VARIABLE SCREENING AND MODEL SELECTION IN CENSORED QUANTILE
REGRESSION VIA SPARSE PENALTIES AND STEPWISE REFINEMENT

BY
LU GAN

DISSE RESS ATION
Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Statistics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:
Professor Douglas Simpson, Chair
Professor Stephen Portnoy, Co-Chair
Professor Roger Koenker
Professor Feng Liang
Abstract

Many variable selection methods are available for linear regression but very little has been developed for quantile regression, especially for the censored problems. This study will look at the possibilities of utilizing some existing penalty variable selection methods on censored quantile regression problems.

In the situation when censored values are not known for each observation, it is common to model the censoring as random. Under the assumption that $y_l$ and $C_l$ are conditionally independent given $x_l$, we use the random censored quantile regression Portnoy estimators (2010). This method simplifies the censored problem into a weight problem. When combined with the penalized regression method: LASSO and SCAD, one can perform variable screening for the censored data at quantiles of interest. Furthermore, we establish the asymptotic property, and illustrate the methodology in the context of ultrasound safety study.
Dedicated to my father.
Acknowledgement

First and foremost, I would like to express my ultimate gratitude towards my beloved advisor Prof. Simpson and co-advisor Prof. Portnoy, for their unbreakable patience and endless support and the wealth of knowledge they have provided. Without them I wouldn’t be able to achieve what I have today.

I would like to thank Prof. Liang and Prof. Koenker for their insightful feedback, which helped me improve my dissertation in many ways.

I would like to thank the rest of the faculty and staff members in the Department of Statistics at the University of Illinois for their assistance throughout my graduate study.

Special thanks goes to my husband Wei and my dearest son Dylon, for the endless joy and love you have given me.

Finally, I want to thank my father for his love and wisdom. Although I will not find you in the crowd when I stand on that stage, but you will always be with me in my heart. I miss you dad.
Table of Contents

List of Tables .................................................................................................................... vii
List of Figures .................................................................................................................. viii
List of Abbreviations ....................................................................................................... ix
Chapter 1 Introduction ...................................................................................................... 1
  1.1 Censored Regression Quantile ................................................................................. 1
  1.2 Penalized Variable Selection Methods .................................................................. 5
Chapter 2 Penalized Quantile Regression for Variable Selection .................................. 7
Chapter 3 Large Sample Theory for Censored Penalized Quantile Regression ............ 13
Chapter 4 Case Study ...................................................................................................... 25
  4.1 Penalized Model Screening - LASSO .................................................................... 28
  4.2 Extension of Portnoy Estimator – Inference Based Refinement of Screened
      Variables ...................................................................................................................... 33
  4.3 Penalized Model Screening - SCAD ........................................................................ 44
  4.4 A Closer Look at Tail Quantiles ............................................................................. 48
  4.5 λ Selection ................................................................................................................ 50
  4.6 Two – Way Interactions ......................................................................................... 51
  4.7 Conclusion ................................................................................................................ 56
Chapter 5 Computing Algorithm ...................................................................................... 59
Appendix A Algorithm ..................................................................................................... 61
  A.1 Model Screening -- LASSO ..................................................................................... 61
  A.2 crq.por.tsp ............................................................................................................ 66
  A.3 crq.fit.por.tsp ....................................................................................................... 68
Appendix B Summary Tables ............................................................................................ 70
  B.1 LASSO 2-way Interaction τ = 0.5: Chest.Wall ..................................................... 70
  B.2 LASSO 2-way Interaction at τ = 0.5: Chest.Atten .............................................. 71
B.3 LASSO 2-way Interaction at $\tau = 0.5$: $Freq$ .......................................................... 72
B.4 LASSO 2-way Interaction at $\tau = 0.5$: $Pd$ .............................................................. 73
B.5 LASSO 2-way Interaction at $\tau = 0.5$: $Prf$ ................................................................. 74
B.6 LASSO 2-way Interaction at $\tau = 0.5$: $ed$ ................................................................. 75
B.7 LASSO 2-way Interaction at $\tau = 0.5$: $beam$ .............................................................. 76
B.8 LASSO 2-way Interaction at $\tau = 0.5$: $Pr$ ................................................................. 77

Bibliography .................................................................................................................................. 78
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Summary of pig age versus threshold peak rear factional pressure in the Gaussian–Tobit analysis of pig data</td>
<td>26</td>
</tr>
<tr>
<td>4.2</td>
<td>Summary table of estimates for ( pr + pr<em>mi + chest.wall</em>mi + pd*mi + prf + ed )</td>
<td>38</td>
</tr>
<tr>
<td>4.3</td>
<td>Summary table of estimates for ( pr<em>mi + chest.wall</em>mi + pd + prf )</td>
<td>41</td>
</tr>
<tr>
<td>4.4</td>
<td>Summary table of estimates for ( pr<em>mi + chest.wall</em>mi + pd + prf + beam )</td>
<td>43</td>
</tr>
<tr>
<td>4.5</td>
<td>Summary table of LASSO and SCAD screening results for a sequence of ( \tau )'s</td>
<td>49</td>
</tr>
<tr>
<td>4.6</td>
<td>Summary of ( \lambda ) vs. ( \tau )</td>
<td>50</td>
</tr>
<tr>
<td>4.7</td>
<td>Summary table of 2-way interaction for LASSO</td>
<td>52</td>
</tr>
<tr>
<td>4.8</td>
<td>Crq results of the full model at ( \tau = 0.5 ), and ( R=10 )</td>
<td>53</td>
</tr>
<tr>
<td>4.9</td>
<td>Crq results of the final model with quadratic term ( pr^2 ) at ( \tau = 0.5 ), and ( R=100 )</td>
<td>54</td>
</tr>
</tbody>
</table>
List of Figures

4.1  Plot of Powell predictions where \( y = \) lesion depth and \( x = \) threshold wave length for two different starting value........................................................................................................28

4.2  Prediction Made Using Portnoy’s and Peng and Huang’s estimates .........................................................29

4.3  Plot of \( \lambda \) versus coefficients of standardized variables .................................................................31

4.4  Plot of \( \lambda \) versus coefficients of original variables (none standardized) ........................................32

4.5  Confidence band of the base model \( pr + chest.wall + prf + mi + freq + ed + pd \), where \( x \)-axis is \( 1-\tau \) and \( y \)-axis is the coefficient .................................................................34

4.6  Confidence band of \( (pr + chest.wall + prf + ed + pd) \)*(\( mi + freq \)), where \( x \)-axis is \( 1-\tau \) and \( y \)-axis is the coefficient ..................................................................................................................................................35

4.7  Confidence band of \( pr + pulse + pr*mi + chest.wall*mi + chest.wall*rw \) \( pr + pd*mi + prf + ed \), where \( x \)-axis is \( 1-\tau \) and \( y \)-axis is the coefficient ..................................................................................................................................................36

4.8  Confidence band of \( rw_pr + pr*mi + chest.wall*mi + pd*mi + prf + ed \), where \( x \)-axis is \( 1-\tau \) and \( y \)-axis is the coefficient ..................................................................................................................................................37

4.9  Confidence band of \( pr*mi + chest.wall*mi + pd + prf \), where \( x \)-axis is \( 1-\tau \) and \( y \)-axis is the coefficient ..................................................................................................................................................40

4.10 Confidence band of \( pr*mi + chest.wall*mi + pd + prf + beam \), where \( x \)-axis is \( 1-\tau \) and \( y \)-axis is the coefficient ..................................................................................................................................................42

4.11 Plot of \( \lambda \) versus coefficients of the standardized variables for SCAD .................................................45

4.12 Confidence band of \( mi + pr + pd + chest.attcn \), where \( x \)-axis is \( 1-\tau \) and \( y \)-axis is the coefficient ..................................................................................................................................................46

4.13 Histogram of \( chest.attcn \) ..................................................................................................................................................47

4.14 Example of widening of Confidence Band .......................................................................................................48

4.15 Summary of \( \lambda \) vs. \( \tau \) ..................................................................................................................................................51

4.16 Confidence plot of the final model at \( \tau = 0.5 \), and \( R=10 \) .................................................................................................................................55
List of Abbreviations

CLAD  Censored Least Absolute deviations
LASSO  Least Absolute Selection and Shrinkage Operator
SCAD  Smoothly Clipped Absolute Deviation
animal  Animal ID
chest.wall  Thickness of Chest Wall to Lung in mm
chest.atten  Attenuation Coefficient in dB/mm
freq  Ultrasound Frequency in MHz
pd  Pulse Duration in mu-sec
prf  Pulse Repetition Frequency
ed  Exposure Duration
beam  Width of Beam at Focus
pr  Calculated in Vivo Acoustic Pressure at the Lung Surface
mi  Mechanical Index
Chapter 1

Introduction

Many variable selection methods are available for linear regression but very little has been developed for quantile regression, especially for the censored problems. This study will look at the possibilities of utilizing some existing penalty variable selection methods on censored quantile regression problems.

1.1 Censored Regression Quantile

Censoring occurs when the recorded data is only partially known or occurs outside the range of measuring instrument (Chay et. Al, 2001). Many statistical methods can be used to fit models with censored data. The traditional statistical analysis used the maximum likelihood procedures. This method requires specification of the error distribution and a wrong error distribution will cause this method to generate an estimate that is not consistent (Honore and Khan, 2002).

The Tobit model in context of quantile regression has been of great interest among the researchers. It is also known as the censored normal regression model and has the following structural equation:

\[ y_i^* = x_i' \beta_0 + \epsilon_i, \quad i = 1, ..., n \]

where \( \epsilon_i \sim N(0, \delta^2) \), \( y^* \) is a latent variable that is observed for values greater than \( C \) and censored otherwise. The observed \( y \) is defined by the following measurement equation
\[ y_i = \begin{cases} y^* & \text{if } y^* > C \\ C_y & \text{if } y^* \leq C \end{cases} \] \tag{2}

In the ultrasound safety example below, the data set is censored at 0 lesion depth. Thus, we have

\[ y_i = \begin{cases} y^* & \text{if } y^* > 0 \\ 0 & \text{if } y^* \leq 0 \end{cases} \] \tag{3}

A required feature of Tobit model is that the censoring points are known for the dependent variables for all the observations. If the distribution of the error term is indeed Normal and homoscedastic, then it is straightforward to derive and maximized the likelihood function. If these two conditions are unsatisfied then the Tobit estimator will be inconsistent (Yu and Stander, 2007). It is obvious these assumptions put stringent restrictions on the application of this model. Nevertheless, it is interesting to study how the Tobit estimator differs from the more generalized estimator such as the Portnoy (2003) censored quantile regression estimator, which requires less assumptions. A small example on the neonate data set was performed. And the results showed that the estimations from these two methods are very similar.

Powell (1986) proposed the censored least absolute deviations (CLAD) estimation method. He noted that the estimator for the Tobit model solves

\[
\min_{\beta} \left\{ \frac{1}{n} \sum_{i=1}^{n} \rho_{\theta} (y_i - \max \{y_i^0, X_i \beta\}) \right\},
\]

\[
\text{Where } \rho_{\theta} (\lambda) = [\theta - I(\lambda < 0)] \lambda \text{ is the check function and } I(\cdot) \text{ is the usual indicator function (Buchinsky and Hahn, 1998). Although, Powell's work opened a new window to solving censored problems, with error distribution that is nonnormal or heteroskedastic, his estimator has some disadvantages. The most notable disadvantage is that the objective function is not convex in } \beta, \text{ therefore, local optimization may not obtain a global minimizer. The algorithm described}
\]
by Fitzenberger (1996) helps solve this problem, but still doesn’t ensure convergence to a local optimum. It should be noted when implementing the Powell's method using the QUANTREG package in R, one can specify the starting value of the coefficients. In the present study, we looked at both the Powell estimator and Portnoy estimator (Portnoy, 2003). When running the Powell method in CRQ program, the default starting regression quantile estimate on the Neonate lung data set had a strong influence on the resulting estimate of the coefficients. The majority of the runs could not be completed and, hence, failed to generate coefficient estimates for any of the quantiles. The reason for such failure is unclear. It is suspected that the large ratio, approximately 0.5, of the censored and uncensored data points may have caused this failure, where the censored value is 0. Because of the problem we experienced with the Powell’s estimator, it was decided that it is not suitable for this particular data and will not be further used in our study. Our focus is put on the Portnoy estimator that has more general censoring assumptions, which we will discuss in more detail later on.

In the situation when censored values are not known for each observation, it is common to model the censoring as random. Under the assumption that $y_i$ and $C_i$ are conditionally independent given $x_i$, there are two commonly used random censored quantile regression estimators: the Portnoy (2003) and Peng and Huang (2008) estimators. Both are incorporated in the CRQ program in R (Koenker, 2008). Recently it has been proven that in large sample environment Peng and Huang's and Portnoy estimator are asymptotically equivalent (Limin Peng, 2009). The present study focuses on the Portnoy estimator. One computational advantage of using Portnoy estimator is it is relatively simpler to retrieve the weights used in this method than in other methods. These weights are crucial when using the LASSO variable selection method (Tibshirani, 1996).
When presented with one-sample problem where \( y_i = \min (y_i^*, c_i) \), Kaplan-Meier can be used to estimate the quantile function instead of the survival function. One method of computing Kaplan-Meier estimator is to recursively reweights the estimates by redistributing the mass of the censored data point, \( P[Y_i > C_i] \), to the observations above \( C_i \). The fundamental problem is finding \( P[Y_i > C_i] \). In terms of quantile regression, this is equivalent to finding the \( \tau \) which \( x_i' \beta(\tau) \) crosses \( C_i \). The weights can then be incorporated into the parametric approach described by Koenker and D’Orey (1987). The disadvantage of reweighting the estimate is the monotonicity condition weakens, where even under the one-sample setting, it is possible to have censored observations crossed by estimate quantile and which will cause it to have negative residuals (Koenker and D’Orey, 1987). In his 2003 paper, Prof. Portnoy introduced an effective method that is similar to the approach described in Koenker and D’Orey (1987), but solves the "recrossing" problem by computing \( \hat{\beta}(\tau) \) along a grid of \( \tau \) values.

From the detailed algorithm described in Portnoy (2003), one can simply retrieve the weights from the crq.fit.por program by adding the \( \tau_i \)'s of the crossed censored data points, named \( \text{splittau} \) in the R-code, to the list of returned variables within the algorithm of the crq.fit.por program. The crossed censored data points that correspond to split \( \tau_i \) and are smaller than \( \tau \) will have its weight split into two parts:

1) \( w_i(\tau) = \frac{\tau - \tau_i}{1 - \tau_i}, \text{ for } (x_i, +\infty) \), \[ 5 \]

2) \( 1 - w_i(\tau) \), otherwise, \[ 6 \]

where \( \tau > \tau_i \). As summarized in Koenker (2008), the intuition of this method is that the quantile depends solely on how much mass is below and how much is above. Assigning \( 1 - w_i(\tau) \) to
some arbitrarily large response and \( w_i(\tau) \) to \( \tau_i \) ensures that the weighted response corresponds to the \( \tau^{th} \) conditional quantile.

An advantage of the weighted censored regression quantile approach is that one can, in principal, use any kind of penalty in combination with the censor-weight modified regression quantile. This is especially helpful computationally, because crq program in the quantreg package (Koenker, 2012) already allows certain penalties. In this study, two penalized variable selection methods are of primary interest. One is the least absolute selection and shrinkage operator (LASSO) Tibshirani (1996) and the other is the smoothly clipped absolute deviation (SCAD) Fan and Li (2001).

1.2 Penalized Variable Selection Methods

The well-known traditional model selection criteria, such as AIC and BIC, are infeasible for variable selection in high-dimensional data. When a large number of predictors are introduced at the initial stage, these methods become prone to suffer from instability and they involve heavy computations (Breiman, 1996). Two feasible alternative methods for model selection of high-dimensional problems are LASSO and SCAD. Both methods are penalized maximum likelihood estimators that can jointly minimize the residual sum of squares plus penalty and can avoid large variation, which occurs in estimating complex models. The computational simplicity of LASSO has attracted much attention of many users. Its computational complexity is similar to just one linear regression. But its dubious consistency property makes it tricky to use (Zou and Li, 2008). Zhao and Yu (2003) showed that if an irrelevant covariate is strongly correlated with the significant covariates, then no matter how large the data set is, LASSO will most likely fail to
distinguish the true covariates. Furthermore, Zhao and Yu showed that there exists an almost necessary and sufficient condition, the irrepresentable condition, for which the consistency of LASSO will hold true. The SCAD estimator of Fan and Li (2001) was introduced to improve the performance. Many scientists have studied its asymptotic properties and have found that under appropriate conditions, the SCAD estimator is consistent for variable selection. In addition, its asymptotic distribution and its 'oracle' property can be computed when the SCAD estimator is consistent.

The LASSO and SCAD methods have been adapted for use when dealing with weighted-regression. The process of retrieving the censor related weights from the Portnoy's algorithm and feeding these weights along with the response and predictors into the algorithm of these methods provides an approach for penalized censored quantile regression. Although, the asymptotic consistency of the LASSO and SCAD methods and the Portnoy estimator (Portnoy and Lin, 2010) have been established, the main theoretic hurdle of this study is to prove that the consistency property does indeed extend to the weights, such that the transformed response and predictor variables will still hold the asymptotic consistency properties of the LASSO and SCAD results.

This paper is organized as follows. In Chapter 2, we describe the proposed penalized quantile regression variable selection method. In Chapter 3, we present the large sample theory. In Chapter 4, we examine the high-dimensional data set using the proposed variable selection method. Finally, in Chapter 5, we provide the computing algorithm used in the penalized quantile regression.
Chapter 2

Penalized Quantile Regression for Variable Selection

There are limited tools for performing variable selection in quantile regression. Although in this study we focus on censored quantile regression, our method can be applied to all quantile regression problems. The basic idea of our method is 1) variable screening; 2) Fine tuning of variables using backwards selection and bidirectional regression. We tested LASSO and SCAD variable screening method for the censored problem. This is more complicated than regular quantile regression, but using Portnoy method, the censored data is simplified into a weighted problem. Then manually applied the backwards and bidirectional regression for fine-tuning of variable and their interactions terms.

The crq function offered in the quantreg package (Koenker, 2012) fits conditional quantile regression model for censored data. User can choose from three methods: Powell (1986), Portnoy (2003), Peng and Huang (2008). We are interested in random censoring in particular the Portnoy’s method. When the Portnoy’s method is specified in crq, this signals the program to call the function crq.fit.por. The original crq.fit.por command in R does not return the index of the crossed censored data points and the crossed $\tau$. But these two pieces of information are vital in the calculation of weights generated by the Portnoy’s random censored quantile regression method at the crossed data points. To retrieve the above information we simply add the variable Isplit, which stores the index of the crossed censored response, and $T = tsp$, which stores the
quantile that corresponds to $Isplit^{th}$ crossed censored response, to the list of the returned variables defined in crq.fit.por algorithm. We name this new version of crq.fit.por that returns these two additional variables, the crq.por.tsp. This new version will replace the original crq.fit.por when Portnoy’s method is called in crq function.

The algorithm for calculating the split weights is rather straightforward. First, we retrieve $Isplit$ and $T$. We then duplicate the entries identified by $Isplit$ for which $\tau_i < \tau$. This creates the additional entries for us to assign the split weight. Third, overwrite the response of the duplicated entry with a much larger value, say 100. Since the quantile can be determined by only knowing the sign of the residual, therefore, assuming the response is the arbitrary value of 100 is equivalent to knowing the true observation because the sign of their residual are the same. Hence, a part of the weight of the crossed censored observation is distributed to the true observation, even though the true value is unknown. Finally, assign the correct weights according to equation [4] and [5]. All other entries that are uncensored or uncrossed have the weight of 1. Multiplying this vector of weights to the response vector and the $X$ matrix (this $X$ matrix includes the duplicated entries with the edits mentioned above), transforms the problem into a weighted regression problem.

Please note the crq function was designed to handle only right random censoring problems until recently. Majority of the analysis were performed before this update of crq. Originally to compute the left censoring problem we had to change the sign of the response variable. Although this tricks the program to compute left censoring as the right censoring problems, it also replaces $\tau$ with $1-\tau$. Therefore, the plots given in this paper actually has x-axis equal to $1-\tau$ rather than $\tau$, but the discussion will assume left censoring. So small responses will
correspond to small $\tau$, and large responses will correspond to large $\tau$. The most recent version of rq package now appears to allow specification of left censoring.

In this study we used two popular variable screening methods: the L1 norm LASSO penalty of Tibshirani and the L1 norm SCAD penalty of Fan and Li. These two methods have conveniently built-in functionality in the quantreg package. The algorithm for analyzing the weighted regression problem using these two methods is rather straightforward. Users only need to plug in the weighted $x$ and $y$ values calculated following the above method into the function rq.fit.lasso and rq.fit.scad and specify the $\lambda$ value. The penalty parameter $\lambda$ determines how much shrinkage is done.

The solution to the LASSO with L1 penalty can be given as

$$\hat{\beta}^{\text{LASSO}} = \underset{\beta}{\text{argmin}} \left( ||Y - X\beta||^2 + \lambda ||\beta||_1 \right), \quad [7]$$

where $\lambda \geq 0$ is the shrinkage parameter (Tibshirani, 1996). When $\lambda$ is large this becomes an ordinary least squares linear regression problem. But for $\lambda$ that is sufficiently small, it would be possible to shift through the variables and separate the significant variables from the rest. It has been shown (Zhao and Yu, 2006) that there exists an irrepresentable condition, which depends mainly on the covariance of the predictor variables. And LASSO selects the true model consistently if and (almost) only if the predictors that are not in the true model are “irrepresentable” (in a sense to be clarified) by predictors that are in the true model. The quantile regression version of LASSO is

$$\hat{\beta}^{\text{LASSO}} = \underset{\beta}{\text{argmin}} \left( \rho_\tau (wY, wX\beta) + \lambda ||\beta||_1 \right). \quad [8]$$
The smoothly slipped absolute deviation (SCAD) was first proposed by Fan and Li (2001). The penalty is given by

$$\hat{\beta}_{SCAD}^{\lambda} = \arg\min_{\beta} \left( \frac{1}{2} \|Y - X\beta\|^2 + \lambda \sum_{j=1}^{d} p_j(|\beta_j|) \right).$$  \[9\]

The solution to the SCAD penalty is given as

$$\hat{\beta}_{j}^{SCAD} = \begin{cases} 
(\lambda - \beta_j) + \text{sign}(\beta_j) & \text{if } |\beta_j| < 2\lambda; \\
\frac{(a-1)\lambda - \text{sign}(\beta_j)a\lambda}{a-2} & \text{if } 2\lambda < |\beta_j| \leq a\lambda; \\
\beta_j & \text{if } |\beta_j| > a\lambda
\end{cases}$$ \[10\]

It has been shown (Huang and Xie, 2007) that under appropriate conditions, the SCAD-penalized least squares estimator is consistent for variable selection. The estimators of nonzero coefficients have the same asymptotic distribution as they would have if the zero coefficients were known in advance. The quantile version of the SCAD is

$$\hat{\beta}_{SCAD}^{\lambda}(\lambda) = \arg\min_{\beta} \left( \frac{1}{2} \rho_{\tau}(wY, wX\beta) + \lambda \sum_{j=1}^{d} p_j(|\beta_j|) \right).$$ \[11\]

Oracle results are available for the SCAD method. Because of the oracle properties of SCAD, it was found the SCAD penalty method outperforms the LASSO in terms of model error (Zou and Li, 2009). But because SCAD is a nonconcave penalty, it is much more difficult to compute than LASSO. Furthermore, Hall and Lee and Park (2009) showed that the bootstrap-based penalty choice for the LASSO can achieve oracle performance.

LASSO and SCAD is different mostly in its penalty term. But similarly the computation of penalties only involve $\beta$, do not work with the weights for the censored data directly. This makes the problem much simpler to handle and relatively easier to prove its consistency. Chapter 3 discusses the proof of consistency in large sample. In general, one would expect
LASSO and SCAD results to differ in varying degrees. It is expected their result to be most similar at the quantile with most data, i.e. the median, and most different when there is least amount of data, i.e. the boundaries. In Chapter 4, we will compare screening results at different levels of \( \tau \) for a specific case study.

How to choose the right \( \lambda \) can be tricky. For example, when \( \lambda \) is too large LASSO becomes an ordinary least squares linear regression problem. When \( \lambda \) is too small, significant variables cannot be separated from the rest. Instinctively we feel this threshold \( \lambda \) is likely subjected to the magnitude of variables and the specific dataset. To remove the magnitude effects, we can simply standardize the variables. The dataset effect on \( \lambda \) is much more complicated to remove if not entirely impossible. But if we can determine a certain pattern in threshold \( \lambda \) and \( \tau \), then we could possibly generalize a method for selecting the appropriate threshold \( \lambda \). But this requires tremendous amount of data. We begin by looking at the lung dataset, which will be elaborated in Chapter 4.

There are some disadvantages in using LASSO and SCAD in crq. Particularly, when dealing with higher order models, performing LASSO and SCAD can be a tedious task. For example, one cannot simply multiplying the variable together to create the interaction term in the command object of crq.LASSO, which is common privilege in other crq fitting tools. Hence, the variables must be manually multiplied together and added to the data frame. So the 2-way variable is treated as a new variable. This can be cumbersome when you have many 2-way variables in your model. Most importantly, when model gets too lengthy, crq can take a very long time to run, and in our experience it may fail to converge. Instead of putting all the 2-way variables in one large model with its main effects, we propose two ways to analyze a complex model. First, screen the main effects model and manually perform bidirectional regression by
entering the 2-way interactions. Second, screen each main effect along with all its 2-way interaction individually, then, combine the results and manually perform a backward regression to fine-tune the model. Surprisingly, two methods produced very similar results. Though, the backward regression may be simpler to program into a function that can be applied automatically in *qr.*
Chapter 3

Large Sample Theory for Censored Penalized Quantile Regression

In this chapter we will establish consistency of the proposed censored penalized quantile regression estimator to provide theoretical backing for the use of the procedure in practice. It has been shown under appropriate conditions, the SCAD-penalized least squares estimator is consistent for variable selection (Huang and Xie, 2007). In addition, we also know the model selection consistency of the LASSO is highly dependent on some irrepresentable condition (Zhao and Yu, 2006). The asymptotic distributions of both methods have been derived and shown to perform consistent modal selection (Potscher and Leeb, 2009). Therefore, in principle, since individual methods of the proposed penalized regression quantile estimator has been proven to be asymptotically consistent under certain conditions, it seems natural for the combination of the LASSO or SCAD with Portnoy estimator to produce an asymptotically consistent results. In this chapter we will show, in finite dimensional problem with large sample, our variable selection method produces consistent results and the asymptotic normal distribution holds.

Penalized and unpenalized censored quantile regression differs only by a penalty term. Naturally, the consistency proofs are very similar. We have adapted previous results from Portnoy and Lin (2010) to establish the consistency for the finite dimensional penalized censored problem. In addition, as an analogy to Portnoy and Lin (2010) work, the asymptotic distribution theory also
holds for penalized version. First, define the subgradient of the penalized quantile loss function 

\[ \Psi_{k+1}^p \] as follows:

\[
\Psi_{k+1}^p (w_l, \beta(t_k)) = \sum_{i=1}^{n} x_i \{ I(\Delta_i = 1)\psi(Y_i - x_i' \beta(t_k), t_k) + I(\Delta_i = 0) \}
\times \left[ w_i(t_k)\psi(C_i - x_i' \beta(t_k), t_k) + \left(1 - w_i(t_k)\right)\psi(Y^* - x_i' \beta(t_k), t_k) \right] + \sum_{i=1}^{d} p_i(t_k) \]

where \( \psi(u, \tau) = \tau - I(u \leq 0) \), and \( Y^* \) is sufficiently large value larger than all observed and fitted values. As described in Koenker (2008), the gradient conditions impose a bound of the form \( \Psi_k(\varnothing, b) = O(1) \) at \( b = \beta(t_k) \) (uniformly in \( k \)), as long as \( \|x_i\| \) remain bounded (see condition (5) below).

Here we will restate the conditions needed for the results in Portnoy and Lin (2012) plus 4 additional conditions (8)-(11):

1. All conditions restrict to \( \epsilon \leq \tau \leq \min\{\bar{\tau}, 1 - \epsilon\} \) where \( \bar{\tau} \) is the largest identifiable \( \tau \)-value.

Furthermore, there is no censoring below \( x' \beta(\epsilon) \). Hence, \( \hat{\beta}(\epsilon') \) can be computed as an unweighted regression quantile for \( \epsilon' < \epsilon \) with probability tending to one.

2. \( (x_i, Y_i, C_i) \) are i.i.d.

3. \( Y_i \) and \( C_i \) are conditionally independent given \( x_i \) and have conditional distribution functions \( F_i \) and \( G_i \), respectively.

4. The conditional densities \( f_i(x_i' \beta(t)) \) and \( g_i(x_i' \beta(t)) \) (conditional on \( \{x_i\} \)) have uniformly
bounded derivatives (with respect to \( t \)) on \( \epsilon \leq \tau \leq \min\{\bar{\tau}, 1 - \epsilon\} \), and are strictly positive on this set.

(5) \( \|x_i\| \) has bounded support.

(6) \( \{t_1, \ldots, t_M\} \) is a grid with mesh \( \delta_n = c_n n^{-a} \) for some \( a \) with \( \frac{1}{4} < a < \frac{1}{2} \) and \( c_n \to c \) (with \( c > 0 \)).

(7) The design matrix, \( X \), satisfies \( n^{-1} X' X \to A \), where \( A \) is invertible.

(8) \( \{x_i\} \) is consider fixed.

(9) \( \beta(t) \) is boundely differentiable on \( \epsilon \leq \tau \leq \min\{\bar{\tau}, 1 - \epsilon\} \).

(10) \( p'_\lambda \leq c\lambda_n \), where \( c > 0 \).

(11) There is a constant \( c^* \) such that \( \lambda_n \) satisfies \( \lambda_n \leq c^* (d_{k,n} \delta_n) \) where \( d_{k,n} \) and \( \delta_n \) are defined in Theorem 3.1.

**Inductive Proof of Consistency**

From this point on, we will follow the exact steps of Portnoy and Lin (2012) inductive proof of consistency.

**THEOREM 3.1.** Let \( \hat{\beta}_M \equiv (\hat{\beta}(t_1)', \ldots, \hat{\beta}(t_M)') \) be the right censored penalized quantile estimator along the grid \( \epsilon = t_1 < t_2 < \cdots < t_M \leq \min\{\bar{\tau}, 1 - \epsilon\} \) (where \( \bar{\tau} \) is the largest identifiable \( \tau \)-value). Under Conditions (1)–(7), we have

\[
\|\hat{\beta}(t_k) - \beta(t_k)\| \leq 2 r_1 n^{-1} d_{k,n}, \quad k=1, \ldots, M, \]

where \( d_{k,n} \) and \( \delta_n \) are defined in Theorem 3.1.
where \( M = o(n^{1/2}) \), \( d_{k,n} = R_n \sqrt{n} (1 + 2r_1 r_2 E_n^* \delta_n)^{k-1} \) with \( E_n^* = O_p(1) \), and \( R_n = O_p(1) \) and is defined by: 
\[
R_n = n^{-1/2} \max_k \left\{ \left\| \Psi_{k+1} \left( w_t, \beta(t_k) \right) \right\| + \left\| E_{n,k} \right\| \right\} .
\]
Here, \( \lambda_n = O(d_{k,n} \delta_n) \), \( \Psi_{k+1} \) is the defined by Equation (1) and \( r_1 \) and \( r_2 \) are defined by Equations (2) and (3); and we show that \( E_{n,k} = O_p \left( n^{1/4} \log n \right) \) uniformly in \( k \), where \( E_{n,k} \) is defined by Equation (4). Recall that \( \beta(t_k) \) is the true regression quantile along the same grid, and \( \delta_n = O(n^{-a}) \) is defined in Condition (6).

Remark: note that since \( d_{k,n} \delta_n = O \left( n^{3/2 - a} \right) \) (by condition (6)), the rate at which \( \lambda_n \) can tend to infinity is controlled by the constant, \( a \). Under the hypothesis of Theorem, the rate can grow at least algebraically (since \( a < 1/2 \)), but can grow at a rate no faster than \( n^{1/4} \).

**Proof** Let \( CI_k = \{ i : Y_i = C_i \text{ and } \max\{ \hat{t}_i, \tau_i \} \leq t_k \} \) be the index set of the crossed censored observations.

We shall use mathematical induction to show that for any \( k = 1, 2, \ldots, M \),
\[
\sum_{i \in CI_k} | \hat{t}_i - \tau_i | \leq d_{k,n}, \text{ and } \| \hat{\beta}(t_k) - \beta(t_k) \| \leq 2r_1 n^{-1} d_{k,n} .
\]
[13]

First let \( k = 1 \), \( \hat{\beta}(t_1) \) is the penalized quantile estimator at \( t_1 \) by applying the usual censored regression quantile. Since we are dealing with a left censoring problem, then there are no censored data to the left of \( t_1 \), hence, it can be treated as an uncensored quantile regression problem. Furthermore, it is known that \( \| \hat{\beta}_U(t_1) - \beta(t_1) \| = O_p(1) \) by Theorem 3.1 in Koenker (2008), where \( \hat{\beta}_U(t_1) \) is the uncensored quantile regression estimator at \( t_1 \). Therefore,
\[
\| \hat{\beta}(t_1) - \beta(t_1) \| = O_p \left( n^{-1/2} \right) \leq 2r_1 n^{-1} d_{k,n} .
\]
Using the fact that there are no censored data to the left of \( X' \beta(t_i) \), it follows the \( \tau_i \) and \( \hat{\tau}_i \) must be greater than \( t_i \). Hence, \( \sum_{i \in C_t} |\hat{\tau}_i - \tau_i| = 0 \), and Equation [13] is true \( k=1 \).

As shown in Portnoy and Lin (2010), assume that for \( k=1 \), Equation [13] is true, then the bound for the difference between the estimated weights and the true weights at \( t_{k+1} \)th quantile can be expressed as follows:

\[
\sum_{i=1}^{n} \left| w_i \left( \hat{\beta}_{i}, t_{i+1} \right) - w_i \left( \beta_{i}, t_{i+1} \right) \right| \leq \frac{1-\epsilon}{\epsilon^2} d_{t_{i}, n} \left( 1 + E_n \delta_n \right). \tag{14}
\]

Define \( \eta_n(\theta, \beta(t_{i+1})) = \Psi_{t_{i+1}} \left( w \left( \hat{\beta}_{i}, t_{i+1} \right), \theta \right) - \Psi_{t_{i+1}} \left( w \left( \beta_{i}, t_{i+1} \right), \beta(t_{i+1}) \right) \), with \( \Psi_{t_{i+1}} \) given by Equation (1). From the proof of Lemma 4.1 in He and Shao (1996), we can see that for any constant \( C^*>0 \), there is a constant \( A_2>0 \), such that for large enough values of \( A_1>0 \) and \( n \),

\[
P \left( \max_{1 \leq i \leq M} \sup_{\theta: \|	heta - \beta(t_{i+1})\| \leq C^* n^{-\frac{1}{2}}} \| \eta_n(\theta, \beta(t_{i+1})) - E \eta_n(\theta, \beta(t_{i+1})) \| > A_1 n^{\frac{1}{2}} \log n \right) \leq MA_2 e^{-A_1 n}. \]

Let \( E_{n,l} \) be defined as

\[
E_{n,l} = \eta_n \left( \theta, \beta(t_{i+1}) \right) - E \eta_n \left( \theta, \beta(t_{i+1}) \right) \tag{15}
\]

on \( \left\{ \theta: \|	heta - \beta(t_{i+1})\| \leq C^* n^{-\frac{1}{2}} \right\} \). As shown in Portnoy and Lin (2010) proof of Theorem 2.1, for none penalized censored problem \( E_{n,l} = O_p(n^{1/4} \log n) \) uniformly in \( l \). For clarification purposes let the superscript \( p \) symbolize the penalty affect. Note, \( \sum_{i=1}^{d} p'_{\lambda} \) is not a function of \( y \), therefore, adding a penalty in the regression will not change this limiting behavior. This is shown below.
\[ E_{n,l}^p = \Psi_{l+1}^p \left( w \left( \beta_i, t_{l+1} \right), \theta \right) - \Psi_{l+1}^p \left( w \left( \beta_i, t_{l+1} \right), \beta(t_{l+1}) \right) \]

\[ -E_Y \left[ \Psi_{l+1}^p \left( w \left( \beta_i, t_{l+1} \right), \theta \right) - \Psi_{l+1}^p \left( w \left( \beta_i, t_{l+1} \right), \beta(t_{l+1}) \right) \right] \]

\[ = \Psi_{l+1} \left( w \left( \beta_i, t_{l+1} \right), \theta \right) - \Psi_{l+1} \left( w \left( \beta_i, t_{l+1} \right), \beta(t_{l+1}) \right) \]

\[ -E_Y \left[ \Psi_{l+1} \left( w \left( \beta_i, t_{l+1} \right), \theta \right) - \Psi_{l+1} \left( w \left( \beta_i, t_{l+1} \right), \beta(t_{l+1}) \right) \right] \]

\[ + \sum_{j=1}^{d} p'_{\lambda} (\theta) - \sum_{j=1}^{d} p'_{j,\lambda} (\beta(t_{l+1})) - E_Y \left[ \sum_{j=1}^{d} p'_{\lambda} (\theta) - \sum_{j=1}^{d} p'_{j,\lambda} (\beta(t_{l+1})) \right] \]

\[ = \eta_n (\theta, \beta(t_{l+1})) - E_Y \eta_n (\theta, \beta(t_{l+1})) \]

\[ = E_{n,l}. \]

Since \( \Psi_{l+1}^p = \Psi_{l+1} + \sum_{j=1}^{d} p'_{\lambda} \) it follows the

\[ \Psi_{l+1}^p \left( w \left( \beta_i, t_{l+1} \right), \theta \right) = \Psi_{l+1} \left( w \left( \beta_i, t_{l+1} \right), \theta \right) + \sum_{j=1}^{d} p'_{j,\lambda} (\theta). \] [16]

Using the results from Portnoy and Lin (2010) proof of Theorem 2.1, \( \Psi_{l+1}^p \left( w \left( \beta_i, t_{l+1} \right), \theta \right) \) can be expressed in the following form:

\[ \sum_{i \in C_l} \left( w_i - w_i \right) I (C_i < x_i^\theta) x_i + \Psi_{l+1} \left( w \left( \beta_i, t_{l+1} \right), \beta(t_{l+1}) \right) \]

\[ + X'VX (\theta - \beta(t_{l+1})) + E_{n,l} + \sum_{j=1}^{d} p'_{j,\lambda} (\theta). \] [17]
Note that, uniformly in $l$ (and with $w$ denoting $w \left( \beta_l, t_{t+1} \right)$),

(a) $\|\Psi_{t+1} (\widehat{\omega}, \theta)\| = O \left( n^{1/2} \right)$,

(b) $\| \sum_l (\widehat{\omega}_l - w_i) I (C_i < x_i' \theta) x_i' I (\Delta_i = 1) \| \leq \sum_i \| x_i \| (\widehat{\omega}_i - w_i) \leq \| x_i \| \left( \frac{1-\epsilon}{(1-\gamma)^2} d_{l,n} \right) = O_p \left( n^{1/2} \right)$,

Note: $\| x_i \|$ has a bounded support. Hence, in (b) we assume $\| x_i \|$ is the max of all possible $\| x_i \|$.

(c) $\| \Psi_{t+1} (w, \beta(t_{t+1})) \| = O_p \left( n^{1/2} \right)$,

(d) $E_{n,l} = O_p \left( n^{1/4} \log n \right)$, and

(e) $\lambda_{max} ((X'VX)^{-1}) \leq a^{-1} \lambda_{max} ((X'X)^{-1}) \leq a_1 n^{-1}$,

where $V_{ii} \geq a > 0$, for some $a_1 > 0$ uniformly in $i$. Because of condition (7), the right side inequality of (e) is true.

The convex property of $\Psi_{t+1} (w, \theta)$ and SCAD penalty $\sum_{j=1}^{q_n} p'_A (\theta)$ allows for $\Psi^p_{t+1} (w, \theta)$ to be convex. Note, $q_n$ is the number of non-zero coefficients and $p'$ is the derivative of the SCAD penalty. It follows the above results may be used without changing the gradient condition of $\{ \theta: \| \theta - \beta(t_{t+1}) \| > C^* n^{-1/2} \}$, under the condition that a large enough $C^*$ is chosen (Theorem 2.1 in Koenker (2008)). Replace $\theta$ with $\beta(t_{t+1})$ in Equation [17] and solve for $\| \beta(t_{t+1}) - \beta(t_{t+1}) \|$ we have
\[ \| \beta(t_{l+1}) - \beta(t_{l+1}) \| \]

\[ = \left\| (X'VX)^{-1} \sum_{i \in C_{l1}} (\bar{w}_i - w_i)(C_i < x'_i \theta)x_i + \Psi_{l+1} \left( w(\beta_i, t_{l+1}), \beta(t_{l+1}) \right) \right\| \]

\[ - \Psi^p_{l+1} \left( w(\beta_i, t_{l+1}), \beta(t_{l+1}) \right) + E_{n,l} + \sum_{j=1}^{q_n} \left| p'_j(\theta) \right| \]

Since \( \|x_i\| \) is bounded, by definition \( \Psi_k(\bar{w}, b) = O(1) \) at \( b = \hat{\beta}(t) \), which is small enough to be
absorbed into \( O_p(n^{1/4} \log n) \) of \( E_{n,i} \). Likewise \( \| x_i \| \) can be absorbed in \( a_1 \). Hence, we can drop \( \Psi_k(\omega, b) \) from the inequality. Furthermore, using the bound for \( \sum_i (\omega_i - w_i) \) in Equation [14] it follows

\[
\| \beta(t_{i+1}) - \beta(t_{l+1}) \| \leq a_1 n^{-1} \left\{ \frac{1 - \epsilon}{\epsilon^2} d_{l,n} (1 + E_n \delta_n) + 
\right.
\]

\[
+ \| \Psi_{l+1} (w(t_{l+1}), \beta(t_{l+1})) \| + \| E_{n,l} \| + 2 \left\{ \sum_{j=1}^{q_n} p'_j(\theta) \right\}
\]

In the Theorem 3.1, it is defined that \( R_n = n^{-1/2} \max_k \left\{ \| \Psi_{k+1} (w_k, \beta(t_k)) \| + \| E_{n,k} \| \right\} \). As defined in Portnoy and Ling (2010) Equation 18, \( r_1 = a_1 \frac{1-\epsilon}{\epsilon^2} \). We can rewrite the inequality in terms of \( r_1 \) and \( R_n \)

\[
\| \beta(t_{i+1}) - \beta(t_{l+1}) \|
\]

\[
\leq r_1 n^{-1} d_{l,n} (1 + E_n \delta_n) + a_1 n^{-1} R_n \sqrt{n} + a_1 n^{-1} 2 \left\{ \sum_{j=1}^{q_n} p'_j(\theta) \right\}
\]

By the definition of \( d_{l,n} \), we know that \( d_{l,n} > R_n \sqrt{n} \). In addition, when \( \epsilon \leq 0.5 \) then \( r_1 > a_1 \), so we get the following equation.

\[
\| \beta(t_{i+1}) - \beta(t_{l+1}) \|
\]

\[
\leq r_1 n^{-1} d_{l,n} (1 + E_n \delta_n) + r_1 n^{-1} d_{l,n} + 2r_1 n^{-1} \left\{ \sum_{j=1}^{q_n} p'_j(\theta) \right\}
\]

Manipulating the equation to combine the terms, the following inequality holds true
\[ \| \beta(t_{i+1}) - \beta(t_{i+1}) \| \]

\[ \leq 2r_1 n^{-1} d_{l,n} \left( 1 + E_n \delta_n \right) + 2r_1 n^{-1} \left\| \sum_{j=1}^{q_n} p'_{\lambda}(\theta) \right\| \]

[18]

Next, we will look at how the penalty term can be combined in the first term of the inequality.

We know by definition \( p'_{\lambda}(\theta) = p'(\lambda_n, a) \) and \( p'(\lambda_n, a) < c\lambda_n \), where \( c \) is some positive constant. Then

\[ \sum_{j=1}^{q_n} p'_{\lambda}(\theta) < q_n c\lambda_n \]

By the hypothesis of the theorem, \( \lambda_n \) satisfies

\[ q_n c\lambda_n \leq c^*(d_{l,n} \delta_n) \]

for some constant \( c^* \) (since \( q \leq d \) and \( c \) and \( d \) are bounded). We can achieve an upper bound for \( \sum_{j=1}^{q_n} p'_{\lambda}(\theta) \) by

\[ \sum_{j=1}^{q_n} p'_{\lambda}(\theta) \leq c^*(d_{l,n} \delta_n) \]

Now, Equation [18] can be written as

\[ \| \beta(t_{i+1}) - \beta(t_{i+1}) \| \]

\[ \leq 2r_1 n^{-1} d_{l,n} \left( 1 + E_n \delta_n \right) + 2r_1 n^{-1} c^*(d_{l,n} \delta_n) \]
\[ \leq 2r_1 n^{-1} d_{l,n} \left( 1 + (E_n + c^*) \delta_n \right) \]

[19]

If we let \((E_n + c^*) = 2r_1 r_2 E_n^*\), then Equation [19] can be written as follows:

\[ \| \beta(t_{l+1}) - \beta(t_{l+1}) \| \]

\[ \leq 2r_1 n^{-1} d_{l,n} \left( 1 + 2r_1 r_2 E_n^* \delta_n \right) = 2r_1 n^{-1} d_{l+1,n} \]

Hence \(E_n^* = (E_n + c^*)/2r_1 r_2\).

Let's look at the upper bound for \(\sum_{i \in C_l+1} |\hat{\tau}_i - \tau_i|\). We can use the none-penalized results given in Portnoy and Lin (2010) Equation [19], because its calculation does not involve any penalty term, hence, their results holds true for the penalized quantile regression as well.

\[ \sum_{i \in C_l+1} |\hat{\tau}_i - \tau_i| \leq d_{l,n} + 2r_1 r_2 d_{l,n} \left( 1 + E_n \delta_n \right) \delta_n \]

\[ \leq d_{l,n} \left( 1 + 2r_1 r_2 E_n^* \delta_n \right) = d_{l+1,n} \]

[20]

where \(E_n^* = 1 + E_n \delta_n\), \(r_2 = \| x_i \| \frac{g_i(x_i' \beta(t_i))}{1 - g_i(x_i' \beta(t_i))}\).

Now, in order to satisfy both Equation [19] and [20] we define

\[ E_n^* = \max \left\{ \frac{(E_n + c^*)}{2r_1 r_2}, 1 + E_n \delta_n \right\} = O_p(1). \] [21]
Note, this result is derived from Portnoy and Lin (2010) Equation [20], where adding a constant to the first term will not change the overall asymptotic behavior.

To deal with asymptotic theory, let \( \hat{\beta}^* \) be the estimate based on the selected variables, and let \( \hat{\theta} \) be the estimate based on the true model. By consistency, \( \hat{\beta}^* = \hat{\theta} \) with probability tending to 1. Thus, \( |\hat{\beta}^* - \hat{\theta}| = O_p(a_n) \), where \( a_n \) is an arbitrary sequence tending to zero. Taking \( a_n \) tending to 0 faster than \( 1/\sqrt{n} \), the equivalence of asymptotic distribution for \( \hat{\beta}^* \) and \( \hat{\theta} \) follows from Slutsky’s theorem. Portnoy and Lin (2010) provide an asymptotic Gaussian representation for the process \( \sqrt{n} (\hat{\beta}^*(t) - \hat{\theta}(t)) \). Though there is no closed form expression for the asymptotic covariance matrix, this provides the following theorem covering the right-censored estimator and justifying the use of the bootstrap for asymptotic inference.

Theorem 3.2. Under the hypotheses of Theorem 3.1, for any \( \tau \) between \( \epsilon \) and \( \min\{\tau, 1 - \epsilon\} \),
\[
\sqrt{n} (\hat{\beta}^*(t) - \hat{\theta}(t)) \text{ converges in distribution to multivariate normal distribution with mean 0 and fixed covariance matrix.}
\]

For details on the proof and computation of the estimated covariance matrix see Portnoy and Lin (2010).
Chapter 4

Case Study

Ultrasound has assumed great importance in medical imaging and it is presumed to be safe at typical diagnostic levels. The safety of ultrasound depends on the level of exposure. Back in the 1980s, this level of exposure was much more conservative and fairly safe. But in 1991 the US government relaxed its regulation allowing the intensity level of ultrasound used to scan the in utero fetus to increase almost eight times over the level that had been allowed previously. The ultrasound (measured in time average-intensity) generated by equipment for obstetrics is about 1000-fold of the level back in the 1980s (Toms, 2007)! In addition, the use of ultrasound has expanded in the prenatal ultrasounds from the 2D to now 4D imaging (Merritt, 1989). So the natural question to ask is how safe is the ultrasound these days?

Ultrasound induced lung hemorrhage is one topic that has been studied extensively since the 90's (Zachary and O'Brien 1995; Baggs et al. 1996; Raeman et al. 1996; Dalecki et al. 1997a, 1997b; O'Brien and Zachary 1997; O'Brien et al. 1001a; Zachary et al. 2001). Different animal-based studies were performed in the interest of investigating the association between the occurrence and size of the US-induced lung hemorrhage, and the factors such as the US pressure, frequency, beam width species and age of animal. The logistic regression analysis and Gaussian-Tobit analysis were used. It was found that the ultrasound-induced lung hemorrhage in adult rats
strongly influenced by situ peak rarefaction pressure (acoustic pressure) and beam width and their interaction (O’Brien et al, 2001). In addition, ultrasound frequency was found to be not significant in the same experiment (O’Brien et. al, 2001). But it was determined using the logistic regression analysis and Gaussian-Tobit analysis methods on the pig data, the occurrence of the lesion is a function of age and the results are summarized in Table 4.1.

<table>
<thead>
<tr>
<th>Pig Age</th>
<th>Threshold Peak Rarefactional Pressure (MPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old</td>
<td>2.87 +/- 0.29</td>
</tr>
<tr>
<td>Middle</td>
<td>5.83 +/- 0.52</td>
</tr>
<tr>
<td>Neonate</td>
<td>3.60 +/- 0.44</td>
</tr>
</tbody>
</table>

Table 4.1: Summary of pig age versus threshold peak rarefactional pressure in the Gaussian-Tobit Analysis of pig data (O’Brien et. al, 2003).

It can be seen in Table 4.1, the lesion threshold is greatly affected by the physiology of the lung (O’Brien et. al, pig 2003). In conclusion, structural differences among mammalian species studied are independent of the biological mechanism that causes ultrasound-induced lung damage; therefore, it is reasonable to speculate the mechanisms that cause lung hemorrhage in laboratory animals may also damage human beings especially small children and patients with pulmonary disorders. It is extremely important to be able to estimate thresholds of mechanisms that cause ultrasound-induced damage (O’Brien et. al, 2007).

The pig data used in this example is the same as the data set used in O’Brien paper (2003). It is a high-dimensional data set with 4 animal related variables 7 instrument-setting variables. The variables are defined and listed below.
Animal related variables:

species

animal = id of the animal

cHEST WALL = thickness of chest wall to lung in mm

cHEST ATTEN = attenuation coefficient in dB/mm

Instrument settings:

freq = ultrasound frequency in MHz

pd = pulse duration in mu-sec

prf = pulse repetition frequency

ed = exposure duration

beam = width of beam at focus

pr = calculated in vivo acoustic pressure at the lung surface

mi = mechanical index

Variable mi is closely related to pr, hence it is not necessary to include both variables. However, out of curiosity we did keep mi in the bidirectional regression analysis.

We begin our analysis using Portnoy’s random censoring estimator. Using crq.fit.por algorithm, the information needed to calculate the weights of each observation can be retrieved as described in Chapter 2. Using these weights we can transform a censored quantile regression problem into a weighted quantile regression problem. These weights are used in the weighted regression part of the LASSO variable screening method. The coefficients estimates generated from the LASSO can be plotted and used for visual diagnostics.
4.1 Penalized Model Screening - LASSO

We begin our analysis by first testing Powell’s, Portnoy’s, and Peng and Huang’s methods on a much smaller data set called the neonates. Our goal was to become familiar with the RQ package through replicating the analysis that was previously done by Simpson et. al using the Tobit method (2004). Unfortunately, we had experienced many problems with the Powell’s estimator. We believe these problems are related to the selection of the starting values (Portnoy, 2010). As seen in Figure 4.1, quantiles generated using the Powell’s estimator are crossed and varies drastically with the starting value.

![Figure 4.1](image)

Figure 4.1: Plot of Powell predictions where $y = \text{lesion depth}$ and $x = \text{threshold wavelength}$ for two different starting value.
Portnoy’s and Peng and Huang’s estimator were both successfully computed. These estimates were very close to the estimate of Tobit’s method. Figure 4.2 is a plot of the prediction of lesion depth computed using the Portnoy’s and Peng and Huang’s and Tobit’s methods.

Figure 4.2: Prediction Made Using Portnoy’s and Peng and Huang’s estimates, where green and yellow lines are Portnoy’s prediction, blue and red are Peng and Huang’s estimates, and Orange and black are Tobit’s estimate.
The data set used for the rest of our study is the pig lung data set that was once studied by O’Brian et. al (2003). It is much larger than the neonate data set mentioned above, and consists of 11 variables and 2483 entries. There are 3 response variable: lesion, depth, and area. The variable lesion is the censoring variable that takes on value of 0 and 1, where 1 means lesion depth is greater than 0, and 0 means no lesion. There were precisely 3 observations with either 0 lesion but a depth greater than 0 or the opposite scenario. These points were corrected via R by changing the value of lesion to match the depth. The response variable area measures the surface area of the lesion. Since it serves the same purpose as depth, which both measure the ultrasound-induced hemorrhage, only depth was used in our analysis. Furthermore, we noticed the observed values of variables vary significantly in magnitude. To avoid any magnitude effects on the estimation calculation, these variables were standardized using mean and standard deviation before performing any analysis,

The preliminary analysis begins with analyzing the full model consists of all variables minus variables species and animal. Intuitively, these two variables incorporate information that pertains to specific species and animals. It is our goal to develop a generalized model that can be used for any species and animals. Hence, we hope to find some combination of the variables other than species and animal that could help explain the lesion depth. In addition, including these two variables in the model will cause the crq program to fail. It is likely due to singularity of the X matrix. After a model is fitted using crq.por.tsp, we can then retrieve the Isplit and T.

We begin our analysis at $\tau = 0.5$. The median is a good starting point since the most abundant information is generally at the median; therefore, the analysis would potentially give us the most accurate model. Note, $\tau$ is used in the calculation of the split weights at the crossed censored observations. These adjusted weights help redistribute the probability to the unknown
true responses as well as the censored responses. Then LASSO method is used to generate coefficients for the weighted problem. A sequence of $\lambda$'s was examined. For each $\lambda$ a set of $\beta$ was recorded. Figure 4.3 and Figure 4.4 are plots of $\lambda$ versus $\beta$. In Figure 4.3, the coefficients were calculated based on the standardized variables.

Figure 4.3: Plot of $\lambda$ versus coefficients of standardized variables.

Figure 4.4 is the plot of $\lambda$ versus coefficients of original variables for none standardized data. As expected, the shape of the LASSO tree is similar, but the variables are separated better, where the threshold $\lambda$ is slightly smaller for the standardized data. In addition, there is less tail affect, hence easier to decipher the more significant variables.
Figure 4.4: Plot of $\lambda$ versus coefficients of original variables (none standardized).

From the above LASSO tree plots, it is noticeable that the estimates calculated based on the standardized variables actually shrink to 0 at a different rate than the none standardized. The calculation based on the original data yields the most significant 7 variables to be (in the order of importance) as follows: $Pr, ed, chest.wall, prf, mi, pd, study$. Note: beam is not listed. Although this variable remained in the model for all the $\lambda$’s, its estimated coefficient is consistently very small, relatively speaking. When all the variables are standardized, the order of importance is listed as following: $pr, chest.wall, prf, mi, freq, pd$, and $ed$. When the variables are standardized, the variable beam shrinks to 0 much quicker than before. One question arises, is the magnitude of the estimates influential in variable selection, in the since that does the shrinking rate overrules the magnitude difference? Intuitively it seems likely a variable with coefficient estimates that is small does not have strong influence on the response. Even though a variable
stays in the model for all $\lambda$’s, this is not an indication that this variable is more important than others. In addition, LASSO is not built to detect interaction effects. But one may be able to pick up some signs of possible interactions by looking at the variable shrinking behavior. If a variable initially shrinks to 0 quickly but later reenters the model when some other variable falls out, then this could be an indication that interaction may exist between these variables. Finally, the selected base model using standardized variables with no two-way interaction is present in R language as follows:

```r
crq(Surv(dept h, lesion) ~ pr+chest.wall+prf+mi+freq+ed+pd, method="P0r", data = Dataset).
```

Only 7 variables were selected because the 8th significant variable is `study`, which similar to `species` and `animals`, it incorporates information of a combination of instruments used during the experiment. It is our goal to generalize the model so that it can apply to any combination of instrument settings. Therefore, we would like to remove experiment specific variables such as `study`, and hence, replace it with more important instrument variables.

We have performed the LASSO variable screening method on the pig lung data set. The SCAD method is another screening method that we can easily access in R. We will compare the results from these two methods later in the paper.

### 4.2 Extension of Portnoy Estimator – Inference Based Refinement of Screened Variables

We begin by examining the confidence plot of the `crq` fit shown in Figure 4.5. Please note the x-axis is $1-\tau$ not $\tau$. In our following discussion, $\tau$ is always the left censoring quantile.
Clearly the estimates of $freq$, $mi$, and $ed$ all cross 0 for significant part of the quantiles spectrum. Variables $freq$ and $mi$ seems to be insignificant for large $\tau$’s, and $ed$ is insignificant for small and large $\tau$’s. We will investigate to see if this is due to interaction between the variables. Note, $ed$ should be kept as an main affects since it is significant for $\tau = 0.5$. To investigate for possible two-way interaction of $freq$ and $mi$ with other terms, we examine the following model

$$\text{crq(Surv(-dept h, lesion) ~ (pr+chest.wall+prf+ed+pd)*(mi+freq), method="Por", data = Dataset )}.$$  

Figure 4.6 is the plot of the two-way interaction model. From the plot we can see the confidence band for $pr$, $chest.wall$, $prf$, $ed$, and $pd$ have changed significantly.
Figure 4.6: Confidence band of \((pr+\text{chest.wall}+prf+ed+pd)*(mi+freq)\), where x-axis is 1-\(\tau\) and y-axis is the coefficient.

Estimates of \(\text{chest.wall}, prf, ed,\) and \(pd\) are now mostly insignificant. And \(mi\) and \(freq\) are now insignificant for almost the entire spectrum of \(\tau\). What’s interesting is the interaction of \(\text{chest.wall} * mi\) is significant and \(\text{chest.wall} * freq\) interaction is also significant but only for large \(\tau\)’s. Furthermore, the addition of 2-way interactions removed all the effects of \(ed\). Note, \(1/frequency = wavelength\). Wavelength can be associated with acoustic pressure, \(pr\), at the lung
Although, \textit{freq} is not significant in the above model, the product of the wavelength and the acoustic pressure may hold potential to reveal some interesting relationships. We define \textit{rw_pr} to be the product of wavelength and acoustic pressure. In addition, \textit{prf} * \textit{ed} is the acoustic pressure pulse, \textit{pulse}, which could also be important in explaining the response. Next we examine the model

\begin{verbatim}
crq(Surv(-dept, lesion) ~ pr+pulse+pr*mi+chest.wall*mi+chest.wall*rw pr+pd*mi+prf+ed, method="Por", data = Dataset).
\end{verbatim}

Figure 4.7 is the plot of the confident bands for the above model with the transformed variables and additional two-way interactions.

Figure 4.7: Confidence band of \textit{pr}+\textit{pulse}+\textit{pr}*\textit{mi}+\textit{chest.wall}*\textit{mi}+\textit{chest.wall}*\textit{rw pr}+\textit{pd}*\textit{mi}+\textit{prf}+\textit{ed}, where \textit{x}-axis is 1-\tau and \textit{y}-axis is the coefficient.
From the plot you can see the interaction term \( rw_{\text{pr}} \times chest\text{.wall} \) and the main affect term \( pulse \) are mostly insignificant. Removing the other 2 terms and keeping \( rw_{\text{pr}} \) by itself, we reduce down to the following model:

\[
\text{crq( Surv(-dept, lesion) ~ pr+rw_{\text{pr}}+pr*mi+chest\text{.wall}*mi+pd*mi+prf+ed, }
\text{method="Por", data = Dataset ).}
\]

The results are plotted in Figure 4.8. It should be noted the confidence band widen for \( pd, prf, \) and \( ed \), in particularly for small \( \tau \)’s, but the significance of the variables remained the same.

Figure 4.8: Confidence band of \( rw_{\text{pr}}+pr*mi+chest\text{.wall}*mi+pd*mi+prf+ed \), where \( x \)-axis is \( 1-\tau \) and \( y \)-axis is the coefficient.
Removing the $rw_pr\ast chest.wall$ did not penalize the model significantly. The significant variables remain significant with small changes in their P-value. More detail is given in Table 4.2. In addition, the estimate for $rw_pr$ crosses 0 for large $\tau$’s. Its P-value is given in the following table. In addition, $ed$ and $mi \ast pd$ are still insignificant for almost the entire spectrum of $\tau$. Its unlikely these two variables will appear in the final model. Furthermore, the interaction term $mi\ast pd$ are mostly insignificant but it is highly significant as main effects; therefore, it is necessary to include $pd$ and $mi$ as a main effects in the final model.

| Coefficients | Value   | Lower Bd | Upper Bd | Std Error | T Value | Pr($>|t|$) |
|--------------|---------|----------|----------|-----------|---------|------------|
| (Intercept)  | 0.47969 | 0.22555  | 0.75424  | 0.12977   | -3.6902 | 0.00022    |
| $rw_pr$      | 0.38123 | -0.73529 | 1.49775  | 0.56966   | -0.6962 | 0.50356    |
| $pr$         | 0.81921 | 0.07271  | 1.56570  | 0.38087   | -2.1508 | 0.03149    |
| $mi$         | 0.21030 | -0.26638 | 0.68717  | 0.24331   | -0.8643 | 0.38741    |
| $chest.wall$ | 0.32510 | 0.14511  | 0.50509  | 0.09183   | -3.5401 | 0.00040    |
| $pd$         | 0.34012 | 0.27244  | 0.40780  | 0.03453   | -9.8501 | 0.00000    |
| $prf$        | 0.00065 | 0.00002  | 0.00008  | 0.00012   | -5.5531 | 0.00000    |
| $ed$         | 0.06961 | 0.00593  | 0.13529  | 0.03249   | -2.1437 | 0.03216    |
| $pr:mi$      | -0.25356| -0.31667 | -0.19044 | 0.03220   | 7.87429 | 0.00000    |
| $mi:chest.wall$ | 0.47641 | 0.27662  | 0.67621  | 0.10194   | -4.6733 | 0.00000    |
| $mi:pd$      | 0.29325 | 0.09950  | 0.48695  | 0.09884   | -2.9668 | 0.00301    |

Table 4.2: Summary table of estimates for $pr + pr\ast mi + chest.wall\ast mi + pd\ast mi + prf + ed$. 
Table 4.2 (cont.): Summary table of estimates for $pr + pr \times mi + chest\_wall \times mi + pd \times mi + prf + ed$.

Table 4.2 shows the P-value for each variable. Its shown $rw\_pr$ has P-value of 0.50 at $\tau = 0.8$, and becomes more significant as $\tau$ decreases. Since we are more interested in high $\tau$ values the variable $rw\_pr$ may be removed. In addition, at $\tau=0.8$, $ed$ has a P-value of 0.032 and its P-value increases as $\tau$ decreases. Even at its prime of $\tau=0.8$, $ed$ is not highly significant, it may be beneficial to remove this variable from the model. Removing these terms will further reduce the model down to

```r
crq(Surv(death, lesion) ~ pr*mi+chest\_wall*mi+pd+prf, method="Por", data = Dataset).
```
The result of this reduced model is plotted in Figure 4.9. The confidence band changes are quite noticeable for most variables. Variables $pr$, $mi$, $pd$, and $pri:mi$ show most improvement. In particularly, the bands are narrower than before for large $\tau$ values.

![Figure 4.9: Confidence band of $pr*mi+chest.wall*mi+pd+prf$, where $x$-axis is $1-\tau$ and $y$-axis is the coefficient.](image)

Table 4.3 tabular summary of the estimates for $pr*mi+chest.wall*mi+pd+prf$. The variables are most significant for $1-\tau = 0.8$. $Chest.wall$ becomes insignificant for smaller $1-\tau$ values. Since its interaction term is significant, it will remain in the model. Note, $pr*mi$ shows signs of decrease in significant for $1-\tau = 0.2$. 

Table 4.3: Summary table of estimates for \( pr \times mi + \text{chest.wall} \times mi + pd + prf \).

| Coefficients:          | Value | Lower Bd | Upper Bd | Std Error | T Value | Pr(>|t|) |
|------------------------|-------|----------|----------|-----------|---------|---------|
| (Intercept)            | 0.40406 | 0.17300 | 0.63511 | 0.11789 | -3.42745 | 0.00061 |
| pr                     | 1.01958 | 0.83595 | 1.20321 | 0.09369 | -10.88252 | 0.00000 |
| mi                     | 0.33371 | 0.13569 | 0.53173 | 0.10103 | -3.30295 | 0.00096 |
| chest.wall             | 0.23456 | 0.07793 | 0.39119 | 0.07992 | -2.93314 | 0.00333 |
| pd                     | 0.33880 | 0.28873 | 0.38886 | 0.02554 | -13.26293 | 0.00000 |
| prf                    | 0.00075 | 0.00053 | 0.00097 | 0.00011 | -6.56728 | 0.00000 |
| pr:mi                  | -0.24795 | -0.32841 | -0.16748 | 0.04105 | 6.03958 | 0.00000 |
| mi:chest.wall          | 0.51335 | 0.34880 | 0.67791 | 0.08396 | -6.11442 | 0.00000 |

| Coefficients:          | Value | Lower Bd | Upper Bd | Std Error | T Value | Pr(>|t|) |
|------------------------|-------|----------|----------|-----------|---------|---------|
| (Intercept)            | -0.53747 | -0.84137 | -0.23357 | 0.15505 | 3.46635 | 0.00053 |
| pr                     | 1.26943 | 1.08881 | 1.45005 | 0.09216 | -13.77483 | 0.00000 |
| mi                     | 0.48017 | 0.27181 | 0.58854 | 0.04364 | -5.32809 | 0.00000 |
| chest.wall             | 0.13000 | 0.05488 | 0.31487 | 0.09432 | -1.37816 | 0.16815 |
| pd                     | 0.41962 | 0.32975 | 0.50949 | 0.04585 | -9.15135 | 0.00000 |
| prf                    | 0.00081 | 0.00055 | 0.00107 | 0.00013 | -6.12152 | 0.00000 |
| pr:mi                  | -0.30342 | -0.37309 | -0.23375 | 0.03555 | 8.53359 | 0.00000 |
| mi:chest.wall          | 0.47273 | 0.31947 | 0.62598 | 0.07819 | -6.04570 | 0.00000 |

| Coefficients:          | Value | Lower Bd | Upper Bd | Std Error | T Value | Pr(>|t|) |
|------------------------|-------|----------|----------|-----------|---------|---------|
| (Intercept)            | -1.62459 | -1.98025 | -1.26893 | 0.18146 | 8.95277 | 0.00000 |
| pr                     | 1.66192 | 1.40811 | 1.91572 | 0.12949 | -12.83395 | 0.00000 |
| mi                     | 0.56543 | 0.36034 | 0.77051 | 0.10464 | -5.40361 | 0.00000 |
| chest.wall             | -0.27925 | -0.72847 | 0.16997 | 0.22920 | 1.21838 | 0.23208 |
| pd                     | 0.44943 | 0.32410 | 0.57475 | 0.06394 | -7.02864 | 0.00000 |
| prf                    | 0.00099 | 0.00076 | 0.00123 | 0.00012 | -8.29791 | 0.00000 |
| pr:mi                  | -0.38068 | -0.48340 | -0.27797 | 0.05240 | 7.26427 | 0.00000 |
| mi:chest.wall          | 0.73032 | 0.40954 | 1.05109 | 0.16367 | -4.46224 | 0.00001 |

| Coefficients:          | Value | Lower Bd | Upper Bd | Std Error | T Value | Pr(>|t|) |
|------------------------|-------|----------|----------|-----------|---------|---------|
| (Intercept)            | -5.32805 | -6.80172 | -3.85438 | 0.75188 | 7.08626 | 0.00000 |
| pr                     | 2.47223 | 1.60854 | 3.33952 | 0.44066 | -5.61023 | 0.00000 |
| mi                     | 1.56233 | 0.98716 | 2.13751 | 0.29346 | -5.32379 | 0.00000 |
| chest.wall             | -1.87491 | -2.62396 | -1.12566 | 0.38218 | 4.90587 | 0.00000 |
| pd                     | 0.43567 | 0.16275 | 0.70860 | 0.13925 | -3.28722 | 0.00176 |
| prf                    | 0.00210 | 0.00139 | 0.00281 | 0.00036 | -5.79847 | 0.00000 |
| pr:mi                  | -0.56125 | -0.94170 | -0.18080 | 0.19411 | 2.89142 | 0.00384 |
| mi:chest.wall          | 1.76093 | 1.34541 | 2.17645 | 0.21200 | -8.30615 | 0.00000 |
Look at Table 4.3, it is clear each variable in this new model is highly significant with an exception of chest.wall. But since mi*chest.wall is highly significant, chest.wall will remain in the final model.

The 7 variable model is reduced down to pr*mi+chest.wall*mi+prf+pd. To see if this model has the best fit, we compare to a slightly larger model. We chose to introduce the variable beam to the model because it is the 8th significant variable screened using LASSO.

Figure 4.10: Confidence band of pr*mi+chest.wall*mi+pd+prf+beam, where x-axis is 1-τ and y-axis is the coefficient.
Table 4.4: Summary table of estimates for $pr*mi+chest.wall*mi+pd+prf+beam$.
It is easy to see in Figure 4.10 that beam is highly significant for all τ’s except when near 1. Table 4.4 shows even at tail regions beam is still highly significant. The two-way interactions of beam were tested but none were significant. In addition, 9th significant variable, chest.attent, was also added to the model but it crashed crq program. It is possible that this may be the result of singularities. Finally, I conclude the best-fitted model has the following variables:

\[ pr*mi, chest.wall*mi, pd, prf, beam. \]

It is interesting to note, that one way to examine the validity of a model is by looking at the number of quantiles can be fitted. A "bad" model is unable to generate coefficient estimates for a large spectrum of τ’s. The model selected above was able to fit τ ranging from 0.88 to 0.01 in increments of 0.01. It may even be possible to fit τ’s less than 0.01. Hence we are confident, our final model is one of the best-fitted model for this data set.

### 4.3 Penalized Model Screening - SCAD

Similar to LASSO, the SCAD variable screening method is also build into the RQ package. The weights are generated the same way as for the LASSO method in Section 4.2. The weighted response and variables are then feed into SCAD as we’ve done for LASSO. The results are surprisingly very different than the LASSO findings.

Figure 4.11 is the plot of λ versus coefficient estimates based on the standardized variables. Just from the first glance, we notice the behavior of the estimates is noticeably different than LASSO.
The variables drop off to 0 where in LASSO they asymptotically approach 0. The advantage of this clear-cut result allows the variable selection to be less subjective. Furthermore, variables are essentially split into two groups at $\lambda = 300$: 1) mi, pr, pd, and chest.attend that never approach 0, therefore, remain highly significant for all $\lambda$’s, 2) variables either already shrunk to 0 before $\lambda$ reach 300 or remain very close to 0 for $\lambda > 300$.

![Figure 4.11: Plot of $\lambda$ versus coefficients of the standardized variables for SCAD.](image)

Next, we discuss the difference of the base model chosen via SCAD vs. LASSO. Surprisingly, chest.attend is most significant for SCAD method. Figure 4.12 is the plot of confidence band of the variables estimates for the following base model

```r
crq(Surv(dept h, lesion) ~ mi + pr + pd + chest.attend, method = "Por", data = Dataset)
```
The chest.attenu is highly significant for the most of $\tau$’s in SCAD results, especially the large $\tau$ ‘s of interest. This is a strong contrast to the LASSO findings. If you recall, chest.attenu was added to the LASSO final model as a check for the best fit. When it was included in the final model $pr*mi$, chest.wall*mi, $pd$, prf, beam the crq program crashed. It appears the variable chest.wall with chest.attenu may have caused the singularity in the X matrix. These two variables are both animal variables measuring characteristics of the animal lung. Hence, it is possible they are correlated.

Although, SCAD method generated 4 distinctively significant variables, for the sake of
comparison let’s be consistent with the LASSO and focus on the 7 significant variables. These variables are: chest.att, pr, mi, pd, beam, prf, and ed (in the order of significance). If you recall, we initially selected the following variables using LASSO method: pr, chest.wall, prf, mi, freq, pd, and ed (in the order of significance). The differences are chest.att replaced the chest.wall and beam replaced freq in the SCAD results. Figure 4.13 is the plot of chest.att histogram. Notice chest.att takes on rather discrete values; this may also be a reason why LASSO crashed. We investigated this possibility by tethering $\lambda$ with a very small random number. But this did not solve the crash of LASSO. Hence, evidence points towards the possibility of correlation.

Figure 4.13: Histogram of chest.att.
4.4 A Closer Look at Tail Quantiles

The confidence band figures of the fitted model have occasionally shown spikes near the extreme \( \tau \) areas. The probability reflects the fact that estimates near the endpoints tend to be much more variable since they are based on very little data. Especially uncensored data may be very scarce in these regions. Near the endpoints, standard asymptotic normality fails and needs to be replaced by some form of extreme value theory. In the lung data set, some variables exhibited more discrete characteristics than others. Next we will examine if this contributed to the irregularity of the confidence bands. As seen in Figure 4.14, \textit{freq}, \textit{chest.attn} and \textit{beam} have spikes in their confidence band.

![Figure 4.14: Example of widening of Confidence Band.](image-url)
To lessen the effects of the discrete characteristics of the above variables, these variables were dithered with a small random variable, $0.001*\text{uniform}(0,1)$. Unfortunately, this change did not help with the removal of the spikes. On another note, the spikes may also be an indication of the inappropriate modeling at the $\tau$'s where spikes occurred, hence, the selection of the relevant variable may vary significantly at these $\tau$'s when compared to, for example, the median. A sequence of $\tau$ from 0.5 to 0.99 was examined. At each $\tau$, LASSO and SCAD were performed and the variable screening results are summarized in a tabular form shown below. Note, $\tau$ is the left censoring quantile.

<table>
<thead>
<tr>
<th>$\tau$ (lasso)</th>
<th>$\tau$ (Scad)</th>
<th>chest.wall</th>
<th>chest.attenu</th>
<th>ed</th>
<th>beam</th>
<th>mi</th>
<th>freq</th>
<th>prf</th>
<th>pr</th>
<th>pd</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>0.8</td>
<td>0.8</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>0.9</td>
<td>0.9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>0.95</td>
<td>0.95</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>0.98</td>
<td>0.98</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>0.99</td>
<td>0.99</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>0.99</td>
<td>0.99</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Table 4.5: Summary table of LASSO and SCAD screening results for a sequence of $\tau$'s.

In Table 4.5, the blue-highlighted rows contain SCAD results and the white rows contain LASSO results. Although, LASSO and SCAD results are not always in agreement, variables $\text{chest.wall}$, $\text{pr}$, $\text{pd}$ and $\text{beam}$ (with exception of $\tau = 0.99$) consistently shown to be significant by LASSO and SCAD at all $\tau$'s. Furthermore, $\text{mi}$ appears to be insignificant for $\tau < 0.8$, $\text{chest.attenu}$ appears to be significant for $\tau > 0.9$, and $\text{prf}$ appears to be significant for $\tau < 0.95$. Note, in
Figure 4.14, it was observed the spikes to have occurred approximately at $1 > \tau > 0.75$. The results in Table 4.5 agrees with our observation that the variable selection for large $\tau$’s, in this case $\tau > 0.8$, is different than $\tau = 0.5$. It is clear, for the lung dataset, the best-fit model for large $\tau$’s will be different than for the median.

### 4.5 $\lambda$ Selection

Selecting the appropriate $\lambda$ is vital for the success of the variable screening method. But how do we decide what $\lambda$ to use? First, let’s define threshold $\lambda$ as the $\lambda$ that marks the natural break of the variables in the LASSO and SCAD tree plot. To begin answering this question we must understand how this $\lambda$ is related to the dataset. It seems logical that this threshold $\lambda$ value is correlated with the characteristics of the variables that are being considered. As we’ve discussed in Chapter 2, significant differences in the magnitude of the variables will affects the results of LASSO and SCAD. Hence, if in the dataset some variable are significantly larger in magnitude than the rest, then it is necessary to standardize all variables prior to running LASSO and SCAD. In addition, after running numerous runs with difference combination of $\lambda$ and $\tau$, we have observed a unique relationship between $\lambda$ and $\tau$. LASSO and SACD were performed on the lung dataset at $\tau = 0.5, 0.9, 0.95, 0.98$, and $0.99$. The threshold $\lambda$ were recorded and summarized in the Table 4.6.

<table>
<thead>
<tr>
<th>Method/Threshold $\lambda$</th>
<th>$\tau=0.5$</th>
<th>$\tau=0.90$</th>
<th>$\tau=0.95$</th>
<th>$\tau=0.98$</th>
<th>$\tau=0.99$</th>
</tr>
</thead>
<tbody>
<tr>
<td>LASSO</td>
<td>181.02</td>
<td>128</td>
<td>90.51</td>
<td>38</td>
<td>13.45</td>
</tr>
<tr>
<td>SCAD</td>
<td>13.45</td>
<td>29.18</td>
<td>128</td>
<td>22.63</td>
<td>11.31</td>
</tr>
</tbody>
</table>

Table 4.6: Summary of $\lambda$ vs. $\tau$. 
As expected with the LASSO method, the natural breaks between the significant and insignificant variable is not as clear-cut as with SCAD method. Hence, the determination of the threshold $\lambda$ is slightly more subjective. Visualization of Table 4.6 information is given in Figure 4.15. Interestingly the threshold $\lambda$ produced using LASSO shows an inverse relationship between $\lambda$ and $\tau$. In other words, one need to penalized less at the higher quantile in order to separate the significant from the insignificant variables. This linear trend makes the $\lambda$ selection quite simple. Any $\lambda$ greater than or equal to the median threshold $\lambda$ will be sufficiently large to handle the job. But the SCAD method is more complicated in the since that it produced a parabolic plot. The appropriate $\lambda$ will have to be larger than the maximum threshold $\lambda$. Nevertheless, it is safe to say LASSO and SCAD each have a unique $\lambda$ and $\tau$ relationship.

![Graph showing $\lambda$ vs. $\tau$](image)

Figure 4.15: Summary of $\lambda$ vs. $\tau$.

### 4.6 Two-Way Interactions

Interesting shrinking behavior was observed in the LASSO tree for the lung data set. At $\tau=0.98$, \textit{beam} disappeared at the threshold $\lambda$, but reentered the model briefly when \textit{ed} dropped out. This was observed in both LASSO and SCAD tree. Which suggested that perhaps without \textit{ed}
influence, the significance of a beam was exposed. Hence, an interaction exists between the two. In addition, prf also returned very briefly but no variables were dropped out of the model during that time. This difference between prf and beam phenomena suggests the prf incident may be due to some glitch in the crq calculation rather than interaction. But the former requires further investigation to verify this interaction theory. As mentioned in the previous sections, pr and mi are essentially explaining the same mechanical information. Therefore, the variable mi was not included in the initial LASSO screening, nor was it part of the inference model selection for the backwards method. We begin by looking at the 2-way interaction terms for all 9 variables to get a full picture of the possible relationships. The most straightforward method is to include all the main effects and interaction terms in one large model and feed it to LASSO. The resulting plot or table would be difficult to analyze since plotting a 36 variable model on one single plot is quite messy. So, instead of one plot of all the variables and their 2-way intersections, the results were summarized in 9 individual tables, one table for every main variable and its interactions with the 8 remaining variables. Note, only the LASSO results are discussed here.

<table>
<thead>
<tr>
<th>Main Effect</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chest.wall</td>
<td></td>
</tr>
</tbody>
</table>
| Chest.attcn | Chest.attcn:Freq  
              Chest.attcn:Prf |
| Freq        | Chest.attcn:Freq  
              Pr:Freq |
| Pd          | Pd:Prf       |
| Prf         | Pd:Prf       
              Chest.attcn:Prf  
              Pr:Prf |
| Ed          |             |
| Beam        | Beam:Pr      |
| Pr          |             |

Table 4.7: Summary table of 2-way interaction for LASSO.
Table 4.7 is a record of the interaction terms that reduced to 0 later than its corresponding main effects variable. The idea is to keep the interaction term that is more influential than its main effects. Since no interactions of \( \text{ed} \) is considered, it will be included in the model as a main effects. Note, in the \( \text{chest.wall} \) 2-way interaction table, \( \text{Chest.wall*pr} \) reduced to 0 early, but reentered when \( \text{chest.wall*ed} \) dropped out. We will look at the 3-way interaction \( \text{Chest.wall*ed*pr} \) in more detail. The complete 2-way interaction tables can be found in Appendix I. We begin with the full model

\[
pd*prf+\text{chest.atten*freq}+\text{chest.atten*prf}+pr*freq+prf+\text{beam*pr}+\text{chest.wall*ed*pr}.
\]

| Coefficients:       | Value | Lower  | Upper  | Std   | T     | Pr(>|t|) |
|---------------------|-------|--------|--------|-------|-------|---------|
| (Intercept)         | -0.1102 | -1.4670 | 1.2466 | 0.6923 | 0.1592 | 0.8735  |
| pd                  | 0.1596  | -0.2205 | 0.5396 | 0.1939 | -0.8228 | 0.4106  |
| prf                 | 0.0000  | -0.0015 | 0.0015 | 0.0008 | 0.0019 | 0.9985  |
| chest.atten         | -1.1106 | -2.8236 | 0.6025 | 0.8740 | 1.2706 | 0.2039  |
| freq                | 0.2133  | -0.2704 | 0.6970 | 0.2468 | -0.8643 | 0.3874  |
| pr                  | 0.6110  | 0.2043  | 1.0178 | 0.2075 | -2.9444 | 0.0032  |
| beam                | 0.2056  | 0.1263  | 0.2848 | 0.0404 | -5.0846 | 0.0000  |
| chest.wall          | 0.5095  | 0.2078  | 0.8112 | 0.1539 | -3.3103 | 0.0009  |
| ed                  | 0.3026  | 0.0362  | 0.5690 | 0.1359 | -2.2266 | 0.0260  |
| pd:prf              | 0.0002  | -0.0001 | 0.0004 | 0.0001 | -1.2650 | 0.2059  |
| chest.atten:freq    | -0.2800 | -0.8616 | 0.3016 | 0.2967 | 0.9437 | 0.3453  |
| prf:chest.atten     | -0.0011 | -0.0016 | 0.0038 | 0.0014 | -0.8050 | 0.4208  |
| freq:pr             | 0.0923  | -0.0757 | 0.2603 | 0.0857 | -1.0772 | 0.2814  |
| prf:pr              | 0.0003  | 0.0000  | 0.0007 | 0.0002 | -1.9508 | 0.0511  |
| pr:beam             | 0.1030  | 0.0166  | 0.1894 | 0.0441 | -2.3377 | 0.0194  |
| pr:chest.wall       | 0.2223  | -0.0747 | 0.5194 | 0.1516 | -1.4669 | 0.1424  |
| chest.wall:ed       | 0.4380  | -0.3319 | 1.2078 | 0.3928 | -1.1150 | 0.2648  |
| pr:ed               | -0.2043 | -1.0887 | 0.6801 | 0.4512 | 0.4527 | 0.6507  |
| pr:chest.wall:ed    | -0.5950 | -1.8940 | 0.7040 | 0.6628 | 0.8977 | 0.3693  |

Table 4.8: Crq results of the full model at \( \tau = 0.5 \), and R=10.
The model selection method involves removing of the highest order of insignificant interaction term one at a time until the model is left with all significant variables. Note, when an interaction term is dropped, the corresponding main effects are reintroduced into the model. As seen in Table 4.8, the 3-way interaction is insignificant; therefore, it is the first variable dropped from the model. Note, $R$ is the number of bootstraps. Larger the $R$, more accurate the results.

Unfortunately we were unable to use a larger $R$-value for some of the more complex models, because many times the program would enter an infinite loop. Even with $R=10$, this problem occurred occasionally. I’ve noticed, reordering the variables in the model sometimes help with this problem. The best model determined using backward model selection method is

$$Pd + beam + prf + chest.wall:pr$$

If you recall we excluded the variable $mi$ from backward analysis because it contained similar information as $pr$. However, in the bidirectional study we did include the $mi$ variable and surprising found $pr:mi$ to be significant. This indicates it is possible the quadratic term $pr^2$ is significant. Hence, adding $pr^2$ to the above model and apply backward regression analysis, we found the best model to be

$$Pd + beam + prf + chest.wall + pr^2$$

If you recall we excluded the variable $mi$ from backward analysis because it contained similar information as $pr$. However, in the bidirectional study we did include the $mi$ variable and surprising found $pr:mi$ to be significant. This indicates it is possible the quadratic term $pr^2$ is significant. Hence, adding $pr^2$ to the above model and apply backward regression analysis, we found the best model to be

$$Pd + beam + prf + chest.wall + pr^2$$

| Coefficients: | Value   | Lower  | Upper  | Std   | T       | Pr(>|t|) |
|---------------|---------|--------|--------|-------|---------|---------|
| (Intercept)   | -1.3497 | -1.6863| -1.0130| 0.1718| 7.8571  | 0.0000  |
| pr            | 0.0009  | 0.0007 | 0.0011 | 0.0001| -9.7805 | 0.0000  |
| pr2           | -2.2028 | -2.5674| -1.8381| 0.1861| 11.8391 | 0.0000  |
| beam          | 0.2085  | 0.1204 | 0.2966 | 0.0450| -4.6385 | 0.0000  |
| pd            | 0.5212  | 0.4510 | 0.5915 | 0.0358| -14.5507| 0.0000  |
| chest.wall    | 0.5130  | 0.2983 | 0.7276 | 0.1095| -4.6831 | 0.0000  |
| pr            | 3.9546  | 3.4425 | 4.4668 | 0.2613| -15.1338| 0.0000  |

Table 4.9: Crq results of the final model with quadratic term $pr^2$ at $\tau = 0.5$, and $R=100$. 

Table 4.9 summarizes the results for the best model, and the variables are all highly significant. Note, the results were generated from 100 bootstrap samples. All the estimates are positive except for pr2. This suggests pr has a positive effect on lesion depth until a turning point is reached. This turning point is located at $3.95/(2*2.20)=0.89$. When the standardized pr is greater than 0.89, pr has negative impact on depth.

Figure 4.16 is the plot of confidence band for the final model. The backward results are very similar to the bidirectional results in Figure 4.10. Hence, we are confident the final model is a good fit for the data set.

![Confidence Plot](image)

Figure 4.16: Confidence plot of the final model at $\tau = 0.5$, and R=10.
Our initial hypothesis about the beam interact with ed is proven to be an insignificant relationship. In addition, the 3-way interaction of chest.wall*ed*pr is also proven to be insignificant. Hence, the LASSO results can only be a guide to what possible relationship exists between variables. A more detailed investigation is always necessary.

4.7 Conclusion

Although LASSO and SCAD variable screening method are powerful tools in initial stage of model selection, but in crq they can be tedious to use when presented with a large model of higher order terms. Unlike the standard crq model fitting function where you can simply multiply variables within the function to form interaction, LASSO and SCAD in crq is not capable of handling multiplications.

Our analysis of the lung data set had shown that when sufficient data is available, SCAD and LASSO produced very similar results. At $\tau = 0.5$ both SCAD and LASSO found chest.wall, beam, mi, prf, pd, pr, and ed to be significant. When data is scarce the results can very significantly, i.e. at $\tau$ near its extremes values ($1-\tau=0$, $1-\tau=\bar{\tau}$). In particular, SCAD and LASSO often disagreed on the variable beam and freq when $\tau$ is near its lower limit ($1-\tau$).

Furthermore, smaller $\tau$ values have different significant variables than the median. For example, chest.attend is not significant for $\tau = 0.5$ but it is significant for $\tau<0.1$ by both LASSO and SCAD.

Choice of $\lambda$ can be tricky for SCAD method, since $\lambda$ and $\tau$ has a parabolic relationship. It may be difficult to determine a sufficiently large penalty to use. On the other hand, LASSO method
has an inverse and approximately linear $\lambda$ and $\tau$ relationship, hence easier to determine the minimum penalty needed.

Model selection methods for censored quantile regression are limited. In this study we introduced the variable screening method, LASSO and SCAD combined with inference model selection. The proposed inference model selection is the bidirectional and backward model selection. This method is not available in crq, so it was performed manually. We found the best-fit model at $\tau = 0.5$ using the bidirectional model selection method to be:

$$Pd + beam + prf + chest.wall:mi + pr:mi.$$  

The best-fit model at $\tau = 0.5$ using the backward model selection method (exclude mi) to be:

$$Pd + beam + prf + chest.wal + pr2,$$

where $pr2$ is the quadratic term of $pr$.

In the ultrasound problem, the most influential variable is $pr$, especially when $pr < 0.89$. Controlling the instrument parameter $pr$ will most effectively control the lesion depth. The only significant animal variable is $chest.wall$. It has a positive influence on lesion depth. Naturally, the chest wall thickness is related to animal’s age and health condition. This should shed some insight to the importance of the biological condition of the subjects.

Threshold values are generally of interest for the ultrasound problems. The censored quantile regression method is highly recommended for this type of problem. Because the best fit model may differ depending on the quantile of interest. The LASSO and SCAD screening method is a good guide for variable selection process. But the true relationship can only be exposed by more
thorough investigation. The bidirectional and backward model selection produced similar results and is believed to be effective in identifying the more complex relationships such as interactions.
Chapter 5

Computing Algorithm

The complete algorithm for the model screening process is provided in Appendix A. Here is an outline of the procedure.

1) Use program crq.por.tsp (see Appendix A.1) to fit the saturated model. This is a modification of the R-program crq in the quantreg package (Koenker, 2014). It will return the index of the crossed censored data points, Isplit, and the crossed $\tau_i$, tsp

2) Retrieve Isplit, and tsp and check for mistakes.

3) Specify the $\tau$ of interest.

4) Identify the $\tau_i$ that are less than $\tau$. Duplicate the entries corresponding to these $\tau_i$ and attach it to the end of the X matrix. This prepares these data points to accept the double weights.

5) Overwrite the $y$ of the first set of cossed $\tau_i$ with a vary large value.

6) Define the weight vector, w, to be a vector of 1’s. This vector will be used to store weights

7) Columns combine w to the new X matrix.

8) Calculate the split weight entries using the following equations:

$$w_i(\tau) = \frac{\tau - \tau_i}{1 - \tau_i}, \text{ for } (x_i, +\infty),$$

1. $1 - w_i(\tau)$, otherwise.

9) Replace the entries in w that corresponds to the duplicated crossed $\tau_i$ with $w_i(\tau)$ and the original entries of crossed $\tau_i$ with $1 - w_i(\tau)$. This completes the weight vector w.
10) To generate the coefficients using lasso, a beta matrix of size $n \times p$ has to be defined, where $n$ is the total number of $\lambda$, and $p$ is the number of variables.

11) Create a $X$ matrix to be used in Lasso and SCAD by simply column-combining the variables of interest. Note, a vector of 1’s must be included for the intercept.

12) Run rq.fit.lasso with $x*w$ and $y*w$.

$$D<- rq.fit.lasso(x*w, y*w, \tau=0.5, \lambda=l[i], \beta=.9995, \epsilon=1e-06)$$

13) Finally write the above lasso file “lasso coef’s” to a csv file at a chosen location.

The stepwise regression methods used for the fine-tuning of variables are described below.

Bidirectional elimination – Begin with a saturated model with only the main effects. The main effect terms are selected using the LASSO or SCAD screening method. Then perform backward elimination. Every time a main effects model is removed, reintroduce the eliminated variable as an interaction with the others back in the model. Continue backward elimination until all variables are significant.

Backward elimination – Begin with the saturated model of the main effects and interaction terms. The main effects and interaction terms are selected using the LASSO or SCAD screening method. Then perform backward elimination, until all variables are significant.
Appendix A

Algorithm

A.1 Model Screening – LASSO

```r
library(survival)
library(quantreg)

#############################################################
### Input data into R ###
#############################################################

tmp <- read.csv("3_29_LungSummaryStats-12-10-05.csv")

names(tmp)
summary(tmp)
Dataset <-
tmp[,c("Study.Name",
   "Species",
   "Animal.Number",
   "Chest.wall.thickness..mm.",
   "Chest.wall.attenuation..dB.mm.",
   "freq..MHz.",
   "PD",
   "PRF",
   "ED",
   "Beamwidth",
   "In.situ.pr..MPa.",
   "MI",
   "Lesion..0.No..1.Yes",
   "Lesion.depth..mm.",
   "Lesion.surface.area...mm.2.")]```

#############################################################
### Rename the variables ###
#############################################################

```r
names(Dataset) <-
c("study",
   "species",
   "species")
```
"animal",
"chest.wall",
"chest.attenu",
"freq",
"pd",
"prf",
"ed",
"beam",
"pr",
"mi",
"lesion",
"depth",
"area")

attach(Dataset)
Dataset<-data.frame(Dataset)
attach(Dataset)

##############################################
#### Correct for miss matched pairs of lesion and depth ####
##############################################

Dataset[depth>0 && lesion==0, 'lesion']=1
Dataset[depth==0 && lesion==1, 'lesion']=0
depth = Dataset$depth
lesion = Dataset$lesion

####################################################
### Standardize variables (mean, sd) ###
####################################################

Dataset$mi<-mi<-c(mi-mean(mi))/(var(mi))^.5
Dataset$freq<-freq<-c(freq-mean(freq))/(var(freq))^.5
Dataset$prf<-prf<-c(prf-mean(prf))/(var(prf))^.5
Dataset$pd<-pd<-c(pd-mean(pd))/(var(pd))^.5
Dataset$pr<-pr<-c(pr-mean(pr))/(var(pr))^.5

#------------------#
# Generate Weights #
#------------------#

####################################################
### Fit model to generate tsp and Isplit ..........................###
### Load before fitting (1) crq.por.tsp (2)crq.fit.por_return_tsp ..........................###
####################################################

fit1 <- crq.por.tsp(Surv(depth, lesion )~mi+freq+prf+pd+pr, method = "Por",
                   data = Dataset)
### Retrieve tau_i (tau of the cross censored y’s) ###

tau_i<-fit1$T

### specify tau of interest ###

tau<-0.5

### Determine the number of tau_i< tau and record each id ###

c<-0
for (i in 1:length(tau_i))
{
  if (tau_i[i]<tau) {c<-c+1}
}
#tau_split is the vector of tau_i<tau
#id is the id of the corresponding tau_split
tau_split<-rep(0,c)
id<-rep(0,c)
j<-1
for (i in 1:length(tau_i))
{
  if(tau_i[i]<tau && j<(c+1))
  {
    tau_split[j]<-tau_i[i]
    id[j]<-fit1$Isplit[i]
    j<-j+1
  }
}

### Duplicate entries identified by id, prepare for split weight assignment ###

xd<-Dataset[id,]

### Replace original censored depth with a large enough y (100mm) ###

for (i in 1:length(Dataset$depth))
{
if (depth[i]==0.00){depth[i]<-100}
}

#########################################################################
### Create vector of weights, set it to 1 ###
#########################################################################

w<-rep(1,(length(depth)+length(tau_split)))

#########################################################################
### Creating new data frame with additional column ###
### of weights and duplicate set of x and y values ###
#########################################################################

Dataset$depth<-depth
Dataset_weight<-data.frame(cbind(data.frame(rbind(Dataset,xd)),w))
attach(Dataset_weight)

#########################################################################
### Must overwrite depth to update the addition of new entries ###
#########################################################################

depth<-Dataset_weight$depth

#########################################################################
### Calculate weights for the original and duplicated tau_i ###
#########################################################################

n<-length(w)-length(tau_split)
dummy<-rep(0,length(tau_split))
j<-1
for(i in 1:length(w))
{
  if(i>n){
    w[i]<-(tau-tau_split[j])/(1-tau_split[j])
    dummy[j]<-w[i]
    j<-j+1
  }
}
w[id]<-1-dummy

        #-------#
        # Lasso #
        #-------#

#########################################################################
### Generate coef for a sequence of lambda (l) in lasso ###
#########################################################################
n<-50
l<-rep(0,n)
beta<-matrix(0,nrow = n, ncol=11,)
x<-
cbind(1,Dataset_weight$study,Dataset_weight$chest.wall,Dataset_weight$chest.attention,Dataset_weight$ed,Dataset_weight$beam,Dataset_weight$mi,Dataset_weight$freq, Dataset_weight$prf, Dataset_weight$pr, Dataset_weight$pd)
y<-Dataset_weight$depth
for(i in 1:n){
l[i]<-2^(i/4)
d<-rq.fit.scad(x*w, y*w,tau=0.5,lambda=l[i], beta = .9995, eps = 1e-06)
beta[i,]<-d$coef
}

########################################
### Export file ###
########################################
write.csv(beta, file="lasso_all_variable_50")
A.2 crq.por.tsp

crq.por.tsp <-
function (formula, taus, data, subset, weights, na.action, method =
c("Powell", "Portnoy", "PengHuang"), contrasts = NULL, ...)
{
  require(survival)
  call <- match.call()
  mf <- match.call(expand.dots = FALSE)
  m <- match(c("formula", "data", "subset", "weights", "na.action"),
             names(mf), 0)
  mf <- mf[c(1, m)]
  mf$drop.unused.levels <- TRUE
  mf[[1]] <- as.name("model.frame")
  mf <- eval.parent(mf)
  if (method == "model.frame")
    return(mf)
  mt <- attr(mf, "terms")
  X <- model.matrix(mt, mf, contrasts)
  weights <- model.weights(mf)
  Y <- model.extract(mf, "response")
  eps <- .Machine$double.eps^(2/3)
  if (!inherits(Y, "Surv"))
    stop("Response must be a survival object")
  method <- match.arg(method)
  if (method == "Powell") {
    type <- attr(Y, "type")
    if (!type %in% c("right", "left"))
      stop("Only right or left censoring Surv objects are allowed")
    left <- (type == "left")
    if (any(taus < -eps) || any(taus > 1 + eps))
      stop("Invalid taus: taus should be >= 0 and <= 1")
    y <- Y[, 1]
    cen <- Y[, 2]
    if (length(taus) > 1) {
      coef <- matrix(0, ncol(X), length(taus))
      fitted <- resid <- matrix(0, nrow(X), length(taus))
      for (i in 1:length(taus))
        z <- crq.fit.pow(X, y, cen, tau = taus[i], weights,
                         left = left, ...)
        coef[, i] <- z$coefficients
        resid[, i] <- z$residuals
        fitted[, i] <- y - z$residuals
      }
    }
  taulabs <- paste("tau=", format(round(taus, 3)))
  dimnames(coef) <- list(dimnames(X)[[2]], taulabs)
  dimnames(resid) <- list(dimnames(X)[[1]], taulabs)
fit <- list(coefficients = coef, residuals = resid,
            fitted.values = fitted)
fit$tau <- taus
class(fit) <- "crqs"
}
else {
    fit <- crq.fit.pow(X, y, cen, tau = taus, weights,
                       left = left, ...)
    fit$tau <- taus
    class(fit) <- "crq"
}
}
else if (method == "Portnoy") {
    if (attr(Y, "type") != "right")
        stop("Only right censoring Surv objects are allowed for Portnoy method")
    y <- Y[, 1]
    cen <- Y[, 2]
    fit <- crq.fit.por.tsp(X, y, cen, weights, ...)
    class(fit) <- "crq"
}
else if (method == "PengHuang") {
    if (attr(Y, "type") != "right")
        stop("Only right censoring Surv objects are allowed for Peng-Huang method")
    y <- Y[, 1]
    cen <- Y[, 2]
    fit <- crq.fit.pen(X, y, cen, weights, ...)
    class(fit) <- "crq"
}
else stop("Method not defined for crq")
fit$terms <- mt
fit$call <- call
fit$formula <- formula(mt)
fit$method <- method
attr(fit, "na.message") <- attr(m, "na.message")
fit
A.3 crq.fit.por.tsp

```r

```n

crq.fit.por.tsp <- function (x, y, cen, weights = NULL, grid)
{
  p <- ncol(x)
  n <- length(y)
  cen <- 1 - cen
  mp <- n + 5 + max(1, sum(cen))
  eps <- 1e-04
  if (length(weights)) {
    if (any(weights < 0))
      stop("negative weights not allowed")
    contr <- attr(x, "contrasts")
    x <- weights * x
    y <- weights * y
  }
  if (missing(grid))
    grid <- seq(1/n, 1 - 1/n, by = 1/(5 + 3 * n^0.4))
  if (is.numeric(grid)) {
    ginit <- min(grid)
    dgrid <- diff(grid)
    gstep <- median(dgrid)
    if (any(dgrid < 0))
      stop("grid is not monotonic")
    if (gstep < eps)
      stop("grid stepsize too small")
    nsol <- 3 * n
    mw = -1
  } else if (grid == "pivot") {
    nsol <- 3 * n
    ginit <- 1/(2 * n)
    gstep <- 1/(2 * n)
    mw <- 20
  } else stop("Invalid grid")
  z <- .Fortran("crq", as.integer(n), as.integer(p), as.integer(mp),
    as.integer(p + 2), as.double(x), as.double(y), as.integer(cen),
    as.double(ginit), as.integer(mw), as.double(gstep), ift = integer(1),
    h = integer(p), xh = double(p * p), wa = double(mp *
    p), wb = double(mp), wc = double(mp * (p + 2)), wd = double(mp),
    we = double(mp), wf = double(p), iflag = integer(mp),
    as.integer(nsol), sol = double(nsol * (p + 2)), lsol = integer(1),
    icen = integer(n), tcen = double(n), lcen = integer(1),
    PACKAGE = "quantreg")
  nw <- z$h[1]
  flag <- z$ift
```n
```
msg <- switch(flag, paste("Error in input dimensions, n,p,mw "),
  paste("Error in input dimensions, n,p,mw "), paste("Error in input dimensions, n,p,mw "),
  paste("Less than p=", p, "observations above tau = 0 solution"),
  paste("Possible degeneracy at", nw, "tau values.", "$tau.degen: first mp =",
    n + 5 + sum(cen), " such tau values"), paste("Number of pivots to be saved in sol > nsol.",
    "Redefine nsol: use nsol < n to save for tau = i/(nsol-1)"),
  paste("Error with partial return: possible degeneracies",
    "Max number of rq calls exceeded: dither x or increase mw"),
  paste("Premature stop: defective conditional distribution"))
if (flag > 0 && flag != 5 && flag < 8)
  ifelse(flag <= 3, stop(msg), warning(msg))
J <- z$lsol
B <- matrix(z$sol, nrow = p + 2, ncol = nsol, byrow = FALSE)[,
  1:J]
dimnames(B) <- list(c("tau", dimnames(x)[[2]], "Qbar"), NULL)
ic <- z$icen
sp <- (1:n)[ic == 1]
tsp <- z$tcen[sp]
t1 <- z$wd[1:nw]
a <- list(sol = B, Isplit = sp, status = ic, T=tsp)
class(a) <- "crq"
return(a)
## Appendix B

### Summary Tables

#### B.1 LASSO 2-way Interaction $\tau = 0.5$: Chest.Wall

<table>
<thead>
<tr>
<th>lambda</th>
<th>chest.wall</th>
<th>cw_ca</th>
<th>cw_f</th>
<th>cw_pd</th>
<th>cw_prf</th>
<th>cw_ed</th>
<th>cw_b</th>
<th>cw_pr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.189</td>
<td>1.993</td>
<td>-3.455</td>
<td>4.554</td>
<td>-1.282</td>
<td>-1.625</td>
<td>0.337</td>
<td>1.899</td>
<td>-0.313</td>
</tr>
<tr>
<td>1.414</td>
<td>2.347</td>
<td>-3.426</td>
<td>4.485</td>
<td>0.000</td>
<td>-1.430</td>
<td>0.307</td>
<td>2.010</td>
<td>-0.190</td>
</tr>
<tr>
<td>1.682</td>
<td>2.265</td>
<td>-3.106</td>
<td>4.029</td>
<td>0.000</td>
<td>-1.425</td>
<td>0.346</td>
<td>1.797</td>
<td>-0.221</td>
</tr>
<tr>
<td>2.000</td>
<td>2.130</td>
<td>-2.635</td>
<td>3.212</td>
<td>0.000</td>
<td>-1.177</td>
<td>0.122</td>
<td>1.389</td>
<td>-0.117</td>
</tr>
<tr>
<td>2.378</td>
<td>1.957</td>
<td>-2.565</td>
<td>3.083</td>
<td>0.000</td>
<td>-0.867</td>
<td>0.046</td>
<td>1.405</td>
<td>0.000</td>
</tr>
<tr>
<td>2.828</td>
<td>1.805</td>
<td>-2.168</td>
<td>2.530</td>
<td>0.000</td>
<td>-0.848</td>
<td>0.036</td>
<td>1.245</td>
<td>0.000</td>
</tr>
<tr>
<td>3.364</td>
<td>1.615</td>
<td>-1.695</td>
<td>1.793</td>
<td>0.000</td>
<td>-0.617</td>
<td>0.000</td>
<td>0.863</td>
<td>0.057</td>
</tr>
<tr>
<td>4.000</td>
<td>1.609</td>
<td>-1.101</td>
<td>1.142</td>
<td>0.000</td>
<td>-0.344</td>
<td>0.000</td>
<td>0.308</td>
<td>0.170</td>
</tr>
<tr>
<td>4.757</td>
<td>1.513</td>
<td>-0.449</td>
<td>0.866</td>
<td>0.000</td>
<td>-0.315</td>
<td>0.000</td>
<td>0.000</td>
<td>0.361</td>
</tr>
<tr>
<td>5.657</td>
<td>1.270</td>
<td>0.000</td>
<td>0.696</td>
<td>0.000</td>
<td>-0.253</td>
<td>0.000</td>
<td>0.000</td>
<td>0.561</td>
</tr>
<tr>
<td>6.727</td>
<td>1.175</td>
<td>0.000</td>
<td>0.530</td>
<td>0.000</td>
<td>-0.055</td>
<td>0.000</td>
<td>0.000</td>
<td>0.569</td>
</tr>
<tr>
<td>8.000</td>
<td>1.037</td>
<td>0.000</td>
<td>0.157</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.566</td>
</tr>
<tr>
<td>9.514</td>
<td>0.981</td>
<td>0.000</td>
<td>0.033</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.489</td>
</tr>
<tr>
<td>11.314</td>
<td>0.987</td>
<td>0.000</td>
<td>0.041</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.485</td>
</tr>
<tr>
<td>13.454</td>
<td>0.945</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.475</td>
</tr>
<tr>
<td>16.000</td>
<td>0.937</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.476</td>
</tr>
<tr>
<td>19.027</td>
<td>0.875</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.426</td>
</tr>
<tr>
<td>22.627</td>
<td>0.839</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.359</td>
</tr>
<tr>
<td>26.909</td>
<td>0.741</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.339</td>
</tr>
<tr>
<td>32.000</td>
<td>0.561</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.382</td>
</tr>
<tr>
<td>38.055</td>
<td>0.499</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.151</td>
</tr>
<tr>
<td>45.255</td>
<td>0.481</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.107</td>
</tr>
<tr>
<td>53.817</td>
<td>0.418</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.019</td>
</tr>
<tr>
<td>64.000</td>
<td>0.352</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>76.109</td>
<td>0.320</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>90.510</td>
<td>0.222</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>107.635</td>
<td>0.135</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>128.000</td>
<td>0.056</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>152.219</td>
<td>0.008</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>181.019</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
### B.2 LASSO 2-way Interaction at $\tau = 0.5$: Chest.Atten

<table>
<thead>
<tr>
<th>lambda</th>
<th>chest.attent</th>
<th>cw_ca</th>
<th>ca_f</th>
<th>ca_pd</th>
<th>ca_prf</th>
<th>ca_ed</th>
<th>ca_b</th>
<th>ca_pr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.189</td>
<td>44.063</td>
<td>-3.455</td>
<td>0.000</td>
<td>0.000</td>
<td>5.249</td>
<td>0.000</td>
<td>-47.100</td>
<td>6.866</td>
</tr>
<tr>
<td>1.414</td>
<td>23.972</td>
<td>-3.426</td>
<td>15.029</td>
<td>0.000</td>
<td>4.807</td>
<td>0.000</td>
<td>-23.939</td>
<td>6.858</td>
</tr>
<tr>
<td>1.682</td>
<td>24.589</td>
<td>-3.106</td>
<td>15.058</td>
<td>0.000</td>
<td>4.758</td>
<td>0.000</td>
<td>-23.832</td>
<td>6.308</td>
</tr>
<tr>
<td>2.000</td>
<td>25.251</td>
<td>-2.635</td>
<td>15.136</td>
<td>0.000</td>
<td>4.811</td>
<td>0.000</td>
<td>-23.711</td>
<td>5.681</td>
</tr>
<tr>
<td>2.378</td>
<td>26.109</td>
<td>-2.565</td>
<td>15.048</td>
<td>0.000</td>
<td>3.883</td>
<td>0.000</td>
<td>-23.584</td>
<td>5.176</td>
</tr>
<tr>
<td>2.828</td>
<td>27.167</td>
<td>-2.168</td>
<td>15.087</td>
<td>0.000</td>
<td>3.726</td>
<td>0.000</td>
<td>-23.421</td>
<td>4.211</td>
</tr>
<tr>
<td>3.364</td>
<td>28.153</td>
<td>-1.695</td>
<td>15.137</td>
<td>0.000</td>
<td>3.498</td>
<td>0.000</td>
<td>-23.147</td>
<td>3.312</td>
</tr>
<tr>
<td>4.000</td>
<td>29.121</td>
<td>-1.101</td>
<td>15.119</td>
<td>0.000</td>
<td>2.915</td>
<td>0.000</td>
<td>-22.878</td>
<td>2.579</td>
</tr>
<tr>
<td>4.757</td>
<td>29.270</td>
<td>-0.449</td>
<td>15.153</td>
<td>0.000</td>
<td>2.719</td>
<td>0.000</td>
<td>-22.411</td>
<td>2.276</td>
</tr>
<tr>
<td>5.657</td>
<td>5.393</td>
<td>0.000</td>
<td>15.840</td>
<td>0.000</td>
<td>2.330</td>
<td>0.000</td>
<td>-0.736</td>
<td>1.987</td>
</tr>
<tr>
<td>6.727</td>
<td>5.473</td>
<td>0.000</td>
<td>15.802</td>
<td>0.000</td>
<td>1.660</td>
<td>0.000</td>
<td>-0.200</td>
<td>1.579</td>
</tr>
<tr>
<td>8.000</td>
<td>5.515</td>
<td>0.000</td>
<td>15.762</td>
<td>0.000</td>
<td>1.477</td>
<td>0.000</td>
<td>0.000</td>
<td>1.379</td>
</tr>
<tr>
<td>9.514</td>
<td>5.805</td>
<td>0.000</td>
<td>15.758</td>
<td>0.000</td>
<td>1.229</td>
<td>0.000</td>
<td>0.000</td>
<td>1.225</td>
</tr>
<tr>
<td>11.314</td>
<td>5.855</td>
<td>0.000</td>
<td>15.709</td>
<td>0.000</td>
<td>1.214</td>
<td>0.000</td>
<td>0.000</td>
<td>1.159</td>
</tr>
<tr>
<td>13.454</td>
<td>5.915</td>
<td>0.000</td>
<td>15.686</td>
<td>0.000</td>
<td>1.166</td>
<td>0.000</td>
<td>0.000</td>
<td>1.106</td>
</tr>
<tr>
<td>16.000</td>
<td>5.960</td>
<td>0.000</td>
<td>15.646</td>
<td>0.000</td>
<td>1.200</td>
<td>0.000</td>
<td>0.000</td>
<td>0.993</td>
</tr>
<tr>
<td>19.027</td>
<td>4.358</td>
<td>0.000</td>
<td>12.561</td>
<td>0.000</td>
<td>0.993</td>
<td>0.000</td>
<td>0.000</td>
<td>0.962</td>
</tr>
<tr>
<td>22.627</td>
<td>3.856</td>
<td>0.000</td>
<td>11.067</td>
<td>0.000</td>
<td>0.326</td>
<td>0.000</td>
<td>0.000</td>
<td>0.944</td>
</tr>
<tr>
<td>26.909</td>
<td>4.111</td>
<td>0.000</td>
<td>10.970</td>
<td>-0.024</td>
<td>0.238</td>
<td>0.000</td>
<td>0.000</td>
<td>0.687</td>
</tr>
<tr>
<td>32.000</td>
<td>3.410</td>
<td>0.000</td>
<td>8.264</td>
<td>-0.113</td>
<td>0.21</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>38.055</td>
<td>0.000</td>
<td>0.000</td>
<td>1.283</td>
<td>0.000</td>
<td>0.041</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>45.255</td>
<td>0.000</td>
<td>0.000</td>
<td>1.007</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>53.817</td>
<td>0.000</td>
<td>0.000</td>
<td>0.651</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.084</td>
<td>0.000</td>
</tr>
<tr>
<td>64.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.385</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.097</td>
<td>0.000</td>
</tr>
<tr>
<td>76.109</td>
<td>0.000</td>
<td>0.000</td>
<td>0.352</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.082</td>
<td>0.000</td>
</tr>
<tr>
<td>90.510</td>
<td>0.000</td>
<td>0.000</td>
<td>0.245</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.076</td>
<td>0.000</td>
</tr>
<tr>
<td>107.635</td>
<td>0.000</td>
<td>0.000</td>
<td>0.163</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.033</td>
<td>0.000</td>
</tr>
<tr>
<td>128.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.050</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>152.219</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
B.3 LASSO 2-way Interaction at $\tau = 0.5$: \textit{Freq}

<table>
<thead>
<tr>
<th>lambda</th>
<th>freq</th>
<th>cw_f</th>
<th>ca_f</th>
<th>f_pd</th>
<th>f_prf</th>
<th>f_ed</th>
<th>f_b</th>
<th>f_pr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.189</td>
<td>-31.543</td>
<td>4.554</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.786</td>
<td>20.679</td>
<td>-3.787</td>
</tr>
<tr>
<td>1.414</td>
<td>-18.635</td>
<td>4.485</td>
<td>15.029</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.742</td>
<td>16.141</td>
<td>-3.873</td>
</tr>
<tr>
<td>1.682</td>
<td>-19.344</td>
<td>4.029</td>
<td>15.058</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.662</td>
<td>16.062</td>
<td>-3.325</td>
</tr>
<tr>
<td>2.000</td>
<td>-20.307</td>
<td>3.212</td>
<td>15.136</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.590</td>
<td>15.977</td>
<td>-2.676</td>
</tr>
<tr>
<td>2.378</td>
<td>-19.990</td>
<td>3.083</td>
<td>15.048</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.566</td>
<td>15.844</td>
<td>-2.805</td>
</tr>
<tr>
<td>2.828</td>
<td>-21.124</td>
<td>2.530</td>
<td>15.087</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.422</td>
<td>15.716</td>
<td>-1.864</td>
</tr>
<tr>
<td>3.364</td>
<td>-22.005</td>
<td>1.793</td>
<td>15.137</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.323</td>
<td>15.551</td>
<td>-1.113</td>
</tr>
<tr>
<td>4.000</td>
<td>-22.442</td>
<td>1.142</td>
<td>15.119</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.233</td>
<td>15.337</td>
<td>-0.763</td>
</tr>
<tr>
<td>4.757</td>
<td>-22.324</td>
<td>0.866</td>
<td>15.153</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.208</td>
<td>15.080</td>
<td>-0.581</td>
</tr>
<tr>
<td>5.657</td>
<td>-5.939</td>
<td>0.696</td>
<td>15.840</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.234</td>
<td>0.656</td>
<td>-0.658</td>
</tr>
<tr>
<td>6.727</td>
<td>-5.876</td>
<td>0.530</td>
<td>15.802</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.096</td>
<td>0.240</td>
<td>-0.357</td>
</tr>
<tr>
<td>8.000</td>
<td>-5.946</td>
<td>0.157</td>
<td>15.762</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.073</td>
<td>0.045</td>
<td>-0.302</td>
</tr>
<tr>
<td>9.514</td>
<td>-6.015</td>
<td>0.033</td>
<td>15.758</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.049</td>
<td>0.031</td>
<td>-0.281</td>
</tr>
<tr>
<td>11.314</td>
<td>-6.038</td>
<td>0.041</td>
<td>15.709</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.043</td>
<td>0.000</td>
<td>-0.226</td>
</tr>
<tr>
<td>13.454</td>
<td>-6.031</td>
<td>0.000</td>
<td>15.686</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.043</td>
<td>0.000</td>
<td>-0.206</td>
</tr>
<tr>
<td>16.000</td>
<td>-5.969</td>
<td>0.000</td>
<td>15.646</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.163</td>
</tr>
<tr>
<td>19.027</td>
<td>-1.926</td>
<td>0.000</td>
<td>12.561</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-2.230</td>
</tr>
<tr>
<td>22.627</td>
<td>0.000</td>
<td>0.000</td>
<td>11.067</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-3.266</td>
</tr>
<tr>
<td>26.909</td>
<td>0.000</td>
<td>0.000</td>
<td>10.970</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-3.181</td>
</tr>
<tr>
<td>32.000</td>
<td>0.000</td>
<td>0.000</td>
<td>8.264</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.213</td>
<td>-2.209</td>
</tr>
<tr>
<td>38.055</td>
<td>0.000</td>
<td>0.000</td>
<td>1.283</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.207</td>
<td>0.000</td>
</tr>
<tr>
<td>45.255</td>
<td>0.000</td>
<td>0.000</td>
<td>1.007</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.092</td>
<td>0.000</td>
</tr>
<tr>
<td>53.817</td>
<td>0.000</td>
<td>0.000</td>
<td>0.651</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>64.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.385</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>76.109</td>
<td>0.000</td>
<td>0.000</td>
<td>0.352</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>90.510</td>
<td>0.000</td>
<td>0.000</td>
<td>0.245</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>107.635</td>
<td>0.000</td>
<td>0.000</td>
<td>0.163</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>128.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.050</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>152.219</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
B.4 LASSO 2-way Interaction at \( \tau = 0.5 \): \( Pd \)

<table>
<thead>
<tr>
<th>lambda</th>
<th>pd</th>
<th>cw_pd</th>
<th>ca_pd</th>
<th>f_pd</th>
<th>pd_prf</th>
<th>pd_ed</th>
<th>pd_b</th>
<th>pd_pr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.189</td>
<td>-2.487</td>
<td>-1.282</td>
<td>0.000</td>
<td>0.000</td>
<td>1.220</td>
<td>0.000</td>
<td>0.000</td>
<td>0.571</td>
</tr>
<tr>
<td>1.414</td>
<td>-1.698</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.062</td>
<td>0.000</td>
<td>0.000</td>
<td>0.477</td>
</tr>
<tr>
<td>1.682</td>
<td>-1.671</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.049</td>
<td>0.000</td>
<td>0.000</td>
<td>0.472</td>
</tr>
<tr>
<td>2.000</td>
<td>-1.662</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.041</td>
<td>0.000</td>
<td>0.000</td>
<td>0.443</td>
</tr>
<tr>
<td>2.378</td>
<td>-1.046</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.681</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>2.828</td>
<td>-1.002</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.648</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>3.364</td>
<td>-0.912</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.610</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>4.000</td>
<td>-0.601</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.407</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>4.757</td>
<td>-0.476</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.348</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>5.657</td>
<td>-0.347</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.287</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>6.727</td>
<td>-0.199</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.091</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>8.000</td>
<td>-0.106</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.046</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>9.514</td>
<td>-0.002</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>11.314</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>13.454</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>16.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>19.027</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>22.627</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>26.909</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>32.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.113</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>38.055</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>45.255</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>53.817</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
### B.5 LASSO 2-way Interaction at $\tau = 0.5$: Prf

<table>
<thead>
<tr>
<th>lambda</th>
<th>prf</th>
<th>cw_prf</th>
<th>ca_prf</th>
<th>f_prf</th>
<th>pd_prf</th>
<th>prf_pr</th>
<th>prf_b</th>
<th>prf_ed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.189</td>
<td>-2.063</td>
<td>-1.625</td>
<td>5.249</td>
<td>0.000</td>
<td>1.220</td>
<td>0.000</td>
<td>0.000</td>
<td>2.529</td>
</tr>
<tr>
<td>1.414</td>
<td>-1.903</td>
<td>-1.430</td>
<td>4.807</td>
<td>0.000</td>
<td>1.062</td>
<td>0.000</td>
<td>0.000</td>
<td>2.239</td>
</tr>
<tr>
<td>1.682</td>
<td>-1.883</td>
<td>-1.425</td>
<td>4.758</td>
<td>0.000</td>
<td>1.049</td>
<td>0.000</td>
<td>0.000</td>
<td>2.209</td>
</tr>
<tr>
<td>2.000</td>
<td>-1.712</td>
<td>-1.177</td>
<td>4.811</td>
<td>0.000</td>
<td>1.041</td>
<td>0.000</td>
<td>0.000</td>
<td>2.235</td>
</tr>
<tr>
<td>2.378</td>
<td>-1.389</td>
<td>-0.867</td>
<td>3.883</td>
<td>0.000</td>
<td>0.681</td>
<td>0.000</td>
<td>0.000</td>
<td>1.776</td>
</tr>
<tr>
<td>2.828</td>
<td>-1.320</td>
<td>-0.848</td>
<td>3.726</td>
<td>0.000</td>
<td>0.648</td>
<td>0.000</td>
<td>0.000</td>
<td>1.700</td>
</tr>
<tr>
<td>3.364</td>
<td>-1.126</td>
<td>-0.617</td>
<td>3.498</td>
<td>0.000</td>
<td>0.610</td>
<td>0.000</td>
<td>0.000</td>
<td>1.565</td>
</tr>
<tr>
<td>4.000</td>
<td>-0.862</td>
<td>-0.344</td>
<td>2.915</td>
<td>0.000</td>
<td>0.407</td>
<td>0.000</td>
<td>0.000</td>
<td>1.316</td>
</tr>
<tr>
<td>4.757</td>
<td>-0.802</td>
<td>-0.315</td>
<td>2.719</td>
<td>0.000</td>
<td>0.348</td>
<td>0.000</td>
<td>0.000</td>
<td>1.222</td>
</tr>
<tr>
<td>5.657</td>
<td>-0.660</td>
<td>-0.253</td>
<td>2.330</td>
<td>0.000</td>
<td>0.287</td>
<td>0.000</td>
<td>0.000</td>
<td>1.051</td>
</tr>
<tr>
<td>6.727</td>
<td>-0.439</td>
<td>-0.055</td>
<td>1.660</td>
<td>0.000</td>
<td>0.091</td>
<td>0.000</td>
<td>0.000</td>
<td>0.667</td>
</tr>
<tr>
<td>8.000</td>
<td>-0.346</td>
<td>0.000</td>
<td>1.477</td>
<td>0.000</td>
<td>0.046</td>
<td>0.000</td>
<td>0.000</td>
<td>0.604</td>
</tr>
<tr>
<td>9.514</td>
<td>-0.276</td>
<td>0.000</td>
<td>1.229</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.497</td>
</tr>
<tr>
<td>11.314</td>
<td>-0.272</td>
<td>0.000</td>
<td>1.214</td>
<td>0.000</td>
<td>-0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.492</td>
</tr>
<tr>
<td>13.454</td>
<td>-0.254</td>
<td>0.000</td>
<td>1.166</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.467</td>
</tr>
<tr>
<td>16.000</td>
<td>-0.056</td>
<td>0.000</td>
<td>1.200</td>
<td>0.000</td>
<td>-0.003</td>
<td>-0.006</td>
<td>0.000</td>
<td>0.394</td>
</tr>
<tr>
<td>19.027</td>
<td>0.000</td>
<td>0.000</td>
<td>0.993</td>
<td>0.000</td>
<td>-0.047</td>
<td>0.000</td>
<td>0.000</td>
<td>0.305</td>
</tr>
<tr>
<td>22.627</td>
<td>0.000</td>
<td>0.000</td>
<td>0.326</td>
<td>0.000</td>
<td>-0.142</td>
<td>-0.027</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>26.909</td>
<td>0.000</td>
<td>0.000</td>
<td>0.238</td>
<td>0.000</td>
<td>-0.104</td>
<td>-0.061</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>32.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.021</td>
<td>0.000</td>
<td>-0.095</td>
<td>-0.068</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>38.055</td>
<td>0.000</td>
<td>0.000</td>
<td>0.041</td>
<td>0.000</td>
<td>-0.047</td>
<td>-0.076</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>45.255</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.040</td>
<td>-0.066</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>53.817</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
B.6 LASSO 2-way Interaction at $\tau = 0.5$: $ed$

<table>
<thead>
<tr>
<th>lambda</th>
<th>ed</th>
<th>cw_ed</th>
<th>ca_ed</th>
<th>f_ed</th>
<th>pd_ed</th>
<th>prf_ed</th>
<th>ed_b</th>
<th>ed_pr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.189</td>
<td>-0.369</td>
<td>0.337</td>
<td>0.000</td>
<td>-0.786</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.307</td>
</tr>
<tr>
<td>1.414</td>
<td>-0.463</td>
<td>0.307</td>
<td>0.000</td>
<td>-0.742</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.368</td>
</tr>
<tr>
<td>1.682</td>
<td>-0.417</td>
<td>0.346</td>
<td>0.000</td>
<td>-0.662</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.394</td>
</tr>
<tr>
<td>2.000</td>
<td>-0.463</td>
<td>0.122</td>
<td>0.000</td>
<td>-0.590</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.440</td>
</tr>
<tr>
<td>2.378</td>
<td>-0.487</td>
<td>0.046</td>
<td>0.000</td>
<td>-0.566</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.477</td>
</tr>
<tr>
<td>2.828</td>
<td>-0.413</td>
<td>0.036</td>
<td>0.000</td>
<td>-0.422</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.399</td>
</tr>
<tr>
<td>3.364</td>
<td>-0.362</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.323</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.387</td>
</tr>
<tr>
<td>4.000</td>
<td>-0.322</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.233</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.342</td>
</tr>
<tr>
<td>4.757</td>
<td>-0.298</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.208</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.279</td>
</tr>
<tr>
<td>5.657</td>
<td>-0.222</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.234</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.226</td>
</tr>
<tr>
<td>6.727</td>
<td>-0.275</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.096</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.072</td>
</tr>
<tr>
<td>8.000</td>
<td>-0.243</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.073</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>9.514</td>
<td>-0.231</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.049</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>11.314</td>
<td>-0.228</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.043</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>13.454</td>
<td>-0.232</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.043</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>16.000</td>
<td>-0.218</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>19.027</td>
<td>-0.199</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>22.627</td>
<td>-0.258</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>26.909</td>
<td>-0.159</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>32.000</td>
<td>-0.052</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>38.055</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
B.7 LASSO 2-way Interaction at $\tau = 0.5$: beam

<table>
<thead>
<tr>
<th>lambda</th>
<th>beam</th>
<th>cw_b</th>
<th>ca_b</th>
<th>f_b</th>
<th>pd_b</th>
<th>prf_b</th>
<th>ed_b</th>
<th>b_pr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.189</td>
<td>-24.809</td>
<td>1.899</td>
<td>-47.100</td>
<td>20.679</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-2.420</td>
</tr>
<tr>
<td>1.414</td>
<td>-7.536</td>
<td>2.010</td>
<td>-23.939</td>
<td>16.141</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-2.499</td>
</tr>
<tr>
<td>1.682</td>
<td>-8.106</td>
<td>1.797</td>
<td>-23.831</td>
<td>16.062</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-1.997</td>
</tr>
<tr>
<td>2.000</td>
<td>-8.847</td>
<td>1.389</td>
<td>-23.711</td>
<td>15.977</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-1.421</td>
</tr>
<tr>
<td>2.378</td>
<td>-8.382</td>
<td>1.405</td>
<td>-23.584</td>
<td>15.844</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-1.762</td>
</tr>
<tr>
<td>2.828</td>
<td>-9.275</td>
<td>1.245</td>
<td>-23.421</td>
<td>15.716</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.893</td>
</tr>
<tr>
<td>3.364</td>
<td>-9.931</td>
<td>0.863</td>
<td>-23.147</td>
<td>15.551</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.270</td>
</tr>
<tr>
<td>4.000</td>
<td>-10.258</td>
<td>0.308</td>
<td>-22.878</td>
<td>15.337</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.099</td>
</tr>
<tr>
<td>4.757</td>
<td>-10.270</td>
<td>0.000</td>
<td>-22.411</td>
<td>15.080</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>5.657</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.736</td>
<td>0.656</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.228</td>
</tr>
<tr>
<td>6.727</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.200</td>
<td>0.240</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>8.000</td>
<td>0.056</td>
<td>0.000</td>
<td>0.000</td>
<td>0.045</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>9.514</td>
<td>0.065</td>
<td>0.000</td>
<td>0.000</td>
<td>0.031</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>11.314</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.048</td>
</tr>
<tr>
<td>13.454</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.051</td>
</tr>
<tr>
<td>16.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.054</td>
</tr>
<tr>
<td>19.027</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.081</td>
</tr>
<tr>
<td>22.627</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.099</td>
</tr>
<tr>
<td>26.909</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.105</td>
</tr>
<tr>
<td>32.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.213</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.011</td>
</tr>
<tr>
<td>38.055</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.207</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
B.8 LASSO 2-way Interaction at $\tau = 0.5$: $Pr$

<table>
<thead>
<tr>
<th>lambda</th>
<th>pr</th>
<th>cw_pr</th>
<th>ca_pr</th>
<th>f_pr</th>
<th>pd_pr</th>
<th>prf_pr</th>
<th>ed_pr</th>
<th>b_pr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.189</td>
<td>-6.000</td>
<td>-0.313</td>
<td>6.866</td>
<td>-3.787</td>
<td>0.571</td>
<td>2.529</td>
<td>-0.307</td>
<td>-2.420</td>
</tr>
<tr>
<td>1.414</td>
<td>-5.719</td>
<td>-0.190</td>
<td>6.858</td>
<td>-3.873</td>
<td>0.477</td>
<td>2.239</td>
<td>-0.368</td>
<td>-2.499</td>
</tr>
<tr>
<td>1.682</td>
<td>-5.562</td>
<td>-0.221</td>
<td>6.308</td>
<td>-3.325</td>
<td>0.472</td>
<td>2.209</td>
<td>-0.394</td>
<td>-1.997</td>
</tr>
<tr>
<td>2.000</td>
<td>-5.337</td>
<td>-0.117</td>
<td>5.681</td>
<td>-2.676</td>
<td>0.443</td>
<td>2.235</td>
<td>-0.440</td>
<td>-1.421</td>
</tr>
<tr>
<td>2.378</td>
<td>-4.570</td>
<td>0.000</td>
<td>5.176</td>
<td>-2.805</td>
<td>0.000</td>
<td>1.776</td>
<td>-0.477</td>
<td>-1.762</td>
</tr>
<tr>
<td>2.828</td>
<td>-4.227</td>
<td>0.000</td>
<td>4.211</td>
<td>-1.864</td>
<td>0.000</td>
<td>1.700</td>
<td>-0.399</td>
<td>-0.893</td>
</tr>
<tr>
<td>3.364</td>
<td>-3.802</td>
<td>0.057</td>
<td>3.312</td>
<td>-1.113</td>
<td>0.000</td>
<td>1.565</td>
<td>-0.387</td>
<td>-0.270</td>
</tr>
<tr>
<td>4.000</td>
<td>-3.115</td>
<td>0.170</td>
<td>2.579</td>
<td>-0.763</td>
<td>0.000</td>
<td>1.316</td>
<td>-0.342</td>
<td>-0.099</td>
</tr>
<tr>
<td>4.757</td>
<td>-2.849</td>
<td>0.361</td>
<td>2.276</td>
<td>-0.581</td>
<td>0.000</td>
<td>1.222</td>
<td>-0.279</td>
<td>0.000</td>
</tr>
<tr>
<td>5.657</td>
<td>-2.313</td>
<td>0.561</td>
<td>1.987</td>
<td>-0.658</td>
<td>0.000</td>
<td>1.051</td>
<td>-0.226</td>
<td>-0.228</td>
</tr>
<tr>
<td>6.727</td>
<td>-1.933</td>
<td>0.569</td>
<td>1.579</td>
<td>-0.357</td>
<td>0.000</td>
<td>0.667</td>
<td>-0.072</td>
<td>0.000</td>
</tr>
<tr>
<td>8.000</td>
<td>-1.695</td>
<td>0.566</td>
<td>1.379</td>
<td>-0.302</td>
<td>0.000</td>
<td>0.604</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>9.514</td>
<td>-1.444</td>
<td>0.489</td>
<td>1.225</td>
<td>-0.281</td>
<td>0.000</td>
<td>0.497</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>11.314</td>
<td>-1.418</td>
<td>0.485</td>
<td>1.159</td>
<td>-0.226</td>
<td>0.000</td>
<td>0.492</td>
<td>0.000</td>
<td>0.048</td>
</tr>
<tr>
<td>13.454</td>
<td>-1.392</td>
<td>0.475</td>
<td>1.106</td>
<td>-0.206</td>
<td>0.000</td>
<td>0.467</td>
<td>0.000</td>
<td>0.051</td>
</tr>
<tr>
<td>16.000</td>
<td>-1.305</td>
<td>0.476</td>
<td>0.993</td>
<td>-0.163</td>
<td>0.000</td>
<td>0.394</td>
<td>0.000</td>
<td>0.054</td>
</tr>
<tr>
<td>19.027</td>
<td>-2.107</td>
<td>0.426</td>
<td>0.962</td>
<td>-2.230</td>
<td>0.000</td>
<td>0.305</td>
<td>0.000</td>
<td>0.081</td>
</tr>
<tr>
<td>22.627</td>
<td>-2.355</td>
<td>0.359</td>
<td>0.944</td>
<td>-3.266</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.099</td>
</tr>
<tr>
<td>26.909</td>
<td>-2.330</td>
<td>0.339</td>
<td>0.687</td>
<td>-3.181</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.105</td>
</tr>
<tr>
<td>32.000</td>
<td>-1.801</td>
<td>0.382</td>
<td>0.000</td>
<td>-2.209</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.011</td>
</tr>
<tr>
<td>38.055</td>
<td>-0.547</td>
<td>0.151</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>45.255</td>
<td>-0.479</td>
<td>0.107</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>53.817</td>
<td>-0.436</td>
<td>0.019</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>64.000</td>
<td>-0.380</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>76.109</td>
<td>-0.371</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>90.510</td>
<td>-0.372</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>107.635</td>
<td>-0.397</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>128.000</td>
<td>-0.390</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>152.219</td>
<td>-0.390</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>181.019</td>
<td>-0.323</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>215.269</td>
<td>-0.274</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>256.000</td>
<td>-0.171</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>304.437</td>
<td>-0.040</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>362.039</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Bibliography


