Variable selection for longitudinal and irregular high-dimensional data

by

Yuwen Liu

M.A., University of Missouri - Columbia, 2017

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2024

Abstract

Variable selection is a commonly used method for analyzing genomic data with high dimensionality. It has been designed to handle complicated data structures and facilitate the identification of crucial genomic features by creating sparse models. Despite the success of existing studies, challenges still remain when the response variables are repeatedly measured or have heavy-tailed distributions. These challenges have motivated the development of novel variable selection methods proposed in the following projects from both frequentist and Bayesian perspectives.

In longitudinal studies, regularized variable selection methods have been extensively developed while accommodating the intra-correlation among repeated measurements. Despite the success, they are limited especially in accommodating structured sparsity. For example, strong correlations generally exist among omics features. Ignoring such a correlation while conducting variable selection in longitudinal studies results in false identification and biased estimation. In the first project, we have proposed a network-based variable selection method under repeatedly measured disease phenotype. The strong interconnections among the omics predictors have been efficiently accommodated while performing variable selection. An efficient Newton-Raphson based algorithm was adopted within the generalized estimating equation (GEE) framework. The advantage of the proposed method has been demonstrated in extensive simulations and a study of the Childhood Asthma Management Program (CAMP) with high dimensional single nucleotide polymorphisms (SNPs) data.

With the vast development of biotechnologies, high-dimensional omics data from different biological systems have been collected and analyzed in order to detect the associated genes responsible for certain diseases. However, this also brings challenges to the process of analysis since the number of candidate genetic features is far greater than the sample size and phenotype data can contain heavy-tailed distributions. In the second project, we proposed

the spike-and-slab quantile LASSO for prediction and variable selection on large-scale genotype data. Our model used a spike-and-slab prior with a mixture of two double-exponential distributions. This allows us to keep the important effects while shrinking the coefficients of irrelevant features to exact zero. The expectation-maximization (EM) algorithm based on cyclic coordinate descent was used to fit the spike-and-slab quantile regression. Performance of the proposed method and other commonly used approach including LASSO, quantile LASSO as well as spike-and-slab LASSO was compared by measuring the identification and estimation accuracy during extensive simulation studies. The result indicates that the proposed approach delivers more precise variable selection results and prediction outcomes. We also applied the proposed procedure to the analysis of the lung adenocarcinomas (LUAD) and skin cutaneous melanoma (SKCM) data from The Cancer Genome Atlas (TCGA). Outcome of the analysis shows that the proposed method selects disease related genes and yields a more accurate prediction.

In the third project, we have developed a nonparametric spike-and-slab quantile group LASSO based mixed model for high-dimensional time-varying SNP effects under longitudinal traits. The nonparametric time varying effects are effectively modeled through varying coefficient models. The splines for the varying coefficients after basis expansion are selected through the spike-and-slab quantile group LASSO that has integrated the merits from Bayesian and frequentist group LASSO, effectively circumventing their individual drawbacks. In addition, the intra-cluster correlations are accommodated through the random effects. We have established the advantage of the proposed model over Bayesian and frequentist alternatives through a diversity of settings under data contamination and model misspecifications. On the longitudinal GWAS data of asthma, the proposed method has yielded better prediction results in time-varying heritability plot along with SNPs with important biological implications.

Overall, for repeated measurement data, we have developed a network-based penalized variable selection method under the GEE framework for analyzing high-dimensional single nucleotide polymorphism (SNP) data combining with the longitudinal features and responses. To accommodate the nonlinear trend, we have also developed a nonparametric

spike-and-slab quantile group LASSO based mixed model for high-dimensional time-varying SNP effects under longitudinal traits. Under univariate setting, we have developed a novel variable selection method applying spike-and-slab LASSO prior to Bayesian quantile regression and constructed the spike-and-slab quantile LASSO (ssQLASSO) to analyze clinical trial data set containing high-dimensional genotype data and heavy tail distributed phenotype data.

Open source R package[1] has been constructed for the projects to enable reproducible research and quick computation. It offers efficient C++ implementation for all the proposed methods and alternative approaches. The package is available on CRAN.

Variable selection for longitudinal and irregular high-dimensional data

by

Yuwen Liu

M.A., University of Missouri - Columbia, 2017

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2024

Approved by:

Major Professor
Cen Wu

Copyright

Abstract

Variable selection is a commonly used method for analyzing genomic data with high dimensionality. It has been designed to handle complicated data structures and facilitate the identification of crucial genomic features by creating sparse models. Despite the success of existing studies, challenges still remain when the response variables are repeatedly measured or have heavy-tailed distributions. These challenges have motivated the development of novel variable selection methods proposed in the following projects from both frequentist and Bayesian perspectives.

In longitudinal studies, regularized variable selection methods have been extensively developed while accommodating the intra-correlation among repeated measurements. Despite the success, they are limited especially in accommodating structured sparsity. For example, strong correlations generally exist among omics features. Ignoring such a correlation while conducting variable selection in longitudinal studies results in false identification and biased estimation. In the first project, we have proposed a network-based variable selection method under repeatedly measured disease phenotype. The strong interconnections among the omics predictors have been efficiently accommodated while performing variable selection. An efficient Newton-Raphson based algorithm was adopted within the generalized estimating equation (GEE) framework. The advantage of the proposed method has been demonstrated in extensive simulations and a study of the Childhood Asthma Management Program (CAMP) with high dimensional single nucleotide polymorphisms (SNPs) data.

With the vast development of biotechnologies, high-dimensional omics data from different biological systems have been collected and analyzed in order to detect the associated genes responsible for certain diseases. However, this also brings challenges to the process of analysis since the number of candidate genetic features is far greater than the sample size and phenotype data can contain heavy-tailed distributions. In the second project, we proposed

the spike-and-slab quantile LASSO for prediction and variable selection on large-scale genotype data. Our model used a spike-and-slab prior with a mixture of two double-exponential distributions. This allows us to keep the important effects while shrinking the coefficients of irrelevant features to exact zero. The expectation-maximization (EM) algorithm based on cyclic coordinate descent was used to fit the spike-and-slab quantile regression. Performance of the proposed method and other commonly used approach including LASSO, quantile LASSO as well as spike-and-slab LASSO was compared by measuring the identification and estimation accuracy during extensive simulation studies. The result indicates that the proposed approach delivers more precise variable selection results and prediction outcomes. We also applied the proposed procedure to the analysis of the lung adenocarcinomas (LUAD) and skin cutaneous melanoma (SKCM) data from The Cancer Genome Atlas (TCGA). Outcome of the analysis shows that the proposed method selects disease related genes and yields a more accurate prediction.

In the third project, we have developed a nonparametric spike-and-slab quantile group LASSO based mixed model for high-dimensional time-varying SNP effects under longitudinal traits. The nonparametric time varying effects are effectively modeled through varying coefficient models. The splines for the varying coefficients after basis expansion are selected through the spike-and-slab quantile group LASSO that has integrated the merits from Bayesian and frequentist group LASSO, effectively circumventing their individual drawbacks. In addition, the intra-cluster correlations are accommodated through the random effects. We have established the advantage of the proposed model over Bayesian and frequentist alternatives through a diversity of settings under data contamination and model misspecifications. On the longitudinal GWAS data of asthma, the proposed method has yielded better prediction results in time-varying heritability plot along with SNPs with important biological implications.

Overall, for repeated measurement data, we have developed a network-based penalized variable selection method under the GEE framework for analyzing high-dimensional single nucleotide polymorphism (SNP) data combining with the longitudinal features and responses. To accommodate the nonlinear trend, we have also developed a nonparametric

spike-and-slab quantile group LASSO based mixed model for high-dimensional time-varying SNP effects under longitudinal traits. Under univariate setting, we have developed a novel variable selection method applying spike-and-slab LASSO prior to Bayesian quantile regression and constructed the spike-and-slab quantile LASSO (ssQLASSO) to analyze clinical trial data set containing high-dimensional genotype data and heavy tail distributed phenotype data.

Open source R package[1] has been constructed for the projects to enable reproducible research and quick computation. It offers efficient C++ implementation for all the proposed methods and alternative approaches. The package is available on CRAN.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my deepest gratitude to my research supervisor, Dr. Cen Wu, whose expertise and knowledge were invaluable during these research projects. Dr. Wu's profound knowledge and dedication to mentorship have been crucial to my development and the success of my research project. He always points me in the right direction when I get stuck on research projects. Without him this dissertation would not have been possible.

My sincere thanks also go to my committee members, Dr. Michael Higgins, Dr. Juan Du and Dr. Weiqun Wang for their advice and support, which made this project possible. Dr. Higgins provided many innovative ideas and suggestions on how to improve my dissertation. My presentation time for Dr. Du was late at night due to the time difference. However, she still managed to make time for me and provided meaningful comments. Dr. Wang provided lots of insights from the real data aspect.

I am deeply thankful to my former group members, Fei Zhou and Jie Ren, for their insightful academic support. Their code and software packages greatly assisted me in developing my own modeling algorithms.

Lastly, I want to acknowledge my family and friends for their unwavering love and encouragement. This project would not have been possible without their kind support.

Dedication

I dedicate this dissertation to my loving wife, Rigele Te. She set an excellent example for me by successfully completing her own dissertation. I really appreciate all the love and support she has given to me over the past years.

And to my parents. I could not accomplish this task without their financial and emotional support.

Chapter 1

Introduction

With the rapid progress made in the advance of biotechnology, massive amount of genomics data were collected and analyzed in order to develop personalized treatments, for example the T-cell gene therapy (Schmitt et al., 2009)[2] for treating cancer. Genotype dataset such as single nucleotide polymorphysm (SNP) data can be viewed as a high-dimensional matrix with its columns corresponding to different SNPs as variables and rows corresponding to subjects. Due to the fact that there are millions of different genetic information containing in a person's genome, the number of genetic features measured in a genotype dataset are almost always much larger than the sample size. The high dimensionality of these high-throughput data became a problem for many statistical analysis methods. However, among these many features, only a small portion is associated with the response. To overcome the curse of dimensionality and correctly choose the important features, regularization method was used to perform variable selection. For example, the least absolute shrinkage and selection operator (LASSO) (Tibshirani, 1996)[3] has been widely used for analyzing high-dimensional data. It can effectively screen out the unimportant features by shrinking all the coefficients towards zero. Inspired by this idea, many other variable selection methods for high-dimensional data were developed in later times. Besides high dimensionality, repeated measurement data structure is another important feature in clinical trial data analysis. Longitudinal data are collected by measuring the same subject repeatedly over time which leads to the cor-

related nature of the data. To account for this correlation, different modeling approaches were proposed. Next, we will briefly present some of the popular longitudinal data analysis and high-dimensional variable selection methods from both frequentist and Bayesian perspectives.

1.1 Penalized Variable Selection

Penalization is one of the most widely used variable selection method in high-dimensional data analysis. The regularized regression model can be expressed as

$$L(\beta; X, Y) + q_\lambda(|\beta|)) \tag{1.1}$$

where Y is the response, usually a continues phenotype response and X denotes a $n \times p$-dimensional design matrix containing genomic features. $L(\beta; X, Y)$ is a loss function that measures the appropriateness of the model in terms of being able to predict the expected outcome. $q_\lambda(|\beta|))$ is a penalty function determined by the tuning parameter λ. It shrinks the coefficient vector $\beta = (\beta_1, \cdots, \beta_p)^\top$ towards zero as λ increases. By selecting the optimal tuning, regression coefficients of less important features shrunk to zero and was automatically removed from the model. The variable selection process can be performed by minimizing the penalized loss function (1.1).

During the modeling process, penalization method can be chosen based on the structure of the design matrix. LASSO is the most commonly used regularization method. It consists of a least square regression and an $L1$ penalty term:

$$\|Y - X\beta\|_2^2 + \lambda \sum_{j=1}^{p} |\beta_j|$$

The LASSO procedure returns a continuous penalized estimator which shrinks small estimated coefficients to zero. Three criteria was proposed by Fan and Li (2001)[4] for defining a good penalty function. A qualified penalty function should result in a continuous estimator

(continuity), while shrinking small estimated coefficients to zero (sparsity). It should also maintain an unbiased estimates for large coefficients (unbiasedness). The LASSO penalty satisfies the property of continuity and sparsity. However, when the magnitude of regression coefficients is large, LASSO yields biased estimates since it penalizes all coefficients with the same intensity. In order to overcome the problem of result being biased, researchers tried various modifications then came up with the smoothly clipped absolute deviation (SCAD) (Fan and Li, 2001)[4] and the minimax concave penalty (MCP) (Zhang, 2010)[5]. They are defined as follows:

$$SCAD: q_{\lambda,\gamma}(\beta_j) = \begin{cases} \lambda|\beta_j| & |\beta_j| \leq \lambda \\ -\frac{\beta_j^2 - 2\gamma\lambda|\beta_j| + \lambda^2}{2(\gamma-1)} & \lambda \leq |\beta_j| \leq \gamma\lambda \\ \frac{1}{2}(\gamma+1)\lambda^2 & |\beta_j| > \gamma\lambda \end{cases}$$

$$MCP: q_{\lambda,\gamma}(\beta_j) = \begin{cases} \lambda|\beta_j| - (2\gamma)^{-1}\beta_j^2 & |\beta_j| \leq \gamma\lambda \\ \frac{1}{2}\gamma\lambda^2 & |\beta_j| > \gamma\lambda \end{cases}$$

where the regularization parameter $\gamma > 2$ for SCAD and $\gamma > 1$ for MCP. These methods also achieve continuity and sparsity. Moreover, they produce nearly unbiased results. Researchers often use LASSO, SCAD and MCP as baseline penalization methods for testing their own models.

Another important factor that requires attention is the group structure. The group LASSO (Yuan and Lin, 2006)[6] can be used to perform group-level variable selection. It is given as:

$$\|Y - X\beta\|_2^2 + \lambda \sum_{g=1}^{G} \sqrt{L_g}\|\beta_g\|_2,$$

where $\beta_g = (\beta_{g1}, \cdots, \beta_{gL_g})^\top$ indicates the coefficient vector of group g of length L_g. The entire coefficient vector $\beta = (\beta_1^\top, \cdots, \beta_G^\top)^\top$.The group LASSO has a "group-in, group-out" manner and accounts for sparsity on a group level.

When analyzing genomics data, other complex data structures were encountered besides

high dimensionality and group structure. One of them is the high network structured correlation which are commonly observed among omics features. Elastics net (Zhou and Hastie, 2005)[7] and fused-lasso (Tibshirani et al., 2005)[8] were developed to account for the correlation. Elastic net has the form $(\lambda_1\|\beta\|_2^2 + \lambda_2|\beta|)$. It consists of both l_1 and l_2 penalties. Fused-lasso has the form $(\lambda_1 \sum_{j=1}^{p} |\beta_j| + \lambda_2 \sum_{j=2}^{p} |\beta_j - \beta_{j-1}|)$. They have the ability to simultaneously select correlated variables. However, in order to make use of the correlation information more efficiently, the network-constrained regularization method was proposed. One of them is the sparse Laplacian shrinkage (SLS) penalty (Huang et al., 2011)[9]. It is a combination of two penalization terms and has the form:

$$\sum_{j=1}^{p} q_{\lambda_1,\gamma}(\beta_j) + \lambda_2 \sum_{1\le j<k\le p} |\alpha_{jk}| \left(\beta_j - \operatorname{sgn}(\alpha_{jk})\,\beta_k\right)^2$$

where $q_{\lambda_1,\gamma}(\beta_j)$ is the MCP penalty with λ_1 and γ being its regularization parameters. $|\alpha_{jk}|$ is the network adjacency that measures the strength of connection between the j-th and k-th variables. The MCP penalty part encourages sparsity and the Laplacian quadratic penalty part promotes smoothness among coefficients of correlated features. In sparse and high-dimensional settings with $p \gg n$ under appropriate assumptions, It is shown that the SLS satisfies model selection consistency and has result equal to the oracle Laplacian shrinkage estimator with high probability.

The penalized likelihood can be solved efficiently by using the cyclic coordinate descent (CD) algorithms. It can also be combined with expectation maximization (EM) algorithm to fit large-scale models.

Robust Variable Selection: Another problem often encountered in genomics data analyzation is that the data set containing outliers and phenotype data following heavy-tailed distributions. Robust methods were developed to combat these situations. For regularization regression, one common technique to achieve robustness is through the form "robust loss function + penalization term". The quantile regression loss function, the check loss function, the rank-based loss function and their variants are the frequently used robust methods.

Quantile regularization regression has loss function:

$$L_\tau = \sum_{i=1}^{n} \rho_\tau(y_i - x_i^\top \beta) + \sum_{j=1}^{p} q_\lambda(|\beta_j|),$$

where x_i is the p-dimensional vector for subject i from design matrix, $\rho_\tau(\cdot)$ is the check loss function at quantile τ and w is the residue:

$$\rho_\tau(w) = \begin{cases} \tau w & w \geq 0 \\ -(1-\tau)w & w < 0 \end{cases}.$$

The least absolute deviation (LAD) regularization regression is a special case of quantile regression. It has the form:

$$\sum_{i=1}^{n} |y_i - x_i^\top \beta| + \sum_{j=1}^{p} q_\lambda(|\beta_j|).$$

The $L1$ based loss function reduces the distance between y_i and $x_i^\top \beta$ when the data is contaminated, while the $L2$ based least square loss function yields results with a much larger deviation. There are many studies on robust penalization methods including robust network-based regularization and variable selection method by (Ren et al., 2019)[10]

1.2 Bayesian Variable Selection

In the mean time, the Bayesian perspective of high-dimensional data analysis also gained progress. Bayesian variable selection approaches were grouped into four categories (O'Hara and Sillanpaa, 2009)[11] (1) adaptive shrinkage, (2) indicator model selection, (3) stochastic search variable selection (SSVS) and (4) model space approach. Among these four classes, adaptive shrinkage is highly related to the variable selection methods from the frequentist perspective.

From the Bayesian perspective, we can view the LASSO estimate as a posterior mode with

the regression coefficients taking the independent and identical Laplace prior (Tibshirani, 1996)[3] given as:

$$\pi(\beta) = \prod_{j=1}^{p} \frac{\lambda}{2} e^{-\lambda|\beta_j|}.$$

After imposing an independent prior on σ^2, the posterior distribution has the form:

$$\pi(\beta, \sigma^2|Y) = \pi\left(\sigma^2\right)\left(\sigma^2\right)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\|Y - X\beta\|_2^2 - \lambda\sum_{j=1}^{p}|\beta_j|\right).$$

For any fixed σ^2, the maximum a posterior probability (MAP) estimate for β is the same as the frequentist LASSO estimate. Later in 2008, Bayesian LASSO (Park and Casella, 2008)[12] was proposed by introducing a conditional Laplace prior:

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

where σ^2 has prior $\pi(\sigma^2)$. Kyung er al. (2010)[13] proposed a prior that works with more general forms of LASSO. Fro example, the Bayesian group LASSO can be constructed by using the prior:

$$\pi(\beta_g|\sigma^2) \propto \exp\left(-\frac{\sqrt{L_g}\lambda}{\sigma}\|\beta_g\|_2\right),$$

where β_g contains the coefficients from group g and L_g is the length of group g.

In order to deal with the sparse structure in high-dimensional data analysis and shrink certain coefficients to exact 0, the spike-and-slab prior (Mitchell and Beauchamp, 1988)[14] was proposed. It has the form:

$$\beta_j|\phi_j \sim (1 - \phi_j)\pi_0(\beta_j) + \phi_j\pi_1(\beta_j)$$

where ϕ_j has value 0 or 1 indicating the prior being π_0 or π_1. π_0 is the spike distribution for modeling irrelevant effects having close to 0 coefficients. π_1 is the slab distribution for modeling important features that have larger coefficients. Kuo and Mallick (1998)[15] proposed to replace π_0 with a point mass δ_0, which allows us to shrink coefficients of unimportant

features to exact 0. Figure 1.1 shows the distribution of a spike-and-slab prior with the form

$$\beta_j \sim 0.5\delta_0 + 0.5N(0,1)$$

where $\delta_0 = 0$.

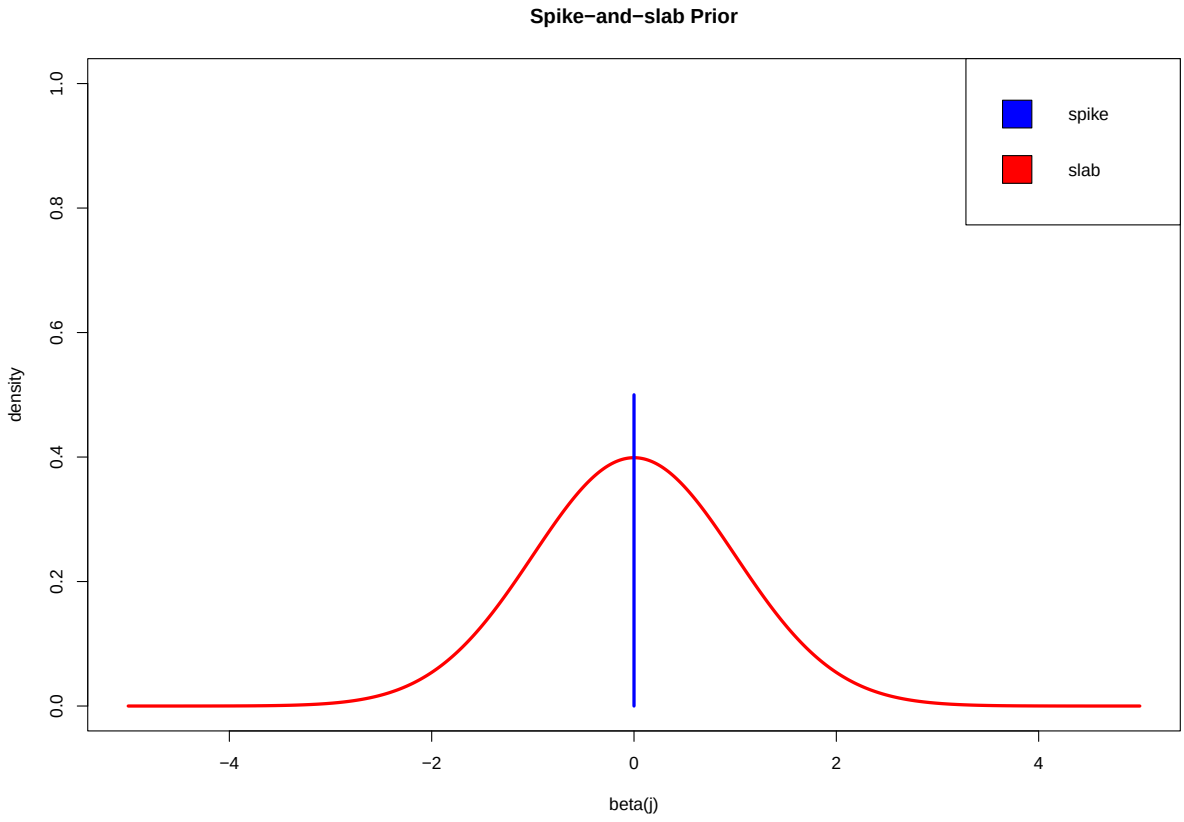

Figure 1.1: Spike-and-slab prior with a point mass and normal distribution

Different distributions can be applied to π_0 and π_1. SSVS (George and McCulloch, 1993)[16] uses a mixture of two normal distributions $N(0, c_j\sigma_j^2)$ and $N(0, \sigma_j^2)$. $N(0, c_j\sigma_j^2)$ will have a variance much smaller than $N(0, \sigma_j^2)$, so that it can be used to model coefficients that are close to zero in a way that resembles the functionality of a point mass. We can also use other distributions in the SSVS method like a mixture of two Laplace distributions (Ročková and George, 2018)[17] which has the form:

$$\beta_j|\phi_j \sim (1-\phi_j)\pi(\beta_j|0, s_0) + \phi_j\pi(\beta_j|0, s_1)$$

where s_0 controls the spike part and s_1 controls the slab part ($s_1 > s_0 > 0$). Laplace distribution has the form:

$$\pi(\beta_j|s) = \frac{1}{2s}\exp\left(-\frac{|\beta_j|}{s}\right).$$

Robust Bayesian Variable Selection: To overcome the challenges of heterogeneity, Li et al. (2010)[18] established a Bayesian framework for regularization in linear quantile regression based on the asymmetric Laplace distribution (ALD) (Yu and Moyeed, 2011)[19] with the form:

$$f(y_i|\mu_i, \sigma, \tau) = \frac{\tau(1-\tau)}{\theta}\exp\left(\rho_\tau\left(\frac{y_i - \mu_i}{\theta}\right)\right)$$

where $\mu_i = x_i^\top \beta$, scale parameter $\theta > 0$, τ is the value of quantile and $\rho_\tau(\cdot)$ is the check loss function

$$\rho_\tau(w) = \begin{cases} \tau w & w \geq 0 \\ -(1-\tau)w & w < 0 \end{cases}.$$

Spike-and-slab LASSO prior can be added to the Bayesian quantile regression to form an hierarchical model. Details will be elaborated in chapter 3. Many other Bayesian shrinkage priors were also developed including Normal Exponential Gamma prior (Griffin and Brown, 2010)[20] and Horseshoe prior (Carvalho et al., 2009)[21]. Besides quantile regression, a Bayesian LAD regression based method was developed by Ren et al. (2023)[22] to accommodate the impact of long-tailed distributions in the phenotype in Gene-environment studies.

1.3 Longitudinal Data Analysis

Longitudinal data are commonly encountered in many scientific areas including economics, biomedical studies and clinical trials. Mixed effects model (Fan and Li, 2012)[23] was proposed with the fixed effects analyzing predictor variables and random effects modeling the correlation structure. It can be achieved through Bayesian hierarchical approach. Another widely used method is the generalized estimating equation (GEE) (Liang and Zeger, 1986)[24].

Generalized estimating equation: Suppose a longitudinal study has n subjects and

m measurements repeatedly taken across time on the i-th subject ($1 \leq i \leq n$). We have the following regression model:

$$Y_{ik} = X_{ik}^{\top}\beta + \epsilon_{ik}$$

Where Y_{ik} is the response of the i-th subject at the k-th time point ($1 \leq j \leq m$). X_{ij} denotes a p-dimensional vector of genetic factors and ϵ is the error term. The estimating equation for coefficient vector β has the form:

$$G(\beta) = \sum_{i=1}^{n} D_i^{\top} V_i^{-1}(Y_i - \mu_i(\beta)),$$

where $D_i^{\top} = \frac{\partial \mu_i(\beta)}{\partial \beta}$, $\mu_i(\beta) = (\mu_{i1}(\beta), \cdots, \mu_{im}(\beta))^{\top}$ and $Y_i = (Y_{i1}, \cdots, Y_{im})^{\top}$. The covariance matrix of the i-th subject is defined as $V_i = A_i^{1/2}(\beta)R_i(\eta)A_i^{1/2}(\beta)$ with $A_i(\beta) = diag\{Var(Y_{i1}), \cdots, Var(Y_{im}))\}$. $R_i(\eta)$ indicates the working correlation matrix (e.g., exchangeable, AR-1, independence...). The solution to equation $G(\beta) = 0$ is the GEE estimator. It is robust to misspecification of correlation due to its sandwich form of the variance estimators. However, GEE is sensitive to outlying observations. To overcome this downside of GEE, Quadratic inference function (QIF) (Qu et al., 2000)[25] was proposed. Its corresponding estimator remains optimal under structure misspecification and is more robust to outliers than GEE. QIF approximates $R(\eta)^{-1}$ using a linear combination of basis matrices with unknown coefficients $b_1, \cdots, b_m$. The score equation of QIF has the form:

$$\sum_{i=1}^{n} D_i^{\top} A_i^{-\frac{1}{2}}(b_1 M_1 + \cdots + b_m M_m) A_i^{-\frac{1}{2}}(Y_i - \mu_i(\beta)),$$

where M_1 is an identity matrix and $M_2, \cdots, M_m$ are a series of symmetric basis matrices. The extended score is defined as

$$Q(\beta) = \frac{1}{n}\sum_{i=1}^{n} Q_i(\beta) = \frac{1}{n}\sum_{i=1}^{n} \begin{pmatrix} D_i^{\top} A_i^{-\frac{1}{2}} M_1 A_i^{-\frac{1}{2}}(Y_i - \mu_i(\beta)) \\ \vdots \\ D_i^{\top} A_i^{-\frac{1}{2}} M_m A_i^{-\frac{1}{2}}(Y_i - \mu_i(\beta)) \end{pmatrix}.$$

The extended score is based on the form of the quasi-score so that we don't have to pick values for parameters $b_1, \cdots, b_m$.

When analyzing high-dimensional data set, longitudinal variable selection can be achieved by adding regularization terms to GEE and QIF score functions. Penalized GEE (Wang et al., 2012)[26] and penalized QIF (Cho and Qu, 2013)[27] has the form:

$$U(\beta) = G(\beta) - q_\lambda(|\beta|)\text{sgn}(\beta)$$

where $q_\lambda(|\beta|) = (q_\lambda(\beta_1), \cdots, q_\lambda(\beta_p))^\top$ and can be any penalization term mentioned above. For example, the MCP penalty is defined as $q(t;\lambda,\gamma) = \lambda \int_0^t \left(1 - \frac{x}{\gamma\lambda}\right)_+ dx$, with tuning parameter λ and regularization parameter γ. Penalized QIF can be achieved through similar procedure. The application has been explored in a Gene-environment interaction variable selection study (Fei et al., 2022)[28].

Mixed model: In matrix notation, a linear mixed model[29] can be represented as

$$Y = X\beta + Zu + \epsilon,$$

where y is a known vector of observation, with mean $E(y) = X\beta$. β is an unknown vector of fixed effects. u is an unknown vector of random effects, with $E(u) = 0$ and $Var(u) = G$. ϵ is an unknown vector of random errors, with $E(\epsilon) = 0$ and $Var(\epsilon) = R$. X is the known design matrix for the fixed effects and Z is the known design matrix for the random effects.

Mixed effect models incorporate both fixed and random effects to accurately represent non-independent data structures. Fixed effects model the across-subject trends. They are assumed to be constant across all subjects. Random effects model the within-subject variability. They are unique to each individual subject. Mixed models are widely used for analyzing repeated measurement data. Repeated measurements taken from the same subject are likely to be more similar to each other than measurements taken from different subjects.

Restricted maximum likelihood (REML) is widely used for fitting mixed models because it provides unbiased estimates of the variance components by accounting for the degrees

of freedom lost by estimating fixed effects. The best linear unbiased prediction (BLUP) is used for the estimation of random effects. It requires the estimation of variance component obtained using REML.

In the Bayesian formulation of mixed model[30], random effects are given prior distributions to form joint posterior.

1.4 Works in this dissertation

In this dissertation, we have developed a suite of novel variable selection methods for longitudinal and irregular high-dimensional data. In literature, although variable selectin methods have been extensively proposed to accommodate a myriad of structured sparsity among omics features, such as strong correlations[7;8], networks[9;31], pathways and gene sets[32–34], only a limited number of studies have been conducted under disease phenotypes with repeated measures and heavy-tailed distributions. To fill the gap, in chapter 2, we have developed a penalized GEE method with network constraints to account for strong correlations among the SNP data in a longitudinal study, which significantly differs and advances from existing network based regularization methods under continuous, categorical and survival responses[10;35;36]. In chapter 3, we have developed the spike-and-slab quantile LASSO for robust variable selection in the presence of irregular disease traits due to cancer heterogeneity. The proposed method integrates the merits of the frequentist quantile LASSO and its Bayesian counterpart (Bayesian quantile LASSO) while overcoming their respective limitations. Estimation of the model parameters have been conducted based on the EM algorithm. One interesting finding from Chapter 3 is that the posterior mode of the high-dimensional sparse coefficient vector corresponding to omics features can be expressed as a closed-form solution incorporating the soft-thresholding rule within the coordinate descent framework, which has rarely been reported in robust penalized regression using non-differentiable robust loss functions. In chapter 4, we have extended the method developed in Chapter 3 to capture sparse nonlinear time-varying interactions with SNPs in longitudinal studies. In Genome-wide association studies (GWAS), identification of non-linear gene-environment interaction studies can

be traced back to[37] where a statistical hypothesis testing framework incorporating the varying coefficient model has been established. For regression based methods, we refer readers to[38] for the general review on variable selection methods for gene-environment interaction studies. In particular, identification of non-linear gene-environment interactions demands sparse varying coefficient models[39–43]. In chapter 4, we have proposed the Spike-and-Slab quantile group LASSO for non-parametric varying coefficient mixed model in longitudinal GWAS to account for time-varying effects with SNPs under heterogeneous repeated measure phenotypes. Numerical analysis suggests the advantage of the proposed method over both the frequentis and Bayesian alternatives.

Chapter 2

Penalized Variable Selection for Network-Based Data in the Longitudinal Study

2.1 Introduction

Longitudinal data are commonly encountered in different scientific studies. Areas like economics, biomedical studies and clinical trials having been using repeated measurement to answer research questions and solve practical problems. Longitudinal data are collected by measuring the same subject repeatedly over time. This leads to the correlated nature of the data. Over the years, different modeling approaches have been proposed to account for the correlation within the repeated measurement data. Different variable selection methods have been developed including the mixed model (Fan and Li[23]), penalized generalized estimating equation (Wang et al.[44]), penalized quadratic inference function (Cho and Qu[45]) and gene-environment interaction variable selection method (Zhou et al.[28]). In genetic association studies where the data is high-dimensional, longitudinal design is acquiring more attention. The main goal of association analysis is to recognize the important signals associated with phenotypes of certain disease from a massive number of genetic variants.

Despite the success in the accommodating of intra-correlation among repeated measurements, there is limited progress in dealing with structured sparsity. Complicated diseases are related to joint effects of multiple genetic factors which are hard to capture by marginal analysis. For example, in cancer related studies, omics features such as gene expressions often contain strong correlations. Overlooking these correlations during the process of variable selection results in identifying the wrong signals. To incorporate the network structure among genetic variants, the network-constrained regularization has been proposed. Some network-based methods are proposed in dealing with different types of responses including continuous (Li and Li[31]; Peng et al.[46]; Shi et al.[47]), binary (Ren et al.[48]; Sun and Wang[49]; Min et al.[50]), multi-nomial (Tian et al.[51]) and survival (Ren et al.[10]). However, it has not been adequately explored in longitudinal studies.

This study was partially motivated by the analysis of high-dimensional SNP data combining with the longitudinal features and responses from Childhood Asthma Management Program (CAMP) data. CAMP starts in early 1990s to study the long-term impacts of anti-inflammatory treatment Budesonide and Nedocromil on children with mild to moderate asthma. It has become the largest randomized longitudinal clinical trial. There are three levels of treatment including placebo. The disease phenotype of main focus is the forced expiratory volume in one second (FEV1). It is a longitudinal indicator response that records whether there being an improvement in children's lung growth. With SNPs as high-dimensional features to be selected and FEV1 as the longitudinal response, we are interested in conducting association analysis using a network based variable selection method.

2.2 Statistical Methods

2.2.1 Data and Model Settings

Consider a longitudinal study with n subjects and m_i measurements repeatedly taken across time on the ith subject ($1 \leqslant i \leqslant n$). Without loss of generality, set $m_i = m$. Let Y_{ij} be the response of the ith subject at the jth time point ($1 \leqslant j \leqslant m$). $Z_{ij} = (z_{ij1}, \ldots, z_{ijp})^\top$ denotes

a p-dimensional vector of genetic factors and $C_{ij} = (1, c_{ij1} \ldots, c_{ijq})^\top$ is a $(q+1)$-dimensional vector of which the last q components are clinical/environmental factors. Consider the following model:

$$Y_{ij} = C_{ij}^\top \alpha + Z_{ij}^\top \theta + \epsilon_{ij} \tag{2.1}$$

where $\alpha = (\alpha_0, \alpha_1 \ldots, \alpha_q)^\top$ is the regression coefficients for the intercept and clinical covariates, and $\theta = (\theta_1 \ldots, \theta_p)^\top$ is the regression coefficient vector for the gene expressions. For convenience, random error $\epsilon_i = (\epsilon_{i1}, ..., \epsilon_{im})^\top$ is assumed to follow a multivariate normal distribution $N_m(0, \Sigma_i)$ with Σ_i as its correlation matrix for the m repeated measurements of the ith subject. The model can also be expressed as:

$$Y_{ij} = X_{ij}^\top \beta + \epsilon_{ij}$$

where $X_{ij} = (C_{ij}^\top, Z_{ij}^\top)^\top$ and $\beta = (\alpha^\top, \theta^\top)^\top$. The observations are assumed to be dependent within the same subject, and independent if they are from different subjects.

2.2.2 Generalized Estimating Equations

For longitudinally clustered response Y_{ij}, it is generally difficult to specify its joint likelihood function. The generalized estimating equation (GEE) was used to account for the intra-cluster correlation by Liang and Zeger[52]. It estimates the average response over the population with the specifying of a working correlation. The estimating function for β is defined as:

$$U(\beta) = \sum_{i=1}^{n} D_i^\top V_i^\top (Y_i - \mu_i(\beta)) \tag{2.2}$$

where $D_i^\top = \frac{\partial \mu_i(\beta)}{\partial \beta}$ in which $\mu_i(\beta) = (\mu_{i1}(\beta), ..., \mu_{im}(\beta))^\top$. $V_i = A_i^{\frac{1}{2}} R(\eta) A_i^{\frac{1}{2}}$ is the covariance matrix of the ith subject, with $A_i = \text{diag}\{\text{Var}(Y_{i1}), ..., \text{Var}(Y_{im})\}$ which is the diagonal marginal covariance matrix for the ith cluster. $R(\eta)$ is a 'working' correlation matrix that describes the pattern of measurements. η is a finite dimensional intra–subject/cluster parameter which fully characterizes $R(\eta)$. The solution to the score equation $U(\beta) = 0$, is the

GEE estimator.

2.2.3 Penalized Identification

Penalized GEE based methods were proposed by Wang et al.[44] and Ma et al.[53] for conducting variable selection under correlated longitudinal responses. Penalized generalized estimating (PGEE) functions are defined as

$$Q_n(\beta_n) = U(\beta_n) - q'_{\lambda_n}(|\beta_n|)sign(\beta_n)$$

$q'_{\lambda_n}(|\beta_n|) = (q'_{\lambda_n}(|\beta_{n1}|), \ldots, q'_{\lambda_n}(|\beta_{np_n}|))^\top$ is a p_n-dimensional vector of penalty functions.

In this article, we consider the network-constrained penalty which is given by

$$\sum_{j=1}^{p} \rho(|\beta_j|; \lambda_1, \gamma) + \frac{1}{2}\lambda_2 \beta^\top L \beta, \tag{2.3}$$

where ρ is a penalty function with tuning parameters λ_1 and γ. The minimum concave penalty (MCP) is chosen to be the penalty, since it achieves a nearly unbiased selection. The two penalty terms in equation (2.3) controls sparsity for variable selection and smoothness of the estimated coefficients of correlated predictors. The first penalty term uses MCP which is defined as

$$\rho(t; \lambda_1, \gamma) = \lambda_1 \int_0^{|t|} (1 - \frac{x}{\gamma\lambda_1})_+ dx.$$

Its derivative is

$$\rho'(t; \lambda_1, \gamma) = \lambda_1 (1 - \frac{|t|}{\gamma\lambda_1})_+ sgn(t).$$

The second penalty term is expressed using a positive semi-definite matrix L. It can be written as

$$\beta^\top L \beta = \sum_{1 \le j < k \le p} |a_{jk}|(\beta_j - s_{jk}\beta_k)^2 \quad \forall \beta \in \mathbb{R}^p.$$

The adjacency measure $|a_{jk}|$ plays a crucial role in quantifying the strength of connection between two predictors. Let $A = (a_{jk},\ 1 \le j, k \le p)$ and $D = diag(d_1, \ldots, d_p)$, where

$d_j = \sum_{k=1}^{p} |a_{jk}|$. We have $L = D - A$ and $\sum_{1 \le j < k \le p} |a_{jk}|(\beta_j - s_{jk}\beta_k)^2 = \beta^\top (D - A)\beta$. Denote r_{jk} as the Pearson's correlation coefficient between x_j and x_k. The adjacency coefficient a_{jk} can be expressed as $a_{jk} = r_{jk}^{\kappa} \cdot I\{|r_{jk}| > \phi\}$ with $\kappa > 0$ being a tuning parameter. The default value for κ is 5. The power transformation and the value of α are chosen based on existing study from Zhang and Horvath[54]. The cutoff value r is determined based on the Fisher transformation $f_{jk} = 0.5 \log((1+r_{jk})/(1-r_{jk}))$. When the correlation between x_j and x_k is zero, $\sqrt{n-3} f_{jk}$ approximately follows $N(0,1)$. It can be used to determine a threshold c for $\sqrt{n-3} f_{jk}$. As a result, we define the threshold for r_{jk} as $\phi = (\exp(2c/\sqrt{n-3}) - 1)/(\exp(2c/\sqrt{n-3}) + 1)$. Adjacency matrix can also be defined using other dissimilarity measures.Different applications may require different adjacency matrices.

Assuming the n PGEE functions are the same, consider the network-constrained regularization estimator for PGEE:

$$q'_\lambda(|\beta_j|) = \begin{cases} \lambda_1 - \frac{|\beta_j|}{\gamma} + \lambda_2 \sum_{k=j+1}^{p} |a_{jk}|(\beta_j - s_{jk}\beta_k) & |\beta_j| \le \gamma\lambda_1 \\ \lambda_2 \sum_{k=j+1}^{p} |a_{jk}|(\beta_j - s_{jk}\beta_k) & |\beta_j| > \gamma\lambda_1 \end{cases}$$

where λ_1, λ_2 and γ are the tuning parameters, $|a_{jk}|$ measures the strength of correlation between x_j and x_k, and $s_{jk} = sgn(a_{jk})$ which is the sign of a_{jk}.

2.2.4 Computational Algorithms

A Newton-Raphson type of algorithm was developed to obtain the penalized estimate $\hat{\beta}$. The iterative algorithm for solving PGEE combines the minorization-maximization algorithm for nonconvex penalty of Hunter and Li[55] with the Newton-Raphson algorithm for the GEE.

For a small $\epsilon > 0$, the minorization-maximization algorithm suggests that the PGEE estimator $\hat{\beta} = (\hat{\beta}_1, \ldots, \hat{\beta}_p)^T$ approximately satisfies

$$U_j(\hat{\beta}) - n q'_\lambda(|\hat{\beta}_j|) sign(\hat{\beta}_j) \frac{|\hat{\beta}_j|}{\epsilon + |\hat{\beta}_j|} = 0, \quad j = 1, \ldots, p, \tag{2.4}$$

where $U_j(\hat{\beta})$ denotes the jth element of $U(\hat{\beta})$. ϵ is taken to be a fixed small number 10^{-6}.

By applying Newton-Raphson algorithm to $U_j(\hat{\beta})$, the iterative algorithm for solving the estimated $\hat{\beta}^d$ in the dth iteration can be obtained by:

$$\hat{\beta}^{(d)} = \hat{\beta}^{(d-1)} + [H(\hat{\beta}^{(d-1)}) + nE(\hat{\beta}^{(d-1)})]^{-1} \times [U(\hat{\beta}^{(d-1)}) - nE(\hat{\beta}^{(d-1)})\hat{\beta}^{(d-1)}], \tag{2.5}$$

where $U(\hat{\beta}^{(d-1)})$ is the score function at the $(d-1)$th iteration and $H(\hat{\beta}^{(d-1)})$ is the corresponding first derivative function of $U(\hat{\beta}^{(d-1)})$. $E(\hat{\beta}^{d-1})$ is a diagonal matrix that contains the first derivative of the network-constrained penalty. They are defined as:

$$\begin{aligned} H(\hat{\beta}^{(d-1)}) &= \sum_{i=1}^{n} D_i^\top V_i^{-1} D_i, \\ E(\hat{\beta}^{(d-1)}) &= diag\{\underbrace{0, ..., 0}_{1+q}, \frac{q'_\lambda(|\hat{\beta}_1^{(d-1)}|+)}{\epsilon + |\hat{\beta}_1^{(d-1)}|}, \ldots, \frac{q'_\lambda(|\hat{\beta}_p^{(d-1)}|+)}{\epsilon + |\hat{\beta}_p^{(d-1)}|}\}. \end{aligned} \tag{2.6}$$

The first $(1+q)$ elements on the diagonal of $E(\hat{\beta}^{d-1})$ are zero, suggesting that there is no shrinkage imposed on the coefficients for the intercept and the environmental factors.

There are two tuning parameters λ_1 and λ_2 along with a regularization parameter γ in the algorithm. Optimal tuning parameters λ_1 and λ_2 were chosen using k-fold cross-validation. Regularization parameter γ is set to 5 based on relevant studies of a sequence of possible values.

For fixed tuning parameters, the algorithm was implemented as follows:

(a) Obtaining the initial coefficient vector $\hat{\beta}^{(0)}$ using LASSO;

(b) At the (d)th iteration, compute $H(\hat{\beta}^{(d-1)})$, $E(\hat{\beta}^{(d-1)})$ and $U(\hat{\beta}^{(d-1)})$;

(c) Update $\hat{\beta}^{(d)}$ based on equation (A.2) ;

(d) Repeat Step (b) until the convergence is achieved.

2.2.5 Unbalanced Data Implementation

In practice, due to missing data, the repeated measurements are unbalanced when cluster sizes vary among different subjects. The proposed method can still be implemented in such a case by introducing a transformation matrix to each subject. Suppose the total number of time points is denoted by k and the ith subject is repeated measured at k_i time points. Let S_i denote a $k \times k_i$ tranformation matrix for the ith subject. Then for the ith subject, the transformation matrix S_i is generated by deleting the columns of the $k \times k$ identity matrix that correspond to the time points with measurement missing. According to this strategy, transformation is performed by letting $D_i^\star = S_i D_i, Y_i^\star = S_i Y_i, \mu_i^\star(\beta) = S_i \mu_i(\beta), A_i^\star = S_i A_i S_i^\top$ and $V_i^\star = (A_i^\star)^{\frac{1}{2}} R(\alpha)(A_i^\star)^{\frac{1}{2}}$. Then we can replace $U_i(\beta)$ in equation (2.2) by the transformed extended score vector $U_i^\star(\beta)$, which is defined as:

$$U_i^\star(\beta) = (D_i^\star)^\top (V_i^\star)^\top (Y_i^\star - \mu_i^\star(\beta)),$$

and the GEE estimator can be further obtained for unbalanced data based on the transformed terms.

2.3 Simulation

To demonstrate the utility of the proposed method, the performance of 5 methods was evaluated through simulation study. We generate the responses from model (2.1) with sample size $n = 200$, and set the number of time points m to 4. The dimension for genetic factors is $p = 750$ and there are $q = 5$ environmental factors. In our simulation, the environmental factors are simulated from a multivariate normal distribution with mean 0 and AR–1 covariance matrix with marginal variance 1 and auto correlation coefficient 0.8. We simulate the random error ϵ for the longitudinal response from a multivariate normal distribution by assuming 0 mean vector and an exchangeable covariance structure with parameter $\tau = 0.8$. Following all these settings, the time-independent genetic factors are simulated in four different scenarios.

In the first scenario, the genetic factors are gene expressions simulated from a multivariate

normal distribution. Two correlation structures are considered. First is the auto-regression (AR) structure, in which gene i and j have correlation coefficient $\rho^{|i-j|}$ when they are in the same cluster. Second correlation structure is the banded structure where gene i and j have correlation coefficient ρ when $|i-j| = 1$ and 0 correlation otherwise. Two different correlation coefficient $\rho = 0.5$ and 0.9 indicating moderate and strong correlation are considered for both structures.

In the second scenario, we consider generating SNP data by dichotomizing the gene expression values from scenario 1 at the 30th and 70th percentiles with respect to each gene, leading to the three categories (0,1,2) for genotypes (aa,Aa,AA).

In the third scenario, we simulate the SNP data using a pairwise linkage disequilibrium (LD) structure. Let q_A and q_B denote the minor allele frequencies (MAFs) for the two risk alleles A and B from two adjacent SNPs, respectively, and δ denote the LD. Then the frequencies of the four haplotypes can be derived as $p_{AB} = q_A q_B + \delta, p_{ab} = (1-q_A)(1-q_B) + \delta, p_{Ab} = q_A(1-q_B) - \delta$, and $p_{aB} = (1-q_A)q_B - \delta$. By assuming Hardy-Weinberg equilibrium, we simulate the SNP genotypes AA, Aa and aa at locus 1 from a multinomial distribution with frequencies $q_A^2, 2q_A(1-q_A)$ and $(1-q_A)^2$. Then the genotypes for SNP at locus 2 can be generated based on the conditional genotype probability matrix[56]. If the MAFs are 0.3 and pairwise correlation r is set to 0.3, we can get $\delta = r\sqrt{q_A(1-q_A)q_B(1-q_B)}$.

Lastly, in scenario four, we adopted a more realistic method to generate the SNP data. 750 consecutive SNPs were extracted from the case study. 200 subjects were randomly selected from the real data to generate repeated measurement responses of 4 time points.

The coefficients were simulated from uniform distribution $U[0.15, 0.25]$ and $U[0.25, 0.5]$ indicating weak and strong signals. The coefficients also have a cluster structure consisting of 150 clusters, each containing 5 elements. Out of the 150 clusters, 6 clusters are designated to have non-zero values.

For comparison with the network approach, we consider four alternative approaches, LASSO, elastic net, SCAD and MCP. LASSO is broadly used in the analysis of genomic data. In contrast to LASSO, elastic net accounts for the correlation among genomic features. SCAD and MCP reduce the weight on λ_1 when the coefficient is large which makes these

methods less biased. The proposed approach is equivalent to MCP when $\lambda_2 = 0$ in (2.3). For convenience, LASSO, elastic net, SCAD, MCP and network approach are termed as A1, A2, A3, A4 and A5, respectively. Simulation results of the five methods (A1-A5) are displayed in Table 2.1-2.6.

Table 2.1: Weak signal coefficients under AR–1 correlation. Mean (sd) based on 100 replicates.

n=200	p=750	A1	A2	A3	A4	A5
			Gene Expression			
Independence	$\rho = 0.5$	0.73(0.07)	0.72(0.08)	0.78(0.07)	0.78(0.06)	0.78(0.06)
	$\rho = 0.9$	0.52(0.06)	0.60(0.06)	0.54(0.06)	0.56(0.06)	0.67(0.07)
AR-1	$\rho = 0.5$	0.77(0.06)	0.77(0.06)	0.81(0.05)	0.81(0.06)	0.79(0.06)
	$\rho = 0.9$	0.58(0.06)	0.66(0.06)	0.60(0.06)	0.62(0.06)	0.76(0.05)
Exchangeable	$\rho = 0.5$	0.74(0.07)	0.75(0.07)	0.78(0.06)	0.79(0.06)	0.76(0.06)
	$\rho = 0.9$	0.56(0.06)	0.67(0.07)	0.57(0.07)	0.58(0.07)	0.76 (0.06)
			SNP			
Independence	$\rho = 0.5$	0.70(0.07)	0.70(0.06)	0.72(0.07)	0.72(0.07)	0.72(0.06)
	$\rho = 0.9$	0.62(0.07)	0.65(0.07)	0.60(0.07)	0.62(0.07)	0.73(0.05)
AR-1	$\rho = 0.5$	0.75(0.06)	0.73(0.06)	0.78(0.05)	0.79(0.06)	0.76(0.05)
	$\rho = 0.9$	0.64(0.07)	0.74(0.07)	0.69(0.07)	0.69(0.07)	0.81(0.05)
Exchangeable	$\rho = 0.5$	0.76(0.07)	0.73(0.07)	0.75(0.09)	0.77(0.07)	0.75(0.07)
	$\rho = 0.9$	0.61(0.06)	0.70(0.06)	0.66(0.05)	0.62(0.08)	0.81(0.08)

Table 2.2: Strong signal coefficients under AR–1 correlation. Mean (sd) based on 100 replicates.

n=200	p=750	A1	A2	A3	A4	A5
			Gene Expression			
Independence	$\rho = 0.5$	0.87(0.06)	0.92(0.04)	0.90(0.04)	0.90(0.04)	0.89(0.05)
	$\rho = 0.9$	0.47(0.07)	0.76(0.06)	0.64(0.07)	0.64(0.06)	0.90(0.05)
AR-1	$\rho = 0.5$	0.87(0.06)	0.92(0.04)	0.90(0.04)	0.90(0.04)	0.89(0.05)
	$\rho = 0.9$	0.47(0.07)	0.76(0.06)	0.64(0.07)	0.64(0.06)	0.90(0.05)
Exchangeable	$\rho = 0.5$	0.87(0.06)	0.92(0.04)	0.90(0.04)	0.90(0.04)	0.89(0.05)
	$\rho = 0.9$	0.47(0.07)	0.76(0.06)	0.64(0.07)	0.64(0.06)	0.90(0.05)
			SNP			
Independence	$\rho = 0.5$	0.87(0.06)	0.92(0.04)	0.90(0.04)	0.90(0.04)	0.89(0.05)
	$\rho = 0.9$	0.54(0.06)	0.69(0.07)	0.66(0.06)	0.64(0.07)	0.88(0.05)
AR-1	$\rho = 0.5$	0.86(0.06)	0.92(0.04)	0.89(0.04)	0.89(0.04)	0.88(0.05)
	$\rho = 0.9$	0.55(0.05)	0.71(0.06)	0.66(0.09)	0.67(0.06)	0.88(0.04)
Exchangeable	$\rho = 0.5$	0.88(0.06)	0.91(0.05)	0.90(0.05)	0.89(0.05)	0.89(0.04)
	$\rho = 0.9$	0.50(0.09)	0.71(0.07)	0.63(0.12)	0.62(0.11)	0.85(0.13)

Table 2.3: Weak signal coefficients under banded correlation. Mean (sd) based on 100 replicates.

n=200	p=750	A1	A2	A3	A4	A5
Gene Expression						
Independence	$\rho = 0.5$	0.72(0.06)	0.73(0.06)	0.73(0.06)	0.74(0.06)	0.75(0.05)
	$\rho = 0.9$	0.46(0.06)	0.49(0.05)	0.46(0.05)	0.47(0.05)	0.56(0.04)
AR-1	$\rho = 0.5$	0.74(0.05)	0.76(0.06)	0.74(0.05)	0.76(0.06)	0.75(0.04)
	$\rho = 0.9$	0.49(0.06)	0.53(0.06)	0.50(0.05)	0.50(0.05)	0.60(0.03)
Exchangeable	$\rho = 0.5$	0.74(0.10)	0.75(0.08)	0.76(0.08)	0.76(0.07)	0.78(0.06)
	$\rho = 0.9$	0.50(0.05)	0.54(0.05)	0.50(0.04)	0.50(0.05)	0.64(0.09)
SNP						
Independence	$\rho = 0.5$	0.72(0.06)	0.72(0.06)	0.72(0.06)	0.73(0.05)	0.72(0.04)
	$\rho = 0.9$	0.53(0.06)	0.57(0.07)	0.53(0.07)	0.52(0.06)	0.65(0.05)
AR-1	$\rho = 0.5$	0.68(0.08)	0.66(0.08)	0.71(0.08)	0.73(0.09)	0.71(0.08)
	$\rho = 0.9$	0.60(0.06)	0.73(0.07)	0.64(0.06)	0.66(0.06)	0.79(0.05)
Exchangeable	$\rho = 0.5$	0.72(0.07)	0.69(0.08)	0.74(0.06)	0.74(0.07)	0.72(0.09)
	$\rho = 0.9$	0.52(0.07)	0.65(0.06)	0.60(0.07)	0.57(0.08)	0.77(0.05)

Table 2.4: Strong signal coefficients under banded correlation. Mean (sd) based on 100 replicates.

n=200	p=750	A1	A2	A3	A4	A5
Gene Expression						
Independence	$\rho = 0.5$	0.74(0.07)	0.88(0.05)	0.87(0.05)	0.88(0.06)	0.86(0.06)
	$\rho = 0.9$	0.46(0.04)	0.68(0.04)	0.58(0.20)	0.65(0.25)	0.71(0.03)
AR-1	$\rho = 0.5$	0.75(0.07)	0.91(0.04)	0.88(0.05)	0.88(0.04)	0.87(0.06)
	$\rho = 0.9$	0.45(0.04)	0.63(0.04)	0.49(0.12)	0.59(0.19)	0.71(0.04)
Exchangeable	$\rho = 0.5$	0.80(0.07)	0.92(0.04)	0.89(0.04)	0.88(0.04)	0.86(0.04)
	$\rho = 0.9$	0.47(0.03)	0.62(0.03)	0.64(0.20)	0.62(0.20)	0.70(0.04)
SNP						
Independence	$\rho = 0.5$	0.79(0.08)	0.82(0.07)	0.85(0.06)	0.84(0.06)	0.81(0.06)
	$\rho = 0.9$	0.60(0.05)	0.79(0.06)	0.68(0.08)	0.66(0.07)	0.86(0.06)
AR-1	$\rho = 0.5$	0.80(0.08)	0.88(0.04)	0.87(0.05)	0.87(0.04)	0.84(0.04)
	$\rho = 0.9$	0.53(0.06)	0.71(0.06)	0.63(0.06)	0.62(0.06)	0.84(0.06)
Exchangeable	$\rho = 0.5$	0.67(0.12)	0.82(0.07)	0.81(0.08)	0.78(0.07)	0.81(0.07)
	$\rho = 0.9$	0.46(0.07)	0.79(0.07)	0.63(0.13)	0.61(0.08)	0.86(0.06)

Table 2.5: LD structure under weak and strong signal coefficients. Mean (sd) based on 100 replicates.

$r = 0.3$	A1	A2	A3	A4	A5
weak signal					
Independence	0.77(0.06)	0.86(0.05)	0.79(0.07)	0.79(0.07)	0.97(0.02)
AR-1	0.82(0.06)	0.89(0.05)	0.84(0.06)	0.83(0.06)	0.99(0.01)
Exchangeable	0.80(0.05)	0.89(0.05)	0.83(0.06)	0.82(0.05)	0.99(0.01)
strong signal					
Independence	0.87(0.04)	0.91(0.03)	0.87(0.04)	0.86(0.04)	0.98(0.02)
AR-1	0.82(0.06)	0.87(0.05)	0.82(0.06)	0.81(0.07)	0.97(0.02)
Exchangeable	0.79(0.05)	0.84(0.04)	0.79(0.06)	0.78 (0.06)	0.96(0.03)

Table 2.6: Real data simulation. Mean (sd) based on 100 replicates.

	A1	A2	A3	A4	A5
weak signal					
Independence	0.77(0.04)	0.82(0.03)	0.80(0.06)	0.81(0.03)	0.89(0.07)
AR-1	0.79(0.06)	0.84(0.05)	0.82(0.08)	0.83(0.06)	0.90(0.05)
Exchangeable	0.80(0.04)	0.86(0.06)	0.83(0.05)	0.83(0.05)	0.93(0.05)
strong signal					
Independence	0.85(0.06)	0.89(0.05)	0.86(0.08)	0.86(0.08)	0.93(0.06)
AR-1	0.79(0.05)	0.87(0.03)	0.82(0.06)	0.82(0.08)	0.96(0.05)
Exchangeable	0.75(0.08)	0.86(0.06)	0.78(0.06)	0.79(0.06)	0.96(0.03)

To evaluate the performance, we computed mean and standard deviation of AUC under different settings. Under medium correlation ($\rho = 0.5$), all 5 methods have similar performance. When $\rho = 0.9$ which indicates strong correlation, network-based penalization method outperforms other methods in distinguishing between zero coefficients and nonzero coefficients. For example in Table 2.1 with gene expression data under exchangeable correlation, when $\rho = 0.5$, the network approach (0.76 (sd 0.06)) has similar performance as LASSO (0.74 (sd 0.07)), elastic net (0.75 (sd 0.07)), SCAD (0.78 (sd 0.06)) and MCP (0.79 (sd 0.06)). When $\rho = 0.9$, the network approach (0.76 (sd 0.06)) has better performance than LASSO (0.56 (sd 0.06)), elastic net (0.67 (sd 0.07)), SCAD (0.57 (sd 0.07)) and MCP (0.58 (sd 0.07)). Elastic net has the second best performance because it has the ability of accommodate some of the correlation. However, in the case of network-structured correlation, the proposed method is capable of fitting the correlation with better accuracy.

In repeated measurement studies, the GEE framework allows the misspecification of working correlations and gives robust results. In our simulation, although correctly specifying working correlation ended up with better outcomes, overall, the results are comparable under different settings. This property is helpful when working on real data sets with unknown working correlations.

2.4 Real Data Analysis

Asthma is a chronic respiratory disease characterized by chronic inflammation of the lungs and reversible obstruction of the airways, leading to breathing difficulties. According to the Centers for Disease Control and Prevention (CDC), over 25 million individuals in the United States suffer from asthma. Asthma affects 7.7% of adults and 8.4% of children in the United States[57;58]. It is the primary chronic disease affecting children. In our case study, data was obtained from the Childhood Asthma Management Program (CAMP)[59–61]. We downloaded and pre-processed the SNP and phenotype datasets from CAMP for our analysis. Subjects between the ages of 5 and 12 years who had been diagnosed with chronic asthma were selected and monitored over 4 years. Prior to treatment, there are three visits scheduled with a time

gap of 1-2 weeks between each visit. Another thirteen visits were made after treatment. The first two visits after treatment were 2 months apart from each other, and the remaining visits were 4 months apart. In order to reduce the impact of variations in the time gap, we chose the final twelve visits that were scheduled 4 months apart after treatment for our study. The phenotype of interest is the forced expiratory volume in one second (FEV1), which represents the total volume of air expelled out of the lung in one second. This measurement was taken multiple times during each visit. The subjects were given two different types of treatments: Budesonide and Nedocromil, with 30% of the subjects being assigned to each of these treatments, while the remaining subjects were given placebo. Treatment, age and gender were considered as clinical factors. Each subject's genotype information comprises more than nine hundred thousand SNPs. We paired the genotypes with the corresponding phenotypes based on the subject IDs and removed SNPs that had a minor allele frequency (MAF) of less than 0.05 or displayed deviation from the Hardy-Weinberg equilibrium. The resulting dataset consisted of 438 subjects and 447,850 SNPs.

For computational convenience in studies involving ultra-high dimensional datasets, such as Genome Wide Association Studies (GWAS) and multi-omics integration studies, it is necessary to conduct marginal feature pre-screening initially to facilitate the application of regularization on datasets with a reasonably large scale[62;63]. For instance, Li et al.[64], Jiang et al.[65] and Wu et al.[66] utilized a single SNP analysis as a pre-screening method before implementing the proposed variable selection techniques in longitudinal and multivariate GWAS studies, respectively. We focused on SNPs in several important pathways potentially related to Asthma including Wnt signaling pathway, p53 signaling pathway, Cell cycle pathway and others. Chromosomes containing SNPs from these pathways were also examined. After cleaning the data by matching phenotypes and genotypes, we obtained working dataset containing 1000 SNPs, 12 time points and 438 subjects. All 5 methods were applied under the exchangeable working correlation. The optimal tuning parameters were selected using 5-fold cross-validation.

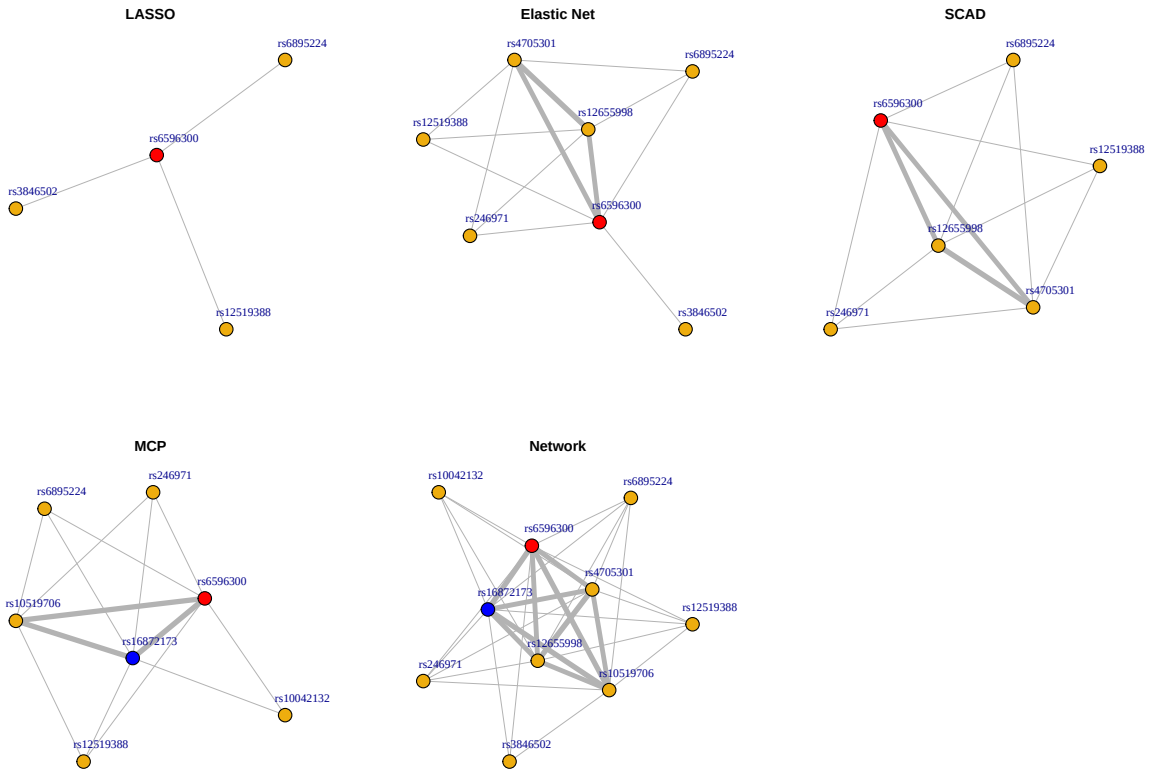

Figure 2.1: Sub-network for TRPC7 on chromosome 5.

On chromosome 5, methods (LASSO (A1), elastic net (A2), SCAD (A3), MCP (A4) and network-based (A5)) A1-5 identified 195, 180, 296, 312 and 391 SNPs associated with Asthma. As an illustrative example, we examined closely SNP rs6596300 and its sub-network. SNP rs6596300 is related to gene TRPC7, which influences calcium ion entry and stimulates the airway muscle cell to increase massively (Wang et al., 2011)[67]. Figure 2.1 shows the sub-network of TRPC7, where the red nodes indicates the SNP at the center. In a sub-network plot, the thickness of an edge represents the strength of the connection between SNPs. The effectiveness of the network-based approach (A5) can be clearly observed as it has identified a significantly higher number of highly correlated SNPs, since the incorporation of network structure information has allowed for a better representation of the interconnections among SNPs. In this sub-network, the network approach selects 10 SNPs and 8 of them belongs to Asthma related genes (TRPC7, HEXB, SH3TC2, C5orf38,

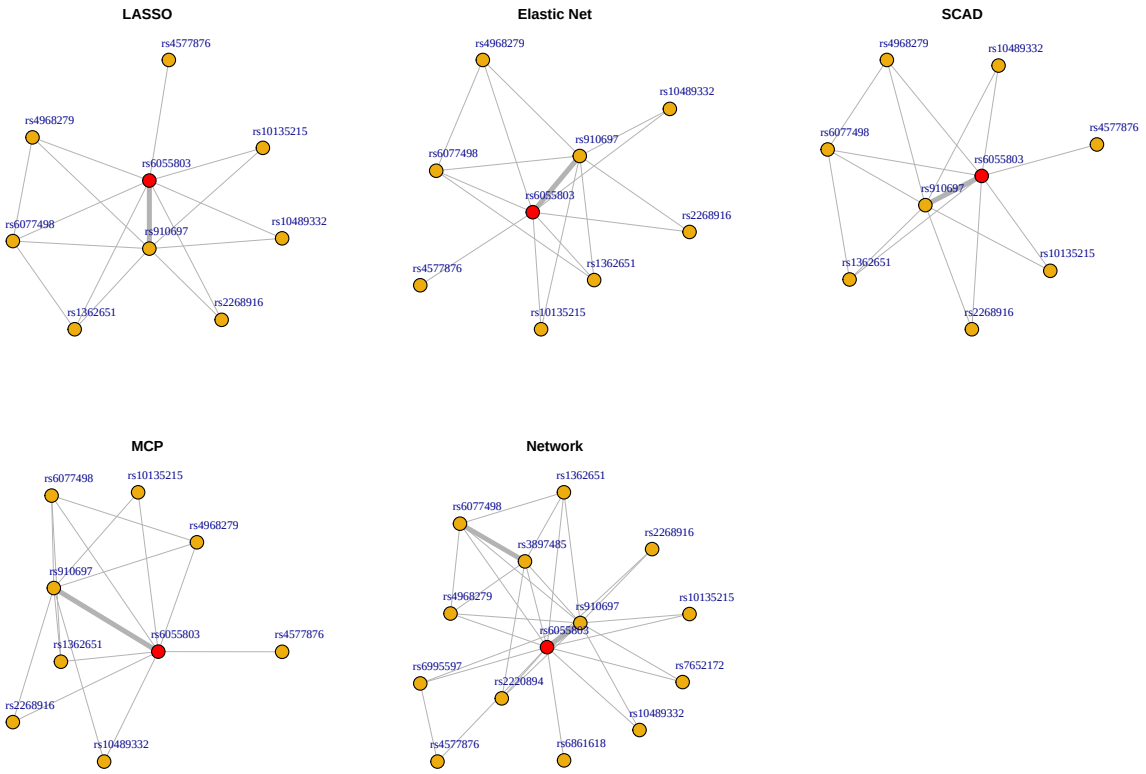

Figure 2.2: Sub-network for ANGPT4 on Wnt signaling pathway.

ZNF366, FBXL17, GABRB2 and SLC38A9), while other 4 approaches only identified 4, 7, 6 and 7 SNPs correspondingly. In addition, A5 and A4 both identified SNP related to gene HEXB, which was not selected by other methods. HEXB activity was significantly higher in patients with severe or moderate asthma[68]. With respect to the intensity of sub-network connections, A5 leads to a more tightly connected network, which is in line with our observations from the simulation study, where it was found to facilitate interconnections among SNPs.

On Wnt signaling pathway, methods A1-5 identified 330, 343, 251, 256 and 428 SNPs associated with Asthma. Figure 2.2 shows the sub-network centered around SNP rs6055803. It is related to gene ANGPT4, which affects the increasing in the vascularity of bronchial mucosa (Pałgan and Bartuzi, 2015)[69]. Although the connection strength is not as good as the previous scenario as the SNPs may not come from the same chromosome, A5 still

identified more strongly related SNPs than other methods.

2.5 Discussion

The network-constrained penalization method within the GEE framework proposed in this study has shown superior performance in performing variable selection on the simulated data in all scenarios. The network-basd penalty term in the regularized loss function accommodates the structured correlation widely occurred among SNPs, which ensures the advantage of the our method over the competing alternative methods.

Generally, penalized regularization methods has best performance when the dimensionality of the data is relatively larger than the sample size. To deal with ultra-high dimensional cases, we have to find a reasonable way to perform marginal screening before applying regularization to the data set. In this study, we implemented marginal linear regression method to conduct screening. Theoretical guaranteed methods like sure independent screening will be applied in the future work. Also, penalized QIF is shown to have a better performance than penalized GEE under various settings. We will explore the possibility of adopting penalized QIF into our study.

Chapter 3

The Spike-and-Slab Quantile Lasso

3.1 Introduction

Modern high-throughput technologies have generated large scale omics data, including gene expressions, copy number variations, singlenucleotide polymorphisms (SNPs), that can be analyzed in cancer research to dissect the genetic and genomic architecture under the complex disease phenotypes. In practice, the heavy-tailed distributions and outliers have been widely observed in those disease traits, reflecting the heterogeneity of cancer. We take the examples from the The Cancer Genomics Atlas (TCGA) to illustrate the irregularities in the cancer phenotypes (Cerami et al., 2012)[70]. In the TCGA lung adenocarcinomas (LUAD) data (Campbell et al., 2011)[71], the phenotype of interest is FEV1, the total volume of air expelled out of the lung in one second. Besides, in TCGA skin cutaneous melanoma (SKCM) study, we are interested in the log of Breslow's thickness (Marghoob et al., 2000)[72], an indicator of the depth that melanoma tumor has grown into skin, as the disease trait. The histograms of the two phenotypes in Figure 3.1 have clearly indicated that they do not follow normal distribution, and are possibly contaminated by the outliers. The deviation from normality can be further confirmed by the p-values from Shapiro-Wilk normality test, which are 0.003384 and 3.562×10^{-14} for FEV1 and log Breslow depth, respectively.

Identifying important disease-gene associations while building powerful predictive mod-

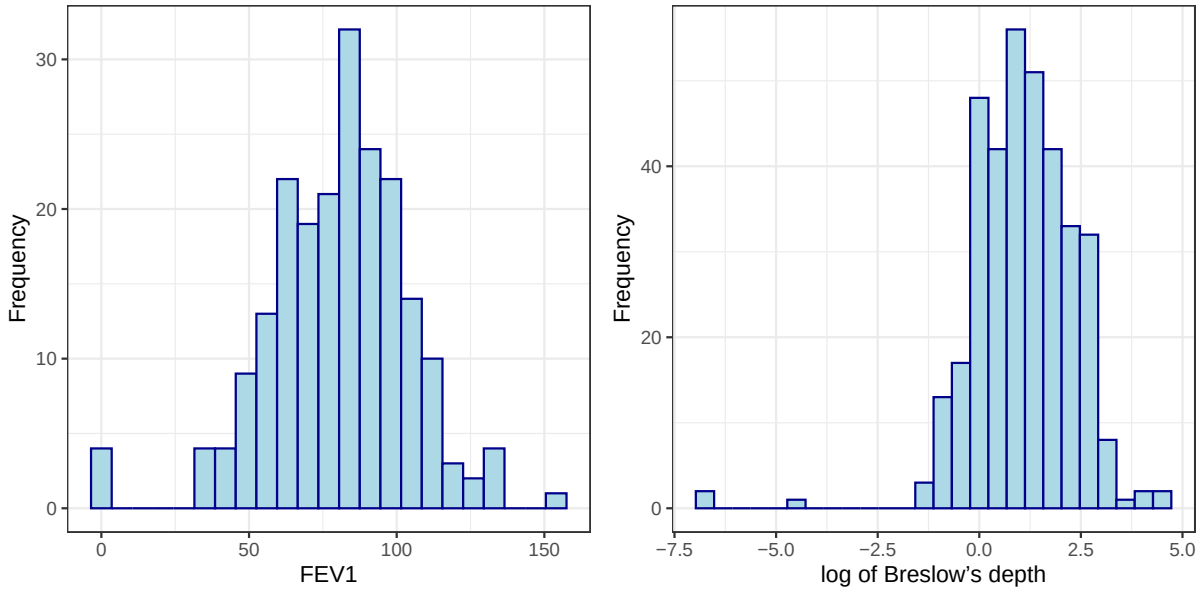

Figure 3.1: Histograms of FEV1 and log Breslow depth.

els under the disease phenotypes is the central task of analyzing high-throughput cancer genomics data. However, with non-robust methods, even a single outlying observation from the phenotypic traits will result in biased estimation and false detection of associated omics features[73]. Our exploratory data analysis has demonstrated that robust sparse regression is in pressing need for performing such a challenging task. In published studies, robust variable selection methods have been extensively developed in bioinformatics studies[73], for the sparse identification of gene expressions[10;74], single nucleotide polymorphisms[75;76 77], DNA copy number variations[78 35;79], and proteins[80;81].

Within the Bayesian framework, robust variable selection has recently been developed in addition to widely examined nonrobust counterpart[11;82–85]. For example, multiple robust Bayesian methods have been proposed to identify important gene–environment (G×E) interactions[22;86;87] and histopathological imaging–environment interactions[88]. In the Bayesian hierarchical model, incorporating the robustness against outliers and performing variable selection can be facilitated by adopting robust likelihood function and sparsity-inducing priors, respectively. For example, using the asymmetric Laplace distribution for the robust likelihood function[19], Li et al. (2010)[18] have established a Bayesian framework for regularization in linear quantile regression using the Laplace shrinkage priors. Other widely used sparse

priors include the spike-and-slab priors[16] and global local shrinkage priors such as the Normal Exponential Gamma prior[20] and Horseshoe prior[21;83]. Our biased literature search suggests that although there has been a surge of sparse Bayesian methods for genomics studies in the past decades[11;82], the robustness in Bayesian analysis of omics data has not received as much attention as in those frequentist studies.

As the frequentist and Bayesian methods have their own pros and cons, the Spike-and-Slab LASSO (ssLASSO) has been proposed to retain remarkable properties of both while overcoming their respective limitations (Ročková and George, 2018)[17]. Arising from a fully Bayes formulation with the spike-and-slab prior, ssLASSO leads to a posterior mode estimation algorithm via incorporating the Expectation-Maximization (EM) steps in the coordinate-descent framework[89], and is thus extremely fast. The spike-and-slab LASSO has been furthered extended to handle survival outcomes with Cox models[90] and generalized responses including multinomial and binary disease outcomes[91], demonstrating great promise in cancer genomics studies.

Despite the success, a major limitation of spike-and-slab LASSO based methods is that they are non-robust and vulnerable to the heterogeneity of cancer, usually in the form of heavy-tailed distributions and outliers in the disease traits. On the other hand, robust variable selection methods face the computational hurdles due to the non-differentiable objective functions (including quantile check loss) for the frequentist ones[73] and computationally intensive MCMC algorithms for the fully Bayesian analysis. These limitations have motivated us to develop spike-and-slab quantile LASSO (ssQLASSO), which has a fully Bayes spike-and-slab formulation with the robust likelihood based on asymmetric Laplace distribution (ALD). The following merits have distinguished ssQLASSO from published studies. First, given a robust likelihood in the hierarchical formulation, the ssQLASSO is robust compared to ssLASSO, its major competitor, and other non-probust variable selection methods. Second, surprisingly, the choice of ALD as the likelihood has equipped the proposed ssQLASSO with an unexpected computational advantage to find the posterior mode, which is revealed by the soft-thresholding rule guided EM steps in the coordinate descent framework, just like the ssLASSO. This is a rare phenomenon since the soft-thresholding rule is generally

not applicable in robust regularization with non-differentiable loss. Last but not least, the ssQLASSO has inherited the prominent properties of selective shrinkage and self-adaptivity to the sparsity pattern from ssLASSO, which has not been observed in any published robust variable selection methods. We have conducted extensive simulation studies under different model settings and a variety of heavy-tailed errors to demonstrate the utility of the proposed methods over alternatives. In the real data analysis, we have analyzed the TCGA LUAD and SKCM data under disease phenotypes with skewed distributions. The results show that the proposed method leads to biologically sensible findings with better predictive performance. The R package *emBayes* is available at https://cran.r-project.org/package=emBayes.

3.2 Statistical Methods

3.2.1 The Quantile Regression with Asymmetric Laplace Distribution

To handle skewness in phenotypic traits and outliers, we considered using quantile regression below. Let y_i be the response of the i-th subject ($1 \leq i \leq n$). We have $\boldsymbol{z}_i = (1, z_{i1}, \ldots, z_{iq})^\top$ being a $(q+1)$-dimensional vector of which the last q components indicate clinical factors. In addition, $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})^\top$ represents a p-dimensional vector of genetic factors. The linear quantile regression model for the τ-th ($0 < \tau < 1$) quantile level is

$$y_i = \boldsymbol{z}_i^\top \boldsymbol{\alpha}_{\boldsymbol{\tau}} + \boldsymbol{x}_i^\top \boldsymbol{\beta}_{\boldsymbol{\tau}} + \epsilon_{i,\tau}, \tag{3.1}$$

where $\boldsymbol{\alpha}_{\boldsymbol{\tau}} = (\alpha_{0\tau}, \cdots, \alpha_{q\tau})^\top$ consists of the regression coefficients corresponding to intercept and clinical covariates, and $\boldsymbol{\beta}_{\boldsymbol{\tau}} = (\beta_{1\tau}, \cdots, \beta_{p\tau})^\top$ are the regression coefficients for genetic factors or other types of omics features. The random error $\epsilon_{i\tau}$'s are independent with the τ-th quantile being 0. For simplicity of notation, the subscript "τ" has been omitted hereafter. The estimation of regression coefficients $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be achieved by minimizing the following

loss function

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{i=1}^{n} \rho_\tau(y_i - \boldsymbol{z}_i^\top \boldsymbol{\alpha} - \boldsymbol{x}_i^\top \boldsymbol{\beta}),$$

in which the check loss function is defined as $\rho_\tau(\epsilon_i) = \epsilon_i\{\tau - I(\epsilon_i < 0)\}$.

In literature, the asymmetric Laplace distribution (ALD) plays a pivotal role in probabilistic approaches to quantile regression, including the maximum likelihood estimation and Bayesian quantile regression (Yu and Moyeed, 2001)[19]. Specifically, with a fixed quantile level τ and scale parameter $\sigma(> 0)$ that controls the skewness of the distribution , the probability density function of ALD can be expressed as

$$f(y_i|\mu_i, \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp\left(-\rho_\tau\left(\frac{y_i - \mu_i}{\sigma}\right)\right),$$

where $\mu_i = \boldsymbol{z}_i^\top \boldsymbol{\alpha} + \boldsymbol{x}_i^\top \boldsymbol{\beta}$. Then the joint distribution of $\boldsymbol{Y} = (y_1, ..., y_n)^\top$ given $\boldsymbol{x} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_n)^\top$ and $\boldsymbol{z} = (\boldsymbol{z}_1, ..., \boldsymbol{z}_n)^\top$ can be written as

$$f(\boldsymbol{Y}|\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\beta}, \boldsymbol{\alpha}, \sigma) = \tau^n(1-\tau)^n\sigma^{-n} \exp\left(-\sum_{i=1}^{n} \rho_\tau\left(\frac{y_i - \mu_i}{\sigma}\right)\right).$$

Therefore, minimizing $L(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is equivalent to maximizing the (log-)likelihood function. The importance of asymmetric Laplace distribution to quantile regression has been further acknowledged since Kozumi and Kobayashi (2009)[92] have demonstrated that it is equivalent to the mixture of a scaled normal distribution and an exponential distribution, which significantly facilitates the formulation of the Bayesian hierarchical model. Define

$$\zeta_1 = \frac{1-2\tau}{\tau(1-\tau)} \text{ and } \zeta_2 = \sqrt{\frac{2}{\tau(1-\tau)}}.$$

Then the response variable has the form:

$$y_i = \mu_i + \zeta_1 v_i + \zeta_2\sqrt{\sigma v_i}u_i,$$

where v_i and u_i follow exponential distribution with a scale parameter $\sigma(> 0)$, Exp(σ), and

the standard normal distribution, $N(0,1)$, respectively. The hierarchical representation of y_i following $ALD(\mu_i, \sigma, \tau)$ has the form as follows:

$$\begin{aligned} y_i|v_i &\overset{ind}{\sim} \mathrm{N}(\mu_i + \zeta_1 v_i, \zeta_2^2 \sigma v_i), \\ v_i &\overset{ind}{\sim} \mathrm{Exp}(\sigma), \quad (i = 1, ..., n), \\ u_i &\overset{ind}{\sim} \mathrm{N}(0,1), \quad (i = 1, ..., n). \end{aligned}$$

3.2.2 The Spike-and-Slab Quantile LASSO

Identifying important genetic factors that are associated with the long-tailed disease traits demands robust variable selection[73]. In the frequentist framework, adding the L1 norm of regression coefficients as a regularizer to the quantile regression model (3.1) leads to the following quantile-LASSO model:

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \sum_{i=1}^{n} \rho_\tau(y_i - \boldsymbol{z}_i^\top \boldsymbol{\alpha} - \boldsymbol{x}_i^\top \boldsymbol{\beta}) + \lambda \sum_{j=1}^{p} |\beta_j|, \tag{3.2}$$

where $\lambda > 0$ is the tuning parameter determining the model sparsity. At the 50% quantile level, model (4.3) reduces to the least absolute deviation (LAD) LASSO[93]. On the other hand, with the aforementioned formulation of quantile regression using asymmetric Laplace distribution, Li et al.[18] have developed the Bayesian quantile LASSO by imposing the i.i.d conditional Laplace prior on regression coefficients corresponding to high-dimensional predictors[12;13]:

$$\beta_j|s \overset{ind}{\sim} \frac{1}{2s} \exp\left(-\frac{|\beta_j|}{s}\right), \tag{3.3}$$

where $s = \frac{\sigma}{\lambda}$. Li et al[18] have further demonstrated that the quantile LASSO estimator can be viewed as a posterior mode of the regression coefficients under such a prior specification.

We consider the following spike-and-slab prior[17;90;91]:

$$\beta_j|\gamma_j, s_0, s_1 \overset{ind}{\sim} (1 - \gamma_j)\pi(\beta_j|s_0) + \gamma_j \pi(\beta_j|s_1), \tag{3.4}$$

where the spike component $\pi(\beta_j|s_0)$ and slab component $\pi(\beta_j|s_1)$ are defined using the conditional Laplace prior in (3.3) with scale parameters $s_1 > s_0 > 0$. When the indicator $\gamma_j = 1$, the regression coefficient β_j is assumed to have coefficients with larger magnitude and be modeled by the slab part. Then the corresponding genetic factor is associated with the disease trait. Otherwise, γ_j being 0 indicates that β_j is small and close-to-0, following the spike distribution. Therefore, the corresponding genetic factor does not contribute to the variation of the phenotype. Alternatively, the spike-and-slab formulation can be written as $\beta_j|\gamma_j, S_j \overset{ind}{\sim} \pi(\beta_j|S_j) = \frac{1}{2S_j}\exp\left(-\frac{|\beta_j|}{S_j}\right)$, where $S_j = (1-\gamma_j)s_0 + \gamma_j s_1$.

In order to integrate the frequentist Bayesian quantile LASSO with the Bayesian quantile spike-and-slab regression for the proposed ssQLASSO, we first specify the following hierarchical representation. We assign the independent Bernoulli prior on γ_j, $(j = 1, ..., p)$:

$$\gamma_j|\theta \overset{ind}{\sim} \text{Ber}(\theta) = \theta^{\gamma_j}(1-\theta)^{1-\gamma_j},$$

with the parameter θ having a uniform prior $\theta \sim U(0,1)$. After θ being treated as unknown and random, it can be viewed as a global shrinkage parameter that equals $p(\gamma_j = 1|\theta)$, the prior probability of nonzero coefficients in $\boldsymbol{\beta}$. The prior expectation of S_j, $E(S_j) = (1-\theta)s_0 + \theta s_1$, suggests that the shrinkage is imposed adaptively on each coordinate of $\boldsymbol{\beta}$.

Besides, the regression coefficients for intercept and clinical factors, α_k, are assumed to have independent normal priors with mean 0 and variance V_k:

$$\alpha_k \overset{ind}{\sim} \text{N}(0, V_k).$$

And the scale parameter $\sigma > 0$ has inverse-gamma prior distribution with $a > 0$ and $b > 0$ being constants:

$$\sigma \sim \text{IG}(a,b) = \sigma^{-a-1}\exp\left(-\frac{b}{\sigma}\right).$$

For computational convenience, we choose $V_k = 10^3$ and $a = b = 1$. More discussions on the sensitivity analysis will be provided in the Section of Simulation.

Remark: Together with the ALD likelihood, the hierarchical model with the prior in

(4.5) leads to the Bayesian quantile LASSO proposed in Li et al.[18] when the two scale parameters are equal, i.e $s_0 = s_1$. A well recognized limitation of Bayesian quantile LASSO is that the posterior estimates of $\boldsymbol{\beta}$ lack exact sparsity, and its frequentist counterpart, the quantile LASSO, incurs estimation bias by over-shrinking coefficients with large magnitude.

The ssQLASSO enjoys the selective shrinkage and self-adaptive properties that are inherited from the ssLASSO[17] to facilitate more accurate and robust identification of important cancer genomcics features. The selective shrinkage can be achieved through a small amount of penalty on large $|\beta_j|$, and a large amount of shrinkage on small $|\beta_j|$. Meanwhile, the self-adaptivity ensures stabilizing coefficients with large magnitude even when the tuning is at its extremes. Therefore regularization of ssQLASSO is essentially different from that of the adaptive (quantile) LASSO[94]. In addition, compared to Bayesian quantile LASSO[18], its prior structure of adaptively mixing the two Laplace distributions leads to the modal estimate of $\boldsymbol{\beta}$ with exact sparsity. When $s_0 \to 0$, the prior specified by (4.5) reduces to the point mass spike-and-slab priors.

3.2.3 Implementation via the EM Algorithm

The elicitation of the continuous spike-and-slab mixture double exponential prior in ssQLASSO is distinct from the point mass spike-and-slab priors in Bayesian regularized quantile regression mainly in that the scale parameters s_0 and s_1 in (4.5) are treated as deterministic, instead of being random and assigned with conjugate priors[22;87]. Such a difference has paved the way for a extremely fast EM algorithm based on the coordinate descent idea to find the modal estimates of $\boldsymbol{\beta}$ with exact sparsity, and freed us from using the computationally intensive MCMC algorithm in the fully robust Bayesian analysis.

The $\boldsymbol{v} = (v_1, \ldots, v_n)^\top$ from the probabilistic formulation of quantile regression and indicator variables $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_p)^\top$ are treated as missing values in the proposed EM algorithm. Recall that $S_j = (1 - \gamma_j)s_0 + \gamma_j s_1$, which also has missingness due to γ_j. Subsequently, at fixed quantile level τ, let $\boldsymbol{\phi} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma, \boldsymbol{\gamma}, \boldsymbol{v}, \theta)$ the log joint posterior density is given as:

$$
\begin{aligned}
Q &= \log p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma, \boldsymbol{\gamma}, \boldsymbol{v}, \theta | \boldsymbol{Y}) \\
&= \log p(\boldsymbol{Y} | \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma, \boldsymbol{v}) + \sum_{k=0}^{q} \log p(\alpha_k) + \sum_{j=1}^{p} \log p(\beta_j | S_j) + \log p(\sigma) + \sum_{i=1}^{n} p(v_i) \\
&+ \sum_{j=1}^{p} \log p(\gamma_j | \theta) + \log p(\theta) \\
&\propto -\frac{3n}{2} \log(\sigma) - \frac{1}{2} \sum_{i=1}^{n} \log(v_i) - \sum_{i=1}^{n} \left(\frac{(y_i - \boldsymbol{z}_i^\top \boldsymbol{\alpha} - \boldsymbol{x}_i^\top \boldsymbol{\beta})^2}{2\zeta_2^2 \sigma} v_i^{-1} - \frac{\zeta_1 (y_i - \boldsymbol{z}_i^\top \boldsymbol{\alpha} - \boldsymbol{x}_i^\top \boldsymbol{\beta})}{\zeta_2^2 \sigma} + \frac{\zeta_1^2}{2\zeta_2^2 \sigma} v_i \right) \\
&\quad - \sum_{k=0}^{q} \frac{\alpha_k^2}{2V_k} - \sum_{j=1}^{p} \frac{1}{S_j} |\beta_j| - \frac{b}{\sigma} - (a+1) \log(\sigma) - \frac{1}{\sigma} \sum_{i=1}^{n} v_i + \sum_{j=1}^{p} \left[\gamma_j \log\theta + (1 - \gamma_j) \log(1 - \theta) \right].
\end{aligned}
\tag{3.5}
$$

As terms involving v_i $(i = 1, ..., n)$ and γ_j $(j = 1, ..., p)$ are treated as missing, the EM algorithm estimates $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma, \theta)$ through averaging the missing values across the parameters' respective posterior distributions. At E-step, we calculate the conditional posterior expectations of v_i^{-1}, v_i, γ_j and S_j^{-1}. In (4.6), all the terms that involve v_i can be collectively written as:

$$
v_i^{-1/2} \exp\left(\frac{1}{2} (w_{1i}^2 v_i^{-1} + w_2^2 v_i) \right), \tag{3.6}
$$

where $w_{1i}^2 = \frac{(y_i - \boldsymbol{z}_i^\top \boldsymbol{\alpha} - \boldsymbol{x}_i^\top \boldsymbol{\beta})^2}{\zeta_2^2 \sigma}$ and $w_2^2 = \frac{2}{\sigma} + \frac{\zeta_1^2}{\zeta_2^2 \sigma}$, which is the kernel of a generalized inverse Gaussian (GIG) distribution denoted as $v_i | y_i, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma \sim GIG\left(\frac{1}{2}, w_{1i}, w_2\right)$. Here, the probability density function of a random variable Ω following the generalized inverse Gaussian (GIG) distribution, $GIG(t, c_1, c_2)$, is given as

$$
f(\Omega | t, c_1, c_2) = \frac{(c_2/c_1)^t}{2K_t(c_1 c_2)} \Omega^{t-1} \exp\left(-\frac{1}{2} (c_1^2 \Omega^{-1} + c_2^2 \Omega) \right),
$$

where $\Omega > 0$, $c_1 > 0$, $c_2 > 0$ and the real number t is a parameter used in function $K(\cdot)$, the modified Bessel function of the third kind (Karlis, 2002)[95]. The moments of Ω are given by

$$
E(\Omega^\kappa) = \left(\frac{c_1}{c_2} \right)^\kappa \frac{K_{t+\kappa}(c_1 c_2)}{K_t(c_1 c_2)}, \kappa \in \mathrm{R}.
$$

When choosing $t = \frac{1}{2}$ in the above pdf under (4.7), we have

$$\begin{aligned} v_{ni}^{(d)} &= E\left((v_i^{-1})^{(d)}\right) = \left(\frac{w_{1i}^{(d)}}{w_2^{(d)}}\right)^{-1} \frac{K_{-1/2}(w_{1i}^{(d)} w_2^{(d)})}{K_{1/2}(w_{1i}^{(d)} w_2^{(d)})}, \\ v_{pi}^{(d)} &= E\left(v_i^{(d)}\right) = \left(\frac{w_{1i}^{(d)}}{w_2^{(d)}}\right) \frac{K_{3/2}(w_{1i}^{(d)} w_2^{(d)})}{K_{1/2}(w_{1i}^{(d)} w_2^{(d)})}. \end{aligned} \tag{3.7}$$

As for γ_j, its conditional posterior expectation at the d-th iteration can be derived as:

$$\begin{aligned} \eta_j^{(d)} &= p(\gamma_j = 1 | \beta_j^{(d)}, \theta^{(d)}, \boldsymbol{Y}) \\ &= \frac{p(\beta_j^{(d)} | \gamma_j = 1, s_1) p(\gamma_j = 1 | \theta^{(d)})}{p(\beta_j^{(d)} | \gamma_j = 0, s_0) p(\gamma_j = 0 | \theta^{(d)}) + p(\beta_j^{(d)} | \gamma_j = 1, s_1) p(\gamma_j = 1 | \theta^{(d)})} \\ &= \frac{\pi(\beta_j^{(d)} | s_1) \theta^{(d)}}{\pi(\beta_j^{(d)} | s_0)(1 - \theta^{(d)}) + \pi(\beta_j^{(d)} | s_1) \theta^{(d)}}, \end{aligned} \tag{3.8}$$

where $\pi(\beta_j^{(d)} | s_0) = \frac{1}{2s_1} \exp\left(-\frac{|\beta_j^{(d)}|}{s_0}\right)$ and $\pi(\beta_j^{(d)} | s_1) = \frac{1}{2s_1} \exp\left(-\frac{|\beta_j^{(d)}|}{s_1}\right)$. Consequently, the conditional expectation of $(S_j^{-1})^{(d)}$ can be computed as:

$$(\tilde{S}_j^{-1})^{(d)} = E((S_j^{-1})^{(d)} | \beta_j^{(d)}) = E\left(\frac{1}{(1 - \gamma_j) s_0 + \gamma_j s_1} \middle| \beta_j^{(d)}\right) = \frac{1 - \eta_j^{(d)}}{s_0} + \frac{\eta_j^{(d)}}{s_1}. \tag{3.9}$$

At the M-step, we update $\boldsymbol{\alpha}^{(d+1)}$, $\boldsymbol{\beta}^{(d+1)}$, $\sigma^{(d+1)}$ and $\theta^{(d+1)}$ by maximizing the log joint posterior density at d-th iteration, or $Q(\boldsymbol{\phi} | \boldsymbol{\phi}^{(d)})$. The coefficient vector $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are updated in a component-wise manner using the cyclic coordinate decent algorithm[89]. By solving $\frac{\partial Q(\boldsymbol{\phi} | \boldsymbol{\phi}^{(d)})}{\partial \alpha_l} = 0$, $\frac{\partial Q(\boldsymbol{\phi} | \boldsymbol{\phi}^{(d)})}{\partial \beta_m} = 0$, $\frac{\partial Q(\boldsymbol{\phi} | \boldsymbol{\phi}^{(d)})}{\partial \sigma} = 0$ and $\frac{\partial Q(\boldsymbol{\phi} | \boldsymbol{\phi}^{(d)})}{\partial \theta} = 0$, where $l = 1, ..., q$ and

$m = 1, ..., p$, we obtain the following expressions

$$
\begin{aligned}
\theta^{(d+1)} &= \frac{1}{p}\sum_{j=1}^{p}\eta_j^{(d)}, \\
\sigma^{(d+1)} &= \frac{\sum_{i=1}^{n}\left(v_{ni}^{(d)}(y_i - \boldsymbol{x}_i^\top\boldsymbol{\beta}^{(d)} - \boldsymbol{z}_i^\top\boldsymbol{\alpha}^{(d)})^2 - 2(y_i - \boldsymbol{x}_i^\top\boldsymbol{\beta}^{(d)} - \boldsymbol{z}_i^\top\boldsymbol{\alpha}^{(d)})\zeta_1 + v_{pi}^{(d)}(\zeta_1^2 + 2\zeta_2^2)\right) + 2\zeta_2^2 b}{(3n + 2a + 2)\zeta_2^2}, \\
\alpha_l^{(d+1)} &= \frac{\sum_{i=1}^{n} v_{ni}^{(d)}\left(y_i - \boldsymbol{x}_i^\top\boldsymbol{\beta}^{(d)} - \sum_{k=1}^{l-1} z_{ik}\alpha_k^{(d+1)} - \sum_{k=l+1}^{q} z_{ik}\alpha_k^{(d)}\right) z_{il} - \zeta_1\sum_{i=1}^{n} z_{il}}{\frac{\zeta_2^2\sigma^{(d+1)}}{V_l} + \sum_{i=1}^{n} v_{ni}^{(d)} z_{il}^2}, \\
\beta_m^{(d+1)} &= \frac{\operatorname{sgn}(T_m^{(d+1)})\left(\left|T_m^{(d+1)}\right| - (\tilde{S}_m^{-1})^{(d)}\right)_+}{\sum_{i=1}^{n}\frac{v_{ni}^{(d)} x_{im}^2}{\zeta_2^2\sigma^{(d+1)}}},
\end{aligned}
\tag{3.10}
$$

where

$$
T_m^{(d+1)} = \sum_{i=1}^{n}\frac{v_{ni}^{(d)}\left(y_i - \sum_{j=1}^{m-1} x_{ij}\beta_j^{(d+1)} - \sum_{j=m+1}^{p} x_{ij}\beta_j^{(d)} - \boldsymbol{z}_i^\top\boldsymbol{\alpha}^{(d+1)}\right) x_{im}}{\zeta_2^2\sigma^{(d+1)}} - \sum_{i=1}^{n}\frac{\zeta_1}{\zeta_2^2\sigma^{(d+1)}} x_{im}.
$$

The ssQLASSO modal estimator on β_m, the regression coefficient corresponding to the mth genetic factor, shares a remarkable resemblance to the LASSO soft thresholding rule[3;89]. This phenomena is rarely observed in robust penalization with non-differentiable loss functions because thresholding rule exists under smooth losses. In high-dimensional quantile regression, more computationally challenging algorithms, including linear programming and weighted median regression among others, are required to accommodate the non-differentiability in optimization[10;77;96;97]. Therefore, ssQLASSO is computationally more appealing over the freuquentist regularized quantile regression because it can naturally fit into the extremely fast coordinate descent algorithm that merely has a complexity of $O(np)$ when updating $\boldsymbol{\beta}$.

The update on β_m from (4.11) further justifies the selective shrinkage property discussed in Section 2.2. Rather than a fixed tuning parameter for all coordinates of $\boldsymbol{\beta}$, $\tilde{S}_m^{-1}$ is coordinate-specific in that the threshold is smaller for larger $|\beta_m|$, leading to different amount of shrinkage for $|\beta_m|$, $(m = 1, ..., p)$, to avoid the estimation bias in quantile LASSO.

In addition, the threshold $\tilde{S}_m^{-1}$ in ssQLASSO can eventually keep large $|\beta_m|$ stable via slight shrinkage even when the tuning parameter s_0 is at its extremes, i.e. close to 0. Therefore the ssQLASSO is self-adaptive to the underlying sparsity pattern, and is essentially different from adaptive (quantile) LASSO which allocate fixed coordinate-specific weight in the penalty for β_m so all coefficients vanish when the tuning parameter, proportional to $1/s_0$ as shown in (3.3), is at infinity. Note that the intercept and clinical covariates are not subject to selection, so the update of regression coefficients $\boldsymbol{\alpha}$ in (4.11) does not follow the soft-thresholding rule.

In published studies, the Schwarz Information Criterion (SIC) has been widely adopted for tuning selection in high-dimensional quantile regression[42;98], which is defined as

$$SIC = \log \sum_{i=1}^{n} L(y_i - \hat{y}_i) + \frac{\log n}{2n} edf,$$

where $L(\cdot)$ is the quantile check loss function, and edf is the number of zero residuals. We adopt SIC to choose the optimal pair of tuning parameters (s_0, s_1). More detailed analyses of tuning selection are provided in the Section of Simulation. Under the fixed tuning parameters, the algorithm proceeds following steps outlined in Table C.1:

Table 3.1: EM algorithm for the spike-and-slab quantile LASSO.

Algorithm
Initialize iteration $d = 0$, quantile τ, coefficients $\boldsymbol{\alpha}^{(0)}$, $\boldsymbol{\beta}^{(0)}$, $\theta^{(0)}$ and $\sigma^{(0)}$;
Repeat
E-step:
Compute the conditional expectation of $(v_i^{-1})^{(d)}$, $v_i^{(d)}$, $\gamma_j^{(d)}$, $(\tilde{S}_j^{-1})^{(d)}$ via (4.8), (4.9), (4.10);
M-step:
Update $\theta^{(d+1)}$, $\sigma^{(d+1)}$, $\alpha_l^{(d+1)}$ and $\beta_m^{(d+1)}$ through (4.11);
Calculate the updated log joint posterior density and check if $\lvert Q^{(d+1)} - Q^{(d)}\rvert < \delta$;
$d \leftarrow d + 1$;
until Convergence ($\lvert Q^{(d+1)} - Q^{(d)}\rvert < \delta$ for a small value $\delta > 0$).

3.3 Simulation Study

We have conducted simulation studies to demonstrate the effectiveness of the proposed spike-and-slab quantile LASSO (ssQLASSO). Three alternatives, quantile LASSO (QLASSO), spike-and-slab LASSO (ssLASSO)[17;90], and LASSO have been included for comparison. Details of the EM algorithm to fit ssLASSO are provided in the Appendix. R packages *glmnet*[89] and *rqPen*[99] have been adopted to fit the LASSO and QLASSO, respectively. For all the alternative methods, the tuning parameters have been chosen via cross-validation under the squared error for ssLASSO and LASSO, and the check loss for QLASSO. Specifically, we have adopted the function *rq.pen.cv* from package *rqPen* to choose tuning parameter in QLASSO.

The data are generated following model (3.1) with sample size $n = 400$. The dimensions are $p = 800$ and $p = 1600$ for genetic factors $\boldsymbol{x}$ which are simulated from multivariate normal distributions with (1) the auto-regression (AR) structure where the ith and jth genes have correlation $\rho^{|i-j|}$, and (2) banded structure where the ith and jth genes have correlation ρ when $|i - j| = 1$ and 0 otherwise. So the diagonal entries in both structures are 1's. The parameter ρ in both cases is set to 0.5. In model (3.1), the clinical covariates $\boldsymbol{z}$ are not subject to selection, so without loss of generality, we set corresponding coefficients to 0, and the intercept α_0 to 2. 15 nonzero components from the coefficient vector $\boldsymbol{\beta}$ are generated from uniform distribution $U[0.6, 0.8]$, with the rest being 0s. Five probability distributions are considered for random error ϵ_i in model (3.1): $N(\mu, 1)$(Error 1), $t(2)$ with mean μ(Error 2), LogNormal(μ,1) (Error 3), $80\%N(\mu,1) + 20\%N(\mu, 3)$ (Error 4) and Laplace(μ,1) (Error 5). All the errors have long-tailed distributions except Error 1. For each error distribution, the location parameter μ has been adjusted so the τth quantile of the distribution is 0. In addition to the homogeneous random error in model (3.1), we have also considered the heterogeneous model:

$$y_i = 2 + \boldsymbol{x}_i^\top \boldsymbol{\beta} + (1 + x_{i2})\epsilon_i, \tag{3.11}$$

where $\boldsymbol{\beta}$ is generated the same as the coefficient vector under the homogeneous model except that β_2, the coefficient corresponding to the 2nd genetic factor, is always nonzero. The error term becomes non-*i.i.d.* in this case.

Therefore there are 8 case scenarios given different dimensionality (p=800 and 1600), correlation structure (AR1 and banded) and model types (heterogeneous and homogeneous). For each setting, the comparison has been conducted under 3 quantile levels (30%,50% and 70%) and 5 error distributions. The number of true positives (TP), false positives (FP), F_1 score (F1) and Matthews correlation coefficient (MCC) have been adopted to evaluate identification performance. In terms of estimation accuracy, we compute $\sum|\hat{\beta}_j - \beta_j|$, the L1 difference between the estimated and true regression coefficients. Additional simulations have been provided at the end of this section.

Table 3.2 shows the identification and estimation results under the homogeneous model with AR-1 correlation and (n, p)= (400,1600) across quantile levels $\tau = 30\%$, 50% and 70%, and 5 error distributions. The advantage of ssQLASSO over alternatives can be convincingly shown when the model errors are heavy-tailed. For example, under $t(2)$ error and quantile level $\tau = 30\%$, it can be observed that ssQLASSO almost identifies all the important signals with a TP of 14.63 (sd0.81) and nearly negligible number of FPs, 0.33 (sd0.84). Although the number of TPs detected by QLASSO is comparable, the FPs, 34.67 (sd5.57), are much higher. The two non-robust approaches are completely dominated by ssQLASSO. In particular, ssLASSO can only discover roughly 60% of true effects with the TP of 9.33 (sd2.75), and LASSO leads to the largest FPs of 70.47 (sd29.01). In terms of both metrics of F1 score and MCC that more comprehensively reflect the identification performance, the proposed ssQLASSO has also demonstrated the best performance. In particular, the F1 score of 0.98 (sd0.04) and MCC of 0.98 (sd0.04) are the largest among the measures of all methods under comparison. The superior performance of ssQLASSO in identification also results in the smallest estimation error of 2.81 (sd1.12) in contrast to the alternatives. The merit of ssQLASSO can be further established under the rest of skewed model errors (Error3–5) at all quantile levels in Table 3.2. On the other hand, with the light-tailed $N(0,1)$ error, ssQLASSO is the 2nd best method which is marginally outperformed by its non-robust counterpart, ssLASSO. For instance, at $\tau = 30\%$, the F1 score and MCC of ssQLASSO are both 0.98 (sd0.03), with an estimation error of 0.69 (sd0.20), while those metrics for ssQLASSO are 0.97 (sd0.04), 0.97 (sd0.04) and 0.87 (sd0.17), respectively. This

phenomenon indicates that although the ssQLASSO has been proposed as a robust method to accommodate long-tailed errors and outlying observations , it is still very competitive under the $N(0,1)$ error.

In addition to tabulating the mean and standard deviation of the measures assessing identification and estimation performance in Table 3.2, we have graphically displayed the box-plots of F_1 score, MCC and estimation error over 100 replications in Figure 3.2, 3.3 and 3.4, respectively, which visualize Table 3.2 from different perspectives.

Figure 3.2 shows the box-plots of F1 scores over 100 replicates across 5 model errors and 3 quantile levels. Under the $N(0,1)$ error, the box-plots of ssQLASSO and ssLASSO are close to the top of all the sub-panels, indicating majority of the F1 scores computed based on the signals identified by the two methods are close to 1. Their performance are comparable although ssLASSO is slightly better, as shown by a smaller interquartile range (IQR) of F1 scores. Besides, the box-plots of QLASSO and LASSO are corresponding to much lower F1 scores, which further demonstrates the advantage of the spike-and-slab (quantile) LASSO prior in identification accuracy. As the model errors become heavy-tailed in the next 4 columns, ssQLASSO has consistently outperformed the non-robust ssLASSO with most of its F1 scores centering around 1, except a few outliers. Interestingly, for all methods under comparison, ssLASSO results in box-plots with the largest IQR of F1 scores in the presence of data heterogeneity and outliers, revealing the instability of ssLASSO and further justifying the pressing need to develop the robust ssQLASSO. In addition to F1 score, MCC has also been frequently used as a criterion to assess identification usually providing more accurate results in cases of imbalanced data, which is applicable in the simulation with sparsity assumption, e.g., only 15 important features out of a total of 1600 gene expressions. By comparing Figure 3.3 to Figure 3.2, we can draw similar conclusions on the superiority of the proposed ssQLASSO in terms of identification.

Figure 3.4 shares a similar layout of Figure 3.2 and 3.3. The only difference is that the box-plots in Figure 3.4 represent estimation error, $\sum |\hat{\beta}_j - \beta_j|$, across 100 replications. We can observe that ssQLASSO has superior prediction performance especially under heavy-tailed errors, as indicated by the lower box-plots with small IQRs, in spite of a few outlying

estimation errors. Meanwhile, ssLASSO appears relatively unstable with much larger IQRs of the box-plots under all the error distributions except Error 1, $N(0,1)$. The strong contrast between the two additionally justifies the merit of ssQLASSO over its non-robust counterpart to handle data contamination.

In the appendix, we have presented results from simulations under the rest 7 out of the 8 case scenarios based on models with different error ($i.i.d$ and non-$i.i.d$), dimensionality (p=800 and 1600), and correlations (AR1 and banded). Under the homogeneous model with the $i.i.d$ error, Table B.2 shows the performance of all 4 methods with AR-1 correlation and p=800, while Table B.3 and B.4 provide the numeric summaries of identification and estimation performance of banded correlation with dimension 800 and 1600, respectively. We can observe patterns similar to those in Table 3.2, and conclude the advantage of the proposed ssQLASSO. All the methods have also been examined when the data are simulated using the heterogeneous model (4.12). With the non-$i.i.d$ model errors, Table B.5–B.8 in the appendix list numeric measures in identification and estimation from settings according to different combination of correlation (AR1 and banded) and number of predictors (p=800 and 1600). The superiority of ssQLASSO over the alternatives can be drawn not merely under Error2–5. Note that even using $N(0,1)$, the random errors are no longer independent in model (4.12). So it is not surprising to observe the better performance of ssQLASSO under Error 1 over non-robust ones. For instance, Table 3.2 and B.6 are corresponding to the settings that differ only in the type of error under the same correlation (AR-1) and dimension (p=1600). With quantile level τ=0.3 and Error1, the estimation errors of ssQLASSO and ssLASSO in Table 3.2 are 0.87 (sd0.17) and 0.69 (sd0.20), respectively. As the errors become non-$i.i.d$ in Table B.6, ssQLASSO has a smaller L1 error of 1.07 (sd0.18), compared to 1.26 (sd0.20) obtained from ssLASSO.

Table 3.2: Evaluation of homogeneous model under AR-1 correlation, (n, p)=(400,1600) with 100 replicates.

$n = 400$	$p = 1600$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	1.10(1.40)	0.97(0.04)	0.97(0.04)	0.87(0.17)
		QLASSO	15.00(0.00)	35.73(7.54)	0.46(0.06)	0.54(0.05)	3.24(0.35)
		ssLASSO	15.00(0.00)	0.70(0.92)	0.98(0.03)	0.98(0.03)	0.69(0.20)
		LASSO	15.00(0.00)	75.87(27.29)	0.30(0.07)	0.41(0.06)	3.47(0.70)
	Error2	ssQLASSO	14.63(0.81)	0.33(0.84)	0.98(0.04)	0.98(0.04)	2.81(1.12)
		QLASSO	15.00(0.00)	34.67(5.57)	0.47(0.04)	0.55(0.03)	5.41(1.00)
		ssLASSO	9.33(2.75)	0.07(0.25)	0.75(0.14)	0.78(0.12)	5.28(1.87)
		LASSO	13.93(3.25)	70.47(29.01)	0.29(0.09)	0.40(0.06)	9.40(2.63)
	Error3	ssQLASSO	14.90(0.40)	0.10(0.31)	0.99(0.02)	0.99(0.02)	2.31(0.92)
		QLASSO	15.00(0.00)	27.80(8.29)	0.53(0.08)	0.60(0.06)	2.76(0.46)
		ssLASSO	10.20(3.14)	0.07(0.25)	0.79(0.15)	0.81(0.13)	4.47(1.88)
		LASSO	14.93(0.37)	71.73(19.48)	0.30(0.05)	0.41(0.05)	7.05(1.57)
	Error4	ssQLASSO	14.10(1.21)	1.13(1.31)	0.93(0.05)	0.93(0.05)	3.48(1.40)
		QLASSO	14.83(0.38)	31.57(4.41)	0.49(0.04)	0.56(0.03)	6.55(0.92)
		ssLASSO	7.03(3.68)	0.17(0.38)	0.60(0.22)	0.65(0.18)	6.28(2.19)
		LASSO	15.00(0.00)	79.73(29.60)	0.29(0.08)	0.40(0.07)	8.41(1.95)
	Error5	ssQLASSO	14.83(0.46)	1.03(1.63)	0.96(0.05)	0.96(0.05)	3.22(1.47)
		QLASSO	14.90(0.40)	34.40(4.21)	0.47(0.03)	0.54(0.02)	6.69(0.90)
		ssLASSO	8.90(3.92)	0.27(0.58)	0.70(0.22)	0.74(0.19)	5.47(2.54)
		LASSO	15.00(0.00)	68.17(19.49)	0.32(0.06)	0.42(0.05)	6.57(1.13)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	1.33(1.75)	0.96(0.05)	0.96(0.05)	0.85(0.17)
		QLASSO	15.00(0.00)	34.83(6.44)	0.47(0.05)	0.55(0.04)	3.28(0.40)
		ssLASSO	15.00(0.00)	0.87(1.25)	0.97(0.04)	0.97(0.04)	0.70(0.24)
		LASSO	15.00(0.00)	76.10(27.42)	0.30(0.07)	0.41(0.06)	3.49(0.78)
	Error2	ssQLASSO	14.73(0.58)	0.70(1.12)	0.97(0.04)	0.97(0.04)	4.77(1.05)
		QLASSO	15.00(0.00)	49.40(5.10)	0.38(0.02)	0.48(0.02)	4.71(0.48)
		ssLASSO	8.80(3.12)	0.23(0.50)	0.71(0.16)	0.74(0.13)	5.54(1.87)
		LASSO	13.67(3.13)	73.27(38.46)	0.28(0.09)	0.38(0.10)	9.97(2.52)
	Error3	ssQLASSO	14.90(0.31)	0.17(0.38)	0.99(0.01)	0.99(0.01)	4.06(0.88)
		QLASSO	15.00(0.00)	50.97(6.81)	0.37(0.03)	0.47(0.03)	4.07(0.60)
		ssLASSO	9.33(3.46)	0.03(0.18)	0.74(0.19)	0.77(0.16)	4.75(2.17)
		LASSO	15.00(0.00)	84.90(30.72)	0.28(0.07)	0.39(0.06)	8.01(2.07)
	Error4	ssQLASSO	14.43(0.90)	1.27(1.53)	0.94(0.05)	0.94(0.04)	5.02(1.03)
		QLASSO	15.00(0.00)	49.50(5.39)	0.38(0.03)	0.48(0.02)	5.55(0.79)
		ssLASSO	8.50(3.70)	0.23(0.50)	0.68(0.21)	0.72(0.17)	5.54(2.29)
		LASSO	15.00(0.00)	71.60(25.68)	0.31(0.07)	0.42(0.06)	7.86(1.52)
	Error5	ssQLASSO	14.57(1.04)	0.90(1.67)	0.96(0.06)	0.96(0.05)	4.07(0.92)
		QLASSO	15.00(0.00)	51.13(6.43)	0.37(0.03)	0.47(0.02)	5.74(0.76)
		ssLASSO	10.20(3.84)	0.67(0.96)	0.76(0.17)	0.79(0.15)	4.39(2.22)
		LASSO	15.00(0.00)	81.53(33.82)	0.29(0.07)	0.40(0.06)	7.16(2.07)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	1.27(1.70)	0.96(0.05)	0.96(0.04)	0.90(0.19)
		QLASSO	15.00(0.00)	36.43(7.11)	0.46(0.05)	0.54(0.04)	3.08(0.31)
		ssLASSO	15.00(0.00)	0.67(1.32)	0.98(0.04)	0.98(0.04)	0.66(0.16)
		LASSO	15.00(0.00)	67.73(22.53)	0.32(0.08)	0.43(0.07)	3.21(0.52)
	Error2	ssQLASSO	14.63(0.61)	0.73(1.01)	0.96(0.03)	0.96(0.03)	3.34(1.29)
		QLASSO	14.97(0.18)	33.73(3.90)	0.47(0.03)	0.55(0.02)	5.41(0.73)
		ssLASSO	8.70(3.60)	0.13(0.35)	0.70(0.21)	0.74(0.17)	5.34(2.29)
		LASSO	14.33(2.59)	64.37(24.21)	0.31(0.08)	0.42(0.07)	8.23(1.71)
	Error3	ssQLASSO	14.93(0.25)	0.30(0.47)	0.99(0.02)	0.99(0.02)	2.87(0.90)
		QLASSO	14.97(0.18)	33.40(5.30)	0.48(0.04)	0.55(0.03)	5.63(0.72)
		ssLASSO	12.33(2.37)	0.30(0.60)	0.88(0.10)	0.89(0.09)	3.11(1.41)
		LASSO	14.87(0.57)	70.40(26.08)	0.31(0.07)	0.42(0.06)	7.31(1.82)
	Error4	ssQLASSO	14.10(1.21)	1.10(2.43)	0.94(0.07)	0.94(0.06)	3.86(1.64)
		QLASSO	14.93(0.25)	34.20(5.16)	0.47(0.03)	0.55(0.03)	6.54(1.07)
		ssLASSO	7.20(3.18)	0.17(0.38)	0.62(0.19)	0.67(0.16)	6.17(1.81)
		LASSO	14.93(0.25)	74.17(32.31)	0.31(0.07)	0.41(0.06)	8.13(2.41)
	Error5	ssQLASSO	14.80(0.41)	1.03(1.47)	0.96(0.05)	0.96(0.05)	3.53(1.10)
		QLASSO	14.80(0.48)	33.90(5.18)	0.47(0.04)	0.54(0.03)	6.66(1.12)
		ssLASSO	9.53(4.11)	0.87(3.10)	0.72(0.24)	0.77(0.16)	4.76(2.44)
		LASSO	15.00(0.00)	73.70(28.40)	0.31(0.07)	0.41(0.06)	6.87(1.77)

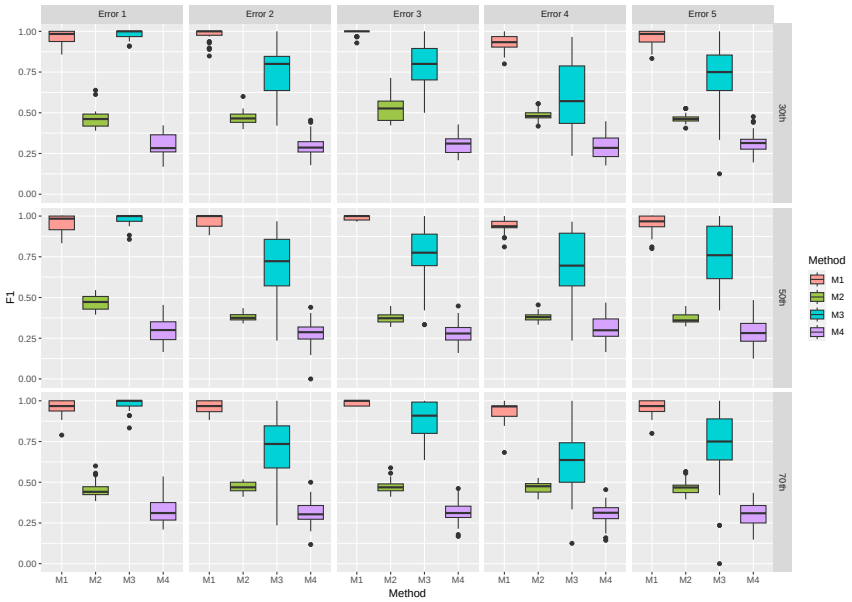

Figure 3.2: F1 score under AR-1 correlation, homogeneous error and $(n, p) = (400, 1600)$. Methods M1-M4 correspond to ssQLASSO, QLASSO, ssLASSO and LASSO, respectively.

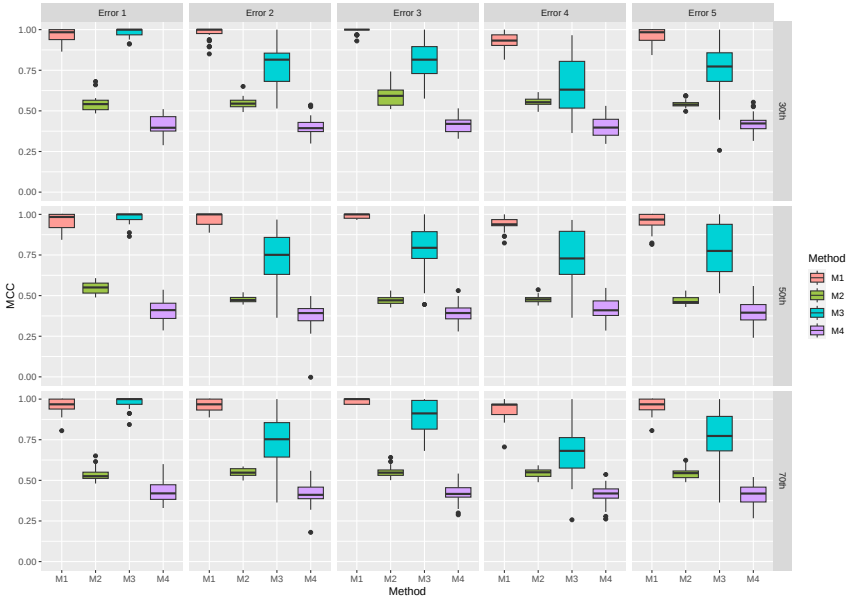

Figure 3.3: MCC score under AR-1 correlation, homogeneous error and $(n, p) = (400, 1600)$. Methods M1-M4 correspond to ssQLASSO, QLASSO, ssLASSO and LASSO, respectively.

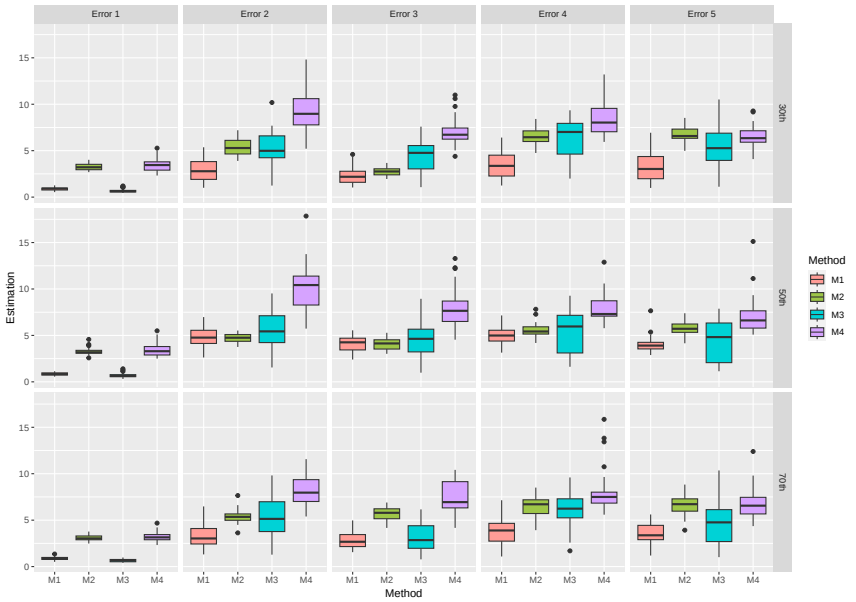

Figure 3.4: Estimation error under AR-1 correlation, homogeneous error and $(n, p) = (400, 1600)$. Methods M1-M4 correspond to ssQLASSO, QLASSO, ssLASSO and LASSO, respectively.

We have performed additional simulations beyond the aforementioned 8 case scenarios. First, we consider generating nonzero coefficients that also include negative ones. In particular, we use $U[0.6, 0.8]$ and $U[-0.8, -0.6]$ to generate 8 and 7 nonzero coefficients in coefficient vector $\boldsymbol{\beta}$, respectively. Under homogeneous model error, AR-1 correlation and p=1600, Table B.9 demonstrates similar patterns in favor of ssQLASSO. Besides, we have investigated the performance when the correlation structure has been extracted from the real genetic data. At each replicate, we randomly draw 1600 gene expressions from the TCGA SKCM data analyzed in the case study. The correlation of those genes are adopted to generate predictors in the homogeneous model. In contrast to Table 3.2 and B.4 that are under AR1 and banded correlation correspondingly, we can observe a clear drop in performance of all methods under comparison in Table B.10. Interestingly, ssQLASSO has even outperformed ssLASSO under *i.i.d.* $N(0,1)$ error in terms of better F1 score, MCC and estimation error. For example, with 50% quantile level, those measures for the proposed ssQLASSO are 0.92 (sd0.06), 0.92 (sd0.05) and 1.58 (sd0.49), respectively. While ssLASSO has yielded an F1 score of 0.45 (sd0.14), an MCC of 0.51 (sd0.12) and L1 error of 8.82 (sd1.51). The ssLASSO has significantly deteriorated even under $N(0,1)$ error. Furthermore, we have conducted

sensitivity analysis on ssQLASSO when using different hyper-parameters (a, b) to specify the inverse-gamma prior on σ. With p=1600 and the AR-1 correlation, we have assessed the performance using $(a, b) = (1, 5)$, $(0.5, 3)$ and $(5, 1)$ for quantile level τ=0.3, 0.5 and 0.7, respectively, in both Table B.11 under *i.i.d.* error and Table B.12 under non-*i.i.d.* error. There is no deviations from the patterns shown in the counterpart Table 3.2 and B.6 under $(a, b) = (1, 1)$. Therefore, the performance of ssQLASSO is not sensitive to different choices of hyper-parameters.

Tuning parameter selection: The performance of ssQLASSO critically depends on the tuning parameters s_0 and s_1, the scales for the spike and slab components, correspondingly. In our simulation study, the optimal pair of (s_0, s_1) has been chosen via the Schwarz Information Criterion (SIC). Figure 3.5 has revealed the V-shape "valley" of the 3D surface plot of SIC. Compared to the spike scale s_0 that determines the sparsity of ssQLASSO, we can see that the choices of s_1 is relatively less important. As the slab scale s_1 changes from 4 to 6, the smallest SIC criterion value on the V-shape "valley" do not have too much variations if the spike scale s_0 has been selected appropriately. Such a pattern is similar to the surface plot of prediction error of ssLASSO in published studies[90;91], and explains why the strategy of fixing the slab scale s_1 (e.g. s_1=2) and conducting a fine tune of s_0 over a grid of values has been adopted. We have found out that there is not much difference between such a strategy and our search of optimal pair of tunings. In the Appendix, we have provided the 3D surface plot of cross-validation error in terms of quantile check loss in Figure B.1, which shows the V-shape "valley" similar to Figure 3.5. Consequently, the superior performance of ssQLASSO observed in the simulation study is insensitive to the strategies adopted to locate optimal tuning parameters.

Solution path and self-adaptivity: The impact of the spike scale s_0 on sparsity can be thoroughly understood by examining the solution path of ssQLASSO. With (n,p)=(400,1600) and AR1 correlation, we generate the response from homogeneous model under Error3 and quantile level $\tau = 0.5$, by setting 5 coefficients in vector $\boldsymbol{\beta}$ as nonzero, i.e., (0.8, 0.6, 0.4, -0.5, -0.6). When fitting the data using ssQLASSO, we fix the slab scale s_1 to 1. The solution paths corresponding to LASSO and ssQLASSO are shown in Figure B.2 and B.3

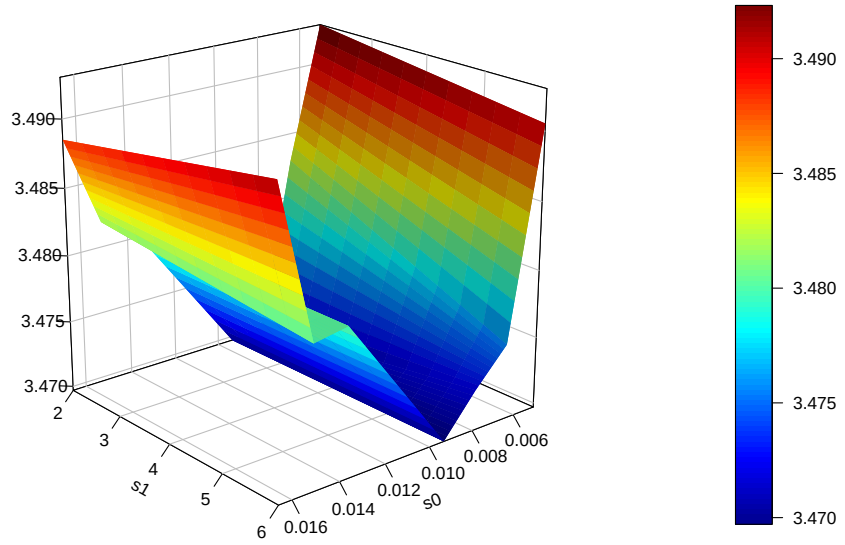

Figure 3.5: The surface plot of Schwarz Information Criterion (SIC) using ssQLASSO under $\tau = 30\%$, AR-1 correlation, $(n = 400, p = 1600)$ and error 3 over 100 replications.

in Appendix, respectively. For illustration purpose, we only present paths corresponding to the aforementioned 5 nonzero coefficients. As λ increases, all the coefficient profiles of LASSO estimates in Figure B.2 converge to the horizontal axis at 0, regardless of the nonzero coefficients or not. On the contrary, Figure B.3 shows that with s_0 growing beyond the optimal tuning around 0.01, those important coefficients are consistently included in the model with little or no shrinkage.

Therefore, unlike LASSO, the ssQLASSO is self-adaptive to underlying sparsity pattern even when an excessive amount of penalty is imposed. The slab component is essential in preventing coefficients with large magnitude from being absorbed by the spike component, so the coefficient trajectories still stabilize as the amount of shrinkage increases. Otherwise the pattern of solution path will be the same as that of Figure B.2. The self-adaptivity

also distinguishes the selective shrinkage mechanism of ssQLASSO from that of adaptive (quantile) LASSO which shrinks all the coefficients to 0 under large penalty.

Computational time: The computational advantage of ssQLASSO over the alternatives can be better understood by assessing the run time using the setting of Table 3.2, that is, a homogeneous data generating model with AR-1 correlation and $(n = 400, p = 1600)$. For a fair comparison, the run time in seconds is recorded under fixed tuning parameters, as shown in Table C.2 from the Appendix. Overall, the computational cost of ssQLASSO is much less than QLASSO and comparable to its non-robust counterpart ssLASSO. For example, at τ=0.3 and the mixture normal error (Error4), the run time of ssQLASSO is 24.19(sd3.84), which is comparable to 10.43(sd0.56) of ssLASSO and much smaller than 264.00(sd10.66) of QLASSO.

Robust regularization with non-differentiable loss function incurs more cost in computation compared to non-robust methods with smooth losses, as shown by the run time difference between QLASSO and LASSO in Table C.2. However, the run time of ssQLASSO and ssLASSO are of the same order, much less than that of the quantile LASSO (QLASSO). Such a phenomenon is due to formulating quantile regression with the asymmetric Laplace distribution which results in the quadratic form of the objective function in equation (4.6). Therefore, the estimation of high-dimensional coefficient vector $\boldsymbol{\beta}$ can be conducted through the fast coordinate descent update of the soft-thresholding rule[89] which is not attainable under non-differentiable loss.

Additional simulations: We have further assessed the performance of ssQLASSO and alternatives through additional simulations in Section F from the Appendix. In Section F.1, we have considered much larger settings $(n, p) = (800, 2400)$ and (800,3200) in Tables B.14 – B.17, as well as the settings using data generating models incorporating the clinical covariates not under selection in Table B.18 and Table B.19. The advantage of ssQLASSO over alternatives still stands. Accordingly, the computational times of the proposed method under growing study sizes have been shown in Figure B.4. In Section F.3, we have evaluated the appropriateness of adopting the asymmetric Laplace distribution (ALD) for the likelihood function. The maximum overlapping density plots in Figure B.5 and Figure B.6 justify

the utility of ALD when distributional assumptions are violated. The summary statistics on number of outliers in simulation are provided in Table B.20 in Section F.5, indicating adaptivity of ALD under a larger numbers of outliers generated from the heterogeneous models. Overall, the additional simulation studies have further demonstrated the superiority of ssQLASSO over competing methods.

3.4 Real Data Analysis

The Cancer Genome Atlas (TCGA, https://cancergenome.nih.gov/) is a joint effort from the National Cancer Institute and the National Human Genome Research Institute to provide high-quality molecular profiling data for 33 cancer types. We have analyzed the Cancer Genome Atlas (TCGA) lung adenocarcinomas (LUAD) and skin cutaneous melanoma (SKCM) data using level 3 gene expression measurements downloaded from the cBio Cancer Genomics Portal (Cerami et al., 2012)[70].

3.4.1 TCGA lung adenocarcinomas (LUAD) data

Lung cancer is the leading cause of cancer-related deaths in the United States. LUAD is one of the most common subtype of lung cancer, accounting for approximately 40% of all lung cancer cases. With the TCGA LUAD data (Campbell et al. (2011)[71]), we treat the variable "forced expiratory volume in 1 second", or FEV1, as the disease phenotype, with the corresponding histogram shown in Figure 3.1. Age, gender, smoking history and pack per year have been included as clinical covariates, where gender and smoking history are categorical, and the other two are continuous. We match the gene expression measurements with the phenotypic trait and clinical covariates through common subjects, resulting in a working dataset with a total of 208 subjects and 20,531 gene expressions. Following published studies ([100;101]), we first conduct pre-screening to reduce the computational cost. The gene expressions have been ranked through coefficient of variations, and we have selected top 1200 genes for downstream analysis.

We have set quantile level $\tau = 50\%$ for ssQLASSO and QLASSO, and applied all four methods to the LUAD dataset. Table 3.3 shows the pairwise overlap of findings, demonstrating the difference in identification results between the proposed method and alternatives. Specifically, the proposed ssQLASSO has identified 17 genes which are LCOR, ALPL, AATBC, AADAT, AGPAT3, POLR3G, C6orf62, UCK2, NUDT16L1, SMC3, SRP14, TMEM217, RPP25 and DDO. In particular, there are only 3 genes commonly discovered by ssQLASSO and SSLASSO, indicating that the spike-and-slab quantile LASSO prior leads to a quite different set of findings when heterogeneity in disease trait exists. Among these genes that have been selected by ssQLASSO, SMC3, SRP14 and RPP25 have not been reported by other approaches. Wang et al., 2018[102] have confirmed that the progression of lung cancer can be inhibited via targeting SMC3, or the structural maintenance of chromosomes 3, through hydrogen gas. Malik et al., 2019[103] have reported SRP14 as a potential biomarker of LUAD, which is consistent with its changed expression levels in cell lines with KRAS mutations. In addition, Mecoli et al., 2021[104] have shown that patients with antibodies against RPP25 are significantly less likely to develop cancer within 2 years of systemic sclerosis (SSc) onset which are frequently caused by lung disease (Steen and Medsger, 2007)[105].

We have evaluated the prediction accuracy based on multi-split of the data. At each split, three-fourth of the subjects are randomly chosen as the training set for model fitting, while the remaining data is utilized as the testing set to assess prediction performance. Then the mean and standard deviation of prediction errors can be computed over the 100 splits. In practice, prediction performance of robust and non-robust methods are usually evaluated using different criteria, such as the prediction mean squared errors (PMSE) for ssLASSO and LASSO, and prediction mean absolute deviations (PMAD) for ssQLASSO and QLASSO at the 50% quantile level. For a fair cross-comparison, we have computed both PMAD and PMSE for all the methods. The PMAD of the robust ssQLASSO and QLASSO are 13.51 (sd 1.84) and 15.61 (sd 1.69), respectively. Their corresponding PMSEs are 327.38 (sd 90.35) and 435.86 (sd 94.78). For non-robust methods ssLASSO and LASSO, the PMADs are 15.65 (sd 2.02) and 17.65 (sd 1.91), and the PMSEs are 419.52 (sd 113.39) and 557.56 (sd 125.24), respectively. The proposed ssQLASSO has the smallest prediction

error in terms of both PMADs and PMSEs. The first row of Figure 3.6 listed the box-plots of PMADs and PMSEs of four methods across 100 splits, which clearly indicate the superior performance of ssQLASSO in terms of prediction. Particularly, it appears that ssQLASSO is more stable as there are no outlying prediction errors as shown in other box-plots. We can obtain the same conclusions regarding the better performance of ssQLASSO in Section F.5 from the Appendix when analyses at quantile levels 0.3 and 0.7 have been conducted with overlapping results shown in Table B.21.

3.4.2 TCGA skin cutaneous melanoma (SKCM) data

Cutaneous melanoma (CM) is the most prevalent form of melanoma across the world. Its incidence has risen steadily, especially among young adults. While CM only comprises of 3% of all skin cancer cases, it causes 65% of skin cancer deaths[106]. In CM, Breslow's thickness is one of the most important prognostic factors indicating the progression of the disease. Figure 3.1 shows that in TCGA SKCM, the Breslow's depth is skewed and has outlying observations even on the log scale. In the case study, we use the log-transformed Breslow's depth as the response variable. The clinical factors are age, initial cancer year and gender, where the first two are continuous and gender is categorical. The initial dataset after matching the phenotype with the genotypes consists of 353 subjects and 20,531 gene expressions. Similar to the LUAD study, we have selected the top 1200 genes for further analysis.

For ssQLASSO and QLASSO, we set the quantile level τ to 50% and analyze the SKCM dataset using all methods under comparison. The number of identified and overlapping genes are shown in the right panel of Table 3.3, revealing smaller numbers of overlapping genes in general, compared to those from analysis of LUAD. In particular, ssQLASSO and QLASSO have only commonly identified 2 genes, which indicates the challenge of analysis in the presence of heavily skewed phenotypic trait and the necessity of developing robust variable selection methods. The proposed ssQLASSO methods have identified 13 important genes: TTC28-AS1, NPR1, CEMIP, FLJ20373, LOC126807, GJA3, INPP1, FMNL2, COQ9, ZFAT, KCTD7, KCTD21 and TFR2. Among these genes, MAP4K4, GJA3, INPP1 and

KCTD7 have not been found by the alternatives. Surman et al., 2023[107] have reported that during the metastasis of cutaneous melanoma (CM), the CM-derived exosomes may have contributions to organ-specific metastasis, and the MAP4K4 are in presence of majority of the CM-derived exosome samples. Orellana et al.2021[108] have found that melanoma patients with higher GJA3 tumor expression levels have a much worse prognosis than those exhibiting lower GJA3 tumor expressions. As for INPP1, Benjamin et al., 2014[109] have reported the up-regulation of INPP1 in melanoma tumors compared to normal skin tissues. Moreover, Angrisani et al., 2021[110] have shown that KCTD7 is up-regulated in tumors influencing skin tissues.

Since it is difficult to have an objective assessment of the selection accuracy in real data, we conduct multiple data splitting to partially justify the utility of ssQLASSO. For robust methods ssQLASSO and QLASSO, the PMADs are 0.76 (sd 0.08) and 0.86 (sd 0.06) and the PMSEs are 1.07 (sd 0.33) and 1.41 (sd 0.33), respectively. For non-robust methods ssLASSO and LASSO, the PMADs are 0.86 (sd 0.07) and 0.88 (sd 0.08) and the PMSEs are 1.41 (sd 0.34) and 1.42 (sd 0.36), respectively. The box-plots of PMSE and PMAD across 100 splits for SKCM data are shown in the second row of Figure 3.6, which has demonstrated the better prediction performance of ssQLASSO over alternatives in both criteria. Consistent observations on the improved performance of ssQLASSO can be observed in Section F.5 from the Appendix under quantile levels 0.3 and 0.7. The overlapping results are provided in Table B.22.

Table 3.3: Identification results for real data analysis. The numbers of genetic effects identified by different approaches and their overlaps.

	LUAD				SKCM			
	ssQLASSO	QLASSO	ssLASSO	LASSO	ssQLASSO	QLASSO	ssLASSO	LASSO
ssQLASSO	17	12	3	7	13	2	4	8
QLASSO		30	6	9		17	1	6
ssLASSO			12	9			8	4
LASSO				51				27

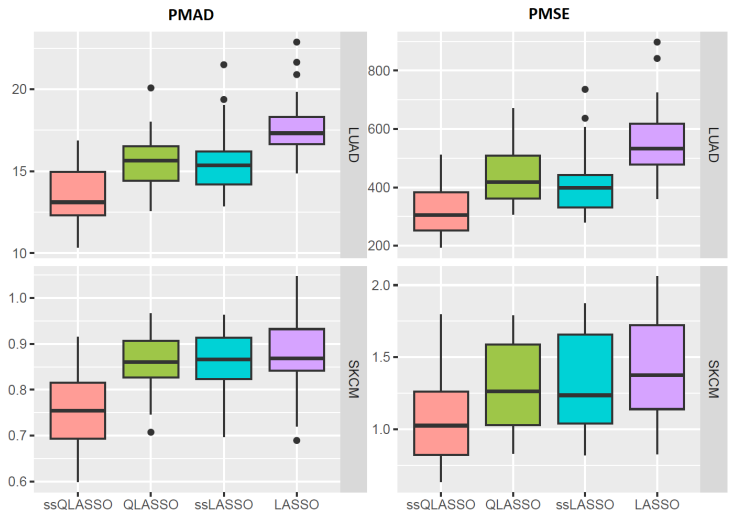

Figure 3.6: PMAD (the first column) and PMSE (the second column) corresponding to the LUAD (the first row) and SKCM (the second row).

3.5 Discussion

The main appeal of the ssQLASSO lies in quickly finding the posterior mode in robust Bayesian analysis, leading to a much smaller computational cost compared to MCMCs. In literature, the variational Bayes methods and related ones also suffice such an analysis goal[111–113]. A direct comparison between the two types of methods worths further explorations, particularly under disease traits with skewed distributions and heavy-tailed errors. Recently, it has been revealed that the spike–and–slab LASSO can be developed for approximate posterior sampling and uncertainty quantification based on Bayesian bootstrap[114], which is potentially applicable for extending ssQLASSO to approximate the posterior distributions beyond the mode detection.

Analysis of convergence rate of EM algorithms in high-dimensional settings remains an open problem with sparse literature available. Although such a rate has not been shown for spike–and–slab LASSO yet, we speculate that ssQLASSO will demonstrate a similar convergence pattern since update of the high-dimensional coefficient vector $\boldsymbol{\beta}$ in the M step can be done using the same coordinate-descent strategy as in ssLASSO. To handle ultra-high-dimensional predictors such as in the GWAS, we suggest a marginal screening of the

predictors to a reasonable scale so ssQLASSO and alternative methods can be applied, which has been widely acknowledged and adopted in ultra-high dimensional studies[115;116].

The robustness of ssQLASSO comes from its resistance to outliers and data contamination in the response variable, such as disease phenotypes. Since it assumes linear structure on predictors, the proposed method may lead to inferior performances when the model is misspecified due to structured sparsity such as interaction or group structures[38;117]. We conjecture that it can be further extended to handle more complicated underlying data structures. For example, a group spike-and-slab quantile LASSO can be developed to incorporate group structures.

Last but not least, quantile regression has been historically proposed to dissect heterogeneous data through estimation at a fixed quantile level[118]. It is a well acknowledged limitation of quantile regression, especially under high-dimensional scenarios, to achieve a final set of predictors across different quantile levels[18;96;98;116]. We follow the common practice of published studies to report results at different values of τ when a specific quantile level of interest is lacking. Such an analysis can be improved by considering composite quantile regression and modeling multiple quantile levels simultaneously[119]. Specifically, we conjecture that the Bayesian composite quantile regression can be incorporated in the proposed ssQLASSO[120].

Chapter 4

The Spike-and-slab Quantile Group LASSO for Non-parametric Varying Coefficient Mixed Model in Longitudinal GWAS

4.1 Introduction

In recent decades, high-dimensional longitudinal data have been widely generated and collected from high-throughput technologies. These data has the following distinct characteristics. First, the disease traits are repeatedly measured across a sequence of time points so there exists intra-cluster correlations among the repeated measurements. Second, due to heterogeneity of complex diseases, long-tailed distributions and heavy-tailed errors have been frequently observed in the disease phenotype. Third, the omics measurements are of high-dimensionality and presents structured sparsity including strong correlations[7;8], networks[9;31], pathways and gene sets[32;33;35], as well as interactions[38]. Therefore, tailored variable selection methods have been demanded to accommodate all these defining characteristics. In published longitudinal studies, regularized variable selection methods have been developed

under main effect models[23;26], interaction models[28;121], as well as non-parametric and semi-parametric models[39;42]. These penalization methods have been mostly developed within either the generalized estimation equation (GEE) or mixed-effects model framework, and usually lacks robustness to outliers and skewed distributions in the response. On the other hand, robust Bayesian mixed effect models have been proposed to account for irregularity in the response variable[122;123]. However, these models can only handle low-dimensional data and cannot perform variable selection. Moreover, these studies are based on fully Bayesian analysis which are computationally intensive even with low-dimensional data.

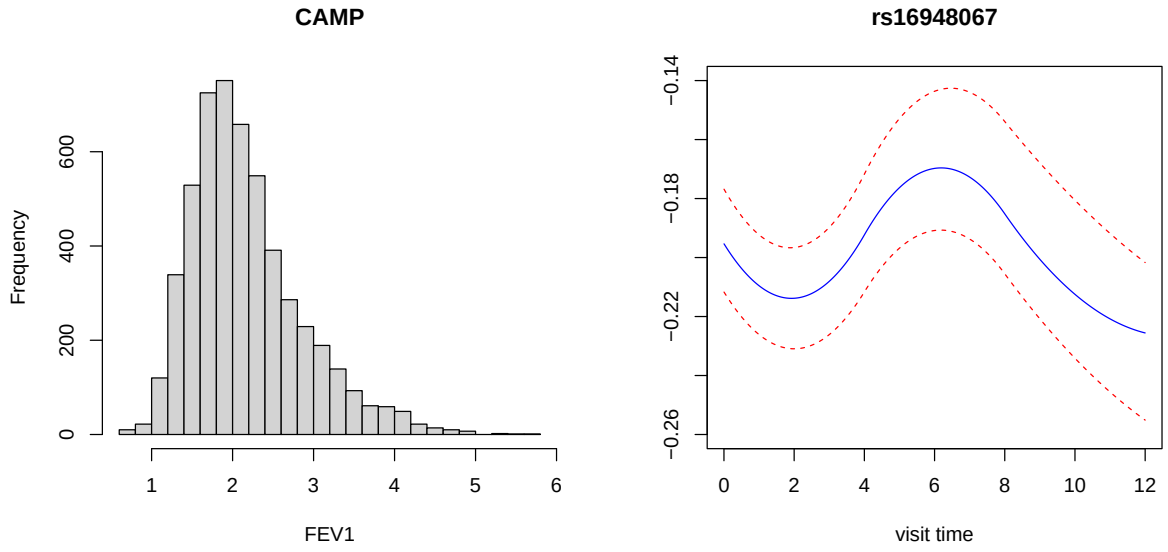

Figure 4.1: Distribution of the FEV1 (left) and non-linear genetic effect of SNP rs16948067 (right) from the NHS data. The red dashed lines denote the 95% confidence interval band.

Our study has been partially motivated by overcoming the methodological limitation of existing frequentist and Bayesian longitudinal analyses, as well as conducting a tailored analysis for the Childhood Asthma Management Program (CAMP) data with longitudinal measurements on the forced expiratory volume in one second (FEV1) and high-dimensional SNPs. The left panel of Figure 4.1 shows skewed distributions of FEV1 collected over all the 12 time points. The right panel shows a significant nonlinear time-varying interactions between SNP rs16948067 and visit time of patients under the longitudinal FEV1 traits. The

nonlinear time-varying effects is expected to be accounted in the high-dimensional longitudinal models.

We have developed the spike-and-slab quantile group LASSO for non-parametric varying coefficient mixed model to analyze the CAMP data. The varying coefficient modelling accounts for nonlinear time varying effects with respect to SNPs. Identification of important varying coefficients amounts to selection of groups of spline coefficients. The proposed model borrows strength from both regularized quantile group LASSO and Bayesian group LASSO while overcoming their respective disadvantages. Estimation of model parameters can be achieved using the EM algorithm. In particular, the posterior mode of high-dimensional regression coefficients can be expressed as a group LASSO solution under the soft-thresholding rule within the coordinate descent framework. To accommodate the long-tailed distributions, we have adopted the robust asymmetric Laplace distribution based likelihood function that has been widely used in Bayesian quantile regression. In addition, the correlations among repeated measured measurements can be handled by specifying the random effects in the mixed model. Recently, multiple software packages have been developed for analyzing high-dimensional longitudinal data[124–126]. Besides methodological development, we have also developed an R package embays to implement all the proposed and alternative methods[1]. Numerical experiments suggest that the proposed method can lead to superior estimation and identification performances in a wide array of simulation scenarios.

4.2 Statistical Methods

4.2.1 The Quantile Varying Coefficient Mixed Model

Let $y_i(t_{ik_i})$ be the response of the i-th subject $(i = 1, \ldots, n)$ at time t_{ik_i} $(k_i = 1, ..., K_i)$. For simplicity, we assume $k_i = k = 1, ..., K$ indicating each subject being repeatedly measured same number of times. For a given quantile level $0 < \tau < 1$, the varying-coefficient data

generating model takes the following form:

$$y_i(t_{ik}) = \sum_{l=1}^{q} c_i^{(l)}(t_{ik})\alpha_{l,\tau}(t_{ik}) + \sum_{j=1}^{p} x_i^{(j)}(t_{ik})\beta_{j,\tau}(t_{ik}) + \boldsymbol{w}_i(t_{ik})^\top \boldsymbol{\gamma}_{i,\tau} + \epsilon_{i,\tau}. \tag{4.1}$$

The fixed effect part consists of $\boldsymbol{c}_i(t_{ik}) = (c_i^{(1)}(t_{ik}), \ldots, c_i^{(q)}(t_{ik}))^\top$ which represents a q-dimensional vector of the intercept as well as clinical factors and $\boldsymbol{x}_i(t_{ik}) = (x_i^{(1)}(t_{ik}), \ldots, x_i^{(p)}(t_{ik}))^\top$ which is a p-dimensional vector of genetic factors. Their corresponding fixed effect varying coefficient smooth function vectors at τ-th quantile are $\boldsymbol{\alpha}_\tau(t) = (\alpha_1(t), ..., \alpha_q(t))^\top$ and $\boldsymbol{\beta}_\tau(t) = (\beta_1(t), ..., \beta_p(t))^\top$. Unknown smooth varying-coefficient functions $\alpha_l(\cdot)$'s and $\beta_j(\cdot)$'s vary with different subject at different time. For the random effect part, $\boldsymbol{w}_i(t_{ik}) = (w_i^{(1)}(t_{ik}), \ldots, w_i^{(p_n)}(t_{ik}))^\top$ is a p_n-dimensional known vector with $\boldsymbol{\gamma}_{i,\tau}$ being the p_n-dimensional random effect coefficient vector. For example, when considering random intercept and time as random slope, $p_n = 2$ and $\boldsymbol{w}_i(t_{ik}) = (w_i^{(1)}(t_{ik}), w_i^{(2)}(t_{ik}))^\top = (1, t_{ik})^\top$ in which $w_i^{(1)}(t_{ik})$ and $w_i^{(2)}(t_{ik})$ represent random intercept and random slope respectively. Random error $\epsilon_{i,\tau}$'s are independent with the τ-th quantile being 0. For simplicity, we omit the subscript "τ" without loss of generality.

4.2.2 The Basis expansion

To construct the Bayesian quantile varying coefficient model, we start by approximating the varying coefficient functions $\alpha_l(\cdot)$ and $\beta_j(\cdot)$ in model (4.1) through basis expansion using polynomial splines. Let $\tilde{\boldsymbol{\pi}}_l(\cdot) = (\tilde{\pi}_{l,1}(\cdot), ..., \tilde{\pi}_{l,d_n}(\cdot))^\top$ and $\boldsymbol{\pi}_j(\cdot) = (\pi_{j,1}(\cdot), ..., \pi_{j,d_n}(\cdot))^\top$ be two sets of normalized B-spline basis functions with $d_n = k_n + D + 1$ (Schumaker, 2007)[127], in which k_n is the number of uniform interior knots and $D = 1$ or 2 denotes linear or quadratic splines respectively. Then, varying coefficient functions $\alpha_l(\cdot)$ and $\beta_j(\cdot)$ can be approximated by a linear combination of $\tilde{\pi}(\cdot)$ and $\pi(\cdot)$:

$$\alpha_l(\cdot) \approx \sum_{m=1}^{d_n} \tilde{\pi}_{l,m}(\cdot)\xi_{l,m} = \boldsymbol{\xi}_l^\top \tilde{\boldsymbol{\pi}}_l(\cdot),$$
$$\beta_j(\cdot) \approx \sum_{m=1}^{d_n} \pi_{j,m}(\cdot)\theta_{j,m} = \boldsymbol{\theta}_j^\top \boldsymbol{\pi}_j(\cdot),$$

where $\boldsymbol{\xi}_l = (\xi_{l,1}, ..., \xi_{l,d_n})^\top$ and $\boldsymbol{\theta}_j = (\theta_{j,1}, ..., \theta_{j,d_n})^\top$ are the spline coefficient vectors with dimensionality equal to q and p respectively. Subsequently, model (4.1) can be rewritten as

$$y_i(t_{ik}) = \boldsymbol{A}_i(t_{ik})^\top \boldsymbol{\xi} + \boldsymbol{B}_i(t_{ik})^\top \boldsymbol{\theta} + \boldsymbol{w}_i(t_{ik})^\top \boldsymbol{\gamma}_i + \epsilon_i(t_{ik}). \tag{4.2}$$

where $\boldsymbol{A}_i^{(l)}(t_{ik}) = \tilde{\boldsymbol{\pi}}_l(t_{ik}) x_i^{(l)}(t_{ik}) = (\tilde{\pi}_{l,1}(t_{ik}) x_i^{(l)}(t_{ik}), ..., \tilde{\pi}_{l,d_n}(t_{ik}) x_i^{(l)}(t_{ik}))^\top$ and $\boldsymbol{B}_i^{(j)}(t_{ik}) = \boldsymbol{\pi}_j(t_{ik}) x_i^{(j)}(t_{ik}) = (\pi_{j,1}(t_{ik}) x_i^{(j)}(t_{ik}), ..., \pi_{j,d_n}(t_{ik}) x_i^{(j)}(t_{ik}))^\top$ which leads to $(q \times d_n)$-dimensional matrix $\boldsymbol{A}_i(t_{ik}) = (\boldsymbol{A}_i^{(1)}(t_{ik})^\top, ..., \boldsymbol{A}_i^{(q)}(t_{ik})^\top)^\top$ and $(p \times d_n)$-dimensional matrix $\boldsymbol{B}_i(t_{ik}) = (\boldsymbol{B}_i^{(1)}(t_{ik})^\top, ..., \boldsymbol{B}_i^{(p)}(t_{ik})^\top)^\top$.

4.2.3 The Quantile Varying Coefficient Model with Asymmetric Laplace Distribution

Consider the quantile varying coefficient model (4.2) after basis expansion was applied, under a fixed quantile level τ, the estimation of regression coefficients $\boldsymbol{\xi}$, $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ can be obtained by minimizing the following loss function

$$L(\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}) = \sum_{i=1}^{n} \sum_{k=1}^{K} \rho_\tau(y_i(t_{ik}) - \boldsymbol{A}_i(t_{ik})^\top \boldsymbol{\xi} - \boldsymbol{B}_i(t_{ik})^\top \boldsymbol{\theta} - \boldsymbol{w}_i(t_{ik})^\top \boldsymbol{\gamma}_i),$$

where $\rho_\tau(\epsilon_i) = \epsilon_i\{\tau - I(\epsilon_i < 0)\}$ is defined as the check loss function.

From the Bayesian perspective, the ALD distribution is widely used in maximum likelihood estimation and Bayesian quantile regression (Yu and Moyeed, 2001)[19]. Under quantile level τ, the probability density function of ALD can be expressed as

$$f(y_i(t_{ik})|\mu_i(t_{ik}), \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp\left(-\rho_\tau\left(\frac{y_i(t_{ik}) - \mu_i(t_{ik})}{\sigma}\right)\right),$$

in which $\mu_i(t_{ik}) = \boldsymbol{A}_i(t_{ik})^\top \boldsymbol{\xi} + \boldsymbol{B}_i(t_{ik})^\top \boldsymbol{\theta} + \boldsymbol{w}_i(t_{ik})^\top \boldsymbol{\gamma}_i$ is the location parameter, $\sigma(> 0)$ is the scale parameter. The joint distribution of $(n \times K)$-dimensional vector $\boldsymbol{Y} = (y_1(t_{11}), ..., y_n(t_{nK}))^\top$ given $\boldsymbol{A} = (\boldsymbol{A}_1(t_{11}), ..., \boldsymbol{A}_n(t_{nK}))^\top$, $\boldsymbol{B} = (\boldsymbol{B}_1(t_{11}), ..., \boldsymbol{B}_n(t_{nK}))^\top$ and $\boldsymbol{w} = (\boldsymbol{w}_1(t_{11}), ..., \boldsymbol{w}_n(t_{nK}))^\top$

can be written as

$$f(\boldsymbol{Y}|\boldsymbol{A},\boldsymbol{B},\boldsymbol{w},\boldsymbol{\xi},\boldsymbol{\theta},\boldsymbol{\gamma},\sigma) = \left(\prod_{i=1}^{n}\prod_{k=1}^{K}\frac{\tau(1-\tau)}{\sigma}\right)\exp\left(-\sum_{i=1}^{n}\sum_{k=1}^{K}\rho_\tau\left(\frac{y_i(t_{ik})-\mu_i(t_{ik})}{\sigma}\right)\right).$$

Therefore, minimizing $L(\boldsymbol{\xi},\boldsymbol{\theta},\boldsymbol{\gamma})$ is equivalent to maximizing the (log-)likelihood function. The application of ALD distribution was further extended after Kozumi and Kobayashi (2011)[92] demonstrated that the distribution can be expressed as a mixture of a scaled normal distribution and an exponential distribution, which leads to a more flexible Bayesian hierarchical model. Let

$$\zeta_1 = \frac{1-2\tau}{\tau(1-\tau)} \text{ and } \zeta_2 = \sqrt{\frac{2}{\tau(1-\tau)}}.$$

Then the response variable will have the following representation:

$$y_i(t_{ik}) = \mu_i(t_{ik}) + \zeta_1 v_i(t_{ik}) + \zeta_2\sqrt{\sigma v_i(t_{ik})}u_i(t_{ik}),$$

where $v_i(t_{ik})$ follows an exponential distribution with a scale parameter $\sigma(>0)$, $\text{Exp}(\sigma)$ and $u_i(t_{ik})$ follows a standard normal distribution, $N(0,1)$. Therefore, a $y_i(t_{ik})$ following $ALD(\mu_i(t_{ik}),\sigma,\tau)$ has the hierarchical representation as follows:

$$\begin{aligned} y_i(t_{ik})|v_i(t_{ik}) &\overset{ind}{\sim} \text{N}(\mu_i(t_{ik}) + \zeta_1 v_i(t_{ik}), \zeta_2^2\sigma v_i(t_{ik})), \\ v_i(t_{ik}) &\overset{ind}{\sim} \text{Exp}(\sigma), \quad (i=1,...,n, k=1,\ldots,K), \\ u_i(t_{ik}) &\overset{ind}{\sim} \text{N}(0,1), \quad (i=1,...,n, k=1,\ldots,K). \end{aligned}$$

4.2.4 The Spike-and-slab Quantile Group LASSO

In the frequentist framework, group penalties are widely utilized to address variable selection problems in non-parametric regression like varying coefficient models. For example, group SCAD (Fan and Li, 2001)[4] was adopted by Wang et al. (2008)[39] as variable selection approach for varying coefficient model. Assuming each group is of the same size, adding

a group LASSO penalty term to the quantile regression model (4.2) gives us the following quantile group LASSO model:

$$\min_{\boldsymbol{\xi},\boldsymbol{\theta},\boldsymbol{\gamma}} \sum_{i=1}^{n}\sum_{k=1}^{K} \rho_\tau(y_i(t_{ik}) - \boldsymbol{A}_i(t_{ik})^\top \boldsymbol{\xi} - \boldsymbol{B}_i(t_{ik})^\top \boldsymbol{\theta} - \boldsymbol{w}_i(t_{ik})^\top \boldsymbol{\gamma}_i) + \lambda \sum_{j=1}^{p} \|\boldsymbol{\theta}_j\|_2, \tag{4.3}$$

where $\lambda > 0$ is the tuning parameter determining sparsity of the model. Under the Bayesian framework, we define the group LASSO density as

$$\boldsymbol{\theta}_j|\lambda \overset{ind}{\sim} \lambda^{d_n} \exp\left(-\lambda \|\boldsymbol{\theta}_j\|_2\right). \tag{4.4}$$

In recent years, the spike-and-slab LASSO (ssLASSO) was developed. It helps integrating the LASSO and Bayesian spike-and-slab variable selection. We extend the idea to the spike-and-slab Quantile Group LASSO (ssQGLASSO) to perform variable selection robustly on the varying coefficient model.

Consider the following spike-and-slab priors:

$$\begin{aligned} \boldsymbol{\theta}_j|\delta_j, \lambda_0, \lambda_1 &\overset{ind}{\sim} (1-\delta_j)\Pi(\boldsymbol{\theta}_j|\lambda_0) + \delta_j \Pi(\boldsymbol{\theta}_j|\lambda_1), \\ \delta_j|\psi &\overset{ind}{\sim} \text{Ber}(\psi) = \psi^{\delta_j}(1-\psi)^{1-\delta_j}, \\ \psi &\sim U(0,1), \end{aligned} \tag{4.5}$$

where the spike component $\Pi(\boldsymbol{\theta}_j|\lambda_0)$ and slab component $\Pi(\boldsymbol{\theta}_j|\lambda_1)$ are both defined using the group LASSO density in (4.4) with scale parameters $\lambda_0 > \lambda_1$ ensuring that the spike is heavily concentrated about zero. Independent Bernoulli priors were assigned to δ_j, with the parameter ψ having a uniform prior.

We assume independent multivariate normal priors on the regression coefficients for the intercept and clinical factors, $\boldsymbol{\xi}_l$ with 0 vector and Σ_l being the mean and covariance matrix. The scale parameter $\sigma > 0$ is assumed to have inverse-gamma prior distribution with $a > 0$

and $b > 0$ being constants. They are displayed as follows:

$$\xi_l \overset{ind}{\sim} \mathrm{N}_{d_n}(0, \Sigma_l),$$
$$\sigma \sim \mathrm{IG}(a, b) = \sigma^{-a-1}\exp\left(-\frac{b}{\sigma}\right).$$

As for the random effect part, we propose the following priors:

$$\boldsymbol{\gamma}_i|\phi \sim N_{p_n}(0, \phi^2\boldsymbol{I})),$$
$$\phi^2 \sim \mathrm{IG}(e, f),$$

where $\boldsymbol{\gamma}_i = (\gamma_{i,1}, ..., \gamma_{i,p_n})^\top$ is a p_n-dimensional vector and random effect $\boldsymbol{\gamma}_i$ follows multivariate normal distribution with mean 0 and covariance matrix $\phi^2\boldsymbol{I}$. ϕ^2 has an inverse-gamma prior with $e > 0$ and $f > 0$ being constants. For computational convenience, we choose $a = b = e = f = 1$.

4.2.5 Implementation via the EM Algorithm

We treat $(n \times K)$-dimensional vector $\boldsymbol{v} = (v_1(t_{11}), \ldots, v_n(t_{nK}))^\top$ and indicator variables $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_p)^\top$ as missing values in the proposed EM algorithm. Let $\Lambda_j = (1-\delta_j)\lambda_0 + \delta_j\lambda_1$,

which is also viewed as missing due to δ_j, the log joint posterior density is given as:

$$
\begin{aligned}
Q &= \log p(\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \sigma, \boldsymbol{\delta}, \boldsymbol{v}, \psi, \phi^2 | \boldsymbol{Y}) \\
&= \log p(\boldsymbol{Y} | \boldsymbol{\theta}, \boldsymbol{\gamma}, \sigma, \boldsymbol{v}, \phi^2) + \sum_{l=1}^{q} \log p(\xi_l) + \sum_{j=1}^{p} \log p(\theta_j | \Lambda_j) + \log p(\sigma) + \sum_{i=1}^{n} \sum_{k=1}^{K} p(v_i(t_{ik})) \\
&+ \sum_{j=1}^{p} \log p(\delta_j | \psi) + \log p(\psi) + \sum_{i=1}^{n} \sum_{g=1}^{p_n} p(\gamma_{ig}) + \log p(\phi^2) \\
&\propto - \sum_{i=1}^{n} \sum_{k=1}^{K} \left(\frac{(y_i(t_{ik}) - \mu_i(t_{ik}))^2}{2\zeta_2^2 \sigma} v_i(t_{ik})^{-1} - \frac{\zeta_1 (y_i(t_{ik}) - \mu_i(t_{ik}))}{\zeta_2^2 \sigma} + \frac{\zeta_1^2}{2\zeta_2^2 \sigma} v_i(t_{ik}) \right) \\
&- \sum_{i=1}^{n} \sum_{k=1}^{K} \frac{3}{2} \log(\sigma) - \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{K} \log(v_i(t_{ik})) - \sum_{l=1}^{q} \frac{\boldsymbol{\xi}_l^\top \Sigma_l^{-1} \boldsymbol{\xi}_l}{2} - \sum_{j=1}^{p} \Lambda_j \|\boldsymbol{\theta}_j\|_2 - \frac{b}{\sigma} \\
&- (a+1) \log(\sigma) - \frac{1}{\sigma} \sum_{i=1}^{n} \sum_{k=1}^{K} v_i(t_{ik}) + \sum_{j=1}^{p} [\delta_j \log \psi + (1 - \delta_j) \log(1 - \psi)] \\
&- \frac{n \times p_n}{2} \log(\phi^2) - \sum_{i=1}^{n} \sum_{g=1}^{p_n} \frac{(\gamma_{ig})^2}{2\phi^2} - \frac{f}{\phi^2} - (e+1) \log(\phi^2).
\end{aligned}
\tag{4.6}
$$

The EM algorithm estimates $(\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \sigma, \psi, \phi^2)$ through averaging the missing values across the parameters' respective posterior distributions with terms involving $v_i(t_{ik})$ $(i = 1, ..., n, k = 1, \ldots, K)$ and δ_j being treated as missing. During E-step, we calculate the conditional posterior expectations of $v_i(t_{ik})^{-1}$, $v_i(t_{ik})$, δ_j and Λ_j^{-1}. Recall $\mu_i(t_{ik}) = \boldsymbol{A}_i(t_{ik})^\top \boldsymbol{\xi} + \boldsymbol{B}_i(t_{ik})^\top \boldsymbol{\theta} + \boldsymbol{w}_i(t_{ik})^\top \boldsymbol{\gamma}_i$, all the terms that involve $v_i(t_{ik})$ in (4.6) can be combined together and written as:

$$
v_i(t_{ik})^{-1/2} \exp\left(\frac{1}{2} (h_i(t_{ik})^2 v_i(t_{ik})^{-1} + \tilde{h}^2 v_i(t_{ik})) \right), \tag{4.7}
$$

where $h_i(t_{ik})^2 = \frac{(y_i(t_{ik}) - \mu_i(t_{ik}))^2}{\zeta_2^2 \sigma}$ and $\tilde{h}^2 = \frac{2}{\sigma} + \frac{\zeta_1^2}{\zeta_2^2 \sigma}$, which is the kernel of a generalized inverse Gaussian (GIG) distribution denoted as $v_i(t_{ik}) | y_i(t_{ik}), \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \sigma \sim GIG\left(\frac{1}{2}, h_i(t_{ik}), \tilde{h}\right)$. For a random variable Ω following the generalized inverse Gaussian (GIG) distribution

$GIG(\omega, o_1, o_2)$, the probability density function is given as

$$f(\Omega|\omega, o_1, o_2) = \frac{(o_2/o_1)^{\omega}}{2K_{\omega}(o_1 o_2)}\Omega^{\omega-1}\exp\left(-\frac{1}{2}(o_1^2\Omega^{-1} + o_2^2\Omega)\right),$$

where $\Omega > 0$, $o_1 > 0$, $o_2 > 0$ and ω is a real number parameter used in function $K(\cdot)$, the modified Bessel function of the third kind (Karlis, 2002)[95]. The moments of Ω are given by

$$E(\Omega^{\kappa}) = \left(\frac{o_1}{o_2}\right)^{\kappa}\frac{K_{\omega+\kappa}(o_1 o_2)}{K_{\omega}(o_1 o_2)}, \kappa \in \mathrm{R}.$$

Based on the information in (4.7), we choose $\omega = \frac{1}{2}$ in the above pdf which gives the conditional posterior expectations of $v_i(t_{ik})^{-1}$ and $v_i(t_{ik})$ at d-th iteration:

$$\begin{aligned} vn_i(t_{ik})^{(d)} &= E\left((v_i(t_{ik})^{-1})^{(d)}\right) = \left(\frac{h_1(it)^{(d)}}{\tilde{h}^{(d)}}\right)^{-1}\frac{K_{-1/2}(h_i(it)^{(d)}\tilde{h}^{(d)})}{K_{1/2}(h_i(it)^{(d)}\tilde{h}^{(d)})}, \\ vp_i(t_{ik})^{(d)} &= E\left(v_i(t_{ik})^{(d)}\right) = \left(\frac{h_1^{(d)}(it)}{\tilde{h}^{(d)}}\right)\frac{K_{3/2}(h_i(it)^{(d)}\tilde{h}^{(d)})}{K_{1/2}(h_i(it)^{(d)}\tilde{h}^{(d)})}. \end{aligned} \tag{4.8}$$

The conditional posterior expectation of δ_j at the d-th iteration can be derived as:

$$\begin{aligned} \eta_j^{(d)} &= p(\delta_j = 1|\boldsymbol{\theta}_j^{(d)}, \psi^{(d)}, \boldsymbol{y}) \\ &= \frac{p(\boldsymbol{\theta}_j^{(d)}|\delta_j = 1, \lambda_1)p(\delta_j = 1|\psi^{(d)})}{p(\boldsymbol{\theta}_j^{(d)}|\delta_j = 0, \lambda_0)p(\delta_j = 0|\psi^{(d)}) + p(\boldsymbol{\theta}_j^{(d)}|\delta_j = 1, \lambda_1)p(\delta_j = 1|\psi^{(d)})} \\ &= \frac{\Pi(\boldsymbol{\theta}_j^{(d)}|\lambda_1)\psi^{(d)}}{\Pi(\boldsymbol{\theta}_j^{(d)}|\lambda_0)(1-\psi^{(d)}) + \Pi(\boldsymbol{\theta}_j^{(d)}|\lambda_1)\psi^{(d)}}, \end{aligned} \tag{4.9}$$

where $\Pi(\boldsymbol{\theta}_j^{(d)}|\lambda_0) = \lambda_0^{d_n}\exp\left(-\lambda_0\|\boldsymbol{\theta}_j\|_2\right)$ and $\Pi(\boldsymbol{\theta}_j^{(d)}|\lambda_1) = \lambda_1^{d_n}\exp\left(-\lambda_1\|\boldsymbol{\theta}_j\|_2\right)$. Then, the conditional expectation of $(\Lambda_j^{-1})^{(d)}$ can be computed as:

$$\tilde{\Lambda}_j^{(d)} = E(\Lambda_j^{(d)}|\boldsymbol{\theta}_j^{(d)}) = E\left((1-\delta_j)\lambda_0 + \delta_j\lambda_1\Big|\boldsymbol{\theta}_j^{(d)}\right) = \left(1-\eta_j^{(d)}\right)\lambda_0 + \left(\eta_j^{(d)}\right)\lambda_1. \tag{4.10}$$

At the M-step, we update $\boldsymbol{\xi}^{(d+1)}$, $\boldsymbol{\theta}^{(d+1)}$, $\boldsymbol{\gamma}^{(d+1)}$, $\sigma^{(d+1)}$, $\psi^{(d+1)}$ and $(\phi^2)^{(d+1)}$ by maximizing the log joint posterior density at d-th iteration. let $\boldsymbol{\Phi} = (\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \sigma, \boldsymbol{\delta}, \boldsymbol{v}, \psi, \phi^2)$, then it is

equivalent to maximizing $Q(\boldsymbol{\Phi}|\boldsymbol{\Phi}^{(d)})$. Coefficient vector $\boldsymbol{\xi}$, $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ are updated component-wise through cyclic coordinate decent algorithm[89]. By solving $\frac{\partial Q(\boldsymbol{\Phi}|\boldsymbol{\Phi}^{(d)})}{\partial \boldsymbol{\xi}_{\tilde{l}}} = 0$, $\frac{\partial Q(\boldsymbol{\Phi}|\boldsymbol{\Phi}^{(d)})}{\partial \boldsymbol{\theta}_{\tilde{j}}} = 0$, $\frac{\partial Q(\boldsymbol{\Phi}|\boldsymbol{\Phi}^{(d)})}{\partial \gamma_{\tilde{i}\tilde{g}}} = 0$, $\frac{\partial Q(\boldsymbol{\Phi}|\boldsymbol{\Phi}^{(d)})}{\partial \sigma} = 0$, $\frac{\partial Q(\boldsymbol{\Phi}|\boldsymbol{\Phi}^{(d)})}{\partial \psi} = 0$ and $\frac{\partial Q(\boldsymbol{\Phi}|\boldsymbol{\Phi}^{(d)})}{\partial \phi^2} = 0$, where $\tilde{l} = 1, ..., q$, $\tilde{j} = 1, ..., p$, $\tilde{i} = 1, ..., n$ and $\tilde{g} = 1, ..., p_n$, we obtain the following expressions

$$
\begin{aligned}
\psi^{(d+1)} &= \frac{1}{p}\sum_{j=1}^{p}\eta_j^{(d)}, \\
\sigma^{(d+1)} &= \frac{\sum_{i=1}^{n}\sum_{k=1}^{K}\left(vn_i(t_{ik})^{(d)}\left(\boldsymbol{R}\mathbf{1}^{(d)}\right)^2 - 2\boldsymbol{R}\mathbf{1}^{(d)}\zeta_1 + vp_i(t_{ik})^{(d)}(\zeta_1^2 + 2\zeta_2^2)\right) + 2\zeta_2^2 b}{(3n + 2a + 2)\zeta_2^2}, \\
(\phi^2)^{(d+1)} &= \frac{\sum_{i=1}^{n}\sum_{g=1}^{p_n}\frac{\left(\gamma_{ig}^{(d)}\right)^2}{2} + f}{\frac{n\cdot p_n}{2} + e + 1}, \\
\gamma_{\tilde{i}\tilde{g}}^{(d+1)} &= \frac{\frac{1}{\zeta_2^2\sigma^{(d+1)}}\sum_{k=1}^{K}\left(\boldsymbol{R}\mathbf{2}_{\tilde{i}\tilde{g}}^{(d+1)}\right)w_{\tilde{i}}^{(\tilde{g})}(t_{\tilde{i}k})vn_{\tilde{i}}(t_{\tilde{i}k})^{(d)} - \frac{\zeta_1}{\zeta_2^2\sigma^{(d+1)}}\sum_{k=1}^{K}w_{\tilde{i}}^{(\tilde{g})}(t_{\tilde{i}k})}{\frac{1}{\zeta_2^2\sigma^{(d+1)}}\sum_{k=1}^{K}\left(w_{\tilde{i}}^{(\tilde{g})}(t_{\tilde{i}k})\right)^2 vn_{\tilde{i}}(t_{\tilde{i}k})^{(d)} + \frac{1}{(\phi^2)^{(d+1)}}}, \\
\boldsymbol{\xi}_{\tilde{l}}^{(d+1)} &= \boldsymbol{V}_{\tilde{l}}^{(d+1)}\sum_{i=1}^{n}\sum_{k=1}^{K}\left(vn_i(t_{ik})^{(d)}\left(\boldsymbol{R}\mathbf{3}_{\tilde{l}}^{(d+1)}\right)A_i^{(\tilde{l})}(t_{ik}) - \zeta_1 A_i^{(\tilde{l})}(t_{ik})\right), \\
\boldsymbol{\theta}_{\tilde{j}}^{(d+1)} &= \left(\frac{\|(\mathrm{diag}\left(\boldsymbol{vn}^{(d)}\right))^{\frac{1}{2}}\boldsymbol{Z}_{\tilde{j}}\boldsymbol{U}_{\tilde{j}}^{(d+1)}\|_2 - \zeta_2^2\sigma^{(d+1)}\tilde{\Lambda}_{\tilde{j}}^{(d)}}{\|(\mathrm{diag}(\boldsymbol{vn}^{(d)}))^{\frac{1}{2}}\boldsymbol{Z}_{\tilde{j}}\boldsymbol{U}_{\tilde{j}}^{(d+1)}\|_2}\right)_+ \boldsymbol{U}_{\tilde{j}}^{(d+1)},
\end{aligned}
\tag{4.11}
$$

where

$$\boldsymbol{R1}^{(d)} = y_i(t_{ik}) - \boldsymbol{A}_i(t_{ik})^\top \boldsymbol{\xi}^{(d)} - \boldsymbol{B}_i(t_{ik})^\top \boldsymbol{\theta}^{(d)} - \boldsymbol{w}_i(t_{ik})^\top \boldsymbol{\gamma}_i^{(d)},$$

$$\boldsymbol{R2}_{\tilde{i}\tilde{g}}^{(d+1)} = y_{\tilde{i}}(t_{\tilde{i}k}) - \boldsymbol{A}_{\tilde{i}}(t_{\tilde{i}k})^\top \boldsymbol{\xi}^{(d)} - \boldsymbol{B}_{\tilde{i}}(t_{\tilde{i}k})^\top \boldsymbol{\theta}^{(d)} - \sum_{g=1}^{\tilde{g}-1} w_{\tilde{i}}^{(g)}(t_{\tilde{i}k})\gamma_{\tilde{i}g}^{(d+1)} - \sum_{g=\tilde{g}+1}^{p_n} w_{\tilde{i}}^{(g)}(t_{\tilde{i}k})\gamma_{\tilde{i}g}^{(d)},$$

$$\boldsymbol{R3}_{\tilde{l}}^{(d+1)} = y_i(t_{ik}) - \sum_{l=1}^{\tilde{l}-1} A_i^{(l)}(t_{ik})\boldsymbol{\xi}_l^{(d+1)} - \sum_{l=\tilde{l}+1}^{q} A_i^{(l)}(t_{ik})\boldsymbol{\xi}_l^{(d)} - \boldsymbol{B}_i(t_{ik})^\top \boldsymbol{\theta}^{(d)} - \boldsymbol{w}_i(t_{ik})^\top \boldsymbol{\gamma}_i^{(d)},$$

$$\boldsymbol{V}_{\tilde{l}}^{(d+1)} = \left(\tilde{\boldsymbol{Z}}_{\tilde{l}}^\top \mathrm{diag}\left(\boldsymbol{vn}^{(d)}\right) \tilde{\boldsymbol{Z}}_{\tilde{l}} + \zeta_2^2 \sigma^{(d+1)} \Sigma_{\tilde{l}}^{-1}\right)^{-1},$$

$$\boldsymbol{U}_{\tilde{j}}^{(d+1)} = \left(\boldsymbol{Z}_{\tilde{j}}^\top \mathrm{diag}\left(\boldsymbol{vn}^{(d)}\right) \boldsymbol{Z}_{\tilde{j}}\right)^{-1} \sum_{i=1}^{n}\sum_{k=1}^{K} \left(vn_i(t_{ik})^{(d)} \left(\boldsymbol{R4}_i^{(\tilde{j})}(t_{ik})^{(d+1)}\right) B_i^{(\tilde{j})}(t_{ik}) - \zeta_1 B_i^{(\tilde{j})}(t_{ik})\right),$$

$$\begin{aligned}\boldsymbol{R4}_i^{(\tilde{j})}(t_{ik})^{(d+1)} = {} & y_i(t_{ik}) - \boldsymbol{A}_i(t_{ik})^\top \boldsymbol{\xi}^{(d+1)} - \sum_{j=1}^{\tilde{j}-1} \boldsymbol{B}_i^{(j)}(t_{ik})^\top \boldsymbol{\theta}_j^{(d+1)} + \sum_{j=\tilde{j}+1}^{p} \boldsymbol{B}_i^{(j)}(t_{ik})^\top \boldsymbol{\theta}_j^{(d)} \\ & - \boldsymbol{w}_i(t_{ik})^\top \boldsymbol{\gamma}_i^{(d+1)},\end{aligned}$$

in which $\mathrm{diag}\left(\boldsymbol{vn}^{(d)}\right)$ is a $(n \times K)$-dimensional diagonal matrix with diagonal elements being vector $\boldsymbol{vn}^{(d)}$, $\boldsymbol{Z}_{\tilde{j}}$ is a $(n \times K) \times d_n$-dimensional matrix with its elements corresponding to the genetic effects of the $\tilde{j}$-th group and $\tilde{\boldsymbol{Z}}_{\tilde{\boldsymbol{l}}}$ is a $(n \times K) \times d_n$-dimensional matrix with its elements corresponding to the intercept and clinical effects of the $\tilde{l}$-th group

Under the fixed tuning parameters λ_0 and λ_1, the algorithm proceeds following steps outlined in Table B.1:

Table 4.1: EM algorithm for the spike-and-slab quantile group LASSO.

Algorithm
Initialize iteration $d = 0$, quantile τ, coefficients $\boldsymbol{\xi}^{(0)}$, $\boldsymbol{\theta}^{(0)}$, $\boldsymbol{\gamma}^{(0)}$, $(\phi^2)^{(0)}$, $\psi^{(0)}$ and $\sigma^{(0)}$;
Repeat
E-step:
Compute the conditional expectation of $(v_i(t_{ik})^{-1})^{(d)}$, $v_i(t_{ik})^{(d)}$, $\delta_j^{(d)}$, $\Lambda_j^{(d)}$ via (4.8), (4.9), (4.10);
M-step:
Update $\psi^{(d+1)}$, $\sigma^{(d+1)}$, $(\phi^2)^{(d+1)}$, $\gamma_{\tilde{i}\tilde{g}}^{(d+1)}$, $\boldsymbol{\xi}_{\tilde{l}}^{(d+1)}$ and $\boldsymbol{\theta}_{\tilde{j}}^{(d+1)}$ through (4.11);
Calculate the updated log joint posterior density and check if $\lvert Q^{(d+1)} - Q^{(d)}\rvert < \varepsilon$;
$d \leftarrow d + 1$;
until Convergence ($\lvert Q^{(d+1)} - Q^{(d)}\rvert < \varepsilon$ for a small value $\varepsilon > 0$).

4.3 Simulation Study

We conduct simulation studies to access the performance of the proposed spike-and-slab quantile group LASSO (SSQVCM)with three alternative method including EM algorithm based spike-and-slab group LASSO (SSVCM) and two frequentist methods, regularized varying coefficient model with adaptive group LASSO under the quantile check loss (QVC) and least square loss (VC) from Tang et al. (2013)[42].

Simulation data are generated according to model (4.1) with sample size $n = 400$. For computational simplicity, number of time points is set to be the same. We consider $K = 5$. The dimensionality of genetic factors is $p = 1000$ after excluding the first column of 1's which represents the intercept. After applying the basis expansion, total dimension of the design matrix including the varying intercept ($(p + 1) \times d_n$) become much larger than the sample size. Under all simulation settings, the number of basis function d_n is set to 5, which enlargers the dimension of regression coefficients to 5005, including the varying intercept. The varying coefficient functions are given by $\alpha_1(t) = 2\cos(2\pi t + 1.5)$, $\beta_1(t) = -8t(1 - t)$, $\beta_2(t) = 4(1-t^3)$, $\beta_3(t) = log(16t+3)$ and $\beta_j(t) = 0$ for $j = 4, ..., p$. Without loss of generality, the clinical factor is omitted in order to bring a fair comparison among the 4 methods as the ability to analyze low dimensional clinical covariates $\boldsymbol{c}$ is not included in QVC and VC. We set $t = 1, ..., 5$, representing the time of post-treatment visits. This allows us to explore the genetic effects change across multiple stages during disease recovery. The random effects are assumed to consist of both random intercept and random slope. To match the correlation strength of CAMP data, we assume the 2 by 2 covariance matrix of random effects $\phi^2\boldsymbol{I}$ to follow a variance component (VC) structure with diagonal elements being 4. We simulate random error ϵ_i in model (4.1) based on five probability distributions: $N(\mu, 1)$(Error 1), $t(2)$ with mean μ(Error 2), LogNormal(μ,1) (Error 3), $80\%N(\mu,1) + 20\%N(\mu, 3)$ (Error 4) and Laplace(μ,1) (Error 5). Except for the standard Normal distribution, all other errors have long-tailed distributions. The location parameter μ has been adjusted so that for each error distribution, the τ-th quantile is 0. Besides conducting simulation studies using the

homogeneous random error in model (4.1), we have also examined a heterogeneous model:

$$y_i(t_{ik}) = c_i^{(1)}(t_{ik})\alpha_{1,\tau}(t_{ik}) + \sum_{j=1}^{p} x_i^{(j)}(t_{ik})\beta_{j,\tau}(t_{ik}) + \boldsymbol{w}_i(t_{ik})^\top \boldsymbol{\gamma}_{i,\tau} + (1 + x_i^{(2)}(t_{ik}))\epsilon_{i,\tau}. \quad (4.12)$$

The error term depend on $x^{(2)}$ in this case.

In the first scenario, genetic factors are gene expressions simulated from a multivariate normal distribution with auto-regression (AR) structure where the ith and jth genes have correlation $\rho^{|i-j|}$ with the parameter ρ set to 0.5. In the second scenario, we simulate the SNP data using a pairwise linkage disequilibrium (LD) structure. Let q_A and q_B denote the minor allele frequencies (MAFs) for the two risk alleles A and B from two adjacent SNPs, respectively. The frequencies of the four haplotypes can be derived as $p_{AB} = q_A q_B + \text{LD}, p_{ab} = (1 - q_A)(1 - q_B) + \text{LD}, p_{Ab} = q_A(1 - q_B) - \text{LD}$, and $p_{aB} = (1 - q_A)q_B - \text{LD}$. By assuming Hardy-Weinberg equilibrium, we simulate the SNP genotypes AA, Aa and aa at locus 1 from a multinomial distribution with frequencies $q_A^2, 2q_A(1 - q_A)$ and $(1 - q_A)^2$. Then the genotypes for SNP at locus 2 can be generated based on the conditional genotype probability matrix[56]. If the MAFs are 0.3 and pairwise correlation is set to 0.3, we can get $\text{LD} = 0.3\sqrt{q_A(1 - q_A)q_B(1 - q_B)}$. Lastly, in scenario three, a more realistic method is adopted to generate the SNP data. 1000 consecutive SNPs were extracted from the case study. 400 subjects were randomly selected from the real data to generate repeated measurement responses of 5 time points.

Under each case scenario, we compare all methods under 3 quantile levels (30%,50% and 70%) and 5 error distributions. For identification performance evaluation, number of true positives (TP), false positives (FP), F_1 score (F1) and Matthews correlation coefficient (MCC) have been adopted and evaluated. As for the estimation accuracy of the varying coefficients, we compare the integrated mean squared error (IMSE). Let $\hat{\beta}_j(\tilde{t})$ denote the estimation of true coefficient $\beta_j(\tilde{t})$ with $(\tilde{t}_1, ..., \tilde{t}_{(n\times K)})$ being the grid of points equally spaced on the interval $[0, 1]$. It can be calculated as $\text{IMSE}(\hat{\beta}_j(\tilde{t})) = \frac{1}{n\times K}\sum_{grid=1}^{n\times K}(\hat{\beta}_j(\tilde{t}_{grid})) - \beta_j(\tilde{t}_{grid}))^2$. Under our simulation setting, $\beta_j(\tilde{t}) = 0$ when $j > 3$. The overall estimation accuracy is measured by the total integrated mean squared error (TIMSE), which is the sum of all the

IMSE's of estimated varying coefficients. Additional simulations are included at the end of this section.

Identification and estimation results under scenario 1 are displayed in Table 4.2. Simulation data are generated using homogeneous model under AR-1 correlation with data size $(n, p, K) = (400, 1000, 5)$. Four different varying coefficient variable selection methods are compared across quantile levels $\tau = 30\%$, 50% and 70% as well as 5 error distributions. From the identification aspect, We can observe that the proposed SSQVCM outperforms all the alternatives, especially when the error distributions are heavy-tailed. For instance, when quantile level $\tau = 30\%$ and error distribution follows $t(2)$, SSQVCM identifies more important signals than other methods with a TP of 2.55 (sd0.51). It also has a close-to-zero value of FP, 0.28 (sd0.45). The three alternative methods including the SSVCM result in comparable number of TPs. However, non robust methods SSVCM and VC have much higher FPs of 3.57 (sd6.57) and 7.00 (sd2.00). The other robust method, QVC, has a lower FP of 1.24 (sd7.11). However, it remains to be dominated by SSQVCM and appears unstable, as indicated by its high standard deviation. As for the F1 score and MCC that give more comprehensive measures of a model's identification performance, the proposed SSQVCM has also demonstrated the best results. In particular, SSQVCM outperforms all alternatives with an F1 score of 0.91 (sd0.08) and a MCC of 0.91 (sd0.07). In terms of the estimation performance, SSQVCM result in a TIMSE value of 1.71 (sd0.53) which is the smallest among the measures of all methods under comparison. The superiority of SSQVCM is further established under the remaining skewed model errors (Error3–5) across all quantile levels, as shown in Table 4.2. Even under light-tailed $N(0, 1)$ error, where all methods show improved performance, SSQVCM still maintains a solid lead over the other methods. Under the setting of $\tau = 30\%$ and Error 1, the F1 score and MCC of SSQVCM are both 0.94 (sd0.08), with a TIMSE value of 1.33 (sd0.49), leading to a dominant performance in both identification and estimation accuracy. From previous studies we know that spike-and-slab robust method is competitive under $N(0, 1)$ error, but is still outperformed by non-robust method like ssLASSO. This phenomenon indicates that the varying-coefficient setting is more unstable than the fixed-coefficient setting, and robust methods have the ability to stabilize

the estimation output of the model.

Apart from assessing the identification and estimation accuracy through presenting the mean and standard deviation of measures in Table 4.2, we have also illustrated the boxplots of the F1 score, MCC, and estimation error across 100 replications in Figure 4.2, 4.3 and 4.4, respectively. These figures provide a visual representation of the data in Table 4.2 from different perspectives.

Figure 4.2 shows the box-plot of F1 scores over 100 replications, covering the setting of all 5 model errors and 3 quantile levels. Under the $N(0, 1)$ error, all methods have better performances, as shown by their box-plots having slightly higher means than those under skewed model errors (Error3–5). As skewness is introduced into model errors in the next 4 columns, SSQVCM consistently outperforms the non-robust SSVCM, VC and robust QVC with most of its F1 scores centering around 0.9. It also establishes its stability across various scenarios, as shown by a small interquartile range (IQR). Among the 4 methods, VC has the largest IQR and the smallest F1 value, suggesting that both mixed model method and robust method contribute to the stability and identification accuracy of variable selection methods. This further justifies the urgent necessity for developing the robust SSQVCM. Besides the F1 score, MCC is also commonly used to evaluate identification performance, often being more effective when handling imbalanced datasets. Unlike the F1 score, which can be biased towards the majority class, MCC gives a more accurate reflection of model performance across all classes. Evaluating using MCC is suitable for simulations with the sparsity assumption, e.g., only 3 out of a total of 1000 gene expressions are selected to be non-zero. The comparison between Figures 4.3 and 4.2 leads to similar conclusions regarding the superiority of the proposed SSQVCM in terms of identification.

Figure 4.4 graphically displays the estimation results of the 4 method under comparison by showing the box-plots of TIMSE across 100 replications. We observe that SSQVCM demonstrates superior prediction performance across all error types indicated by the lower box-plots with small IQRs, despite a few estimation results that barely fall into the outlier category. Meanwhile, SSVCM shows the second-best performance in TIMSE, indicating the importance of estimating random effects during model fitting. However, the lack of

robustness causes it to be outperformed by SSQVCM in handling data contamination.

In the appendix, we have included results from the rest two scenarios (LD structure and real data correlation) under homogeneous errors and all three scenarios under heterogeneous errors. Table C.7 and C.8 shows the identification and estimation summaries of all 4 methods with $i.i.d$ homogeneous model errors under LD structure and real data correlation structure, respectively. Patterns similar to those in Table 4.2 can be observe, which further justifies the superiority of the proposed SSQVCM. Methods are also tested using heterogeneous model (4.12). Table C.9-C.11 in the appendix provide numeric measures of identification and estimation performances of the 4 methods under all three scenarios (AR-1, LD structure and real data correlation). The observation of the results shows that although the accuracy of all 4 methods decreases comparing to those under heterogeneous errors, SSQVCM remains to have higher F1 score, MCC value and lower TIMSE than the alternatives across Error1-5. Moreover, due to the robustness property of the proposed SSQVCM, it suffers less performance drops than non-robust method like SSVCM. For example, in Table C.7 under LD structure, quantile level $\tau = 0.3$ and Error 2, SSQVCM and SSVCM has TIMSE values of 0.86 (sd0.26) and 1.52(0.77), respectively. As the model error become heterogeneous in Table C.10, SSQVCM results in a TIMSE value of 1.38 (sd0.38) indicating a 60% increase. However, non-robust SSVCM obtains a value of 3.23 (sd5.46) under the same setting, which reflects a 113% increase in TIMSE.

Table 4.2: Evaluation of homogeneous model under AR-1 correlation, $(n, p) = (400, 1000)$ with 100 replicates.

$n = 400$	$p = 1000$		TP	FP	F1	MCC	TIMSE
$\tau = 0.3$	Error1	ssQLMM	2.83(0.38)	0.33(0.66)	0.94(0.08)	0.94(0.08)	1.33(0.49)
		SSVCM	2.56(0.50)	3.00(1.70)	0.69(0.15)	0.71(0.14)	2.24(1.03)
		QVC	2.62(0.49)	1.20(4.77)	0.83(0.12)	0.84(0.11)	9.20(2.72)
		VC	2.64(0.48)	5.02(1.36)	0.64(0.20)	0.67(0.17)	3.33(1.54)
	Error2	SSQVCM	2.55(0.51)	0.28(0.45)	0.91(0.08)	0.91(0.07)	1.71(0.53)
		SSVCM	2.47(0.50)	3.57(6.57)	0.70(0.17)	0.72(0.15)	3.81(7.86)
		QVC	2.47(0.54)	1.24(7.11)	0.82(0.14)	0.83(0.13)	9.34(2.47)
		VC	2.47(0.54)	7.00(2.00)	0.56(0.20)	0.61(0.17)	4.95(2.68)
	Error3	SSQVCM	2.73(0.45)	0.43(0.73)	0.92(0.10)	0.92(0.09)	1.84(0.75)
		SSVCM	2.52(0.50)	2.98(1.56)	0.68(0.12)	0.70(0.12)	3.01(1.33)
		QVC	2.62(0.49)	1.26(4.03)	0.84(0.13)	0.84(0.12)	7.35(2.39)
		VC	2.56(0.50)	5.42(1.84)	0.58(0.13)	0.62(0.11)	4.71(1.58)
	Error4	SSQVCM	2.70(0.47)	0.50(0.73)	0.90(0.08)	0.91(0.08)	1.61(0.55)
		SSVCM	2.46(0.50)	2.68(1.49)	0.69(0.12)	0.71(0.11)	2.41(0.96)
		QVC	2.40(0.49)	0.82(4.05)	0.84(0.12)	0.84(0.12)	9.72(2.29)
		VC	2.42(0.50)	4.30(1.14)	0.63(0.18)	0.66(0.16)	3.70(1.70)
	Error5	SSQVCM	2.79(0.41)	0.33(0.64)	0.94(0.08)	0.94(0.08)	1.53(0.57)
		SSVCM	2.48(0.51)	2.75(1.57)	0.69(0.11)	0.71(0.10)	2.31(0.92)
		QVC	2.59(0.50)	1.86(5.89)	0.79(0.15)	0.81(0.14)	9.16(2.18)
		VC	2.64(0.49)	6.98(2.35)	0.56(0.17)	0.61(0.14)	4.03(1.59)
$\tau = 0.5$	Error1	SSQVCM	2.83(0.38)	0.33(0.71)	0.94(0.08)	0.95(0.07)	0.95(0.53)
		SSVCM	2.58(0.50)	3.32(1.94)	0.67(0.14)	0.70(0.13)	1.94(1.03)
		QVC	2.68(0.47)	1.46(5.54)	0.82(0.13)	0.83(0.12)	1.52(0.85)
		VC	2.62(0.49)	5.88(1.58)	0.59(0.17)	0.64(0.14)	3.14(1.61)
	Error2	SSQVCM	2.47(0.51)	0.40(0.62)	0.88(0.08)	0.89(0.08)	1.63(0.62)
		SSVCM	2.30(0.46)	3.36(2.56)	0.64(0.13)	0.66(0.11)	3.03(2.55)
		QVC	2.36(0.56)	1.44(4.23)	0.78(0.13)	0.79(0.13)	2.20(0.99)
		VC	2.42(0.64)	5.66(1.86)	0.56(0.15)	0.61(0.13)	4.16(2.01)
	Error3	SSQVCM	2.70(0.47)	0.53(0.86)	0.90(0.08)	0.91(0.08)	1.44(0.58)
		SSVCM	2.54(0.50)	2.82(1.69)	0.70(0.12)	0.72(0.11)	2.43(0.99)
		QVC	2.56(0.50)	1.20(4.37)	0.83(0.13)	0.83(0.12)	2.04(1.05)
		VC	2.58(0.50)	6.64(1.43)	0.55(0.15)	0.60(0.13)	4.05(1.62)
	Error4	SSQVCM	2.67(0.48)	0.23(0.50)	0.93(0.08)	0.93(0.07)	1.30(0.62)
		SSVCM	2.48(0.50)	2.68(1.70)	0.70(0.12)	0.72(0.12)	2.06(0.89)
		QVC	2.64(0.48)	1.70(5.89)	0.80(0.14)	0.81(0.13)	1.81(0.88)
		VC	2.60(0.49)	6.14(1.79)	0.59(0.18)	0.63(0.15)	3.57(1.87)
	Error5	SSQVCM	2.93(0.27)	0.37(0.69)	0.95(0.07)	0.95(0.07)	1.15(0.40)
		SSVCM	2.60(0.50)	2.62(1.80)	0.72(0.14)	0.74(0.13)	1.86(0.83)
		QVC	2.74(0.44)	1.38(4.98)	0.84(0.14)	0.85(0.13)	1.71(0.84)
		VC	2.68(0.47)	5.98(1.79)	0.60(0.18)	0.64(0.15)	3.36(1.81)
$\tau = 0.7$	Error1	SSQVCM	2.67(0.48)	0.17(0.38)	0.93(0.07)	0.94(0.06)	1.43(0.55)
		SSVCM	2.48(0.50)	2.62(1.40)	0.70(0.11)	0.72(0.10)	2.08(0.88)
		QVC	2.46(0.50)	0.78(4.84)	0.85(0.11)	0.85(0.11)	9.07(2.14)
		VC	2.52(0.50)	5.50(1.00)	0.59(0.15)	0.63(0.13)	3.49(1.44)
	Error2	SSQVCM	2.67(0.48)	0.43(0.73)	0.91(0.08)	0.91(0.08)	1.51(0.57)
		SSVCM	2.48(0.50)	3.12(2.00)	0.67(0.13)	0.70(0.12)	2.44(1.01)
		QVC	2.48(0.50)	1.10(4.75)	0.82(0.11)	0.83(0.10)	9.38(2.26)
		VC	2.54(0.50)	5.06(1.22)	0.61(0.16)	0.65(0.14)	3.82(1.91)
	Error3	SSQVCM	2.67(0.48)	0.40(0.67)	0.91(0.08)	0.92(0.08)	1.39(0.50)
		SSVCM	2.42(0.50)	3.00(1.64)	0.67(0.13)	0.69(0.12)	2.14(0.85)
		QVC	2.48(0.50)	1.48(4.72)	0.80(0.14)	0.81(0.12)	12.54(3.17)
		VC	2.50(0.51)	5.06(2.06)	0.61(0.17)	0.65(0.15)	3.30(1.44)
	Error4	SSQVCM	2.77(0.43)	0.47(0.78)	0.92(0.09)	0.92(0.08)	1.63(0.60)
		SSVCM	2.52(0.50)	3.46(1.70)	0.65(0.12)	0.68(0.11)	2.50(0.97)
		QVC	2.50(0.51)	1.36(4.29)	0.81(0.11)	0.82(0.10)	9.87(2.67)
		VC	2.56(0.50)	6.48(1.70)	0.54(0.14)	0.59(0.11)	4.25(1.37)
	Error5	SSQVCM	2.72(0.46)	0.28(0.46)	0.93(0.08)	0.93(0.08)	1.61(0.65)
		SSVCM	2.51(0.51)	3.09(1.92)	0.68(0.14)	0.70(0.13)	2.37(1.07)
		QVC	2.56(0.50)	1.27(4.73)	0.83(0.12)	0.84(0.11)	8.75(1.99)
		VC	2.60(0.50)	5.42(1.74)	0.61(0.17)	0.65(0.14)	3.77(1.55)

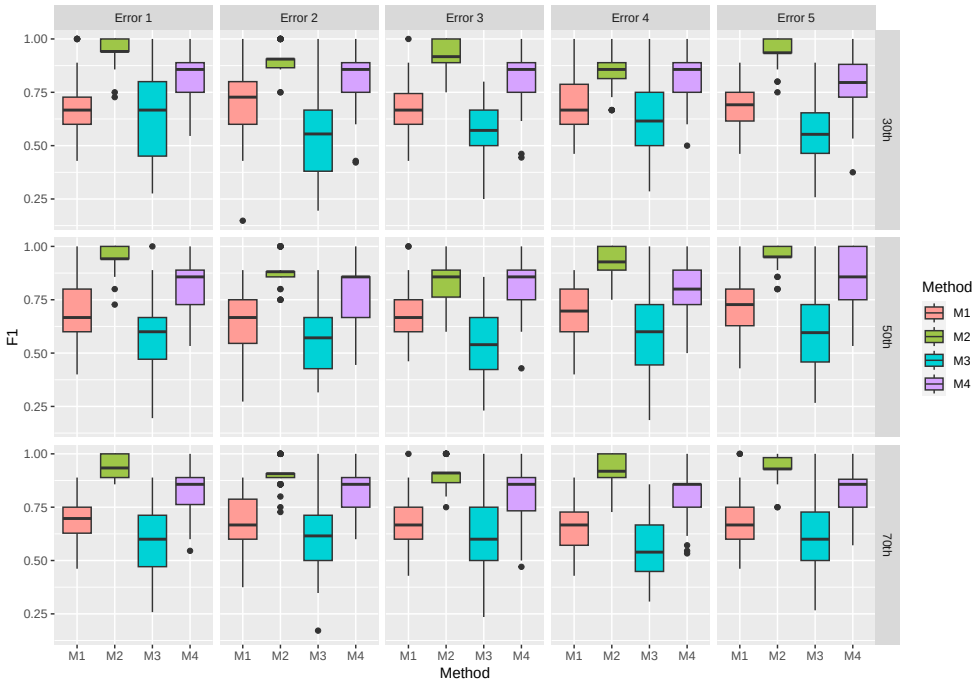

Figure 4.2: F1 score under AR-1 correlation, homogeneous error and (n, p) = (400, 1000). Methods M1-M4 correspond to SSVCM, SSQVCM, VC and QVC, respectively.

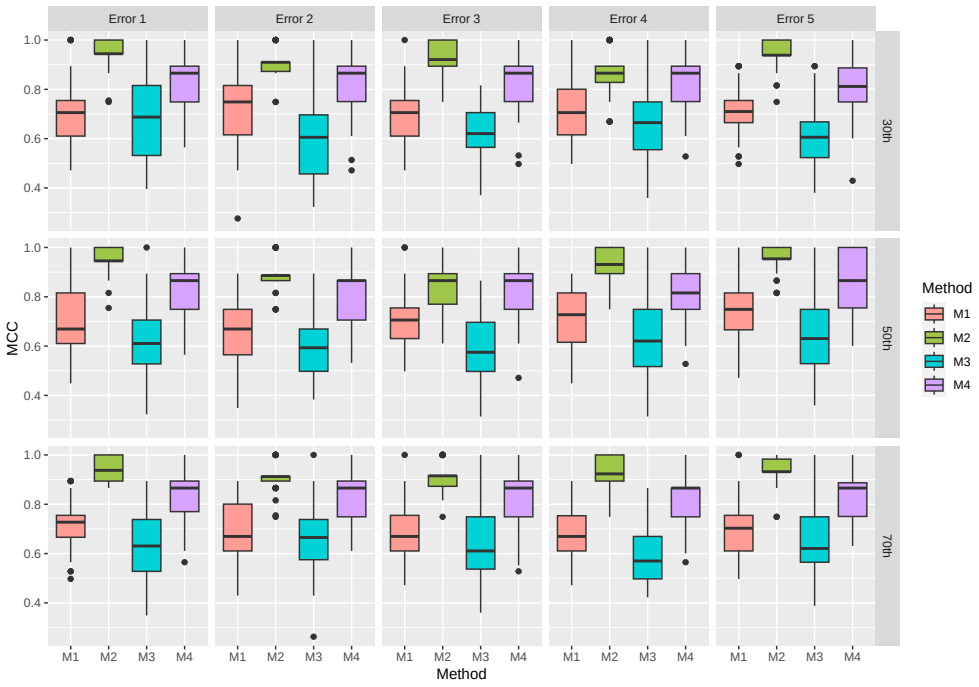

Figure 4.3: MCC score under AR-1 correlation, homogeneous error and (n, p) = (400, 1000). Methods M1-M4 correspond to SSVCM, SSQVCM, VC and QVC, respectively.

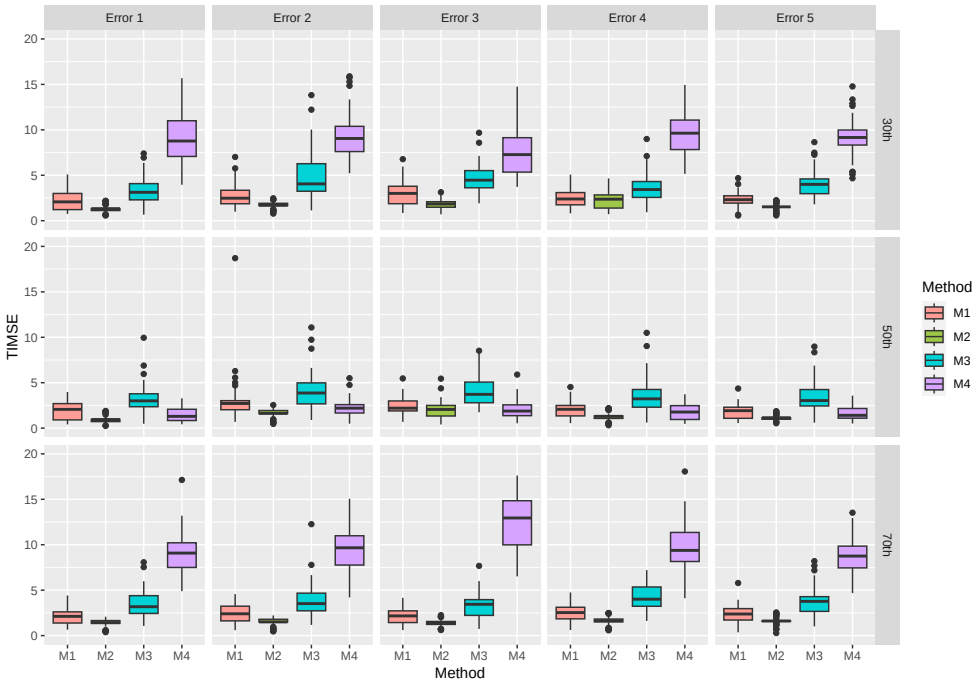

Figure 4.4: TIMSE under AR-1 correlation, homogeneous error and (n, p) = (400, 1000). Methods M1-M4 correspond to SSVCM, SSQVCM, VC and QVC, respectively.

Model misspecification: To further examine the influence of model misspecification, additional simulations have been performed under more complicated settings. First, we consider generating random effect coefficients under variance component structure with different diagonal elements. In particular, We select $(20, 2)$ and $(40, 0.4)$ as two pairs of candidates to represent the relationship between the coefficients of the random intercept and random slope, with ratios of 10 to 1 and 100 to 1, respectively. Under homogeneous model error and LD structure setting, Table C.12 and C.13 show similar patterns as SSQVCM exceeds alternatives in terms of identification and prediction performance. However, it does affect the prediction performance. As the difference of the two components gets larger, TIMSE tend to increase for the same method. Another kind of misspecification of the model involves uneven number of measurements on each subject, which can be encountered in real life situations. In this setting, we assume that each subject is to be measured 5 times, but after the first measurement, there is a 60% chance that each subsequent scheduled time will be skipped. This results in each subject being measured different number of times as well as less number

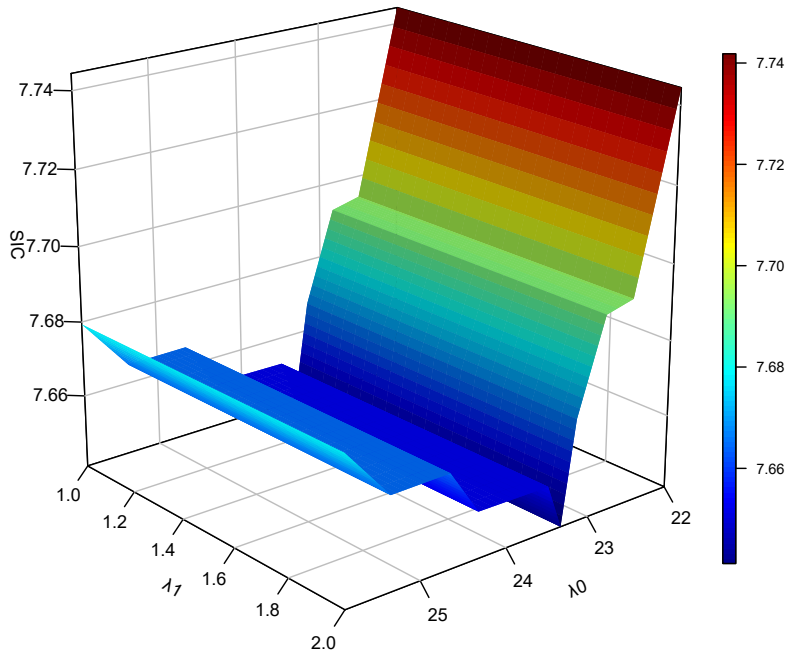

Figure 4.5: The SIC plot of SSQVCM under $\tau = 0.3$, AR-1 correlation, $(n = 400, p = 1000)$ and error 2 over 100 replications.

of total measurements. In Table C.16 with heterogeneous error and LD structure, we can clearly observe significant drop in both identification and estimation accuracy. For example, in Table C.10 with same non $i.i.d$ model error and correlation structure, the proposed SSQVCM has MCC value 0.95(sd0.06) and TIMSE value 1.38 (sd0.38) under $\tau = 0.3$ and Error 2. However, under same setting, SSQVCM yields a MCC score of 0.87 (sd0.08) and TIMSE value 2.29 (sd0.56) in Table C.16. Interestingly, with this amount of performance decrease, SSQVCM still outperforms the alternatives. Also being a robust method, QVC has a MCC value of 0.73 (sd 0.14) and a TIMSE value of 7.23 (sd2.22). Its extremely low estimation performance may be caused by the lack of random effect measuring ability.

Tuning parameter selection: Tuning parameters λ_0 and λ_1 represent the scales of the spike and slab components, respectively. They control the shrinkage level of SSVCM and robust method SSQVCM. For SSVCM, the optimal tuning are selected based on least

squares loss using cross-validation. For SSQVCM, the Schwarz Information Criterion (SIC) is used to determine the proper tuning parameters. Figure B.1 shows that the SIC surface plot reaches minimum when the spike scale λ_0 is between 10 and 11. However, the SIC value seems not affected much by the choice of the slab scale λ_1 as it increases from 1 to 2. This means we can achieve a good enough SIC value when the appropriate λ_1 is selected. The surface plot has similar pattern as the prediction error surface plot of ssLASSO in existing studies[90;91]. Under this observation, we are able to locate an appropriate pair of tuning parameters by performing grid search on more λ_0 values and less λ_1 values. Consequently, the superior performance of ssQLASSO observed in the simulation study is insensitive to the strategies adopted to locate optimal tuning parameters. As a result, the superiority of the proposed SSQVCM in the simulation studies can be established across a broad range of suitable tuning parameters.

Solution path and self-adaptivity: To better understand the impact of the spike scale λ_0 on model performance, we examine the solution path of group LASSO and SSQVCM under homogeneous model with AR-1 correlation and $(n,p)=(200, 400)$ under quantile level $\tau = 0.5$ and Error 2. One coefficient group vector $\boldsymbol{\theta}$ of size 5 is set to be nonzero, i.e., $(-0.7, -0.8, 0.6, 0.8, 0.9)$. For the proposed SSQVCM, we fixed the slab scale λ_1 to 1. The solution paths corresponding to group LASSO and SSQVCM are shown in Figure C.2 and C.1 in Appendix, respectively. For group LASSO, we observe that as λ increases, all non-zero coefficients in the group LASSO estimates in Figure C.2 converge to 0. There is also sign of over shrinkage at the optimal λ around 0.7. However, Figure C.1 shows that when spike scale λ_0 grows beyond optimal tuning around 15.3, SSQVCM consistently include all the important coefficients of the non-zero group in the model with minimal or no shrinkage. From this result, we can observe that the SSQVCM has a self-adaptive ability to capture the sparsity pattern and retain all the important coefficients even when the penalty imposed on the model is large. The using of slab component prevents non-zero coefficients from being overshrunk by a large spike scale penalty. It also helps the model to achieve a stable identification and estimation performance as the trajectories of the non-zero coefficients have little variation as the spike scale λ_0 increases. Without the slab component, variable selection

models tend to overshrunk their coefficients like group LASSO did in Figure C.2.

4.4 Real Data Analysis

Asthma is a persistent respiratory condition with the symptom being chronic inflammation of the lungs and reversible constriction of the airways, leading to breathing issues. According to the Centers for Disease Control and Prevention (CDC), it is the primary chronic disease affecting children. 7.7% of adults and 8.4% of children in the United States[57;58] are affected by asthma. Our case study is motivated by the Childhood Asthma Management Program (CAMP)[59–61]. It is a longitudinal clinical trial that explored and monitored impact of different asthma treatments (Budesonide and Nedocromil) on lung development among subjects between the ages of 5 and 12 over 4 years. Each subject received up to a total of 16 visits, including 3 pre-treatment visits with a time gap of 1-2 weeks, and 13 post-treatment visits. The first 2 post-treatment visits are 2 months apart, and the subsequent visits are 4 months apart. To minimize the influence of variations in the length of time gap and number of visits, we select the last twelve visits that were scheduled 4 months apart after treatment for our analysis. The phenotype variable of interest that were monitored is the lung growth, which is measured by the change in the forced expiratory volume in one second (FEV1). The genome-wide SNP data contains genotype information which comprises more than nine hundred thousand SNPs in total. We remove SNPs with a minor allele frequency (MAF) of less than 0.05 or displayed deviation from the Hardy-Weinberg equilibrium. After matching genotype information with the corresponding phenotypes based on the subject IDs, we result in a longitudinal dataset consisted of 438 subjects, 12 time points and 447,850 SNPs.

To further reduce the number of variables in the genotype dataset, we conducted marginal feature pre-screening which is often utilized in studies involving ultra-high dimensional datasets, such as Genome Wide Association Studies (GWAS). It helps to improve the application of regularization on datasets with a reasonably large scale[62;63]. For instance, single SNP analysis was utilized as a pre-screening method before implementing variable selection techniques in the studies of Li et al.[64], Jiang et al.[65] and Wu et al.[66] in longitudinal and

multivariate GWAS studies, respectively. For our study, marginal pre-screening is conducted using method SSVCM and SSQVCM under $\tau = 0.3$, 0.5 and 0.7. With gender being considered as a varying clinical factor, we combine the results and obtained a working dataset containing 1 clinical factor, 521 SNPs, 12 time points and 438 subjects.

Table 4.3: Identification results for real data analysis. The numbers of genes identified by different methods and their overlaps.

CAMP						
	SSQVCM(0.3)	SSQVCM(0.5)	SSQVCM(0.7)	SSVCM	QVC	VC
SSQVCM(0.3)	26	15	5	2	11	11
SSQVCM(0.5)		25	10	4	9	11
SSQVCM(0.7)			26	3	13	11
SSVCM				37	4	11
QVC					49	22
VC						48

Table 4.3 shows the number of overlapping SNPs identified by SSQVCM under $\tau = 0.3$, 0.5 and 0.7, SSVCM, QVC and VC. Among all the SNPs selected by the proposed SSQVCM method, 7 of them are not identified by other methods. Their associated genes are MIR4527HG, DCAF12, DNAJC13, MIR802, ATP8B1, PHLDB2 and PRKCE. Chen et al., 2023[128] have found that MIR4527HG is significantly expressed among patients with non-small cell lung cancer (NSCLC) and asthma is known to be closely associated with lung cancer (Santillan et al., 2003)[129]. Yan et al., 2017[130] have reported that the expression level of DCAF12 is significantly lower in normal lung tissue than in lung adenocarcinoma which is one of the most common subtype lung cancer. Che et al., 2022[131] have shown that DNAJC13 has strong positive correlation with PIK3CB, a gene closely related to the prognosis of LUAD patients. Zhang et al., 2019[132] have reported that MIR802 has the ability to restrain the growth, migration and invasion of NSCLC cells by targeting FGFR1, of which the expression levels is high in NSCLC tissues. Cooke et al., 2016[133] have found that ATP8B1 compromises lung epithelial cells, causing pulmonary fibrosis and damaging lung tissue. Ge et al., 2021 have reported that PHLDB2 level is significantly lower among lung cancer patients. It also affects the survival time of the patients. Moreover, Hwang et al., 2008[134] have shown that PRKCE causes airway hyper-responsiveness (AHR). It increases the sensitivity of the airway,

causing various symptoms of Asthma.

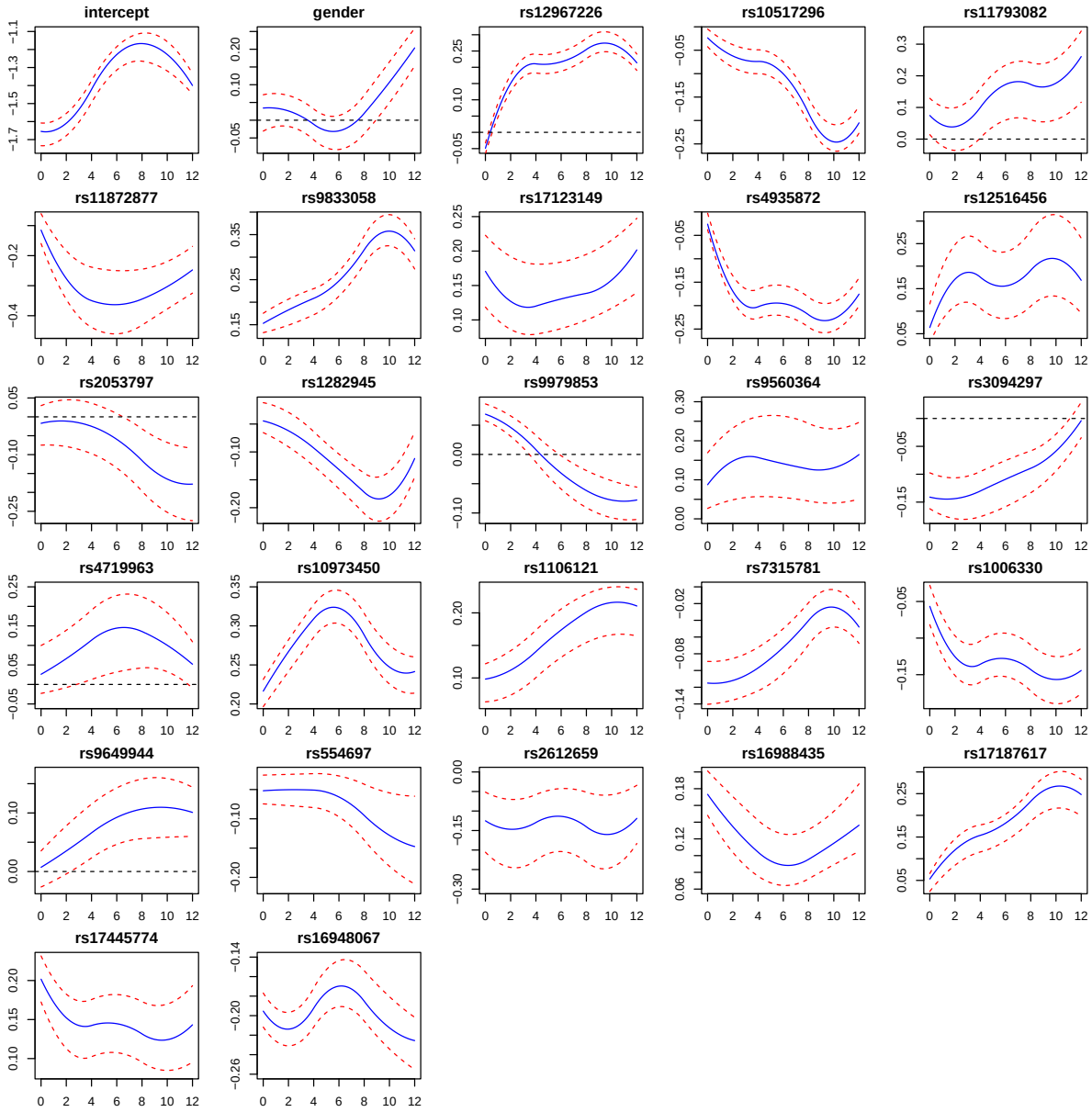

Figure 4.6: Estimated varying coefficient functions for SNPs selected by SSQVCM at $\tau = 0.5$.

Table 4.4: Information of the 25 SNPs selected by SSQVCM under $\tau = 0.5$.

SNP name	Chromosome	Chr. position	Gene	Risk allele	MAF
rs12967226	18	47611845	MIR4527HG	A	14.04%
rs10517296	4	53390964	FIP1L1	T	19.18%
rs11793082	9	34088703	DCAF12	T	5.82%
rs11872877	18	57861027	ATP8B1	C	16.32%
rs9833058	3	132493149	DNAJC13	A	15.41%
rs17123149	1	100673223	LOC124904231	A	27.28%
rs4935872	11	124205288	OR8G2P	G	16.10%
rs4719963	7	28971448	CPVL-AS2	T	43.26%
rs1106121	4	37483651	C4orf19	A	14.95%
rs1006330	11	101755870	PLS1P1	G	13.70%
rs554697	6	167794176	LINC01558	T	25.00%
rs16988435	21	31686258	SCAF4	T	23.06%
rs12516456	5	71626789	MCCC2	G	9.25%
rs2053797	2	46143753	PRKCE	A	18.95%
rs1282945	3	111916272	PHLDB2	A	9.82%
rs9979853	21	35669231	MIR802	G	43.26%
rs9560364	13	89499000	LINC01040	C	10.84%
rs3094297	3	101680416	RPL24	C	36.76%
rs10973450	9	37642237	FRMPD1	C	10.84%
rs7315781	12	41049832	CNTN1	C	42.81%
rs9649944	8	118048882	EXT1	T	43.15%
rs2612659	4	174512187	HPGD	C	7.76%
rs17187617	1	208603220	RPS26P13	T	13.13%
rs16948067	15	93463669	LOC124900612	G	8.11%

Figure 4.6 shows the estimated coefficient functions of the varying intercept and gender effect as well as the 25 estimated coefficient functions of the SNPs selected using SSQVCM under $\tau = 0.5$. Table C.6 provides detailed information about the selected 25 SNPs. The intercept has an increasing effect, representing the FEV1 increase due to the lung development and treatments taking effect through time. As for the gender effect, the decrease before visit time 6 is due to young females growing faster than young males before the age of 12. The increase in the second half indicates the boys starting to develop more lung functions along with other body functions. The SNP functions take various forms. Most of them are strictly above or below 0, showing significant effects.

To assess the prediction performance, we have examined the heritability explained by the selected SNPs by different methods. Heritability summarizes the proportion of variation in

a trait that is explainable by genetic factors. In our setting, we evaluate the heritability of FEV1 given the SNPs selected by our methods. For the proposed SSQVCM, it is calculated by

$$H(FEV1) = \frac{R(FEV1|Gender)}{R(FEV1|Gender)} - \frac{R(FEV1|Gender, SNP_1, ..., SNP_{25})}{R(FEV1|Gender)}.$$

For the residue estimator R, we consider residue sum of squares and sum of residue absolute deviation to evaluate the result from both non-robust and robust perspective, respectively. For non-robust evaluation, we have

$$R = \sum_{i=1}^{n} \sum_{k=1}^{K} \left(y_i(t_{ik}) - \tilde{y}_i(t_{ik})\right)^2.$$

The total heritability for SSQVCM at $\tau = 0.3$, 0.5, 0.7, SSVCM, VC and QVC are 56.22%, 55.64%, 56.02%, 49.06%, 52.94% and 46.92%, respectively. For robust evaluation, we have

$$R = \sum_{i=1}^{n} \sum_{k=1}^{K} \left|y_i(t_{ik}) - \tilde{y}_i(t_{ik})\right|.$$

Total heritability for all of the methods are 35.08%, 34.93%, 35.67%, 30.27%, 33.88% and 30.96%, respectively. We also obtained the corresponding heritability at each time point to estimate the time-varying heritability for all SNPs. The results are shown in Figure 4.7 and Figure 4.8 for non-robust and robust evaluation. The proposed SSQVCM outperforms the alternatives under both circumstances.

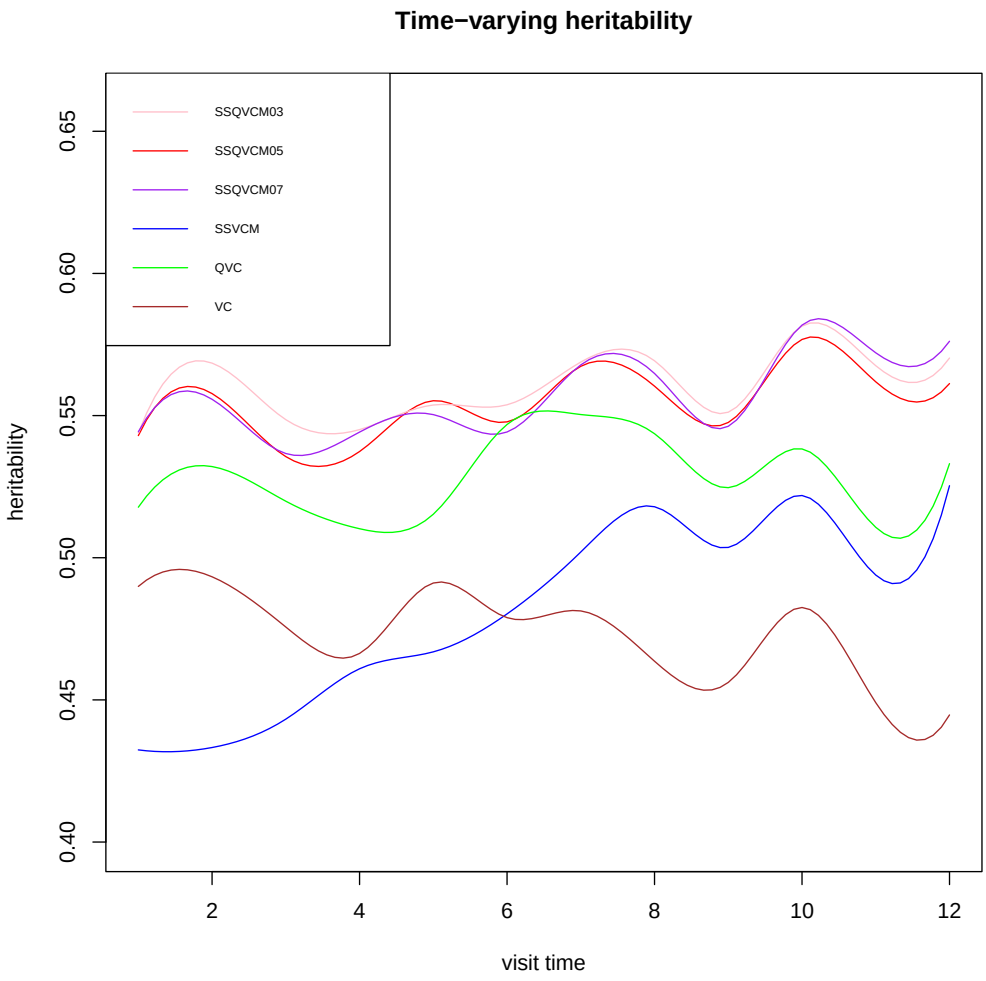

Figure 4.7: Non-robust time-varying heritability of 4 methods.

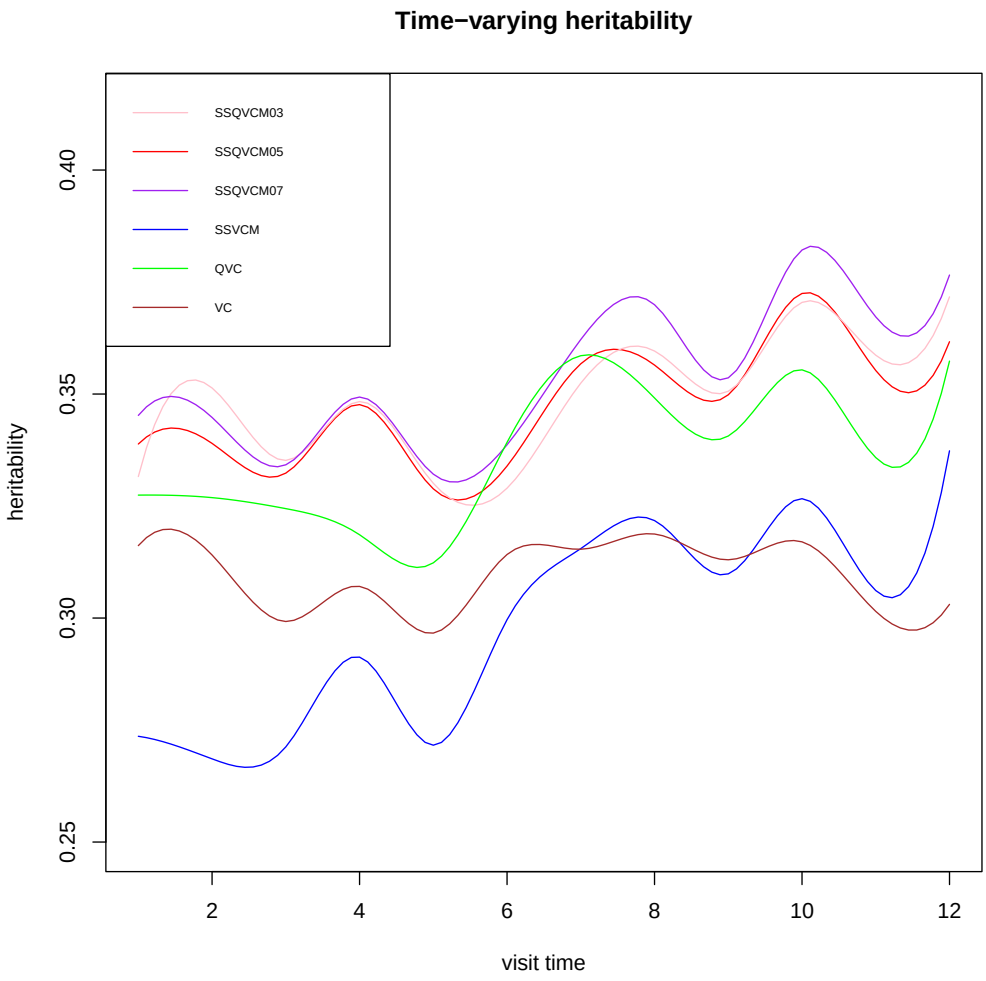

Figure 4.8: Robust time-varying heritability of 4 methods.

Bibliography

[1] Yuwen Liu and Cen Wu. *emBayes: Robust Bayesian Variable Selection via Expectation-Maximization*, 2023. URL https://CRAN.R-project.org/package=emBayes. R package version 0.1.2.

[2] Thomas M Schmitt, Gunnar B Ragnarsson, and Philip D Greenberg. T cell receptor gene therapy for cancer. *Human gene therapy*, 20(11):1240–1248, 2009.

[3] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

[4] Jianqing Fan and Runze Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, 96(456): 1348–1360, 2001.

[5] Cun-Hui Zhang. Nearly unbiased variable selection under minimax concave penalty. *The Annals of statistics*, 38(2):894–942, 2010.

[6] Ming Yuan and Yi Lin. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(1):49–67, 2006.

[7] Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, 67(2):301–320, 2005.

[8] Robert Tibshirani, Michael Saunders, Saharon Rosset, Ji Zhu, and Keith Knight. Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(1):91–108, 2005.

[9] Jian Huang, Shuangge Ma, Hongzhe Li, and Cun-Hui Zhang. The sparse laplacian shrinkage estimator for high-dimensional regression. *Annals of statistics*, 39(4):2021, 2011.

[10] Jie Ren, Yinhao Du, Shaoyu Li, Shuangge Ma, Yu Jiang, and Cen Wu. Robust network-based regularization and variable selection for high-dimensional genomic data in cancer prognosis. *Genetic epidemiology*, 43(3):276–291, 2019.

[11] Robert B O'Hara and Mikko J Sillanpää. A review of bayesian variable selection methods: what, how and which. *Bayesian analysis*, 4(1):85–117, 2009.

[12] Trevor Park and George Casella. The bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686, 2008.

[13] George Casella, Malay Ghosh, Jeff Gill, and Minjung Kyung. Penalized regression, standard errors, and bayesian lassos. 2010.

[14] Toby J Mitchell and John J Beauchamp. Bayesian variable selection in linear regression. *Journal of the american statistical association*, 83(404):1023–1032, 1988.

[15] Lynn Kuo and Bani Mallick. Variable selection for regression models. *Sankhyā: The Indian Journal of Statistics, Series B*, pages 65–81, 1998.

[16] Edward I George and Robert E McCulloch. Variable selection via gibbs sampling. *Journal of the American Statistical Association*, 88(423):881–889, 1993.

[17] Veronika Ročková and Edward I George. The spike-and-slab lasso. *Journal of the American Statistical Association*, 113(521):431–444, 2018.

[18] Qing Li, Nan Lin, and Ruibin Xi. Bayesian regularized quantile regression. 2010.

[19] Keming Yu and Rana A Moyeed. Bayesian quantile regression. *Statistics & Probability Letters*, 54(4):437–447, 2001.

[20] Philip J Brown and Jim E Griffin. Inference with normal-gamma prior distributions in regression problems. 2010.

[21] Carlos M Carvalho, Nicholas G Polson, and James G Scott. Handling sparsity via the horseshoe. In *Artificial intelligence and statistics*, pages 73–80. PMLR, 2009.

[22] Jie Ren, Fei Zhou, Xiaoxi Li, Shuangge Ma, Yu Jiang, and Cen Wu. Robust bayesian variable selection for gene–environment interactions. *Biometrics*, 79(2):684–694, 2023.

[23] Yingying Fan and Runze Li. Variable selection in linear mixed effects models. *Annals of statistics*, 40(4):2043, 2012.

[24] Scott L Zeger, Kung-Yee Liang, and Paul S Albert. Models for longitudinal data: a generalized estimating equation approach. *Biometrics*, pages 1049–1060, 1988.

[25] Annie Qu, Bruce G Lindsay, and Bing Li. Improving generalised estimating equations using quadratic inference functions. *Biometrika*, 87(4):823–836, 2000.

[26] Lan Wang, Jianhui Zhou, and Annie Qu. Penalized generalized estimating equations for high-dimensional longitudinal data analysis. *Biometrics*, 68(2):353–360, 2012.

[27] Hyunkeun Cho and Annie Qu. Model selection for correlated data with diverging number of parameters. *Statistica Sinica*, pages 901–927, 2013.

[28] Fei Zhou, Xi Lu, Jie Ren, Kun Fan, Shuangge Ma, and Cen Wu. Sparse group variable selection for gene–environment interactions in the longitudinal study. *Genetic Epidemiology*, 46(5-6):317–340, 2022.

[29] R.C. Littell, G.A. Milliken, W.W. Stroup, R.D. Wolfinger, and O. Schabenberger. *SAS for Mixed Models, Second Edition*. SAS Institute, 2007. ISBN 9781599940786. URL https://books.google.com/books?id=z9qv32OyEu4C.

[30] Youyi Fong, Håvard Rue, and Jon Wakefield. Bayesian inference for generalized linear mixed models. *Biostatistics*, 11(3):397–412, 2010.

[31] Caiyan Li and Hongzhe Li. Network-constrained regularization and variable selection for analysis of genomic data. *Bioinformatics*, 24(9):1175–1182, 2008.

[32] Cen Wu and Yuehua Cui. Boosting signals in gene-based association studies via efficient snp selection. *Briefings in bioinformatics*, 15(2):279–291, 2014.

[33] Yu Jiang, Yuan Huang, Yinhao Du, Yinjun Zhao, Jie Ren, Shuangge Ma, and Cen Wu. Identification of prognostic genes and pathways in lung adenocarcinoma using a bayesian approach. *Cancer informatics*, 16:1176935116684825, 2017.

[34] Cen Wu, Shaoyu Li, and Yuehua Cui. Genetic association studies: an information content perspective. *Current genomics*, 13(7):566–573, 2012.

[35] Cen Wu, Qingzhao Zhang, Yu Jiang, and Shuangge Ma. Robust network-based analysis of the associations between (epi) genetic measurements. *Journal of Multivariate Analysis*, 168:119–130, 2018.

[36] Jie Ren, Tao He, Ye Li, Sai Liu, Yinhao Du, Yu Jiang, and Cen Wu. Network-based regularization for high dimensional snp data in the case–control study of type 2 diabetes. *BMC genetics*, 18:1–12, 2017.

[37] Cen Wu and Yuehua Cui. A novel method for identifying nonlinear gene–environment interactions in case–control association studies. *Human genetics*, 132:1413–1425, 2013.

[38] Fei Zhou, Jie Ren, Xi Lu, Shuangge Ma, and Cen Wu. Gene–environment interaction: A variable selection perspective. *Epistasis: Methods and Protocols, Springer*, pages 191–223, 2021.

[39] Lifeng Wang, Hongzhe Li, and Jianhua Z Huang. Variable selection in nonparametric varying-coefficient models for analysis of repeated measurements. *Journal of the American Statistical Association*, 103(484):1556–1569, 2008.

[40] Cen Wu, Yu Jiang, Jie Ren, Yuehua Cui, and Shuangge Ma. Dissecting gene-environment interactions: A penalized robust approach accounting for hierarchical structures. *Statistics in medicine*, 37(3):437–456, 2018.

[41] Cen Wu, Ping-Shou Zhong, and Yuehua Cui. Additive varying-coefficient model for nonlinear gene-environment interactions. *Statistical applications in genetics and molecular biology*, 17(2):20170008, 2018.

[42] Yanlin Tang, Huixia Judy Wang, and Zhongyi Zhu. Variable selection in quantile varying coefficient models with longitudinal data. *Computational Statistics & Data Analysis*, 57(1):435–449, 2013.

[43] Jie Ren, Fei Zhou, Xiaoxi Li, Qi Chen, Hongmei Zhang, Shuangge Ma, Yu Jiang, and Cen Wu. Semiparametric bayesian variable selection for gene-environment interactions. *Statistics in medicine*, 39(5):617–638, 2020.

[44] Lan Wang, Jianhui Zhou, and Annie Qu. Penalized generalized estimating equations for high-dimensional longitudinal data analysis. *Biometrics*, 68(2):353–360, 2012. ISSN 0006-341X. doi: 10.1111/j.1541-0420.2011.01678.x. URL https://doi.org/10.1111/j.1541-0420.2011.01678.x.

[45] Hyunkeun Cho and Annie Qu. Model selection for correlated data with diverging number of parameters. *Statistica Sinica*, 23(2):901–927, 2013. doi: http://dx.doi.org/10.5705/ss.2011.058. URL http://www.jstor.org/stable/24310367.

[46] Jie Peng, Ji Zhu, Anna Bergamaschi, Wonshik Han, Dong-Young Noh, Jonathan R Pollack, and Pei Wang. Regularized multivariate regression for identifying master predictors with application to integrative genomics study of breast cancer. *The annals of applied statistics*, 4(1):53, 2010.

[47] Xingjie Shi, Qing Zhao, Jian Huang, Yang Xie, and Shuangge Ma. Deciphering the associations between gene expression and copy number alteration using a sparse double laplacian shrinkage approach. *Bioinformatics*, 31(24):3977–3983, 2015.

[48] Jie Ren, Tao He, Ye Li, Sai Liu, Yinhao Du, and Cen Wu. Network-based regularization for high dimensional SNP data in the case–control study of type 2 diabetes. *BMC Genetics*, 18(1), 2017. doi: 10.1186/s12863-017-0495-5. URL https://doi.org/10.1186/s12863-017-0495-5.

[49] Hokeun Sun and Shuang Wang. Penalized logistic regression for high-dimensional dna methylation data with case-control studies. *Bioinformatics*, 28(10):1368–1375, 2012.

[50] Wenwen Min, Juan Liu, and Shihua Zhang. Network-regularized sparse logistic regression models for clinical risk prediction and biomarker discovery. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 15(3):944–953, 2016.

[51] Xinyu Tian, Xuefeng Wang, and Jun Chen. Network-constrained group lasso for high-dimensional multinomial classification with application to cancer subtype prediction. *Cancer informatics*, 13:CIN–S17686, 2014.

[52] KUNG-YEE Liang and SCOTT L. Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22, 1986. ISSN 0006-3444. doi: 10.1093/biomet/73.1.13. URL https://doi.org/10.1093/biomet/73.1.13.

[53] Cen Wu and Shuangge Ma. A selective review of robust variable selection with applications in bioinformatics. *Briefings in Bioinformatics*, 16(5):873–883, 2015. ISSN 1467-5463. doi: 10.1093/bib/bbu046. URL https://doi.org/10.1093/bib/bbu046.

[54] Bin Zhang and Steve Horvath. A general framework for weighted gene co-expression network analysis. *Statistical applications in genetics and molecular biology*, 4(1), 2005.

[55] David R Hunter and Runze Li. Variable selection using mm algorithms. *Annals of statistics*, 33(4):1617, 2005.

[56] Yuehua Cui, Guolian Kang, Kelian Sun, Minping Qian, Roberto Romero, and Wenjiang Fu. Gene-centric genomewide association study via entropy. *Genetics*, 179(1):637–650, 2008. doi: 10.1534/genetics.107.082370. URL https://doi.org/10.1534/genetics.107.082370.

[57] Data, statistics, and surveillance: asthma surveillance data [Internet]. Atlanta (GA): Centers for Disease Control and Prevention; 2020 [cited 2020 jan 23].

[58] Omolara Jean Akinbami. The state of childhood asthma: United states, 1980-2005. (381), 2006.

[59] Childhood Asthma Management Program Research Group. The childhood asthma management program (CAMP): design, rationale, and methods. *Controlled clinical trials*, 20(1):91–120, 1999.

[60] Childhood Asthma Management Program Research Group. Long-term effects of budesonide or nedocromil in children with asthma. *New England Journal of Medicine*, 343 (15):1054–1063, 2000.

[61] Ronina A Covar, Anne L Fuhlbrigge, Paul Williams, and H William Kelly. The childhood asthma management program (camp): contributions to the understanding of therapy and the natural history of childhood asthma. *Current Respiratory Care Reports*, 1(4):243–250, 2012.

[62] Jianqing Fan and Jinchi Lv. A selective overview of variable selection in high dimensional feature space. *Statistica Sinica*, 20(1):101, 2010.

[63] Cen Wu, Fei Zhou, Jie Ren, Xiaoxi Li, Yu Jiang, and Shuangge Ma. A selective review of multi-level omics data integration using variable selection. *High-throughput*, 8(1):4, 2019.

[64] Jiahan Li, Zhong Wang, Runze Li, and Rongling Wu. Bayesian group lasso for nonparametric varying-coefficient models with application to functional genome-wide association studies. *Ann. Appl. Stat.*, 9(2):640–664, 2015. doi: 10.1214/15-AOAS808. URL https://doi.org/10.1214/15-AOAS808.

[65] Libo Jiang, Jingyuan Liu, Xuli Zhu, Meixia Ye, Lidan Sun, Xavier Lacaze, and Rongling Wu. 2HiGWAS: a unifying high-dimensional platform to infer the global

genetic architecture of trait development. *Briefings in Bioinformatics*, 16(6):905–911, 2015.

[66] Cen Wu, Yuehua Cui, and Shuangge Ma. Integrative analysis of gene-environment interactions under a multi-response partially linear varying coefficient model. *Statistics in Medicine*, 33(28):4988–4998, 2014. doi: 10.1002/sim.6287. URL https://doi.org/10.1002/sim.6287.

[67] Yong-Xiao Wang and Yun-Min Zheng. Molecular expression and functional role of canonical transient receptor potential channels in airway smooth muscle cells. *Transient Receptor Potential Channels*, pages 731–747, 2011.

[68] Nobuyuki Fukuishi, Shinya Murakami, Akane Ohno, Naoya Yamanaka, Nobuaki Matsui, Kenji Fukutsuji, Sakuo Yamada, Kouji Itoh, and Masaaki Akagi. Does β-hexosaminidase function only as a degranulation indicator in mast cells? the primary role of β-hexosaminidase in mast cell granules. *The Journal of Immunology*, 193(4): 1886–1894, 2014.

[69] Krzysztof Pałgan and Zbigniew Bartuzi. Angiogenesis in bronchial asthma. *International journal of immunopathology and pharmacology*, 28(3):415–420, 2015.

[70] Ethan Cerami, Jianjiong Gao, Ugur Dogrusoz, Benjamin E Gross, Selcuk Onur Sumer, Bülent Arman Aksoy, Anders Jacobsen, Caitlin J Byrne, Michael L Heuer, Erik Larsson, et al. The cbio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer discovery*, 2(5):401–404, 2012.

[71] Joshua D Campbell, Anton Alexandrov, Jaegil Kim, Jeremiah Wala, Alice H Berger, Chandra Sekhar Pedamallu, Sachet A Shukla, Guangwu Guo, Angela N Brooks, Bradley A Murray, et al. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nature genetics*, 48(6):607–616, 2016.

[72] Ashfaq A Marghoob, Karen Koenig, Flavia V Bittencourt, Alfred W Kopf, and Robert S Bart. Breslow thickness and clark level in melanoma: support for including

level in pathology reports and in american joint committee on cancer staging. *Cancer*, 88(3):589–595, 2000.

[73] Cen Wu and Shuangge Ma. A selective review of robust variable selection with applications in bioinformatics. *Briefings in bioinformatics*, 16(5):873–883, 2015.

[74] Andreas Alfons, Christophe Croux, and Sarah Gelper. Sparse least trimmed squares regression for analyzing high-dimensional large data sets. *The Annals of Applied Statistics*, pages 226–248, 2013.

[75] Jelena Bradic, Jianqing Fan, and Weiwei Wang. Penalized composite quasi-likelihood for ultrahigh dimensional variable selection. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 73(3):325–349, 2011.

[76] Z John Daye, Jinbo Chen, and Hongzhe Li. High-dimensional heteroscedastic regression with an application to eqtl data analysis. *Biometrics*, 68(1):316–326, 2012.

[77] Bo Peng and Lan Wang. An iterative coordinate descent algorithm for high-dimensional nonconvex penalized quantile regression. *Journal of Computational and Graphical Statistics*, 24(3):676–694, 2015.

[78] Xiaoli Gao and Jian Huang. A robust penalized method for the analysis of noisy dna copy number data. *BMC genomics*, 11(1):1–10, 2010.

[79] Huixia Judy Wang and Jianhua Hu. Identification of differential aberrations in multiple-sample array cgh studies. *Biometrics*, 67(2):353–362, 2011.

[80] Gabriela V Cohen Freue, David Kepplinger, Matías Salibián-Barrera, and Ezequiel Smucler. Robust elastic net estimators for variable selection and identification of proteomic biomarkers. *Annals of Applied Statistics*, 13(4):2065–2090, 2019.

[81] David Kepplinger and Gabriela V Cohen Freue. Robust prediction and protein selection with adaptive pense. In *Statistical Analysis of Proteomic Data: Methods and Tools*, pages 315–331. Springer, 2021.

[82] Katja Ickstadt, Martin Schäfer, and Manuela Zucknick. Toward integrative bayesian analysis in molecular biology. *Annual Review of Statistics and Its Application*, 5:141–167, 2018.

[83] James Johndrow, Paulo Orenstein, and Anirban Bhattacharya. Scalable approximate mcmc algorithms for the horseshoe prior. *Journal of Machine Learning Research*, 21 (73):1–61, 2020.

[84] Xitong Liang, Samuel Livingstone, and Jim Griffin. Adaptive random neighbourhood informed markov chain monte carlo for high-dimensional bayesian variable selection. *Statistics and Computing*, 32(5):84, 2022.

[85] Xitong Liang, Samuel Livingstone, and Jim Griffin. Adaptive mcmc for bayesian variable selection in generalised linear models and survival models. *Entropy*, 25(9): 1310, 2023.

[86] Xi Lu, Kun Fan, Jie Ren, and Cen Wu. Identifying gene–environment interactions with robust marginal bayesian variable selection. *Frontiers in Genetics*, 12:667074, 2021.

[87] Fei Zhou, Jie Ren, Shuangge Ma, and Cen Wu. The bayesian regularized quantile varying coefficient model. *Computational Statistics & Data Analysis*, page 107808, 2023.

[88] Yunju Im, Yuan Huang, Aixin Tan, and Shuangge Ma. Bayesian finite mixture of regression analysis for cancer based on histopathological imaging–environment interactions. *Biostatistics*, 24(2):425–442, 2023.

[89] Jerome Friedman, Trevor Hastie, and Rob Tibshirani. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1):1, 2010.

[90] Zaixiang Tang, Yueping Shen, Xinyan Zhang, and Nengjun Yi. The spike-and-slab lasso cox model for survival prediction and associated genes detection. *Bioinformatics*, 33(18):2799–2807, 2017.

[91] Zaixiang Tang, Yueping Shen, Xinyan Zhang, and Nengjun Yi. The spike-and-slab lasso generalized linear models for prediction and associated genes detection. *Genetics*, 205 (1):77–88, 2017.

[92] Hideo Kozumi and Genya Kobayashi. Gibbs sampling methods for bayesian quantile regression. *Journal of Statistical Computation and Simulation*, 81(11):1565–1578, 2011.

[93] Hansheng Wang, Guodong Li, and Guohua Jiang. Robust regression shrinkage and consistent variable selection through the lad-lasso. *Journal of Business & Economic Statistics*, 25(3):347–355, 2007.

[94] Hui Zou. The adaptive lasso and its oracle properties. *Journal of the American Statistical Association*, 101(476):1418–1429, 2006.

[95] Dimitris Karlis. An em type algorithm for maximum likelihood estimation of the normal–inverse gaussian distribution. *Statistics & probability letters*, 57(1):43–52, 2002.

[96] Yichao Wu and Yufeng Liu. Variable selection in quantile regression. *Statistica Sinica*, 19(2):801–817, 2009.

[97] Youjuan Li and Ji Zhu. L 1-norm quantile regression. *Journal of Computational and Graphical Statistics*, 17(1):163–185, 2008.

[98] Huixia Judy Wang, Zhongyi Zhu, and Jianhui Zhou. Quantile regression in partially linear varying coefficient models. *The Annals of Statistics*, pages 3841–3866, 2009.

[99] Ben Sherwood, Adam Maidman, Maintainer Ben Sherwood, and TRUE ByteCompile. Package 'rqpen'. *Penalized Quantile Regression. In*, 1, 2023.

[100] Jiahan Li, Zhong Wang, Runze Li, and Rongling Wu. Bayesian group lasso for non-parametric varying-coefficient models with application to functional genome-wide association studies. *The annals of applied statistics*, 9(2):640, 2015.

[101] Cen Wu, Yuehua Cui, and Shuangge Ma. Integrative analysis of gene–environment interactions under a multi-response partially linear varying coefficient model. *Statistics in medicine*, 33(28):4988–4998, 2014.

[102] Dongchang Wang, Lifei Wang, Yu Zhang, Yunxia Zhao, and Gang Chen. Hydrogen gas inhibits lung cancer progression through targeting smc3. *Biomedicine & Pharmacotherapy*, 104:788–797, 2018.

[103] Sumaria Malik, Rehan Zafar Paracha, Maryam Khalid, Maryum Nisar, Amnah Siddiqa, Zamir Hussain, Raheel Nawaz, Amjad Ali, and Jamil Ahmad. Micrornas and their target mrnas as potential biomarkers among smokers and non-smokers with lung adenocarcinoma. *IET Systems Biology*, 13(2):69–76, 2019.

[104] Christopher A Mecoli, Brittany L Adler, Qingyuan Yang, Laura K Hummers, Antony Rosen, Livia Casciola-Rosen, and Ami A Shah. Cancer in systemic sclerosis: analysis of antibodies against components of the th/to complex. *Arthritis & Rheumatology*, 73 (2):315–323, 2021.

[105] Virginia D Steen and Thomas A Medsger. Changes in causes of death in systemic sclerosis, 1972–2002. *Annals of the rheumatic diseases*, 66(7):940–944, 2007.

[106] Piyu Parth Naik. Cutaneous malignant melanoma: A review of early diagnosis and management. *World journal of oncology*, 12(1):7, 2021.

[107] Magdalena Surman, Urszula Jankowska, Magdalena Wilczak, and Małgorzata Przybyło. Similarities and differences in the protein composition of cutaneous melanoma cells and their exosomes identified by mass spectrometry. *Cancers*, 15(4): 1097, 2023.

[108] Viviana P Orellana, Andrés Tittarelli, and Mauricio A Retamal. Connexins in melanoma: Potential role of cx46 in its aggressiveness. *Pigment Cell & Melanoma Research*, 34(5):853–868, 2021.

[109] Daniel I Benjamin, Sharon M Louie, Melinda M Mulvihill, Rebecca A Kohnz, Daniel S Li, Lauryn G Chan, Antonio Sorrentino, Sourav Bandyopadhyay, Alyssa Cozzo, Anayo Ohiri, et al. Inositol phosphate recycling regulates glycolytic and lipid metabolism that drives cancer aggressiveness. *ACS chemical biology*, 9(6):1340–1350, 2014.

[110] Annapaola Angrisani, Annamaria Di Fiore, Enrico De Smaele, and Marta Moretti. The emerging role of the kctd proteins in cancer. *Cell Communication and Signaling*, 19(1):56, 2021.

[111] Peter Carbonetto and Matthew Stephens. Scalable variational inference for bayesian variable selection in regression, and its accuracy in genetic association studies. 2012.

[112] Michael Komodromos, Eric O Aboagye, Marina Evangelou, Sarah Filippi, and Kolyan Ray. Variational bayes for high-dimensional proportional hazards models with applications within gene expression. *Bioinformatics*, 38(16):3918–3926, 2022.

[113] Gao Wang, Abhishek Sarkar, Peter Carbonetto, and Matthew Stephens. A simple new approach to variable selection in regression, with application to genetic fine mapping. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 82(5):1273–1300, 2020.

[114] Lizhen Nie and Veronika Ročková. Bayesian bootstrap spike-and-slab lasso. *Journal of the American Statistical Association*, 118(543):2013–2028, 2023.

[115] Jianqing Fan and Jinchi Lv. Sure independence screening for ultrahigh dimensional feature space. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(5):849–911, 2008.

[116] Lan Wang, Yichao Wu, and Runze Li. Quantile regression for analyzing heterogeneity in ultra-high dimension. *Journal of the American Statistical Association*, 107(497): 214–222, 2012.

[117] Jian Huang, Patrick Breheny, and Shuangge Ma. A selective review of group selection

in high-dimensional models. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 27(4), 2012.

[118] Roger Koenker and Gilbert Bassett Jr. Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50, 1978.

[119] Hui Zou and Ming Yuan. Composite quantile regression and the oracle model selection theory. *The Annals of Statistics*, 36(3):1108–1126, 2008.

[120] Hanwen Huang and Zhongxue Chen. Bayesian composite quantile regression. *Journal of Statistical Computation and Simulation*, 85(18):3744–3754, 2015.

[121] Fei Zhou, Jie Ren, Gengxin Li, Yu Jiang, Xiaoxi Li, Weiqun Wang, and Cen Wu. Penalized variable selection for lipid–environment interactions in a longitudinal lipidomics study. *Genes*, 10(12):1002, 2019.

[122] Tao Lu and Yangxin Huang. Bayesian inference on mixed-effects varying-coefficient joint models with skew-t distribution for longitudinal data with multiple features. *Statistical methods in medical research*, 26(3):1146–1164, 2017.

[123] Xiaosun Lu and Yangxin Huang. Bayesian analysis of nonlinear mixed-effects mixture models for longitudinal data with heterogeneity and skewness. *Statistics in medicine*, 33(16):2830–2849, 2014.

[124] Gul Inan and Lan Wang. Pgee: an r package for analysis of longitudinal data with high-dimensional covariates. *R J.*, 9(1):393, 2017.

[125] Fei Zhou, Yuwen Liu, Jie Ren, Weiqun Wang, and Cen Wu. Springer: An r package for bi-level variable selection of high-dimensional longitudinal data. *Frontiers in Genetics*, 14:1088223, 2023.

[126] Fei Zhou, Jie Ren, Yuwen Liu, Xiaoxi Li, Weiqun Wang, and Cen Wu. Interep: An r package for high-dimensional interaction analysis of the repeated measurement data. *Genes*, 13(3):544, 2022.

[127] Larry Schumaker. *Spline functions: basic theory*. Cambridge university press, 2007.

[128] Ling Chen, Lina Wang, Zhuolong Xiong, Xiao Zhu, and Lianzhou Chen. Exploring the prognostic significance and immunotherapeutic potential of single-cell sequencing-identified long noncoding rna (lncrna) in patients with non-small cell lung cancer. *Cureus*, 15(11), 2023.

[129] Alfredo A Santillan, Carlos A Camargo, and Graham A Colditz. A meta-analysis of asthma and risk of lung cancer (united states). *Cancer Causes & Control*, 14:327–334, 2003.

[130] Hong Yan, Lei Bi, Yunshan Wang, Xia Zhang, Zhibo Hou, Qian Wang, Antoine M Snijders, and Jian-Hua Mao. Integrative analysis of multi-omics data reveals distinct impacts of ddb1-cul4 associated factors in human lung adenocarcinomas. *Scientific reports*, 7(1):333, 2017.

[131] Na Che, Nan Zhao, Xiulan Zhao, Shuai Su, Yanhui Zhang, Xiaoyu Bai, Fan Li, Danfang Zhang, and Yanlei Li. The expression and prognostic significance of pik3cb in lung adenocarcinoma. *Annals of Diagnostic Pathology*, 60:152001, 2022.

[132] Jiexia Zhang, Jun Li, Shiyue Li, Chengzhi Zhou, Yinyin Qin, and Xiaoxiang Li. mir-802 inhibits the aggressive behaviors of non-small cell lung cancer cells by directly targeting fgfr1 retraction in/10.3892/ijo. 2022.5389. *International Journal of Oncology*, 54(6): 2211–2221, 2019.

[133] Andrew J Cooke, Jutaro Fukumoto, Lee C Tan, Richard F Lockey, and Narasaiah Kolliputi. Atp8b1 deficiency enhances hyperoxia-induced lipid oxidation and apoptotic response in lung epithelial cells. *Journal of Allergy and Clinical Immunology*, 137(2): AB68, 2016.

[134] Sohyun Hwang, Seung-Woo Son, Sang Cheol Kim, Young Joo Kim, Hawoong Jeong, and Doheon Lee. A protein interaction network associated with asthma. *Journal of theoretical biology*, 252(4):722–731, 2008.

[135] Yunwen Yang, Huixia Judy Wang, and Xuming He. Posterior inference in bayesian quantile regression with asymmetric laplace likelihood. *International Statistical Review*, 84(3):327–344, 2016.

[136] Daniëlle Bodmer, Marc Eleveld, Ellen Kater-Baats, Irene Janssen, Bert Janssen, Marian Weterman, Eric Schoenmakers, Michael Nickerson, Marston Linehan, Berton Zbar, et al. Disruption of a novel mfs transporter gene, dirc2, by a familial renal cell carcinoma-associated t (2; 3)(q35; q21). *Human molecular genetics*, 11(6):641–649, 2002.

[137] Lalu Rudyat Telly Savalas, Bruno Gasnier, Markus Damme, Torben Lübke, Christian Wrocklage, Cécile Debacker, Adrien Jézégou, Thomas Reinheckel, Andrej Hasilik, Paul Saftig, et al. Disrupted in renal carcinoma 2 (dirc2), a novel transporter of the lysosomal membrane, is proteolytically processed by cathepsin l. *Biochemical Journal*, 439(1):113–128, 2011.

[138] Tingting Deng, Yan Huang, Kai Weng, Song Lin, Yujing Li, Guang Shi, Yali Chen, Junjiu Huang, Dan Liu, Wenbin Ma, et al. Toe1 acts as a 3' exonuclease for telomerase rna and regulates telomere maintenance. *Nucleic acids research*, 47(1):391–405, 2019.

[139] Rea M Lardelli, Ashleigh E Schaffer, Veerle RC Eggens, Maha S Zaki, Stephanie Grainger, Shashank Sathe, Eric L Van Nostrand, Zinayida Schlachetzki, Basak Rosti, Naiara Akizu, et al. Biallelic mutations in the 3' exonuclease toe1 cause pontocerebellar hypoplasia and uncover a role in snrna processing. *Nature genetics*, 49(3):457–464, 2017.

[140] TingAn Wang, HuaGe Zhong, YuZhou Qin, WeiYuan Wei, Zhao Li, MingWei Huang, XiaoLing Luo, et al. Armcx family gene expression analysis and potential prognostic biomarkers for prediction of clinical outcome in patients with gastric carcinoma. *BioMed Research International*, 2020, 2020.

[141] Constanze Zeller, Wei Dai, Nicola L Steele, Afshan Siddiq, Andrew J Walley, CSM

Wilhelm-Benartzi, S Rizzo, A Van Der Zee, JA Plumb, and R Brown. Candidate dna methylation drivers of acquired cisplatin resistance in ovarian cancer identified by methylome and expression profiling. *Oncogene*, 31(42):4567–4576, 2012.

Appendix A

Appendices for Chapter 2

A.1 Derivations of Alternative Methods

A.1.1 PGEE Derivation

Alternative methods involve applying different penalties including Lasso, Elastic net, SCAD and MCP to the PGEE functions.

Penalized generalized estimating (PGEE) functions are defined as

$$Q_n(\beta_n) = U(\beta_n) - q'_{\lambda_n}(|\beta_n|)sign(\beta_n)$$

$q'_{\lambda_n}(|\beta_n|) = (q'_{\lambda_n}(|\beta_{n1}|), \ldots, q'_{\lambda_n}(|\beta_{np_n}|))^\top$ is a p_n-dimensional vector of penalty functions.

LASSO estimator for PGEE:

$$q'_\lambda(|\beta_j|) = \lambda$$

where λ is tuning parameter.

Elastic net estimator for PGEE:

$$q'_\lambda(|\beta_j|) = \lambda_1 + 2\lambda_2|\beta_j|$$

where λ_1 and λ_2 are tuning parameters for L1 and L2 term, respectively.

SCAD estimator for PGEE:

$$q'_\lambda(|\beta_j|) = \begin{cases} \lambda & |\beta_j| \leq \lambda \\ \frac{\zeta\lambda - |\beta_j|}{\zeta - 1} & \lambda < |\beta_j| \leq \zeta\lambda \\ 0 & |\beta_j| \geq \zeta\lambda \end{cases}$$

where $\zeta > 2$.

MCP estimator for PGEE:

$$q'_\lambda(|\beta_j|) = \begin{cases} \lambda - \frac{|\beta_j|}{\zeta} & |\beta_j| \leq \zeta\lambda \\ 0 & |\beta_j| > \zeta\lambda \end{cases}$$

where $\zeta > 1$.

A.1.2 Computational Algorithms

A Newton-Raphson type of algorithm was developed to obtain the penalized estimate $\hat{\beta}$. The iterative algorithm for solving PGEE combines the minorization-maximization algorithm for nonconvex penalty of Hunter and Li[55] with the Newton-Raphson algorithm for the GEE.

For a small $\epsilon > 0$, the minorization-maximization algorithm suggests that the PGEE estimator $\hat{\beta} = (\hat{\beta}_1, \ldots, \hat{\beta}_p)^T$ approximately satisfies

$$U_j(\hat{\beta}) - nq'_\lambda(|\hat{\beta}_j|)sign(\hat{\beta}_j)\frac{|\hat{\beta}_j|}{\epsilon + |\hat{\beta}_j|} = 0, \quad j = 1, \ldots, p, \tag{A.1}$$

where $U_j(\hat{\beta})$ denotes the jth element of $U(\hat{\beta})$. ϵ is taken to be a fixed small number 10^{-6}.

By applying Newton-Raphson algorithm to $U_j(\hat{\beta})$, the iterative algorithm for solving the estimated $\hat{\beta}^d$ in the dth iteration can be obtained by:

$$\hat{\beta}^{(d)} = \hat{\beta}^{(d-1)} + [H(\hat{\beta}^{(d-1)}) + nE(\hat{\beta}^{(d-1)})]^{-1} \times [U(\hat{\beta}^{(d-1)}) - nE(\hat{\beta}^{(d-1)})\hat{\beta}^{(d-1)}], \tag{A.2}$$

where $U(\hat{\beta}^{(d-1)})$ is the score function at the $(d-1)$th iteration and $H(\hat{\beta}^{(d-1)})$ is the corresponding first derivative function of $U(\hat{\beta}^{(d-1)})$. $E(\hat{\beta}^{d-1})$ is a diagonal matrix that contains the first derivative of the network-constrained penalty. They are defined as:

$$\begin{aligned} H(\hat{\beta}^{(d-1)}) &= \sum_{i=1}^{n} D_i^\top V_i^{-1} D_i, \\ E(\hat{\beta}^{(d-1)}) &= diag\{\underbrace{0, ..., 0}_{1+q}, \frac{q'_\lambda(|\hat{\beta}_1{}^{(d-1)}|+)}{\epsilon + |\hat{\beta}_1{}^{(d-1)}|}, \ldots, \frac{q'_\lambda(|\hat{\beta}_p{}^{(d-1)}|+)}{\epsilon + |\hat{\beta}_p{}^{(d-1)}|}\}. \end{aligned} \tag{A.3}$$

The first $(1+q)$ elements on the diagonal of $E(\hat{\beta}^{d-1})$ are zero, suggesting that there is no shrinkage imposed on the coefficients for the intercept and the environmental factors.

Appendix B

Appendices for Chapter 3

B.1 Details on Alternative Methods

Alternative methods under comparison are LASSO, quantile LASSO and the spike-and-slab LASSO (ssLASSO). The first two methods have been implemented in R packages *glmnet*[89] and *rqPen*[99], respectively. Below we provide the details of ssLASSO. Note that the derivations for the original version of Spike-and-Slab LASSO are avaiable from published studies[17;90;91]. The ssLASSO adopted in this studies slightly differs in including an term corresponding to low-dimensional clinical and environmental variables that are not subject to selection.

B.2 The Spike-and-Slab LASSO

Denote y_i as the response of the i-th subject ($1 \leqslant i \leqslant n$). Let $\boldsymbol{z}_i = (1, z_{i1}, \ldots, z_{iq})^\top$ be a a $(q+1)$-dimensional vector for the intercept and q clinical factors. In addition, we denote $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})^\top$ as a p-dimensional vector for high-dimensional genetic factors. Consider the following linear regression model:

$$y_i = \boldsymbol{z}_i^\top \boldsymbol{\alpha} + \boldsymbol{x}_i^\top \boldsymbol{\beta} + \epsilon_i,$$

where coefficient vector $\boldsymbol{\alpha} = (\alpha_0, \cdots, \alpha_q)^\top$ represents regression coefficients for the intercept and clinical covariates, respectively. The coefficient vector $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_p)^\top$ is corresponding to the p dimensional omics features. We assume that the random error ϵ_i follows a normal distribution with standard deviation $\tilde{\sigma}$.

The spike-and-slab mixture double-exponential prior is assigned on each component of $\boldsymbol{\beta}$ as:

$$\beta_j|\gamma_j, s_0, s_1 \overset{ind}{\sim} (1-\gamma_j)\pi(\beta_j|s_0) + \gamma_j\pi(\beta_j|s_1),$$

with $\gamma_j = 0$ or 1 indicating the corresponding coefficient β_j being 0 or not. With $s_1 > s_0 > 0$, the scale parameter s_0 controls the spike part which models small and negligible coefficients that are close to 0, and s_1 controls the diffuse slab distribution which models important effects corresponding to regression coefficients with large magnitude. The above prior can also be viewed as $\beta_j|S_j \overset{ind}{\sim} \frac{1}{2S_j}\exp\left(-\frac{|\beta_j|}{S_j}\right)$, where $S_j = (1-\gamma_j)s_0 + \gamma_j s_1$. Regular double exponential prior is a special case when $s_0 = s_1$ or γ_j is set to be 0 or 1.

The indicator variable γ_j is assumed to independently follow a Bernoulli distribution:

$$\gamma_j|\theta \overset{ind}{\sim} \text{Ber}(\theta) = \theta^{\gamma_j}(1-\theta)^{1-\gamma_j},$$

with the probability parameter θ being assigned a uniform prior $\theta \sim U(0,1)$. It can be viewed as probability $p(\gamma_j = 1|\theta)$, a global shrinkage parameter indicating the prior probability of nonzero components in $\boldsymbol{\beta}$. Estimating different S_js will lead to different level of regularization for each coefficient.

The intercept and each clinical factor $\boldsymbol{\alpha}_k$ are assumed to have independent normal priors with mean 0 and variance V_k:

$$\alpha_k \overset{ind}{\sim} \text{N}(0, V_k),$$

where V_k is set to 10^3 for computational convenience. For variance $\tilde{\sigma}^2$, we adopt the limiting improper prior:

$$\pi(\tilde{\sigma}^2) = \tilde{\sigma}^{-2}.$$

Recall that $S_j = (1-\gamma_j)s_0 + \gamma_j s_1$, and denote $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_p)^\top$. After omitting the irrelevant terms, the log joint posterior density function is given by

$$\begin{aligned}
Q &= \log p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \tilde{\sigma}^2, \boldsymbol{\gamma}, \theta | \boldsymbol{Y}) \\
&= \log p(\boldsymbol{Y} | \boldsymbol{\alpha}, \boldsymbol{\beta}, \tilde{\sigma}^2) + \sum_{k=0}^{q} \log p(\alpha_k) + \sum_{j=1}^{p} \log p(\beta_j | S_j) + \log p(\tilde{\sigma}^2) + \sum_{j=1}^{p} \log p(\gamma_j | \theta) + \log p(\theta) \\
&\propto -\frac{n}{2}\log(\tilde{\sigma}^2) - \frac{1}{2\tilde{\sigma}^2}\sum_{i=1}^{n}(y_i - \boldsymbol{z}_i^\top \boldsymbol{\alpha} - \boldsymbol{x}_i^\top \boldsymbol{\beta})^2 - \sum_{k=0}^{q}\frac{\alpha_k^2}{2V_k} - \sum_{j=1}^{p}\frac{1}{S_j}|\beta_j| - \log(\tilde{\sigma}^2) \\
&\quad + \sum_{j=1}^{p}\left[\gamma_j \log\theta + (1-\gamma_j)\log(1-\theta)\right].
\end{aligned} \tag{B.1}$$

At E-step, we calculate the conditional expectations of γ_j at d-th iteration:

$$\begin{aligned}
\eta_j^{(d)} &= p(\gamma_j = 1 | \beta_j^{(d)}, \theta^{(d)}, \boldsymbol{Y}) \\
&= \frac{p(\beta_j^{(d)} | \gamma_j = 1, s_1) p(\gamma_j = 1 | \theta^{(d)})}{p(\beta_j^{(d)} | \gamma_j = 0, s_0) p(\gamma_j = 0 | \theta^{(d)}) + p(\beta_j^{(d)} | \gamma_j = 1, s_1) p(\gamma_j = 1 | \theta^{(d)})} \\
&= \frac{\pi(\beta_j^{(d)} | s_1)\theta^{(d)}}{\pi(\beta_j^{(d)} | s_0)(1-\theta^{(d)}) + \pi(\beta_j^{(d)} | s_1)\theta^{(d)}}.
\end{aligned} \tag{B.2}$$

Therefore, the conditional expectation of $(S_j^{-1})^{(d)}$ can be computed as:

$$(\tilde{S}_j^{-1})^{(d)} = E((S_j^{-1})^{(d)} | \beta_j^{(d)}) = E\left(\frac{1}{(1-\gamma_j)s_0 + \gamma_j s_1}\middle| \beta_j^{(d)}\right) = \frac{1-\eta_j^{(d)}}{s_0} + \frac{\eta_j^{(d)}}{s_1}. \tag{B.3}$$

At M-step, we update $\boldsymbol{\alpha}^{(d+1)}$, $\boldsymbol{\beta}^{(d+1)}$, $(\tilde{\sigma}^2)^{(d+1)}$ and $\theta^{(d+1)}$ by maximizing the log joint posterior density at the d-th iteration, or $Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(d)})$ where $\boldsymbol{\phi} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \tilde{\sigma}^2, \boldsymbol{\gamma}, \theta)$. In addition, we update the coefficient vector $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ component-wisely by following the coordinate descent strategy[89]. By solving $\frac{\partial Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(d)})}{\partial \alpha_l} = 0$, $\frac{\partial Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(d)})}{\partial \beta_m} = 0$, $\frac{\partial Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(d)})}{\partial \tilde{\sigma}^2} = 0$ and $\frac{\partial Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(d)})}{\partial \theta} = 0$, , where $l = 1, ..., q$ and $m = 1, ..., p$, we obtain the following equations:

$$
\begin{aligned}
\theta^{(d+1)} &= \frac{1}{p}\sum_{j=1}^{p}\eta_j^{(d)}, \\
(\tilde{\sigma}^2)^{(d+1)} &= \frac{\sum_{i=1}^{n}(y_i - \boldsymbol{z}_i^\top \boldsymbol{\alpha}^{(d)} - \boldsymbol{x}_i^\top \boldsymbol{\beta}^{(d)})^2}{n+2}, \\
\alpha_l^{(d+1)} &= \frac{\sum_{i=1}^{n}\left(y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}^{(d)} - \sum_{k=0}^{l-1} z_{ik}\alpha_k^{(d+1)} - \sum_{k=l+1}^{q} z_{ik}\alpha_k^{(d)}\right) z_{il}}{\frac{2(\tilde{\sigma}^2)^{(d+1)}}{V_l} + \sum_{i=1}^{n} z_{il}^2}, \\
\beta_m^{(d+1)} &= \frac{\mathrm{sgn}(T_m^{(d+1)})\left(\left|T_m^{(d+1)}\right| - (\tilde{S}_m^{-1})^{(d)}\right)_+}{\sum_{i=1}^{n} x_{im}^2/(\tilde{\sigma}^2)^{(d+1)}},
\end{aligned}
\tag{B.4}
$$

where

$$
T_m^{(d+1)} = \frac{1}{(\tilde{\sigma}^2)^{(d+1)}}\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{m-1} x_{ij}\beta_j^{(d+1)} - \sum_{j=m+1}^{p} x_{ij}\beta_j^{(d)} - \boldsymbol{z}_i^\top \boldsymbol{\alpha}^{(d+1)}\right) x_{im}.
$$

We do not perform variable selection on the clinical factors. Thus, shrinkage is only imposed on β. A k-fold cross-validation is used to choose the optimal tuning parameters s_0 and s_1. For fixed tuning parameters, the model fitting algorithm is provided in the table below:

Table B.1: EM algorithm for the spike-and-slab LASSO.

Algorithm
Initialize iteration $d = 0$, coefficients $\boldsymbol{\alpha}^{(0)}$, $\boldsymbol{\beta}^{(0)}$, $\theta^{(0)}$ and $(\tilde{\sigma}^2)^{(0)}$;
Repeat
E-step:
Compute the conditional expectation of $\gamma_j^{(d)}$ and $(\tilde{S}_j^{-1})^{(d)}$ via (C.2) and (C.3);
M-step:
Update $\theta^{(d+1)}$, $(\tilde{\sigma}^2)^{(d+1)}$, $\alpha_l^{(d+1)}$ and $\beta_m^{(d+1)}$ through (C.4);
Calculate the updated log joint posterior density and check if $\lvert Q^{(d+1)} - Q^{(d)}\rvert < \delta$;
$d \leftarrow d + 1$;
until convergence ($\lvert Q^{(d+1)} - Q^{(d)}\rvert < \delta$ for a small value $\delta > 0$).

B.3 Additional Simulations

All the tables listed in this Section have been referred in the Section of Simulation.

Table B.2: Evaluation of homogeneous model under AR-1 correlation, (n, p)=(400,800) with 100 replicates.

$n = 400$	$p = 800$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	1.53(2.03)	0.95(0.06)	0.96(0.05)	0.81(0.18)
		QLASSO	15.00(0.00)	19.47(4.44)	0.61(0.06)	0.66(0.05)	3.11(0.28)
		ssLASSO	15.00(0.00)	0.23(0.63)	0.99(0.02)	0.99(0.02)	0.63(0.17)
		LASSO	15.00(0.00)	58.40(24.56)	0.36(0.09)	0.45(0.07)	2.97(0.72)
	Error2	ssQLASSO	14.90(0.40)	0.87(1.11)	0.97(0.04)	0.97(0.03)	2.53(1.30)
		QLASSO	15.00(0.00)	20.03(4.55)	0.60(0.05)	0.65(0.04)	4.93(0.85)
		ssLASSO	10.03(3.27)	0.23(0.50)	0.77(0.17)	0.79(0.14)	4.60(1.96)
		LASSO	14.30(2.53)	56.50(21.79)	0.35(0.07)	0.43(0.07)	7.89(1.96)
	Error3	ssQLASSO	15.00(0.00)	0.63(0.89)	0.98(0.03)	0.98(0.03)	2.32(0.86)
		QLASSO	15.00(0.00)	19.40(4.57)	0.61(0.06)	0.66(0.05)	2.32(0.38)
		ssLASSO	12.60(2.98)	0.17(0.38)	0.89(0.14)	0.90(0.12)	2.95(1.83)
		LASSO	14.87(0.57)	58.50(19.83)	0.35(0.06)	0.44(0.06)	6.38(1.66)
	Error4	ssQLASSO	14.83(0.46)	0.97(1.43)	0.96(0.05)	0.97(0.05)	2.91(1.24)
		QLASSO	14.93(0.25)	18.60(4.61)	0.62(0.06)	0.66(0.05)	5.45(0.78)
		ssLASSO	11.40(4.09)	0.70(1.02)	0.81(0.21)	0.83(0.18)	4.00(2.47)
		LASSO	14.97(0.18)	55.80(17.40)	0.36(0.07)	0.45(0.06)	6.82(1.73)
	Error5	ssQLASSO	14.97(0.18)	1.73(2.80)	0.95(0.07)	0.95(0.07)	2.34(0.91)
		QLASSO	14.97(0.18)	18.93(3.96)	0.62(0.05)	0.66(0.04)	5.36(0.93)
		ssLASSO	13.57(2.21)	1.13(2.76)	0.91(0.10)	0.92(0.09)	2.46(1.44)
		LASSO	15.00(0.00)	53.87(19.38)	0.38(0.10)	0.47(0.08)	5.34(1.15)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	1.37(2.13)	0.96(0.06)	0.96(0.06)	0.86(0.23)
		QLASSO	15.00(0.00)	17.83(5.63)	0.64(0.07)	0.68(0.06)	2.81(0.37)
		ssLASSO	15.00(0.00)	0.50(0.86)	0.98(0.03)	0.98(0.03)	0.69(0.20)
		LASSO	15.00(0.00)	53.07(17.57)	0.38(0.08)	0.47(0.07)	2.71(0.50)
	Error2	ssQLASSO	14.93(0.25)	1.00(1.14)	0.97(0.04)	0.97(0.04)	3.70(0.91)
		QLASSO	15.00(0.00)	28.93(3.82)	0.51(0.03)	0.58(0.03)	4.25(0.92)
		ssLASSO	9.60(3.47)	0.27(0.64)	0.75(0.19)	0.77(0.16)	4.85(2.26)
		LASSO	13.77(3.22)	57.87(26.98)	0.33(0.10)	0.42(0.06)	8.33(2.40)
	Error3	ssQLASSO	15.00(0.00)	0.47(0.63)	0.99(0.02)	0.99(0.02)	2.66(0.62)
		QLASSO	15.00(0.00)	27.27(6.35)	0.53(0.06)	0.59(0.05)	2.92(0.47)
		ssLASSO	13.43(2.25)	0.17(0.38)	0.93(0.09)	0.94(0.09)	2.36(1.46)
		LASSO	14.80(0.92)	55.17(18.68)	0.36(0.07)	0.45(0.07)	6.06(2.34)
	Error4	ssQLASSO	14.87(0.43)	1.03(0.85)	0.96(0.03)	0.96(0.03)	3.93(0.75)
		QLASSO	15.00(0.00)	30.30(5.47)	0.50(0.04)	0.57(0.04)	4.69(0.80)
		ssLASSO	10.53(4.03)	0.70(1.44)	0.77(0.21)	0.79(0.18)	4.43(2.51)
		LASSO	14.97(0.18)	59.27(20.64)	0.35(0.07)	0.44(0.06)	7.08(1.40)
	Error5	ssQLASSO	14.97(0.18)	0.93(1.26)	0.97(0.04)	0.97(0.04)	3.48(0.65)
		QLASSO	15.00(0.00)	27.97(3.85)	0.52(0.03)	0.58(0.03)	4.58(0.58)
		ssLASSO	12.30(2.68)	0.30(1.12)	0.88(0.12)	0.89(0.11)	3.09(1.63)
		LASSO	15.00(0.00)	63.83(28.88)	0.35(0.09)	0.44(0.08)	6.15(1.91)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	1.27(1.76)	0.96(0.05)	0.96(0.05)	0.83(0.20)
		QLASSO	15.00(0.00)	19.70(5.15)	0.61(0.06)	0.65(0.05)	3.11(0.32)
		ssLASSO	15.00(0.00)	0.63(1.07)	0.98(0.03)	0.98(0.03)	0.72(0.21)
		LASSO	15.00(0.00)	57.17(17.08)	0.36(0.07)	0.45(0.06)	2.99(0.43)
	Error2	ssQLASSO	14.90(0.40)	0.53(0.82)	0.98(0.03)	0.98(0.03)	2.11(1.00)
		QLASSO	14.97(0.18)	17.67(3.40)	0.63(0.05)	0.67(0.04)	4.57(0.62)
		ssLASSO	11.27(3.33)	0.10(0.31)	0.84(0.17)	0.85(0.14)	4.18(2.17)
		LASSO	13.70(3.71)	50.87(21.93)	0.34(0.10)	0.42(0.11)	8.25(3.38)
	Error3	ssQLASSO	15.00(0.00)	0.97(1.30)	0.97(0.04)	0.97(0.04)	2.34(0.70)
		QLASSO	15.00(0.00)	19.03(4.18)	0.62(0.05)	0.66(0.04)	5.18(0.89)
		ssLASSO	13.37(2.51)	0.40(1.16)	0.92(0.11)	0.93(0.10)	2.45(1.76)
		LASSO	15.00(0.00)	59.57(20.32)	0.35(0.07)	0.44(0.06)	6.04(2.01)
	Error4	ssQLASSO	14.83(0.38)	1.03(1.75)	0.96(0.05)	0.96(0.05)	2.78(1.13)
		QLASSO	14.93(0.37)	19.73(3.31)	0.60(0.04)	0.65(0.04)	5.63(0.88)
		ssLASSO	8.67(4.00)	0.27(0.69)	0.69(0.22)	0.72(0.18)	5.26(2.32)
		LASSO	15.00(0.00)	62.20(25.40)	0.34(0.08)	0.44(0.07)	7.28(2.13)
	Error5	ssQLASSO	14.97(0.18)	1.20(1.24)	0.96(0.04)	0.96(0.04)	2.47(0.93)
		QLASSO	14.83(0.46)	18.50(4.45)	0.62(0.06)	0.66(0.05)	5.39(0.87)
		ssLASSO	13.40(1.99)	0.57(0.94)	0.92(0.07)	0.92(0.07)	2.49(1.19)
		LASSO	15.00(0.00)	55.33(19.49)	0.37(0.09)	0.46(0.07)	5.69(1.20)

Table B.3: Evaluation of homogeneous model under banded correlation, (n, p)=(400,800) with 100 replicates.

$n = 400$	$p = 800$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	0.80(1.45)	0.98(0.04)	0.98(0.04)	0.89(0.21)
		QLASSO	15.00(0.00)	18.47(5.62)	0.63(0.07)	0.67(0.06)	2.95(0.37)
		ssLASSO	15.00(0.00)	0.47(0.78)	0.99(0.02)	0.99(0.02)	0.71(0.20)
		LASSO	15.00(0.00)	59.03(23.25)	0.36(0.09)	0.45(0.08)	3.00(0.67)
	Error2	ssQLASSO	14.83(0.38)	0.50(0.78)	0.98(0.02)	0.98(0.02)	2.43(0.99)
		QLASSO	14.97(0.18)	19.77(4.12)	0.61(0.05)	0.65(0.04)	4.69(0.80)
		ssLASSO	10.63(2.66)	0.10(0.31)	0.81(0.13)	0.83(0.11)	4.24(1.64)
		LASSO	14.33(1.77)	52.57(22.28)	0.37(0.08)	0.45(0.08)	8.08(2.70)
	Error3	ssQLASSO	15.00(0.00)	0.27(0.78)	0.99(0.02)	0.99(0.02)	1.99(0.69)
		QLASSO	15.00(0.00)	17.97(4.15)	0.63(0.05)	0.67(0.04)	2.32(0.36)
		ssLASSO	12.63(2.82)	0.10(0.31)	0.90(0.13)	0.91(0.11)	2.77(1.84)
		LASSO	15.00(0.00)	56.47(14.48)	0.36(0.06)	0.45(0.05)	5.90(1.38)
	Error4	ssQLASSO	14.87(0.43)	1.23(1.63)	0.96(0.05)	0.96(0.05)	2.81(1.11)
		QLASSO	14.80(0.48)	18.53(4.38)	0.62(0.06)	0.66(0.05)	5.54(0.92)
		ssLASSO	10.03(3.39)	0.20(0.41)	0.77(0.17)	0.80(0.15)	4.60(2.19)
		LASSO	14.97(0.18)	61.07(18.14)	0.34(0.07)	0.44(0.06)	7.02(1.43)
	Error5	ssQLASSO	14.97(0.18)	1.23(1.65)	0.96(0.05)	0.96(0.04)	2.59(0.99)
		QLASSO	14.90(0.40)	19.20(3.50)	0.61(0.05)	0.65(0.04)	6.03(0.96)
		ssLASSO	11.47(3.43)	0.60(1.19)	0.83(0.16)	0.84(0.13)	3.48(2.06)
		LASSO	15.00(0.00)	62.20(19.55)	0.34(0.06)	0.43(0.05)	6.11(1.25)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	1.83(2.64)	0.95(0.07)	0.95(0.07)	0.83(0.17)
		QLASSO	15.00(0.00)	17.43(4.45)	0.64(0.06)	0.68(0.05)	3.03(0.33)
		ssLASSO	15.00(0.00)	0.47(0.63)	0.99(0.02)	0.99(0.02)	0.64(0.18)
		LASSO	15.00(0.00)	57.53(20.45)	0.36(0.08)	0.45(0.07)	2.86(0.53)
	Error2	ssQLASSO	14.97(0.18)	0.97(1.13)	0.97(0.03)	0.97(0.03)	3.97(0.62)
		QLASSO	15.00(0.00)	29.40(5.93)	0.51(0.05)	0.57(0.04)	4.42(0.73)
		ssLASSO	8.20(3.81)	0.20(0.48)	0.67(0.22)	0.73(0.16)	5.75(2.25)
		LASSO	13.43(4.06)	54.00(25.87)	0.32(0.10)	0.42(0.09)	8.05(2.12)
	Error3	ssQLASSO	15.00(0.00)	0.53(0.82)	0.98(0.03)	0.98(0.02)	3.04(0.52)
		QLASSO	15.00(0.00)	29.90(4.16)	0.50(0.04)	0.57(0.03)	3.23(0.46)
		ssLASSO	12.10(2.95)	0.07(0.25)	0.88(0.14)	0.89(0.12)	3.16(1.93)
		LASSO	15.00(0.00)	57.30(18.64)	0.36(0.08)	0.45(0.07)	6.25(1.51)
	Error4	ssQLASSO	14.67(0.66)	1.73(1.44)	0.94(0.04)	0.94(0.04)	4.42(0.97)
		QLASSO	15.00(0.00)	29.17(4.39)	0.51(0.04)	0.57(0.03)	4.87(0.83)
		ssLASSO	8.07(3.50)	0.47(0.63)	0.66(0.19)	0.69(0.16)	5.94(2.27)
		LASSO	15.00(0.00)	67.57(21.02)	0.32(0.07)	0.42(0.06)	7.75(1.25)
	Error5	ssQLASSO	14.80(0.66)	1.30(1.53)	0.95(0.05)	0.95(0.05)	3.75(0.91)
		QLASSO	14.97(0.18)	30.03(5.61)	0.50(0.05)	0.57(0.04)	4.88(0.98)
		ssLASSO	12.87(2.84)	0.80(1.27)	0.89(0.13)	0.89(0.12)	2.77(1.77)
		LASSO	15.00(0.00)	60.23(23.67)	0.35(0.08)	0.44(0.07)	5.99(1.54)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	1.60(1.69)	0.95(0.05)	0.95(0.05)	0.84(0.22)
		QLASSO	15.00(0.00)	17.50(4.64)	0.64(0.07)	0.68(0.06)	3.06(0.40)
		ssLASSO	15.00(0.00)	0.43(0.73)	0.99(0.02)	0.99(0.02)	0.64(0.17)
		LASSO	15.00(0.00)	64.67(21.48)	0.33(0.07)	0.43(0.06)	3.11(0.69)
	Error2	ssQLASSO	14.97(0.18)	0.50(0.97)	0.98(0.03)	0.98(0.03)	2.21(0.97)
		QLASSO	14.97(0.18)	19.17(3.71)	0.61(0.05)	0.66(0.04)	5.05(0.90)
		ssLASSO	9.73(3.89)	0.13(0.35)	0.75(0.22)	0.78(0.18)	4.78(2.46)
		LASSO	14.47(1.94)	60.30(26.94)	0.34(0.08)	0.43(0.07)	8.06(2.75)
	Error3	ssQLASSO	15.00(0.00)	0.20(0.55)	0.99(0.02)	0.99(0.02)	2.11(0.70)
		QLASSO	14.93(0.25)	19.57(4.04)	0.61(0.05)	0.65(0.04)	5.46(0.88)
		ssLASSO	11.10(2.96)	0.07(0.25)	0.83(0.14)	0.85(0.13)	3.77(1.87)
		LASSO	15.00(0.00)	58.50(23.39)	0.36(0.07)	0.45(0.06)	6.24(2.03)
	Error4	ssQLASSO	14.87(0.35)	1.47(2.30)	0.95(0.06)	0.95(0.06)	2.98(1.40)
		QLASSO	14.70(0.75)	17.87(3.66)	0.62(0.06)	0.66(0.05)	5.72(1.10)
		ssLASSO	9.23(4.06)	0.33(0.71)	0.71(0.22)	0.75(0.18)	5.00(2.45)
		LASSO	14.97(0.18)	59.23(21.12)	0.35(0.09)	0.45(0.07)	6.88(1.46)
	Error5	ssQLASSO	15.00(0.00)	1.70(1.78)	0.95(0.05)	0.95(0.05)	2.67(1.03)
		QLASSO	14.87(0.35)	18.73(3.49)	0.61(0.04)	0.66(0.04)	5.62(0.87)
		ssLASSO	12.80(3.23)	1.47(3.41)	0.86(0.16)	0.88(0.14)	2.92(1.91)
		LASSO	15.00(0.00)	58.87(17.59)	0.35(0.06)	0.44(0.06)	5.79(1.10)

Table B.4: Evaluation of homogeneous model under banded correlation, (n, p)=(400,1600) with 100 replicates.

$n = 400$	$p = 1600$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	1.77(2.92)	0.95(0.08)	0.95(0.07)	0.86(0.18)
		QLASSO	15.00(0.00)	37.20(7.61)	0.45(0.05)	0.53(0.04)	3.31(0.49)
		ssLASSO	15.00(0.00)	0.87(1.04)	0.97(0.03)	0.97(0.03)	0.69(0.18)
		LASSO	15.00(0.00)	69.07(22.24)	0.32(0.07)	0.42(0.06)	3.38(0.69)
	Error2	ssQLASSO	14.60(0.89)	1.23(2.25)	0.95(0.06)	0.95(0.06)	3.39(1.41)
		QLASSO	15.00(0.00)	32.33(5.77)	0.49(0.05)	0.56(0.04)	5.60(1.09)
		ssLASSO	8.57(3.65)	0.27(0.58)	0.69(0.21)	0.72(0.18)	5.68(2.35)
		LASSO	14.37(1.07)	70.10(22.25)	0.30(0.07)	0.40(0.06)	9.56(2.73)
	Error3	ssQLASSO	14.93(0.25)	0.37(0.61)	0.99(0.02)	0.99(0.02)	2.70(1.09)
		QLASSO	15.00(0.00)	27.40(9.59)	0.54(0.09)	0.60(0.07)	2.75(0.50)
		ssLASSO	10.20(3.22)	0.27(1.28)	0.78(0.17)	0.80(0.14)	4.37(1.96)
		LASSO	14.80(1.10)	78.10(42.92)	0.30(0.08)	0.41(0.07)	7.95(3.29)
	Error4	ssQLASSO	14.53(0.86)	1.13(1.98)	0.95(0.06)	0.95(0.06)	3.18(1.70)
		QLASSO	14.90(0.31)	33.50(4.86)	0.47(0.04)	0.55(0.03)	6.32(0.89)
		ssLASSO	7.53(3.72)	0.33(0.84)	0.62(0.22)	0.67(0.18)	5.91(1.99)
		LASSO	15.00(0.00)	83.70(33.33)	0.28(0.07)	0.39(0.06)	8.65(2.59)
	Error5	ssQLASSO	14.83(0.46)	1.17(1.32)	0.96(0.04)	0.96(0.04)	3.05(1.00)
		QLASSO	14.90(0.31)	34.57(5.16)	0.46(0.04)	0.54(0.03)	6.72(0.95)
		ssLASSO	9.33(4.06)	0.30(0.79)	0.72(0.21)	0.76(0.18)	4.93(2.45)
		LASSO	15.00(0.00)	75.27(28.47)	0.30(0.07)	0.41(0.06)	7.08(1.70)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	1.03(1.54)	0.97(0.04)	0.97(0.04)	0.89(0.20)
		QLASSO	15.00(0.00)	33.87(8.13)	0.48(0.06)	0.55(0.05)	3.35(0.46)
		ssLASSO	15.00(0.00)	0.70(1.86)	0.98(0.05)	0.98(0.05)	0.72(0.34)
		LASSO	15.00(0.00)	68.97(24.67)	0.32(0.07)	0.43(0.06)	3.37(0.63)
	Error2	ssQLASSO	13.90(1.37)	1.20(1.30)	0.92(0.06)	0.92(0.06)	5.20(1.18)
		QLASSO	15.00(0.00)	50.33(4.96)	0.37(0.02)	0.47(0.02)	5.19(0.69)
		ssLASSO	6.80(3.13)	0.07(0.25)	0.59(0.20)	0.65(0.16)	6.31(1.89)
		LASSO	13.00(3.97)	73.77(36.18)	0.26(0.11)	0.36(0.10)	9.69(3.07)
	Error3	ssQLASSO	15.00(0.00)	0.33(0.55)	0.99(0.02)	0.99(0.02)	4.00(0.96)
		QLASSO	15.00(0.00)	49.90(6.57)	0.38(0.03)	0.48(0.03)	3.86(0.80)
		ssLASSO	10.03(3.77)	0.13(0.35)	0.77(0.21)	0.79(0.18)	4.49(2.38)
		LASSO	14.90(0.40)	68.63(21.91)	0.32(0.08)	0.42(0.06)	7.16(2.03)
	Error4	ssQLASSO	13.77(1.92)	1.10(1.86)	0.92(0.09)	0.92(0.08)	5.39(1.32)
		QLASSO	15.00(0.00)	48.70(7.43)	0.38(0.04)	0.48(0.03)	5.66(0.79)
		ssLASSO	6.40(2.76)	0.17(0.46)	0.57(0.18)	0.63(0.14)	6.81(1.66)
		LASSO	15.00(0.00)	72.70(29.23)	0.31(0.08)	0.42(0.07)	7.98(1.59)
	Error5	ssQLASSO	14.83(0.46)	1.30(1.60)	0.95(0.05)	0.96(0.05)	4.20(0.81)
		QLASSO	15.00(0.00)	50.90(4.98)	0.37(0.03)	0.47(0.02)	5.78(1.09)
		ssLASSO	9.37(4.03)	0.73(1.31)	0.71(0.19)	0.75(0.16)	4.95(2.40)
		LASSO	15.00(0.00)	70.97(20.85)	0.31(0.06)	0.42(0.05)	6.79(1.10)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	1.47(2.10)	0.96(0.06)	0.96(0.05)	0.88(0.15)
		QLASSO	15.00(0.00)	35.00(6.87)	0.47(0.05)	0.55(0.04)	3.27(0.35)
		ssLASSO	15.00(0.00)	0.57(0.94)	0.98(0.03)	0.98(0.03)	0.65(0.11)
		LASSO	15.00(0.00)	67.07(19.21)	0.32(0.06)	0.43(0.05)	3.35(0.53)
	Error2	ssQLASSO	14.37(1.35)	0.70(1.12)	0.95(0.06)	0.96(0.06)	3.26(1.63)
		QLASSO	14.97(0.18)	32.57(4.07)	0.48(0.03)	0.56(0.02)	5.70(0.95)
		ssLASSO	7.43(3.72)	0.20(0.41)	0.62(0.23)	0.67(0.19)	6.42(2.16)
		LASSO	14.00(1.95)	60.23(27.32)	0.33(0.08)	0.43(0.06)	9.03(1.87)
	Error3	ssQLASSO	14.87(0.57)	0.17(0.38)	0.99(0.02)	0.99(0.02)	2.81(0.74)
		QLASSO	14.97(0.18)	32.87(4.80)	0.48(0.04)	0.55(0.03)	5.92(0.93)
		ssLASSO	9.93(3.66)	0.07(0.25)	0.77(0.19)	0.80(0.16)	4.28(2.22)
		LASSO	15.00(0.00)	72.37(24.99)	0.31(0.08)	0.42(0.06)	6.74(1.75)
	Error4	ssQLASSO	14.57(1.14)	0.87(1.14)	0.96(0.05)	0.96(0.05)	3.37(1.37)
		QLASSO	14.80(0.41)	33.53(4.79)	0.47(0.04)	0.55(0.03)	6.83(1.08)
		ssLASSO	6.83(3.30)	0.27(0.45)	0.59(0.19)	0.64(0.16)	6.39(1.93)
		LASSO	15.00(0.00)	77.57(28.77)	0.30(0.08)	0.41(0.07)	8.17(2.01)
	Error5	ssQLASSO	14.70(0.84)	1.33(1.54)	0.95(0.05)	0.95(0.05)	3.27(1.27)
		QLASSO	14.83(0.46)	31.93(4.76)	0.48(0.04)	0.56(0.03)	6.64(0.89)
		ssLASSO	8.33(3.84)	0.27(0.74)	0.67(0.21)	0.71(0.18)	5.67(2.47)
		LASSO	15.00(0.00)	82.00(27.70)	0.29(0.08)	0.40(0.07)	7.26(1.57)

Table B.5: Evaluation of heterogeneous model under AR-1 correlation, (n, p)=(400,800) with 100 replicates.

$n = 400$	$p = 800$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	0.63(1.16)	0.98(0.03)	0.98(0.03)	1.12(0.12)
		QLASSO	15.00(0.00)	18.70(5.13)	0.62(0.06)	0.66(0.05)	3.22(0.47)
		ssLASSO	15.00(0.00)	0.07(0.25)	1.00(0.01)	1.00(0.01)	1.39(0.22)
		LASSO	15.00(0.00)	58.77(21.41)	0.35(0.08)	0.45(0.07)	4.46(1.07)
	Error2	ssQLASSO	14.97(0.18)	0.30(0.99)	0.99(0.03)	0.99(0.03)	1.27(0.59)
		QLASSO	14.97(0.18)	19.27(3.77)	0.61(0.05)	0.66(0.04)	4.35(1.21)
		ssLASSO	12.07(2.84)	0.20(0.76)	0.87(0.14)	0.88(0.12)	4.39(2.29)
		LASSO	12.97(3.63)	52.17(27.01)	0.33(0.10)	0.40(0.11)	9.90(3.13)
	Error3	ssQLASSO	15.00(0.00)	0.50(1.07)	0.98(0.03)	0.98(0.03)	0.80(0.14)
		QLASSO	15.00(0.00)	17.13(4.12)	0.64(0.06)	0.68(0.05)	2.45(0.50)
		ssLASSO	14.27(1.36)	0.00(0.00)	0.97(0.05)	0.97(0.05)	2.91(1.09)
		LASSO	14.80(0.76)	59.63(19.97)	0.35(0.08)	0.44(0.07)	8.41(2.07)
	Error4	ssQLASSO	14.70(0.84)	0.30(0.47)	0.98(0.04)	0.98(0.04)	1.58(0.76)
		QLASSO	14.87(0.35)	20.20(4.57)	0.60(0.05)	0.64(0.04)	5.41(1.02)
		ssLASSO	11.70(2.94)	0.40(0.62)	0.85(0.14)	0.86(0.12)	4.56(1.60)
		LASSO	14.40(1.00)	57.10(23.67)	0.35(0.08)	0.44(0.06)	9.49(1.97)
	Error5	ssQLASSO	14.70(0.65)	0.27(0.45)	0.98(0.03)	0.98(0.03)	1.64(0.65)
		QLASSO	14.90(0.31)	19.70(5.36)	0.61(0.07)	0.65(0.05)	5.15(0.83)
		ssLASSO	13.17(1.72)	0.23(0.50)	0.92(0.07)	0.93(0.07)	3.44(1.04)
		LASSO	14.77(0.63)	60.57(23.23)	0.35(0.09)	0.44(0.08)	8.48(2.24)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	0.53(1.14)	0.98(0.03)	0.98(0.03)	0.63(0.20)
		QLASSO	15.00(0.00)	17.87(3.95)	0.63(0.05)	0.67(0.04)	3.13(0.53)
		ssLASSO	15.00(0.00)	0.10(0.31)	1.00(0.01)	1.00(0.01)	0.94(0.22)
		LASSO	15.00(0.00)	63.30(23.12)	0.34(0.08)	0.43(0.07)	4.25(0.98)
	Error2	ssQLASSO	14.93(0.25)	0.10(0.40)	0.99(0.01)	0.99(0.01)	0.95(0.45)
		QLASSO	15.00(0.00)	28.77(5.84)	0.52(0.05)	0.58(0.04)	3.78(0.72)
		ssLASSO	11.37(2.88)	0.17(0.38)	0.84(0.13)	0.86(0.11)	4.26(2.00)
		LASSO	12.00(4.39)	45.97(24.93)	0.33(0.10)	0.39(0.11)	9.66(2.28)
	Error3	ssQLASSO	15.00(0.00)	0.13(0.35)	1.00(0.01)	1.00(0.01)	1.65(0.25)
		QLASSO	15.00(0.00)	26.77(6.97)	0.54(0.07)	0.60(0.05)	3.68(0.67)
		ssLASSO	13.33(2.75)	0.17(0.38)	0.92(0.13)	0.93(0.11)	4.31(1.77)
		LASSO	14.50(1.17)	61.33(20.16)	0.33(0.06)	0.42(0.06)	10.03(2.17)
	Error4	ssQLASSO	14.50(1.20)	0.33(0.66)	0.97(0.05)	0.97(0.05)	1.46(1.06)
		QLASSO	15.00(0.00)	29.20(4.16)	0.51(0.04)	0.57(0.03)	4.18(0.97)
		ssLASSO	12.00(2.74)	0.27(0.58)	0.87(0.14)	0.88(0.12)	3.83(1.75)
		LASSO	14.47(0.73)	54.30(20.25)	0.36(0.08)	0.45(0.06)	9.11(2.59)
	Error5	ssQLASSO	14.73(0.64)	0.63(0.85)	0.97(0.04)	0.97(0.04)	1.38(0.81)
		QLASSO	15.00(0.00)	28.97(3.61)	0.51(0.03)	0.57(0.02)	4.68(0.93)
		ssLASSO	12.60(2.01)	0.50(0.86)	0.89(0.09)	0.90(0.08)	3.43(1.36)
		LASSO	14.83(0.38)	57.50(14.97)	0.35(0.06)	0.44(0.05)	8.50(1.48)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	0.50(1.55)	0.99(0.04)	0.99(0.04)	1.07(0.18)
		QLASSO	15.00(0.00)	18.30(4.47)	0.63(0.06)	0.67(0.05)	3.65(0.44)
		ssLASSO	15.00(0.00)	0.07(0.25)	1.00(0.01)	1.00(0.01)	1.35(0.23)
		LASSO	15.00(0.00)	53.17(17.13)	0.38(0.08)	0.46(0.07)	4.35(0.66)
	Error2	ssQLASSO	13.67(1.21)	0.13(0.43)	0.95(0.05)	0.95(0.05)	2.07(0.89)
		QLASSO	14.63(0.56)	19.27(3.88)	0.60(0.05)	0.64(0.04)	5.55(1.08)
		ssLASSO	10.53(2.61)	0.23(0.50)	0.81(0.13)	0.82(0.11)	4.69(1.79)
		LASSO	11.67(3.24)	50.70(22.31)	0.31(0.09)	0.36(0.10)	10.45(2.31)
	Error3	ssQLASSO	14.13(0.35)	0.10(0.40)	0.97(0.02)	0.97(0.02)	1.45(0.30)
		QLASSO	14.60(0.50)	18.87(4.29)	0.61(0.06)	0.65(0.05)	5.09(0.98)
		ssLASSO	12.07(2.91)	0.10(0.31)	0.87(0.14)	0.88(0.13)	3.48(1.82)
		LASSO	14.63(0.67)	62.47(21.11)	0.33(0.06)	0.42(0.06)	9.07(3.17)
	Error4	ssQLASSO	13.60(0.93)	0.27(0.58)	0.94(0.05)	0.94(0.05)	2.04(0.88)
		QLASSO	14.67(0.48)	19.03(3.37)	0.61(0.05)	0.65(0.04)	5.65(0.95)
		ssLASSO	10.83(2.52)	0.30(0.53)	0.82(0.13)	0.83(0.12)	4.50(1.62)
		LASSO	13.77(0.82)	51.93(25.82)	0.37(0.10)	0.44(0.08)	9.60(2.90)
	Error5	ssQLASSO	13.90(0.31)	0.57(1.14)	0.94(0.04)	0.94(0.04)	1.85(0.55)
		QLASSO	14.07(0.78)	19.20(4.33)	0.59(0.06)	0.62(0.05)	6.37(1.05)
		ssLASSO	11.00(3.44)	0.57(1.01)	0.81(0.17)	0.82(0.15)	4.14(1.99)
		LASSO	14.07(0.64)	61.83(18.46)	0.32(0.07)	0.41(0.06)	8.74(1.70)

Table B.6: Evaluation of heterogeneous model under AR-1 correlation, (n, p)=(400,1600) with 100 replicates.

$n = 400$	$p = 1600$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	0.87(1.50)	0.97(0.04)	0.97(0.04)	1.07(0.18)
		QLASSO	15.00(0.00)	34.53(7.38)	0.47(0.06)	0.55(0.05)	3.39(0.39)
		ssLASSO	15.00(0.00)	0.07(0.37)	1.00(0.01)	1.00(0.01)	1.26(0.20)
		LASSO	15.00(0.00)	67.17(21.86)	0.32(0.07)	0.43(0.06)	4.61(0.78)
	Error2	ssQLASSO	14.10(1.75)	0.23(0.43)	0.96(0.07)	0.96(0.07)	1.89(1.33)
		QLASSO	15.00(0.00)	33.00(4.53)	0.48(0.03)	0.55(0.03)	5.29(1.02)
		ssLASSO	6.90(2.81)	0.07(0.25)	0.61(0.17)	0.66(0.14)	7.10(1.56)
		LASSO	12.60(4.02)	64.83(33.67)	0.28(0.09)	0.37(0.10)	11.27(2.99)
	Error3	ssQLASSO	15.00(0.00)	0.43(0.77)	0.99(0.02)	0.99(0.02)	0.85(0.22)
		QLASSO	15.00(0.00)	29.37(8.03)	0.51(0.07)	0.58(0.05)	2.74(0.61)
		ssLASSO	10.03(3.18)	0.07(0.25)	0.78(0.16)	0.80(0.13)	5.52(2.05)
		LASSO	14.67(0.99)	71.03(20.69)	0.30(0.06)	0.41(0.05)	9.30(2.05)
	Error4	ssQLASSO	13.73(1.82)	0.27(0.78)	0.94(0.08)	0.95(0.07)	2.43(1.44)
		QLASSO	14.97(0.18)	33.30(4.63)	0.48(0.03)	0.55(0.03)	5.56(0.90)
		ssLASSO	8.07(2.84)	0.10(0.31)	0.68(0.16)	0.72(0.13)	6.50(1.83)
		LASSO	13.93(0.91)	64.50(32.13)	0.32(0.08)	0.41(0.07)	10.11(2.15)
	Error5	ssQLASSO	14.30(1.39)	0.33(0.55)	0.96(0.06)	0.96(0.05)	1.95(1.07)
		QLASSO	14.93(0.25)	33.23(4.88)	0.48(0.04)	0.55(0.03)	6.26(1.04)
		ssLASSO	8.17(4.32)	0.37(0.85)	0.65(0.24)	0.70(0.20)	6.58(2.51)
		LASSO	14.83(0.38)	78.63(34.84)	0.29(0.07)	0.40(0.06)	10.17(2.93)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	1.10(1.60)	0.97(0.05)	0.97(0.04)	0.70(0.19)
		QLASSO	15.00(0.00)	33.40(7.30)	0.48(0.05)	0.56(0.04)	3.55(0.55)
		ssLASSO	15.00(0.00)	0.07(0.25)	1.00(0.01)	1.00(0.01)	0.88(0.15)
		LASSO	15.00(0.00)	68.10(24.09)	0.32(0.08)	0.43(0.07)	4.71(0.95)
	Error2	ssQLASSO	13.67(1.97)	0.40(1.35)	0.94(0.09)	0.94(0.08)	2.03(1.65)
		QLASSO	15.00(0.00)	50.37(6.98)	0.38(0.03)	0.47(0.03)	4.76(0.75)
		ssLASSO	6.80(3.90)	0.13(0.35)	0.58(0.25)	0.63(0.21)	6.86(2.38)
		LASSO	12.83(3.43)	60.00(30.17)	0.31(0.10)	0.39(0.10)	10.85(3.23)
	Error3	ssQLASSO	15.00(0.00)	0.57(1.22)	0.98(0.04)	0.98(0.03)	1.61(0.25)
		QLASSO	15.00(0.00)	41.83(11.11)	0.43(0.07)	0.51(0.05)	4.05(0.61)
		ssLASSO	9.53(2.70)	5.43(28.82)	0.73(0.16)	0.76(0.14)	8.06(7.11)
		LASSO	14.63(0.61)	74.17(21.59)	0.29(0.06)	0.40(0.05)	10.43(2.06)
	Error4	ssQLASSO	13.17(2.21)	0.40(0.72)	0.92(0.09)	0.92(0.09)	2.44(1.66)
		QLASSO	15.00(0.00)	48.30(6.35)	0.39(0.03)	0.48(0.02)	5.37(0.74)
		ssLASSO	5.87(3.40)	0.23(0.68)	0.52(0.21)	0.59(0.17)	7.43(1.96)
		LASSO	14.20(1.13)	61.13(24.83)	0.34(0.09)	0.43(0.07)	10.37(2.05)
	Error5	ssQLASSO	13.90(1.37)	0.27(0.45)	0.95(0.06)	0.95(0.06)	1.97(1.03)
		QLASSO	15.00(0.00)	50.67(4.70)	0.37(0.02)	0.47(0.02)	5.28(0.89)
		ssLASSO	6.93(3.35)	0.23(0.50)	0.59(0.22)	0.64(0.18)	6.47(1.79)
		LASSO	14.73(0.45)	66.93(23.52)	0.32(0.08)	0.43(0.07)	9.08(1.69)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	0.43(0.82)	0.99(0.03)	0.99(0.02)	1.16(0.21)
		QLASSO	15.00(0.00)	35.57(7.10)	0.46(0.05)	0.54(0.04)	3.84(0.46)
		ssLASSO	15.00(0.00)	0.03(0.18)	1.00(0.01)	1.00(0.01)	1.36(0.21)
		LASSO	15.00(0.00)	79.63(25.17)	0.29(0.07)	0.40(0.06)	5.47(1.06)
	Error2	ssQLASSO	12.63(2.98)	0.20(0.48)	0.89(0.18)	0.92(0.07)	2.88(1.82)
		QLASSO	14.60(0.50)	34.50(5.43)	0.46(0.04)	0.53(0.03)	6.29(1.13)
		ssLASSO	5.73(2.36)	0.17(0.38)	0.53(0.17)	0.59(0.14)	7.39(1.64)
		LASSO	10.33(4.96)	51.33(39.07)	0.26(0.12)	0.35(0.12)	10.85(3.25)
	Error3	ssQLASSO	14.03(0.76)	0.17(0.38)	0.96(0.03)	0.96(0.03)	1.67(0.56)
		QLASSO	14.90(0.31)	34.73(4.46)	0.46(0.04)	0.54(0.03)	5.99(1.04)
		ssLASSO	8.67(3.11)	0.13(0.43)	0.71(0.18)	0.74(0.15)	5.56(1.87)
		LASSO	14.33(1.67)	70.63(22.30)	0.30(0.07)	0.40(0.06)	9.49(1.83)
	Error4	ssQLASSO	12.67(1.63)	0.13(0.51)	0.91(0.07)	0.91(0.06)	2.61(1.03)
		QLASSO	14.37(0.56)	34.20(4.10)	0.45(0.03)	0.53(0.03)	6.50(0.88)
		ssLASSO	6.70(2.55)	0.30(0.53)	0.59(0.15)	0.64(0.13)	6.70(1.59)
		LASSO	13.43(0.63)	61.30(30.48)	0.33(0.09)	0.41(0.08)	10.16(2.59)
	Error5	ssQLASSO	12.93(1.51)	0.97(1.50)	0.89(0.08)	0.90(0.07)	2.79(1.22)
		QLASSO	14.00(0.69)	34.67(6.20)	0.44(0.05)	0.51(0.04)	7.18(1.21)
		ssLASSO	8.33(3.27)	0.37(0.85)	0.68(0.19)	0.71(0.16)	5.71(1.86)
		LASSO	13.47(0.94)	63.27(25.19)	0.32(0.09)	0.40(0.08)	9.45(1.94)

Table B.7: Evaluation of heterogeneous model under banded correlation, (n, p)=(400,800) with 100 replicates.

$n = 400$	$p = 800$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	0.53(0.73)	0.98(0.02)	0.98(0.02)	1.14(0.20)
		QLASSO	15.00(0.00)	17.27(3.90)	0.64(0.05)	0.68(0.04)	3.28(0.48)
		ssLASSO	15.00(0.00)	0.03(0.18)	1.00(0.01)	1.00(0.01)	1.38(0.24)
		LASSO	15.00(0.00)	67.00(17.35)	0.32(0.06)	0.42(0.05)	4.71(0.90)
	Error2	ssQLASSO	14.97(0.18)	0.27(0.78)	0.99(0.02)	0.99(0.02)	1.24(0.38)
		QLASSO	15.00(0.00)	19.10(4.48)	0.62(0.05)	0.66(0.04)	4.68(0.83)
		ssLASSO	11.50(3.10)	0.30(0.47)	0.84(0.15)	0.85(0.13)	4.51(1.76)
		LASSO	13.43(2.71)	52.00(19.63)	0.34(0.06)	0.42(0.06)	9.42(1.80)
	Error3	ssQLASSO	15.00(0.00)	0.50(1.04)	0.98(0.03)	0.98(0.03)	0.87(0.17)
		QLASSO	15.00(0.00)	18.80(5.28)	0.62(0.08)	0.67(0.06)	2.44(0.43)
		ssLASSO	14.23(1.70)	0.03(0.18)	0.97(0.07)	0.97(0.06)	3.02(1.31)
		LASSO	14.83(0.75)	61.23(21.67)	0.34(0.09)	0.44(0.07)	8.39(2.00)
	Error4	ssQLASSO	14.67(0.71)	0.23(0.63)	0.98(0.04)	0.98(0.04)	1.68(0.87)
		QLASSO	14.70(0.70)	17.80(3.12)	0.62(0.05)	0.66(0.04)	5.07(1.08)
		ssLASSO	11.10(2.75)	0.23(0.50)	0.83(0.13)	0.84(0.12)	4.87(1.69)
		LASSO	14.10(1.09)	50.60(18.81)	0.37(0.08)	0.45(0.07)	9.28(2.00)
	Error5	ssQLASSO	14.87(0.35)	0.53(0.73)	0.98(0.03)	0.98(0.03)	1.58(0.55)
		QLASSO	14.80(0.41)	20.43(4.01)	0.59(0.05)	0.64(0.04)	5.84(1.01)
		ssLASSO	11.63(2.41)	0.70(1.09)	0.84(0.10)	0.85(0.10)	4.44(1.44)
		LASSO	14.93(0.37)	61.73(25.32)	0.35(0.09)	0.44(0.08)	8.83(2.09)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	0.60(0.93)	0.98(0.03)	0.98(0.03)	0.61(0.15)
		QLASSO	15.00(0.00)	19.53(5.30)	0.61(0.06)	0.66(0.05)	3.27(0.60)
		ssLASSO	15.00(0.00)	0.10(0.40)	1.00(0.01)	1.00(0.01)	0.90(0.22)
		LASSO	15.00(0.00)	60.93(23.20)	0.35(0.07)	0.44(0.06)	4.26(1.01)
	Error2	ssQLASSO	14.63(1.10)	0.30(0.60)	0.98(0.05)	0.98(0.05)	1.26(1.04)
		QLASSO	15.00(0.00)	29.43(4.68)	0.51(0.04)	0.57(0.03)	4.16(0.79)
		ssLASSO	11.20(2.34)	0.30(0.53)	0.84(0.11)	0.85(0.10)	4.25(1.57)
		LASSO	12.93(2.91)	54.27(27.05)	0.33(0.08)	0.40(0.07)	10.21(2.44)
	Error3	ssQLASSO	15.00(0.00)	0.20(0.61)	0.99(0.02)	0.99(0.02)	1.55(0.25)
		QLASSO	15.00(0.00)	25.20(5.68)	0.55(0.06)	0.61(0.05)	3.58(0.54)
		ssLASSO	13.10(2.86)	0.13(0.43)	0.92(0.13)	0.92(0.12)	4.43(2.07)
		LASSO	14.07(2.82)	52.63(20.73)	0.35(0.08)	0.43(0.08)	9.12(2.32)
	Error4	ssQLASSO	14.57(1.04)	0.70(1.58)	0.96(0.06)	0.96(0.06)	1.57(1.15)
		QLASSO	15.00(0.00)	30.03(4.34)	0.50(0.04)	0.57(0.03)	4.46(0.76)
		ssLASSO	10.97(2.51)	0.53(0.90)	0.82(0.11)	0.83(0.10)	4.39(1.44)
		LASSO	14.33(0.88)	58.83(22.65)	0.34(0.08)	0.43(0.06)	9.90(2.08)
	Error5	ssQLASSO	14.67(0.88)	0.67(1.18)	0.97(0.04)	0.97(0.04)	1.40(0.93)
		QLASSO	15.00(0.00)	28.00(5.65)	0.52(0.05)	0.58(0.04)	4.61(0.93)
		ssLASSO	11.30(3.01)	0.37(0.72)	0.83(0.14)	0.85(0.12)	3.90(1.77)
		LASSO	14.83(0.38)	55.63(16.72)	0.36(0.07)	0.45(0.06)	8.02(1.71)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	0.50(0.78)	0.98(0.02)	0.98(0.02)	1.16(0.21)
		QLASSO	15.00(0.00)	20.37(3.81)	0.60(0.04)	0.65(0.04)	3.74(0.70)
		ssLASSO	15.00(0.00)	0.00(0.00)	1.00(0.00)	1.00(0.00)	1.33(0.19)
		LASSO	15.00(0.00)	62.90(25.24)	0.35(0.09)	0.44(0.08)	4.85(1.10)
	Error2	ssQLASSO	13.90(0.55)	0.20(0.41)	0.96(0.02)	0.96(0.02)	1.68(0.58)
		QLASSO	14.43(0.57)	19.97(4.73)	0.59(0.06)	0.63(0.05)	5.38(1.24)
		ssLASSO	10.83(3.30)	0.10(0.31)	0.81(0.18)	0.83(0.15)	4.32(1.98)
		LASSO	11.67(4.36)	41.40(23.15)	0.34(0.12)	0.42(0.08)	9.24(2.16)
	Error3	ssQLASSO	14.07(0.37)	0.13(0.43)	0.96(0.02)	0.96(0.02)	1.44(0.34)
		QLASSO	14.63(0.49)	19.63(3.92)	0.60(0.05)	0.64(0.05)	5.06(0.83)
		ssLASSO	12.20(2.63)	0.10(0.31)	0.88(0.13)	0.89(0.11)	3.20(1.74)
		LASSO	14.63(0.67)	56.53(18.70)	0.35(0.07)	0.44(0.06)	7.76(1.51)
	Error4	ssQLASSO	13.67(1.18)	0.37(0.96)	0.94(0.06)	0.94(0.05)	2.10(1.08)
		QLASSO	14.27(0.58)	19.50(5.03)	0.59(0.06)	0.63(0.05)	6.39(1.06)
		ssLASSO	9.37(2.58)	0.17(0.38)	0.75(0.13)	0.77(0.11)	5.27(1.47)
		LASSO	13.33(1.27)	53.97(20.59)	0.34(0.07)	0.41(0.06)	10.33(1.82)
	Error5	ssQLASSO	13.93(0.87)	0.43(0.68)	0.95(0.04)	0.95(0.04)	2.13(0.76)
		QLASSO	14.40(0.67)	19.33(3.62)	0.59(0.05)	0.63(0.05)	6.43(0.94)
		ssLASSO	12.57(1.89)	0.87(1.50)	0.88(0.08)	0.88(0.08)	3.63(1.31)
		LASSO	14.07(0.52)	54.37(18.04)	0.35(0.08)	0.43(0.07)	8.62(1.66)

Table B.8: Evaluation of heterogeneous model under banded correlation, (n, p)=(400,1600) with 100 replicates.

$n = 400$	$p = 1600$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	0.73(1.17)	0.98(0.04)	0.98(0.03)	1.07(0.26)
		QLASSO	15.00(0.00)	33.73(6.06)	0.47(0.05)	0.55(0.04)	3.47(0.48)
		ssLASSO	15.00(0.00)	0.07(0.25)	1.00(0.01)	1.00(0.01)	1.27(0.22)
		LASSO	15.00(0.00)	71.73(21.72)	0.31(0.07)	0.42(0.06)	5.02(0.93)
	Error2	ssQLASSO	13.80(1.75)	0.10(0.31)	0.95(0.07)	0.95(0.06)	2.01(1.32)
		QLASSO	14.97(0.18)	32.93(4.49)	0.48(0.03)	0.55(0.03)	5.34(1.01)
		ssLASSO	6.60(2.81)	0.00(0.00)	0.59(0.18)	0.65(0.15)	7.21(1.66)
		LASSO	12.90(3.11)	64.20(32.42)	0.30(0.09)	0.38(0.09)	10.89(2.63)
	Error3	ssQLASSO	15.00(0.00)	0.93(2.10)	0.97(0.06)	0.97(0.05)	0.88(0.24)
		QLASSO	15.00(0.00)	25.40(8.74)	0.55(0.08)	0.61(0.07)	2.63(0.50)
		ssLASSO	9.20(3.34)	0.13(0.35)	0.73(0.18)	0.76(0.15)	6.11(2.11)
		LASSO	14.30(1.32)	65.10(26.78)	0.32(0.08)	0.42(0.07)	9.52(2.99)
	Error4	ssQLASSO	13.07(2.05)	0.37(0.67)	0.91(0.09)	0.92(0.08)	2.90(1.43)
		QLASSO	14.80(0.48)	33.37(4.21)	0.47(0.03)	0.55(0.02)	6.33(0.90)
		ssLASSO	6.13(3.26)	0.13(0.43)	0.55(0.21)	0.61(0.17)	7.39(1.93)
		LASSO	13.90(1.03)	60.23(24.49)	0.34(0.10)	0.43(0.07)	10.20(2.03)
	Error5	ssQLASSO	14.17(1.58)	0.40(0.56)	0.96(0.06)	0.96(0.06)	2.18(1.23)
		QLASSO	14.77(0.43)	33.43(3.86)	0.47(0.03)	0.54(0.03)	6.06(1.16)
		ssLASSO	8.27(3.71)	0.20(0.48)	0.67(0.21)	0.71(0.17)	6.34(2.13)
		LASSO	14.80(0.48)	72.53(31.36)	0.31(0.08)	0.42(0.06)	9.68(2.51)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	1.07(1.44)	0.97(0.04)	0.97(0.04)	0.69(0.17)
		QLASSO	15.00(0.00)	34.17(4.90)	0.47(0.04)	0.55(0.03)	3.16(0.46)
		ssLASSO	15.00(0.00)	0.03(0.18)	1.00(0.01)	1.00(0.01)	0.88(0.18)
		LASSO	15.00(0.00)	75.43(22.94)	0.30(0.06)	0.41(0.05)	4.64(1.03)
	Error2	ssQLASSO	14.17(1.80)	0.33(0.80)	0.96(0.08)	0.96(0.08)	1.74(1.57)
		QLASSO	15.00(0.00)	49.80(7.64)	0.38(0.04)	0.48(0.03)	4.68(0.95)
		ssLASSO	6.93(3.32)	0.13(0.35)	0.60(0.20)	0.65(0.17)	6.85(2.01)
		LASSO	12.17(4.21)	64.60(38.92)	0.28(0.10)	0.37(0.09)	11.62(3.85)
	Error3	ssQLASSO	14.87(0.57)	0.80(2.02)	0.97(0.06)	0.97(0.05)	1.80(0.58)
		QLASSO	15.00(0.00)	44.27(12.30)	0.42(0.08)	0.51(0.06)	4.31(0.70)
		ssLASSO	8.33(3.57)	0.10(0.31)	0.68(0.20)	0.72(0.17)	7.13(2.29)
		LASSO	13.97(2.06)	66.40(25.76)	0.31(0.07)	0.40(0.06)	10.17(2.32)
	Error4	ssQLASSO	12.27(2.66)	0.73(1.41)	0.87(0.12)	0.87(0.11)	3.29(1.83)
		QLASSO	15.00(0.00)	51.73(6.33)	0.37(0.03)	0.47(0.02)	5.62(0.61)
		ssLASSO	5.63(3.24)	0.23(0.50)	0.51(0.23)	0.59(0.18)	7.57(1.90)
		LASSO	13.87(1.22)	66.03(30.61)	0.31(0.07)	0.41(0.06)	10.80(2.43)
	Error5	ssQLASSO	13.70(1.32)	0.40(0.72)	0.94(0.06)	0.94(0.05)	2.08(1.18)
		QLASSO	15.00(0.00)	50.50(5.94)	0.37(0.03)	0.47(0.02)	5.42(0.86)
		ssLASSO	8.33(3.82)	0.20(0.48)	0.67(0.23)	0.71(0.20)	5.65(2.30)
		LASSO	14.60(0.86)	66.13(23.30)	0.32(0.07)	0.42(0.06)	9.01(1.68)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	0.80(1.13)	0.98(0.03)	0.98(0.03)	1.15(0.19)
		QLASSO	15.00(0.00)	35.33(5.34)	0.46(0.04)	0.54(0.03)	4.14(0.53)
		ssLASSO	15.00(0.00)	0.13(0.35)	1.00(0.01)	1.00(0.01)	1.38(0.23)
		LASSO	15.00(0.00)	75.83(26.12)	0.30(0.07)	0.41(0.06)	5.47(1.03)
	Error2	ssQLASSO	12.73(2.29)	0.47(0.73)	0.90(0.10)	0.90(0.09)	3.01(1.55)
		QLASSO	14.77(0.50)	33.77(5.73)	0.47(0.04)	0.54(0.04)	6.39(1.02)
		ssLASSO	6.27(2.79)	0.27(0.45)	0.56(0.18)	0.62(0.15)	7.20(1.66)
		LASSO	11.63(3.50)	58.83(31.39)	0.28(0.09)	0.35(0.10)	11.34(2.66)
	Error3	ssQLASSO	13.83(0.59)	0.10(0.40)	0.96(0.03)	0.96(0.03)	1.64(0.67)
		QLASSO	14.53(0.57)	33.83(4.02)	0.46(0.04)	0.53(0.03)	6.24(1.20)
		ssLASSO	7.07(3.15)	0.03(0.18)	0.61(0.20)	0.67(0.16)	6.17(1.95)
		LASSO	13.87(2.24)	73.47(32.26)	0.29(0.09)	0.39(0.08)	10.13(2.69)
	Error4	ssQLASSO	12.70(1.78)	0.30(0.53)	0.90(0.08)	0.91(0.07)	2.61(1.20)
		QLASSO	14.40(0.56)	33.23(4.98)	0.46(0.04)	0.53(0.04)	6.26(1.05)
		ssLASSO	7.17(2.98)	0.13(0.35)	0.62(0.18)	0.67(0.15)	6.40(1.90)
		LASSO	13.50(1.01)	67.80(28.11)	0.31(0.09)	0.40(0.07)	10.65(2.22)
	Error5	ssQLASSO	12.87(1.70)	0.27(0.52)	0.91(0.07)	0.91(0.07)	2.55(1.26)
		QLASSO	14.10(0.66)	34.27(6.38)	0.45(0.05)	0.52(0.04)	6.47(1.20)
		ssLASSO	7.83(3.64)	0.23(0.63)	0.65(0.21)	0.69(0.17)	5.95(2.06)
		LASSO	13.77(0.68)	69.23(27.82)	0.30(0.09)	0.40(0.07)	9.36(2.06)

Table B.9: Evaluation of homogeneous model under AR-1 correlation with 8 positive and 7 negative coefficients, (n, p)=(400,1600).

$n = 400$	$p = 1600$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	1.27(1.91)	0.96(0.05)	0.96(0.05)	0.89(0.22)
		QLASSO	15.00(0.00)	34.20(6.01)	0.47(0.04)	0.55(0.03)	3.27(0.39)
		ssLASSO	15.00(0.00)	0.40(0.86)	0.99(0.03)	0.99(0.03)	0.64(0.12)
		LASSO	15.00(0.00)	71.40(30.91)	0.32(0.09)	0.42(0.07)	3.39(0.84)
	Error2	ssQLASSO	14.40(1.10)	0.70(1.12)	0.96(0.05)	0.96(0.05)	3.42(1.36)
		QLASSO	15.00(0.00)	34.13(3.82)	0.47(0.03)	0.55(0.02)	5.82(0.98)
		ssLASSO	7.63(3.55)	0.17(0.38)	0.64(0.22)	0.68(0.18)	6.12(2.20)
		LASSO	13.27(4.14)	63.27(28.64)	0.29(0.09)	0.40(0.08)	8.85(2.51)
	Error3	ssQLASSO	14.97(0.18)	0.10(0.31)	1.00(0.01)	1.00(0.01)	2.58(1.19)
		QLASSO	15.00(0.00)	29.33(6.84)	0.51(0.07)	0.58(0.05)	2.72(0.44)
		ssLASSO	11.17(3.07)	0.00(0.00)	0.84(0.15)	0.85(0.13)	3.75(2.02)
		LASSO	14.97(0.18)	71.37(28.05)	0.31(0.07)	0.42(0.06)	6.98(1.84)
	Error4	ssQLASSO	14.50(0.86)	0.70(1.37)	0.96(0.05)	0.96(0.04)	3.62(1.33)
		QLASSO	14.87(0.35)	33.37(6.24)	0.47(0.05)	0.55(0.04)	6.70(1.06)
		ssLASSO	6.30(3.58)	0.10(0.31)	0.55(0.23)	0.63(0.17)	6.76(2.20)
		LASSO	14.97(0.18)	71.47(25.64)	0.31(0.08)	0.42(0.07)	8.07(1.85)
	Error5	ssQLASSO	14.83(0.38)	1.07(2.07)	0.96(0.05)	0.96(0.05)	3.23(1.56)
		QLASSO	14.90(0.31)	31.73(5.37)	0.49(0.05)	0.56(0.04)	6.63(1.06)
		ssLASSO	8.20(4.48)	0.17(0.53)	0.65(0.25)	0.70(0.21)	5.86(2.82)
		LASSO	15.00(0.00)	74.43(26.20)	0.31(0.08)	0.41(0.07)	6.93(1.41)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	1.33(2.17)	0.96(0.06)	0.96(0.06)	0.84(0.25)
		QLASSO	15.00(0.00)	33.70(5.89)	0.48(0.05)	0.55(0.04)	3.23(0.38)
		ssLASSO	15.00(0.00)	0.60(0.86)	0.98(0.03)	0.98(0.03)	0.72(0.19)
		LASSO	15.00(0.00)	71.47(19.52)	0.31(0.05)	0.41(0.05)	3.39(0.57)
	Error2	ssQLASSO	14.63(0.61)	0.90(0.96)	0.96(0.03)	0.96(0.03)	4.86(0.91)
		QLASSO	15.00(0.00)	51.50(6.34)	0.37(0.03)	0.47(0.02)	4.97(0.80)
		ssLASSO	8.57(3.51)	0.20(0.48)	0.69(0.20)	0.73(0.17)	5.64(2.10)
		LASSO	13.57(3.71)	71.87(29.57)	0.27(0.06)	0.37(0.07)	9.19(1.64)
	Error3	ssQLASSO	14.97(0.18)	0.50(1.04)	0.98(0.03)	0.98(0.03)	3.90(0.82)
		QLASSO	15.00(0.00)	50.37(8.68)	0.38(0.05)	0.47(0.04)	3.90(0.73)
		ssLASSO	10.37(3.55)	0.03(0.18)	0.79(0.18)	0.82(0.15)	4.08(2.22)
		LASSO	14.67(1.47)	78.40(29.26)	0.29(0.07)	0.39(0.06)	7.37(2.23)
	Error4	ssQLASSO	14.57(0.90)	1.30(1.37)	0.94(0.04)	0.95(0.04)	4.77(0.89)
		QLASSO	15.00(0.00)	52.93(5.88)	0.36(0.03)	0.46(0.02)	5.66(0.67)
		ssLASSO	7.40(3.37)	0.23(0.77)	0.62(0.19)	0.67(0.16)	6.28(2.08)
		LASSO	15.00(0.00)	73.27(17.54)	0.30(0.05)	0.41(0.04)	7.96(1.38)
	Error5	ssQLASSO	14.90(0.40)	1.23(1.22)	0.96(0.04)	0.96(0.04)	4.07(0.85)
		QLASSO	15.00(0.00)	52.00(5.63)	0.37(0.03)	0.47(0.02)	6.00(0.79)
		ssLASSO	10.73(3.49)	1.07(1.89)	0.78(0.16)	0.80(0.14)	4.34(2.05)
		LASSO	15.00(0.00)	74.80(21.97)	0.30(0.06)	0.41(0.05)	6.91(1.15)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	1.53(1.85)	0.95(0.05)	0.96(0.05)	0.84(0.19)
		QLASSO	15.00(0.00)	36.47(6.77)	0.46(0.04)	0.54(0.03)	3.31(0.36)
		ssLASSO	15.00(0.00)	0.63(0.81)	0.98(0.03)	0.98(0.02)	0.60(0.20)
		LASSO	15.00(0.00)	77.47(21.93)	0.29(0.06)	0.40(0.05)	3.57(0.67)
	Error2	ssQLASSO	14.63(0.61)	0.43(0.77)	0.97(0.03)	0.97(0.03)	3.63(1.42)
		QLASSO	15.00(0.00)	33.03(4.77)	0.48(0.04)	0.55(0.03)	5.61(0.68)
		ssLASSO	7.70(2.81)	0.13(0.35)	0.66(0.15)	0.70(0.13)	6.10(1.69)
		LASSO	14.37(1.77)	63.00(25.64)	0.33(0.08)	0.43(0.07)	8.27(2.00)
	Error3	ssQLASSO	14.97(0.18)	0.17(0.38)	0.99(0.01)	0.99(0.01)	2.75(0.92)
		QLASSO	14.90(0.40)	33.17(5.17)	0.48(0.04)	0.55(0.03)	5.92(0.86)
		ssLASSO	11.10(3.56)	0.00(0.00)	0.83(0.17)	0.85(0.15)	3.69(2.15)
		LASSO	15.00(0.00)	77.60(29.32)	0.30(0.07)	0.41(0.06)	7.12(2.44)
	Error4	ssQLASSO	14.00(1.62)	1.70(3.24)	0.92(0.09)	0.92(0.08)	3.56(1.29)
		QLASSO	14.80(0.48)	31.87(4.86)	0.48(0.05)	0.56(0.04)	6.68(0.91)
		ssLASSO	7.27(3.03)	0.40(0.81)	0.62(0.16)	0.67(0.13)	6.23(1.58)
		LASSO	14.90(0.31)	73.37(24.48)	0.31(0.08)	0.41(0.07)	8.02(1.38)
	Error5	ssQLASSO	14.70(0.65)	1.13(1.48)	0.95(0.04)	0.96(0.04)	3.52(1.21)
		QLASSO	14.73(0.45)	33.00(4.90)	0.47(0.04)	0.55(0.03)	6.36(1.09)
		ssLASSO	9.90(4.12)	0.60(1.10)	0.74(0.21)	0.77(0.18)	4.61(2.46)
		LASSO	15.00(0.00)	75.97(26.09)	0.30(0.07)	0.41(0.06)	6.86(1.47)

Table B.10: Evaluation of homogeneous model under real data correlation, (n, p)=(400,1600) with 100 replicates.

$n = 400$	$p = 1600$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	3.28(1.93)	0.90(0.05)	0.91(0.05)	1.84(0.59)
		QLASSO	14.50(0.76)	25.75(6.94)	0.53(0.07)	0.59(0.07)	6.85(1.25)
		ssLASSO	4.67(2.20)	0.44(0.51)	0.45(0.16)	0.52(0.13)	8.35(1.32)
		LASSO	15.00(0.00)	57.40(9.60)	0.34(0.04)	0.44(0.03)	4.55(0.65)
	Error2	ssQLASSO	12.78(1.80)	3.28(1.23)	0.82(0.08)	0.82(0.08)	4.35(1.57)
		QLASSO	11.90(1.55)	23.65(4.39)	0.47(0.06)	0.50(0.07)	8.52(1.22)
		ssLASSO	4.78(2.69)	0.94(0.94)	0.44(0.19)	0.49(0.16)	8.58(1.19)
		LASSO	12.60(2.60)	53.45(16.08)	0.32(0.07)	0.40(0.12)	10.34(2.14)
	Error3	ssQLASSO	14.89(0.32)	3.61(1.65)	0.89(0.05)	0.89(0.05)	2.29(1.18)
		QLASSO	14.95(0.22)	23.65(5.01)	0.57(0.05)	0.61(0.05)	4.42(1.12)
		ssLASSO	5.39(2.95)	0.44(0.86)	0.49(0.21)	0.55(0.17)	7.93(1.64)
		LASSO	14.35(1.31)	58.00(13.51)	0.33(0.05)	0.42(0.06)	8.94(1.63)
	Error4	ssQLASSO	12.89(1.37)	2.72(1.45)	0.84(0.08)	0.84(0.08)	4.23(1.42)
		QLASSO	12.10(1.77)	23.95(6.99)	0.47(0.08)	0.51(0.10)	8.75(1.57)
		ssLASSO	6.28(2.42)	1.06(0.94)	0.54(0.15)	0.59(0.12)	8.28(1.04)
		LASSO	13.80(1.15)	54.95(12.08)	0.34(0.04)	0.42(0.04)	9.65(1.07)
	Error5	ssQLASSO	13.22(1.86)	3.00(1.78)	0.85(0.09)	0.85(0.09)	4.07(1.75)
		QLASSO	10.80(1.61)	23.95(5.32)	0.44(0.07)	0.46(0.08)	9.75(1.04)
		ssLASSO	6.17(2.50)	0.83(1.04)	0.54(0.17)	0.59(0.14)	8.12(1.22)
		LASSO	14.20(1.06)	59.75(14.78)	0.32(0.05)	0.41(0.05)	8.94(1.54)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	2.89(2.54)	0.92(0.06)	0.92(0.05)	1.58(0.49)
		QLASSO	14.65(0.49)	31.05(7.27)	0.50(0.06)	0.55(0.05)	6.39(1.13)
		ssLASSO	4.78(2.10)	0.94(1.06)	0.45(0.14)	0.51(0.12)	8.82(1.51)
		LASSO	14.95(0.22)	58.55(13.34)	0.35(0.06)	0.44(0.05)	4.70(0.76)
	Error2	ssQLASSO	13.83(1.34)	2.78(1.63)	0.88(0.08)	0.88(0.08)	3.66(1.73)
		QLASSO	14.05(1.10)	29.15(5.18)	0.48(0.05)	0.54(0.06)	7.27(1.04)
		ssLASSO	5.22(1.90)	1.28(1.07)	0.47(0.12)	0.52(0.09)	8.61(0.87)
		LASSO	13.05(2.01)	53.30(12.60)	0.32(0.05)	0.40(0.08)	10.25(1.90)
	Error3	ssQLASSO	14.33(0.84)	4.06(1.70)	0.86(0.06)	0.86(0.06)	2.66(1.18)
		QLASSO	14.60(0.50)	29.50(6.45)	0.50(0.06)	0.56(0.05)	6.14(1.45)
		ssLASSO	4.78(2.51)	1.17(1.42)	0.43(0.16)	0.49(0.13)	9.19(1.10)
		LASSO	13.65(1.18)	54.50(11.02)	0.34(0.05)	0.41(0.05)	9.17(2.04)
	Error4	ssQLASSO	13.44(1.50)	2.56(1.72)	0.87(0.09)	0.87(0.09)	3.59(1.81)
		QLASSO	13.90(1.12)	25.60(5.20)	0.52(0.05)	0.57(0.06)	7.45(1.52)
		ssLASSO	5.50(2.46)	0.83(0.99)	0.50(0.17)	0.55(0.14)	8.34(1.51)
		LASSO	14.15(1.04)	54.00(12.23)	0.35(0.06)	0.43(0.05)	9.21(1.82)
	Error5	ssQLASSO	13.89(1.02)	3.22(1.52)	0.87(0.06)	0.87(0.06)	3.85(1.21)
		QLASSO	13.35(1.04)	27.00(3.73)	0.48(0.04)	0.53(0.04)	8.42(1.10)
		ssLASSO	5.78(2.71)	0.67(0.69)	0.52(0.18)	0.57(0.14)	8.30(1.57)
		LASSO	14.70(0.47)	55.40(16.62)	0.36(0.06)	0.45(0.05)	8.71(1.38)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	2.94(1.06)	0.91(0.03)	0.91(0.03)	1.91(0.55)
		QLASSO	14.35(1.09)	25.45(4.47)	0.53(0.06)	0.57(0.06)	6.53(1.40)
		ssLASSO	5.33(2.40)	0.33(0.49)	0.50(0.17)	0.56(0.14)	8.01(1.57)
		LASSO	15.00(0.00)	56.15(11.04)	0.35(0.04)	0.45(0.04)	4.50(0.73)
	Error2	ssQLASSO	14.06(1.06)	2.61(1.65)	0.89(0.06)	0.89(0.06)	3.32(1.53)
		QLASSO	13.25(1.33)	22.95(4.22)	0.52(0.06)	0.55(0.06)	7.94(1.31)
		ssLASSO	5.06(2.80)	0.89(0.90)	0.46(0.17)	0.52(0.13)	8.50(1.31)
		LASSO	14.10(0.97)	51.70(14.90)	0.36(0.05)	0.44(0.06)	9.77(2.46)
	Error3	ssQLASSO	14.17(0.92)	3.00(1.61)	0.88(0.06)	0.88(0.07)	3.10(1.10)
		QLASSO	12.00(1.45)	24.25(4.55)	0.48(0.05)	0.50(0.07)	8.89(1.34)
		ssLASSO	3.67(2.52)	0.89(1.13)	0.35(0.18)	0.42(0.14)	9.73(0.47)
		LASSO	14.00(0.92)	55.05(11.63)	0.34(0.05)	0.42(0.04)	8.99(1.33)
	Error4	ssQLASSO	12.61(1.42)	3.39(0.92)	0.81(0.06)	0.81(0.06)	4.85(1.13)
		QLASSO	11.20(1.44)	21.95(3.94)	0.46(0.05)	0.49(0.08)	9.60(1.19)
		ssLASSO	4.28(1.56)	0.89(1.13)	0.41(0.11)	0.48(0.08)	9.45(0.98)
		LASSO	13.60(0.99)	57.50(10.10)	0.32(0.04)	0.41(0.04)	10.26(1.32)
	Error5	ssQLASSO	14.17(1.04)	2.56(1.15)	0.89(0.06)	0.89(0.06)	3.22(1.05)
		QLASSO	11.15(1.93)	22.30(4.47)	0.45(0.06)	0.48(0.09)	9.11(1.04)
		ssLASSO	5.67(2.68)	0.78(1.06)	0.51(0.17)	0.56(0.13)	8.37(1.19)
		LASSO	14.40(0.88)	59.35(8.52)	0.33(0.03)	0.42(0.04)	8.40(1.17)

Table B.11: Sensitivity evaluation of homogeneous model under AR-1 correlation, (n, p)=(400,1600) with 100 replicates.

$n = 400$	$p = 1600$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	1.70(2.09)	0.95(0.06)	0.95(0.05)	0.83(0.15)
$(a, b) = (1, 5)$		QLASSO	15.00(0.00)	35.33(7.75)	0.47(0.05)	0.54(0.04)	3.27(0.38)
		ssLASSO	15.00(0.00)	0.67(0.99)	0.98(0.03)	0.98(0.03)	0.65(0.20)
		LASSO	15.00(0.00)	75.33(23.92)	0.30(0.08)	0.41(0.06)	3.44(0.69)
	Error2	ssQLASSO	14.93(0.37)	0.27(0.78)	0.99(0.03)	0.99(0.03)	1.09(0.36)
		QLASSO	15.00(0.00)	32.03(3.51)	0.49(0.03)	0.56(0.02)	5.38(0.61)
		ssLASSO	8.60(3.47)	0.03(0.18)	0.70(0.20)	0.74(0.17)	5.52(2.20)
		LASSO	14.17(1.74)	66.53(22.74)	0.31(0.06)	0.41(0.05)	8.96(2.23)
	Error3	ssQLASSO	15.00(0.00)	0.50(1.33)	0.99(0.04)	0.99(0.03)	0.70(0.15)
		QLASSO	15.00(0.00)	28.17(6.57)	0.52(0.06)	0.59(0.05)	2.60(0.40)
		ssLASSO	9.47(3.35)	0.03(0.18)	0.75(0.17)	0.78(0.14)	4.68(2.17)
		LASSO	14.80(1.10)	65.47(24.03)	0.33(0.07)	0.43(0.06)	7.12(1.95)
	Error4	ssQLASSO	14.43(1.22)	0.40(0.77)	0.97(0.05)	0.97(0.05)	1.65(0.93)
		QLASSO	14.77(0.77)	33.20(3.75)	0.47(0.04)	0.55(0.04)	6.15(1.25)
		ssLASSO	7.23(3.19)	0.07(0.25)	0.62(0.20)	0.67(0.16)	6.07(1.96)
		LASSO	14.90(0.31)	74.20(27.88)	0.31(0.08)	0.41(0.07)	8.08(2.04)
	Error5	ssQLASSO	14.73(0.64)	0.73(1.23)	0.97(0.04)	0.97(0.04)	1.50(0.63)
		QLASSO	14.80(0.41)	33.37(6.29)	0.47(0.05)	0.55(0.04)	6.51(0.91)
		ssLASSO	7.93(3.89)	0.17(0.46)	0.65(0.22)	0.70(0.19)	5.61(2.44)
		LASSO	15.00(0.00)	73.90(25.79)	0.30(0.06)	0.41(0.06)	6.92(1.52)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	1.23(1.81)	0.96(0.05)	0.96(0.05)	0.81(0.20)
$(a, b) = (0.5, 3)$		QLASSO	15.00(0.00)	34.90(7.42)	0.47(0.06)	0.55(0.04)	3.15(0.41)
		ssLASSO	15.00(0.00)	11.00(55.54)	0.94(0.16)	0.95(0.15)	1.14(2.25)
		LASSO	15.00(0.00)	79.77(24.34)	0.29(0.06)	0.40(0.06)	3.55(0.78)
	Error2	ssQLASSO	14.77(0.77)	0.40(0.72)	0.98(0.04)	0.98(0.04)	1.28(0.69)
		QLASSO	15.00(0.00)	52.80(5.59)	0.36(0.02)	0.46(0.02)	5.12(0.73)
		ssLASSO	7.27(2.65)	0.17(0.38)	0.63(0.16)	0.67(0.14)	6.40(1.71)
		LASSO	14.37(1.47)	75.83(25.78)	0.29(0.06)	0.39(0.05)	8.92(1.92)
	Error3	ssQLASSO	15.00(0.00)	0.73(1.60)	0.98(0.04)	0.98(0.04)	0.85(0.17)
		QLASSO	15.00(0.00)	47.23(7.79)	0.39(0.04)	0.49(0.03)	3.67(0.62)
		ssLASSO	12.93(2.92)	0.30(1.47)	0.91(0.13)	0.91(0.11)	2.71(1.97)
		LASSO	15.00(0.00)	67.77(20.74)	0.32(0.06)	0.43(0.05)	6.39(1.49)
	Error4	ssQLASSO	14.60(0.77)	0.53(0.63)	0.97(0.03)	0.97(0.03)	1.62(0.70)
		QLASSO	15.00(0.00)	52.67(5.70)	0.36(0.03)	0.46(0.02)	5.42(0.63)
		ssLASSO	8.20(3.96)	0.27(0.52)	0.66(0.22)	0.70(0.18)	5.64(2.38)
		LASSO	14.97(0.18)	72.93(21.98)	0.30(0.07)	0.41(0.06)	8.00(1.73)
	Error5	ssQLASSO	14.53(1.48)	0.50(0.68)	0.97(0.07)	0.97(0.06)	1.65(1.09)
		QLASSO	14.97(0.18)	49.83(6.06)	0.38(0.03)	0.47(0.02)	5.95(0.90)
		ssLASSO	9.30(5.05)	0.70(1.74)	0.69(0.28)	0.73(0.23)	4.83(2.96)
		LASSO	15.00(0.00)	80.57(27.31)	0.29(0.07)	0.40(0.06)	7.19(1.67)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	1.73(2.70)	0.95(0.07)	0.95(0.07)	0.85(0.20)
$(a, b) = (5, 1)$		QLASSO	15.00(0.00)	36.67(6.10)	0.45(0.04)	0.54(0.03)	3.27(0.34)
		ssLASSO	15.00(0.00)	1.07(1.44)	0.97(0.04)	0.97(0.04)	0.70(0.17)
		LASSO	15.00(0.00)	77.00(25.80)	0.30(0.07)	0.41(0.06)	3.47(0.72)
	Error2	ssQLASSO	14.70(0.79)	0.77(1.17)	0.97(0.04)	0.97(0.04)	1.63(0.95)
		QLASSO	14.93(0.25)	32.83(4.24)	0.48(0.03)	0.55(0.03)	5.71(0.88)
		ssLASSO	8.47(2.62)	0.13(0.35)	0.70(0.15)	0.74(0.12)	5.60(1.80)
		LASSO	14.23(2.75)	68.90(25.53)	0.29(0.08)	0.41(0.05)	8.47(2.09)
	Error3	ssQLASSO	15.00(0.00)	0.20(0.61)	0.99(0.02)	0.99(0.02)	1.06(0.30)
		QLASSO	14.93(0.25)	33.67(4.82)	0.47(0.04)	0.55(0.03)	5.96(0.77)
		ssLASSO	10.30(3.05)	0.17(0.46)	0.79(0.14)	0.81(0.13)	4.24(1.93)
		LASSO	14.97(0.18)	74.83(31.91)	0.31(0.08)	0.41(0.07)	7.63(2.01)
	Error4	ssQLASSO	14.17(1.39)	0.27(0.58)	0.96(0.05)	0.96(0.05)	1.84(1.09)
		QLASSO	14.83(0.38)	33.13(4.73)	0.47(0.04)	0.55(0.03)	6.51(0.98)
		ssLASSO	7.43(2.94)	0.37(0.61)	0.63(0.17)	0.67(0.14)	6.05(1.75)
		LASSO	14.97(0.18)	76.13(24.27)	0.29(0.05)	0.40(0.05)	8.04(1.66)
	Error5	ssQLASSO	14.97(0.18)	1.07(1.74)	0.97(0.05)	0.97(0.04)	1.51(0.60)
		QLASSO	14.83(0.38)	32.93(4.99)	0.48(0.04)	0.55(0.03)	6.59(0.81)
		ssLASSO	9.67(3.90)	0.80(1.54)	0.73(0.19)	0.76(0.16)	4.78(2.26)
		LASSO	15.00(0.00)	75.97(40.22)	0.31(0.08)	0.42(0.07)	7.16(2.45)

Table B.12: Sensitivity evaluation of heterogeneous model under AR-1 correlation, (n, p)=(400,1600) with 100 replicates.

$n = 400$	$p = 1600$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	0.77(0.94)	0.98(0.03)	0.98(0.03)	1.07(0.19)
$(a, b) = (1, 5)$		QLASSO	15.00(0.00)	35.53(6.74)	0.46(0.05)	0.54(0.04)	3.70(0.59)
		ssLASSO	15.00(0.00)	0.10(0.40)	1.00(0.01)	1.00(0.01)	1.31(0.20)
		LASSO	15.00(0.00)	76.83(30.14)	0.30(0.09)	0.41(0.08)	5.34(1.25)
	Error2	ssQLASSO	13.47(2.06)	0.33(0.66)	0.93(0.09)	0.93(0.08)	2.33(1.51)
		QLASSO	14.97(0.18)	35.43(4.57)	0.46(0.03)	0.54(0.03)	5.52(0.94)
		ssLASSO	6.80(2.92)	0.47(1.66)	0.59(0.17)	0.64(0.14)	7.32(1.69)
		LASSO	12.63(2.89)	59.03(28.63)	0.31(0.09)	0.39(0.09)	10.52(2.98)
	Error3	ssQLASSO	15.00(0.00)	1.20(2.47)	0.97(0.06)	0.97(0.06)	0.86(0.20)
		QLASSO	15.00(0.00)	27.00(8.03)	0.54(0.07)	0.60(0.06)	2.58(0.50)
		ssLASSO	8.97(3.42)	0.07(0.25)	0.72(0.18)	0.75(0.15)	6.05(1.90)
		LASSO	14.57(0.68)	67.57(18.21)	0.31(0.07)	0.41(0.05)	9.70(1.88)
	Error4	ssQLASSO	13.50(2.27)	0.40(0.77)	0.93(0.10)	0.93(0.10)	2.53(1.55)
		QLASSO	14.97(0.18)	33.70(5.21)	0.47(0.04)	0.55(0.03)	5.85(1.33)
		ssLASSO	5.90(2.63)	0.23(0.63)	0.54(0.17)	0.60(0.14)	7.58(1.41)
		LASSO	14.03(1.75)	73.93(28.03)	0.29(0.07)	0.39(0.06)	11.06(2.93)
	Error5	ssQLASSO	14.40(1.10)	0.37(0.76)	0.97(0.05)	0.97(0.05)	1.94(0.98)
		QLASSO	14.87(0.43)	32.70(4.55)	0.48(0.04)	0.55(0.03)	5.78(1.23)
		ssLASSO	8.60(3.63)	0.30(0.65)	0.69(0.20)	0.73(0.16)	6.23(2.19)
		LASSO	14.80(0.41)	64.00(27.63)	0.34(0.08)	0.44(0.07)	8.83(2.39)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	0.73(1.28)	0.98(0.04)	0.98(0.04)	0.67(0.20)
$(a, b) = (0.5, 3)$		QLASSO	15.00(0.00)	36.60(4.99)	0.45(0.04)	0.53(0.03)	3.70(0.56)
		ssLASSO	15.00(0.00)	0.03(0.18)	1.00(0.01)	1.00(0.01)	0.92(0.22)
		LASSO	15.00(0.00)	75.73(28.34)	0.30(0.07)	0.41(0.06)	5.08(1.37)
	Error2	ssQLASSO	13.17(2.26)	0.30(0.47)	0.92(0.09)	0.92(0.09)	2.22(1.56)
		QLASSO	15.00(0.00)	49.37(6.77)	0.38(0.04)	0.48(0.03)	4.67(0.80)
		ssLASSO	6.77(3.36)	0.17(0.38)	0.59(0.22)	0.66(0.16)	6.88(2.22)
		LASSO	12.43(4.29)	53.77(23.38)	0.30(0.09)	0.38(0.11)	9.65(1.75)
	Error3	ssQLASSO	15.00(0.00)	0.20(0.41)	0.99(0.01)	0.99(0.01)	1.58(0.24)
		QLASSO	15.00(0.00)	43.60(9.95)	0.41(0.06)	0.50(0.05)	4.06(0.57)
		ssLASSO	9.30(3.96)	6.03(32.67)	0.71(0.24)	0.74(0.21)	9.29(14.38)
		LASSO	14.23(2.61)	75.03(32.47)	0.29(0.08)	0.39(0.07)	10.71(2.69)
	Error4	ssQLASSO	12.13(2.29)	0.47(0.73)	0.87(0.10)	0.88(0.09)	3.19(1.71)
		QLASSO	14.97(0.18)	49.70(6.34)	0.38(0.03)	0.47(0.02)	5.52(1.00)
		ssLASSO	6.73(3.65)	0.47(0.90)	0.57(0.23)	0.62(0.19)	6.91(2.08)
		LASSO	13.90(0.96)	73.73(29.12)	0.29(0.09)	0.39(0.07)	11.38(2.44)
	Error5	ssQLASSO	13.33(1.99)	0.70(1.18)	0.91(0.09)	0.92(0.08)	2.33(1.53)
		QLASSO	14.93(0.25)	50.63(6.13)	0.37(0.03)	0.47(0.02)	5.51(1.10)
		ssLASSO	7.27(3.90)	0.17(0.38)	0.61(0.24)	0.66(0.20)	6.19(2.21)
		LASSO	14.30(0.95)	71.80(29.79)	0.30(0.08)	0.40(0.07)	9.49(2.08)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	0.67(0.99)	0.98(0.03)	0.98(0.03)	1.18(0.23)
$(a, b) = (5, 1)$		QLASSO	15.00(0.00)	35.83(8.11)	0.46(0.05)	0.54(0.04)	4.12(0.50)
		ssLASSO	15.00(0.00)	0.10(0.40)	1.00(0.01)	1.00(0.01)	1.41(0.20)
		LASSO	15.00(0.00)	82.13(31.20)	0.29(0.08)	0.40(0.07)	5.77(1.33)
	Error2	ssQLASSO	13.33(1.77)	0.13(0.43)	0.93(0.08)	0.94(0.07)	2.16(1.29)
		QLASSO	14.43(0.50)	33.17(5.04)	0.46(0.04)	0.54(0.03)	5.42(1.09)
		ssLASSO	6.93(2.95)	0.07(0.37)	0.61(0.19)	0.66(0.15)	6.84(1.87)
		LASSO	9.90(4.71)	43.47(29.48)	0.29(0.10)	0.34(0.11)	10.61(1.86)
	Error3	ssQLASSO	14.30(0.53)	0.17(0.38)	0.97(0.02)	0.97(0.02)	1.62(0.34)
		QLASSO	14.93(0.25)	31.53(4.90)	0.49(0.04)	0.56(0.03)	5.97(1.01)
		ssLASSO	9.00(3.38)	0.00(0.00)	0.72(0.19)	0.76(0.16)	5.30(2.11)
		LASSO	14.10(2.38)	62.47(27.96)	0.33(0.08)	0.42(0.08)	9.23(2.36)
	Error4	ssQLASSO	11.10(2.31)	0.60(0.77)	0.82(0.11)	0.83(0.10)	3.79(1.48)
		QLASSO	13.97(0.72)	33.40(6.09)	0.45(0.05)	0.52(0.04)	7.05(1.35)
		ssLASSO	6.10(3.34)	0.40(0.72)	0.53(0.22)	0.59(0.18)	6.94(1.78)
		LASSO	12.50(1.46)	64.67(22.63)	0.29(0.07)	0.37(0.06)	11.02(2.07)
	Error5	ssQLASSO	12.40(2.28)	0.30(0.53)	0.89(0.11)	0.89(0.10)	2.94(1.62)
		QLASSO	14.00(1.14)	34.27(5.51)	0.45(0.05)	0.52(0.05)	7.12(1.38)
		ssLASSO	6.70(2.69)	0.17(0.38)	0.59(0.18)	0.64(0.15)	6.76(1.65)
		LASSO	13.83(0.83)	62.97(24.70)	0.33(0.10)	0.42(0.08)	9.45(2.12)

B.4 Tuning Selection

In the numeric study, the tuning parameters of ssQLASSO have been chosen using the Schwarz Information Criterion (SIC). We have also explored the selection of tuning parameter (s_0,s_1) under cross validation (CV) in terms of the quantile check loss.

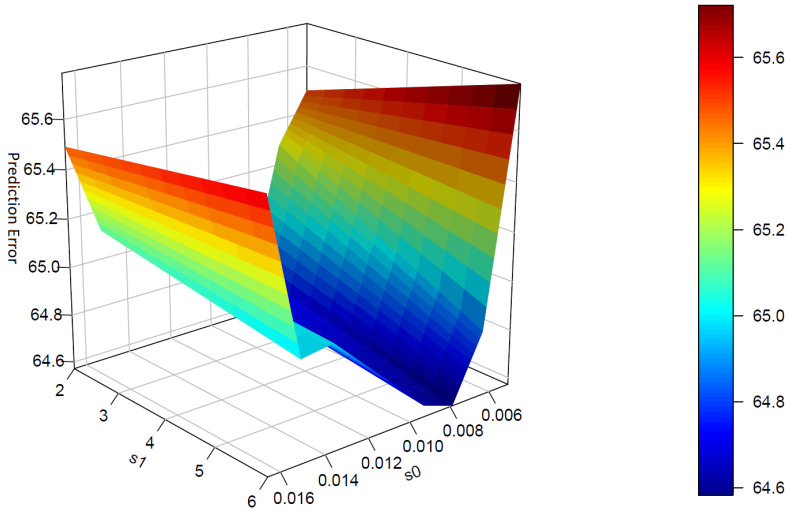

Figure B.1: The surface plot of prediction error using ssQLASSO based on quantile check loss under $\tau = 70\%$, AR-1 correlation, $(n = 400, p = 1600)$ and error 3 over 100 replications.

Figure B.1 shows the 3D surface plot of CV error when ssQLASSO is adopted for model fitting. The pattern is similar to that in Figure 3.5 when best tunings are selected using SIC. Please refer to the subsection of **Tuning parameter selection** in the Section of Simulation for more details of the interpretation.

Besides, for ssQLASSO and LASSO, we have compared their solution paths in Figure B.2 and Figure B.3, respectively. The response variable has been generated under the setting of (n, p)=(400,1600), AR1 correlation and homogeneous Error 3. For better illustration, only 5 nonzero components, $(0.8, 0.6, 0.4, -0.5, -0.6)$, in the coefficient vector $\boldsymbol{\beta}$ are assumed. We refer you to the subsection of **Solution path and self-adaptivity** in the Section of Simulation for more details in interpreting Figure B.2 and Figure B.3 to understand the selective shrinkage and self-adaptivity properties of ssQLASSO.

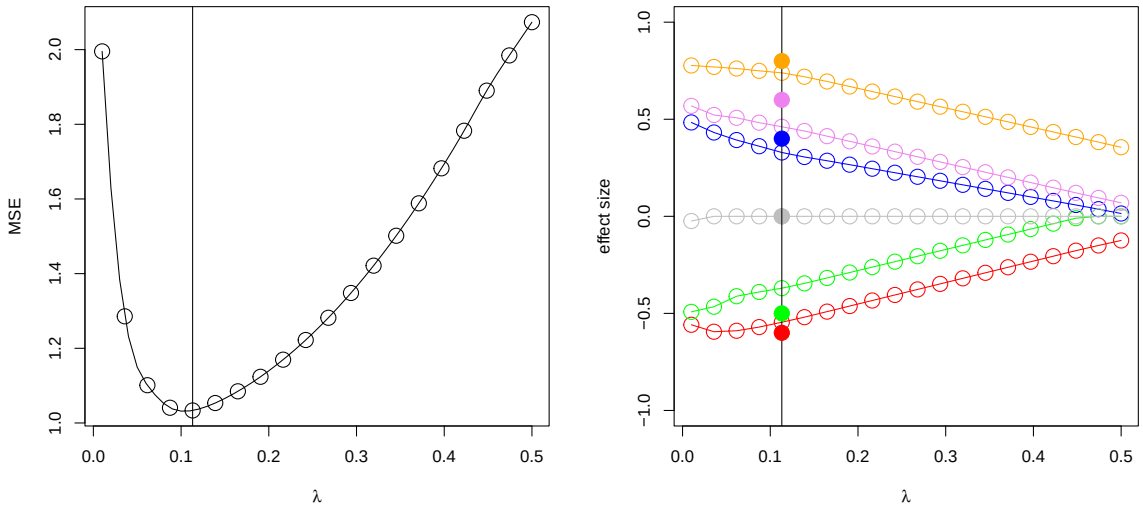

Figure B.2: The MSE profile (left) and solution path (right) of LASSO under $\tau = 50\%$, AR-1 correlation, $(n = 400, p = 1600)$ and Error3. The colored circles on the solution path denote the estimated values with respect to the 5 non-zero coefficients represented by filled circles. The vertical lines correspond to the optimal tuning parameter.

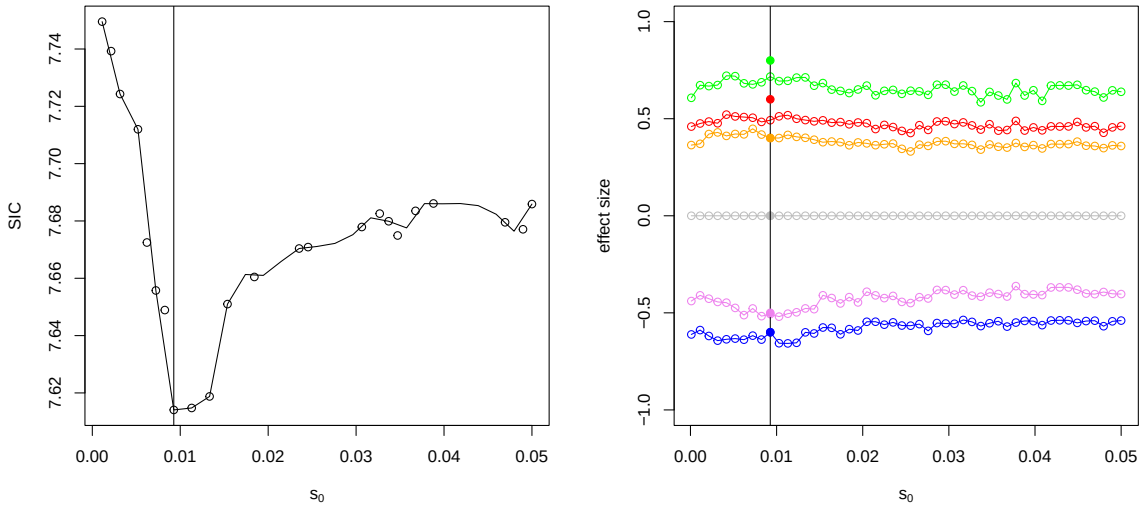

Figure B.3: The SIC profile (left) and solution path (right) of ssQLASSO under $\tau = 50\%$, AR-1 correlation, $(n = 400, p = 1600)$ and Error3. The colored circles on the solution path denote the estimated values with respect to the 5 non-zero coefficients represented by filled circles. The vertical lines correspond to the optimal tuning parameter.

B.5 Computational Time

Table B.13: Computational time (in seconds) of homogeneous model under AR-1 correlation, (n, p)=(400,1600) with 100 replicates.

Run time		$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$
Error1	ssQLASSO	24.27(4.33)	28.85(7.65)	25.27(6.69)
	QLASSO	281.23(10.61)	183.50(7.73)	286.97(11.01)
	ssLASSO	9.12(0.43)	11.66(0.64)	9.19(0.58)
	LASSO	0.01(0.00)	0.02(0.00)	0.01(0.00)
Error2	ssQLASSO	23.08(3.28)	22.51(3.39)	24.47(3.97)
	QLASSO	248.85(9.24)	261.97(13.67)	275.52(11.42)
	ssLASSO	10.34(0.56)	9.70(0.68)	10.80(0.55)
	LASSO	0.02(0.01)	0.02(0.02)	0.01(0.00)
Error3	ssQLASSO	27.87(5.30)	26.14(4.47)	23.08(3.40)
	QLASSO	293.26(10.98)	286.14(13.28)	268.87(10.66)
	ssLASSO	10.82(0.59)	10.14(0.53)	10.31(0.48)
	LASSO	0.02(0.02)	0.02(0.00)	0.01(0.00)
Error4	ssQLASSO	24.19(3.84)	23.53(3.38)	23.90(4.40)
	QLASSO	264.00(10.66)	295.07(13.58)	273.10(10.12)
	ssLASSO	10.43(0.56)	10.44(0.47)	10.37(0.25)
	LASSO	0.01(0.00)	0.02(0.00)	0.01(0.00)
Error5	ssQLASSO	23.27(3.52)	24.07(4.34)	22.94(3.89)
	QLASSO	267.89(10.37)	292.75(12.99)	253.00(9.73)
	ssLASSO	10.30(0.35)	10.36(0.53)	9.66(0.41)
	LASSO	0.01(0.00)	0.01(0.00)	0.02(0.00)

B.6 More Numerical Results

B.6.1 Simulation under large scale settings

We have conducted additional simulation studies to demonstrate that the ssQLASSO can consistently outperform alternatives with much larger sample size and dimensionality in addition to the cases of $(n, p) = (400, 800)\text{and}(400, 1600)$ examined before. Table B.14 and B.15 have shown the identification and estimation results when predictors are generated with AR-1 correlation and $(n, p) = (800, 2400)$ from the following homogeneous and heterogeneous data generating models:

$$\text{homogeneous }: y_i = 2 + \boldsymbol{x}_i^\top \boldsymbol{\beta} + \epsilon_i,$$

$$\text{heterogeneous }: y_i = 2 + \boldsymbol{x}_i^\top \boldsymbol{\beta} + (1 + x_{i2})\epsilon_i,$$

where details of the data generating mechanism have been described in the Section of Simulation. In addition, using the same data generating model under $(n, p) = (800, 3200)$, we have generated the corresponding assessment results in Table B.16 and B.17 for homogeneous and heterogeneous models, respectively.

In all the aforementioned simulation, we have not included the clinical covariates that are not subject to selection in the data generating model since typically inclusion of low dimensional covariates not subject to selection will not affect the estimation and identification performance on high-dimensional omics predictors. To assess the model performance when clinical covariates are incorporated, we then consider the following data generating models with predictors being simulated with AR-1 correlation and $(n, p) = (800, 3200)$ as follows.

$$\text{homogeneous }: y_i = \boldsymbol{z}_i^\top \boldsymbol{\alpha} + \boldsymbol{x}_i^\top \boldsymbol{\beta} + \epsilon_i,$$

$$\text{heterogeneous }: y_i = \boldsymbol{z}_i^\top \boldsymbol{\alpha} + \boldsymbol{x}_i^\top \boldsymbol{\beta} + (1 + x_{i2})\epsilon_i,$$

where $\boldsymbol{z}_i = (1, z_{i1}, z_{i2}, z_{i3})^\top$ corresponds to the intercept and 3 clinical covariates, and $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2, \alpha_3)^\top$ with α_0=2 . The 3 clinical covariates are simulated from multivariate

normal distribution and the associated coefficients are generated from the uniform distribution $U[0.6,0.8]$. Table B.18 and Table B.19 show the results under homogeneous and heterogeneous errors, respectively. The patterns of model performance in terms of identification and estimations under these additional large scale settings in Tables B.14 – B.19 are consistent with our previous findings, indicating the superiority of ssQLASSO over alternative methods.

Table B.14: Evaluation of homogeneous model under AR-1 correlation, (n, p)=(800,2400) with 100 replicates.

$n = 800$	$p = 2400$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	4.33(0.58)	0.87(0.01)	0.88(0.01)	0.67(0.06)
		QLASSO	15.00(0.00)	2.47(1.63)	0.93(0.05)	0.93(0.04)	2.83(0.30)
		ssLASSO	15.00(0.00)	0.57(0.77)	0.98(0.02)	0.98(0.02)	0.43(0.08)
		LASSO	15.00(0.00)	83.43(35.24)	0.29(0.07)	0.40(0.06)	2.45(0.67)
	Error2	ssQLASSO	15.00(0.00)	1.13(1.78)	0.97(0.05)	0.97(0.05)	0.71(0.14)
		QLASSO	15.00(0.00)	5.07(2.52)	0.86(0.06)	0.87(0.05)	4.11(0.53)
		ssLASSO	13.77(2.21)	0.00(0.00)	0.95(0.09)	0.95(0.08)	2.03(1.71)
		LASSO	14.60(2.19)	80.53(38.22)	0.29(0.09)	0.40(0.08)	6.99(2.24)
	Error3	ssQLASSO	15.00(0.00)	3.23(1.96)	0.91(0.05)	0.91(0.05)	0.52(0.11)
		QLASSO	15.00(0.00)	5.40(2.14)	0.85(0.05)	0.86(0.05)	1.85(0.24)
		ssLASSO	15.00(0.00)	0.23(0.57)	0.99(0.02)	0.99(0.02)	0.90(0.23)
		LASSO	15.00(0.00)	75.23(30.18)	0.30(0.07)	0.41(0.06)	4.95(1.27)
	Error4	ssQLASSO	15.00(0.00)	1.70(2.64)	0.95(0.07)	0.95(0.07)	0.75(0.14)
		QLASSO	15.00(0.00)	5.03(2.27)	0.86(0.06)	0.87(0.05)	4.25(0.49)
		ssLASSO	14.60(1.45)	0.03(0.18)	0.98(0.06)	0.98(0.06)	1.28(0.97)
		LASSO	15.00(0.00)	75.93(28.82)	0.30(0.09)	0.42(0.07)	5.37(1.14)
	Error5	ssQLASSO	15.00(0.00)	0.67(1.09)	0.98(0.03)	0.98(0.03)	0.81(0.19)
		QLASSO	15.00(0.00)	4.80(1.97)	0.86(0.05)	0.87(0.04)	4.78(0.63)
		ssLASSO	14.93(0.37)	0.20(0.41)	0.99(0.02)	0.99(0.02)	0.90(0.26)
		LASSO	15.00(0.00)	81.53(27.30)	0.29(0.07)	0.40(0.06)	4.92(0.92)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	2.50(3.01)	0.93(0.08)	0.93(0.07)	0.91(0.21)
		QLASSO	15.00(0.00)	2.17(1.95)	0.94(0.05)	0.94(0.05)	2.87(0.28)
		ssLASSO	15.00(0.00)	0.53(0.78)	0.98(0.02)	0.98(0.02)	0.45(0.12)
		LASSO	15.00(0.00)	79.97(29.67)	0.29(0.08)	0.41(0.07)	2.42(0.46)
	Error2	ssQLASSO	15.00(0.00)	1.27(2.60)	0.97(0.07)	0.97(0.06)	0.66(0.14)
		QLASSO	15.00(0.00)	11.03(3.12)	0.74(0.06)	0.76(0.05)	2.89(0.37)
		ssLASSO	14.57(1.10)	0.03(0.18)	0.98(0.04)	0.98(0.04)	1.50(0.96)
		LASSO	14.63(1.03)	65.67(24.17)	0.32(0.08)	0.43(0.06)	6.47(2.38)
	Error3	ssQLASSO	15.00(0.00)	0.77(1.61)	0.98(0.05)	0.98(0.04)	0.58(0.12)
		QLASSO	15.00(0.00)	11.00(2.82)	0.74(0.05)	0.76(0.04)	2.37(0.33)
		ssLASSO	14.90(0.40)	0.27(0.64)	0.99(0.02)	0.99(0.02)	0.97(0.43)
		LASSO	15.00(0.00)	75.47(23.86)	0.30(0.06)	0.41(0.05)	4.96(1.34)
	Error4	ssQLASSO	15.00(0.00)	1.07(2.08)	0.97(0.06)	0.97(0.05)	0.79(0.19)
		QLASSO	15.00(0.00)	11.03(2.28)	0.73(0.04)	0.76(0.03)	3.41(0.50)
		ssLASSO	14.67(1.15)	0.00(0.00)	0.99(0.05)	0.99(0.04)	1.47(0.95)
		LASSO	15.00(0.00)	84.47(44.41)	0.29(0.08)	0.40(0.07)	5.66(1.74)
	Error5	ssQLASSO	15.00(0.00)	1.00(2.41)	0.97(0.06)	0.97(0.06)	0.80(0.20)
		QLASSO	15.00(0.00)	10.87(2.90)	0.74(0.05)	0.76(0.04)	3.71(0.53)
		ssLASSO	15.00(0.00)	0.30(0.60)	0.99(0.02)	0.99(0.02)	0.91(0.21)
		LASSO	15.00(0.00)	81.50(53.39)	0.31(0.10)	0.42(0.08)	5.04(1.90)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	4.89(2.35)	0.86(0.06)	0.87(0.05)	0.69(0.10)
		QLASSO	15.00(0.00)	2.57(1.87)	0.92(0.05)	0.93(0.05)	2.78(0.26)
		ssLASSO	15.00(0.00)	0.57(0.68)	0.98(0.02)	0.98(0.02)	0.48(0.11)
		LASSO	15.00(0.00)	74.30(27.11)	0.31(0.08)	0.42(0.07)	2.34(0.51)
	Error2	ssQLASSO	15.00(0.00)	1.27(2.15)	0.96(0.06)	0.97(0.06)	0.71(0.15)
		QLASSO	15.00(0.00)	5.73(2.23)	0.84(0.05)	0.85(0.05)	3.88(0.39)
		ssLASSO	14.27(1.55)	0.00(0.00)	0.97(0.07)	0.97(0.06)	1.75(1.16)
		LASSO	14.70(1.15)	73.00(27.36)	0.30(0.06)	0.41(0.06)	6.90(1.63)
	Error3	ssQLASSO	15.00(0.00)	0.90(1.97)	0.97(0.05)	0.98(0.05)	0.73(0.13)
		QLASSO	15.00(0.00)	5.40(2.19)	0.85(0.05)	0.86(0.04)	4.26(0.42)
		ssLASSO	14.90(0.55)	0.23(0.50)	0.99(0.02)	0.99(0.02)	0.95(0.49)
		LASSO	15.00(0.00)	84.90(21.32)	0.27(0.05)	0.39(0.04)	5.09(0.94)
	Error4	ssQLASSO	15.00(0.00)	1.17(2.18)	0.97(0.06)	0.97(0.05)	0.74(0.17)
		QLASSO	15.00(0.00)	5.37(2.53)	0.85(0.06)	0.86(0.05)	4.09(0.51)
		ssLASSO	14.57(1.33)	0.03(0.18)	0.98(0.06)	0.98(0.05)	1.37(1.04)
		LASSO	15.00(0.00)	70.83(22.19)	0.31(0.07)	0.42(0.06)	5.15(0.77)
	Error5	ssQLASSO	15.00(0.00)	0.83(1.72)	0.98(0.05)	0.98(0.05)	0.79(0.19)
		QLASSO	15.00(0.00)	4.47(2.15)	0.87(0.05)	0.88(0.05)	4.75(0.76)
		ssLASSO	15.00(0.00)	0.17(0.46)	0.99(0.01)	0.99(0.01)	0.91(0.20)
		LASSO	15.00(0.00)	70.73(23.97)	0.31(0.07)	0.42(0.06)	4.48(0.94)

Table B.15: Evaluation of heterogeneous model under AR-1 correlation, (n, p)=(800,2400) with 100 replicates.

$n = 800$	$p = 2400$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	3.57(2.53)	0.90(0.07)	0.90(0.06)	0.95(0.14)
		QLASSO	15.00(0.00)	2.70(1.49)	0.92(0.04)	0.92(0.04)	2.81(0.36)
		ssLASSO	15.00(0.00)	0.13(0.35)	1.00(0.01)	1.00(0.01)	1.11(0.14)
		LASSO	15.00(0.00)	67.53(25.34)	0.33(0.08)	0.43(0.07)	3.36(0.59)
	Error2	ssQLASSO	15.00(0.00)	0.93(2.12)	0.97(0.06)	0.98(0.05)	0.85(0.18)
		QLASSO	15.00(0.00)	4.97(1.90)	0.86(0.05)	0.87(0.04)	3.42(0.54)
		ssLASSO	13.47(2.11)	0.00(0.00)	0.94(0.09)	0.94(0.08)	2.89(1.41)
		LASSO	14.17(2.68)	68.50(26.50)	0.30(0.07)	0.40(0.07)	8.39(1.98)
	Error3	ssQLASSO	15.00(0.00)	1.30(2.07)	0.96(0.06)	0.96(0.05)	0.59(0.12)
		QLASSO	15.00(0.00)	4.47(2.00)	0.87(0.05)	0.88(0.04)	1.75(0.27)
		ssLASSO	15.00(0.00)	0.03(0.18)	1.00(0.01)	1.00(0.01)	2.36(0.42)
		LASSO	14.93(0.37)	79.00(23.33)	0.29(0.06)	0.40(0.05)	7.69(1.46)
	Error4	ssQLASSO	15.00(0.00)	0.17(0.38)	0.99(0.01)	0.99(0.01)	0.96(0.18)
		QLASSO	14.97(0.18)	4.97(2.39)	0.86(0.06)	0.87(0.05)	4.13(0.69)
		ssLASSO	13.73(1.87)	0.00(0.00)	0.95(0.08)	0.95(0.07)	2.77(1.14)
		LASSO	14.93(0.25)	80.30(27.17)	0.29(0.08)	0.40(0.07)	8.34(1.48)
	Error5	ssQLASSO	15.00(0.00)	0.30(0.84)	0.99(0.03)	0.99(0.02)	0.96(0.15)
		QLASSO	15.00(0.00)	5.17(2.09)	0.86(0.05)	0.87(0.05)	4.28(0.69)
		ssLASSO	13.97(2.17)	0.03(0.18)	0.96(0.10)	0.96(0.09)	2.33(1.25)
		LASSO	15.00(0.00)	75.73(23.69)	0.30(0.06)	0.41(0.05)	6.75(1.16)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	3.47(3.79)	0.91(0.09)	0.91(0.09)	0.47(0.15)
		QLASSO	15.00(0.00)	2.73(1.53)	0.92(0.04)	0.92(0.04)	2.77(0.43)
		ssLASSO	15.00(0.00)	0.10(0.55)	1.00(0.02)	1.00(0.02)	0.63(0.12)
		LASSO	15.00(0.00)	75.17(22.35)	0.30(0.07)	0.41(0.06)	3.27(0.57)
	Error2	ssQLASSO	15.00(0.00)	0.73(1.55)	0.98(0.04)	0.98(0.04)	0.48(0.13)
		QLASSO	15.00(0.00)	11.30(3.00)	0.73(0.05)	0.76(0.04)	2.61(0.49)
		ssLASSO	13.97(1.56)	0.10(0.31)	0.96(0.06)	0.96(0.06)	2.25(1.06)
		LASSO	14.13(2.66)	67.90(25.94)	0.30(0.08)	0.41(0.07)	8.24(2.28)
	Error3	ssQLASSO	15.00(0.00)	1.03(2.36)	0.97(0.06)	0.97(0.06)	1.31(0.10)
		QLASSO	15.00(0.00)	9.43(2.64)	0.76(0.05)	0.79(0.04)	2.52(0.23)
		ssLASSO	15.00(0.00)	0.00(0.00)	1.00(0.00)	1.00(0.00)	2.86(0.43)
		LASSO	14.97(0.18)	75.30(28.87)	0.30(0.07)	0.41(0.06)	7.79(1.56)
	Error4	ssQLASSO	15.00(0.00)	0.57(0.97)	0.98(0.03)	0.98(0.03)	0.53(0.15)
		QLASSO	15.00(0.00)	11.77(2.74)	0.72(0.05)	0.75(0.04)	3.20(0.45)
		ssLASSO	12.90(2.02)	0.00(0.00)	0.92(0.08)	0.92(0.08)	2.64(1.22)
		LASSO	14.93(0.25)	77.13(32.82)	0.30(0.08)	0.41(0.07)	8.01(1.67)
	Error5	ssQLASSO	15.00(0.00)	0.23(0.50)	0.99(0.02)	0.99(0.02)	0.52(0.11)
		QLASSO	15.00(0.00)	11.30(3.67)	0.73(0.07)	0.76(0.06)	3.24(0.50)
		ssLASSO	14.47(1.36)	0.07(0.37)	0.98(0.05)	0.98(0.05)	1.64(0.79)
		LASSO	15.00(0.00)	75.63(25.25)	0.30(0.07)	0.41(0.06)	6.49(1.36)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	2.80(3.48)	0.92(0.09)	0.93(0.08)	1.00(0.22)
		QLASSO	15.00(0.00)	2.63(1.88)	0.92(0.05)	0.93(0.05)	3.12(0.35)
		ssLASSO	15.00(0.00)	0.37(0.61)	0.99(0.02)	0.99(0.02)	1.12(0.19)
		LASSO	15.00(0.00)	66.20(25.25)	0.33(0.08)	0.44(0.07)	3.57(0.56)
	Error2	ssQLASSO	14.90(0.31)	1.30(2.42)	0.96(0.06)	0.96(0.06)	1.10(0.14)
		QLASSO	14.77(0.43)	5.47(2.30)	0.84(0.05)	0.85(0.05)	4.13(0.58)
		ssLASSO	13.53(1.07)	0.00(0.00)	0.95(0.04)	0.95(0.04)	2.28(0.72)
		LASSO	14.00(0.64)	80.27(30.04)	0.28(0.09)	0.38(0.08)	9.26(2.49)
	Error3	ssQLASSO	15.00(0.00)	0.20(0.48)	0.99(0.02)	0.99(0.02)	0.92(0.21)
		QLASSO	15.00(0.00)	4.83(2.46)	0.87(0.06)	0.87(0.05)	4.10(0.60)
		ssLASSO	14.57(0.63)	0.00(0.00)	0.98(0.02)	0.99(0.02)	1.86(0.70)
		LASSO	14.67(1.49)	70.20(26.04)	0.31(0.08)	0.42(0.07)	6.87(1.52)
	Error4	ssQLASSO	14.50(0.51)	0.47(1.07)	0.97(0.03)	0.97(0.03)	1.30(0.15)
		QLASSO	14.10(0.55)	5.47(2.30)	0.82(0.06)	0.83(0.06)	5.10(0.76)
		ssLASSO	12.50(2.19)	0.03(0.18)	0.90(0.11)	0.91(0.09)	2.95(1.36)
		LASSO	14.13(0.35)	74.83(28.65)	0.29(0.08)	0.39(0.07)	8.33(1.86)
	Error5	ssQLASSO	14.13(0.35)	0.60(1.43)	0.95(0.04)	0.95(0.03)	1.19(0.12)
		QLASSO	14.30(0.47)	5.70(2.23)	0.82(0.05)	0.83(0.05)	4.72(0.44)
		ssLASSO	13.57(0.82)	0.00(0.00)	0.95(0.03)	0.95(0.03)	1.92(0.58)
		LASSO	14.10(0.31)	78.50(29.01)	0.28(0.09)	0.39(0.07)	6.87(1.29)

Table B.16: Evaluation of homogeneous model under AR-1 correlation, (n, p)=(800,3200) with 100 replicates.

$n = 800$	$p = 3200$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	1.70(1.70)	0.95(0.05)	0.95(0.04)	0.78(0.19)
		QLASSO	15.00(0.00)	3.27(1.98)	0.90(0.05)	0.91(0.05)	2.92(0.27)
		ssLASSO	15.00(0.00)	0.40(0.72)	0.99(0.02)	0.99(0.02)	0.44(0.07)
		LASSO	15.00(0.00)	81.90(28.37)	0.28(0.07)	0.40(0.06)	2.45(0.49)
	Error2	ssQLASSO	15.00(0.00)	0.50(0.97)	0.98(0.03)	0.98(0.03)	0.71(0.16)
		QLASSO	15.00(0.00)	6.47(2.05)	0.83(0.05)	0.84(0.04)	3.92(0.48)
		ssLASSO	13.50(2.03)	0.00(0.00)	0.94(0.08)	0.95(0.08)	2.25(1.45)
		LASSO	14.23(2.18)	80.93(37.59)	0.28(0.07)	0.39(0.06)	7.57(2.53)
	Error3	ssQLASSO	15.00(0.00)	1.30(1.32)	0.96(0.04)	0.96(0.04)	0.53(0.11)
		QLASSO	15.00(0.00)	6.57(2.28)	0.82(0.05)	0.84(0.05)	1.93(0.20)
		ssLASSO	14.93(0.25)	0.33(1.12)	0.99(0.03)	0.99(0.03)	1.01(0.37)
		LASSO	15.00(0.00)	90.37(30.07)	0.26(0.07)	0.38(0.06)	5.64(1.82)
	Error4	ssQLASSO	15.00(0.00)	0.30(0.70)	0.99(0.02)	0.99(0.02)	0.78(0.15)
		QLASSO	15.00(0.00)	6.20(2.19)	0.83(0.05)	0.84(0.04)	4.54(0.63)
		ssLASSO	14.10(2.23)	0.03(0.18)	0.96(0.10)	0.96(0.09)	1.69(1.65)
		LASSO	15.00(0.00)	80.27(29.04)	0.29(0.06)	0.40(0.05)	5.73(1.20)
	Error5	ssQLASSO	15.00(0.00)	0.73(1.44)	0.98(0.04)	0.98(0.04)	0.78(0.17)
		QLASSO	15.00(0.00)	6.60(2.99)	0.82(0.07)	0.84(0.06)	4.97(0.60)
		ssLASSO	14.83(0.65)	0.37(0.85)	0.98(0.03)	0.98(0.03)	1.09(0.49)
		LASSO	15.00(0.00)	81.37(24.89)	0.28(0.06)	0.40(0.05)	4.91(0.77)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	1.73(1.80)	0.95(0.05)	0.95(0.04)	0.80(0.22)
		QLASSO	15.00(0.00)	3.33(1.90)	0.90(0.05)	0.91(0.05)	3.35(0.42)
		ssLASSO	15.00(0.00)	0.77(1.04)	0.98(0.03)	0.98(0.03)	0.43(0.07)
		LASSO	15.00(0.00)	101.43(39.50)	0.25(0.06)	0.37(0.06)	2.87(0.71)
	Error2	ssQLASSO	15.00(0.00)	0.90(1.30)	0.97(0.04)	0.97(0.04)	0.66(0.14)
		QLASSO	15.00(0.00)	13.57(4.05)	0.69(0.06)	0.73(0.05)	3.11(0.31)
		ssLASSO	13.63(2.46)	0.00(0.00)	0.94(0.13)	0.95(0.11)	2.03(1.54)
		LASSO	14.40(2.63)	79.80(32.30)	0.27(0.07)	0.39(0.07)	7.00(1.80)
	Error3	ssQLASSO	15.00(0.00)	2.33(3.76)	0.94(0.08)	0.94(0.08)	0.63(0.14)
		QLASSO	15.00(0.00)	15.50(3.68)	0.66(0.05)	0.70(0.04)	2.42(0.30)
		ssLASSO	14.63(2.01)	0.47(1.78)	0.97(0.11)	0.97(0.10)	1.14(1.28)
		LASSO	15.00(0.00)	77.07(31.74)	0.30(0.08)	0.41(0.06)	5.36(1.39)
	Error4	ssQLASSO	15.00(0.00)	1.93(3.71)	0.95(0.09)	0.95(0.08)	0.76(0.23)
		QLASSO	15.00(0.00)	13.40(3.49)	0.70(0.06)	0.73(0.05)	3.33(0.45)
		ssLASSO	14.57(1.36)	0.00(0.00)	0.98(0.05)	0.98(0.05)	1.29(0.84)
		LASSO	15.00(0.00)	81.27(29.44)	0.29(0.07)	0.40(0.06)	5.64(1.08)
	Error5	ssQLASSO	15.00(0.00)	1.30(2.81)	0.96(0.07)	0.97(0.06)	0.78(0.19)
		QLASSO	15.00(0.00)	14.77(3.59)	0.67(0.06)	0.71(0.05)	3.83(0.50)
		ssLASSO	14.93(0.37)	0.30(0.65)	0.99(0.02)	0.99(0.02)	0.87(0.32)
		LASSO	15.00(0.00)	82.63(34.83)	0.28(0.07)	0.40(0.06)	5.02(1.19)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	2.40(2.92)	0.93(0.07)	0.94(0.06)	0.86(0.20)
		QLASSO	15.00(0.00)	3.10(1.81)	0.91(0.05)	0.91(0.04)	2.88(0.26)
		ssLASSO	15.00(0.00)	2.77(2.99)	0.92(0.07)	0.93(0.07)	0.48(0.11)
		LASSO	15.00(0.00)	85.93(53.53)	0.29(0.08)	0.40(0.07)	2.57(0.95)
	Error2	ssQLASSO	15.00(0.00)	0.90(2.67)	0.98(0.07)	0.98(0.06)	0.71(0.13)
		QLASSO	15.00(0.00)	6.80(2.93)	0.82(0.06)	0.83(0.06)	3.86(0.56)
		ssLASSO	13.30(3.17)	0.00(0.00)	0.92(0.16)	0.93(0.14)	2.40(2.26)
		LASSO	14.10(2.96)	80.53(34.06)	0.27(0.08)	0.39(0.06)	7.29(2.32)
	Error3	ssQLASSO	15.00(0.00)	1.03(2.04)	0.97(0.05)	0.97(0.05)	0.76(0.19)
		QLASSO	15.00(0.00)	6.87(2.87)	0.82(0.06)	0.83(0.06)	4.24(0.51)
		ssLASSO	15.00(0.00)	0.20(0.41)	0.99(0.01)	0.99(0.01)	0.91(0.23)
		LASSO	15.00(0.00)	79.53(24.94)	0.29(0.07)	0.40(0.06)	5.09(0.89)
	Error4	ssQLASSO	15.00(0.00)	0.80(1.45)	0.98(0.04)	0.98(0.04)	0.71(0.16)
		QLASSO	15.00(0.00)	6.37(2.55)	0.83(0.06)	0.84(0.05)	4.42(0.55)
		ssLASSO	14.87(0.73)	0.00(0.00)	0.99(0.03)	1.00(0.03)	1.21(0.96)
		LASSO	15.00(0.00)	82.90(32.78)	0.29(0.09)	0.40(0.08)	5.76(1.04)
	Error5	ssQLASSO	15.00(0.00)	0.70(1.24)	0.98(0.04)	0.98(0.04)	0.72(0.14)
		QLASSO	15.00(0.00)	6.50(2.35)	0.83(0.05)	0.84(0.05)	4.91(0.59)
		ssLASSO	14.47(2.10)	0.37(0.61)	0.96(0.10)	0.97(0.09)	1.22(1.34)
		LASSO	15.00(0.00)	101.70(49.70)	0.26(0.09)	0.37(0.08)	5.57(1.71)

Table B.17: Evaluation of heterogeneous model under AR-1 correlation, (n, p)=(800,3200) with 100 replicates.

$n = 800$	$p = 3200$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	2.67(2.47)	0.92(0.07)	0.93(0.06)	0.93(0.12)
		QLASSO	15.00(0.00)	3.47(1.78)	0.90(0.05)	0.90(0.04)	2.82(0.32)
		ssLASSO	15.00(0.00)	0.30(0.79)	0.99(0.02)	0.99(0.02)	1.11(0.15)
		LASSO	15.00(0.00)	82.40(30.43)	0.29(0.08)	0.40(0.07)	3.64(0.69)
	Error2	ssQLASSO	15.00(0.00)	0.40(1.04)	0.99(0.03)	0.99(0.03)	0.84(0.14)
		QLASSO	15.00(0.00)	6.60(2.50)	0.82(0.06)	0.84(0.05)	3.69(0.44)
		ssLASSO	13.37(1.97)	0.00(0.00)	0.94(0.08)	0.94(0.07)	2.93(1.31)
		LASSO	13.17(3.80)	79.23(50.98)	0.26(0.10)	0.37(0.08)	9.37(2.47)
	Error3	ssQLASSO	15.00(0.00)	1.53(2.29)	0.96(0.06)	0.96(0.06)	0.64(0.15)
		QLASSO	15.00(0.00)	6.70(3.51)	0.82(0.08)	0.84(0.07)	1.90(0.29)
		ssLASSO	14.93(0.25)	0.00(0.00)	1.00(0.01)	1.00(0.01)	2.28(0.45)
		LASSO	14.97(0.18)	91.90(31.16)	0.26(0.07)	0.38(0.06)	8.67(2.66)
	Error4	ssQLASSO	15.00(0.00)	0.20(0.61)	0.99(0.02)	0.99(0.02)	0.96(0.22)
		QLASSO	15.00(0.00)	6.17(2.97)	0.83(0.06)	0.85(0.05)	4.08(0.63)
		ssLASSO	11.93(2.02)	0.00(0.00)	0.88(0.09)	0.89(0.08)	3.74(1.23)
		LASSO	14.87(0.35)	75.93(24.18)	0.30(0.08)	0.41(0.06)	8.04(1.42)
	Error5	ssQLASSO	15.00(0.00)	0.60(1.43)	0.98(0.04)	0.98(0.04)	0.96(0.14)
		QLASSO	14.97(0.18)	6.53(2.46)	0.82(0.06)	0.84(0.05)	4.44(0.83)
		ssLASSO	14.00(1.53)	0.00(0.00)	0.96(0.06)	0.96(0.06)	2.71(1.13)
		LASSO	15.00(0.00)	92.07(45.26)	0.28(0.10)	0.39(0.09)	7.58(2.06)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	0.90(1.21)	0.97(0.04)	0.97(0.04)	0.50(0.14)
		QLASSO	15.00(0.00)	3.60(1.35)	0.89(0.04)	0.90(0.03)	2.78(0.36)
		ssLASSO	15.00(0.00)	0.03(0.18)	1.00(0.01)	1.00(0.01)	0.61(0.14)
		LASSO	15.00(0.00)	95.03(40.26)	0.26(0.06)	0.38(0.06)	3.81(1.09)
	Error2	ssQLASSO	15.00(0.00)	0.67(1.47)	0.98(0.04)	0.98(0.04)	0.45(0.12)
		QLASSO	15.00(0.00)	14.73(3.50)	0.67(0.05)	0.71(0.04)	2.73(0.45)
		ssLASSO	13.30(1.82)	0.00(0.00)	0.94(0.08)	0.94(0.07)	2.66(1.35)
		LASSO	14.10(2.81)	73.93(30.11)	0.28(0.08)	0.39(0.09)	8.65(1.57)
	Error3	ssQLASSO	15.00(0.00)	1.80(2.99)	0.95(0.08)	0.95(0.07)	1.33(0.18)
		QLASSO	15.00(0.00)	12.93(4.36)	0.71(0.07)	0.74(0.06)	2.69(0.29)
		ssLASSO	14.93(0.25)	0.07(0.37)	1.00(0.01)	1.00(0.01)	2.82(0.42)
		LASSO	15.00(0.00)	82.10(31.68)	0.29(0.07)	0.40(0.06)	8.23(1.98)
	Error4	ssQLASSO	15.00(0.00)	0.37(0.96)	0.99(0.03)	0.99(0.03)	0.46(0.15)
		QLASSO	15.00(0.00)	14.50(3.63)	0.68(0.06)	0.72(0.04)	3.08(0.51)
		ssLASSO	12.57(2.80)	0.03(0.18)	0.90(0.12)	0.91(0.11)	2.82(1.52)
		LASSO	15.00(0.00)	80.03(28.25)	0.29(0.09)	0.41(0.08)	7.96(1.48)
	Error5	ssQLASSO	15.00(0.00)	0.90(1.71)	0.97(0.05)	0.97(0.05)	0.50(0.12)
		QLASSO	15.00(0.00)	14.40(3.21)	0.68(0.05)	0.72(0.04)	3.17(0.47)
		ssLASSO	13.63(2.17)	0.00(0.00)	0.95(0.10)	0.95(0.08)	2.11(1.37)
		LASSO	15.00(0.00)	86.20(35.19)	0.28(0.08)	0.40(0.07)	7.11(1.70)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	2.23(2.13)	0.93(0.06)	0.94(0.06)	0.96(0.17)
		QLASSO	15.00(0.00)	3.13(1.61)	0.91(0.04)	0.91(0.04)	3.25(0.26)
		ssLASSO	15.00(0.00)	0.13(0.43)	1.00(0.01)	1.00(0.01)	1.08(0.13)
		LASSO	15.00(0.00)	76.90(26.19)	0.30(0.06)	0.41(0.06)	3.83(0.63)
	Error2	ssQLASSO	15.00(0.00)	1.60(2.37)	0.95(0.06)	0.96(0.06)	1.06(0.27)
		QLASSO	14.90(0.31)	7.33(2.55)	0.80(0.05)	0.82(0.04)	4.07(0.52)
		ssLASSO	13.47(0.86)	0.03(0.18)	0.94(0.03)	0.95(0.03)	2.72(0.87)
		LASSO	12.77(3.09)	58.67(27.69)	0.31(0.07)	0.39(0.08)	8.40(1.74)
	Error3	ssQLASSO	14.90(0.31)	1.73(3.44)	0.95(0.08)	0.95(0.08)	1.12(0.21)
		QLASSO	15.00(0.00)	5.77(2.31)	0.84(0.06)	0.85(0.05)	4.26(0.44)
		ssLASSO	14.83(0.38)	0.30(1.64)	0.99(0.04)	0.99(0.04)	1.79(0.79)
		LASSO	15.00(0.00)	82.77(32.25)	0.29(0.08)	0.40(0.06)	7.04(1.72)
	Error4	ssQLASSO	14.90(0.31)	1.43(3.33)	0.96(0.07)	0.96(0.07)	1.27(0.32)
		QLASSO	14.73(0.45)	6.70(2.87)	0.81(0.07)	0.83(0.06)	4.87(0.75)
		ssLASSO	12.90(1.24)	0.00(0.00)	0.92(0.05)	0.93(0.05)	2.74(0.73)
		LASSO	14.00(0.37)	79.53(34.95)	0.28(0.08)	0.39(0.07)	8.45(1.79)
	Error5	ssQLASSO	14.77(0.43)	1.27(2.29)	0.96(0.06)	0.96(0.06)	1.30(0.27)
		QLASSO	14.70(0.53)	6.80(2.43)	0.81(0.06)	0.82(0.05)	5.14(0.62)
		ssLASSO	12.60(1.83)	1.33(7.30)	0.89(0.12)	0.90(0.10)	3.15(2.12)
		LASSO	14.20(0.41)	79.50(31.38)	0.28(0.08)	0.39(0.07)	7.29(1.60)

Table B.18: Evaluation of homogeneous model with clinical factors under AR-1 correlation, (n, p)=(800,3200), 100 replicates.

$n = 800$	$p = 3200$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	3.70(1.97)	0.89(0.05)	0.90(0.05)	0.76(0.15)
		QLASSO	15.00(0.00)	3.87(2.01)	0.89(0.05)	0.89(0.05)	2.85(0.24)
		ssLASSO	15.00(0.00)	0.37(0.76)	0.99(0.02)	0.99(0.02)	0.40(0.08)
		LASSO	15.00(0.00)	82.80(22.16)	0.28(0.05)	0.39(0.05)	2.47(0.38)
	Error2	ssQLASSO	15.00(0.00)	0.43(0.97)	0.99(0.03)	0.99(0.03)	0.72(0.15)
		QLASSO	15.00(0.00)	6.47(2.76)	0.83(0.06)	0.84(0.05)	4.11(0.54)
		ssLASSO	13.63(2.44)	0.00(0.00)	0.94(0.11)	0.95(0.10)	2.14(1.83)
		LASSO	14.43(2.74)	75.37(29.12)	0.28(0.09)	0.41(0.06)	6.91(2.07)
	Error3	ssQLASSO	15.00(0.00)	1.03(1.10)	0.97(0.03)	0.97(0.03)	0.51(0.10)
		QLASSO	15.00(0.00)	7.17(2.04)	0.81(0.04)	0.82(0.04)	1.84(0.24)
		ssLASSO	14.93(0.25)	0.13(0.43)	0.99(0.02)	0.99(0.02)	1.01(0.40)
		LASSO	15.00(0.00)	88.60(38.09)	0.28(0.08)	0.39(0.07)	5.69(1.73)
	Error4	ssQLASSO	15.00(0.00)	0.90(1.67)	0.97(0.05)	0.97(0.04)	0.76(0.16)
		QLASSO	15.00(0.00)	6.33(2.41)	0.83(0.05)	0.84(0.05)	4.58(0.58)
		ssLASSO	14.57(1.22)	0.00(0.00)	0.98(0.05)	0.98(0.04)	1.32(0.80)
		LASSO	15.00(0.00)	86.70(29.56)	0.27(0.07)	0.39(0.06)	5.74(1.13)
	Error5	ssQLASSO	15.00(0.00)	1.80(5.70)	0.96(0.10)	0.96(0.08)	0.85(0.29)
		QLASSO	15.00(0.00)	7.03(2.25)	0.81(0.05)	0.83(0.04)	4.70(0.47)
		ssLASSO	15.00(0.00)	0.20(0.55)	0.99(0.02)	0.99(0.02)	0.89(0.23)
		LASSO	15.00(0.00)	85.77(36.11)	0.28(0.07)	0.40(0.06)	4.94(1.17)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	3.63(2.36)	0.90(0.06)	0.90(0.05)	0.83(0.17)
		QLASSO	15.00(0.00)	3.13(2.00)	0.91(0.05)	0.91(0.05)	2.68(0.26)
		ssLASSO	15.00(0.00)	0.77(0.97)	0.98(0.03)	0.98(0.03)	0.46(0.10)
		LASSO	15.00(0.00)	87.23(28.33)	0.27(0.06)	0.39(0.06)	2.44(0.47)
	Error2	ssQLASSO	15.00(0.00)	0.70(1.42)	0.98(0.04)	0.98(0.04)	0.70(0.14)
		QLASSO	15.00(0.00)	13.73(3.15)	0.69(0.05)	0.72(0.04)	2.98(0.36)
		ssLASSO	14.37(1.73)	0.20(1.10)	0.97(0.08)	0.97(0.07)	1.71(1.32)
		LASSO	15.00(0.00)	81.97(37.71)	0.30(0.10)	0.41(0.08)	6.44(1.88)
	Error3	ssQLASSO	15.00(0.00)	1.13(1.85)	0.97(0.05)	0.97(0.05)	0.63(0.13)
		QLASSO	15.00(0.00)	14.80(3.25)	0.67(0.05)	0.71(0.04)	2.37(0.25)
		ssLASSO	14.97(0.18)	0.30(1.29)	0.99(0.04)	0.99(0.03)	0.87(0.23)
		LASSO	15.00(0.00)	86.23(29.05)	0.27(0.06)	0.39(0.05)	5.21(1.18)
	Error4	ssQLASSO	15.00(0.00)	1.17(2.48)	0.97(0.07)	0.97(0.06)	0.80(0.20)
		QLASSO	15.00(0.00)	13.57(4.45)	0.70(0.07)	0.73(0.06)	3.45(0.44)
		ssLASSO	13.97(2.31)	0.07(0.25)	0.95(0.10)	0.96(0.09)	1.92(1.49)
		LASSO	15.00(0.00)	82.77(27.07)	0.28(0.07)	0.40(0.06)	5.65(1.02)
	Error5	ssQLASSO	15.00(0.00)	1.73(3.10)	0.95(0.08)	0.96(0.07)	0.80(0.18)
		QLASSO	15.00(0.00)	14.37(3.35)	0.68(0.05)	0.72(0.04)	3.71(0.36)
		ssLASSO	14.80(1.10)	0.43(1.10)	0.98(0.05)	0.98(0.05)	1.05(0.86)
		LASSO	15.00(0.00)	87.77(25.43)	0.27(0.06)	0.39(0.05)	5.00(0.81)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	1.83(2.39)	0.95(0.06)	0.95(0.06)	0.74(0.20)
		QLASSO	15.00(0.00)	3.17(1.51)	0.91(0.04)	0.91(0.04)	2.72(0.20)
		ssLASSO	15.00(0.00)	0.63(0.89)	0.98(0.03)	0.98(0.03)	0.44(0.09)
		LASSO	15.00(0.00)	86.73(39.39)	0.29(0.10)	0.40(0.08)	2.49(0.59)
	Error2	ssQLASSO	15.00(0.00)	1.00(1.58)	0.97(0.05)	0.97(0.04)	0.69(0.11)
		QLASSO	15.00(0.00)	7.27(2.50)	0.81(0.05)	0.82(0.05)	3.86(0.66)
		ssLASSO	12.40(3.22)	0.00(0.00)	0.89(0.15)	0.90(0.13)	2.73(1.98)
		LASSO	14.90(0.40)	88.97(27.19)	0.26(0.06)	0.38(0.05)	7.54(2.13)
	Error3	ssQLASSO	15.00(0.00)	1.20(2.34)	0.97(0.06)	0.97(0.06)	0.73(0.17)
		QLASSO	15.00(0.00)	7.17(2.12)	0.81(0.05)	0.82(0.04)	4.44(0.57)
		ssLASSO	14.47(1.68)	0.23(0.57)	0.97(0.07)	0.97(0.06)	1.37(1.41)
		LASSO	14.93(0.37)	74.53(23.36)	0.30(0.06)	0.41(0.05)	5.28(1.42)
	Error4	ssQLASSO	15.00(0.00)	1.13(2.45)	0.97(0.06)	0.97(0.06)	0.81(0.17)
		QLASSO	15.00(0.00)	5.90(2.04)	0.84(0.05)	0.85(0.04)	4.54(0.69)
		ssLASSO	14.63(1.22)	0.03(0.18)	0.98(0.05)	0.99(0.05)	1.42(0.78)
		LASSO	15.00(0.00)	85.73(31.70)	0.28(0.08)	0.40(0.07)	5.89(1.23)
	Error5	ssQLASSO	15.00(0.00)	0.70(1.39)	0.98(0.04)	0.98(0.04)	0.80(0.20)
		QLASSO	15.00(0.00)	7.10(2.84)	0.81(0.06)	0.83(0.05)	4.71(0.64)
		ssLASSO	15.00(0.00)	0.20(0.41)	0.99(0.01)	0.99(0.01)	0.92(0.15)
		LASSO	15.00(0.00)	83.20(31.07)	0.28(0.07)	0.40(0.06)	4.92(0.93)

Table B.19: Evaluation of heterogeneous model with clinical factors under AR-1 correlation, (n, p)=(800,3200), 100 replicates.

$n = 800$	$p = 3200$		TP	FP	F1	MCC	Estimation
$\tau = 0.3$	Error1	ssQLASSO	15.00(0.00)	3.63(4.80)	0.91(0.10)	0.91(0.09)	0.88(0.21)
		QLASSO	15.00(0.00)	3.33(1.24)	0.90(0.03)	0.91(0.03)	2.67(0.29)
		ssLASSO	15.00(0.00)	0.07(0.25)	1.00(0.01)	1.00(0.01)	1.11(0.16)
		LASSO	15.00(0.00)	79.57(33.40)	0.30(0.08)	0.41(0.07)	3.65(0.79)
	Error2	ssQLASSO	15.00(0.00)	0.87(2.18)	0.98(0.06)	0.98(0.05)	0.87(0.14)
		QLASSO	15.00(0.00)	7.33(2.97)	0.81(0.06)	0.82(0.06)	3.60(0.52)
		ssLASSO	13.87(1.91)	0.07(0.25)	0.95(0.09)	0.96(0.08)	2.92(1.50)
		LASSO	14.13(2.18)	74.60(24.61)	0.28(0.05)	0.39(0.05)	9.04(2.00)
	Error3	ssQLASSO	15.00(0.00)	0.80(1.73)	0.98(0.05)	0.98(0.04)	0.61(0.14)
		QLASSO	15.00(0.00)	6.13(2.78)	0.83(0.06)	0.85(0.06)	1.75(0.27)
		ssLASSO	15.00(0.00)	0.00(0.00)	1.00(0.00)	1.00(0.00)	2.38(0.39)
		LASSO	15.00(0.00)	86.13(30.41)	0.27(0.07)	0.39(0.06)	8.17(1.81)
	Error4	ssQLASSO	15.00(0.00)	0.33(1.12)	0.99(0.03)	0.99(0.03)	0.93(0.16)
		QLASSO	15.00(0.00)	6.70(2.61)	0.82(0.06)	0.83(0.05)	4.14(0.53)
		ssLASSO	12.50(2.11)	0.00(0.00)	0.90(0.09)	0.91(0.08)	3.56(1.37)
		LASSO	14.90(0.31)	93.27(37.21)	0.26(0.08)	0.38(0.07)	8.91(2.14)
	Error5	ssQLASSO	15.00(0.00)	0.13(0.43)	1.00(0.01)	1.00(0.01)	1.02(0.17)
		QLASSO	14.97(0.18)	7.03(3.01)	0.81(0.06)	0.83(0.06)	4.41(0.76)
		ssLASSO	13.47(2.42)	0.00(0.00)	0.94(0.11)	0.94(0.10)	2.90(1.55)
		LASSO	15.00(0.00)	80.77(29.00)	0.29(0.07)	0.40(0.06)	7.23(1.18)
$\tau = 0.5$	Error1	ssQLASSO	15.00(0.00)	2.70(3.34)	0.93(0.09)	0.93(0.08)	0.43(0.13)
		QLASSO	15.00(0.00)	3.83(2.34)	0.89(0.06)	0.90(0.05)	2.68(0.37)
		ssLASSO	15.00(0.00)	0.27(0.78)	0.99(0.02)	0.99(0.02)	0.60(0.13)
		LASSO	15.00(0.00)	92.53(25.26)	0.26(0.06)	0.38(0.05)	3.63(0.61)
	Error2	ssQLASSO	15.00(0.00)	0.70(1.44)	0.98(0.04)	0.98(0.04)	0.50(0.10)
		QLASSO	15.00(0.00)	14.50(3.52)	0.68(0.05)	0.72(0.04)	2.91(0.52)
		ssLASSO	12.73(2.26)	0.07(0.25)	0.91(0.10)	0.91(0.09)	2.97(1.48)
		LASSO	13.73(2.64)	72.67(37.03)	0.29(0.08)	0.39(0.08)	9.03(2.58)
	Error3	ssQLASSO	15.00(0.00)	0.60(1.10)	0.98(0.03)	0.98(0.03)	1.30(0.11)
		QLASSO	15.00(0.00)	15.30(3.60)	0.67(0.05)	0.71(0.04)	2.70(0.23)
		ssLASSO	14.97(0.18)	0.00(0.00)	1.00(0.01)	1.00(0.01)	2.81(0.40)
		LASSO	14.90(0.40)	75.33(18.20)	0.29(0.05)	0.41(0.04)	7.91(1.36)
	Error4	ssQLASSO	15.00(0.00)	0.40(1.16)	0.99(0.03)	0.99(0.03)	0.50(0.13)
		QLASSO	15.00(0.00)	15.07(4.62)	0.67(0.07)	0.71(0.06)	3.35(0.53)
		ssLASSO	12.93(2.20)	0.00(0.00)	0.92(0.09)	0.92(0.08)	2.61(1.30)
		LASSO	14.97(0.18)	74.37(18.13)	0.30(0.05)	0.41(0.05)	7.74(1.25)
	Error5	ssQLASSO	15.00(0.00)	0.33(0.99)	0.99(0.03)	0.99(0.03)	0.52(0.14)
		QLASSO	15.00(0.00)	14.63(3.45)	0.68(0.05)	0.71(0.04)	3.30(0.40)
		ssLASSO	12.50(2.93)	3.90(21.36)	0.87(0.18)	0.88(0.16)	3.81(4.66)
		LASSO	14.97(0.18)	90.60(37.38)	0.27(0.09)	0.39(0.07)	7.27(1.77)
$\tau = 0.7$	Error1	ssQLASSO	15.00(0.00)	2.13(2.42)	0.94(0.06)	0.94(0.06)	0.91(0.17)
		QLASSO	14.27(0.45)	3.87(1.81)	0.86(0.04)	0.87(0.04)	3.18(0.31)
		ssLASSO	14.53(0.51)	1.03(1.54)	0.95(0.04)	0.95(0.04)	1.22(0.14)
		LASSO	14.70(0.47)	77.13(34.37)	0.30(0.08)	0.41(0.07)	3.82(0.82)
	Error2	ssQLASSO	15.00(0.00)	0.90(1.58)	0.97(0.05)	0.97(0.04)	1.04(0.25)
		QLASSO	14.97(0.18)	7.47(2.78)	0.80(0.06)	0.82(0.05)	4.52(0.61)
		ssLASSO	12.97(1.56)	0.00(0.00)	0.92(0.06)	0.93(0.06)	3.07(1.21)
		LASSO	12.67(3.92)	68.73(29.95)	0.26(0.07)	0.36(0.09)	9.26(2.21)
	Error3	ssQLASSO	14.97(0.18)	1.23(1.98)	0.96(0.06)	0.96(0.05)	1.11(0.19)
		QLASSO	14.93(0.25)	5.83(1.82)	0.84(0.04)	0.85(0.04)	4.26(0.71)
		ssLASSO	14.57(0.73)	0.00(0.00)	0.98(0.03)	0.99(0.03)	1.89(0.67)
		LASSO	14.80(0.66)	83.10(34.75)	0.28(0.07)	0.40(0.06)	7.86(2.65)
	Error4	ssQLASSO	14.67(0.48)	1.30(1.95)	0.95(0.05)	0.95(0.05)	1.30(0.20)
		QLASSO	14.53(0.51)	6.07(1.98)	0.82(0.04)	0.83(0.04)	4.90(0.74)
		ssLASSO	12.33(2.20)	0.00(0.00)	0.89(0.11)	0.90(0.09)	2.97(1.35)
		LASSO	14.00(0.37)	78.90(25.38)	0.27(0.06)	0.38(0.05)	8.41(1.64)
	Error5	ssQLASSO	14.50(0.51)	1.47(2.39)	0.94(0.06)	0.94(0.05)	1.31(0.16)
		QLASSO	14.47(0.51)	7.20(2.99)	0.79(0.07)	0.81(0.06)	4.97(0.48)
		ssLASSO	13.17(1.32)	0.00(0.00)	0.93(0.05)	0.94(0.05)	2.41(1.01)
		LASSO	14.03(0.18)	83.37(34.40)	0.27(0.07)	0.38(0.06)	7.32(1.63)

B.6.2 Computational time under different study sizes

The computational times (in minutes) under different combinations of n and p have been demonstrated in Figure B.4.

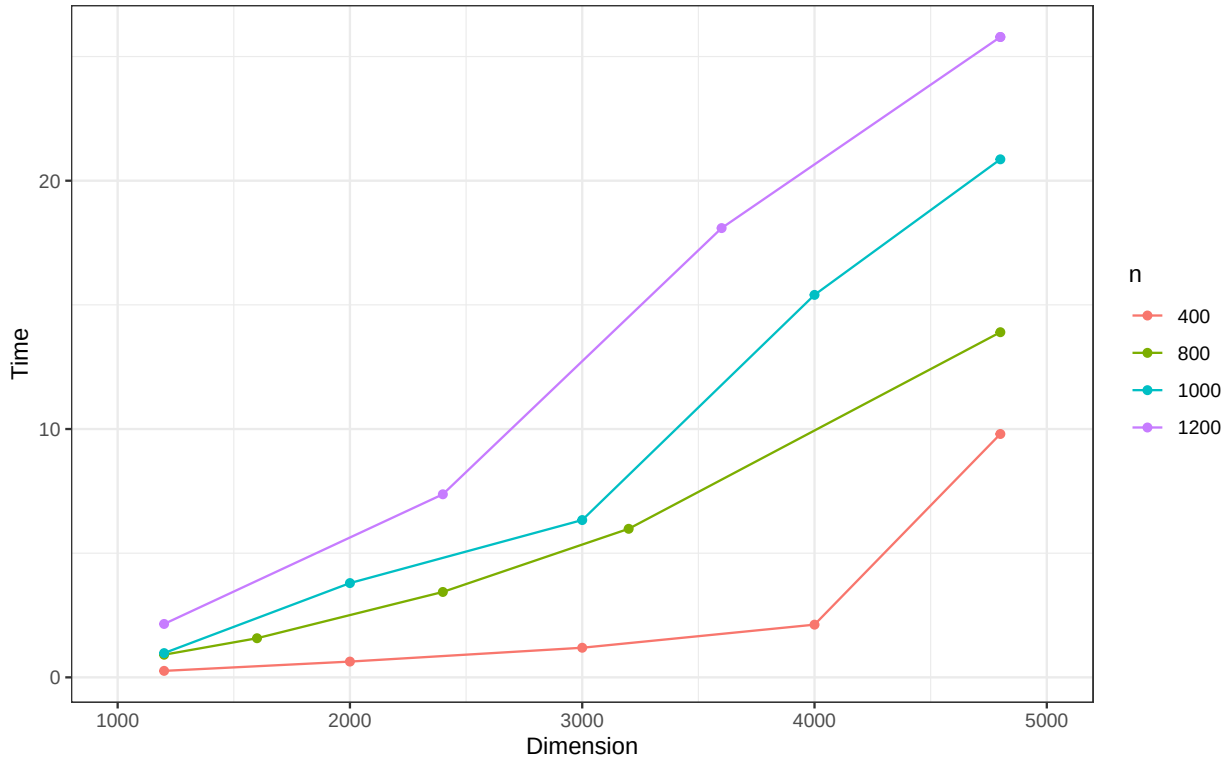

Figure B.4: The computational time (in minutes averaged across 100 replicates) of ssQLASSO under $\tau = 30\%$, AR-1 correlation and Error2 in terms of n and p.

B.6.3 Adaptivity of Asymmetric Laplace Distribution

The proposed ssQLASSO has rooted in quantile regression that have been widely examined in both the frequentist and Bayesian frameworks. In published studies from both frameworks, quantile regression based variable selection methods generally do not require a test to check whether it can be applied since model assessment in high-dimensional settings are mainly based on prediction and identification, as shown in the simulation study and case studies. Nevertheless, it is still necessary to evaluate how well the asymmetric Laplace distribution (ALD) adapts to the widely adopted heavy-tailed errors from published literature.

As described in Section 2.1, we assume that the model error ϵ_i $(i = 1, ..., n)$ follows the asymmetric Laplace distribution (ALD) with a fixed quantile level τ and the quantile check loss function[19]:

$$f(\epsilon_i|\mu_i, \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma}\exp\left(-\rho_\tau\left(\frac{\epsilon_i - \mu}{\sigma}\right)\right),$$

where μ is the location parameter and $\sigma(> 0)$ is the scale parameter to determine the the skewness of ALD. In simulation and case studies, although the distributional assumptions on ALD are violated, the ssQLASSO has still demonstrated superior performance, which is consistent with the observations on robustness of ALD–based likelihood to violations of a diversity of distributional assumptions from published studies[135].

To further justify that the asymmetric Laplace distribution (ALD) is appropriate to modeling a variety of heavy-tailed errors, we have performed a heuristic analysis by checking the maximum overlapping density between the heavy-tailed errors and best fitted ALD which is determined by finding the largest overlapping areas (OV). Figure B.5 and Figure B.6 show that regardless of the type of distribution and quantile level, the overlapping density between ALD and other errors are around or above 90%, indicating that ALD can satisfactorily approximate the target heavy-tailed errors. Therefore, if the location and scale parameters can be accurately estimated, ssQLASSO is expected to yield superior performance as shown in the simulation and case studies. Please note that, Figure B.6 is corresponding to the heterogeneous errors which are highly skewed and not $i.i.d$ even under Error 1 (the normal error). We have provided the R codes to reproduce the analysis.

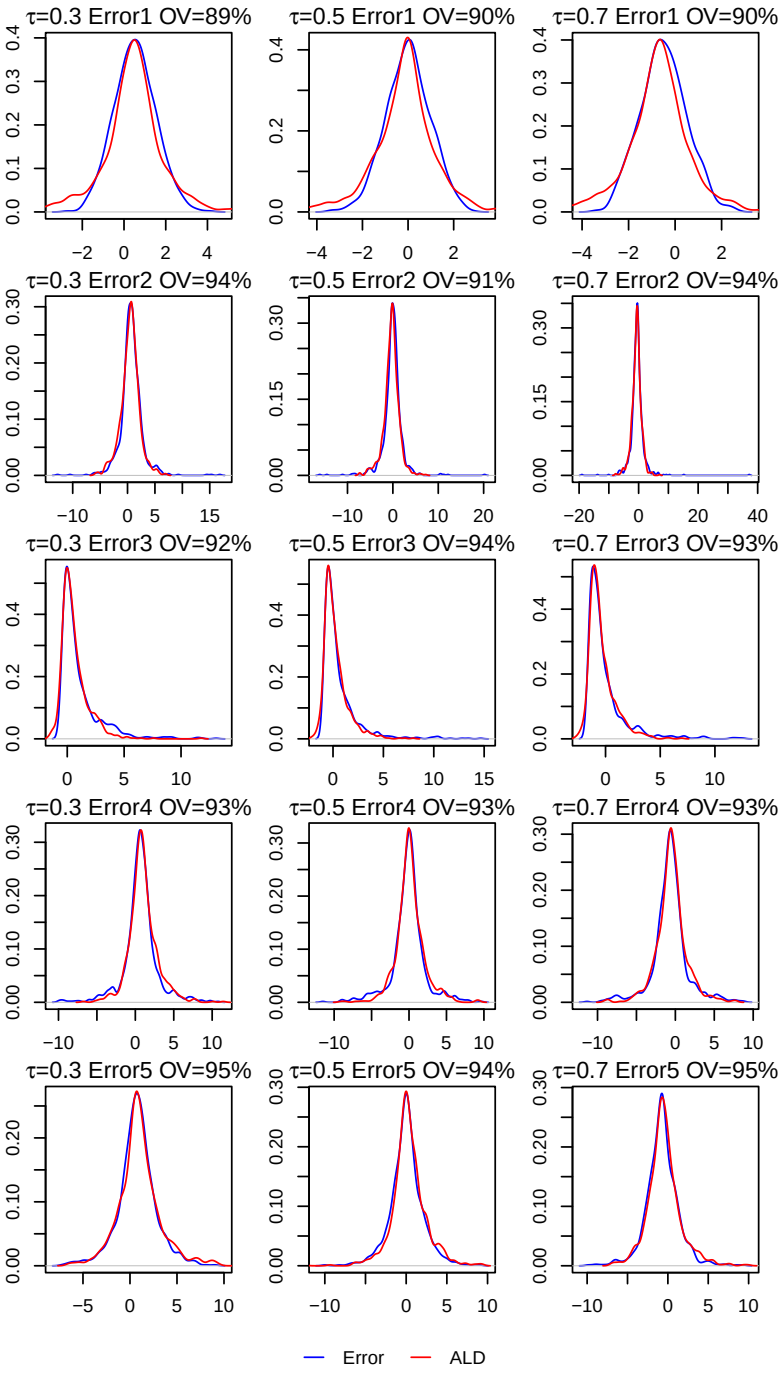

Figure B.5: Density plots and overlapping areas of homogeneous error distributions and the corresponding best fitted ALD based on 800 random samples.

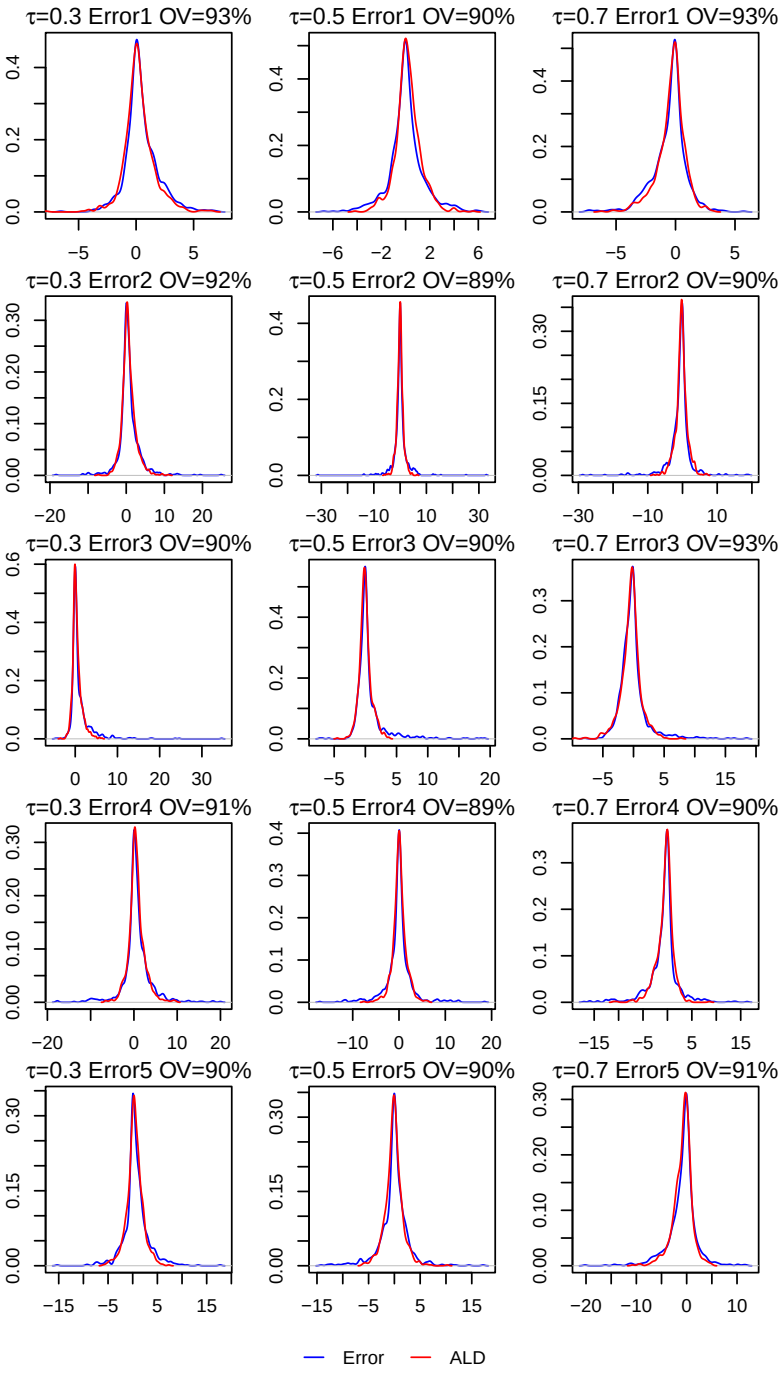

Figure B.6: Density plots and overlapping areas of heterogeneous error distributions and the corresponding best fitted ALD based on 800 random samples.

```
library(ald)
library(rmutil)
library(overlapping)
library(lattice)

set.seed(1)
n = 800
quant = 0.7

# Laplace distribution & homogeneous model
x0 = rlaplace(n, 0, sqrt(2))
x1 = x0 − quantile(x0,probs = quant)

l1 <− seq(0.1,0.9,length.out=20)
l2 <− seq(0.1,0.9,length.out=20)

# Select proper sigma and p at max overlapping area
M <− matrix(0,20,20)
for(i in 1:20){
  for(j in 1:20){
    x2 <− rALD(n,mu=mean(x1),sigma=l1[i],p=l2[j])
    x <− list(x1,x2)
    M[i,j] <− as.numeric(overlap(x))
  }
}
indices <− which(M == max(M), arr.ind=TRUE)
lambda1 <− l1[indices[1]]
```

```
lambda2 <- l2[indices[2]]


# Plot density
x2 <- rALD(n,mu=mean(x1),sigma=lambda1,p=lambda2)
plot( density(x1), col="blue",main=paste("OV=",ceiling(100*M[indices]),"%"),xlab="")
lines( density(x2), col="red")
```

B.6.4 Summary statistics of outliers in simulation

The model errors generated from the simulation have been widely adopted in published literature on high-dimensional robust methods including those based on quantile regressions, such as[18;22;96;116;135] among many others. Therefore, it is critical to assess the performance of ssQLASSO on these representative settings which have been used as a common testbed for high-dimensional robust methods. Here, Table B.20 provides the summary statistics on outliers from data generated from the simulation, where the outliers have been determined by using the built-in R function *boxplot.stats*. We can observe that heterogeneous settings have produced a much larger number of outliers compared to their counterpart settings under the homogeneous errors, including the normal distributions.

Table B.20: Number of outliers under 5 error distributions in homogeneous and heterogeneous models across 100 replicates.

Count		$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$
homogeneous	Error1	2.98(1.89)	2.92(2.14)	2.90(1.93)
n=400	Error2	32.49(4.91)	32.83(5.45)	32.84(5.56)
	Error3	31.01(4.62)	30.40(4.84)	30.68(4.82)
	Error4	44.04(5.40)	43.06(5.93)	43.76(5.66)
	Error5	25.05(5.70)	25.24(6.04)	24.80(5.61)
n=800	Error1	5.82(2.87)	5.54(2.66)	5.82(3.06)
	Error2	66.01(7.74)	64.74(8.41)	67.40(8.96)
	Error3	62.57(6.65)	61.19(7.17)	61.81(7.09)
	Error4	87.22(7.06)	87.03(7.77)	87.97(9.80)
	Error5	49.42(8.42)	50.97(7.21)	49.69(8.15)
heterogeneous	Error1	31.41(6.45)	42.54(7.49)	31.99(6.19)
n=400	Error2	48.06(7.30)	58.89(7.60)	49.20(7.17)
	Error3	49.16(6.57)	50.82(6.12)	32.29(6.08)
	Error4	50.46(7.27)	58.30(6.63)	50.71(7.94)
	Error5	45.27(7.99)	58.49(8.09)	45.56(7.54)
n=800	Error1	61.86(10.01)	83.32(11.07)	63.45(9.38)
	Error2	96.09(10.77)	116.64(10.83)	96.54(10.17)
	Error3	97.06(9.38)	100.53(8.88)	71.90(7.90)
	Error4	100.12(10.18)	119.98(10.90)	102.78(11.72)
	Error5	92.71(10.81)	117.15(12.53)	92.24(10.21)

B.6.5 Additional real data analysis

We have also applied the proposed ssQLASSO under quantile level $\tau = 0.3$ and $\tau = 0.7$. The number of overlapping findings have been provided in Table B.21 and B.22 for LUAD and SKCM, respectively. For LUAD, ssQLASSO results in an PMAD of 14.03 (sd1.87), and an PMSE of 334.73 (sd83.50) under $\tau = 0.3$. It leads to an PMAD of 13.97 (sd2.77), and an PMSE of 341.34 (sd128.87) under $\tau = 0.7$. For SKCM, the PMAD and PMSE of ssQLASSO (τ=0.3) are 0.75 (sd0.09) and 1.10 (sd0.46), respectively. Under $\tau = 0.7$, PMAD and PMSE are 0.77 (sd0.09) and 1.19 (sd0.49). The prediction performances are comparable to ssQLASSO under $\tau = 0.5$.

Analysis of LUAD shows that the non-robust ssLASSO has identified 12 genes, ALPL, MSANTD4, IKZF4, SBDS, S100A1, RRN3P3, AGPAT3, DIRC2, POLR3G, FNBP1, ACSF2, TOE1, which have also been detected by at least one of the other 3 methods except DIRC2 and TOE1. However, DIRC2, or Disrupted in renal carcinoma 2, has been found to be associated with the development of renal cancer (Bodmer et al.,[136] and Savalas et al.,[137]). And TOE1 has been well acknowledged to play a key role in the progression of Pontocerebella Hypoplasia Type 7 (PCH7) (Deng et al.,[138] and Lardelli et al.[139]). None of the two genes is related to lung functions. On the other hand, among genes only detected by ssQLASSO method, LUAD related gene SRP14 and SMC3 have been identified by ssQLASSO across all the 3 quantile levels.

For SKCM, genes MAP4K4, GJA3, INPP1, and KCTD7 have been uniquely found by the proposed ssQLASSO. Among them, the first three (MAP4K4, GJA3 and INPP1) have been detected at $\tau = 0.3$, and the last two (INPP1 and KCTD7) have been identified at $\tau = 0.7$. As we have analyze in Section of Real Data Analysis, these genes have been verified in scientific literature regarding their roles in the formation and progression of melanoma. The non-robust ssLASSO results in the discovery of 8 genes (ADA, FLJ20373, PER1, LOC126807, COQ9, ARMCX2, ZFAT, KCTD21), including the uniquely identified PER1 and ARMCX2. The plausible associations between the two genes with melanoma have been rarely reported. For example, ARMCX2 has been found as a prognostic marker for gastric cancer[140] and a

key player in the acquired methylation in ovarian tumors[141].

Table B.21: Identification results for real data analysis. The numbers of genes identified by different methods and their overlaps.

LUAD

	ssQLASSO(0.3)	ssQLASSO(0.5)	ssQLASSO(0.7)	QLASSO	ssLASSO	LASSO
ssQLASSO(0.3)	15	5	1	4	6	7
ssQLASSO(0.5)		17	4	12	3	7
ssQLASSO(0.7)			16	6	1	3
QLASSO				30	6	9
ssLASSO					12	9
LASSO						51

Table B.22: Identification results for real data analysis. The numbers of genes identified by different methods and their overlaps.

SKCM

	ssQLASSO(0.3)	ssQLASSO(0.5)	ssQLASSO(0.7)	QLASSO	ssLASSO	LASSO
ssQLASSO(0.3)	10	4	3	1	2	4
ssQLASSO(0.5)		13	5	2	4	8
ssQLASSO(0.7)			13	1	3	4
QLASSO				17	1	6
ssLASSO					8	4
LASSO						27

Appendix C

Appendices for Chapter 4

C.1 Details on Alternative Methods

Alternative methods under comparison are spike-and-slab group LASSO (SSVCM), regularized varying coefficient model with adaptive group LASSO under the quantile check loss (QRVC-adp) and least square loss (VC-adp). The details of the derivation of SSVCM are provided in the following section.

C.2 The Spike-and-Slab Group LASSO

Denote $y_i(t_{ik})$ as the response of the i-th subject ($i = 1, \ldots, n$) at time t_{ik} ($k = 1, \ldots, K$), consider the following varying-coefficient data generating model:

$$y_i(t_{ik}) = \sum_{l=1}^{q} c_i^{(l)}(t_{ik})\alpha_l(t_{ik}) + \sum_{j=1}^{p} x_i^{(j)}(t_{ik})\beta_j(t_{ik}) + \boldsymbol{w}_i(t_{ik})^\top \boldsymbol{\gamma}_i + \epsilon_i,$$

where $\boldsymbol{c}_i(t_{ik}) = (c_i^{(1)}(t_{ik}), \ldots, c_i^{(q)}(t_{ik}))^\top$ is a q-dimensional vector representing the intercept and clinical factors, $\boldsymbol{x}_i(t_{ik}) = (x_i^{(1)}(t_{ik}), \ldots, x_i^{(p)}(t_{ik}))^\top$ is a p-dimensional vector of genetic factors with $\boldsymbol{\alpha}(t)$ and $\boldsymbol{\beta}(t)$ being their corresponding fixed effect varying coefficient vectors. For the random effect part, $\boldsymbol{w}_i(t_{ik}) = (w_i^{(1)}(t_{ik}), \ldots, w_i^{(r)}(t_{ik}))^\top$ is a p_n-dimensional vector

with $\boldsymbol{\gamma}_i$ being the random effect coefficients. Random error $\epsilon_i(t_{ik})$'s are assumed to be independent.

For the genetic factor part, group spike-and-slab prior is assigned on each component of $\boldsymbol{\theta}$ as:

$$\boldsymbol{\theta}_j|\delta_j, \lambda_0, \lambda_1 \overset{ind}{\sim} (1-\delta_j)\Pi(\boldsymbol{\theta}_j|\lambda_0) + \delta_j\Pi(\boldsymbol{\theta}_j|\lambda_1),$$

with the spike-and-slab components $\Pi(\boldsymbol{\theta}_j|\lambda_0)$ and $\Pi(\boldsymbol{\theta}_j|\lambda_1)$ both defined based on the group LASSO density in (4.4). With $\lambda_0 > \lambda_1$, the scale parameter s_0 controls the spike part which models irrelevant coefficients that have close-to-0 coefficients, and s_1 controls the slab distribution which models important effects with large coefficients magnitude.

The indicator variable δ_j is assumed to follow independent Bernoulli distributions :

$$\delta_j|\psi \overset{ind}{\sim} \text{Ber}(\psi) = \psi^{\delta_j}(1-\psi)^{1-\delta_j},$$

with probability parameter ψ being assigned a uniform prior $\psi \sim U(0,1)$. It indicates the prior probability of nonzero components in $\boldsymbol{\theta}$.

The intercept and each clinical factor $\boldsymbol{\xi}_l$ $(l = 1, ..., q)$ is assumed to have independent multivariate normal priors with mean 0 and covariance Σ_l:

$$\xi_l \overset{ind}{\sim} \text{N}_{d_n}(0, \Sigma_l).$$

For variance $\tilde{\sigma}^2$, we assume an improper prior:

$$\pi(\tilde{\sigma}^2) = \tilde{\sigma}^{-2}.$$

The random effect related priors are proposed as follows:

$$\boldsymbol{\gamma}_i|\phi \sim N_{p_n}(0, \phi^2\boldsymbol{I})),$$
$$\phi^2 \sim \text{IG}(e, f),$$

where p_n-dimensional random effect vector $\boldsymbol{\gamma}_i = (\gamma_{i,1}, ..., \gamma_{i,p_n})^\top$ follows multivariate normal distribution with mean 0 and covariance matrix $\phi^2 \boldsymbol{I}$. An inverse-gamma prior is assigned to ϕ^2 with $e > 0$ and $f > 0$ being constants.

Let δ_j and $\Lambda_j = (1-\delta_j)\lambda_0 + \delta_j \lambda_1$ being treated as missing, the log joint posterior density function of SSVCM after omitting the irrelevant terms is given by

$$
\begin{aligned}
Q &= \log p(\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \tilde{\sigma}^2, \boldsymbol{\delta}, \psi, \phi^2 | \boldsymbol{Y}) \\
&= \log p(\boldsymbol{Y} | \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \tilde{\sigma}^2, \phi^2) + \sum_{l=1}^{q} \log p(\xi_l) + \sum_{j=1}^{p} \log p(\theta_j | \Lambda_j) + \log p(\tilde{\sigma}^2) + \sum_{j=1}^{p} \log p(\delta_j | \psi) \\
&+ \log p(\psi) + \sum_{i=1}^{n} \sum_{g=1}^{p_n} p(\gamma_{ig}) + \log p(\phi^2) \\
&\propto -\frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{K} \log(\tilde{\sigma}^2) - \frac{1}{2\tilde{\sigma}^2} \sum_{i=1}^{n} \sum_{k=1}^{K} \left(y_i(t_{ik} - \boldsymbol{A}_i(t_{ik})^\top \boldsymbol{\xi} - \boldsymbol{B}_i(t_{ik})^\top \boldsymbol{\theta} - \boldsymbol{w}_i(t_{ik})^\top \boldsymbol{\gamma}_i \right) \\
&\quad - \sum_{l=1}^{q} \frac{\boldsymbol{\xi}_l^\top \Sigma_l^{-1} \boldsymbol{\xi}_l}{2} - \sum_{j=1}^{p} \Lambda_j \|\boldsymbol{\theta}_j\|_2 + \sum_{j=1}^{p} [\delta_j \log\psi + (1-\delta_j)\log(1-\psi)] - \frac{n \times p_n}{2} \log(\phi^2) \\
&\quad - \log(\tilde{\sigma}^2) - \sum_{i=1}^{n} \sum_{g=1}^{p_n} \frac{(\gamma_{ig})^2}{2\phi^2} - \frac{f}{\phi^2} - (e+1)\log(\phi^2).
\end{aligned}
\tag{C.1}
$$

At E-step, we calculate the conditional posterior expectations of δ_j at d-th iteration:

$$
\begin{aligned}
\eta_j^{(d)} &= p(\delta_j = 1 | \boldsymbol{\theta}_j^{(d)}, \psi^{(d)}, \boldsymbol{y}) \\
&= \frac{p(\boldsymbol{\theta}_j^{(d)} | \delta_j = 1, \lambda_1) p(\delta_j = 1 | \psi^{(d)})}{p(\boldsymbol{\theta}_j^{(d)} | \delta_j = 0, \lambda_0) p(\delta_j = 0 | \psi^{(d)}) + p(\boldsymbol{\theta}_j^{(d)} | \delta_j = 1, \lambda_1) p(\delta_j = 1 | \psi^{(d)})} \\
&= \frac{\Pi(\boldsymbol{\theta}_j^{(d)} | \lambda_1) \psi^{(d)}}{\Pi(\boldsymbol{\theta}_j^{(d)} | \lambda_0)(1 - \psi^{(d)}) + \Pi(\boldsymbol{\theta}_j^{(d)} | \lambda_1) \psi^{(d)}},
\end{aligned}
\tag{C.2}
$$

Therefore, the conditional expectation of $(\Lambda_j^{-1})^{(d)}$ can be computed as:

$$
\tilde{\Lambda}_j^{(d)} = E(\Lambda_j^{(d)} | \boldsymbol{\theta}_j^{(d)}) = E\left((1-\delta_j)\lambda_0 + \delta_j \lambda_1 \middle| \boldsymbol{\theta}_j^{(d)} \right) = \left(1 - \eta_j^{(d)}\right) \lambda_0 + \left(\eta_j^{(d)}\right) \lambda_1. \tag{C.3}
$$

At M-step, we update $\boldsymbol{\xi}^{(d+1)}$, $\boldsymbol{\theta}^{(d+1)}$, $\boldsymbol{\gamma}^{(d+1)}$, $(\tilde{\sigma}^2)^{(d+1)}$, $\psi^{(d+1)}$ and $(\phi^2)^{(d+1)}$ by maximizing the log joint posterior density at the d-th iteration, or $Q(\boldsymbol{\Phi}|\boldsymbol{\Phi}^{(d)})$ where $\boldsymbol{\Phi} = (\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \tilde{\sigma}^2, \boldsymbol{\delta}, \psi, \phi^2)$. By solving $\frac{\partial Q(\boldsymbol{\Phi}|\boldsymbol{\Phi}^{(d)})}{\partial \boldsymbol{\xi}_{\tilde{l}}} = 0$, $\frac{\partial Q(\boldsymbol{\Phi}|\boldsymbol{\Phi}^{(d)})}{\partial \boldsymbol{\theta}_{\tilde{j}}} = 0$, $\frac{\partial Q(\boldsymbol{\Phi}|\boldsymbol{\Phi}^{(d)})}{\partial \gamma_{\tilde{i}\tilde{g}}} = 0$, $\frac{\partial Q(\boldsymbol{\Phi}|\boldsymbol{\Phi}^{(d)})}{\partial \tilde{\sigma}^2} = 0$, $\frac{\partial Q(\boldsymbol{\Phi}|\boldsymbol{\Phi}^{(d)})}{\partial \psi} = 0$ and $\frac{\partial Q(\boldsymbol{\Phi}|\boldsymbol{\Phi}^{(d)})}{\partial \phi^2} = 0$, where $\tilde{l} = 1, ..., q$, $\tilde{j} = 1, ..., p$, $\tilde{i} = 1, ..., n$ and $\tilde{g} = 1, ..., p_n$, we obtain the following updating equations:

$$
\begin{aligned}
\psi^{(d+1)} &= \frac{1}{p}\sum_{j=1}^{p}\eta_j^{(d)}, \\
(\tilde{\sigma}^2)^{(d+1)} &= \frac{\sum_{i=1}^{n}\sum_{k=1}^{K}\left(y_i(t_{ik}) - \boldsymbol{A}_i(t_{ik})^\top\boldsymbol{\xi}^{(d)} - \boldsymbol{B}_i(t_{ik})^\top\boldsymbol{\theta}^{(d)} - \boldsymbol{w}_i(t_{ik})^\top\boldsymbol{\gamma}_i^{(d)}\right)}{\sum_{i=1}^{n}\sum_{k=1}^{K}1 + 2}, \\
(\phi^2)^{(d+1)} &= \frac{\sum_{i=1}^{n}\sum_{g=1}^{p_n}\frac{\left(\gamma_{ig}^{(d)}\right)^2}{2} + f}{\frac{n\cdot p_n}{2} + e + 1}, \\
\gamma_{\tilde{i}\tilde{g}}^{(d+1)} &= \frac{\frac{1}{(\tilde{\sigma}^2)^{(d+1)}}\sum_{k=1}^{K}\left(\boldsymbol{R2}_{\tilde{i}\tilde{g}}^{(d+1)}\right)w_{\tilde{i}}^{(\tilde{g})}(t_{\tilde{i}k})}{\frac{1}{(\tilde{\sigma}^2)^{(d+1)}}\sum_{k=1}^{K}\left(w_{\tilde{i}}^{(\tilde{g})}(t_{\tilde{i}k})\right)^2 + \frac{1}{(\phi^2)^{(d+1)}}}, \\
\boldsymbol{\xi}_{\tilde{l}}^{(d+1)} &= \left(\tilde{\boldsymbol{Z}}_{\tilde{l}}^\top\tilde{\boldsymbol{Z}}_{\tilde{l}} + 2(\tilde{\sigma}^2)^{(d+1)}\Sigma_{\tilde{l}}^{-1}\right)^{-1}\sum_{i=1}^{n}\sum_{k=1}^{K}\left(\boldsymbol{R3}_{\tilde{l}}^{(d+1)}\right)A_i^{(\tilde{l})}(t_{ik}), \\
\boldsymbol{\theta}_{\tilde{j}}^{(d+1)} &= \left(\frac{\|\boldsymbol{z}_{\tilde{j}}\boldsymbol{U}_{\tilde{j}}^{(d+1)}\|_2 - (\tilde{\sigma}^2)^{(d+1)}\tilde{\Lambda}_{\tilde{j}}^{(d)}}{\|\boldsymbol{Z}_{\tilde{j}}\boldsymbol{U}_{\tilde{j}}^{(d+1)}\|_2}\right)_+\boldsymbol{U}_{\tilde{j}}^{(d+1)},
\end{aligned}
\tag{C.4}
$$

where

$$
\begin{aligned}
\boldsymbol{R2}_{\tilde{i}\tilde{g}}^{(d+1)} &= y_{\tilde{i}}(t_{\tilde{i}k}) - \boldsymbol{A}_{\tilde{i}}(t_{\tilde{i}k})^\top\boldsymbol{\xi}^{(d)} - \boldsymbol{B}_{\tilde{i}}(t_{\tilde{i}k})^\top\boldsymbol{\theta}^{(d)} - \sum_{g=1}^{\tilde{g}-1}w_{\tilde{i}}^{(g)}(t_{\tilde{i}k})\gamma_{\tilde{i}g}^{(d+1)} - \sum_{g=\tilde{g}+1}^{p_n}w_{\tilde{i}}^{(g)}(t_{\tilde{i}k})\gamma_{\tilde{i}g}^{(d)}, \\
\boldsymbol{R3}_{\tilde{l}}^{(d+1)} &= y_i(t_{ik}) - \sum_{l=1}^{\tilde{l}-1}A_i^{(l)}(t_{ik})\boldsymbol{\xi}_l^{(d+1)} - \sum_{l=\tilde{l}+1}^{q}A_i^{(l)}(t_{ik})\boldsymbol{\xi}_l^{(d)} - \boldsymbol{B}_i(t_{ik})^\top\boldsymbol{\theta}^{(d)} - \boldsymbol{w}_i(t_{ik})^\top\boldsymbol{\gamma}_i^{(d)}, \\
\boldsymbol{U}_{\tilde{j}}^{(d+1)} &= \left(\boldsymbol{Z}_{\tilde{j}}^\top\boldsymbol{Z}_{\tilde{j}}\right)^{-1}\sum_{i=1}^{n}\sum_{k=1}^{K}\left(\left(\boldsymbol{R4}_i^{(\tilde{j})}(t_{ik})^{(d+1)}\right)B_i^{(\tilde{j})}(t_{ik})\right), \\
\boldsymbol{R4}_i^{(\tilde{j})}(t_{ik})^{(d+1)} &= y_i(t_{ik}) - \boldsymbol{A}_i(t_{ik})^\top\boldsymbol{\xi}^{(d+1)} - \sum_{j=1}^{\tilde{j}-1}\boldsymbol{B}_i^{(j)}(t_{ik})^\top\boldsymbol{\theta}_j^{(d+1)} + \sum_{j=\tilde{j}+1}^{p}\boldsymbol{B}_i^{(j)}(t_{ik})^\top\boldsymbol{\theta}_j^{(d)} \\
&\quad - \boldsymbol{w}_i(t_{ik})^\top\boldsymbol{\gamma}_i^{(d+1)},
\end{aligned}
$$

in which $\boldsymbol{Z}_{\tilde{j}}$ is a $(n \times K) \times d_n$-dimensional matrix representing the $\tilde{j}$-th group of genetic effects. Similarly, $\tilde{\boldsymbol{Z}}_{\tilde{l}}$ is also a $(n \times K) \times d_n$-dimensional matrix denoting the $\tilde{l}$-th group of the intercept and clinical effects.

Variable selection is not performed on the varying intercept and clinical factors. Thus, shrinkage is only imposed on $\boldsymbol{\theta}_j$. A k-fold cross-validation is used to choose the optimal tuning parameters. Under the fixed tuning parameters λ_0 and λ_1, the model fitting algorithm is provided in the table below:

Table C.1: EM algorithm for the spike-and-slab group LASSO.

Algorithm
Initialize iteration $d = 0$, coefficients $\boldsymbol{\xi}^{(0)}$, $\boldsymbol{\theta}^{(0)}$, $(\tilde{\sigma}^2)^{(0)}$, $(\phi^2)^{(0)}$, $\psi^{(0)}$ and $\sigma^{(0)}$;
Repeat
E-step:
Compute the conditional expectation of $\delta_j^{(d)}$ and $\Lambda_j^{(d)}$ via (C.2) and (C.3);
M-step:
Update $\psi^{(d+1)}$, $(\tilde{\sigma}^2)^{(d+1)}$, $(\phi^2)^{(d+1)}$, $\gamma_{\tilde{i}\tilde{g}}^{(d+1)}$, $\boldsymbol{\xi}_{\tilde{l}}^{(d+1)}$ and $\boldsymbol{\theta}_{\tilde{j}}^{(d+1)}$ through (C.4);
Calculate the updated log joint posterior density and check if $\lvert Q^{(d+1)} - Q^{(d)} \rvert < \varepsilon$;
$d \leftarrow d + 1$;
until Convergence ($\lvert Q^{(d+1)} - Q^{(d)} \rvert < \varepsilon$ for a small value $\varepsilon > 0$).

C.3 Additional Plots

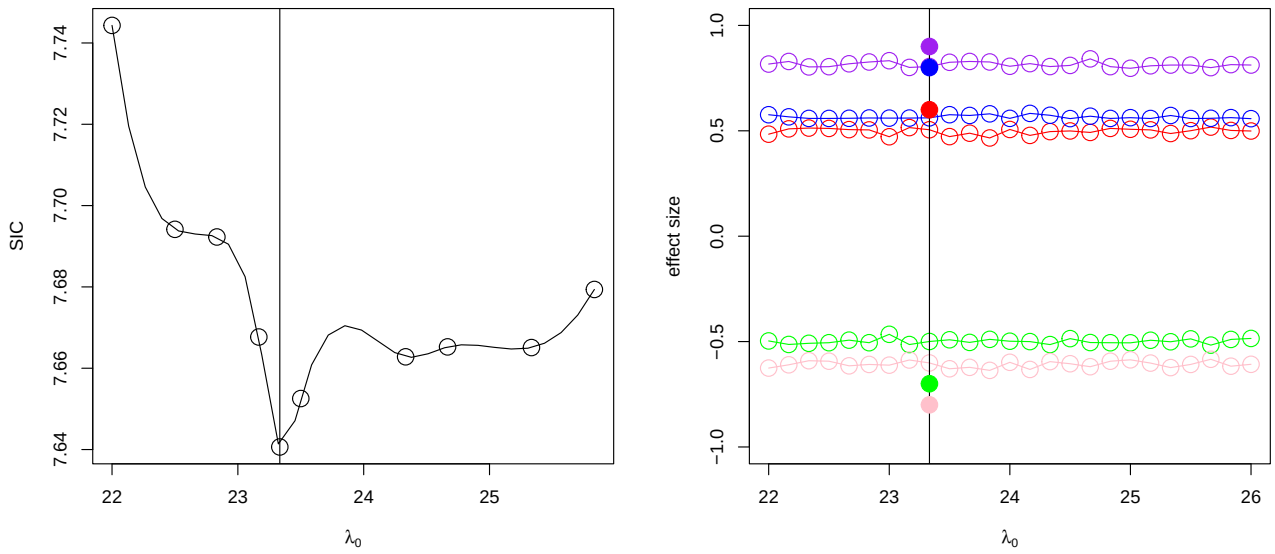

Figure C.1: The solution path of SSQVCM under $\tau = 0.3$, AR-1 correlation, $n = 200$,$p = 400$, group size equal to 5.

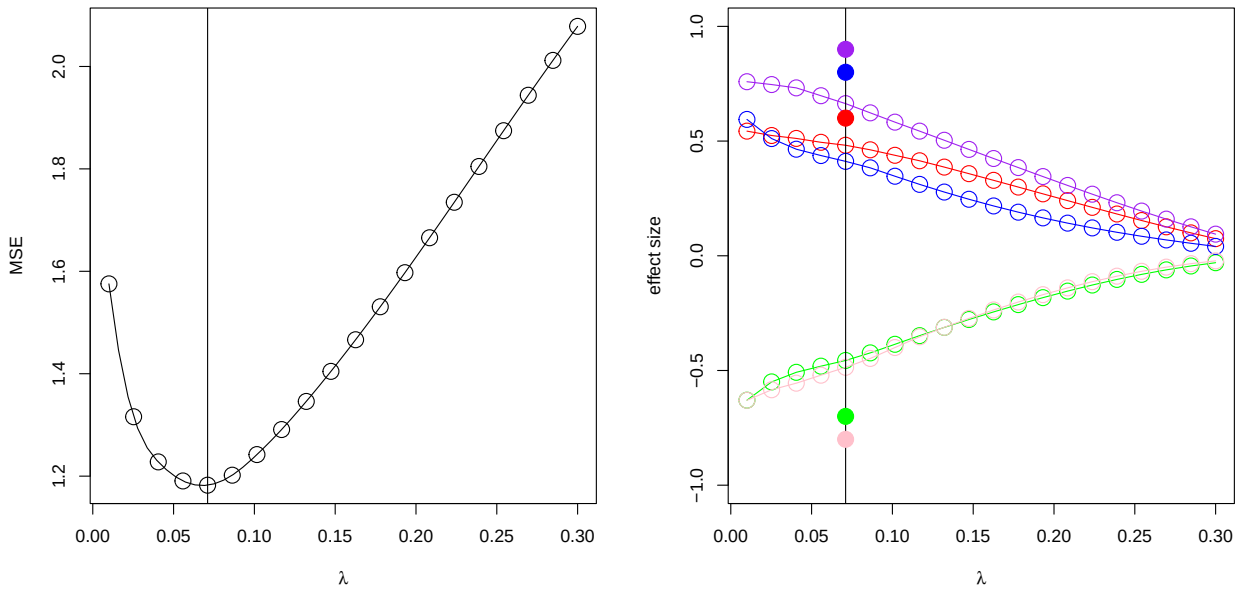

Figure C.2: The MSE profile (left) and solution path (right) of group LASSO under $\tau = 0.5\%$, AR-1 correlation, $n = 200$,$p = 400$, group size equal to 5.

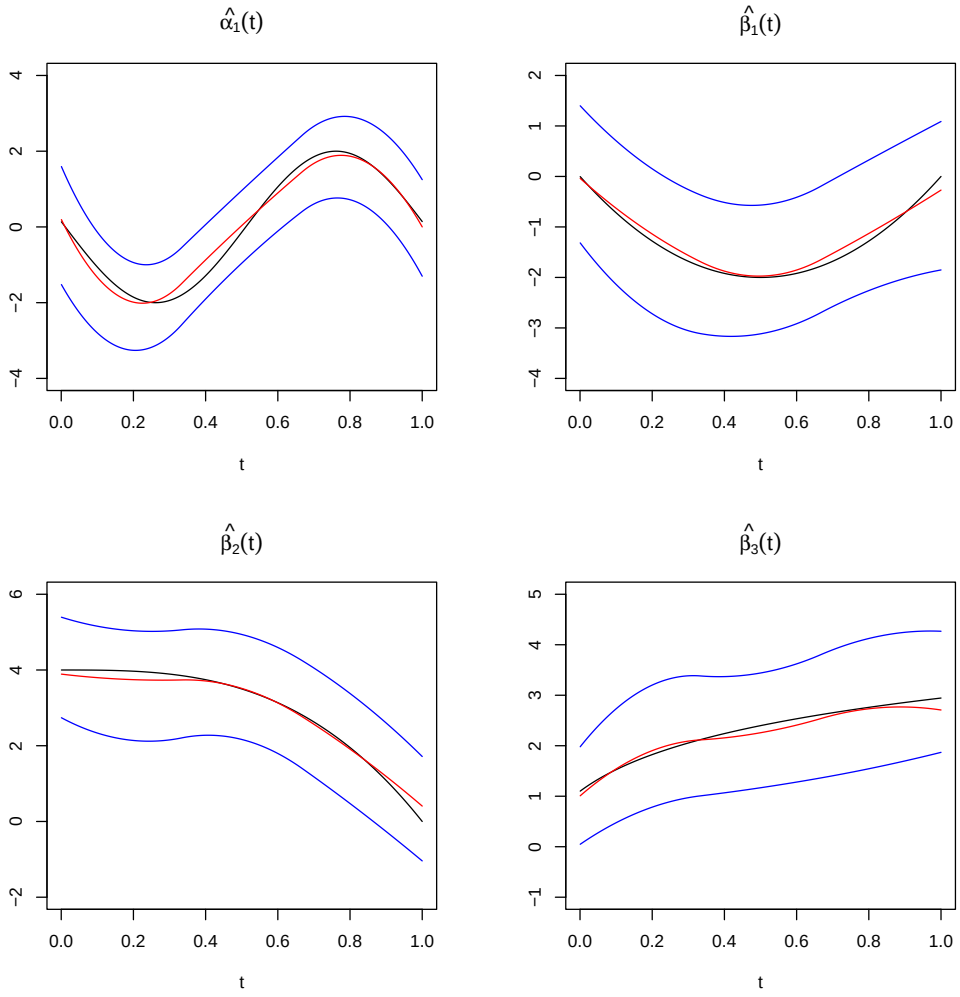

Figure C.3: Estimation of non-zero varying coefficients under Error 2, $\tau = 0.3$, AR-1 correlation and $(n, p) = (400, 1000)$ with 100 replicates. Red line: true varying coefficients. Black line: estimated varying coefficients from SSQVCM. Blue lines: 95% bootstrap confidence intervals for the estimated varying coefficients.

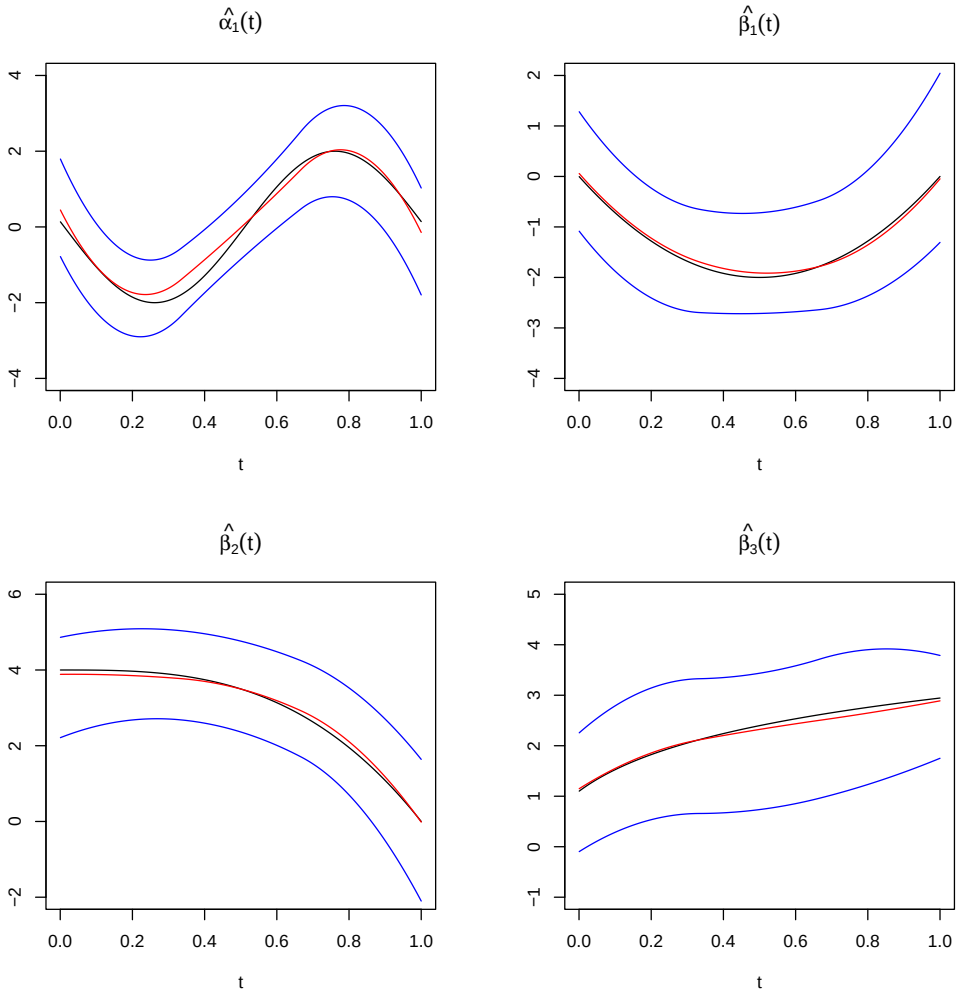

Figure C.4: Estimation of non-zero varying coefficients under Error 2, $\tau = 0.3$, LD structure and $(n, p) = (400, 1000)$ with 100 replicates. Red line: true varying coefficients. Black line: estimated varying coefficients from SSQVCM. Blue lines: 95% bootstrap confidence intervals for the estimated varying coefficients.

C.4 Computational Time

Table C.2: Computational time (in minutes) of homogeneous model under AR-1 correlation, (n, p)=(400,1000) with 100 replicates.

Run time		$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$
Error1	sQGL	11.19(0.54)	9.44(2.51)	13.51(0.21)
	SSVCM	7.59(7.57)	8.88(4.63)	8.12(5.04)
	QVC	14.07(0.54)	13.66(0.68)	17.04(0.65)
	VC	6.74(0.87)	7.93(0.77)	7.32(0.93)
Error2	sQGL	12.05(0.67)	8.39(2.40)	12.87(0.86)
	SSVCM	8.11(3.72)	8.65(3.81)	9.89(4.69)
	QVC	14.85(0.62)	12.02(0.79)	16.76(0.76)
	VC	7.13(1.09)	7.24(0.81)	8.01(1.05)
Error3	sQGL	11.50(0.45)	8.95(2.52)	13.23(0.47)
	SSVCM	7.32(8.59)	7.97(5.02)	9.63(4.34)
	QVC	14.16(0.44)	13.09(0.63)	16.53(0.60)
	VC	7.01(0.66)	7.79(1.07)	7.57(0.88)
Error4	sQGL	12.72(0.59)	9.34(2.53)	12.84(0.67)
	SSVCM	7.45(4.56)	9.77(3.73)	8.50(4.13)
	QVC	15.63(1.02)	12.61(0.79)	16.92(1.74)
	VC	7.23(1.00)	7.67(0.86)	8.02(0.91)
Error5	sQGL	19.92(4.18)	15.50(4.49)	20.64(5.27)
	SSVCM	12.88(8.39)	17.95(8.48)	12.41(8.60)
	QVC	25.07(4.36)	21.01(2.87)	25.38(4.38)
	VC	11.31(2.03)	12.33(2.25)	12.58(2.82)

C.5 Sensitivity Analysis

Table C.3: Sensitivity analysis of homogeneous model under LD correlation, and $\tau = 0.3$ with 100 replicates.

$D = 2$	k_n	1	2	3	4	5
$\tau = 0.3$	TP	2.83(0.38)	2.83(0.38)	2.90(0.30)	2.88(0.33)	2.86(0.35)
Normal	FP	0.12(0.46)	0.33(0.66)	0.10(0.30)	0.10(0.30)	0.10(0.30)
	F1	0.96(0.07)	0.94(0.08)	0.97(0.05)	0.97(0.05)	0.97(0.06)
	MCC	0.96(0.06)	0.94(0.08)	0.98(0.05)	0.97(0.05)	0.97(0.05)
	TIMSE	1.12(0.51)	1.33(0.49)	0.99(0.42)	1.02(0.45)	1.05(0.48)
$D = 3$	k_n	1	2	3	4	5
$\tau = 0.3$	TP	2.98(0.16)	2.95(0.22)	2.95(0.22)	2.93(0.26)	2.91(0.29)
Normal	FP	0.32(0.57)	0.95(1.04)	0.51(0.78)	0.50(0.77)	0.49(0.77)
	F1	0.96(0.06)	0.90(0.09)	0.94(0.08)	0.94(0.08)	0.94(0.08)
	MCC	0.96(0.06)	0.90(0.08)	0.94(0.07)	0.94(0.07)	0.94(0.07)
	TIMSE	1.27(0.46)	1.50(0.47)	1.26(0.47)	1.28(0.48)	1.30(0.51)
$D = 2$	k_n	1	2	3	4	5
$\tau = 0.3$	TP	2.85(0.36)	2.79(0.41)	2.81(0.40)	2.79(0.42)	2.79(0.41)
Laplace	FP	0.15(0.36)	0.33(0.64)	0.37(0.63)	0.36(0.62)	0.38(0.62)
	F1	0.96(0.06)	0.94(0.08)	0.93(0.08)	0.93(0.08)	0.93(0.08)
	MCC	0.96(0.06)	0.94(0.08)	0.94(0.08)	0.93(0.08)	0.93(0.08)
	TIMSE	1.42(0.53)	1.53(0.57)	1.52(0.59)	1.55(0.61)	1.59(0.62)
$D = 3$	k_n	1	2	3	4	5
$\tau = 0.3$	TP	2.85(0.37)	2.88(0.33)	2.85(0.36)	2.82(0.39)	2.80(0.41)
Laplace	FP	0.38(0.64)	0.46(0.90)	0.44(0.89)	0.43(0.88)	0.43(0.86)
	F1	0.94(0.08)	0.94(0.09)	0.93(0.09)	0.93(0.09)	0.93(0.09)
	MCC	0.94(0.08)	0.94(0.08)	0.94(0.08)	0.94(0.08)	0.93(0.08)
	TIMSE	1.48(0.56)	1.47(0.56)	1.51(0.58)	1.54(0.60)	1.62(0.66)

Table C.4: Sensitivity analysis of homogeneous model under LD correlation, $(n, p) = (400, 1000)$ and $\tau = 0.5$ with 100 replicates.

$D = 2$	k_n	1	2	3	4	5
$\tau = 0.5$	TP	2.88(0.33)	2.83(0.38)	2.85(0.36)	2.83(0.38)	2.81(0.39)
Normal	FP	0.17(0.45)	0.33(0.71)	0.17(0.44)	0.17(0.44)	0.16(0.43)
	F1	0.96(0.06)	0.94(0.08)	0.96(0.06)	0.96(0.06)	0.96(0.07)
	MCC	0.97(0.06)	0.95(0.07)	0.96(0.06)	0.96(0.06)	0.96(0.06)
	TIMSE	0.92(0.42)	0.95(0.53)	0.94(0.44)	0.96(0.45)	0.98(0.47)
$D = 3$	k_n	1	2	3	4	5
$\tau = 0.5$	TP	2.95(0.22)	2.92(0.27)	2.93(0.26)	2.90(0.30)	2.88(0.32)
Normal	FP	0.45(0.68)	0.67(0.90)	0.44(0.67)	0.43(0.67)	0.42(0.66)
	F1	0.95(0.07)	0.92(0.09)	0.94(0.07)	0.94(0.07)	0.94(0.07)
	MCC	0.95(0.07)	0.93(0.08)	0.95(0.07)	0.94(0.07)	0.94(0.07)
	TIMSE	1.16(0.47)	1.39(0.48)	1.18(0.50)	1.21(0.52)	1.24(0.55)
$D = 2$	k_n	1	2	3	4	5
$\tau = 0.5$	TP	2.89(0.32)	2.93(0.27)	2.86(0.35)	2.84(0.37)	2.82(0.39)
Laplace	FP	0.14(0.35)	0.37(0.69)	0.14(0.35)	0.13(0.34)	0.13(0.34)
	F1	0.97(0.06)	0.95(0.07)	0.97(0.06)	0.96(0.06)	0.96(0.06)
	MCC	0.97(0.05)	0.95(0.07)	0.97(0.05)	0.96(0.06)	0.96(0.06)
	TIMSE	0.94(0.49)	1.15(0.40)	0.98(0.52)	1.01(0.56)	1.05(0.59)
$D = 3$	k_n	1	2	3	4	5
$\tau = 0.5$	TP	2.92(0.28)	3.00(0.00)	2.88(0.33)	2.89(0.32)	2.86(0.36)
Laplace	FP	0.32(0.63)	1.06(0.68)	0.31(0.62)	0.30(0.61)	0.29(0.60)
	F1	0.95(0.07)	0.89(0.07)	0.95(0.07)	0.95(0.07)	0.95(0.07)
	MCC	0.96(0.07)	0.90(0.06)	0.95(0.07)	0.96(0.07)	0.95(0.07)
	TIMSE	1.07(0.43)	1.17(0.26)	1.11(0.47)	1.15(0.50)	1.18(0.52)

C.6 Additional Real Data Results

Table C.5: Information of the 25 SNPs selected by SSQVCM under $\tau = 0.3$.

SNP name	Chromosome	Chr. position	Gene	Risk allele	MAF
rs12967226	18	47611845	MIR4527HG	A	14.04%
rs10517296	4	53390964	FIP1L1	T	19.18%
rs11793082	9	34088703	DCAF12	T	5.82%
rs9992485	4	112390016	ALPK1	G	36.76%
rs9833058	3	132493149	DNAJC13	A	15.41%
rs17123149	1	100673223	LOC124904231	A	27.28%
rs4935872	11	124205288	OR8G2P	G	16.10%
rs10744193	12	125085651	AACS	C	15.30%
rs1106121	4	37483651	C4orf19	A	14.95%
rs1006330	11	101755870	PLS1P1	G	13.70%
rs11134686	5	157322168	CYFIP2	G	40.75%
rs16988435	21	31686258	SCAF4	T	23.06%
rs7874261	9	112210999	SUSD1	A	6.51%
rs2053797	2	46143753	PRKCE	A	18.95%
rs1282945	3	111916272	PHLDB2	A	9.82%
rs9979853	21	35669231	MIR802	G	43.26%
rs1557812	17	37674363	LOC105371755	A	49.32%
rs3094297	3	101680416	RPL24	C	36.76%
rs16950876	16	52176886	LOC105371261	A	49.32%
rs16855688	3	171054889	SLC2A2	G	49.77%
rs1567664	20	53307536	TSHZ2	G	21.12%
rs10498717	6	23725576	LOC105374976	G	22.03%
rs17187617	1	208603220	RPS26P13	T	13.13%
rs16948067	15	93463669	LOC124900612	G	8.11%
rs9945084	18	43142930	RIT2	A	41.21%

Table C.6: Information of the 25 SNPs selected by SSQVCM under $\tau = 0.7$.

SNP name	Chromosome	Chr. position	Gene	Risk allele	MAF
rs12967226	18	47611845	MIR4527HG	A	14.04%
rs10517296	4	53390964	FIP1L1	T	19.18%
rs1542434	12	114026868	LOC105369993	T	11.19%
rs11872877	18	57861027	ATP8B1	C	16.32%
rs10502644	18	35161529	MAPRE2	A	13.70%
rs1949060	10	65162957	MYL6P3	C	49.66%
rs12054225	3	3622669	LOC100130207	C	23.97%
rs4719963	7	28971448	CPVL-AS2	T	43.26%
rs1537852	1	222186904	LOC124904518	A	26.94%
rs2293504	7	70784922	AUTS2	T	31.19%
rs554697	6	167794176	LINC01558	T	25.00%
rs17128993	10	111523013	RPS6P15	T	14.38%
rs12516456	5	71626789	MCCC2	G	9.25%
rs925078	1	91196463	HFM1	G	49.54%
rs2738148	8	6897397	RPL23AP96	C	24.20%
rs10965890	9	23516290	LOC101929563	C	44.86%
rs17742449	3	7909261	LOC124909342	G	36.53%
rs3094297	3	101680416	RPL24	C	36.76%
rs10973450	9	37642237	FRMPD1	C	10.84%
rs6874971	5	16156431	MARCHF11	C	16.44%
rs16860073	3	112811940	CD200R1L-AS1	T	28.88%
rs2612659	4	174512187	HPGD	C	7.76%
rs17445774	2	199994526	LOC124906112	T	13.70%
rs16948067	15	93463669	LOC124900612	G	8.11%
rs12027986	1	119028642	WARS2-IT1	G	37.90%

C.7 Additional Numerical Results

Table C.7: Evaluation of homogeneous model under LD correlation, $(n, p) = (400, 1000)$ with 100 replicates.

$n = 400$	$p = 1000$		TP	FP	F1	MCC	TIMSE
$\tau = 0.3$	Error1	SSQVCM	3.00(0.00)	0.13(0.35)	0.99(0.04)	0.99(0.04)	0.73(0.19)
		SSVCM	3.00(0.00)	3.02(1.72)	0.74(0.12)	0.77(0.10)	1.29(0.67)
		QVC	3.00(0.00)	0.86(2.62)	0.92(0.10)	0.92(0.09)	8.03(1.88)
		VC	3.00(0.00)	4.74(1.23)	0.65(0.14)	0.70(0.11)	2.61(1.08)
	Error2	SSQVCM	3.00(0.00)	0.13(0.35)	0.99(0.04)	0.99(0.04)	0.86(0.26)
		SSVCM	2.92(0.27)	2.88(1.55)	0.74(0.11)	0.76(0.10)	1.52(0.77)
		QVC	2.92(0.27)	1.18(4.98)	0.89(0.14)	0.90(0.12)	9.12(2.39)
		VC	2.94(0.24)	6.32(1.84)	0.62(0.19)	0.67(0.16)	3.56(1.84)
	Error3	SSQVCM	3.00(0.00)	0.17(0.38)	0.98(0.04)	0.98(0.04)	1.36(0.39)
		SSVCM	2.90(0.30)	3.30(1.98)	0.72(0.13)	0.74(0.12)	2.26(0.92)
		QVC	2.92(0.27)	0.64(3.26)	0.93(0.10)	0.93(0.09)	6.26(1.71)
		VC	2.98(0.14)	4.82(1.12)	0.66(0.16)	0.70(0.13)	3.77(1.51)
	Error4	SSQVCM	3.00(0.00)	0.08(0.28)	0.99(0.03)	0.99(0.03)	0.91(0.23)
		SSVCM	2.93(0.25)	3.14(2.26)	0.73(0.13)	0.76(0.11)	1.62(0.80)
		QVC	2.93(0.25)	1.05(4.73)	0.90(0.13)	0.91(0.11)	9.02(2.47)
		VC	2.93(0.25)	5.09(1.78)	0.66(0.18)	0.71(0.15)	3.18(1.73)
	Error5	SSQVCM	3.00(0.00)	0.21(0.41)	0.98(0.05)	0.98(0.04)	1.07(0.31)
		SSVCM	2.93(0.25)	2.75(1.74)	0.75(0.11)	0.78(0.10)	1.60(0.54)
		QVC	2.95(0.21)	0.86(5.77)	0.91(0.11)	0.92(0.10)	8.54(2.51)
		VC	2.98(0.15)	6.86(1.41)	0.60(0.18)	0.65(0.15)	3.50(1.73)
$\tau = 0.5$	Error1	SSQVCM	3.00(0.00)	0.23(0.43)	0.97(0.05)	0.98(0.05)	0.69(0.19)
		SSVCM	2.94(0.24)	3.24(1.82)	0.72(0.13)	0.75(0.11)	1.26(0.64)
		QVC	2.94(0.24)	0.82(3.62)	0.91(0.10)	0.92(0.09)	1.06(0.56)
		VC	2.94(0.24)	4.52(1.37)	0.67(0.15)	0.71(0.12)	2.33(1.13)
	Error2	SSQVCM	3.00(0.00)	0.30(0.53)	0.97(0.06)	0.97(0.05)	0.79(0.24)
		SSVCM	2.94(0.24)	3.36(3.23)	0.73(0.14)	0.76(0.12)	1.71(1.51)
		QVC	2.96(0.20)	1.08(4.87)	0.90(0.13)	0.91(0.12)	1.14(0.60)
		VC	2.96(0.20)	5.62(1.66)	0.64(0.17)	0.68(0.14)	3.19(2.11)
	Error3	SSQVCM	3.00(0.00)	0.27(0.52)	0.97(0.06)	0.97(0.05)	0.82(0.29)
		SSVCM	2.98(0.14)	3.30(2.15)	0.73(0.13)	0.76(0.11)	1.58(0.82)
		QVC	3.00(0.00)	1.28(4.94)	0.88(0.13)	0.89(0.11)	1.31(0.63)
		VC	3.00(0.00)	5.88(1.67)	0.63(0.18)	0.68(0.15)	3.17(1.62)
	Error4	SSQVCM	3.00(0.00)	0.07(0.25)	0.99(0.03)	0.99(0.03)	0.62(0.19)
		SSVCM	2.98(0.14)	3.22(2.22)	0.74(0.14)	0.77(0.12)	1.14(0.61)
		QVC	2.98(0.14)	1.14(5.35)	0.90(0.13)	0.91(0.12)	1.10(0.64)
		VC	2.98(0.14)	6.70(1.82)	0.61(0.21)	0.67(0.17)	3.07(1.78)
	Error5	SSQVCM	3.00(0.00)	0.11(0.32)	0.99(0.04)	0.99(0.03)	0.67(0.17)
		SSVCM	2.96(0.20)	2.91(1.75)	0.75(0.12)	0.77(0.11)	1.22(0.53)
		QVC	2.98(0.15)	1.06(4.33)	0.90(0.12)	0.91(0.11)	1.09(0.47)
		VC	3.00(0.00)	5.15(1.44)	0.67(0.19)	0.71(0.16)	2.45(1.38)
$\tau = 0.7$	Error1	SSQVCM	3.00(0.00)	0.23(0.43)	0.97(0.05)	0.98(0.05)	0.75(0.28)
		SSVCM	2.96(0.20)	3.12(1.94)	0.74(0.13)	0.76(0.11)	1.33(0.57)
		QVC	3.00(0.00)	1.02(4.23)	0.91(0.12)	0.91(0.11)	8.55(2.18)
		VC	3.00(0.00)	5.16(1.49)	0.66(0.17)	0.70(0.13)	2.69(1.31)
	Error2	SSQVCM	3.00(0.00)	0.32(0.56)	0.97(0.06)	0.97(0.06)	0.91(0.25)
		SSVCM	2.96(0.21)	3.71(3.37)	0.71(0.14)	0.74(0.12)	2.12(2.14)
		QVC	2.93(0.25)	0.93(6.81)	0.90(0.11)	0.91(0.10)	8.79(1.99)
		VC	2.93(0.25)	6.89(1.25)	0.61(0.20)	0.67(0.17)	4.18(3.33)
	Error3	SSQVCM	3.00(0.00)	0.07(0.25)	0.99(0.03)	0.99(0.03)	0.57(0.19)
		SSVCM	2.94(0.24)	2.86(1.64)	0.75(0.13)	0.77(0.11)	1.14(0.54)
		QVC	2.96(0.20)	0.96(3.69)	0.91(0.12)	0.91(0.11)	11.75(2.51)
		VC	2.98(0.14)	5.04(1.43)	0.65(0.14)	0.69(0.12)	2.55(1.20)
	Error4	SSQVCM	3.00(0.00)	0.17(0.39)	0.98(0.04)	0.98(0.04)	1.02(0.26)
		SSVCM	2.91(0.29)	3.02(1.95)	0.73(0.12)	0.76(0.11)	1.84(0.88)
		QVC	3.00(0.00)	1.23(4.79)	0.89(0.13)	0.90(0.11)	8.37(2.04)
		VC	3.00(0.00)	5.65(1.66)	0.64(0.18)	0.69(0.14)	3.39(1.63)
	Error5	SSQVCM	3.00(0.00)	0.10(0.40)	0.99(0.04)	0.99(0.04)	1.06(0.32)
		SSVCM	2.96(0.20)	2.50(1.74)	0.78(0.13)	0.80(0.11)	1.51(0.64)
		QVC	2.96(0.20)	0.80(4.25)	0.92(0.11)	0.92(0.10)	8.36(2.36)
		VC	2.98(0.14)	4.70(1.26)	0.69(0.19)	0.73(0.16)	2.90(1.55)

Table C.8: Evaluation of homogeneous model under real data correlation, $(n, p) = (400 \times 5, 1000 \times 5)$ with 100 replicates.

$n = 400 \times 5$	$p = 1000 \times 5$		TP	FP	F1	MCC	TIMSE
$\tau = 0.3$	Error1	SSQVCM	3.00(0.00)	0.27(0.45)	0.97(0.05)	0.97(0.05)	0.83(0.27)
		SSVCM	2.98(0.14)	3.40(2.34)	0.73(0.15)	0.76(0.12)	1.41(0.62)
		QVC	2.98(0.14)	0.94(3.70)	0.91(0.11)	0.92(0.10)	8.38(2.41)
		VC	2.98(0.14)	4.90(1.35)	0.66(0.17)	0.71(0.14)	2.72(1.23)
	Error2	SSQVCM	3.00(0.00)	0.04(0.20)	0.99(0.02)	0.99(0.02)	0.87(0.28)
		SSVCM	2.91(0.29)	3.23(2.96)	0.74(0.16)	0.77(0.14)	1.80(1.44)
		QVC	2.98(0.15)	0.86(4.65)	0.92(0.11)	0.92(0.10)	8.76(1.77)
		VC	2.98(0.15)	5.66(1.42)	0.64(0.18)	0.69(0.15)	3.37(2.10)
	Error3	SSQVCM	3.00(0.00)	0.37(0.49)	0.96(0.05)	0.96(0.05)	1.26(0.36)
		SSVCM	3.00(0.00)	2.10(1.53)	0.81(0.12)	0.83(0.10)	1.83(0.67)
		QVC	2.98(0.14)	1.04(3.76)	0.91(0.13)	0.92(0.12)	6.85(1.74)
		VC	3.00(0.00)	4.84(1.82)	0.67(0.18)	0.72(0.15)	3.39(1.41)
	Error4	SSQVCM	3.00(0.00)	0.24(0.44)	0.97(0.05)	0.97(0.05)	1.01(0.33)
		SSVCM	2.96(0.21)	2.76(1.75)	0.76(0.12)	0.78(0.11)	1.75(0.88)
		QVC	2.98(0.15)	0.89(7.15)	0.91(0.11)	0.92(0.10)	9.15(2.29)
		VC	2.96(0.21)	7.29(1.50)	0.62(0.24)	0.67(0.20)	3.56(2.03)
	Error5	SSQVCM	3.00(0.00)	0.30(0.47)	0.97(0.05)	0.97(0.05)	1.00(0.32)
		SSVCM	3.00(0.00)	2.88(2.13)	0.76(0.13)	0.79(0.11)	1.68(0.81)
		QVC	2.98(0.14)	1.64(8.11)	0.87(0.16)	0.88(0.14)	7.98(2.19)
		VC	3.00(0.00)	7.42(2.60)	0.60(0.20)	0.66(0.16)	3.64(2.32)
$\tau = 0.5$	Error1	SSQVCM	3.00(0.00)	0.17(0.38)	0.98(0.04)	0.98(0.04)	0.54(0.15)
		SSVCM	3.00(0.00)	3.02(1.76)	0.74(0.11)	0.77(0.09)	1.19(0.57)
		QVC	3.00(0.00)	0.76(4.36)	0.92(0.10)	0.93(0.09)	0.90(0.47)
		VC	3.00(0.00)	5.78(1.06)	0.63(0.18)	0.68(0.15)	2.62(1.53)
	Error2	SSQVCM	3.00(0.00)	0.20(0.42)	0.98(0.05)	0.98(0.04)	0.68(0.21)
		SSVCM	2.83(0.58)	3.33(3.52)	0.73(0.19)	0.75(0.18)	2.71(5.83)
		QVC	2.92(0.29)	1.33(3.32)	0.87(0.13)	0.88(0.12)	1.09(0.62)
		VC	2.83(0.58)	5.58(1.72)	0.60(0.15)	0.65(0.13)	4.02(3.88)
	Error3	SSQVCM	3.00(0.00)	0.20(0.41)	0.98(0.05)	0.98(0.04)	0.73(0.21)
		SSVCM	3.00(0.00)	2.68(1.38)	0.76(0.10)	0.79(0.08)	1.43(0.59)
		QVC	2.98(0.14)	0.86(4.29)	0.91(0.11)	0.92(0.10)	1.10(0.51)
		VC	2.98(0.14)	5.60(1.20)	0.64(0.19)	0.69(0.15)	3.04(1.47)
	Error4	SSQVCM	3.00(0.00)	0.23(0.43)	0.97(0.05)	0.98(0.05)	0.76(0.23)
		SSVCM	2.98(0.14)	3.16(1.97)	0.74(0.13)	0.76(0.11)	1.41(0.75)
		QVC	3.00(0.00)	1.08(5.40)	0.89(0.10)	0.90(0.09)	1.21(0.64)
		VC	3.00(0.00)	6.90(1.18)	0.60(0.19)	0.66(0.15)	3.18(1.87)
	Error5	SSQVCM	3.00(0.00)	0.15(0.37)	0.98(0.04)	0.98(0.04)	0.63(0.24)
		SSVCM	2.98(0.15)	3.39(1.95)	0.72(0.12)	0.75(0.10)	1.30(0.77)
		QVC	2.98(0.15)	1.28(6.27)	0.88(0.13)	0.89(0.12)	1.14(0.76)
		VC	2.98(0.15)	6.63(1.63)	0.61(0.18)	0.66(0.15)	2.89(1.69)
$\tau = 0.7$	Error1	SSQVCM	3.00(0.00)	0.27(0.45)	0.97(0.05)	0.97(0.05)	0.81(0.23)
		SSVCM	3.00(0.00)	2.76(1.70)	0.76(0.12)	0.79(0.10)	1.32(0.54)
		QVC	3.00(0.00)	0.74(4.58)	0.93(0.10)	0.93(0.09)	8.72(2.63)
		VC	3.00(0.00)	5.44(1.12)	0.66(0.19)	0.70(0.16)	2.77(1.45)
	Error2	SSQVCM	3.00(0.00)	0.30(0.60)	0.97(0.06)	0.97(0.06)	0.87(0.30)
		SSVCM	2.98(0.14)	3.18(2.87)	0.75(0.16)	0.78(0.14)	1.92(1.51)
		QVC	2.92(0.27)	1.12(6.28)	0.88(0.12)	0.89(0.11)	9.16(1.84)
		VC	2.94(0.24)	6.78(1.47)	0.61(0.20)	0.66(0.16)	3.94(2.61)
	Error3	SSQVCM	3.00(0.00)	0.47(0.57)	0.95(0.06)	0.95(0.06)	0.76(0.26)
		SSVCM	2.96(0.20)	3.36(2.03)	0.72(0.12)	0.75(0.11)	1.45(0.69)
		QVC	2.90(0.30)	1.20(5.35)	0.87(0.12)	0.88(0.11)	11.70(3.16)
		VC	2.96(0.20)	7.10(1.40)	0.59(0.21)	0.65(0.18)	3.27(1.77)
	Error4	SSQVCM	3.00(0.00)	0.30(0.47)	0.97(0.05)	0.97(0.05)	0.96(0.32)
		SSVCM	3.00(0.00)	2.65(1.70)	0.77(0.12)	0.79(0.10)	1.61(0.64)
		QVC	3.00(0.00)	0.70(4.19)	0.93(0.10)	0.94(0.09)	9.16(2.07)
		VC	3.00(0.00)	5.74(1.10)	0.63(0.18)	0.68(0.14)	3.20(1.52)
	Error5	SSQVCM	3.00(0.00)	0.16(0.37)	0.98(0.04)	0.98(0.04)	0.95(0.28)
		SSVCM	3.00(0.00)	2.96(2.03)	0.75(0.13)	0.78(0.11)	1.60(0.76)
		QVC	3.00(0.00)	1.00(6.52)	0.91(0.12)	0.92(0.11)	8.41(2.02)
		VC	3.00(0.00)	6.69(1.48)	0.63(0.22)	0.68(0.18)	3.48(2.22)

Table C.9: Evaluation of heterogeneous model under AR-1 correlation, $(n, p) = (400, 1000)$ with 100 replicates.

$n = 400$	$p = 1000$		TP	FP	F1	MCC	TIMSE
$\tau = 0.3$	Error1	SSQVCM	3.00(0.00)	0.13(0.35)	0.99(0.04)	0.99(0.04)	1.03(0.20)
		SSVCM	2.86(0.35)	2.66(1.73)	0.75(0.13)	0.77(0.11)	1.32(0.58)
		QVC	2.86(0.35)	1.22(3.67)	0.88(0.14)	0.89(0.12)	8.85(1.99)
		VC	2.94(0.24)	5.26(1.97)	0.64(0.16)	0.68(0.13)	2.26(0.82)
	Error2	SSQVCM	2.67(0.48)	0.33(0.61)	0.92(0.07)	0.92(0.07)	1.75(0.27)
		SSVCM	2.64(0.48)	3.00(2.61)	0.71(0.14)	0.73(0.12)	2.41(1.86)
		QVC	2.64(0.48)	0.94(4.28)	0.86(0.11)	0.86(0.10)	9.82(2.76)
		VC	2.60(0.49)	5.28(1.15)	0.60(0.15)	0.64(0.13)	3.80(1.51)
	Error3	SSQVCM	2.63(0.49)	0.17(0.38)	0.93(0.08)	0.93(0.08)	2.00(0.35)
		SSVCM	2.58(0.50)	2.74(1.74)	0.71(0.13)	0.73(0.12)	2.82(0.77)
		QVC	2.64(0.48)	1.10(3.26)	0.85(0.12)	0.86(0.11)	7.84(1.93)
		VC	2.70(0.46)	4.82(1.76)	0.62(0.14)	0.67(0.12)	4.14(1.06)
	Error4	SSQVCM	2.37(0.49)	0.13(0.35)	0.89(0.08)	0.90(0.07)	1.91(0.30)
		SSVCM	2.26(0.44)	2.76(1.97)	0.67(0.13)	0.69(0.13)	2.50(0.81)
		QVC	2.40(0.49)	1.36(4.73)	0.80(0.12)	0.81(0.11)	10.48(2.41)
		VC	2.44(0.50)	5.90(1.74)	0.57(0.17)	0.61(0.15)	4.57(1.84)
	Error5	SSQVCM	2.47(0.51)	0.40(0.72)	0.88(0.09)	0.89(0.09)	1.99(0.47)
		SSVCM	2.44(0.50)	2.98(1.81)	0.68(0.13)	0.70(0.12)	2.53(0.93)
		QVC	2.50(0.51)	1.26(5.66)	0.82(0.13)	0.83(0.12)	9.50(2.49)
		VC	2.56(0.50)	6.44(1.61)	0.57(0.17)	0.61(0.14)	4.66(2.21)
$\tau = 0.5$	Error1	SSQVCM	2.50(0.51)	0.40(0.77)	0.89(0.08)	0.89(0.07)	1.48(0.55)
		SSVCM	2.32(0.47)	2.92(1.99)	0.67(0.12)	0.69(0.12)	2.33(1.00)
		QVC	2.50(0.51)	1.06(5.14)	0.83(0.11)	0.84(0.10)	1.81(0.85)
		VC	2.48(0.50)	5.70(1.25)	0.59(0.18)	0.63(0.15)	3.47(1.64)
	Error2	SSQVCM	2.60(0.50)	0.50(0.90)	0.89(0.09)	0.90(0.08)	1.54(0.63)
		SSVCM	2.40(0.53)	5.22(10.19)	0.62(0.17)	0.65(0.16)	5.61(19.06)
		QVC	2.48(0.65)	1.42(3.97)	0.80(0.14)	0.81(0.13)	2.12(1.67)
		VC	2.30(0.76)	5.24(1.79)	0.56(0.16)	0.60(0.14)	5.08(3.02)
	Error3	SSQVCM	2.63(0.49)	0.33(0.61)	0.91(0.08)	0.92(0.07)	1.48(0.42)
		SSVCM	2.46(0.50)	2.82(1.92)	0.69(0.13)	0.71(0.12)	2.22(0.85)
		QVC	2.64(0.48)	0.96(4.91)	0.86(0.10)	0.87(0.10)	1.86(0.65)
		VC	2.54(0.50)	6.14(1.23)	0.57(0.17)	0.61(0.14)	4.23(1.58)
	Error4	SSQVCM	2.87(0.35)	0.43(0.77)	0.94(0.08)	0.94(0.07)	1.18(0.54)
		SSVCM	2.58(0.50)	3.20(2.17)	0.69(0.15)	0.71(0.14)	2.18(1.10)
		QVC	2.54(0.50)	1.66(4.67)	0.79(0.14)	0.80(0.13)	2.09(1.01)
		VC	2.58(0.50)	6.64(1.84)	0.54(0.14)	0.59(0.12)	4.30(1.97)
	Error5	SSQVCM	2.73(0.45)	0.43(0.63)	0.92(0.08)	0.92(0.07)	1.38(0.61)
		SSVCM	2.52(0.50)	2.64(1.66)	0.71(0.14)	0.73(0.13)	2.24(0.94)
		QVC	2.52(0.50)	1.30(4.59)	0.82(0.13)	0.82(0.12)	1.99(1.03)
		VC	2.56(0.50)	6.10(1.66)	0.57(0.16)	0.61(0.14)	3.87(1.82)
$\tau = 0.7$	Error1	SSQVCM	2.93(0.25)	0.53(0.82)	0.94(0.08)	0.94(0.08)	1.71(0.53)
		SSVCM	2.64(0.48)	2.90(1.43)	0.70(0.11)	0.72(0.11)	2.82(1.14)
		QVC	2.70(0.46)	0.84(4.82)	0.87(0.11)	0.88(0.10)	9.10(2.18)
		VC	2.74(0.44)	4.64(1.04)	0.67(0.20)	0.71(0.17)	3.51(1.65)
	Error2	SSQVCM	2.93(0.25)	0.90(0.96)	0.90(0.09)	0.91(0.08)	2.22(0.65)
		SSVCM	2.66(0.48)	4.10(2.82)	0.65(0.16)	0.68(0.14)	3.77(1.84)
		QVC	2.70(0.51)	1.16(5.55)	0.84(0.11)	0.85(0.10)	10.45(2.48)
		VC	2.64(0.60)	6.30(1.11)	0.58(0.19)	0.63(0.16)	5.56(2.84)
	Error3	SSQVCM	2.83(0.38)	0.43(0.68)	0.93(0.08)	0.93(0.08)	1.44(0.63)
		SSVCM	2.60(0.49)	2.52(1.85)	0.73(0.13)	0.75(0.12)	2.09(0.88)
		QVC	2.52(0.50)	1.46(5.58)	0.81(0.14)	0.82(0.13)	12.42(3.07)
		VC	2.60(0.49)	6.06(2.38)	0.59(0.18)	0.64(0.15)	3.76(1.79)
	Error4	SSQVCM	2.60(0.50)	0.47(0.73)	0.89(0.08)	0.90(0.07)	2.90(0.80)
		SSVCM	2.38(0.49)	2.12(1.29)	0.72(0.11)	0.73(0.11)	3.99(1.21)
		QVC	2.52(0.50)	1.14(4.22)	0.83(0.13)	0.84(0.12)	10.48(2.10)
		VC	2.62(0.49)	4.56(1.58)	0.65(0.19)	0.69(0.16)	4.73(1.83)
	Error5	SSQVCM	2.77(0.43)	0.30(0.60)	0.93(0.08)	0.94(0.08)	2.58(0.84)
		SSVCM	2.50(0.51)	2.92(1.88)	0.69(0.14)	0.71(0.13)	3.48(1.25)
		QVC	2.58(0.50)	0.82(4.11)	0.87(0.13)	0.87(0.12)	10.99(3.60)
		VC	2.78(0.42)	4.30(1.44)	0.68(0.19)	0.72(0.16)	4.18(1.71)

Table C.10: Evaluation of heterogeneous model under LD correlation, $(n, p) = (400, 1000)$ with 100 replicates.

$n = 400$	$p = 1000$		TP	FP	F1	MCC	TIMSE
$\tau = 0.3$	Error1	SSQVCM	3.00(0.00)	0.10(0.31)	0.99(0.03)	0.99(0.03)	0.83(0.26)
		SSVCM	3.00(0.00)	2.84(1.84)	0.76(0.12)	0.78(0.10)	1.43(0.80)
		QVC	3.00(0.00)	1.14(3.36)	0.90(0.13)	0.91(0.11)	8.89(2.64)
		VC	3.00(0.00)	4.70(1.55)	0.67(0.15)	0.71(0.13)	2.95(1.26)
	Error2	SSQVCM	2.83(0.38)	0.27(0.52)	0.95(0.07)	0.95(0.06)	1.38(0.38)
		SSVCM	2.86(0.35)	3.86(4.32)	0.71(0.17)	0.74(0.15)	3.23(5.46)
		QVC	2.84(0.51)	0.40(4.47)	0.93(0.11)	0.94(0.10)	9.00(2.50)
		VC	2.86(0.50)	5.78(0.70)	0.62(0.19)	0.67(0.16)	5.02(3.18)
	Error3	SSQVCM	2.93(0.25)	0.33(0.61)	0.95(0.07)	0.96(0.06)	1.51(0.41)
		SSVCM	2.98(0.14)	3.04(1.67)	0.74(0.11)	0.77(0.10)	2.81(0.95)
		QVC	2.98(0.14)	1.04(4.57)	0.90(0.12)	0.91(0.10)	8.24(2.31)
		VC	3.00(0.00)	5.82(1.50)	0.63(0.16)	0.68(0.13)	5.29(2.30)
	Error4	SSQVCM	2.87(0.35)	0.30(0.60)	0.95(0.07)	0.95(0.07)	1.52(0.42)
		SSVCM	2.90(0.30)	2.82(1.56)	0.74(0.12)	0.77(0.11)	1.94(0.78)
		QVC	2.96(0.20)	0.76(5.06)	0.92(0.11)	0.93(0.10)	10.13(2.81)
		VC	2.98(0.14)	6.68(1.17)	0.60(0.18)	0.65(0.14)	4.27(1.86)
	Error5	SSQVCM	2.97(0.18)	0.33(0.55)	0.96(0.06)	0.96(0.06)	1.40(0.40)
		SSVCM	2.98(0.14)	3.32(1.98)	0.73(0.13)	0.76(0.11)	2.17(0.91)
		QVC	2.98(0.14)	0.98(5.28)	0.90(0.11)	0.91(0.09)	9.51(2.16)
		VC	3.00(0.00)	7.04(1.17)	0.60(0.20)	0.65(0.16)	4.51(2.06)
$\tau = 0.5$	Error1	SSQVCM	3.00(0.00)	0.13(0.35)	0.99(0.04)	0.99(0.04)	0.62(0.17)
		SSVCM	2.94(0.24)	2.80(1.70)	0.75(0.13)	0.78(0.11)	1.15(0.56)
		QVC	2.98(0.14)	0.82(2.67)	0.91(0.09)	0.92(0.08)	0.92(0.38)
		VC	2.98(0.14)	4.08(0.92)	0.69(0.14)	0.73(0.12)	2.23(1.00)
	Error2	SSQVCM	3.00(0.00)	0.63(0.81)	0.93(0.08)	0.94(0.07)	0.94(0.28)
		SSVCM	2.94(0.24)	3.72(2.19)	0.70(0.13)	0.73(0.11)	2.01(1.08)
		QVC	2.94(0.24)	0.96(3.55)	0.90(0.11)	0.91(0.10)	1.23(0.49)
		VC	2.94(0.24)	5.42(1.18)	0.63(0.16)	0.68(0.13)	3.65(1.91)
	Error3	SSQVCM	3.00(0.00)	0.10(0.31)	0.99(0.03)	0.99(0.03)	0.84(0.28)
		SSVCM	2.94(0.24)	2.82(1.96)	0.75(0.13)	0.78(0.12)	1.84(1.04)
		QVC	2.98(0.14)	0.82(4.36)	0.91(0.10)	0.92(0.09)	1.31(0.52)
		VC	2.96(0.20)	5.74(0.98)	0.63(0.18)	0.68(0.14)	3.91(1.95)
	Error4	SSQVCM	3.00(0.00)	0.20(0.41)	0.98(0.05)	0.98(0.04)	0.73(0.21)
		SSVCM	2.90(0.30)	2.86(1.76)	0.74(0.12)	0.77(0.11)	1.34(0.60)
		QVC	2.96(0.20)	1.20(3.58)	0.88(0.10)	0.89(0.09)	1.15(0.45)
		VC	2.98(0.14)	4.80(1.23)	0.66(0.15)	0.71(0.12)	2.82(1.41)
	Error5	SSQVCM	3.00(0.00)	0.17(0.38)	0.98(0.04)	0.98(0.04)	0.77(0.25)
		SSVCM	2.96(0.20)	3.18(1.66)	0.73(0.11)	0.75(0.10)	1.47(0.84)
		QVC	2.98(0.14)	1.30(4.42)	0.88(0.13)	0.89(0.11)	1.33(0.68)
		VC	2.98(0.14)	5.52(1.49)	0.64(0.18)	0.69(0.14)	3.05(1.65)
$\tau = 0.7$	Error1	SSQVCM	3.00(0.00)	0.07(0.25)	0.99(0.03)	0.99(0.03)	1.23(0.27)
		SSVCM	3.00(0.00)	2.64(1.55)	0.77(0.11)	0.79(0.10)	1.94(0.79)
		QVC	3.00(0.00)	0.68(3.86)	0.94(0.10)	0.94(0.09)	8.85(2.34)
		VC	3.00(0.00)	4.22(1.20)	0.71(0.18)	0.74(0.15)	2.86(1.40)
	Error2	SSQVCM	3.00(0.00)	0.37(0.72)	0.96(0.07)	0.96(0.07)	1.59(0.35)
		SSVCM	2.90(0.30)	4.92(9.48)	0.71(0.17)	0.74(0.15)	5.01(11.08)
		QVC	2.80(0.64)	1.00(4.03)	0.88(0.16)	0.89(0.14)	9.89(2.39)
		VC	2.82(0.63)	3.96(1.63)	0.70(0.19)	0.73(0.15)	4.32(2.46)
	Error3	SSQVCM	2.92(0.28)	0.46(0.78)	0.94(0.08)	0.94(0.08)	1.17(0.42)
		SSVCM	2.94(0.24)	3.33(1.85)	0.72(0.13)	0.75(0.11)	1.57(0.75)
		QVC	2.94(0.24)	0.89(5.10)	0.92(0.14)	0.93(0.12)	11.43(1.94)
		VC	3.00(0.00)	7.11(1.91)	0.58(0.18)	0.64(0.15)	3.49(1.78)
	Error4	SSQVCM	3.00(0.00)	0.43(0.73)	0.95(0.07)	0.96(0.07)	1.72(0.36)
		SSVCM	2.96(0.20)	2.72(1.97)	0.76(0.13)	0.79(0.12)	2.62(1.00)
		QVC	2.98(0.14)	0.88(7.50)	0.91(0.11)	0.92(0.10)	9.99(2.55)
		VC	3.00(0.00)	5.56(1.27)	0.69(0.21)	0.73(0.18)	3.89(2.37)
	Error5	SSQVCM	2.97(0.19)	0.28(0.53)	0.97(0.06)	0.97(0.06)	1.91(0.42)
		SSVCM	2.96(0.20)	2.69(1.64)	0.76(0.12)	0.78(0.10)	2.83(0.93)
		QVC	2.96(0.20)	1.08(4.21)	0.89(0.12)	0.90(0.10)	9.78(2.54)
		VC	2.98(0.14)	4.90(1.38)	0.67(0.19)	0.72(0.16)	3.93(1.63)

Table C.11: Evaluation of heterogeneous model under real data correlation, $(n, p) = (400, 1000)$ with 100 replicates.

$n = 400$	$p = 1000$		TP	FP	F1	MCC	TIMSE
$\tau = 0.3$	Error1	SSQVCM	3.00(0.00)	0.37(0.61)	0.96(0.06)	0.96(0.06)	1.08(0.30)
		SSVCM	3.00(0.00)	3.10(1.96)	0.74(0.13)	0.77(0.11)	1.54(0.63)
		QVC	3.00(0.00)	1.14(5.71)	0.89(0.11)	0.90(0.10)	9.09(2.10)
		VC	3.00(0.00)	5.42(1.54)	0.66(0.18)	0.70(0.15)	2.98(1.64)
	Error2	SSQVCM	2.97(0.18)	0.30(0.60)	0.96(0.06)	0.97(0.06)	1.21(0.34)
		SSVCM	2.80(0.40)	3.64(3.43)	0.71(0.19)	0.74(0.17)	3.67(6.91)
		QVC	2.72(0.64)	1.08(5.36)	0.86(0.16)	0.87(0.15)	10.24(2.65)
		VC	2.76(0.66)	6.60(1.96)	0.57(0.17)	0.63(0.14)	5.64(5.63)
	Error3	SSQVCM	2.97(0.18)	0.57(0.82)	0.94(0.08)	0.94(0.08)	1.91(0.30)
		SSVCM	2.96(0.20)	2.82(1.57)	0.75(0.10)	0.77(0.09)	3.07(0.84)
		QVC	2.92(0.27)	0.76(5.65)	0.91(0.10)	0.92(0.10)	8.15(1.74)
		VC	2.96(0.20)	7.30(1.04)	0.58(0.19)	0.64(0.15)	5.51(1.75)
	Error4	SSQVCM	2.97(0.18)	0.20(0.41)	0.97(0.06)	0.97(0.06)	1.38(0.41)
		SSVCM	2.98(0.14)	2.72(1.93)	0.77(0.13)	0.79(0.11)	2.02(0.76)
		QVC	3.00(0.00)	1.08(4.54)	0.90(0.12)	0.91(0.10)	9.48(2.35)
		VC	3.00(0.00)	5.56(1.38)	0.65(0.18)	0.69(0.15)	3.84(1.89)
	Error5	SSQVCM	3.00(0.00)	0.27(0.52)	0.97(0.06)	0.97(0.05)	1.31(0.30)
		SSVCM	2.98(0.14)	2.96(1.93)	0.75(0.13)	0.78(0.11)	1.88(0.81)
		QVC	2.96(0.20)	0.98(5.33)	0.90(0.12)	0.91(0.11)	9.73(2.12)
		VC	2.96(0.20)	5.60(1.49)	0.66(0.21)	0.70(0.17)	3.62(1.69)
$\tau = 0.5$	Error1	SSQVCM	3.00(0.00)	0.07(0.25)	0.99(0.03)	0.99(0.03)	0.55(0.15)
		SSVCM	2.98(0.14)	2.76(1.90)	0.76(0.13)	0.79(0.11)	1.17(0.67)
		QVC	3.00(0.00)	0.82(5.14)	0.92(0.11)	0.93(0.10)	0.99(0.56)
		VC	3.00(0.00)	5.70(1.26)	0.66(0.22)	0.71(0.18)	2.60(1.65)
	Error2	SSQVCM	2.96(0.20)	0.28(0.54)	0.96(0.06)	0.97(0.06)	1.18(0.31)
		SSVCM	2.78(0.42)	3.44(3.55)	0.72(0.19)	0.75(0.17)	3.90(7.25)
		QVC	2.69(0.67)	0.80(5.55)	0.87(0.15)	0.88(0.14)	10.25(2.64)
		VC	2.73(0.69)	6.71(1.16)	0.57(0.17)	0.62(0.14)	5.83(5.90)
	Error3	SSQVCM	3.00(0.00)	0.33(0.61)	0.96(0.06)	0.97(0.06)	0.70(0.26)
		SSVCM	2.92(0.27)	3.22(2.61)	0.73(0.15)	0.76(0.13)	1.90(1.67)
		QVC	2.96(0.20)	1.34(5.29)	0.88(0.14)	0.89(0.13)	1.15(0.78)
		VC	2.88(0.39)	7.32(1.97)	0.57(0.18)	0.62(0.15)	4.08(2.04)
	Error4	SSQVCM	3.00(0.00)	0.30(0.47)	0.97(0.05)	0.97(0.05)	0.73(0.22)
		SSVCM	2.98(0.14)	2.70(1.93)	0.77(0.13)	0.79(0.12)	1.35(0.54)
		QVC	2.94(0.24)	1.14(3.76)	0.89(0.12)	0.89(0.11)	1.24(0.63)
		VC	2.94(0.24)	5.10(1.50)	0.65(0.18)	0.70(0.15)	2.80(1.41)
	Error5	SSQVCM	3.00(0.00)	0.13(0.35)	0.99(0.04)	0.99(0.04)	0.66(0.20)
		SSVCM	2.96(0.20)	2.86(1.97)	0.75(0.13)	0.78(0.11)	1.24(0.73)
		QVC	2.92(0.27)	0.98(5.00)	0.90(0.12)	0.90(0.11)	1.19(0.66)
		VC	2.96(0.20)	6.94(1.35)	0.59(0.20)	0.65(0.16)	3.17(1.62)
$\tau = 0.7$	Error1	SSQVCM	3.00(0.00)	0.20(0.48)	0.98(0.05)	0.98(0.05)	1.17(0.23)
		SSVCM	3.00(0.00)	2.38(2.12)	0.80(0.14)	0.82(0.12)	1.85(0.62)
		QVC	3.00(0.00)	1.22(5.08)	0.90(0.15)	0.91(0.13)	9.46(2.72)
		VC	3.00(0.00)	5.84(2.01)	0.64(0.19)	0.69(0.16)	3.24(1.64)
	Error2	SSQVCM	2.97(0.18)	0.33(0.66)	0.96(0.07)	0.96(0.06)	1.40(0.40)
		SSVCM	2.86(0.40)	3.82(3.14)	0.70(0.17)	0.73(0.15)	3.14(2.81)
		QVC	2.82(0.48)	0.92(5.44)	0.90(0.15)	0.91(0.14)	9.89(2.38)
		VC	2.84(0.47)	5.16(2.06)	0.67(0.21)	0.71(0.18)	4.33(2.91)
	Error3	SSQVCM	3.00(0.00)	0.60(0.62)	0.93(0.07)	0.94(0.06)	0.96(0.30)
		SSVCM	2.94(0.24)	3.94(2.26)	0.68(0.12)	0.72(0.10)	1.69(0.90)
		QVC	2.90(0.30)	1.22(6.53)	0.87(0.13)	0.88(0.12)	11.53(2.45)
		VC	2.94(0.24)	6.50(1.53)	0.62(0.20)	0.67(0.16)	3.34(2.33)
	Error4	SSQVCM	2.60(0.50)	0.47(0.73)	0.89(0.08)	0.90(0.07)	2.90(0.80)
		SSVCM	2.38(0.49)	2.12(1.29)	0.72(0.11)	0.73(0.11)	3.99(1.21)
		QVC	2.52(0.50)	1.14(4.22)	0.83(0.13)	0.84(0.12)	10.48(2.10)
		VC	2.62(0.49)	4.56(1.58)	0.65(0.19)	0.69(0.16)	4.73(1.83)
	Error5	SSQVCM	2.77(0.43)	0.30(0.60)	0.93(0.08)	0.94(0.08)	2.58(0.84)
		SSVCM	2.50(0.51)	2.92(1.88)	0.69(0.14)	0.71(0.13)	3.48(1.25)
		QVC	2.58(0.50)	0.82(4.11)	0.87(0.13)	0.87(0.12)	10.99(3.60)
		VC	2.78(0.42)	4.30(1.44)	0.68(0.19)	0.72(0.16)	4.18(1.71)

Table C.12: Evaluation of homogeneous model under AR-1 correlation, with the diagonal elements of the random correlation structure being 20 and 2 (100 replicates).

$n = 400$	$p = 1000$		TP	FP	F1	MCC	TIMSE
$\tau = 0.3$	Error1	SSQVCM	2.93(0.25)	0.13(0.35)	0.98(0.05)	0.98(0.05)	0.95(0.43)
		SSVCM	2.70(0.46)	2.92(1.94)	0.71(0.12)	0.74(0.11)	1.50(0.88)
		QVC	2.76(0.43)	1.00(3.52)	0.88(0.12)	0.89(0.10)	9.08(2.57)
		VC	2.78(0.42)	4.62(1.58)	0.65(0.16)	0.69(0.14)	2.19(1.03)
	Error2	SSQVCM	2.90(0.31)	0.10(0.31)	0.97(0.05)	0.98(0.05)	1.21(0.46)
		SSVCM	2.58(0.61)	2.72(1.85)	0.71(0.15)	0.73(0.15)	1.96(1.79)
		QVC	2.64(0.53)	0.62(3.81)	0.89(0.12)	0.90(0.12)	9.55(2.16)
		VC	2.62(0.60)	3.96(1.10)	0.67(0.17)	0.70(0.15)	2.70(1.72)
	Error3	SSQVCM	2.97(0.18)	0.20(0.48)	0.97(0.06)	0.98(0.05)	1.61(0.51)
		SSVCM	2.70(0.46)	3.20(5.08)	0.73(0.16)	0.75(0.14)	2.62(2.91)
		QVC	2.70(0.46)	0.80(7.46)	0.88(0.12)	0.89(0.11)	6.97(1.64)
		VC	2.74(0.44)	6.28(1.43)	0.62(0.20)	0.66(0.16)	3.64(2.88)
	Error4	SSQVCM	3.00(0.00)	0.13(0.35)	0.99(0.04)	0.99(0.04)	1.19(0.40)
		SSVCM	2.74(0.44)	2.68(1.43)	0.73(0.11)	0.75(0.11)	1.65(0.77)
		QVC	2.84(0.37)	1.10(3.61)	0.88(0.13)	0.89(0.11)	8.28(1.84)
		VC	2.90(0.30)	4.56(1.73)	0.67(0.16)	0.71(0.14)	2.26(0.79)
	Error5	SSQVCM	3.00(0.00)	0.10(0.40)	0.99(0.04)	0.99(0.04)	1.18(0.31)
		SSVCM	2.78(0.42)	2.84(1.68)	0.73(0.13)	0.75(0.12)	1.70(0.88)
		QVC	2.82(0.39)	1.22(4.33)	0.87(0.14)	0.88(0.12)	8.73(2.09)
		VC	2.84(0.37)	5.18(1.94)	0.65(0.19)	0.69(0.16)	2.54(1.01)
$\tau = 0.5$	Error1	SSQVCM	3.00(0.00)	0.10(0.31)	0.99(0.03)	0.99(0.03)	0.74(0.22)
		SSVCM	2.84(0.37)	2.48(1.52)	0.76(0.11)	0.78(0.10)	1.11(0.61)
		QVC	2.84(0.37)	0.80(3.45)	0.90(0.12)	0.91(0.11)	1.31(0.71)
		VC	2.86(0.35)	4.42(1.36)	0.67(0.16)	0.71(0.13)	1.73(0.85)
	Error2	SSQVCM	2.93(0.25)	0.23(0.43)	0.96(0.06)	0.97(0.05)	1.00(0.40)
		SSVCM	2.66(0.48)	3.36(5.58)	0.72(0.15)	0.74(0.14)	2.08(3.33)
		QVC	2.72(0.45)	1.64(4.72)	0.83(0.15)	0.84(0.14)	1.71(0.88)
		VC	2.72(0.57)	5.24(2.64)	0.63(0.18)	0.67(0.15)	2.74(1.96)
	Error3	SSQVCM	3.00(0.00)	0.13(0.35)	0.99(0.04)	0.99(0.04)	0.90(0.18)
		SSVCM	2.72(0.45)	2.62(1.37)	0.73(0.13)	0.75(0.12)	1.55(0.93)
		QVC	2.90(0.30)	0.90(3.66)	0.90(0.11)	0.91(0.10)	1.43(0.60)
		VC	2.94(0.24)	5.40(1.31)	0.63(0.16)	0.68(0.13)	2.27(0.95)
	Error4	SSQVCM	3.00(0.00)	0.10(0.31)	0.99(0.03)	0.99(0.03)	0.86(0.28)
		SSVCM	2.82(0.39)	2.60(1.59)	0.75(0.12)	0.77(0.11)	1.30(0.67)
		QVC	2.86(0.35)	1.50(2.84)	0.85(0.14)	0.86(0.12)	1.60(0.77)
		VC	2.88(0.33)	3.90(1.95)	0.69(0.13)	0.72(0.11)	1.86(0.85)
	Error5	SSQVCM	3.00(0.00)	0.10(0.31)	0.99(0.03)	0.99(0.03)	0.74(0.18)
		SSVCM	2.76(0.43)	2.78(1.57)	0.73(0.12)	0.75(0.11)	1.18(0.81)
		QVC	2.82(0.39)	1.12(3.69)	0.87(0.12)	0.88(0.11)	1.49(0.69)
		VC	2.86(0.35)	4.92(1.36)	0.65(0.18)	0.69(0.15)	1.87(0.89)
$\tau = 0.7$	Error1	SSQVCM	3.00(0.00)	0.13(0.35)	0.99(0.04)	0.99(0.04)	0.93(0.27)
		SSVCM	2.80(0.40)	2.40(1.86)	0.76(0.14)	0.78(0.13)	1.30(0.64)
		QVC	2.86(0.35)	0.74(2.93)	0.91(0.11)	0.92(0.10)	8.58(1.98)
		VC	2.86(0.35)	3.52(1.21)	0.72(0.17)	0.75(0.15)	1.82(0.92)
	Error2	SSQVCM	2.87(0.35)	0.33(0.80)	0.95(0.08)	0.95(0.08)	1.31(0.44)
		SSVCM	2.62(0.49)	3.56(2.58)	0.67(0.14)	0.70(0.12)	2.03(1.65)
		QVC	2.80(0.40)	1.02(4.05)	0.88(0.12)	0.89(0.11)	9.29(1.96)
		VC	2.82(0.39)	5.44(1.49)	0.62(0.15)	0.66(0.13)	2.68(1.28)
	Error3	SSQVCM	3.00(0.00)	0.20(0.48)	0.98(0.05)	0.98(0.05)	0.95(0.35)
		SSVCM	2.76(0.43)	3.12(2.03)	0.71(0.12)	0.73(0.10)	1.39(0.73)
		QVC	2.78(0.42)	0.96(4.76)	0.88(0.12)	0.89(0.11)	12.27(2.66)
		VC	2.78(0.42)	5.06(1.50)	0.64(0.17)	0.68(0.14)	2.12(1.14)
	Error4	SSQVCM	2.83(0.38)	0.20(0.41)	0.95(0.06)	0.96(0.06)	1.35(0.46)
		SSVCM	2.66(0.48)	2.46(1.75)	0.74(0.12)	0.76(0.11)	1.87(0.84)
		QVC	2.74(0.44)	1.14(4.64)	0.86(0.11)	0.87(0.10)	9.58(2.15)
		VC	2.82(0.39)	5.36(1.50)	0.64(0.18)	0.68(0.15)	2.54(1.18)
	Error5	SSQVCM	2.97(0.18)	0.33(0.66)	0.96(0.07)	0.96(0.07)	1.33(0.39)
		SSVCM	2.72(0.45)	2.60(1.64)	0.74(0.13)	0.76(0.12)	1.79(0.85)
		QVC	2.84(0.37)	1.42(4.65)	0.86(0.15)	0.87(0.13)	8.79(1.94)
		VC	2.88(0.33)	6.10(2.24)	0.61(0.17)	0.66(0.13)	2.66(1.05)

Table C.13: Evaluation of homogeneous model under AR-1 correlation, with the diagonal elements of the random correlation structure being 40 and 0.4 (100 replicates).

$n = 400$	$p = 1000$		TP	FP	F1	MCC	TIMSE
$\tau = 0.3$	Error1	SSQVCM	3.00(0.00)	0.17(0.38)	0.98(0.04)	0.98(0.04)	1.06(0.38)
		SSVCM	2.70(0.46)	3.08(1.61)	0.70(0.11)	0.72(0.10)	1.49(0.82)
		QVC	2.74(0.44)	1.42(3.77)	0.84(0.15)	0.86(0.13)	11.07(2.09)
		VC	2.80(0.40)	4.82(1.98)	0.64(0.16)	0.69(0.13)	2.18(0.98)
	Error2	SSQVCM	3.83(0.38)	0.33(0.55)	0.94(0.07)	0.94(0.06)	1.51(0.39)
		SSVCM	3.58(0.50)	2.70(1.67)	0.71(0.12)	0.73(0.11)	2.08(0.99)
		QVC	3.62(0.49)	1.06(3.37)	0.85(0.13)	0.85(0.13)	11.15(2.74)
		VC	3.70(0.46)	4.56(1.25)	0.64(0.17)	0.68(0.15)	2.97(1.28)
	Error3	SSQVCM	2.87(0.35)	0.20(0.41)	0.96(0.06)	0.96(0.06)	1.65(0.47)
		SSVCM	2.70(0.46)	2.80(1.44)	0.72(0.12)	0.74(0.11)	2.37(1.07)
		QVC	2.70(0.46)	1.40(3.17)	0.84(0.13)	0.85(0.12)	8.63(3.02)
		VC	2.82(0.39)	4.26(2.07)	0.67(0.15)	0.70(0.12)	3.22(1.06)
	Error4	SSQVCM	2.77(0.43)	0.17(0.46)	0.95(0.08)	0.95(0.07)	1.63(0.46)
		SSVCM	2.60(0.49)	2.42(1.68)	0.73(0.12)	0.75(0.11)	2.03(0.93)
		QVC	2.64(0.48)	1.28(3.99)	0.84(0.12)	0.85(0.10)	10.56(2.25)
		VC	2.76(0.43)	3.90(1.75)	0.70(0.17)	0.73(0.14)	2.51(1.09)
	Error5	SSQVCM	3.00(0.00)	0.13(0.35)	0.99(0.04)	0.99(0.04)	1.14(0.31)
		SSVCM	2.78(0.42)	2.66(1.57)	0.74(0.14)	0.76(0.13)	1.58(0.86)
		QVC	2.84(0.37)	1.44(3.84)	0.86(0.15)	0.87(0.14)	11.15(2.70)
		VC	2.92(0.27)	5.00(2.16)	0.65(0.16)	0.69(0.13)	2.51(1.26)
$\tau = 0.5$	Error1	SSQVCM	2.87(0.35)	0.20(0.41)	0.96(0.06)	0.96(0.06)	1.05(0.38)
		SSVCM	2.58(0.50)	2.42(1.74)	0.73(0.14)	0.75(0.13)	1.41(0.78)
		QVC	2.68(0.47)	1.48(3.48)	0.83(0.13)	0.84(0.11)	1.97(0.89)
		VC	2.68(0.47)	3.58(2.48)	0.69(0.15)	0.72(0.12)	1.92(0.86)
	Error2	SSQVCM	2.90(0.31)	0.27(0.58)	0.96(0.07)	0.96(0.06)	1.24(0.45)
		SSVCM	2.56(0.61)	4.28(7.35)	0.67(0.17)	0.70(0.16)	19.01(122.57)
		QVC	2.64(0.60)	1.44(3.92)	0.81(0.13)	0.82(0.12)	2.27(1.60)
		VC	2.68(0.59)	5.40(1.40)	0.60(0.16)	0.65(0.14)	3.27(3.55)
	Error3	SSQVCM	2.83(0.38)	0.10(0.31)	0.97(0.06)	0.97(0.06)	1.39(0.49)
		SSVCM	2.70(0.46)	2.70(1.71)	0.73(0.12)	0.75(0.11)	1.86(0.87)
		QVC	2.78(0.42)	1.24(3.97)	0.85(0.12)	0.86(0.12)	2.24(1.03)
		VC	2.76(0.43)	5.04(1.44)	0.63(0.16)	0.67(0.13)	2.80(1.13)
	Error4	SSQVCM	2.90(0.31)	0.30(0.47)	0.95(0.06)	0.95(0.06)	1.23(0.44)
		SSVCM	2.58(0.50)	2.68(1.38)	0.71(0.11)	0.72(0.10)	1.62(0.81)
		QVC	2.72(0.45)	1.50(3.69)	0.83(0.14)	0.84(0.13)	2.02(1.05)
		VC	2.74(0.44)	4.60(2.20)	0.65(0.16)	0.69(0.14)	2.22(1.00)
	Error5	SSQVCM	3.00(0.00)	0.20(0.48)	0.98(0.05)	0.98(0.05)	1.00(0.27)
		SSVCM	2.78(0.42)	2.34(1.48)	0.76(0.12)	0.78(0.11)	1.32(0.70)
		QVC	2.76(0.43)	1.02(4.22)	0.87(0.13)	0.88(0.12)	1.79(0.78)
		VC	2.78(0.42)	4.08(1.49)	0.69(0.18)	0.73(0.16)	1.97(0.91)
$\tau = 0.7$	Error1	SSQVCM	3.00(0.00)	0.13(0.35)	0.99(0.04)	0.99(0.04)	1.04(0.32)
		SSVCM	2.82(0.39)	2.62(1.64)	0.75(0.13)	0.77(0.12)	1.29(0.59)
		QVC	2.86(0.35)	1.70(4.43)	0.85(0.16)	0.86(0.14)	10.93(2.57)
		VC	2.92(0.27)	5.04(2.53)	0.67(0.19)	0.71(0.16)	2.11(0.96)
	Error2	SSQVCM	2.83(0.38)	0.47(1.38)	0.94(0.11)	0.94(0.09)	1.62(0.58)
		SSVCM	2.60(0.61)	5.56(8.14)	0.63(0.17)	0.66(0.16)	4.36(9.94)
		QVC	2.62(0.64)	1.54(5.85)	0.80(0.14)	0.81(0.13)	10.62(2.20)
		VC	2.66(0.69)	5.98(1.53)	0.60(0.18)	0.65(0.15)	3.90(2.57)
	Error3	SSQVCM	2.97(0.18)	0.13(0.35)	0.98(0.04)	0.98(0.04)	1.16(0.33)
		SSVCM	2.64(0.48)	2.38(1.46)	0.74(0.14)	0.76(0.13)	1.53(0.78)
		QVC	2.62(0.49)	1.28(3.83)	0.83(0.13)	0.85(0.12)	15.40(3.28)
		VC	2.74(0.44)	4.36(1.77)	0.67(0.18)	0.70(0.16)	2.18(0.97)
	Error4	SSQVCM	2.87(0.35)	0.20(0.41)	0.96(0.06)	0.96(0.06)	1.53(0.45)
		SSVCM	2.62(0.49)	2.64(1.86)	0.73(0.14)	0.75(0.13)	1.95(0.81)
		QVC	2.64(0.48)	0.92(3.12)	0.87(0.14)	0.88(0.13)	10.70(2.48)
		VC	2.72(0.45)	3.56(1.98)	0.70(0.18)	0.73(0.15)	2.53(1.04)
	Error5	SSQVCM	2.83(0.38)	0.17(0.46)	0.96(0.07)	0.96(0.06)	1.50(0.55)
		SSVCM	2.64(0.48)	2.98(2.06)	0.70(0.12)	0.72(0.11)	1.88(0.92)
		QVC	2.74(0.44)	2.08(3.92)	0.80(0.16)	0.82(0.14)	10.64(2.78)
		VC	2.74(0.44)	4.76(2.61)	0.65(0.18)	0.69(0.15)	2.76(1.05)

Table C.14: Evaluation of homogeneous model under AR-1 correlation, with unstructured random correlation structure (matrix(c(20,0.5,0.5,0.2),2,2) over 100 replicates).

$n = 400$	$p = 1000$		TP	FP	F1	MCC	TIMSE
$\tau = 0.3$	Error1	SSQVCM	2.75(0.44)	0.10(0.38)	0.95(0.07)	0.96(0.06)	1.40(0.62)
		SSVCM	2.64(0.48)	1.82(1.49)	0.78(0.12)	0.80(0.11)	1.51(0.76)
		QVC	2.72(0.45)	1.14(3.77)	0.86(0.14)	0.87(0.12)	9.55(2.27)
		VC	2.76(0.43)	4.16(1.76)	0.68(0.19)	0.72(0.16)	2.13(1.01)
	Error2	SSQVCM	2.49(0.51)	0.49(0.82)	0.88(0.09)	0.88(0.09)	1.86(0.52)
		SSVCM	2.38(0.60)	3.92(3.92)	0.64(0.17)	0.67(0.16)	3.01(2.70)
		QVC	2.50(0.54)	1.00(5.20)	0.84(0.12)	0.84(0.11)	10.98(2.18)
		VC	2.50(0.65)	6.52(1.40)	0.55(0.18)	0.60(0.15)	4.75(2.85)
	Error3	SSQVCM	2.80(0.41)	0.29(0.62)	0.94(0.07)	0.94(0.07)	1.64(0.53)
		SSVCM	2.70(0.46)	2.66(1.65)	0.73(0.14)	0.75(0.13)	2.18(0.98)
		QVC	2.74(0.44)	0.92(3.25)	0.88(0.13)	0.89(0.12)	8.21(2.31)
		VC	2.78(0.42)	4.50(1.64)	0.65(0.16)	0.69(0.14)	3.08(1.22)
	Error4	SSQVCM	2.71(0.46)	0.17(0.45)	0.94(0.08)	0.94(0.07)	1.53(0.50)
		SSVCM	2.58(0.50)	2.58(1.50)	0.72(0.13)	0.73(0.13)	1.87(0.91)
		QVC	2.78(0.42)	1.84(3.89)	0.83(0.16)	0.84(0.14)	10.56(2.55)
		VC	2.76(0.43)	4.50(2.74)	0.66(0.16)	0.69(0.14)	2.65(1.28)
	Error5	SSQVCM	2.86(0.36)	0.23(0.49)	0.95(0.07)	0.96(0.06)	1.32(0.57)
		SSVCM	2.66(0.48)	3.14(1.80)	0.69(0.12)	0.72(0.11)	1.79(0.92)
		QVC	2.74(0.44)	1.18(4.07)	0.86(0.15)	0.88(0.13)	9.14(1.68)
		VC	2.74(0.44)	4.90(2.03)	0.64(0.18)	0.68(0.15)	2.74(1.21)
$\tau = 0.5$	Error1	SSQVCM	2.86(0.36)	0.26(0.51)	0.95(0.07)	0.95(0.06)	1.05(0.47)
		SSVCM	2.66(0.48)	2.66(1.77)	0.73(0.14)	0.75(0.13)	1.37(0.74)
		QVC	2.62(0.49)	1.10(3.37)	0.85(0.13)	0.85(0.12)	1.68(0.78)
		VC	2.70(0.46)	4.44(1.61)	0.65(0.15)	0.68(0.12)	2.06(0.75)
	Error2	SSQVCM	2.77(0.43)	0.46(0.82)	0.92(0.08)	0.93(0.08)	1.35(0.61)
		SSVCM	2.60(0.49)	3.50(2.40)	0.67(0.13)	0.70(0.12)	2.03(1.67)
		QVC	2.74(0.49)	1.80(4.58)	0.82(0.15)	0.84(0.13)	1.84(0.98)
		VC	2.78(0.46)	6.28(2.81)	0.59(0.19)	0.64(0.16)	3.18(2.08)
	Error3	SSQVCM	2.94(0.24)	0.11(0.32)	0.98(0.05)	0.98(0.04)	1.15(0.47)
		SSVCM	2.72(0.45)	2.90(1.53)	0.71(0.12)	0.73(0.11)	1.59(0.83)
		QVC	2.82(0.39)	1.22(5.31)	0.86(0.13)	0.87(0.11)	1.81(0.72)
		VC	2.82(0.39)	5.74(1.43)	0.63(0.20)	0.68(0.17)	2.76(1.47)
	Error4	SSQVCM	2.86(0.36)	0.23(0.43)	0.95(0.06)	0.96(0.06)	1.20(0.50)
		SSVCM	2.72(0.45)	2.78(1.87)	0.73(0.14)	0.75(0.12)	1.46(0.74)
		QVC	2.74(0.44)	1.34(5.66)	0.84(0.10)	0.85(0.09)	1.88(0.67)
		VC	2.84(0.37)	5.62(1.39)	0.64(0.20)	0.69(0.16)	2.42(1.25)
	Error5	SSQVCM	2.86(0.36)	0.09(0.28)	0.97(0.06)	0.97(0.05)	0.93(0.45)
		SSVCM	2.66(0.48)	2.08(1.65)	0.77(0.13)	0.78(0.12)	1.30(0.77)
		QVC	2.74(0.44)	0.50(3.65)	0.91(0.09)	0.92(0.08)	1.34(0.73)
		VC	2.78(0.42)	3.92(0.86)	0.69(0.16)	0.72(0.13)	1.89(0.89)
$\tau = 0.7$	Error1	SSQVCM	2.97(0.17)	0.23(0.43)	0.97(0.05)	0.97(0.05)	0.98(0.43)
		SSVCM	2.74(0.44)	2.68(1.80)	0.74(0.14)	0.76(0.13)	1.40(0.80)
		QVC	2.74(0.44)	1.24(3.79)	0.86(0.14)	0.87(0.13)	9.77(2.51)
		VC	2.80(0.40)	4.28(2.12)	0.67(0.16)	0.71(0.13)	2.17(0.96)
	Error2	SSQVCM	2.80(0.41)	0.20(0.53)	0.95(0.07)	0.95(0.07)	1.42(0.49)
		SSVCM	2.62(0.49)	2.54(2.14)	0.74(0.16)	0.76(0.14)	1.82(0.79)
		QVC	2.64(0.48)	1.16(5.03)	0.85(0.14)	0.86(0.13)	10.26(2.50)
		VC	2.74(0.44)	5.42(1.87)	0.63(0.18)	0.67(0.15)	3.03(1.71)
	Error3	SSQVCM	2.97(0.17)	0.17(0.38)	0.98(0.05)	0.98(0.04)	0.87(0.36)
		SSVCM	2.76(0.43)	2.60(1.78)	0.74(0.13)	0.76(0.11)	1.28(0.75)
		QVC	2.76(0.43)	1.36(3.38)	0.84(0.12)	0.85(0.11)	12.81(2.17)
		VC	2.78(0.42)	5.10(1.45)	0.62(0.14)	0.66(0.12)	2.27(0.75)
	Error4	SSQVCM	2.86(0.36)	0.31(0.58)	0.95(0.07)	0.95(0.07)	1.52(0.48)
		SSVCM	2.72(0.45)	2.42(1.50)	0.75(0.13)	0.77(0.12)	1.81(0.82)
		QVC	2.74(0.44)	1.22(4.31)	0.86(0.14)	0.87(0.12)	9.96(2.19)
		VC	2.76(0.43)	4.98(1.99)	0.64(0.17)	0.68(0.14)	2.81(1.14)
	Error5	SSQVCM	2.86(0.36)	0.51(0.85)	0.93(0.09)	0.93(0.08)	1.66(0.48)
		SSVCM	2.70(0.46)	2.68(1.56)	0.72(0.12)	0.74(0.11)	1.89(0.79)
		QVC	2.78(0.42)	1.24(4.08)	0.86(0.14)	0.88(0.12)	9.23(1.77)
		VC	2.82(0.39)	5.04(1.90)	0.65(0.18)	0.69(0.15)	2.69(1.07)

Table C.15: Evaluation of homogeneous model under AR-1 correlation, with unstructured random correlation structure (matrix(c(20,1,1,0.2),2,2) over 100 replicates).

$n = 400$	$p = 1000$		TP	FP	F1	MCC	TIMSE
$\tau = 0.3$	Error1	SSQVCM	3.00(0.00)	0.07(0.25)	0.99(0.03)	0.99(0.03)	0.78(0.17)
		SSVCM	2.98(0.16)	2.38(1.53)	0.78(0.11)	0.80(0.10)	0.88(0.41)
		QVC	2.95(0.22)	0.38(2.65)	0.95(0.07)	0.96(0.07)	6.30(1.50)
		VC	3.00(0.00)	2.65(0.67)	0.79(0.16)	0.81(0.14)	1.05(0.45)
	Error2	SSQVCM	2.78(0.42)	0.15(0.36)	0.95(0.07)	0.95(0.07)	1.30(0.43)
		SSVCM	2.76(0.43)	3.24(3.28)	0.73(0.18)	0.75(0.16)	2.24(3.08)
		QVC	2.84(0.44)	0.76(4.78)	0.90(0.11)	0.91(0.11)	6.81(1.39)
		VC	2.76(0.49)	5.05(1.12)	0.64(0.17)	0.68(0.14)	2.91(1.68)
	Error3	SSQVCM	2.97(0.18)	0.23(0.50)	0.97(0.06)	0.97(0.05)	1.27(0.38)
		SSVCM	2.93(0.26)	2.93(1.72)	0.74(0.13)	0.77(0.11)	1.60(0.56)
		QVC	2.93(0.26)	0.68(2.48)	0.93(0.11)	0.93(0.09)	4.86(1.24)
		VC	3.00(0.00)	3.24(1.21)	0.74(0.15)	0.77(0.12)	1.93(0.46)
	Error4	SSQVCM	3.00(0.00)	0.10(0.30)	0.99(0.03)	0.99(0.03)	1.12(0.33)
		SSVCM	2.96(0.20)	2.80(1.23)	0.75(0.10)	0.77(0.09)	1.25(0.57)
		QVC	2.92(0.27)	0.92(2.73)	0.91(0.12)	0.91(0.11)	6.29(1.38)
		VC	2.96(0.20)	3.44(1.58)	0.73(0.15)	0.76(0.12)	1.64(0.58)
	Error5	SSQVCM	3.00(0.00)	0.07(0.27)	0.99(0.03)	0.99(0.03)	0.98(0.42)
		SSVCM	2.94(0.24)	2.66(1.60)	0.76(0.11)	0.78(0.09)	1.11(0.67)
		QVC	2.98(0.14)	0.88(2.82)	0.91(0.11)	0.92(0.10)	5.84(1.23)
		VC	2.96(0.20)	3.54(1.27)	0.72(0.15)	0.75(0.13)	1.50(0.64)
$\tau = 0.5$	Error1	SSQVCM	3.00(0.00)	0.09(0.29)	0.99(0.03)	0.99(0.03)	0.66(0.25)
		SSVCM	2.89(0.32)	3.00(1.64)	0.73(0.11)	0.75(0.10)	0.84(0.55)
		QVC	2.86(0.35)	0.34(2.26)	0.94(0.07)	0.95(0.06)	1.02(0.58)
		VC	2.93(0.25)	2.23(0.61)	0.80(0.15)	0.82(0.13)	0.93(0.52)
	Error2	SSQVCM	3.00(0.00)	0.07(0.25)	0.99(0.03)	0.99(0.03)	0.69(0.23)
		SSVCM	2.90(0.30)	5.25(9.60)	0.70(0.19)	0.73(0.17)	2.37(6.40)
		QVC	2.85(0.53)	1.12(3.72)	0.87(0.14)	0.88(0.12)	1.31(1.79)
		VC	2.83(0.59)	4.30(1.38)	0.67(0.18)	0.72(0.15)	1.97(2.40)
	Error3	SSQVCM	3.00(0.00)	0.10(0.31)	0.99(0.03)	0.99(0.03)	0.79(0.31)
		SSVCM	2.92(0.27)	3.00(1.70)	0.73(0.12)	0.76(0.10)	1.10(0.69)
		QVC	2.92(0.27)	0.90(2.28)	0.91(0.12)	0.92(0.11)	1.37(0.82)
		VC	2.97(0.16)	3.03(1.67)	0.75(0.13)	0.78(0.11)	1.42(0.71)
	Error4	SSQVCM	3.00(0.00)	0.12(0.34)	0.99(0.04)	0.99(0.04)	0.59(0.21)
		SSVCM	2.95(0.22)	2.57(1.67)	0.77(0.12)	0.79(0.10)	0.71(0.43)
		QVC	2.95(0.22)	1.10(2.98)	0.89(0.12)	0.90(0.11)	0.99(0.55)
		VC	2.95(0.22)	3.38(1.53)	0.74(0.17)	0.77(0.14)	1.03(0.55)
	Error5	SSQVCM	2.86(0.35)	0.06(0.32)	0.97(0.06)	0.97(0.05)	0.80(0.51)
		SSVCM	2.88(0.33)	3.06(1.60)	0.72(0.12)	0.75(0.11)	0.74(0.49)
		QVC	2.90(0.30)	0.80(2.67)	0.91(0.11)	0.92(0.10)	0.99(0.52)
		VC	2.94(0.24)	3.36(1.31)	0.73(0.15)	0.76(0.13)	1.03(0.54)
$\tau = 0.7$	Error1	SSQVCM	2.92(0.28)	0.08(0.28)	0.98(0.06)	0.98(0.06)	0.97(0.49)
		SSVCM	2.95(0.22)	2.71(2.04)	0.77(0.14)	0.79(0.12)	0.99(0.61)
		QVC	2.93(0.26)	0.88(2.14)	0.91(0.12)	0.91(0.11)	6.20(1.50)
		VC	2.98(0.16)	2.15(1.33)	0.82(0.15)	0.83(0.14)	1.05(0.51)
	Error2	SSQVCM	2.90(0.31)	0.21(0.62)	0.96(0.07)	0.97(0.06)	1.19(0.57)
		SSVCM	2.88(0.33)	4.79(9.83)	0.72(0.19)	0.75(0.16)	2.53(7.32)
		QVC	2.79(0.59)	0.71(3.42)	0.89(0.13)	0.90(0.11)	6.71(1.71)
		VC	2.76(0.61)	4.06(1.00)	0.68(0.20)	0.72(0.17)	2.22(2.15)
	Error3	SSQVCM	2.97(0.17)	0.09(0.29)	0.99(0.05)	0.99(0.05)	0.65(0.34)
		SSVCM	2.97(0.17)	2.26(1.48)	0.79(0.10)	0.81(0.09)	0.60(0.41)
		QVC	2.97(0.17)	0.74(2.78)	0.93(0.11)	0.93(0.10)	7.99(1.46)
		VC	2.97(0.17)	3.14(1.40)	0.75(0.14)	0.77(0.12)	0.88(0.40)
	Error4	SSQVCM	2.84(0.37)	0.08(0.28)	0.97(0.06)	0.97(0.06)	1.34(0.68)
		SSVCM	2.88(0.33)	2.80(1.68)	0.74(0.13)	0.77(0.12)	1.23(0.67)
		QVC	2.86(0.35)	0.80(2.60)	0.91(0.12)	0.92(0.11)	6.40(1.48)
		VC	2.96(0.20)	3.10(1.48)	0.75(0.15)	0.77(0.13)	1.54(0.69)
	Error5	SSQVCM	2.91(0.28)	0.06(0.24)	0.98(0.05)	0.98(0.04)	1.25(0.56)
		SSVCM	2.94(0.23)	2.86(1.59)	0.74(0.11)	0.77(0.10)	1.15(0.60)
		QVC	2.92(0.28)	0.81(2.96)	0.91(0.10)	0.92(0.09)	5.81(1.28)
		VC	2.97(0.17)	3.36(1.21)	0.74(0.16)	0.77(0.14)	1.51(0.52)

Table C.16: Evaluation of heterogeneous model under uneven repeated measurements with LD correlation, $(n, p) = (400, 1000)$ for 100 replicates.

$n = 400$	$p = 1000$		TP	FP	F1	MCC	TIMSE
$\tau = 0.3$	Error1	SSQVCM	2.37(0.49)	0.47(0.82)	0.86(0.07)	0.87(0.06)	2.03(0.55)
		SSVCM	2.22(0.42)	2.58(2.01)	0.68(0.13)	0.69(0.12)	3.13(1.11)
		QVC	2.26(0.49)	2.36(5.66)	0.72(0.17)	0.74(0.15)	7.07(1.91)
		VC	2.26(0.49)	6.50(3.06)	0.54(0.19)	0.58(0.16)	5.34(2.32)
	Error2	SSQVCM	2.37(0.49)	0.43(0.68)	0.86(0.09)	0.87(0.08)	2.29(0.56)
		SSVCM	2.26(0.49)	3.44(2.87)	0.64(0.16)	0.66(0.15)	3.98(2.65)
		QVC	2.32(0.55)	2.30(6.49)	0.71(0.15)	0.73(0.14)	7.23(2.22)
		VC	2.26(0.63)	8.04(2.01)	0.50(0.20)	0.55(0.17)	6.77(3.16)
	Error3	SSQVCM	2.33(0.48)	0.23(0.43)	0.88(0.08)	0.88(0.08)	2.35(0.62)
		SSVCM	2.26(0.44)	2.04(1.62)	0.72(0.13)	0.73(0.12)	3.49(1.31)
		QVC	2.42(0.50)	2.60(6.20)	0.72(0.15)	0.74(0.14)	5.65(1.70)
		VC	2.44(0.50)	6.82(2.81)	0.55(0.17)	0.60(0.14)	5.71(2.25)
	Error4	SSQVCM	2.33(0.48)	0.83(0.91)	0.82(0.10)	0.83(0.10)	2.43(0.71)
		SSVCM	2.26(0.44)	2.96(2.21)	0.66(0.14)	0.68(0.13)	3.60(1.32)
		QVC	2.32(0.47)	3.24(4.34)	0.67(0.16)	0.69(0.15)	8.47(2.39)
		VC	2.28(0.45)	6.48(3.02)	0.52(0.16)	0.56(0.14)	5.63(2.02)
	Error5	SSQVCM	2.37(0.49)	0.30(0.65)	0.88(0.06)	0.89(0.06)	2.28(0.65)
		SSVCM	2.24(0.43)	2.40(1.86)	0.69(0.13)	0.70(0.12)	3.23(1.05)
		QVC	2.36(0.48)	1.96(5.58)	0.75(0.14)	0.76(0.12)	7.08(1.82)
		VC	2.32(0.51)	5.46(2.35)	0.58(0.18)	0.62(0.16)	4.78(2.08)
$\tau = 0.5$	Error1	SSQVCM	2.40(0.50)	0.73(1.11)	0.84(0.08)	0.85(0.07)	1.88(0.51)
		SSVCM	2.26(0.44)	2.78(1.50)	0.66(0.09)	0.67(0.08)	2.83(0.82)
		QVC	2.40(0.49)	3.14(4.73)	0.67(0.12)	0.69(0.11)	2.61(0.83)
		VC	2.42(0.50)	6.46(2.78)	0.54(0.16)	0.58(0.13)	4.51(1.61)
	Error2	SSQVCM	2.43(0.50)	0.47(0.73)	0.87(0.09)	0.88(0.09)	1.76(0.60)
		SSVCM	2.28(0.50)	2.34(2.32)	0.71(0.15)	0.72(0.14)	3.12(2.25)
		QVC	2.28(0.54)	2.50(5.15)	0.70(0.16)	0.72(0.14)	2.71(1.08)
		VC	2.18(0.69)	6.60(2.44)	0.51(0.18)	0.56(0.14)	5.86(3.50)
	Error3	SSQVCM	2.63(0.49)	0.53(0.78)	0.89(0.09)	0.90(0.09)	1.82(0.58)
		SSVCM	2.44(0.50)	2.64(1.78)	0.70(0.12)	0.71(0.12)	3.14(1.47)
		QVC	2.60(0.49)	3.00(5.80)	0.71(0.14)	0.73(0.13)	2.88(1.08)
		VC	2.56(0.50)	7.18(2.40)	0.54(0.17)	0.59(0.14)	5.40(2.20)
	Error4	SSQVCM	2.40(0.50)	0.47(0.73)	0.87(0.09)	0.87(0.09)	1.93(0.61)
		SSVCM	2.28(0.45)	2.40(1.53)	0.69(0.13)	0.70(0.12)	2.93(1.15)
		QVC	2.40(0.49)	2.58(5.54)	0.71(0.14)	0.73(0.13)	2.82(1.16)
		VC	2.34(0.48)	6.44(2.41)	0.55(0.17)	0.59(0.14)	4.96(2.20)
	Error5	SSQVCM	2.70(0.47)	0.37(0.67)	0.92(0.08)	0.92(0.07)	1.56(0.47)
		SSVCM	2.46(0.50)	1.96(1.35)	0.75(0.12)	0.76(0.12)	2.46(0.87)
		QVC	2.52(0.50)	2.80(5.72)	0.71(0.15)	0.73(0.13)	2.63(0.98)
		VC	2.52(0.50)	6.34(2.39)	0.57(0.17)	0.61(0.15)	4.51(1.94)
$\tau = 0.7$	Error1	SSQVCM	2.70(0.47)	0.47(0.68)	0.91(0.08)	0.91(0.08)	1.55(0.66)
		SSVCM	2.46(0.50)	2.44(1.64)	0.71(0.13)	0.73(0.12)	2.51(1.05)
		QVC	2.60(0.49)	2.20(4.77)	0.76(0.14)	0.77(0.13)	6.19(1.91)
		VC	2.64(0.48)	6.76(1.86)	0.55(0.16)	0.60(0.13)	4.28(1.78)
	Error2	SSQVCM	2.17(0.38)	0.77(1.50)	0.81(0.10)	0.82(0.09)	2.49(0.56)
		SSVCM	2.06(0.37)	3.84(3.86)	0.61(0.15)	0.63(0.14)	4.85(4.41)
		QVC	2.16(0.58)	3.22(6.62)	0.65(0.17)	0.67(0.15)	8.06(2.55)
		VC	2.08(0.70)	7.44(3.66)	0.48(0.18)	0.53(0.14)	7.10(3.56)
	Error3	SSQVCM	2.60(0.50)	0.53(0.86)	0.89(0.08)	0.90(0.08)	1.67(0.58)
		SSVCM	2.44(0.50)	2.58(1.68)	0.70(0.12)	0.72(0.12)	2.63(1.01)
		QVC	2.56(0.50)	2.88(5.44)	0.72(0.16)	0.74(0.14)	9.16(1.57)
		VC	2.50(0.51)	7.14(2.86)	0.53(0.19)	0.58(0.16)	4.85(2.23)
	Error4	SSQVCM	3.00(0.00)	0.47(0.73)	0.95(0.07)	0.95(0.07)	1.88(0.43)
		SSVCM	2.96(0.20)	3.10(2.21)	0.74(0.14)	0.77(0.12)	2.60(0.92)
		QVC	2.94(0.24)	1.06(4.70)	0.90(0.13)	0.90(0.11)	10.31(1.92)
		VC	2.96(0.20)	5.12(1.53)	0.67(0.20)	0.72(0.17)	3.78(2.00)
	Error5	SSQVCM	3.00(0.00)	0.43(0.50)	0.95(0.06)	0.95(0.05)	1.88(0.49)
		SSVCM	3.00(0.00)	2.82(1.72)	0.76(0.12)	0.78(0.11)	2.60(0.92)
		QVC	2.96(0.20)	1.02(4.46)	0.90(0.12)	0.91(0.11)	9.25(2.01)
		VC	2.98(0.14)	5.44(1.52)	0.65(0.19)	0.70(0.16)	3.87(1.71)