 $corr(x_{14}, x_{15})$ =0.7001														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_1	x_{13}	x_{14}	x_{12}	-	-	x_8	-	x_7	x_9	-	-	-	-
ENETCV	x_{11}	x_1	x_{13}	x_{14}	-	-	x_4	x_8	x_2	x_7	x_9	x_5	x_3	-	x_6
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_1	x_{13}	x_{14}	x_{12}	x_{15}	x_4	x_8	x_2	x_7	x_9	x_5	x_3	x_{10}	-
ENETBF	x_{11}	x_1	x_{13}	x_{14}	x_{12}	x_{15}	x_4	x_8	x_2	x_7	x_9	x_5	x_3	x_{10}	x_6

Data no. 66															
Predictor variables [a]	x_1	x_{14}	x_{11}	x_{13}	x_{12}	x_{15}	x_7	x_2	x_9	x_5	x_3	x_4	x_6	x_8	x_{10}
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.8290 $corr(x_{12}, x_{14})$ =0.7442 $corr(x_{12}, x_{15})$ =0.7416														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{14}	x_{11}	x_{13}	-	-	x_7	-	x_9	-	-	-	-	-	-
ENETCV	x_1	x_{14}	x_{11}	x_{13}	-	-	x_7	x_2	x_9	-	-	-	-	-	x_{10}
λ_2 based on Bayes factor															
BVSBF	x_1	x_{14}	x_{11}	x_{13}	x_{12}	x_{15}	x_7	x_2	x_9	-	-	-	-	-	-
ENETBF	x_1	x_{14}	x_{11}	x_{13}	x_{12}	x_{15}	x_7	x_2	x_9	x_5	x_3	-	-	x_8	x_{10}

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 67															
Predictor variables [a]	x_1	x_{11}	x_{15}	x_{13}	x_{12}	x_7	x_{14}	x_5	x_6	x_2	x_4	x_9	x_3	x_{10}	x_8
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7856 $corr(x_{12}, x_{14})$ =0.7257 $corr(x_{12}, x_{15})$ =0.7169														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{15}	x_{13}	x_{12}	x_7	x_{14}	x_5	x_6	x_2	x_4	x_9	-	-	-
ENETCV	x_1	x_{11}	x_{15}	x_{13}	x_{12}	x_7	x_{14}	x_5	x_6	x_2	-	x_9	-	x_{10}	x_8
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{15}	x_{13}	x_{12}	x_7	x_{14}	x_5	x_6	x_2	x_4	x_9	-	-	-
ENETBF	x_1	x_{11}	x_{15}	x_{13}	x_{12}	x_7	x_{14}	x_5	x_6	x_2	-	x_9	x_3	x_{10}	x_8

Data no. 68															
Predictor variables [a]	x_1	x_{11}	x_7	x_3	x_{15}	x_{13}	x_8	x_{14}	x_{12}	x_6	x_4	x_5	x_{10}	x_9	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7494 $corr(x_{14}, x_{15})$ =0.7467 $corr(x_{12}, x_{14})$ =0.7357														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_7	x_3	x_{15}	x_{13}	x_8	-	-	-	-	-	-	-	-
ENETCV	x_1	x_{11}	x_7	x_3	-	x_{13}	x_8	-	-	-	x_4	x_5	x_{10}	-	x_2
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_7	x_3	x_{15}	x_{13}	x_8	-	-	-	-	-	-	-	-
ENETBF	x_1	x_{11}	x_7	x_3	x_{15}	x_{13}	x_8	x_{14}	x_{12}	x_6	x_4	x_5	x_{10}	x_9	x_2

Data no. 69															
Predictor variables [a]	x_1	x_{11}	x_7	x_{14}	x_{12}	x_{13}	x_4	x_{15}	x_2	x_6	x_{10}	x_8	x_3	x_5	x_9
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.8165 $corr(x_{14}, x_{15})$ =0.7702 $corr(x_{12}, x_{15})$ =0.7428														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_7	x_{14}	x_{12}	x_{13}	x_4	x_{15}	x_2	x_6	x_{10}	-	-	-	-
ENETCV	x_1	x_{11}	x_7	x_{14}	x_{12}	-	x_4	-	x_2	x_6	x_{10}	x_8	-	x_5	-
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_7	x_{14}	x_{12}	x_{13}	x_4	x_{15}	x_2	x_6	x_{10}	-	-	-	-
ENETBF	x_1	x_{11}	x_7	x_{14}	x_{12}	x_{13}	x_4	-	x_2	x_6	x_{10}	x_8	x_3	x_5	x_9

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 70															
Predictor variables [a]	x_7	x_{12}	x_{13}	x_{11}	x_1	x_{14}	x_3	x_6	x_{15}	x_4	x_8	x_2	x_5	x_{10}	x_9
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7313 $corr(x_{12}, x_{15})$ =0.7197 $corr(x_{14}, x_{15})$ =0.6913														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_7	x_{12}	x_{13}	x_{11}	x_1	x_{14}	x_3	x_6	-	-	-	-	-	-	-
ENETCV	x_7	-	x_{13}	x_{11}	x_1	-	x_3	x_6	x_{15}	x_4	x_8	-	x_5	x_{10}	x_9
λ_2 based on Bayes factor															
BVSBF	x_7	x_{12}	x_{13}	x_{11}	x_1	x_{14}	x_3	x_6	x_{15}	x_4	-	-	-	-	-
ENETBF	x_7	x_{12}	x_{13}	x_{11}	x_1	x_{14}	x_3	x_6	x_{15}	x_4	x_8	x_2	x_5	x_{10}	x_9

Data no. 71															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_7	x_3	x_4	x_2	x_8	x_6	x_9	x_{10}	x_5
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.6934 $corr(x_{14}, x_{15})$ =0.6919 $corr(x_{12}, x_{15})$ =0.6094														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_7	x_3	x_4	-	-	-	-	-	x_5
ENETCV	x_1	x_{11}	x_{13}	x_{15}	x_{14}	-	x_7	x_3	x_4	x_2	-	x_6	x_9	-	x_5
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_7	x_3	x_4	-	-	-	-	-	-
ENETBF	x_1	x_{11}	x_{13}	x_{15}	x_{14}	x_{12}	x_7	x_3	x_4	x_2	x_8	x_6	x_9	-	x_5

Data no. 72															
Predictor variables [a]	x_{14}	x_{12}	x_{11}	x_{13}	x_1	x_3	x_{15}	x_4	x_{10}	x_5	x_7	x_2	x_9	x_8	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7762 $corr(x_{14}, x_{15})$ =0.7286 $corr(x_{12}, x_{15})$ =0.7260														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{14}	x_{12}	x_{11}	x_{13}	x_1	x_3	x_{15}	x_4	x_{10}	-	x_7	-	-	-	-
ENETCV	x_{14}	x_{12}	x_{11}	x_{13}	x_1	x_3	x_{15}	-	x_{10}	x_5	x_7	-	-	x_8	x_6
λ_2 based on Bayes factor															
BVSBF	x_{14}	x_{12}	x_{11}	x_{13}	x_1	x_3	x_{15}	x_4	x_{10}	-	x_7	-	-	-	-
ENETBF	x_{14}	x_{12}	x_{11}	x_{13}	x_1	x_3	x_{15}	-	x_{10}	x_5	x_7	-	-	x_8	x_6

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 73															
Predictor variables [a]	x_{11}	x_{15}	x_{14}	x_{12}	x_1	x_{13}	x_4	x_3	x_7	x_2	x_5	x_8	x_{10}	x_9	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7076 $corr(x_{14}, x_{15})$ =0.6934 $corr(x_{12}, x_{15})$ =0.6280														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{15}	x_{14}	x_{12}	x_1	x_{13}	x_4	x_3	x_7	x_2	-	x_8	x_{10}	-	-
ENETCV	x_{11}	x_{15}	x_{14}	x_{12}	x_1	x_{13}	x_4	-	x_7	x_2	x_5	x_8	x_{10}	x_9	x_6
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{15}	x_{14}	x_{12}	x_1	x_{13}	x_4	x_3	x_7	x_2	x_5	x_8	x_{10}	x_9	-
ENETBF	x_{11}	x_{15}	x_{14}	x_{12}	x_1	x_{13}	-	x_3	x_7	-	x_5	x_8	x_{10}	x_9	x_6

Data no. 74															
Predictor variables [a]	x_{15}	x_1	x_{13}	x_{11}	x_7	x_{14}	x_{12}	x_2	x_4	x_3	x_9	x_8	x_6	x_5	x_{10}
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7503 $corr(x_{12}, x_{15})$ =0.6811 $corr(x_{14}, x_{15})$ =0.6318														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{15}	x_1	x_{13}	x_{11}	x_7	x_{14}	x_{12}	x_2	x_4	x_3	-	-	-	-	-
ENETCV	x_{15}	x_1	x_{13}	x_{11}	x_7	-	x_{12}	-	x_4	x_3	x_9	x_8	x_6	x_5	x_{10}
λ_2 based on Bayes factor															
BVSBF	x_{15}	x_1	x_{13}	x_{11}	x_7	x_{14}	x_{12}	x_2	x_4	x_3	-	-	-	-	-
ENETBF	x_{15}	x_1	x_{13}	x_{11}	x_7	-	x_{12}	x_2	x_4	x_3	x_9	x_8	x_6	x_5	x_{10}

Data no. 75															
Predictor variables [a]	x_1	x_7	x_3	x_{11}	x_{14}	x_{13}	x_{15}	x_{12}	x_8	x_4	x_2	x_{10}	x_9	x_5	x_6
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7848 $corr(x_{12}, x_{15})$ =0.7729 $corr(x_{12}, x_{14})$ =0.7719														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_7	x_3	x_{11}	x_{14}	x_{13}	x_{15}	x_{12}	x_8	x_4	-	x_{10}	-	x_5	-
ENETCV	x_1	x_7	x_3	x_{11}	-	x_{13}	x_{15}	x_{12}	x_8	x_4	x_2	x_{10}	x_9	x_5	-
λ_2 based on Bayes factor															
BVSBF	x_1	x_7	x_3	x_{11}	x_{14}	x_{13}	x_{15}	x_{12}	x_8	x_4	x_2	x_{10}	-	-	-
ENETBF	x_1	x_7	x_3	x_{11}	x_{14}	x_{13}	x_{15}	x_{12}	x_8	x_4	x_2	-	x_9	x_5	-

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 76															
Predictor variables [a]	x_7	x_1	x_{11}	x_{14}	x_3	x_{15}	x_{13}	x_{12}	x_8	x_6	x_2	x_5	x_{10}	x_9	x_4
Correlation coefficient between predictors	$corr(x_{12},x_{15})$ =0.8863 $corr(x_{12},x_{14})$ =0.7998 $corr(x_{14},x_{15})$ =0.7927														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_7	x_1	x_{11}	x_{14}	x_3	x_{15}	x_{13}	-	x_8	x_6	x_2	-	-	-	-
ENETCV	x_7	x_1	x_{11}	x_{14}	x_3	-	x_{13}	x_{12}	x_8	x_6	x_2	x_5	x_{10}	-	-
λ_2 based on Bayes factor															
BVSBF	x_7	x_1	x_{11}	x_{14}	x_3	x_{15}	x_{13}	x_{12}	x_8	x_6	x_2	-	-	-	-
ENETBF	x_7	x_1	x_{11}	x_{14}	x_3	-	x_{13}	x_{12}	x_8	x_6	x_2	x_5	x_{10}	x_9	-

Data no. 77															
Predictor variables [a]	x_{11}	x_{13}	x_1	x_{15}	x_{12}	x_{14}	x_3	x_2	x_7	x_5	x_4	x_6	x_{10}	x_9	x_8
Correlation coefficient between predictors	$corr(x_{12},x_{14})$ =0.7790 $corr(x_{12},x_{15})$ =0.7541 $corr(x_{14},x_{15})$ =0.6842														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{13}	x_1	x_{15}	x_{12}	x_{14}	x_3	x_2	x_7	-	-	-	-	-	-
ENETCV	x_{11}	x_{13}	x_1	x_{15}	x_{12}	x_{14}	x_3	x_2	x_7	x_5	x_4	x_6	x_{10}	x_9	x_8
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{13}	x_1	x_{15}	x_{12}	x_{14}	x_3	x_2	x_7	-	-	-	-	-	-
ENETBF	x_{11}	x_{13}	x_1	x_{15}	x_{12}	x_{14}	x_3	x_2	x_7	x_5	x_4	x_6	x_{10}	x_9	x_8

Data no. 78															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_7	x_{15}	x_2	x_{12}	x_{14}	x_3	x_4	x_9	x_6	x_{10}	x_8	x_5
Correlation coefficient between predictors	$corr(x_{12},x_{15})$ =0.8185 $corr(x_{14},x_{15})$ =0.7735 $corr(x_{12},x_{14})$ =0.7631														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{13}	x_7	x_{15}	x_2	x_{12}	x_{14}	x_3	x_4	-	x_6	-	-	-
ENETCV	x_1	x_{11}	x_{13}	x_7	-	x_2	-	-	x_3	x_4	x_9	x_6	-	x_8	x_5
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{13}	x_7	x_{15}	x_2	x_{12}	x_{14}	x_3	x_4	x_9	x_6	-	-	-
ENETBF	x_1	x_{11}	x_{13}	x_7	x_{15}	x_2	-	x_{14}	x_3	x_4	x_9	x_6	x_{10}	x_8	x_5

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 79															
Predictor variables [a]	x_{11}	x_{13}	x_1	x_{15}	x_{12}	x_{14}	x_3	x_4	x_5	x_2	x_{10}	x_7	x_9	x_8	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7802 $corr(x_{14}, x_{15})$ =0.7647 $corr(x_{12}, x_{15})$ =0.7528														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{13}	x_1	x_{15}	x_{12}	-	-	-	-	-	x_{10}	-	-	-	-
ENETCV	x_{11}	x_{13}	x_1	-	-	x_{14}	-	-	x_5	x_2	x_{10}	x_7	-	-	-
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{13}	x_1	x_{15}	x_{12}	x_{14}	x_3	x_4	x_5	x_2	-	-	-	-	-
ENETBF	x_{11}	x_{13}	x_1	-	x_{12}	x_{14}	x_3	x_4	x_5	x_2	x_{10}	x_7	x_9	x_8	x_6

Data no. 80															
Predictor variables [a]	x_{15}	x_{12}	x_{13}	x_{11}	x_1	x_{14}	x_7	x_4	x_8	x_5	x_2	x_{10}	x_3	x_6	x_9
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7836 $corr(x_{12}, x_{14})$ =0.7775 $corr(x_{14}, x_{15})$ =0.7684														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{15}	x_{12}	x_{13}	x_{11}	x_1	x_{14}	x_7	x_4	x_8	-	x_2	x_{10}	-	-	-
ENETCV	x_{15}	x_{12}	x_{13}	x_{11}	x_1	-	x_7	x_4	x_8	x_5	-	x_{10}	x_3	x_6	x_9
λ_2 based on Bayes factor															
BVSBF	x_{15}	x_{12}	x_{13}	x_{11}	x_1	x_{14}	x_7	x_4	x_8	x_5	x_2	x_{10}	-	-	-
ENETBF	x_{15}	x_{12}	x_{13}	x_{11}	x_1	x_{14}	x_7	x_4	x_8	x_5	x_2	x_{10}	x_3	x_6	x_9

Data no. 81															
Predictor variables [a]	x_{11}	x_{15}	x_{13}	x_{14}	x_1	x_3	x_{12}	x_{10}	x_9	x_4	x_8	x_2	x_5	x_7	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7695 $corr(x_{14}, x_{15})$ =0.7236 $corr(x_{12}, x_{14})$ =0.6547														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{15}	x_{13}	x_{14}	x_1	x_3	x_{12}	x_{10}	x_9	x_4	x_8	-	-	x_7	-
ENETCV	x_{11}	x_{15}	x_{13}	-	x_1	x_3	-	x_{10}	x_9	-	x_8	-	x_5	x_7	-
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{15}	x_{13}	x_{14}	x_1	x_3	x_{12}	x_{10}	x_9	x_4	x_8	x_2	x_5	x_7	-
ENETBF	x_{11}	x_{15}	x_{13}	x_{14}	x_1	x_3	x_{12}	x_{10}	x_9	x_4	x_8	-	x_5	x_7	x_6

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\dots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 82															
Predictor variables [a]	x_1	x_{13}	x_{11}	x_{15}	x_{14}	x_{12}	x_7	x_3	x_4	x_2	x_8	x_6	x_{10}	x_5	x_9
Correlation coefficient between predictors	$corr(x_{12},x_{14})$ =0.7583 $corr(x_{12},x_{15})$ =0.7574 $corr(x_{14},x_{15})$ =0.7495														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{13}	x_{11}	x_{15}	x_{14}	x_{12}	x_7	x_3	-	x_2	x_8	x_6	-	-	-
ENETCV	x_1	x_{13}	x_{11}	x_{15}	x_{14}	x_{12}	x_7	x_3	-	x_2	x_8	x_6	-	x_5	x_9
λ_2 based on Bayes factor															
BVSBF	x_1	x_{13}	x_{11}	x_{15}	x_{14}	x_{12}	x_7	x_3	x_4	x_2	x_8	x_6	-	-	-
ENETBF	x_1	x_{13}	x_{11}	x_{15}	x_{14}	x_{12}	x_7	x_3	x_4	x_2	x_8	x_6	-	x_5	x_9

Data no. 83															
Predictor variables [a]	x_1	x_{13}	x_{14}	x_{11}	x_{15}	x_4	x_{12}	x_3	x_9	x_7	x_8	x_6	x_5	x_{10}	x_2
Correlation coefficient between predictors	$corr(x_{12},x_{14})$ =0.7606 $corr(x_{12},x_{15})$ =0.7367 $corr(x_{14},x_{15})$ =0.7347														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{13}	x_{14}	x_{11}	x_{15}	x_4	x_{12}	x_3	x_9	x_7	x_8	-	-	-	-
ENETCV	x_1	x_{13}	x_{14}	x_{11}	x_{15}	x_4	x_{12}	x_3	x_9	x_7	x_8	x_6	x_5	x_{10}	-
λ_2 based on Bayes factor															
BVSBF	x_1	x_{13}	x_{14}	x_{11}	x_{15}	x_4	x_{12}	x_3	x_9	x_7	x_8	-	-	-	-
ENETBF	x_1	x_{13}	x_{14}	x_{11}	x_{15}	x_4	x_{12}	x_3	x_9	x_7	x_8	x_6	x_5	x_{10}	-

Data no. 84															
Predictor variables [a]	x_1	x_7	x_{11}	x_{13}	x_6	x_{12}	x_4	x_3	x_{14}	x_{15}	x_2	x_9	x_8	x_{10}	x_5
Correlation coefficient between predictors	$corr(x_{12},x_{14})$ =0.6786 $corr(x_{14},x_{15})$ =0.6756 $corr(x_{12},x_{15})$ =0.6694														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_7	x_{11}	x_{13}	x_6	x_{12}	x_4	x_3	x_{14}	x_{15}	x_2	-	-	-	-
ENETCV	x_1	x_7	x_{11}	x_{13}	x_6	-	x_4	x_3	x_{14}	x_{15}	x_2	x_9	x_8	x_{10}	x_5
λ_2 based on Bayes factor															
BVSBF	x_1	x_7	x_{11}	x_{13}	x_6	x_{12}	x_4	x_3	x_{14}	x_{15}	x_2	-	-	-	-
ENETBF	x_1	x_7	x_{11}	x_{13}	x_6	x_{12}	x_4	x_3	x_{14}	x_{15}	x_2	x_9	x_8	x_{10}	x_5

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,...,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 85															
Predictor variables [a]	x_{11}	x_4	x_{13}	x_{15}	x_{14}	x_{12}	x_1	x_2	x_9	x_5	x_{10}	x_7	x_3	x_6	x_8
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7490 $corr(x_{12}, x_{14})$ =0.6986 $corr(x_{12}, x_{15})$ =0.6668														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_4	x_{13}	x_{15}	x_{14}	x_{12}	x_1	x_2	x_9	-	-	x_7	-	-	-
ENETCV	x_{11}	x_4	x_{13}	x_{15}	-	-	x_1	x_2	x_9	x_5	x_{10}	x_7	x_3	x_6	-
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_4	x_{13}	x_{15}	x_{14}	x_{12}	x_1	x_2	x_9	-	-	x_7	-	-	-
ENETBF	x_{11}	x_4	x_{13}	x_{15}	-	-	x_1	x_2	x_9	x_5	x_{10}	x_7	x_3	x_6	-

Data no. 86															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{12}	x_{15}	x_{14}	x_4	x_8	x_5	x_9	x_7	x_2	x_6	x_{10}	x_3
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.8083 $corr(x_{14}, x_{15})$ =0.7597 $corr(x_{12}, x_{15})$ =0.7551														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{13}	x_{12}	x_{15}	-	x_4	x_8	-	x_9	x_7	-	-	-	-
ENETCV	x_1	x_{11}	-	x_{12}	x_{15}	-	x_4	x_8	x_5	x_9	x_7	-	x_6	-	x_3
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{13}	x_{12}	x_{15}	x_{14}	x_4	x_8	x_5	x_9	x_7	-	x_6	-	-
ENETBF	x_1	x_{11}	-	x_{12}	x_{15}	x_{14}	-	x_8	x_5	x_9	x_7	x_2	x_6	-	x_3

Data no. 87															
Predictor variables [a]	x_{11}	x_1	x_{13}	x_7	x_3	x_{15}	x_4	x_{12}	x_{10}	x_5	x_8	x_9	x_{14}	x_6	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7991 $corr(x_{12}, x_{15})$ =0.6806 $corr(x_{14}, x_{15})$ =0.6619														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_1	x_{13}	x_7	x_3	x_{15}	x_4	x_{12}	x_{10}	x_5	-	-	-	-	-
ENETCV	x_{11}	x_1	x_{13}	x_7	x_3	x_{15}	x_4	x_{12}	x_{10}	x_5	x_8	-	x_{14}	x_6	x_2
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_1	x_{13}	x_7	x_3	x_{15}	x_4	x_{12}	x_{10}	x_5	-	-	-	-	-
ENETBF	x_{11}	x_1	x_{13}	x_7	x_3	x_{15}	x_4	x_{12}	x_{10}	x_5	x_8	-	x_{14}	x_6	x_2

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 88															
Predictor variables [a]	x_{11}	x_{13}	x_1	x_{15}	x_{12}	x_7	x_{14}	x_3	x_4	x_5	x_9	x_8	x_{10}	x_6	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7340 $corr(x_{12}, x_{15})$ =0.7138 $corr(x_{14}, x_{15})$ =0.6629														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{13}	x_1	x_{15}	x_{12}	x_7	x_{14}	x_3	x_4	-	-	-	-	-	-
ENETCV	x_{11}	x_{13}	x_1	x_{15}	x_{12}	x_7	x_{14}	x_3	-	x_5	-	-	x_{10}	x_6	x_2
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{13}	x_1	x_{15}	x_{12}	x_7	x_{14}	x_3	x_4	-	x_9	-	-	-	-
ENETBF	x_{11}	x_{13}	x_1	x_{15}	x_{12}	x_7	x_{14}	x_3	-	x_5	x_9	-	x_{10}	x_6	x_2

Data no. 89															
Predictor variables [a]	x_1	x_{11}	x_{12}	x_{13}	x_{15}	x_{14}	x_4	x_3	x_7	x_8	x_{10}	x_6	x_5	x_9	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7155 $corr(x_{14}, x_{15})$ =0.7080 $corr(x_{12}, x_{14})$ =0.6974														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{12}	x_{13}	x_{15}	x_{14}	x_4	x_3	x_7	x_8	x_{10}	-	-	-	-
ENETCV	x_1	x_{11}	x_{12}	-	x_{15}	-	x_4	x_3	x_7	x_8	x_{10}	-	x_5	x_9	-
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{12}	x_{13}	x_{15}	x_{14}	x_4	x_3	x_7	x_8	-	-	-	-	-
ENETBF	x_1	x_{11}	x_{12}	x_{13}	x_{15}	x_{14}	x_4	x_3	x_7	x_8	x_{10}	x_6	x_5	x_9	-

Data no. 90															
Predictor variables [a]	x_1	x_{13}	x_{11}	x_{12}	x_7	x_{14}	x_{15}	x_5	x_2	x_8	x_9	x_3	x_{10}	x_4	x_6
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ = 0.7600 $corr(x_{12}, x_{15})$ =0.7033 $corr(x_{14}, x_{15})$ =0.5595														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{13}	x_{11}	x_{12}	x_7	x_{14}	x_{15}	x_5	x_2	x_8	x_9	-	-	-	-
ENETCV	x_1	x_{13}	x_{11}	-	x_7	x_{14}	x_{15}	-	x_2	x_8	x_9	-	-	x_4	-
λ_2 based on Bayes factor															
BVSBF	x_1	x_{13}	x_{11}	x_{12}	x_7	x_{14}	x_{15}	x_5	x_2	x_8	x_9	x_3	-	-	-
ENETBF	x_1	x_{13}	x_{11}	x_{12}	x_7	x_{14}	x_{15}	x_5	x_2	x_8	x_9	x_3	x_{10}	x_4	x_6

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 91															
Predictor variables [a]	x_7	x_1	x_{13}	x_{11}	x_{12}	x_{15}	x_{14}	x_6	x_5	x_4	x_2	x_8	x_9	x_3	x_{10}
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7712 $corr(x_{12}, x_{15})$ =0.6881 $corr(x_{12}, x_{14})$ =0.6499														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_7	x_1	x_{13}	x_{11}	x_{12}	x_{15}	x_{14}	x_6	-	-	-	-	-	-	-
ENETCV	x_7	x_1	x_{13}	x_{11}	x_{12}	x_{15}	x_{14}	x_6	-	-	x_2	-	x_9	x_3	-
λ_2 based on Bayes factor															
BVSBF	x_7	x_1	x_{13}	x_{11}	x_{12}	x_{15}	x_{14}	x_6	x_5	x_4	x_2	-	-	-	-
ENETBF	x_7	x_1	x_{13}	x_{11}	x_{12}	x_{15}	x_{14}	x_6	x_5	x_4	x_2	x_8	x_9	x_3	x_{10}

Data no. 92															
Predictor variables [a]	x_1	x_7	x_{13}	x_{11}	x_3	x_{14}	x_{15}	x_{12}	x_4	x_2	x_5	x_8	x_6	x_{10}	x_9
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.8425 $corr(x_{12}, x_{15})$ =0.8133 $corr(x_{14}, x_{15})$ =0.7763														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_7	x_{13}	x_{11}	x_3	x_{14}	x_{15}	x_{12}	x_4	x_2	-	-	-	-	-
ENETCV	x_1	x_7	x_{13}	x_{11}	x_3	-	x_{15}	-	x_4	-	x_5	-	x_6	x_{10}	x_9
λ_2 based on Bayes factor															
BVSBF	x_1	x_7	x_{13}	x_{11}	x_3	x_{14}	x_{15}	x_{12}	x_4	x_2	-	-	-	-	-
ENETBF	x_1	x_7	x_{13}	x_{11}	x_3	x_{14}	x_{15}	-	x_4	x_2	x_5	-	x_6	x_{10}	x_9

Data no. 93															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_{12}	x_2	x_3	x_{10}	x_4	x_6	x_5	x_8	x_9	x_7
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7083 $corr(x_{12}, x_{14})$ =0.6954 $corr(x_{12}, x_{15})$ =0.6772														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_{12}	x_2	x_3	x_{10}	x_4	x_6	-	-	-	-
ENETCV	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_{12}	x_2	x_3	x_{10}	-	-	x_5	x_8	x_9	x_7
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_{12}	x_2	x_3	x_{10}	x_4	x_6	x_5	-	-	-
ENETBF	x_1	x_{11}	x_{13}	x_{14}	x_{15}	x_{12}	x_2	x_3	x_{10}	x_4	-	x_5	x_8	x_9	x_7

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3,\ldots,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 94															
Predictor variables [a]	x_1	x_{11}	x_{13}	x_4	x_{14}	x_{12}	x_{15}	x_7	x_9	x_3	x_{10}	x_6	x_8	x_5	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7612 $corr(x_{12}, x_{14})$ =0.6948 $corr(x_{14}, x_{15})$ =0.6609														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{13}	x_4	x_{14}	x_{12}	x_{15}	x_7	x_9	x_3	x_{10}	x_6	-	-	-
ENETCV	x_1	x_{11}	x_{13}	x_4	x_{14}	-	-	x_7	x_9	x_3	-	x_6	x_8	x_5	x_2
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{13}	x_4	x_{14}	x_{12}	x_{15}	x_7	x_9	x_3	x_{10}	x_6	-	-	-
ENETBF	x_1	x_{11}	x_{13}	x_4	-	-	x_{15}	x_7	x_9	x_3	x_{10}	x_6	x_8	x_5	x_2

Data no. 95															
Predictor variables [a]	x_1	x_{12}	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	x_9	x_5	x_3	x_8	x_6	x_2	x_{10}
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.7725 $corr(x_{12}, x_{15})$ =0.7652 $corr(x_{12}, x_{14})$ =0.7060														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{12}	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	x_9	-	-	-	-	-	-
ENETCV	x_1	x_{12}	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	-	x_5	-	x_8	-	x_2	x_{10}
λ_2 based on Bayes factor															
BVSBF	x_1	x_{12}	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	x_9	x_5	x_3	-	-	-	-
ENETBF	x_1	x_{12}	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	x_9	x_5	x_3	x_8	x_6	x_2	x_{10}

Data no.96															
Predictor variables [a]	x_1	x_7	x_{11}	x_{13}	x_2	x_{12}	x_{14}	x_6	x_4	x_9	x_{15}	x_{10}	x_8	x_3	x_5
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7393 $corr(x_{12}, x_{15})$ =0.6927 $corr(x_{14}, x_{15})$ =0.6895														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_7	x_{11}	x_{13}	x_2	x_{12}	x_{14}	x_6	x_4	x_9	-	-	-	-	-
ENETCV	x_1	x_7	x_{11}	x_{13}	x_2	-	x_{14}	x_6	x_4	x_9	-	x_{10}	x_8	-	x_5
λ_2 based on Bayes factor															
BVSBF	x_1	x_7	x_{11}	x_{13}	x_2	x_{12}	x_{14}	x_6	x_4	x_9	x_{15}	-	-	-	-
ENETBF	x_1	x_7	x_{11}	x_{13}	x_2	x_{12}	x_{14}	x_6	x_4	x_9	x_{15}	x_{10}	x_8	x_3	x_5

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 97															
Predictor variables [a]	x_1	x_{11}	x_{12}	x_{13}	x_{15}	x_7	x_4	x_{14}	x_2	x_6	x_3	x_5	x_9	x_{10}	x_8
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.8360 $corr(x_{12}, x_{15})$ =0.7898 $corr(x_{14}, x_{15})$ =0.6442														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{12}	x_{13}	x_{15}	x_7	x_4	x_{14}	x_2	x_6	-	-	-	-	-
ENETCV	x_1	x_{11}	x_{12}	x_{13}	-	x_7	x_4	x_{14}	x_2	x_6	x_3	x_5	x_9	x_{10}	-
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{12}	x_{13}	x_{15}	x_7	x_4	x_{14}	x_2	x_6	-	-	-	-	-
ENETBF	x_1	x_{11}	x_{12}	x_{13}	x_{15}	x_7	x_4	x_{14}	x_2	x_6	x_3	x_5	x_9	x_{10}	-

Data no. 98															
Predictor variables [a]	x_{11}	x_{14}	x_1	x_{13}	x_{15}	x_3	x_8	x_{12}	x_4	x_7	x_{10}	x_6	x_9	x_5	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{15})$ =0.7449 $corr(x_{12}, x_{14})$ =0.6878 $corr(x_{14}, x_{15})$ =0.6876														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_{11}	x_{14}	x_1	x_{13}	x_{15}	x_3	x_8	-	-	x_7	-	-	-	-	-
ENETCV	x_{11}	x_{14}	x_1	x_{13}	-	x_3	x_8	-	-	x_7	-	x_6	x_9	x_5	-
λ_2 based on Bayes factor															
BVSBF	x_{11}	x_{14}	x_1	x_{13}	x_{15}	x_3	x_8	x_{12}	x_4	x_7	-	-	-	-	-
ENETBF	x_{11}	x_{14}	x_1	x_{13}	-	x_3	x_8	x_{12}	-	x_7	x_{10}	x_6	x_9	x_5	x_2

Data no. 99															
Predictor variables [a]	x_1	x_{12}	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	x_5	x_6	x_3	x_{10}	x_9	x_8	x_2
Correlation coefficient between predictors	$corr(x_{12}, x_{14})$ =0.7771 $corr(x_{14}, x_{15})$ =0.7615 $corr(x_{12}, x_{15})$ =0.7713														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{12}	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	-	x_6	-	-	-	-	-
ENETCV	x_1	x_{12}	x_{11}	x_{13}	x_{14}	-	x_4	x_7	-	x_6	x_3	-	x_9	x_8	-
λ_2 based on Bayes factor															
BVSBF	x_1	x_{12}	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	x_5	x_6	x_3	x_{10}	-	-	-
ENETBF	x_1	x_{12}	x_{11}	x_{13}	x_{14}	x_{15}	x_4	x_7	x_5	x_6	x_3	x_{10}	x_9	x_8	-

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.

Table J Variable selection results of BVSCV, BVSBF, and elastic net (Cont.)

Data no. 100															
Predictor variables [a]	x_1	x_{11}	x_{14}	x_{12}	x_{13}	x_{15}	x_7	x_5	x_{10}	x_4	x_6	x_9	x_8	x_3	x_2
Correlation coefficient between predictors	$corr(x_{14}, x_{15})$ =0.8077 $corr(x_{12}, x_{15})$ =0.7433 $corr(x_{12}, x_{14})$ =0.7397														
λ_1 and λ_2 are estimated by the 10-fold cross-validation method															
BVSCV	x_1	x_{11}	x_{14}	x_{12}	x_{13}	-	x_7	-	x_{10}	-	-	-	-	-	-
ENETCV	x_1	x_{11}	x_{14}	x_{12}	x_{13}	-	x_7	x_5	x_{10}	x_4	-	-	-	-	-
λ_2 based on Bayes factor															
BVSBF	x_1	x_{11}	x_{14}	x_{12}	x_{13}	x_{15}	x_7	x_5	x_{10}	x_4	x_6	-	-	-	-
ENETBF	x_1	x_{11}	x_{14}	-	x_{13}	x_{15}	x_7	x_5	x_{10}	x_4	x_6	x_9	x_8	-	x_2

[a] Predictor variables are in descending order according to the correlation coefficient between y and x_j ; $j = 1,2,3, \ldots ,15$.

Appendix K

Pearson correlation and partial correlation coefficients for the structure of the simulation dataset 1

Table K True values of the Pearson correlation and the partial correlation coefficients (in absolute values) for the structure of the simulation dataset 1

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
$corr(y, x_j)$	0.5511	0.1641	0.2385	0.2763	0.1691	0.1243	0.3150	0.1234	0.1147	0.1128	0.5507	0.4149	0.4863	0.4374	0.4318
$corr(y, x_j \mid x_{\setminus j})$	0.5554	0.1482	0.1436	0.1471	0.1864	0.1385	0.4856	0.1457	0.1233	0.1389	0.3921	0.1530	0.2134	0.1246	0.1359

The resulting pairwise true correlation between the predictor variables are:

$corr(x_1, x_{11})$ = 0.1762, $corr(x_1, x_{12})$ = 0.1673, $corr(x_1, x_{13})$ = 0.1670,
$corr(x_1, x_{14})$ = 0.1779, $corr(x_1, x_{15})$ = 0.1843,

$corr(x_2, x_{11})$ = 0.2545, $corr(x_2, x_{12})$ = 0.2607, $corr(x_2, x_{13})$ = 0.2566,
$corr(x_2, x_{14})$ = 0.2554, $corr(x_2, x_{15})$ = 0.2511,

$corr(x_3, x_{11})$ = 0.3191, $corr(x_3, x_{12})$ = 0.3237, $corr(x_3, x_{13})$ = 0.3352,
$corr(x_3, x_{14})$ = 0.3418, $corr(x_3, x_{15})$ = 0.3268,

$corr(x_4, x_{11})$ = 0.4339, $corr(x_4, x_{12})$ = 0.4420, $corr(x_4, x_{13})$ = 0.4426,
$corr(x_4, x_{14})$ = 0.4428, $corr(x_4, x_{15})$ = 0.4522,

$corr(x_5, x_{11})$ = 0.5626, $corr(x_5, x_{12})$ = 0.5660, $corr(x_5, x_{13})$ = 0.5574,
$corr(x_5, x_{14})$ = 0.5409, $corr(x_5, x_{15})$ = 0.5454,

$corr(x_{11}, x_{12})$ = 0.7253, $corr(x_{11}, x_{13})$ = 0.7309, $corr(x_{11}, x_{14})$ = 0.7300,
$corr(x_{11}, x_{15})$ = 0.7247,

$corr(x_{12}, x_{13})$ = 0.7332, $corr(x_{12}, x_{14})$ = 0.7365, $corr(x_{12}, x_{15})$ = 0.7321,

$corr(x_{13}, x_{14})$ = 0.7383, $corr(x_{13}, x_{15})$ = 0.7324, $corr(x_{14}, x_{15})$ = 0.7276.

Appendix L

Pearson correlation and partial correlation coefficients for real datasets

The two datasets are the diabetes dataset and prostate cancer data which are used in elastic net literature and related methods.

L.1 Diabetes data ($p = 10, n = 442$)

The response variable (y) is a quantitative measure of disease progression one year after baseline for 442 diabetes patients. The dataset contains 10 baseline predictor variables: AGE, SEX, body mass index (BMI), average blood pressure (BP), and six blood serum measurements: tc(S1), ldl(S2), hdl(S3), tch(S4), ltg(S5), glu(S6).

Table L.1 Correlation coefficients between the response and predictor variables for diabetes data

	y	AGE	SEX	BMI	BP	S1	S2	S3	S4	S5	S6
y	1										
AGE	0.188	1									
SEX	0.043	0.174	1								
BMI	0.586	0.185	0.088	1							
BP	0.441	0.335	0.241	0.395	1						
S1	0.212	0.260	0.035	0.250	0.242	1					
S2	0.174	0.219	0.143	0.261	0.186	0.897	1				
S3	-0.395	-0.075	-0.379	-0.367	-0.179	0.052	-0.196	1			
S4	0.430	0.204	0.332	0.414	0.258	0.542	0.660	-0.738	1		
S5	0.566	0.271	0.150	0.446	0.393	0.516	0.318	-0.398	0.618	1	
S6	0.382	0.302	0.208	0.389	0.390	0.326	0.291	-0.274	0.417	0.465	1

Table L.2 Partial correlation coefficients between the response and predictor variables for diabetes data

control variables	Partial correlation coefficients	
	predictors	y
SEX, BMI, BP, S1, S2, S3, S4, S5, S6	AGE	-0.008
AGE, BMI, BP, S1, S2, S3, S4, S5, S6	SEX	-0.185
AGE, SEX, BP, S1, S2, S3, S4, S5, S6	BMI	0.352
AGE, SEX, BMI, S1, S2, S3, S4, S5, S6	BP	0.232
AGE, SEX, BMI, BP, S2, S3, S4, S5, S6	S1	-0.092
AGE, SEX, BMI, BP, S1, S3, S4, S5, S6	S2	0.068
AGE, SEX, BMI, BP, S1, S2, S4, S5, S6	S3	0.023
AGE, SEX, BMI, BP, S1, S2, S3, S5, S6	S4	0.053
AGE, SEX, BMI, BP, S1, S2, S3, S4, S6	S5	0.206
AGE, SEX, BMI, BP, S1, S2, S3, S4, S5	S6	0.049

L.2 Prostate cancer data ($p = 8, n = 97$)

The response variable (y) is the logarithm of prostate specific antigen (lpsa). The predictor variables are eight clinical measures:

X1 - the logarithm of cancer volume (lcavol),

X2 - the logarithm of prostate weight (lweight),

X3 - age,

X4 - the logarithm of the amount of benign prostatic hyperplasia (lbph),

X5 - seminal vesicle invasion (svi),

X6 - the logarithm of capsular penetration (lcp),

X7 - the Gleason score (gleason), and

X8 - the percentage Gleason score 4 or 5 (pgg45).

Table L.3 Correlation coefficients between the response and predictor variables for prostate cancer data

	y	X1	X2	X3	X4	X5	X6	X7	X8
y	1								
X1	0.7345	1							
X2	0.4333	0.2805	1						
X3	0.1696	0.2250	0.3480	1					
X4	0.1798	0.0273	0.4423	0.3502	1				
X5	0.5662	0.5388	0.1554	0.1177	-0.0858	1			
X6	0.5488	0.6753	0.1645	0.1277	-0.0070	0.6731	1		
X7	0.3690	0.4324	0.0565	0.2689	0.0778	0.3204	0.5148	1	
X8	0.4223	0.4337	0.1074	0.2761	0.0785	0.4576	0.6315	0.7519	1

Table L.4 Partial correlation coefficients between the response and predictor variables for prostate cancer data

control variables	Partial correlation coefficients	
	predictors	y
X2, X3, X4, X5, X6, X7, X8	X1	0.565
X1, X3, X4, X5, X6, X7, X8	X2	0.313
X1, X2, X4, X5, X6, X7, X8	X3	-0.200
X1, X2, X3, X5, X6, X7, X8	X4	0.175
X1, X2, X3, X4, X6, X7, X8	X5	0.319
X1, X2, X3, X4, X5, X7, X8	X6	-0.125
X1, X2, X3, X4, X5, X6, X8	X7	0.034
X1, X2, X3, X4, X5, X6, X7	X8	0.108

Appendix M

Sample correlation coefficients for the structure of the simulation dataset 2

For simulation dataset 2, there are six cases for combination of n = 100, 200, 400 and ρ = 0.5, 0.75. The simulation method is repeated 100 times.

The percentage of the difference between the values of ρ and sample correlation coefficients is

$$\frac{|\rho_i - r_i|}{\rho_i} \times 100 ,$$

where ρ_i be the population correlation coefficient of the ith dataset, and r_i be the sample correlation coefficient of the ith dataset. Table M shows the sample correlation coefficients for the structure of the simulation dataset 2.

Table M Sample correlation coefficients for the structure of the simulation dataset 2

n	ρ	Average of sample correlation coefficients	Average of the percentage of the difference between the values of ρ and sample correlation coefficients
100	0.5	0.4900	12.1899
100	0.75	0.7500	5.1098
200	0.5	0.4939	9.0304
200	0.75	0.7493	3.2725
400	0.5	0.4974	5.7847
400	0.75	0.7525	2.3785

BIOGRAPHY

Name	Miss Kanyalin Jiratchayut
Date of birth	October 13, 1973
Educational Attainment	Academic Year 1994: Bachelor of Science (Mathematics), Mahidol University, Thailand Academic Year 1998: Master of Science (Applied Statistics), Thammasat University, Thailand
Work Position	Assistant Professor at Faculty of Science and Arts, Burapha University, Chantaburi Campus, Thailand
Scholarships	Fiscal Year 2014: financial support provided by Thammasat University Research Fund under the TU Research Scholar, Contract No.30/2557 Academic Year 2010 – 2012: scholarship provided by Faculty of Science and Arts, Burapha University, Chantaburi Campus, Thailand Academic Year 1998: scholarship provided by the University Affairs, Thailand

Publications

Jiratchayut, K., & Bumrungsup, C. (in press). A study of adaptive elastic net estimators with different adaptive weights. *Thammasat International Journal of Science and Technology, 20*(3).

Jiratchayut, K., & Leadprathom, N. (2011). *The survey on pesticide using in difference land use in Chantaburi river area (phase II)*. Research report (published in Thai language), Burapha University.

Jiratchayut, K., & Leadprathom, N. (2010). *The survey on pesticide using in difference land use in Chantaburi river area*. Research report (published in Thai language), Burapha University.

Jiratchayut, K. (2009). *Factors affecting the first year students in selecting to study at Faculty of Science and Arts, Burapha University, Chantaburi Campus, academic year 2008*. Research report (published in Thai language), Burapha University, Faculty of Science and Arts.

Jiratchayut, K. (2006). *Elementary Statistics*. n.p. (published in Thai language.)

Jiratchayut, K. (2005). *Biostatistics*. n.p. (published in Thai language.)

Work Experiences	2006 – Present: Assistant Professor at Faculty of Science and Arts, Burapha University, Chantaburi Campus, Thailand 2000 – 2008: Part time Lecturer at Faculty of Science and Technology, Rambhai Barni Rajabhat University, Chantaburi, Thailand 1999 – 2006: Lecturer at Faculty of Science and Arts, Burapha University, Chantaburi Campus, Thailand