PENALIZED, SHRINKAGE, AND PRELIMINARY TEST STRATEGIES IN NONLINEAR AND PROPORTIONAL HAZARD REGRESSION MODELS FOR LOW AND HIGH-DIMENSIONAL DATA

BY

JANJIRA PILADAENG

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (STATISTICS)

DEPARTMENT OF MATHEMATICS AND STATISTICS
FACULTY OF SCIENCE AND TECHNOLOGY
THAMMASAT UNIVERSITY

ACADEMIC YEAR 2021

THAMMASAT UNIVERSITY
FACULTY OF SCIENCE AND TECHNOLOGY

DISSERTATION

BY

JANJIRA PILADAENG

ENTITLED

PENALIZED, SHRINKAGE, AND PRELIMINARY TEST STRATEGIES IN NONLINEAR AND PROPORTIONAL HAZARD REGRESSION MODELS FOR LOW AND HIGH-DIMENSIONAL DATA

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

on June 23, 2022

Chairman	(Professor Saowanit Sukparungsee, Ph.D.)
Member and Advisor	(Associate Professor Supranee Lisawadi, Ph.D.)
Member and Co-Advisor	(Professor Syed Ejaz Ahmed, Ph.D.)
Member	(Assistant Professor Patchanok Srisuradetchai, Ph.D.)
Member	(Assistant Professor Teerawat Simmachan, Ph.D.)
Dean	(Associate Professor Nuttanont Hongwarittorrn, Ph.D.)

Dissertation Title	PENALIZED, SHRINKAGE, AND PRELIMINARY TEST STRATEGIES IN NONLINEAR AND PROPORTIONAL HAZARD REGRESSION MODELS FOR LOW AND HIGH-DIMENSIONAL DATA
Author	Janjira Piladaeng
Degree	Doctor of Philosophy (Statistics)
Department /Faculty/University	Mathematics and Statistics Faculty of Science and Technology Thammasat University
Dissertation Advisor	Associate Professor Supranee Lisawadi, Ph.D.
Dissertation Co-Advisor	Professor Syed Ejaz Ahmed, Ph.D.
Academic Year	2021

ABSTRACT

In regression analysis with many regressors, it is expected that some may not significantly contribute to predicting the response variable. This is called uncertain prior information (UPI) and may be obtained by variable selection. The use of UPI to produce a submodel or restricted model has received increasing attention in statistical models. In practice, the full, or unrestricted, model may be overfitted, with too many predictors included, when prior information uncertainty proves correct. A submodel including only important regressors may be more practical and feasible, but concerns remain that the submodel may be inappropriate when UPI is incorrect.

The objective of this study was to propose novel estimators that are more efficient in estimation than the classical estimator. In addition, the study also attempted to optimally incorporate the full model and submodel for parameter estimation using preliminary test and shrinkage strategies. This idea will improve regression parameter estimation efficiency, even with uncertain prior information accuracy. Proposed estimators were applied with the Cobb-Douglas, exponential, and monomolecular multiple nonlinear regression models and the Cox proportional hazards regression model (special

chapter) under UPI in low-dimensional and high-dimensional data regimes.

The proposed estimator's performance was compared theoretically by deriving asymptotic distributional quadratic bias and risk under the sequence of local alternatives. In addition, Monte Carlo simulations were conducted to evaluate the numerical proposed estimator performance. Numerical comparisons were also made with penalty estimation strategies: least absolute shrinkage and selection operator (LASSO) and adaptive LASSO. Finally, the proposed estimators were applied to real data examples to verify their usefulness. Regardless of prior information correctness, the proposed estimators were shown to perform significantly better than classical estimators which are severely affected by information uncertainty.

Keywords: asymptotic properties, linear shrinkage, positive-part shrinkage, semiparametric hazards regression model, shrinkage preliminary test

ACKNOWLEDGEMENTS

I wish to express my grateful and profound thanks to the following who have assisted me during my Ph.D. advanced degree program and the writing of this dissertation:

- My advisor, Associate Professor Supranee Lisawadi, for her insightful advice, encouragement, and invaluable support throughout this research.
- Professor Syed Ejaz Ahmed, Dean of the Faculty of Mathematics and Science at Brock University, Canada. His guidance, immense knowledge, and patience have played a pivotal role in the completion of my Ph.D. program.
- Committee chairperson Professor Saowanit Sukparungsee, King Mongkut's University of Technology North Bangkok, and committee members Assistant Professors Patchanok Srisuradetchai and Teerawat Simmachan for their effective advice and valued suggestions which improved my dissertation and enabled me to complete it.

I gratefully acknowledge the financial support provided by the Human Resource Development in Science Project (Science Achievement Scholarship of Thailand, SAST) and Canada-ASEAN Scholarships and Educational Exchanges for Development (SEED). I am deeply grateful to the Royal Irrigation Department of Thailand for real dataset support on the socio-economic situation monitoring and evaluation, Kwae Noy Dam under His Majesty the King's initiation, budget year 2015.

My sincere thanks go to Professor John Winward, who assisted me in proofreading my proposal and research papers, and Assistant Professor Saengla Chaimongkol, who provided steadfast support. I am very thankful to all those who gave me their assistance and encouragement, including the faculty and staff members of the Department of Mathematics and Statistics, Thammasat University. In addition, I would like to thank all my seniors, juniors, and fellow Ph.D. students at the Department of Mathematics and Statistics, Thammasat University, for their companionship during my Ph.D. studies.

Finally, without the unwavering understanding and encouragement of my parents and relatives during the past many years, it would have been impossible for me to complete my studies. I am indebted to them all.

Janjira Piladaeng

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviations	Terms
ADB	Asymptotic distributional bias
ADMSE	Asymptotic distributional mean squared error
ADQB	Asymptotic distributional quadratic bias
ADQR	Asymptotic distributional quadratic risk
ADR	Asymptotic distributional risk
AIC	Akaike information criterion
aLASSO	Adaptive least absolute shrinkage and selection operator
BIC	Bayesian information criterion
CDF	Cumulative distribution function
LASSO	Least absolute shrinkage and selection operator
LS	Linear shrinkage
MSE	Mean squared error
NSI	Non-sample information
PH	Proportional hazards
PT	Preliminary test
RE	Restricted estimator
RMSE	Relative mean squared error
RMSPE	Relative mean squared prediction error
S	Shrinkage
S^+	Positive-part shrinkage
SI	Sample information
SP	Shrinkage preliminary test
UE	Unrestricted estimator
UPI	Uncertain Prior information

LIST OF SYMBOLS

Symbols	Terms
$^\top$	Transpose of vector/matrix
$\in$	Belong to
$\xrightarrow{D}$	Converges in distribution to
$\xrightarrow{P}$	Converges in probability to
$H_\nu(\cdot;\Delta)$	Cumulative distribution function of non-central chi-square with ν degrees of freedom and non-centrality parameter Δ
$\mathbb{P}$	Probability
$\mathbb{E}$	Expectation of a random variable
$\mathbb{V}$	Variance of a random variable
Cov	Covariance of two random variables
$\mathbb{R}^{p_2}$	p_2-dimensional real vector space
m	Bootstrap sample size
n	Sample size
p	Number of all parameters
p_1	Number of active parameters
p_2	Number of inactive parameters
p_s	Number of parameters with strong signal
p_w	Number of parameters with weak-to-moderate signal
p_n	Number of parameters with no signal
k	Number of all predictors
k_1	Number of active predictors
k_2	Number of inactive predictors
k_s	Number of predictors with strong signal
k_w	Number of predictors with weak-to-moderate signal
k_n	Number of predictors with no signal

$\boldsymbol{\beta}$	Vector of unknown regression coefficients associated with all p parameters
$\boldsymbol{\beta}_1$	Vector of unknown regression coefficients associated with p_1 active parameters
$\boldsymbol{\beta}_2$	Vector of unknown regression coefficients associated with p_2 inactive parameters
$\boldsymbol{\beta}_s$	Vector of unknown regression coefficients associated with p_s parameters with strong signal
$\boldsymbol{\beta}_w$	Vector of unknown regression coefficients associated with p_w parameters with weak-to-moderate signal
$\boldsymbol{\beta}_n$	Vector of unknown regression coefficients associated with p_n parameters with no signal
H_0	Null hypothesis
H_1	Alternative hypothesis
$\{K_n\}$	Sequence of local alternatives
$\mathcal{I}$	Information matrix
$\boldsymbol{I}$	Identity matrix
$\boldsymbol{W}$	Positive semi-definite weighting matrix
tr	Trace
$\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$	Unrestricted estimator of $\boldsymbol{\beta}_1$
$\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$	Restricted estimator of $\boldsymbol{\beta}_1$
$\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$	Linear shrinkage estimator of $\boldsymbol{\beta}_1$
$\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$	Preliminary test estimator of $\boldsymbol{\beta}_1$
$\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$	Shrinkage preliminary test estimator of $\boldsymbol{\beta}_1$
$\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$	Shrinkage estimator of $\boldsymbol{\beta}_1$
$\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$	Positive-part shrinkage estimator of $\boldsymbol{\beta}_1$
$\widehat{\boldsymbol{\beta}}_1^{\mathrm{LASSO}}$	Least absolute shrinkage and selection operator estimator of $\boldsymbol{\beta}_1$
$\widehat{\boldsymbol{\beta}}_1^{\mathrm{aLASSO}}$	Adaptive least absolute shrinkage and selection operator estimator of $\boldsymbol{\beta}_1$
α	Significance level

π	Shrinkage intensity
Λ_n	Wald statistic
$\mathscr{L}_n$	Likelihood ratio test statistic
λ_α	α-level critical value of the chi-square distribution
$I(\cdot)$	Indicator function
$ch_{\min}(\cdot)$	Smallest eigenvalues of matrix
$ch_{\max}(\cdot)$	Largest eigenvalues of matrix

CHAPTER 1
INTRODUCTION

1.1 Problem Statement and Importance of the Study

In statistical inferences, the use of prior information on some or all of the parameters in the model often improves inference procedures. Regression analysis is a statistical technique used to model and analyze the relationships between variables. In many cases, however, the investigator begins the statistical analysis with both sample information (SI) and non-sample information (NSI).

Non-sample information (or prior information on the parameters) may be derived from theoretical arguments or from past experience, or else a variable selection method may be used. The prior information may be certain or uncertain. When known prior information is incorporated into the model in the form of a constraint, this gives rise to a restricted model (submodel) that can be used to increase the precision of estimates and to reduce computational load. A restricted model therefore gives a superior statistical analysis when compared to an unrestricted model (full model).

In an unrestricted model, if uncertain prior information (UPI) is used as a constraint on the parameters, we may consider the uncertain constraints as nuisances or inactive parameters in the statistical analysis. For this reason, an unrestricted model that considers all parameters leads to overfitting, as too many nuisance parameters are included. If it is a *priori* known or suspected that some regressors do not significantly contribute to predicting the response variable, a restricted model excluding these regression coefficients and containing only active parameters may be adequate. It is often true that a restricted estimator provides a considerable improvement over an unrestricted estimator, and can address overfitting. However, if the UPI is incorrect, we may encounter underfitting. This means that the estimators based on a restricted model may be biased and inefficient, and the model may not truly representative of the data.

In addition, we are surrounded by a world of information, especially big data. For years, big data analytics has been a hot topic in computer science and statistics, and a focus for many researchers. Compared to ordinary datasets, or low-dimensional (LD) datasets, big data represents large and diverse sets of information, otherwise known as high-dimensional (HD) data, which may be structured, semi-structured or unstructured. High-dimensional statistics focus on datasets where the number of features or independent variables (k) is larger than the number of observations (n). Since classical theory and methodology become unreliable or impossible to compute if k exceeds n, big data analysis requires methods outside of traditional frameworks.

For example, many methods have been developed to estimate regression coefficients in the high-dimensional linear regression model, such as the least absolute shrinkage and selection operator (LASSO) and adaptive LASSO (aLASSO), as well as other penalty methods. However, results from using different variable selection methods may have different subsets of predictors, leading to the consideration of two models. Using the LASSO method, a model can include predictors that produce strong signals and possibly some predictors with weak-to-moderate signals, which is the main cause of overfitting problems. Conversely, the aLASSO method establishes another model that includes predictors with strong signals while leaving out predictors with weak-to-moderate signals, resulting in underfitting problems. Therefore, the uncertainty in the correctness of variable selection results makes the estimation based on both competitive models inefficient, as with LD data.

To eliminate the uncertainty of the prior information, a preliminary test (PT) procedure is used that removes the inactive parameters. The choice of an unrestricted or restricted model is based on the validity of the UPI. The result is a compromise between the two extremes. The choice between unrestricted or restricted models can also be made using linear shrinkage (LS), shrinkage preliminary test (SP), shrinkage (S), or positive-part shrinkage (S) estimations. These are extensions of preliminary test estimation. Under UPI, the full parameter vector $\boldsymbol{\beta}$ can be divided into $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$. We are interested in the estimation of the active parameter sub-vectors $\boldsymbol{\beta}_1$ when information about $\boldsymbol{\beta}_2$ is available, that is $\boldsymbol{\beta}_2 = \mathbf{0}$.

Moreover, there have been many studies of UPI in linear and generalized linear regressions. However, in many situations, a linear regression model cannot adequately represent the relationships between variables. In these cases, a more suitable approach is a nonlinear regression model, and the parameters appear in the form of an exponential or are multiplied or divided by other parameters. Such models are an important tool in several research areas, including agriculture, forestry, econometrics, and biology.

In the same way, data that measure lifetime or the length of time until the occurrence of an event are also attractive to studying the UPI. This data is called lifetime, failure time, or survival data, and it is an essential topic in many areas, including the biomedical, engineering, and social sciences. A statistical technique for analyzing the likely duration until the occurrence of an event of interest is called survival analysis. One of the widely used functions in survival analysis is the hazards function, which is a measure of risk at time t. Unfortunately, this analysis is usually a difficult process due to censoring. This problem causes a lack of information; therefore, the ordinary regression models cannot be used for survival data. As a result, the Cox proportional hazards (PH) regression model is the most commonly applied hazard model in survival data.

In this dissertation, we will estimate parameters in nonlinear regression and Cox PH regression (special chapter) models under UPI using the above estimation strategies in the context of low-dimensional and high-dimensional regimes. The unrestricted (overfitted) model contains all independent variables, and the restricted (underfitted) model is from a variable selection using the Akaike information criterion (AIC) and Bayesian information criterion (BIC) for low-dimensional data. We also applied LASSO and aLASSO to build the overfitted and underfitted models, respectively, for high-dimensional data.

The organization of this dissertation is as follows:

Chapter 1 presents the problem statement and importance of the study, research objectives, and the scope and benefits of the study.

Chapter 2 reviews the theoretical background of nonlinear regression models, the development of estimation strategies, and relevant literature on both low- and high-dimensional data.

Chapter 3 introduces the nonlinear least squares estimator as well as various suggested estimators along with penalized estimators, and presents the concepts of asymptotic distributional bias (ADB), asymptotic distributional quadratic bias (ADQB), asymptotic mean squared error matrix (AMSEM), and asymptotic distributional quadratic risk (ADQR). The theorems to explain the theoretical aspects, the dimensionality of data, and the measures of estimator performance are also described here.

Chapter 4 demonstrates the derived asymptotic properties of the suggested estimators, and their performance is compared. Numerical analyses are conducted through Monte Carlo simulations, and study results are discussed. Finally, real data examples are analyzed to display the practicality of the suggested and penalized estimators.

Chapter 5 discusses the special topic of parameter estimations in the Cox proportional hazards regression model under the uncertainty of prior information in both low- and high-dimensional settings. The asymptotic properties, Monte Carlo simulations, and real data examples for the Cox model are also represented here.

Finally, Chapter 6 presents conclusions and recommendations for future research work.

1.2 Research Objectives

The objectives of this dissertation are:

1.2.1 To propose preliminary test and shrinkage estimation strategies for the Cobb-Douglas, exponential, and monomolecular nonlinear models, allowing parameter estimation of $\boldsymbol{\beta}$ under UPI.

1.2.2 To derive and assess the asymptotic properties of the proposed estimators from their asymptotic distributional quadratic bias (ADQB) and asymptotic distributional quadratic risk (ADQR).

1.2.3 To compare the performance of the proposed estimators using Monte Carlo simulations and application to real data in low-dimensional and high-dimensional statistical models.

1.3 Scope of the Study

The aim of this research is to propose estimators for nonlinear models under uncertain prior information.

1.3.1 The evaluation of performance is divided into two parts:

1.3.1.1 A theoretical part, in which the performance of the proposed estimators will be evaluated by comparing the ADQB and ADQR under local alternatives.

1.3.1.2 A computational part, using Monte-Carlo simulations and application to real data in both low- and high-dimensional contexts. Performance will be evaluated by the relative mean squares error (RMSE) in the Monte Carlo simulations and by the relative mean squares prediction error (RMSPE) in the real data example.

1.3.2 The R program will be used in the computational part of this study.

1.4 Benefits of the Study

The benefits of this research are:

1.4.1 The introduction of a robust process for overcoming overfitting or underfitting in parameter estimation under UPI.

1.4.2 The provision of guidelines for applying the proposed estimators to other nonlinear and proportional hazards regression models in which uncertain subspace information is available.

CHAPTER 2
THEORIES AND RESEARCH OF RELEVANT

In this chapter, we review theoretical and research studies related to preliminary test and shrinkage strategies for nonlinear regression models.

2.1 Theoretical Background

Two theories underpin this research. The first concerns general nonlinear regression models and the second the specific nonlinear models considered in the dissertation.

2.1.1 Nonlinear Regression Models

Nonlinear regression is a form of regression analysis. The response variable is modeled using a nonlinear function, and depends on one or more predictors. Any model that is nonlinear in the unknown parameters will be called a nonlinear regression model.

Let $y_1, y_2, ..., y_n$ be independent variables, and $\boldsymbol{x}_i = (x_{i1}, x_{i2}, ..., x_{ik})^\top$ be a $k \times 1$ vector of predictors for the ith subject, where $i = 1, 2, ..., n$. The general form of the nonlinear regression model is

$$y_i = f(\boldsymbol{x}_i, \boldsymbol{\beta}) + \varepsilon_i, \quad i = 1, 2, ..., n, \tag{2.1}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_j)^\top$ is a $p \times 1$ vector of regression coefficients and $\boldsymbol{\varepsilon}$ represents independent and identically distributed random errors. The distributional assumption is that $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n)^\top$ has a cumulative distribution function $F(\boldsymbol{\varepsilon})$ with mean zero and variance $\sigma^2\boldsymbol{I}$, where σ^2 is finite and $\boldsymbol{I}$ is an $n \times n$ identity matrix. The estimators of β_j are calculated by minimizing the sum of squares error function:

$$\begin{aligned} S(\boldsymbol{\beta}) &= \sum_{i=1}^{n} [y_i - f(\boldsymbol{x}_i, \boldsymbol{\beta})]^2 \\ &= [\boldsymbol{y} - f(\boldsymbol{x}, \boldsymbol{\beta})]^\top [\boldsymbol{y} - f(\boldsymbol{x}, \boldsymbol{\beta})]. \end{aligned} \tag{2.2}$$

To find the nonlinear least squares estimators, the derivative of $S(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$ is obtained by solving the normal equation:

$$\frac{\partial S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} [y_i - f(\boldsymbol{x}_i, \boldsymbol{\beta})] \left[\frac{\partial f(\boldsymbol{x}_i, \boldsymbol{\beta})}{\partial \beta_j} \right] = \mathbf{0}, \quad j = 1, 2, ..., p. \tag{2.3}$$

It is impossible to attain a closed form solution to the least squares estimate of the parameters by solving the p normal equations. An iterative method must be applied to minimize the sum of squares. In this dissertation, the iterative Gauss-Newton method is used.

2.1.1.1 Linearization and the Gauss-Newton Method

Some nonlinear regression problems can be linearized by an appropriate transformation of the model formulation. However, the use of a nonlinear transformation requires caution. When the data are transformed, the data units and the influence of the data values will be changed, as will the error structure of the model and the interpretation of any inferential results, as the regression coefficients will need to be interpreted with respect to the transformed scale. There is no straightforward way of back-transforming them to values that can be interpreted according to the original scale.

It should also be noted that the parameters of a linearized model are often not as interesting or important as the original parameters. In physical and chemical models, the original parameters usually have a physical meaning, e.g., rate constants, so that estimates for these parameters are still required. Therefore, given the availability of efficient nonlinear algorithms, the usefulness of linearization is somewhat diminished (Seber & Wild, 2003, p. 17). However, the linearization transformations can be of use in providing good initial estimates for the iterative techniques used with nonlinear regression (Dotson, 1966).

As in linear regression, nonlinear regression is based on determining the values of the model parameters that minimize the sum of the squares of residuals. However, in this case, iterative methods are required. Here we apply a Gauss-Newton iteration method that uses Taylor series expansion to express the original nonlinear equation in approximately linear form. This is not the same as linearization transformation since the original equation and the associated data are not transformed. Least squares theory can be used for new parameter estimation by minimizing the residual.

This method will establish an iterative technique, but we need to make an initial guess for the parameters. The initial parameters are used as a center of expansion, then new parameters are estimated and used as centers of expansion in subsequent stages. If we carry out a Taylor series expansion of $f(\boldsymbol{x}, \boldsymbol{\beta})$ about the point $\widehat{\boldsymbol{\beta}}_{(0)}$, where $\widehat{\boldsymbol{\beta}}_{(0)} = (\widehat{\beta}_{1(0)}, \widehat{\beta}_{2(0)}, ..., \widehat{\beta}_{p(0)})$ is usually an initial estimate or a set of starting values for the model parameters $\boldsymbol{\beta}$, and stop after the first derivatives, we have the linear Taylor series expansion

$$f(\boldsymbol{x}_i, \boldsymbol{\beta}) = f(\boldsymbol{x}_i, \widehat{\boldsymbol{\beta}}_{(0)}) + \sum_{j=1}^{p} \left[\frac{\partial f(\boldsymbol{x}_i, \boldsymbol{\beta})}{\partial \beta_j}\right]_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}_{(0)}} (\beta_j - \widehat{\beta}_{j(0)}). \tag{2.4}$$

Substituting Equation (2.4) in Equation (2.1), we get

$$y_i = f(\boldsymbol{x}_i, \widehat{\boldsymbol{\beta}}_{(0)}) + \sum_{j=1}^{p} \left[\frac{\partial f(\boldsymbol{x}_i, \boldsymbol{\beta})}{\partial \beta_j}\right]_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}_{(0)}} (\beta_j - \widehat{\beta}_{j(0)}) + \varepsilon_i. \tag{2.5}$$

If we set $f_i^{(0)} = f(\boldsymbol{x}_i, \widehat{\boldsymbol{\beta}}_{(0)})$, $z_i^{(0)} = y_i - f_i^{(0)}$, $D_{ij}^{(0)} = \left[\frac{\partial f(\boldsymbol{x}_i, \boldsymbol{\beta})}{\partial \beta_j}\right]_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}_{(0)}}$, and $\theta_j^{(0)} = (\beta_j - \widehat{\beta}_{j(0)})$, we can rewrite Equation (2.5) as

$$\begin{aligned} y_i &= f_i^{(0)} + \sum_{j=1}^{p} D_{ij}^{(0)} \theta_j^{(0)} + \varepsilon_i \\ y_i - f_i^{(0)} &= \sum_{j=1}^{p} D_{ij}^{(0)} \theta_j^{(0)} + \varepsilon_i \\ z_i^{(0)} &= \sum_{j=1}^{p} D_{ij}^{(0)} \theta_j^{(0)} + \varepsilon_i \end{aligned} \tag{2.6}$$

In matrix notation, Equation (2.6) is

$$\boldsymbol{z}_{(0)} = \boldsymbol{D}_{(0)} \boldsymbol{\theta}_{(0)} + \boldsymbol{\epsilon}_{(0)}, \tag{2.7}$$

where $\boldsymbol{z}_{(0)}$, $\boldsymbol{\theta}_{(0)}$, $\boldsymbol{D}_{(0)}$, and $\boldsymbol{\epsilon}_{(0)}$ are $n \times 1$, $n \times p$, $p \times 1$, and $n \times 1$ matrices, respectively. This is a linear regression model with unknown parameters $\boldsymbol{\theta}_{(0)}$. The sum of squares error function is then

$$SS(\boldsymbol{\beta}) = (\boldsymbol{z}_{(0)} - \boldsymbol{D}_{(0)} \boldsymbol{\theta}_{(0)})^\top (\boldsymbol{z}_{(0)} - \boldsymbol{D}_{(0)} \boldsymbol{\theta}_{(0)}), \tag{2.8}$$

where $\boldsymbol{z}_{(0)} = \boldsymbol{y} - \boldsymbol{f}_{(0)}$. Note the difference between the sum of squares $S(\boldsymbol{\beta})$ in Equation (2.2), where the appropriate nonlinear model is used, and the sum of squares $SS(\boldsymbol{\beta})$

in Equation (2.8), where the approximating linear expansion of the model is employed (Draper and Smith, 1998, p. 509). To find the least squares estimates, we must differentiate Equation (2.8) with respect to $\boldsymbol{\theta}_0$ and the normal equation, as shown below

$$\frac{\partial SS(\boldsymbol{\beta})}{\partial \boldsymbol{\theta}_{(0)}} = \frac{\partial}{\partial \boldsymbol{\theta}_{(0)}} \left[(\boldsymbol{z}_{(0)} - \boldsymbol{D}_{(0)}\boldsymbol{\theta}_{(0)})^\top (\boldsymbol{z}_{(0)} - \boldsymbol{D}_{(0)}\boldsymbol{\theta}_{(0)}) \right] = 0. \tag{2.9}$$

Therefore,

$$\begin{aligned} \boldsymbol{D}_{(0)}^\top \boldsymbol{D}_{(0)} \boldsymbol{\theta}_{(0)} &= \boldsymbol{D}_{(0)}^\top \boldsymbol{z}_{(0)}. \\ \underbrace{(\boldsymbol{D}_{(0)}^\top \boldsymbol{D}_{(0)})^{-1} \boldsymbol{D}_{(0)}^\top \boldsymbol{D}_{(0)}}_{\mathrm{I}_p} \boldsymbol{\theta}_{(0)} &= (\boldsymbol{D}_{(0)}^\top \boldsymbol{D}_{(0)})^{-1} \boldsymbol{D}_{(0)}^\top \boldsymbol{z}_{(0)} \\ \therefore \quad \widehat{\boldsymbol{\theta}}_{(0)} &= (\boldsymbol{D}_{(0)}^\top \boldsymbol{D}_{(0)})^{-1} \boldsymbol{D}_{(0)}^\top (\boldsymbol{y} - \boldsymbol{f}_{(0)}), \end{aligned} \tag{2.10}$$

where $\widehat{\boldsymbol{\theta}}_{(0)}$ is the least squares estimator of $\boldsymbol{\theta}_{(0)}$. Because $\boldsymbol{\theta}_{(0)} = \boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}_{(0)}$, then $\widehat{\boldsymbol{\theta}}_{(0)} = \widehat{\boldsymbol{\beta}}_{(1)} - \widehat{\boldsymbol{\beta}}_{(0)}$ and we can use $\widehat{\boldsymbol{\beta}}_{(1)} = \widehat{\boldsymbol{\beta}}_0 + \widehat{\boldsymbol{\theta}}_{(0)}$ as a revised best estimate of the unknown parameters $\boldsymbol{\beta}$. We call $\widehat{\boldsymbol{\theta}}_{(0)}$ the vector of increments.

We can now place the revised parameter estimates $\widehat{\boldsymbol{\beta}}_{(1)}$ in (2.5) in the same role originally played by the starting values $\widehat{\boldsymbol{\beta}}_{(0)}$, then apply the procedure described above, but replacing all zero subscripts with one. This will yield a second set of revised estimates, $\widehat{\boldsymbol{\beta}}_{(2)}$, and so on. In general, at the qth of these iterations we have

$$\begin{aligned} \widehat{\boldsymbol{\beta}}_{(q+1)} &= \widehat{\boldsymbol{\beta}}_{(q)} + \widehat{\boldsymbol{\theta}}_{(q)} \\ &= \widehat{\boldsymbol{\beta}}_{(q)} + (\boldsymbol{D}_{(q)}^\top \boldsymbol{D}_{(q)})^{-1} \boldsymbol{D}_{(q)}^\top \boldsymbol{z}_{(q)} \\ &= \left(\boldsymbol{D}_{(q)}^\top \boldsymbol{D}_{(q)} \right)^{-1} \boldsymbol{D}_{(q)}^\top \boldsymbol{F}_{(q)}. \end{aligned} \tag{2.11}$$

where $\boldsymbol{F}_{(q)} = \boldsymbol{D}_{(q)}\widehat{\boldsymbol{\beta}}_{(q)} + \boldsymbol{z}_{(q)}$. The iterative procedure continues until convergence, that is until there is little meaningful change in the estimates of the parameters. Commonly, the convergence criteria are based on

$$\left| \frac{\widehat{\beta}_{j(q+1)} - \widehat{\beta}_{j(q)}}{\widehat{\beta}_{j(q)}} \right| < \xi, \tag{2.12}$$

where ξ is some small number, e.g., 10^{-6}. At each stage of the iteration, the residual sum of squares is measured to ensure that its value has actually reduced.

The error variance (σ^2) is estimated from the residual mean square ($\widehat{\sigma}^2$) when the estimation process converges to a final vector of parameter estimate $\widehat{\boldsymbol{\beta}}$. The residual mean square ($\mathrm{MS}_{\mathrm{res}}$) is

$$\widehat{\sigma}^2 = \frac{\left[\boldsymbol{y} - f(\boldsymbol{x}, \widehat{\boldsymbol{\beta}})\right]^\top \left[\boldsymbol{y} - f(\boldsymbol{x}, \widehat{\boldsymbol{\beta}})\right]}{n - p} \tag{2.13}$$

where p is the number of all parameters. The asymptotic (large-sample) covariance matrix of the parameter vector $\widehat{\boldsymbol{\beta}}$ can be estimated by

$$\mathrm{Var}(\widehat{\boldsymbol{\beta}}) = \mathbb{V}(\widehat{\boldsymbol{\beta}}) = \widehat{\sigma}^2(\boldsymbol{D}^\top \boldsymbol{D})^{-1}, \tag{2.14}$$

where, $\boldsymbol{D}$ is the matrix of partial derivatives evaluated at the final-iteration least squares parameter estimates $\widehat{\boldsymbol{\beta}}$. More detailed information about the nonlinear regression model can be found in Bates and Watts (1988), Draper and Smith (1998), Seber and Wild (2003), and Myers et al. (2010), among others.

2.1.2 Nonlinear Models of Interest

From the multiple nonlinear models available, we chose the Cobb-Douglas, exponential, and monomolecular models, as these are used in many research areas. We discuss these models below.

2.1.2.1 Cobb-Douglas Model

The Cobb-Douglas function is perhaps the most widely-used model in economics. It has been applied to econometric analysis of utility and production functions in growth, development, macroeconomics, public finance, labour, and many other areas of application. In unrestricted form, the Cobb-Douglas function can be written as

$$f(\boldsymbol{x}) = A \prod_{j=1}^{p-1} x_j^{a_j}, \tag{2.15}$$

where A is an efficiency parameter, a_j is the elasticity of $f(\boldsymbol{x})$ with respect to x_j, and $\boldsymbol{x}$ is confined to $\Re^+$. Here, x_j may refer to goods consumed, goods produced, etc. More detailed information can be found in Durlauf and Blume (2008, pp. 862–863). Many studies have applied the Cobb-Douglas model, including Hossain and Al-Amri (2010), Yasar et al. (2012), Hossain et al. (2012), Cheng and Han (2013), etc.

2.1.2.2 Exponential Model

The simplest model considered here is generally called the exponential model, and is one of the most widely-used nonlinear models. The exponential model is often used in growth curve analysis and applied to exponential growth and decay. Examples include investment growth, radioactive decay, and bacterial growth. Studies using this model include Annamalai (2010), Abichou et al. (2012), Archontoulis and Miguez (2015), Wang et al. (2017), etc. This growth model is represented by a basic differential equation:

$$\frac{dy}{dt} = ry, \tag{2.16}$$

in which r is a rate parameter. The integrated form of Equation (2.16) is written as

$$y = y_0 e^{rt}, \tag{2.17}$$

where y_0 is a constant of integration. More information about the exponential model can be found in Madden and Campbell (1990, pp. 190–191) and Pal (2016, p. 112).

2.1.2.3 Monomolecular Model

The monomolecular model is an alternative to the exponential model. The monomolecular model is one of the basic growth models and was originally derived in physical chemistry, where it represents monomolecular chemical reactions of the first order. It has been used to elucidate several other phenomena including cell expansion, the response of crops to nutrients, and animal growth. For plant growth and nutrient supply, this model is called the Mitscherlich or Mitscherlich-Bray model. The rate equation is written as

$$\frac{dy}{dt} = r(K - y), \tag{2.18}$$

where K is a parameter representing a maximum, or $K = y_{max}$. The integrated form of Equation (2.18) is

$$y = K(1 - Be^{-rt}), \tag{2.19}$$

in which B is a constant of integration. More information about this model can be found in Madden and Campbell (1990, pp. 191–193) and Gowariker et al. (2009, p. 104). Studies that have applied this model include Ware et al. (1982), Rajarshi (1995), Fekedulegn et al. (1999), Khamis et al. (2005), etc.

2.2 Literature Review

In this section, we discuss the development of estimation strategies and review the relevant literature.

2.2.1 Development of Estimation Strategies

One estimation strategy is the preliminary test (PT) or pretest, which was introduced by Bancroft (1944). This test statistic is used to decide the estimator, based on either the unrestricted or restricted model. Bancroft applied the preliminary test strategy to two seemingly unrelated problems: a data pooling problem and simultaneous model selection and pretest estimation in a regression model.

The shrinkage (S) strategy or James-Stein type shrinkage strategy was presented in Stein (1956) and James and Stein (1961), and addressed the admissibility of the sample mean vector in a multivariate normal distribution. They concluded that the estimator has a smaller risk than the ordinary least squares (OLS) estimator for $p \geq 3$ when the quadratic loss function is considered. This technique may be regarded as a smoothed version of the preliminary test estimation strategy.

An alternative strategy was that of Thompson (1968). He considered the advisability of shrinking a usual estimator for θ, say $\widehat{\theta}$, towards the approximation θ_0 by multiplying it by a shrinking factor (c). The estimator was constructed as linear combination of other estimators along with a shrinkage coefficient, and can be interpreted as a linear shrinkage (LS) estimator.

The shrinkage preliminary test (SP) strategy was introduced by Ahmed (1992). He concluded that the shrinkage pretest estimator is an improved version of the preliminary test estimator with respect to the size of alpha for estimating the mean vector of a multivariate normal distribution.

As a penalty strategy, Tibshirani (1996) proposed the least absolute shrinkage and selection operator (LASSO), which uses an L_1-norm in linear regression. This strategy forces some coefficients to shrink to exactly zero, providing simultaneous parameter selection and estimation. Yang et al. (2015) studied sparse nonlinear regression models. They proposed an efficient algorithm for estimating the parameter by solving the L_1-regularized nonlinear least-squares problem.

Another penalty strategy is adaptive LASSO (aLASSO) which was proposed by Zou (2006). This is a new version of LASSO where adaptive weights were used to penalize different coefficients in the L_1 penalty. Zou (2006) also showed that adaptive LASSO enjoys the oracle properties; namely, it performs as well as if the true underlying model were given in advance. Furthermore, the aLASSO can be solved by the same efficient algorithm for solving the LASSO.

2.2.2 Relevant Literature

In the past few decades, many studies have examined the use of preliminary test and shrinkage estimation strategies under uncertain prior information in linear regression, partial linear regression, nonlinear regression (i.e., demand-for-money), and generalized linear regression (i.e., Poisson, logistic, multinomial logistic, negative binomial) models. Moreover, there are also numerous studies on regression models (i.e., exponential, Cox PH, Weibull, lognormal, exponentiated Weibull) with censored or time-to-event data.

Ahmed (1997) introduced estimation into preliminary test and shrinkage strategies. For a linear regression model, the selected estimators were shrinkage restricted, standard preliminary test, shrinkage preliminary test, standard shrinkage, improved preliminary test, and positive-part shrinkage. The proposed estimators were applied to regression coefficient estimation when prior information was regarded as uncertain. Ahmed also made a comprehensive study of the asymptotic distributional quadratic bias and risk, and compared the performance of the estimators. He concluded that the positive-part shrinkage estimator was more efficient than the usual shrinkage estimator, and that pretest estimators had good control over the magnitude of the risk.

Ahmed and Saleh (1999) and Khan (2002) considered parameter estimations in an exponential model with censoring under non-sample information (NSI). The James-Stein (JS) type and positive-part JS-type estimators were proposed in Ahmed and Saleh (1999) and a preliminary test estimator was proposed in Khan (2002). The performance of these estimators was compared with that of the maximum likelihood estimators using quadratic distribution risk under local alternatives. The performance of these estimators heavily depends on the quality of the UPI. The positive-part JS-type estimator dominated the usual JS-type, and both performed well relative to the unrestricted estimator of the parameter vector in the entire parameter space.

Sapra (2003) studied preliminary test maximum likelihood estimation in a Poisson regression model. He also compared the performance of the pretest estimator with the unrestricted and restricted estimators, using Monte Carlo simulations. The results showed that the pretest estimator dominated the unrestricted estimator. The restricted estimator was outperformed by both the pretest and unrestricted estimators when the values of the regression coefficients were close to the restrictions. In contrast, the pretest and unrestricted estimators eventually converged and dominated the restricted estimator when the regression coefficients were far from the restrictions.

Raheem and Ahmed (2011) examined the preliminary test and shrinkage strategies in a multiple regression model through a Monte Carlo study and application to real datasets. The study confirmed the asymptotic properties of the estimators. Ahmed and Raheem (2012) also compared the performance of shrinkage estimators with the absolute penalty estimators LASSO, adaptive LASSO, and smoothly clipped absolute deviation (SCAD), in multiple linear regression. The Monte Carlo simulation reconfirmed the dominance of S^+ over LASSO for moderate to large values of inactive parameters. However, SCAD and adaptive LASSO outperformed S^+ when the number of inactive parameters was large relative to the sample size.

Ahmed and Nicol (2012) applied the preliminary test, JS type, and positive-part JS type estimators to large sample sizes in a nonlinear regression. They used a demand-for-money model for when it is *a priori* suspected that the coefficients may be restricted to a subspace. In simulations, the positive-part James-Stein shrinkage estimator was superior to the other estimators. They also compared the performance in terms of asymptotic distributional quadratic bias and risk. The results showed that the preliminary test strategy dominated the JS type and positive-part JS type and also outperformed the unrestricted estimator when the uncertain prior information was correct.

Raheem et al. (2012) proposed shrinkage semiparametric estimation based on the Stein rule in the context of a partially linear regression model with a nonparametric component based on the B-spline basis function. In Monte Carlo simulations, the proposed estimators were compared with an absolute penalty estimator. LASSO and aLASSO were implemented for simultaneous model selection and parameter estimation. They reported that the performance of shrinkage and absolute penalty estimators may vary depending on the number of active parameters.

Ahmed et al. (2012), Hossain and Howlader (2017), and Hossain and Khan (2020) addressed the problem of estimating parameters in Weibull regression, lognormal regression, and exponentiated Weibull regression models, respectively, for time-to-event data (censored data) involving many predictors, some of which may not have any influence on the response of interest. The shrinkage and positive-part shrinkage estimation methods were introduced in their study. The properties of these estimators used the notion of asymptotic distributional bias and risk. Moreover, the LASSO estimation strategy was suggested in Ahmed et al. (2012), and both LASSO and adaptive LASSO estimation methods were recommended in Hossain and Howlader (2017). Furthermore, the relative performance of both shrinkage estimators was compared with that of the classical maximum likelihood estimators and penalty estimators. Their studies suggested that the shrinkage estimators perform well in terms of statistical efficiency. Finally, real-life data examples were used to illustrate the performance of their suggested estimators.

Hossain et al. (2014) studied the shrinkage, positive-part shrinkage, and LASSO strategies for simultaneous model selection and parameter estimation in a multinomial regression model when some of the predictors may or may not be active for the response of interest. The asymptotic distributional bias and risk were used to compare the performance of the shrinkage and positive-part shrinkage estimators. In the Monte Carlo simulations and real data example they compared the performance of all three estimators. The results showed that, when there were many inactive parameters, all proposed estimators were superior to the unrestricted estimator, and that both shrinkage estimators dominated LASSO.

Hossain and Ahmed (2014) studied the shrinkage and positive-part shrinkage estimation strategies in the Cox proportional hazards regression model for when it is suspected that some of the parameters may be restricted to a subspace. They developed the statistical properties of the shrinkage estimators including asymptotic distributional biases and risks. A Monte Carlo simulation and two real datasets were used to evaluate the performance of two shrinkage estimators, and two penalty estimators were also considered. Their result showed that the performance of the shrinkage estimators was superior to that of the classical estimators and was comparable to that of the penalty estimators when the number of irrelevant predictors in the model was relatively large.

Lisawadi et al. (2016) studied parameter estimation in the logistic regression model under subspace information, using the linear shrinkage, preliminary test, and shrinkage preliminary test strategies. They compared the performance of the proposed estimators by the asymptotic distributional bias and risk. The proposed estimators dominated the traditional maximum likelihood estimator. Simulations were run to evaluate the performance of the proposed estimators in terms of relative efficiency. The results were in strong agreement with the asymptotic results.

Reangsephet et al. (2018) proposed estimators based on the preliminary test and Stein-type strategies for parameter estimation in a logistic regression model. The LASSO and ridge regression were also considered. They recommended the positive-part Stein-type shrinkage estimator because its performance was robust regardless of the reliability of the subspace information. Reangsephet et al. (2019) introduced a pretest estimator in a negative binomial model for parameter estimation when subspace information is unknown. Monte Carlo simulations evaluated the performance of the proposed estimator and demonstrated that estimation based on a preliminary test strategy is appropriate.

Moreover, Phukongtong et al. (2020) also addressed the problem of estimating regression coefficients for partially linear models, where the nonparametric component is approximated using smoothing splines. The pretest and shrinkage estimation strategies were proposed. The results found that the positive-part shrinkage estimator was shown to be more efficient than the other estimators, and also superior to penalty estimators such that LASSO, aLASSO, and elastic net (Enet).

In a high-dimensional situation, Ahmed and Yüzbaşı (2016) suggested penalty and shrinkage estimation procedures for sparse multiple linear models when some of the predictors may have a very weak influence on the response of interest. They applied penalty methods, i.e., LASSO, aLASSO, and SCAD, to provide two submodels selected by such methods, which are the overfitted and underfitted models. The two models were combined to improve estimation and prediction performance, that is the positive part of the shrinkage estimation. Monte Carlo simulation studies appraised the performance of the penalty and positive-part shrinkage estimators. The shrinkage estimation strategy in the positive part provided better results whether or not the resulting submodel was appropriate.

Further, Yüzbaşı et al. (2017) also applied shrinkage strategies in the form of a family of double shrunken estimators (FS) to estimate regression coefficients efficiently for the high-dimensional sparse multiple regression model. The working models provided by LASSO and aLASSO variable selection methods were called overfitted and underfitted models. Monte Carlo simulation studies and real examples confirmed the superior performance of the double shrunken estimators in the high-dimensional regression model after the screening of variables.

Gao et al. (2017) studied the post-selection shrinkage strategy to improve the prediction performance of a selected subset model. Their post-selection shrinkage estimator combined a post-selection weighted ridge estimator with a post-selection least squares estimator. Under an asymptotic distributional quadratic risk criterion, the prediction performance was explored analytically, and they reported that the post-selection shrinkage estimator outperformed the post-selection weighted ridge estimator.

More recently, Reangsephet et al. (2020), Lisawadi et al. (2021), and Reangsephet et al. (2021) have applied linear shrinkage, preliminary test, shrinkage preliminary test, Jame-Stein shrinkage, and PMLE strategies to parameter estimation when subspace information on the parameters is available to logistic regression, negative binomial regression, and Poisson regression models. The results were similarly consistent with those of previous studies, suggesting the use of estimators based on the Jame-Stein shrinkage procedure. Moreover, they also show that the shrinkage pretest strategy was suitable for post-selection parameter estimation if the condition of Jame-Stein shrinkage procedure was not satisfied.

The literature evaluating preliminary test and shrinkage strategies has been extended to nonlinear regression and Cox PH regression models in the low-dimensional setting. However, there are no complete reviews, especially of the linear shrinkage and preliminary test estimation strategies. Still, previous estimators have not been applied in the context of nonlinear regression and Cox PH regression models for the high-dimensional setting. Further, the LASSO and aLASSO estimators have been proposed for linear, generalized linear, and partially linear models but have not been extended to nonlinear and Cox PH models. In this dissertation, we propose parameter estimation based on linear shrinkage, preliminary test, shrinkage preliminary test, shrinkage, positive-part shrinkage, LASSO, and aLASSO in the context of nonlinear regression models of the Cobb-Douglas, exponential, and monomolecular types and Cox PH regression model (special chapter) in the low-dimensional regime and also extend the work to high-dimensional problems.

CHAPTER 3
RESEARCH METHODOLOGY

Our research includes analysis of model and estimation strategies and asymptotic properties. Monte-Carlo simulations were also employed.

3.1 Nonlinear Regression Model and Least Squares Estimation Strategies

The nonlinear models that we consider are of the Cobb-Douglas, exponential, and monomolecular types. These are as follows:

(a) Cobb-Douglas model: $$y_i = \beta_1 (x_{i1}^{\beta_2})(x_{i2}^{\beta_3}) \cdots (x_{i,p-1}^{\beta_p}) + \varepsilon_i. \tag{3.1}$$

(b) Exponential model: $$y_i = \beta_1 e^{\beta_2 x_{i1} + \beta_3 x_{i2} + \cdots + \beta_p x_{i,p-1}} + \varepsilon_i. \tag{3.2}$$

(c) Monomolecular model: $$y_i = \beta_1 (1 - \beta_2 e^{-\beta_3 x_{i1} - \beta_4 x_{i2} - \cdots - \beta_p x_{i,p-2}}) + \varepsilon_i. \tag{3.3}$$

Suppose that we partition $\boldsymbol{\beta}$ into $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$, where the sub-vectors $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are assumed to have dimensions $p_1 \times 1$ and $p_2 \times 1$ respectively, and $p = p_1 + p_2$. We also let the partitioned matrix be $\boldsymbol{D} = [\boldsymbol{D}_1, \boldsymbol{D}_2]$, which has dimensions $n \times p$. The product matrix is

$$\boldsymbol{D}^\top \boldsymbol{D} = \begin{bmatrix} \boldsymbol{D}_1^\top \boldsymbol{D}_1 & \boldsymbol{D}_1^\top \boldsymbol{D}_2 \\ \boldsymbol{D}_2^\top \boldsymbol{D}_1 & \boldsymbol{D}_2^\top \boldsymbol{D}_2 \end{bmatrix}. \tag{3.4}$$

Therefore, by using the Gauss-Newton method, we can write

$$\boldsymbol{z} = \boldsymbol{D}_1 \boldsymbol{\theta}_1 + \boldsymbol{D}_2 \boldsymbol{\theta}_2 + \boldsymbol{\epsilon}. \tag{3.5}$$

Here, $\boldsymbol{D}_1 = \left[\frac{\partial f(\boldsymbol{x},\boldsymbol{\beta})}{\partial \beta_1}\right]_{\beta=\widehat{\beta}_1}$ and $\boldsymbol{D}_2 = \left[\frac{\partial f(\boldsymbol{x},\boldsymbol{\beta})}{\partial \beta_2}\right]_{\beta=\widehat{\beta}_2}$ are $n \times p_1$ and $n \times p_2$ matrices of derivatives of $f(\boldsymbol{x}, \boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, respectively.

As $\boldsymbol{C} = \boldsymbol{D}^\top \boldsymbol{D}$ is a $p \times p$ matrix, we represent the product matrix decomposition in Equation (3.4) as follows:

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}_1^\top \boldsymbol{D}_1 & \boldsymbol{D}_1^\top \boldsymbol{D}_2 \\ \boldsymbol{D}_2^\top \boldsymbol{D}_1 & \boldsymbol{D}_2^\top \boldsymbol{D}_2 \end{bmatrix} = \boldsymbol{D}^\top \boldsymbol{D}. \tag{3.6}$$

We assume that $\boldsymbol{Q} = \lim\limits_{n\to\infty} \frac{1}{n}\boldsymbol{C}$ as $n \to \infty$, where $\boldsymbol{Q}$ is a $p \times p$ positive definite matrix decomposed as

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{Q}_{11} & \boldsymbol{Q}_{12} \\ \boldsymbol{Q}_{21} & \boldsymbol{Q}_{22} \end{bmatrix}, \tag{3.7}$$

where $\boldsymbol{Q}_{ij} = \lim\limits_{n\to\infty} \frac{1}{n}\boldsymbol{C}_{ij}$ and $i, j = 1, 2$.

3.1.1 Unrestricted Estimator

The unrestricted estimator (UE) or full model estimator of $\boldsymbol{\beta}$, denoted as $\widehat{\boldsymbol{\beta}}^{\mathrm{UE}}$, is the final-iteration nonlinear least squares estimator, which is obtained by solving the Gauss-Newton iterative equation.

Theorem 3.1.1. *Given $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}_n)$ and assuming appropriate regularity conditions of nonlinear least squares. Therefore, for a large n, we have approximately $\widehat{\boldsymbol{\beta}}^{UE} \xrightarrow{D} \mathcal{N}_p\left(\boldsymbol{\beta}, \frac{\sigma^2}{n}\boldsymbol{Q}^{-1}\right)$, where $\boldsymbol{Q} = \lim\limits_{n\to\infty} \frac{1}{n}(\boldsymbol{D}^\top \boldsymbol{D})$ and $\xrightarrow{D}$ indicates convergence in distribution.*

Proof. See Seber and Wild (2003, p. 25) for detailed proof. □

In the final iteration, the UE of $\boldsymbol{\beta}$ can be evaluated as

$$\widehat{\boldsymbol{\beta}}^{\mathrm{UE}} = \left(\boldsymbol{D}^\top \boldsymbol{D}\right)^{-1} \boldsymbol{D}^\top \boldsymbol{F}, \tag{3.8}$$

which can be rewritten as

$$\boldsymbol{D}^\top \boldsymbol{D} \widehat{\boldsymbol{\beta}}^{\mathrm{UE}} = \boldsymbol{D}^\top \boldsymbol{F}, \tag{3.9}$$

where $\boldsymbol{F} = \boldsymbol{D}\widehat{\boldsymbol{\beta}} + \boldsymbol{z}$. Here, $\boldsymbol{D}$ is the matrix of derivatives of $f(\boldsymbol{x}, \boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$, $\widehat{\boldsymbol{\beta}}$ is the estimation of $\boldsymbol{\beta}$, and $\boldsymbol{z} = \boldsymbol{y} - \boldsymbol{f}$. Also, $\widehat{\boldsymbol{\beta}}^{\mathrm{UE}}$ is a consistent estimator of $\boldsymbol{\beta}$ and asymptotically normally distributed.

In later steps, the partition matrix of Equation (3.9) is

$$\begin{bmatrix} \boldsymbol{D}_1^\top \boldsymbol{D}_1 & \boldsymbol{D}_1^\top \boldsymbol{D}_2 \\ \boldsymbol{D}_2^\top \boldsymbol{D}_1 & \boldsymbol{D}_2^\top \boldsymbol{D}_2 \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} \\ \widehat{\boldsymbol{\beta}}_2^{\mathrm{UE}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}_1^\top \boldsymbol{F} \\ \boldsymbol{D}_2^\top \boldsymbol{F} \end{bmatrix}. \tag{3.10}$$

We can write the equations without the surrounding matrix braces as

$$\boldsymbol{D}_1^\top \boldsymbol{D}_1 \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} + \boldsymbol{D}_1^\top \boldsymbol{D}_2 \widehat{\boldsymbol{\beta}}_2^{\mathrm{UE}} = \boldsymbol{D}_1^\top \boldsymbol{F} \tag{3.11}$$

and

$$\boldsymbol{D}_2^\top \boldsymbol{D}_1 \widehat{\boldsymbol{\beta}}_1^{\text{UE}} + \boldsymbol{D}_2^\top \boldsymbol{D}_2 \widehat{\boldsymbol{\beta}}_2^{\text{UE}} = \boldsymbol{D}_2^\top \boldsymbol{F}. \tag{3.12}$$

From Equation (3.11), we get

$$\widehat{\boldsymbol{\beta}}_1^{\text{UE}} = \left(\boldsymbol{D}_1^\top \boldsymbol{D}_1\right)^{-1} \boldsymbol{D}_1^\top \left(\boldsymbol{F} - \boldsymbol{D}_2 \widehat{\boldsymbol{\beta}}_2^{\text{UE}}\right), \tag{3.13}$$

and then substituting for $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ using Equation (3.13) into Equation (3.12). This obtains

$$\begin{aligned} \boldsymbol{D}_2^\top \boldsymbol{D}_1 \left(\boldsymbol{D}_1^\top \boldsymbol{D}_1\right)^{-1} \boldsymbol{D}_1^\top \left(\boldsymbol{F} - \boldsymbol{D}_2 \widehat{\boldsymbol{\beta}}_2^{\text{UE}}\right) + \boldsymbol{D}_2^\top \boldsymbol{D}_2 \widehat{\boldsymbol{\beta}}_2^{\text{UE}} &= \boldsymbol{D}_2^\top \boldsymbol{F} \\ \boldsymbol{D}_2^\top \boldsymbol{D}_1 \left(\boldsymbol{D}_1^\top \boldsymbol{D}_1\right)^{-1} \boldsymbol{D}_1^\top \boldsymbol{F} - \boldsymbol{D}_2^\top \boldsymbol{D}_1 \left(\boldsymbol{D}_1^\top \boldsymbol{D}_1\right)^{-1} \boldsymbol{D}_1^\top \boldsymbol{D}_2 \widehat{\boldsymbol{\beta}}_2^{\text{UE}} + \boldsymbol{D}_2^\top \boldsymbol{D}_2 \widehat{\boldsymbol{\beta}}_2^{\text{UE}} &= \boldsymbol{D}_2^\top \boldsymbol{F}, \end{aligned}$$

which simplifies to

$$\begin{aligned} \boldsymbol{D}_2^\top \left[\boldsymbol{I} - \boldsymbol{D}_1 \left(\boldsymbol{D}_1^\top \boldsymbol{D}_1\right)^{-1} \boldsymbol{D}_1^\top\right] \boldsymbol{D}_2 \widehat{\boldsymbol{\beta}}_2^{\text{UE}} &= \boldsymbol{D}_2^\top \left[\boldsymbol{I} - \boldsymbol{D}_1 \left(\boldsymbol{D}_1^\top \boldsymbol{D}_1\right)^{-1} \boldsymbol{D}_1^\top\right] \boldsymbol{F} \\ \widehat{\boldsymbol{\beta}}_2^{\text{UE}} &= \left(\boldsymbol{D}_2^\top \boldsymbol{M}_1 \boldsymbol{D}_2\right)^{-1} \left(\boldsymbol{D}_2^\top \boldsymbol{M}_1 \boldsymbol{F}\right), \end{aligned} \tag{3.14}$$

where $\boldsymbol{M}_1 = \boldsymbol{I} - \boldsymbol{D}_1 \left(\boldsymbol{D}_1^\top \boldsymbol{D}_1\right)^{-1} \boldsymbol{D}_1^\top$. This is the unrestricted estimator of $\boldsymbol{\beta}_2$. Using the same method, the unrestricted estimator of $\boldsymbol{\beta}_1$ is

$$\widehat{\boldsymbol{\beta}}_1^{\text{UE}} = \left(\boldsymbol{D}_1^\top \boldsymbol{M}_2 \boldsymbol{D}_1\right)^{-1} \left(\boldsymbol{D}_1^\top \boldsymbol{M}_2 \boldsymbol{F}\right), \tag{3.15}$$

where $\boldsymbol{M}_2 = \boldsymbol{I} - \boldsymbol{D}_2 \left(\boldsymbol{D}_2^\top \boldsymbol{D}_2\right)^{-1} \boldsymbol{D}_2^\top$.

Theorem 3.1.2. *If the usual regularity conditions and Theorem 3.1.1 hold, as $n \longrightarrow \infty$, the marginal distribution of $\widehat{\boldsymbol{\beta}}_1^{UE} \xrightarrow{D} \mathcal{N}_{p_1}\left(\boldsymbol{\beta}_1, \frac{\sigma^2}{n}\boldsymbol{Q}_{11.2}^{-1}\right)$ and of $\widehat{\boldsymbol{\beta}}_2^{UE} \xrightarrow{D} \mathcal{N}_{p_2}\left(\boldsymbol{\beta}_2, \frac{\sigma^2}{n}\boldsymbol{Q}_{22.1}^{-1}\right)$. Here, $\boldsymbol{Q} = \lim\limits_{n\to\infty} \frac{1}{n}(\boldsymbol{D}^\top \boldsymbol{D})$, and*

$$\boldsymbol{Q}^{-1} = \begin{bmatrix} \boldsymbol{Q}_{11} & \boldsymbol{Q}_{12} \\ \boldsymbol{Q}_{21} & \boldsymbol{Q}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{Q}_{11.2}^{-1} & -\boldsymbol{Q}_{11.2}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22}^{-1} \\ -\boldsymbol{Q}_{22}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11.2}^{-1} & \boldsymbol{Q}_{22.1}^{-1} \end{bmatrix} \tag{3.16}$$

where $\boldsymbol{Q}_{11.2}^{-1} = \left(\boldsymbol{Q}_{11} - \boldsymbol{Q}_{12}\boldsymbol{Q}_{22}^{-1}\boldsymbol{Q}_{21}\right)^{-1}$, $\boldsymbol{Q}_{22.1}^{-1} = \left(\boldsymbol{Q}_{22} - \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\right)^{-1}$, and $\xrightarrow{D}$ indicates convergence in distribution.

Proof. See Ravishanker and Dey (2001, p. 155) for detailed proof. □

3.1.2 Restricted Estimator

In this study, we first consider the subspace information in the form of a general linear hypothesis, which is

$$\mathrm{H}_0 : \boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{r} \quad \text{versus} \quad \mathrm{H}_1 : \boldsymbol{R}\boldsymbol{\beta} \neq \boldsymbol{r}, \tag{3.17}$$

where $\boldsymbol{R}$ is a known $p_2 \times p$ matrix such that $p_2 \leq p$, and $\boldsymbol{r}$ is $p_2 \times 1$ vector of known constants. Applying the Lagrange multiplier, the restricted (submodel) estimator (RE) of $\boldsymbol{\beta}$ or $\widehat{\boldsymbol{\beta}}^{\mathrm{RE}}$ is derived under the linear restriction $\boldsymbol{R}\boldsymbol{\beta} - \boldsymbol{r} = \boldsymbol{0}$ as follows:

$$\widehat{\boldsymbol{\beta}}^{\mathrm{RE}} = \widehat{\boldsymbol{\beta}}^{\mathrm{UE}} - \boldsymbol{C}^{-1}\boldsymbol{R}^{\top}\left(\boldsymbol{R}\boldsymbol{C}^{-1}\boldsymbol{R}^{\top}\right)^{-1}(\boldsymbol{R}\widehat{\boldsymbol{\beta}}^{\mathrm{UE}} - \boldsymbol{r}), \tag{3.18}$$

which is a linear function of the UE and $\boldsymbol{C} = \boldsymbol{D}^{\top}\boldsymbol{D}$.

Under uncertainty of the prior information, the regression parameter vector $\boldsymbol{\beta}$ can be partitioned as $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^{\top}, \boldsymbol{\beta}_2^{\top})^{\top}$, where $\boldsymbol{\beta}_1$ is a $p_1 \times 1$ subvector of active parameters and $\boldsymbol{\beta}_2$ is a $p_2 \times 1$ subvector of inactive parameters, with $p_1 < p_2$ and $p_1 + p_2 = p$, and there is a possibility that $\boldsymbol{\beta}_2$ is near zero. Thus, we consider the restriction $\boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{0}$ with $\boldsymbol{R} = [\boldsymbol{0}, \boldsymbol{I}]$ and $\boldsymbol{r} = \boldsymbol{0}$, where $\boldsymbol{0}$ is a $p_2 \times p_1$ matrix of zeroes, $\boldsymbol{I}$ is a $p_2 \times p_2$ identity matrix, and $\boldsymbol{r}$ is a $p_2 \times 1$ vector of zero. The relevant hypothesis is

$$\mathrm{H}_0 : \boldsymbol{\beta}_2 = \boldsymbol{0} \quad \text{versus} \quad \mathrm{H}_1 : \boldsymbol{\beta}_2 \neq \boldsymbol{0}. \tag{3.19}$$

Under a submodel in which $\boldsymbol{\beta}_2 = \boldsymbol{0}$ contains only active parameters. The RE of $\boldsymbol{\beta}$ is

$$\begin{aligned} \widehat{\boldsymbol{\beta}}^{\mathrm{RE}} &= (\widehat{\beta}_1^{\mathrm{RE}}, \widehat{\beta}_2^{\mathrm{RE}}, \dots, \widehat{\beta}_{p_1}^{\mathrm{RE}}, \underbrace{0, 0, \dots, 0}_{p_2}) \\ &= \left((\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})^{\top}, \boldsymbol{0}^{\top}\right). \end{aligned} \tag{3.20}$$

The restricted estimator in Equation (3.18) can be rewritten in partitioned form as follows:

$$\begin{bmatrix} \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} \\ \widehat{\boldsymbol{\beta}}_2^{\mathrm{RE}} \end{bmatrix} = \begin{bmatrix} \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} \\ \widehat{\boldsymbol{\beta}}_2^{\mathrm{UE}} \end{bmatrix} - \begin{bmatrix} \boldsymbol{C}_{11}^{-1} + \boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}\boldsymbol{C}_{22.1}^{-1}\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1} & -\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}\boldsymbol{C}_{22.1}^{-1} \\ -\boldsymbol{C}_{22.1}^{-1}\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1} & \boldsymbol{C}_{22.1}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{I} \end{bmatrix}$$
$$\left(\begin{bmatrix} \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{C}_{11}^{-1} + \boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}\boldsymbol{C}_{22.1}^{-1}\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1} & -\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}\boldsymbol{C}_{22.1}^{-1} \\ -\boldsymbol{C}_{22.1}^{-1}\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1} & \boldsymbol{C}_{22.1}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{I} \end{bmatrix}\right)^{-1}$$
$$\left(\begin{bmatrix} \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} \\ \widehat{\boldsymbol{\beta}}_2^{\mathrm{UE}} \end{bmatrix} - \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}\right)$$

$$\begin{bmatrix}\widehat{\boldsymbol{\beta}}_1^{\text{RE}}\\ \widehat{\boldsymbol{\beta}}_2^{\text{RE}}\end{bmatrix} = \begin{bmatrix}\widehat{\boldsymbol{\beta}}_1^{\text{UE}}\\ \widehat{\boldsymbol{\beta}}_2^{\text{UE}}\end{bmatrix} - \begin{bmatrix}-\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}\boldsymbol{C}_{22.1}^{-1}\\ \boldsymbol{C}_{22.1}^{-1}\end{bmatrix}\left(\begin{bmatrix}\boldsymbol{0} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}-\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}\boldsymbol{C}_{22.1}^{-1}\\ \boldsymbol{C}_{22.1}^{-1}\end{bmatrix}\right)^{-1}\widehat{\boldsymbol{\beta}}_2^{\text{UE}}$$
$$= \begin{bmatrix}\widehat{\boldsymbol{\beta}}_1^{\text{UE}}\\ \widehat{\boldsymbol{\beta}}_2^{\text{UE}}\end{bmatrix} - \begin{bmatrix}-\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}\boldsymbol{C}_{22.1}^{-1}\\ \boldsymbol{C}_{22.1}^{-1}\end{bmatrix}(\boldsymbol{C}_{22.1}^{-1})^{-1}\widehat{\boldsymbol{\beta}}_2^{\text{UE}}$$
$$= \begin{bmatrix}\widehat{\boldsymbol{\beta}}_1^{\text{UE}}\\ \widehat{\boldsymbol{\beta}}_2^{\text{UE}}\end{bmatrix} - \begin{bmatrix}-\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}\boldsymbol{C}_{22.1}^{-1}\boldsymbol{C}_{22.1}\widehat{\boldsymbol{\beta}}_2^{\text{UE}}\\ \widehat{\boldsymbol{\beta}}_2^{\text{UE}}\end{bmatrix}$$
$$= \begin{bmatrix}\widehat{\boldsymbol{\beta}}_1^{\text{UE}} + \boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}\boldsymbol{C}_{22.1}^{-1}\boldsymbol{C}_{22.1}\widehat{\boldsymbol{\beta}}_2^{\text{UE}}\\ \boldsymbol{0}\end{bmatrix}.$$

For deriving mathematical results (Lawless & Singhal, 1978), the RE of $\boldsymbol{\beta}_1$, denoted as $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$ can be written as

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_1^{\text{RE}} &= \widehat{\boldsymbol{\beta}}_1^{\text{UE}} - (-\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}\widehat{\boldsymbol{\beta}}_2^{\text{UE}})\\ &= \widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \boldsymbol{\omega}_n\widehat{\boldsymbol{\beta}}_2^{\text{UE}},\end{aligned} \tag{3.21}$$

where $\boldsymbol{\omega}_n = -\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}$. We further suppose that $\boldsymbol{\omega}_n \xrightarrow{P} \boldsymbol{\omega} = -\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}$ as $n \to \infty$, where $\xrightarrow{P}$ indicates convergence in probability. In the simulation, $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$ can also be obtained using the Gauss-Newton iterative method, in which the condition $\boldsymbol{\beta}_2 = \boldsymbol{0}$ is defined as a constraint.

3.2 Suggested Estimation Strategies

As the accuracy of prior information is unknown, then selecting either UE or RE as the estimator of $\boldsymbol{\beta}_1$ may not be a good idea. To avoid this issue, we suggest the linear shrinkage, preliminary test, shrinkage preliminary test, shrinkage, and positive-part shrinkage estimators as in the following subsections.

3.2.1 Linear Shrinkage Estimator

The LS estimator of $\boldsymbol{\beta}_1$, denoted $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, is a linear function of both UE and RE, so that

$$\widehat{\boldsymbol{\beta}}_1^{\text{LS}} = \pi\,\widehat{\boldsymbol{\beta}}_1^{\text{RE}} + (1-\pi)\widehat{\boldsymbol{\beta}}_1^{\text{UE}}, \tag{3.22}$$

or in alternative form

$$\widehat{\boldsymbol{\beta}}_1^{\text{LS}} = \widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \pi(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \widehat{\boldsymbol{\beta}}_1^{\text{RE}}), \tag{3.23}$$

where π is defined as the degree of confidence in the given UPI. The values of π are between 0 and 1, and may be set based on the researcher's belief in the accuracy of the subspace information (Lisawadi et al., 2016). If $\pi = 0$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ and if $\pi = 1$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$.

3.2.2 Preliminary Test Estimator

The PT estimator of $\boldsymbol{\beta}_1$, denoted $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, is a discontinuous function of UE and RE and is defined as follows:

$$\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - (\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) I(\Lambda_n \leq \lambda_\alpha), \tag{3.24}$$

where $I(\cdot)$ is an indicator function, Λ_n is a suitable test statistic for $\mathrm{H}_0 : \boldsymbol{\beta}_2 = 0$, and λ_α is the α-level critical value of the exact distribution of Λ_n. Clearly, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ when H_0 is rejected, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ otherwise.

The PT estimator is limited by the size of α. This limitation is relaxed by defining a shrinkage technique that is a smooth function of PT.

3.2.3 Shrinkage Preliminary Test Estimator

The SP estimator of $\boldsymbol{\beta}_1$, denoted $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, is defined by replacing $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ by $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ in (3.24). This combines the pretest and linear shrinkage strategies in an optimal way. This yields

$$\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - (\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) I(\Lambda_n \leq \lambda_\alpha). \tag{3.25}$$

$\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ is a significant improvement on $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$ in the size of the test and dominates $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ in a large portion of the parameter space. When H_0 is accepted, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ when H_0 is rejected. An alternative form of the SP estimator is given as:

$$\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \pi\,(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) I(\Lambda_n \leq \lambda_\alpha). \tag{3.26}$$

In the case of $\pi = 1$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$.

3.2.4 Shrinkage Estimator

The shrinkage estimator of $\boldsymbol{\beta}_1$, denoted as $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, is a smooth version of the pretest estimator which takes a mixed approach by shrinking the unrestricted model estimator to a plausible alternative estimator or restricted model estimator. This estimator is derived as follows:

$$\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} + \{1 - c\Lambda_n^{-1}\}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}), \tag{3.27}$$

where $c = p_2 - 2$ for the asymptotic case and $p_2 \geq 3$. Here, c is a shrinkage constant chosen in an interval such that $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$ dominates $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ and Λ_n is a test statistic. If $c\Lambda_n^{-1}$ is larger than one, the shrinkage factor $\{1 - c\Lambda_n^{-1}\}$ will be negative, causing the sign of some coefficients to reverse. This is an indication of over-shrinkage. To moderate this effect, the positive-part shrinkage estimator, introduced below, has been suggested.

3.2.5 Positive-Part Shrinkage Estimator

The positive-part shrinkage estimator of $\boldsymbol{\beta}_1$, denoted as $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$, is an improved shrinkage estimator, defined by

$$\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} + \{1 - c\Lambda_n^{-1}\}^+(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}), \tag{3.28}$$

where $c = p_2 - 2$, $p_2 \geq 3$, and $\{1 - c\Lambda_n^{-1}\}^+ = \max(0, 1 - c\Lambda_n^{-1})$. $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$ can be also written in alternative form as

$$\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{S}} - \{1 - c\Lambda_n^{-1}\}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})I(\Lambda_n \leq c), \tag{3.29}$$

where $I(\cdot)$ is an indicator function which is 1 if $\Lambda_n \leq c$, and zero otherwise.

3.3 Penalty Estimation Strategies

The two most widely used penalized estimations were applied to detect significant predictors and reduce the dimensions to a low-dimensional setting.

3.3.1 Least Absolute Shrinkage and Selection Operator Estimator

In statistics, the L_1-norm regularized least squares is known as the least absolute shrinkage and selection operator or LASSO, which is a member of a wide class in the absolute penalty estimation (APE) family. The LASSO estimator of $\boldsymbol{\beta}$ is denoted as $\widehat{\boldsymbol{\beta}}^{\mathrm{LASSO}}$, and is given by

$$\widehat{\boldsymbol{\beta}}^{\mathrm{LASSO}} = \arg\min_{\boldsymbol{\beta}} \left\{ \sum_{i=1}^{n} [y_i - f(\boldsymbol{x}_i, \boldsymbol{\beta})]^2 + \tau \sum_{j=1}^{p} |\beta_j| \right\}, \tag{3.30}$$

where $\tau > 0$ and $\sum_{j=1}^{p} |\beta_j| = \|\boldsymbol{\beta}\|_1$ is the vector L_1-norm. Here, τ is a tuning parameter. If τ is very large, all coefficients are equal to zero. If τ is very small, then the LASSO estimate is equal to the least squares estimate. For nonlinear regression, the algorithm in Yang et al. (2015) is used for finding the LASSO solutions or solving the L_1-norm nonlinear least squares in Equation (3.30).

3.3.2 Adaptive Least Absolute Shrinkage and Selection Operator Estimator

The adaptive least absolute shrinkage and selection operator (adaptive LASSO or aLASSO) is an improvement on LASSO. It provides small weights to active predictors thereby shrinking their associated coefficients a little. On the other hand, it provides large weights to inactive predictors and thus shrinks their related coefficients to exactly zero. The adaptive LASSO estimator of $\boldsymbol{\beta}$ is defined as

$$\widehat{\boldsymbol{\beta}}^{\text{aLASSO}} = \arg\min_{\boldsymbol{\beta}} \left\{ \sum_{i=1}^{n} [y_i - f(\boldsymbol{x}_i, \boldsymbol{\beta})]^2 + \tau \sum_{j=1}^{p} \widehat{w}_j |\beta_j| \right\}, \tag{3.31}$$

where $\tau > 0$ is a tuning parameter, $\widehat{w}_j = |\widehat{\beta}_j|^{-\gamma}$ is is the adaptive weight with $\gamma > 0$, and $\widehat{\beta}_j$ is an initial estimator of β_j for $j = 1, 2, ..., p$. The adaptive LASSO was introduced by Zou (2006) for linear regression. For our study, we applied the algorithm in Yang et al. (2015) by using adaptive weights for the L_1 penalty on the regression coefficients for solving the aLASSO solution in a nonlinear regression.

3.4 Large Sample Test

In this section, an appropriate large sample test statistic is desired for testing the hypothesis $\text{H}_0 : \boldsymbol{\beta}_2 = \mathbf{0}$ against $\text{H}_1 : \boldsymbol{\beta}_2 \neq \mathbf{0}$. The Wald statistic (Ahmed and Nicol, 2012; Harrell, 2015, p. 189) for the null hypothesis is asymptotically valid under non-normality, which is

$$\Lambda_n = \left(\widehat{\boldsymbol{\beta}}_2^{\text{UE}}\right)^\top \left[\mathbb{V}\left(\widehat{\boldsymbol{\beta}}_2\right)\right]^{-1} \widehat{\boldsymbol{\beta}}_2^{\text{UE}}, \tag{3.32}$$

where $\mathbb{V}(\widehat{\boldsymbol{\beta}}_2)$ is the variance-covariance matrix of $\widehat{\boldsymbol{\beta}}_2$. Here, $\mathbb{V}(\widehat{\boldsymbol{\beta}}_2) = \widehat{\sigma}^2 \boldsymbol{C}_{22.1}^{-1}$, $\widehat{\sigma}^2 = [\boldsymbol{y} - f(\boldsymbol{x}, \widehat{\boldsymbol{\beta}})]^\top [\boldsymbol{y} - f(\boldsymbol{x}, \widehat{\boldsymbol{\beta}})]/(n-p)$, and $\boldsymbol{C}_{22.1} = \boldsymbol{C}_{22} - \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}$. Under the null hypothesis, as $n \to \infty$, the test statistic is asymptotically distributed as chi-square with p_2 degrees of freedom, where p_2 is the number of restrictions on $\widehat{\boldsymbol{\beta}}^{\text{RE}}$.

3.5 Asymptotic Properties

The goal here is to derive the asymptotic distributional bias (ADB), asymptotic distributional quadratic bias (ADQB), and asymptotic distributional quadratic risk (ADQR) of the estimators. To achieve this, we define a sequence of local alternatives $\{K_n\}$ as follows:

$$\{K_n\} : \boldsymbol{\beta}_2 = \frac{\boldsymbol{\delta}}{\sqrt{n}}, \tag{3.33}$$

where $\boldsymbol{\delta} = (\delta_1, \delta_2, ..., \delta_{p_2})^\top \in \mathbb{R}^{p_2}$ is a $p_2 \times 1$ fixed vector. Note that $\boldsymbol{\delta} = \mathbf{0}$ implies that the null hypothesis in (3.19) is a special case of $\{K_n\}$.

We measure the performance of each estimator under local alternatives. We are mainly interested in estimating the unknown parameter vector $\boldsymbol{\beta}_1$ by means of an estimator $\widehat{\boldsymbol{\beta}}_1^*$. To begin the process of computing ADB, ADQB, and ADQR, we suppose that the asymptotic cumulative distribution function (CDF) of $\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^* - \boldsymbol{\beta}_1)$ is a limiting distribution under $\{K_n\}$, and is defined as $F(\boldsymbol{y}) = \lim_{n\to\infty} \mathbb{P}\left[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^* - \boldsymbol{\beta}_1) \leq \boldsymbol{y}\right]$. This is known as the asymptotic distribution function (ADF) of $\widehat{\boldsymbol{\beta}}_1^*$.

3.5.1 Asymptotic Distributional Quadratic Bias

First, we present the expressions for the asymptotic distributional bias (ADB) of the proposed estimator. The ADB of an estimator $\widehat{\boldsymbol{\beta}}_1^*$ is defined as

$$\text{ADB}(\widehat{\boldsymbol{\beta}}_1^*) = \lim_{n\to\infty} \mathbb{E}\left[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^* - \boldsymbol{\beta}_1)\right], \tag{3.34}$$

where the result is in vector form. Then, we convert these functions into the scalar (quadratic) form, using the asymptotic distributional quadratic bias (ADQB). The ADQB of estimator $\widehat{\boldsymbol{\beta}}_1^*$ of parameter vector $\boldsymbol{\beta}_1$ is defined as

$$\text{ADQB}(\widehat{\boldsymbol{\beta}}_1^*) = \left[\text{ADB}(\widehat{\boldsymbol{\beta}}_1^*)\right]^\top \psi_{11.2}^{-1} \left[\text{ADB}(\widehat{\boldsymbol{\beta}}_1^*)\right], \tag{3.35}$$

where $\psi_{11.2} = \sigma^2 \boldsymbol{Q}_{11.2}^{-1} = \sigma^2(\boldsymbol{Q}_{11} - \boldsymbol{Q}_{12}\boldsymbol{Q}_{22}^{-1}\boldsymbol{Q}_{21})^{-1}$ is the asymptotic variance-covariance matrix of $\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \boldsymbol{\beta}_1)$.

3.5.2 Asymptotic Distributional Quadratic Risk

A loss function $\mathcal{L}(\widehat{\boldsymbol{\beta}}_1^*, \boldsymbol{\beta}_1)$ represents the loss incurred if a wrong decision is made about $\boldsymbol{\beta}_1$ using the estimator $\widehat{\boldsymbol{\beta}}_1^*$. We define the weighted quadratic loss function in the form

$$\mathcal{L}(\widehat{\boldsymbol{\beta}}_1^*, \boldsymbol{\beta}_1; \boldsymbol{W}) = \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^* - \boldsymbol{\beta}_1)^\top \boldsymbol{W} \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^* - \boldsymbol{\beta}_1), \tag{3.36}$$

where $\boldsymbol{W}$ is a positive semi-definite weighting matrix and $\widehat{\boldsymbol{\beta}}_1^*$ can be any of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$. Then, the asymptotic distributional quadratic risk (ADQR) of $\widehat{\boldsymbol{\beta}}_1^*$ is given by

$$\begin{aligned}
\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^*) &= \lim_{n\to\infty} \mathbb{E}[\mathcal{L}(\widehat{\boldsymbol{\beta}}_1^*, \boldsymbol{\beta}_1; \boldsymbol{W})] \\
&= \lim_{n\to\infty} \mathbb{E}\left[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^* - \boldsymbol{\beta}_1)\boldsymbol{W}\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^* - \boldsymbol{\beta}_1)^\top\right] \\
&= \mathrm{tr}[\boldsymbol{W}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^*)],
\end{aligned} \tag{3.37}$$

where $\mathrm{tr}(\cdot)$ is the trace of the matrix. Here, we let

$$\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^*) = \lim_{n\to\infty} \mathbb{E}\left[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^* - \boldsymbol{\beta}_1)\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^* - \boldsymbol{\beta}_1)^\top\right], \tag{3.38}$$

be the asymptotic mean squared error matrix (AMSEM) of an estimator $\widehat{\boldsymbol{\beta}}_1^*$.

The way to evaluate performance is to compare the ADQR of two estimators with a suitable matrix $\boldsymbol{W}$. Normally, a preferable estimator will have a smaller ADQR. The estimator $\widehat{\boldsymbol{\beta}}_1^*$ will be termed inadmissible if there exists another estimator $\widehat{\boldsymbol{\beta}}_1^0$, such that $\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^0) \leq \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^*)$. In these cases, we can say that the estimator $\widehat{\boldsymbol{\beta}}_1^0$ dominates $\widehat{\boldsymbol{\beta}}_1^*$.

3.6 Related Theorems

Theorem 3.6.1. *Under $\{K_n\}$ and the usual regularity conditions, as n increases, the test statistic Λ_n converges to a non-central chi-squared distribution with p_2 degrees of freedom and the non-centrality parameter $\Delta = \boldsymbol{\delta}^\top \psi_{22.1}^{-1}\boldsymbol{\delta}$. Here, $\psi_{22.1} = \sigma^2\boldsymbol{Q}_{22.1}^{-1}$ is the asymptotic variance-covariance matrix of $\sqrt{n}(\widehat{\boldsymbol{\beta}}_2^{UE} - \boldsymbol{\beta}_2)$.*

Proof. See Davidson and Lever (1970) for detailed proof. □

Theorem 3.6.2. *Let $\boldsymbol{y}$ be a p_2-dimensional random vector that follows a multivariate normal distribution with mean $\boldsymbol{\mu_y}$ and variance $\boldsymbol{I_y}$. Then, for any measurable function φ, we have*

$$\mathbb{E}\left[\boldsymbol{y}\varphi\left(\boldsymbol{y}^\top\boldsymbol{y}\right)\right] = \boldsymbol{\mu_y}\mathbb{E}\left[\varphi\left(\chi^2_{p_2+2}(\Delta)\right)\right], \tag{3.39}$$

$$\mathbb{E}\left[\boldsymbol{y}\boldsymbol{y}^\top\varphi\left(\boldsymbol{y}^\top\boldsymbol{y}\right)\right] = \boldsymbol{I_y}\mathbb{E}\left[\varphi\left(\chi^2_{p_2+2}(\Delta)\right)\right] + \boldsymbol{\mu_y}\boldsymbol{\mu_y}^\top\mathbb{E}\left[\varphi\left(\chi^2_{p_2+4}(\Delta)\right)\right]. \tag{3.40}$$

Proof. See Judge and Bock (1978, p. 322) for detailed proof. □

Theorem 3.6.3. *Let $\boldsymbol{y} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Suppose we partition $\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix}$, where $\boldsymbol{y}_1$ is a $p_1 \times 1$ vector, $\boldsymbol{y}_2$ is a $p_2 \times 1$ vector, and assume that $\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}$ and $\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$. The conditional distribution of $\boldsymbol{y}_1$ given that $\boldsymbol{y}_2 = \boldsymbol{c}_2$ is a multivariate normal with mean vector*

$$\mathbb{E}[\boldsymbol{y}_1|\boldsymbol{y}_2 = \boldsymbol{c}_2] = \boldsymbol{\mu}_{\boldsymbol{y}_1|\boldsymbol{y}_2=\boldsymbol{c}_2} = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{c}_2 - \boldsymbol{\mu}_2) \tag{3.41}$$

and variance-covariance matrix

$$\mathbb{V}[\boldsymbol{y}_1|\boldsymbol{y}_2 = \boldsymbol{c}_2] = \boldsymbol{\Sigma}_{\boldsymbol{y}_1|\boldsymbol{y}_2=\boldsymbol{c}_2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}. \tag{3.42}$$

Proof. See Ravishanker and Dey (2001, pp. 156–157) for detailed proof. □

3.7 Dimensionality of Data

Dimensionality in statistics refers to how many attributes a dataset has. This dissertation classified two types of data: low-dimensional and high-dimensional data.

3.7.1 Low-Dimensional Data Setting

In ordinary or low-dimensional (LD) data, the sample size (n) exceeds the number of predictors (k), so $k < n$. Under uncertainty of the prior information, the predictors can be categorized into two groups:

Group 1: Active predictors, which we believed to be significantly associated with the response variable.

Group 2: Inactive predictors, which are gently or not significant for predictive modelling. They should be removed from the model.

For this reason, the unrestricted model contains all k predictors, both k_1 active and k_2 inactive, such that $k = k_1 + k_2$, while the restricted model contains only k_1 active predictors. Therefore, we partitioned the regression coefficients as

$$\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top = (\beta_1, \beta_2, ..., \beta_{p_1}, \beta_{p_1+1}, \beta_{p_1+2}, ..., \beta_{p_1+p_2})^\top,$$

where $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are active and inactive parameters with dimensions $p_1 \times 1$ and $p_2 \times 1$, respectively. As some parameters may be eliminated from the model, it is plausible that $\boldsymbol{\beta}_2$ can be set to a zero vector as follows:

$$\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \mathbf{0}_{p_2}^\top)^\top = (\underbrace{\beta_1, \beta_2, ..., \beta_{p_1}}_{p_1}, \underbrace{0, 0, ..., 0}_{p_2})^\top, \tag{3.43}$$

which is the coefficient vector under subspace information.

3.7.2 High-Dimensional Data Setting

Nowadays, many researchers are focused on the analysis and modeling of high-dimensional (HD) data analysis, where the sample size (n) is smaller than the number of predictors (k). To obtain a meaningful statistical output, the classical parameter estimation methods, i.e. least squares and maximum likelihood, are based on the assumption that $k < n$, and become unreliable or impossible to directly compute when $k \geq n$. In this context, we construct possibly sparse models in high-dimensional settings when k is not fixed (often written as $k \geq n$). This sparsity means that some explanatory variables do not affect the response variable, or else some regression coefficients in the model are exactly zero.

Following Ahmed and Yüzbaşı (2016), Gao et al. (2017), and Yüzbaşı et al. (2017), the estimation problem of regression parameters is seen when there are many predictors in the model. The predictors can be characterized into the following three groups:

Group 1: Predictors with strong influence (strong signals) on the response variable and $|\beta_j| > c\sqrt{\log(p)/n}$ for some $c > 0$ and $1 \leq j \leq k$.

Group 2: Predictors with weak-to-moderate influence (weak-to-moderate signals) which may or may not contribute to explaining the response variable and $0 < |\beta_j| < c\sqrt{\log(p)/n}$ for some $c > 0$ and $1 \leq j \leq k$.

Group 3: Predictors with no influence (sparse or no signals) on the response variable in which their related regression coefficients are exactly zero.

In this work, we consider the sparse regression models when there are many predictors that have a weak influence on the response variable. We are still interested in cases where predictors with strong signals are stored in the model, and some or all predictors with weak-to-moderate signals are also included in the model, even though they may not be useful for prediction purposes. This leads to the consideration of two models. One is an overfitted (OF) model that includes predictors with strong signals and possibly some predictors with weak-to-moderate signals which may be produced by using a variable selection strategy. Conversely, other methods may establish an underfitted (UF) model that possibly includes predictors with strong signals while leaving out predictors with weak-to-moderate signals.

Generally, the k predictors can be featured as k_s strong, k_w weak-to-moderate, and k_n no signals, such that $k_s + k_w + k_n = k$. We partition $\boldsymbol{x} = (\boldsymbol{x}_s, \boldsymbol{x}_w, \boldsymbol{x}_n)$ and $\boldsymbol{\beta} = (\boldsymbol{\beta}_s^\top, \boldsymbol{\beta}_w^\top, \boldsymbol{\beta}_n^\top)^\top$. Further, $\boldsymbol{x}_s$ is the $n \times k_s$, $\boldsymbol{x}_w$ is the $n \times k_w$, and $\boldsymbol{x}_n$ is the $n \times k_n$ submatrix of the predictors, respectively. Here, $\boldsymbol{\beta}_s$ is the $p_s \times 1$, $\boldsymbol{\beta}_w$ is the $p_w \times 1$, and $\boldsymbol{\beta}_n$ is the $p_n \times 1$ subvector of the parameters, respectively. The resulting models produced by using LASSO and aLASSO strategies are rewritten as

$$\boldsymbol{y} = f(\boldsymbol{x}_s, \boldsymbol{x}_w, \boldsymbol{\beta}_s, \boldsymbol{\beta}_w) + \varepsilon,$$

where $k_s \leq k_w < n$, and

$$\boldsymbol{y} = f(\boldsymbol{x}_s, \boldsymbol{\beta}_s,) + \varepsilon,$$

where $k_s < n$, respectively. The sparsity structure in the high dimensional setting was assumed to be known and reordered the coefficients of $\boldsymbol{\beta}$ according to the signal strength as follows:

$$\boldsymbol{\beta} = (\underbrace{\beta_1, \beta_2, ..., \beta_{p_s}}_{p_s}, \underbrace{\kappa, \kappa, ..., \kappa}_{p_w}, \underbrace{0, 0, ..., 0}_{p_n})^\top, \tag{3.44}$$

having strong, weak-to-moderate, and no signals, respectively. For this study, we separated the procedure under a high-dimensional setting into two steps, namely:

1. A variable or dimensional reduction step to detect significant predictors and to reduce the dimensions to a low-dimensional model.

2. A post-selection parameter estimation step, using the resulting model attained from step 1 above.

In the dimensional reduction step, we applied the LASSO method to eliminate predictors with no signals and to keep predictors with both strong and weak-to-moderate signals. Therefore, this may be considered to be an overfitted (OF) model. The other model is from the aLASSO method, which retains the predictors with strong signals and produces a lower-dimensional model as compared to the LASSO method. Therefore, this model may be termed an underfitted (UF) model. In this step, we assumed that the subset of predictors using the LASSO strategy contained $k_1 + k_2$ selected variables, containing $x_{i1}, x_{i2}, ..., x_{i(k_1+k_2)}$. While the aLASSO strategy chose only k_1 relevant variables, which was $x_{i1}, x_{i2}, ..., x_{ik_1}$, where $k_1 < k_1 + k_2 < k$.

For the post-selection parameter estimation step, the estimation strategies that we suggested in Section 3.2 were applied after performing variable selection for estimating $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$ when $\boldsymbol{\beta}_2$ may be a zero vector. From the partition of regression coefficient vector, $\boldsymbol{\beta}_1 = (\beta_1, \beta_2, ..., \beta_{p_1})^\top$ is a $p_1 \times 1$ vector of parameters associated with k_1 predictors selected by LASSO and aLASSO methods, and $\boldsymbol{\beta}_2 = (\beta_{p_1+1}, \beta_{p_1+2}, ..., \beta_{p_1+p_2})^\top$ is $p_2 \times 1$ vector of parameters related with k_2 predictors chosen by only LASSO. To allow post-selection parameter estimation, the UE and the RE were the nonlinear least square estimators based on LASSO (OF) and aLASSO (UF) models. For testing $\mathrm{H}_0 : \boldsymbol{\beta}_2 = \mathbf{0}_{p_2}$, the distribution of Λ_n in Equation (3.32) converges to a chi-square distribution with p_2 degrees of freedom as $n \to \infty$.

3.8 Measures of Estimator Performance

The relative mean squares error evaluated the performance of the proposed estimators in the Monte Carlo simulation. In contrast, the relative mean squares prediction error assessed the performance of the suggested estimators in the real data example.

3.8.1 Monte Carlo Simulations

Monte Carlo simulations are used to assess the performance of proposed estimators under different situations. The performance of the estimator $\widehat{\boldsymbol{\beta}}_1^*$ will be measured by its mean square error (MSE). The formula for computing the simulated

MSE of an estimator is defined as

$$\begin{aligned}\text{MSE}(\widehat{\boldsymbol{\beta}}_1^*) &= \frac{1}{n}(\boldsymbol{\beta}_1 - \widehat{\boldsymbol{\beta}}_1^*)^\top(\boldsymbol{\beta}_1 - \widehat{\boldsymbol{\beta}}_1^*) \\ &= \frac{1}{n}\sum_{i=1}^{p}(\boldsymbol{\beta}_1 - \widehat{\boldsymbol{\beta}}_1^*)^2\end{aligned}$$

where $\widehat{\boldsymbol{\beta}}_1^*$ is one of $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}, \widehat{\boldsymbol{\beta}}_1^{\text{RE}}, \widehat{\boldsymbol{\beta}}_1^{\text{LS}}, \widehat{\boldsymbol{\beta}}_1^{\text{PT}}, \widehat{\boldsymbol{\beta}}_1^{\text{SP}}, \widehat{\boldsymbol{\beta}}_1^{\text{S}}, \widehat{\boldsymbol{\beta}}_1^{\text{S+}}$, or $\widehat{\boldsymbol{\beta}}_1^{\text{LASSO}}$. The simulated relative mean square error (RMSE) was used to compare the performance of the estimator $\widehat{\boldsymbol{\beta}}_1^*$ with benchmark $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$. The RMSE is the ratio of simulated MSE of two estimators and is given as

$$\text{RMSE}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}, \widehat{\boldsymbol{\beta}}_1^*) = \frac{\text{MSE}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}})}{\text{MSE}(\widehat{\boldsymbol{\beta}}_1^*)}.$$

A value of RMSE larger than one indicates $\widehat{\boldsymbol{\beta}}_1^*$ superiority over $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$.

3.8.2 Real Data Example

We compared the performance of the proposed estimators with respect to the benchmark estimator $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ using the simulated relative mean squares prediction error (RMSPE), defined as

$$\text{RMSPE}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}, \widehat{\boldsymbol{\beta}}_1^*) = \frac{\text{MSPE}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}})}{\text{MSPE}(\widehat{\boldsymbol{\beta}}_1^*)}. \tag{3.45}$$

Note that an RMSPE larger than one means that $\widehat{\boldsymbol{\beta}}_1^*$ outperforms $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$. The mean squared prediction error (MSPE) was obtained by taking the squared deviation of the observed and predicted values, divided by the number of bootstrap samples (m), as follows:

$$\text{MSPE}(\widehat{\boldsymbol{\beta}}_1^*) = \frac{1}{m}(\boldsymbol{y} - f(\boldsymbol{x}, \widehat{\boldsymbol{\beta}}_1^*))^\top(\boldsymbol{y} - f(\boldsymbol{x}, \widehat{\boldsymbol{\beta}}_1^*)),$$

where $\widehat{\boldsymbol{\beta}}_1^*$ is one of $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}, \widehat{\boldsymbol{\beta}}_1^{\text{RE}}, \widehat{\boldsymbol{\beta}}_1^{\text{LS}}, \widehat{\boldsymbol{\beta}}_1^{\text{PT}}, \widehat{\boldsymbol{\beta}}_1^{\text{SP}}, \widehat{\boldsymbol{\beta}}_1^{\text{S}}, \widehat{\boldsymbol{\beta}}_1^{\text{S+}}, \widehat{\boldsymbol{\beta}}_1^{\text{LASSO}}$, or $\widehat{\boldsymbol{\beta}}_1^{\text{aLASSO}}$.

We used bootstrap methodology to assess the performance of the proposed estimators. For every bootstrapped sample, we computed RMSPE for every estimator by employing the formula (3.45) relative to the UE.

CHAPTER 4
RESEARCH RESULTS

This chapter presents the theoretical and numerical analysis and describes the results for each nonlinear regression model (i.e., Cobb-Douglas, exponential, and monomolecular). The results of the data analysis are divided into three parts, including asymptotic results, simulation results, and application to real data.

4.1 Asymptotic Results

We first display the asymptotic distribution of UE and RE and their joint distributions. Next, the asymptotic properties of the suggested estimators are derived in the context of ADQB and ADQR. We then summarize the asymptotic results for the estimators to compare their performance. In this study, we do not examine the behavior of the two penalty strategies since these do not make use of subspace information, and their performance depends on the tuning parameter.

To achieve the asymptotic properties of the proposed estimators and of the test statistic Λ_n, the following theorem facilitates computation of asymptotic properties under local alternatives $\{K_n\}$.

Theorem 4.1.1. *Under local alternatives* $\{K_n\}$ *and the usual regularity conditions, as* $n \to \infty$ *we have the following:*

$$\boldsymbol{T}_n = \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{UE} - \boldsymbol{\beta}_1) \xrightarrow{D} \boldsymbol{T} \sim \mathcal{N}_{p_1}(\boldsymbol{0}, \sigma^2 \boldsymbol{Q}_{11.2}^{-1}),$$
$$\boldsymbol{U}_n = \sqrt{n}(\widehat{\boldsymbol{\beta}}_2^{UE} - \boldsymbol{\beta}_2) \xrightarrow{D} \boldsymbol{U} \sim \mathcal{N}_{p_2}(\boldsymbol{0}, \sigma^2 \boldsymbol{Q}_{22.1}^{-1}),$$
$$\boldsymbol{W}_n = \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{RE} - \boldsymbol{\beta}_1) \xrightarrow{D} \boldsymbol{W} \sim \mathcal{N}_{p_1}(-\boldsymbol{\omega}\boldsymbol{\delta}, \sigma^2 \boldsymbol{Q}_{11}^{-1}),$$
$$\boldsymbol{Z}_n = \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{UE} - \widehat{\boldsymbol{\beta}}_1^{RE}) \xrightarrow{D} \boldsymbol{Z} \sim \mathcal{N}_{p_1}(\boldsymbol{\omega}\boldsymbol{\delta}, \sigma^2 \boldsymbol{\Omega}),$$
$$\begin{bmatrix} \boldsymbol{T}_n \\ \boldsymbol{Z}_n \end{bmatrix} \xrightarrow{D} \begin{bmatrix} \boldsymbol{T} \\ \boldsymbol{Z} \end{bmatrix} \sim \mathcal{N}_{2p_1}\left(\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{\omega}\boldsymbol{\delta} \end{bmatrix}, \sigma^2 \begin{bmatrix} \boldsymbol{Q}_{11.2}^{-1} & \boldsymbol{\Omega} \\ \boldsymbol{\Omega} & \boldsymbol{\Omega} \end{bmatrix} \right),$$
$$\begin{bmatrix} \boldsymbol{W}_n \\ \boldsymbol{Z}_n \end{bmatrix} \xrightarrow{D} \begin{bmatrix} \boldsymbol{W} \\ \boldsymbol{Z} \end{bmatrix} \sim \mathcal{N}_{2p_1}\left(\begin{bmatrix} -\boldsymbol{\omega}\boldsymbol{\delta} \\ \boldsymbol{\omega}\boldsymbol{\delta} \end{bmatrix}, \sigma^2 \begin{bmatrix} \boldsymbol{Q}_{11}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Omega} \end{bmatrix} \right),$$

where $\boldsymbol{\omega} = -\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}$*,* $\boldsymbol{\Omega} = \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}$.

Proof. Under the sequence of local alternatives, $\{K_n\} : \boldsymbol{\beta}_2 = \boldsymbol{\beta}_2^0 + \frac{\boldsymbol{\delta}}{\sqrt{n}}$, we obtained $\boldsymbol{\delta}_{p_2\times 1} = \sqrt{n}(\boldsymbol{\beta}_2 - \mathbf{0})$, where $\boldsymbol{\beta}_2^0 = \mathbf{0}$.

(1) The asymptotic distribution of $\boldsymbol{T}_n$ and $\boldsymbol{U}_n$ can be directly attained from Theorem 3.1.2 in Chapter 3.

(2) Let $\boldsymbol{W}_n = \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}} - \boldsymbol{\beta}_1)$ be a $p_1 \times 1$ matrix and we obtain

$$\begin{aligned}
\boldsymbol{W}_n &= \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}} - \boldsymbol{\beta}_1) \\
&= \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \boldsymbol{\omega}_n\widehat{\boldsymbol{\beta}}_2^{\text{UE}} - \boldsymbol{\beta}_1) \\
&= \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \boldsymbol{\beta}_1) - \boldsymbol{\omega}_n\sqrt{n}\widehat{\boldsymbol{\beta}}_2^{\text{UE}} \\
&= \boldsymbol{T}_n - \boldsymbol{\omega}_n\sqrt{n}\widehat{\boldsymbol{\beta}}_2^{\text{UE}},
\end{aligned}$$

which is a linear function of $\boldsymbol{T}_n$. By Slutsky's theorem and $n \to \infty$, then $\boldsymbol{W}_n \xrightarrow{d} \boldsymbol{W} \sim \mathcal{N}_{p_1}(\mu_{\boldsymbol{W}}, \Sigma_{\boldsymbol{W}})$, where

$$\begin{aligned}
\mu_{\boldsymbol{W}} &= \mathbb{E}(\boldsymbol{W}) = \mathbb{E}\left[\boldsymbol{T} - \boldsymbol{\omega}\sqrt{n}\widehat{\boldsymbol{\beta}}_2^{\text{UE}}\right] \\
&= \mathbb{E}\left[\boldsymbol{T} - \boldsymbol{\omega}\sqrt{n}\widehat{\boldsymbol{\beta}}_2^{\text{UE}} - \boldsymbol{\omega}\sqrt{n}\boldsymbol{\beta}_2 + \boldsymbol{\omega}\sqrt{n}\boldsymbol{\beta}_2\right] \\
&= \mathbb{E}\left[\boldsymbol{T} - \boldsymbol{\omega}\sqrt{n}(\widehat{\boldsymbol{\beta}}_2^{\text{UE}} - \boldsymbol{\beta}_2) - \boldsymbol{\omega}\sqrt{n}\frac{\boldsymbol{\delta}}{\sqrt{n}}\right] \\
&= \mathbb{E}\left[\boldsymbol{T} - \boldsymbol{\omega}\boldsymbol{U} - \boldsymbol{\omega}\boldsymbol{\delta}\right] \\
&= \mathbb{E}(\boldsymbol{T}) - \boldsymbol{\omega}\mathbb{E}(\boldsymbol{U}) - \boldsymbol{\omega}\boldsymbol{\delta} \\
&= -\boldsymbol{\omega}\boldsymbol{\delta}
\end{aligned}$$

and

$$\begin{aligned}
\Sigma_{\boldsymbol{W}} &= \mathbb{V}(\boldsymbol{W}) = \mathbb{V}[\boldsymbol{T} - \boldsymbol{\omega}\boldsymbol{U} - \boldsymbol{\omega}\boldsymbol{\delta}] \\
&= \mathbb{V}(\boldsymbol{T}) - \boldsymbol{\omega}\mathbb{V}(\boldsymbol{U})\boldsymbol{\omega}^\top \\
&= \sigma^2\boldsymbol{Q}_{11.2}^{-1} - \boldsymbol{\omega}\sigma^2\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{\omega}^\top \\
&= \sigma^2(\boldsymbol{Q}_{11}^{-1} + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}) - \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\sigma^2\boldsymbol{Q}_{22.1}^{-1}(\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12})^\top \\
&= \sigma^2\left[\boldsymbol{Q}_{11}^{-1} + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} - \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}(\boldsymbol{Q}_{11}^{-1})^\top\right] \\
&= \sigma^2[\boldsymbol{Q}_{11}^{-1} + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} - \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}] \\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}.
\end{aligned}$$

(3) Let $\boldsymbol{Z}_n = \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \widehat{\boldsymbol{\beta}}_1^{\text{RE}})$ be a $p_1 \times 1$ matrix and we get

$$\begin{aligned}
\boldsymbol{Z}_n &= \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \widehat{\boldsymbol{\beta}}_1^{\text{RE}}) \\
&= \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \widehat{\boldsymbol{\beta}}_1^{\text{RE}} - \boldsymbol{\beta}_1 + \boldsymbol{\beta}_1) \\
&= \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \boldsymbol{\beta}_1) - \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}} - \boldsymbol{\beta}_1) \\
&= \boldsymbol{T}_n - \boldsymbol{W}_n \\
&= \boldsymbol{T}_n - (\boldsymbol{T}_n - \boldsymbol{\omega}_n \sqrt{n}\widehat{\boldsymbol{\beta}}_2^{\text{UE}})
\end{aligned}$$

which is also a linear function of $\boldsymbol{T}_n$. Hence, by Slutsky's theorem and $\boldsymbol{\omega}_n \xrightarrow{P} \boldsymbol{\omega}$, we obtain $\boldsymbol{Z}_n \xrightarrow{D} \boldsymbol{Z} \sim \mathcal{N}_{p_1}(\mu_{\boldsymbol{Z}}, \Sigma_{\boldsymbol{Z}})$ as $n \to \infty$, where

$$\begin{aligned}
\mu_{\boldsymbol{Z}} = \mathbb{E}(\boldsymbol{Z}) &= \mathbb{E}\left[\boldsymbol{T} - (\boldsymbol{T} - \boldsymbol{\omega}\sqrt{n}\widehat{\boldsymbol{\beta}}_2^{\text{UE}})\right] \\
&= \mathbb{E}\left[\boldsymbol{\omega}\sqrt{n}\widehat{\boldsymbol{\beta}}_2^{\text{UE}} - \boldsymbol{\omega}\sqrt{n}\boldsymbol{\beta}_2 + \boldsymbol{\omega}\sqrt{n}\boldsymbol{\beta}_2\right] \\
&= \mathbb{E}\left[\boldsymbol{\omega}\sqrt{n}(\widehat{\boldsymbol{\beta}}_2^{\text{UE}} - \boldsymbol{\beta}_2) + \boldsymbol{\omega}\sqrt{n}\frac{\boldsymbol{\delta}}{\sqrt{n}}\right] \\
&= \mathbb{E}[\boldsymbol{\omega}\boldsymbol{U} + \boldsymbol{\omega}\boldsymbol{\delta}] \\
&= \boldsymbol{\omega}\mathbb{E}(\boldsymbol{U}) + \boldsymbol{\omega}\boldsymbol{\delta} \\
&= \boldsymbol{\omega}\boldsymbol{\delta},
\end{aligned}$$

and

$$\begin{aligned}
\Sigma_{\boldsymbol{Z}} = \mathbb{V}(\boldsymbol{Z}) &= \mathbb{V}[\boldsymbol{\omega}\boldsymbol{U} + \boldsymbol{\omega}\boldsymbol{\delta}] \\
&= \boldsymbol{\omega}\mathbb{V}(\boldsymbol{U})\boldsymbol{\omega}^\top \\
&= (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12})\sigma^2\boldsymbol{Q}_{22.1}^{-1}(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12})^\top \\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} \\
&= \sigma^2\boldsymbol{\Omega},
\end{aligned}$$

where $\boldsymbol{\Omega} = \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}$.

(4) Now, we consider

$$\begin{bmatrix}\boldsymbol{T}_n \\ \boldsymbol{Z}_n\end{bmatrix} = \begin{bmatrix}\boldsymbol{T}_n \\ \boldsymbol{T}_n - (\boldsymbol{T}_n - \boldsymbol{\omega}_n\sqrt{n}\widehat{\boldsymbol{\beta}}_2^{\text{UE}})\end{bmatrix} = \begin{bmatrix}\boldsymbol{T}_n \\ \boldsymbol{\omega}_n\sqrt{n}\widehat{\boldsymbol{\beta}}_2^{\text{UE}}\end{bmatrix} = \boldsymbol{T}_n\begin{bmatrix}\boldsymbol{I}_{p_1} \\ \boldsymbol{0}_{p_1}\end{bmatrix} + \begin{bmatrix}\boldsymbol{0}_{p_1} \\ \boldsymbol{\omega}_n\sqrt{n}\widehat{\boldsymbol{\beta}}_2^{\text{UE}}\end{bmatrix},$$

which is a linear function of $\boldsymbol{T}_n$. Also, by Slutsky's theorem, the joint distribution of $\boldsymbol{T}_n$ and $\boldsymbol{Z}_n$, as $n \to \infty$ is $\left[\begin{smallmatrix}\boldsymbol{T}_n \\ \boldsymbol{Z}_n\end{smallmatrix}\right] \xrightarrow{d} \left[\begin{smallmatrix}\boldsymbol{T} \\ \boldsymbol{Z}\end{smallmatrix}\right] \sim \mathcal{N}_{2p_1}(\mu_{\boldsymbol{TZ}}, \Sigma_{\boldsymbol{TZ}})$, where

$$\mu_{\boldsymbol{TZ}} = \mathbb{E}\begin{bmatrix}\boldsymbol{T} \\ \boldsymbol{Z}\end{bmatrix} = \begin{bmatrix}\mathbb{E}(\boldsymbol{T}) \\ \mathbb{E}(\boldsymbol{Z})\end{bmatrix} = \begin{bmatrix}\boldsymbol{0} \\ \boldsymbol{\omega\delta}\end{bmatrix}$$

and

$$\begin{aligned}
\Sigma_{\boldsymbol{TZ}} &= \begin{bmatrix}\mathbb{V}(\boldsymbol{T}) & Cov(\boldsymbol{T},\boldsymbol{Z}) \\ Cov(\boldsymbol{Z},\boldsymbol{T}) & \mathbb{V}(\boldsymbol{Z})\end{bmatrix} \\
&= \begin{bmatrix}\sigma^2\boldsymbol{Q}_{11.2}^{-1} & \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} \\ \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} & \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\end{bmatrix} \\
&= \sigma^2\begin{bmatrix}\boldsymbol{Q}_{11.2}^{-1} & \boldsymbol{\Omega} \\ \boldsymbol{\Omega} & \boldsymbol{\Omega}\end{bmatrix}.
\end{aligned}$$

Here,

$$\begin{aligned}
Cov(\boldsymbol{T},\boldsymbol{Z}) &= Cov[\boldsymbol{T}, \boldsymbol{\omega}(\boldsymbol{U}+\boldsymbol{\delta})] \\
&= Cov[\boldsymbol{T}, \boldsymbol{U}+\boldsymbol{\delta}]\boldsymbol{\omega}^\top \\
&= Cov[\boldsymbol{T}, \boldsymbol{U}]\boldsymbol{\omega}^\top + Cov[\boldsymbol{T}, \boldsymbol{\delta}]\boldsymbol{\omega}^\top \\
&= \sigma^2(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1})\boldsymbol{\omega}^\top + [\mathbb{E}(\boldsymbol{T\delta}) - \mathbb{E}(\boldsymbol{T})\mathbb{E}(\boldsymbol{\delta})]\boldsymbol{\omega}^\top \\
&= \sigma^2(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1})(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12})^\top + \boldsymbol{0} \\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} = \sigma^2\boldsymbol{\Omega}.
\end{aligned}$$

(5) Lastly, we consider

$$\begin{bmatrix}\boldsymbol{W}_n \\ \boldsymbol{Z}_n\end{bmatrix} = \begin{bmatrix}\boldsymbol{T}_n - \boldsymbol{\omega}_n\sqrt{n}\widehat{\boldsymbol{\beta}}_2^{\text{UE}} \\ \boldsymbol{T}_n - (\boldsymbol{T}_n - \boldsymbol{\omega}_n\sqrt{n}\widehat{\boldsymbol{\beta}}_2^{\text{UE}})\end{bmatrix} = \boldsymbol{T}_n\begin{bmatrix}\boldsymbol{I}_{p_1} \\ \boldsymbol{0}_{p_1}\end{bmatrix} + \begin{bmatrix}-\boldsymbol{\omega}_n\sqrt{n}\widehat{\boldsymbol{\beta}}_2^{\text{UE}} \\ \boldsymbol{\omega}_n\sqrt{n}\widehat{\boldsymbol{\beta}}_2^{\text{UE}}\end{bmatrix},$$

which is a linear function of $\boldsymbol{T}_n$. Again, by Slutsky's theorem, the joint distribution of $\boldsymbol{W}_n$ and $\boldsymbol{Z}_n$ as $n \to \infty$ is $\left[\begin{smallmatrix}\boldsymbol{W}_n \\ \boldsymbol{Z}_n\end{smallmatrix}\right] \xrightarrow{d} \left[\begin{smallmatrix}\boldsymbol{W} \\ \boldsymbol{Z}\end{smallmatrix}\right] \sim \mathcal{N}_{2p_1}(\mu_{\boldsymbol{WZ}}, \Sigma_{\boldsymbol{WZ}})$, where

$$\mu_{\boldsymbol{YZ}} = \mathbb{E}\begin{bmatrix}\boldsymbol{W} \\ \boldsymbol{Z}\end{bmatrix} = \begin{bmatrix}\mathbb{E}(\boldsymbol{W}) \\ \mathbb{E}(\boldsymbol{Z})\end{bmatrix} = \begin{bmatrix}-\boldsymbol{\omega\delta} \\ \boldsymbol{\omega\delta}\end{bmatrix}$$

and

$$\Sigma_{\boldsymbol{WZ}} = \begin{bmatrix} \mathbb{V}(\boldsymbol{W}) & Cov(\boldsymbol{W},\boldsymbol{Z}) \\ Cov(\boldsymbol{Z},\boldsymbol{W}) & \mathbb{V}(\boldsymbol{Z}) \end{bmatrix} = \begin{bmatrix} \sigma^2\boldsymbol{Q}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} \end{bmatrix}$$
$$= \sigma^2 \begin{bmatrix} \boldsymbol{Q}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega} \end{bmatrix}.$$

Here,

$$\begin{aligned} Cov(\boldsymbol{W},\boldsymbol{Z}) &= Cov[\boldsymbol{T} - \boldsymbol{\omega}(\boldsymbol{U}+\boldsymbol{\delta}), \boldsymbol{\omega}(\boldsymbol{U}+\boldsymbol{\delta})] \\ &= Cov[\boldsymbol{T}, \boldsymbol{\omega}(\boldsymbol{U}+\boldsymbol{\delta})] - Cov[\boldsymbol{\omega}(\boldsymbol{U}+\boldsymbol{\delta}), \boldsymbol{\omega}(\boldsymbol{U}+\boldsymbol{\delta})] \\ &= Cov[\boldsymbol{T},\boldsymbol{U}](\boldsymbol{\omega})^\top - \mathbb{V}[\boldsymbol{\omega}(\boldsymbol{U}+\boldsymbol{\delta})] \\ &= Cov[\boldsymbol{T},\boldsymbol{U}](\boldsymbol{\omega})^\top - \boldsymbol{\omega}\mathbb{V}[\boldsymbol{U}+\boldsymbol{\delta}](\boldsymbol{\omega})^\top \\ &= \sigma^2(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1})(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12})^\top \\ &\quad - (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12})\sigma^2\boldsymbol{Q}_{22.1}^{-1}(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12})^\top \\ &= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} - \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} \\ &= \mathbf{0}. \end{aligned}$$

□

Corollary 1. *Under usual regularity conditions and the sequence of local alternatives, as* $n \to \infty$*,*

$$\boldsymbol{Z}_n^* = \sqrt{n}\sigma^{-1}\boldsymbol{\Omega}_n^{-\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{UE} - \widehat{\boldsymbol{\beta}}_1^{RE}) \xrightarrow{D} \boldsymbol{Z}^* \sim \mathcal{N}_{p_1}(\sigma^{-1}\boldsymbol{\Omega}^{-\frac{1}{2}}\boldsymbol{\omega\delta}, \boldsymbol{I}_{p_1}),$$

where $\boldsymbol{\Omega}_n = \boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}\boldsymbol{C}_{22.1}^{-1}\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}$ *and* $\boldsymbol{\Omega}_n \xrightarrow{P} \boldsymbol{\Omega}$.

Proof. Note that

$$\begin{aligned} \boldsymbol{Z}_n^* &= \sqrt{n}\sigma^{-1}\boldsymbol{\Omega}_n^{-\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \widehat{\boldsymbol{\beta}}_1^{\text{RE}}) \\ &= \sigma^{-1}\boldsymbol{\Omega}_n^{-\frac{1}{2}}\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \widehat{\boldsymbol{\beta}}_1^{\text{RE}}) \\ &= \sigma^{-1}\boldsymbol{\Omega}_n^{-\frac{1}{2}}\boldsymbol{Z}_n, \end{aligned}$$

which is a linear function of $\boldsymbol{Z}_n$, by $\boldsymbol{\Omega}_n^{-\frac{1}{2}}$ being a continuous function of $\boldsymbol{\Omega}_n$ converging in probability to $\boldsymbol{\Omega}^{-\frac{1}{2}}$, or $\boldsymbol{\Omega}_n^{-\frac{1}{2}} \xrightarrow{P} \boldsymbol{\Omega}^{-\frac{1}{2}} = (\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1})^{-\frac{1}{2}}$. Therefore, as

$n \to \infty$ and by Slutsky's theorem, $\boldsymbol{Z}_n^* \xrightarrow{D} \boldsymbol{Z}^* \sim \mathcal{N}_{p_2}(\mu_{\boldsymbol{Z}^*}, \Sigma_{\boldsymbol{Z}^*})$. Here

$$\begin{aligned}
\mu_{\boldsymbol{Z}^*} &= \mathbb{E}(\boldsymbol{Z}^*) = \mathbb{E}[\sigma^{-1}\boldsymbol{\Omega}^{-\frac{1}{2}}\boldsymbol{Z}] \\
&= \sigma^{-1}\boldsymbol{\Omega}^{-\frac{1}{2}}\mathbb{E}(\boldsymbol{Z}) \\
&= \sigma^{-1}\boldsymbol{\Omega}^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta} \\
&= (\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta},
\end{aligned}$$

and

$$\begin{aligned}
\Sigma_{\boldsymbol{Z}^*} &= \mathbb{V}(\boldsymbol{Z}^*) = \mathbb{V}[\sigma^{-1}\boldsymbol{\Omega}_n^{-\frac{1}{2}}\boldsymbol{Z}] \\
&= \sigma^{-1}\boldsymbol{\Omega}^{-\frac{1}{2}}\mathbb{V}(\boldsymbol{Z})(\sigma^{-1}\boldsymbol{\Omega}^{-\frac{1}{2}})^\top \\
&= (\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\sigma^2\boldsymbol{\Omega}((\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}})^\top \\
&= (\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\sigma^2\boldsymbol{\Omega}(\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}} \\
&= \boldsymbol{I}_{p_1}.
\end{aligned}$$

□

It can be noted that the asymptotic distribution of $\boldsymbol{Z}_n$ has covariance matrix $\sigma^2\boldsymbol{\Omega}$, while $\boldsymbol{Z}_n^*$ has covariance matrix $\boldsymbol{I}_{p_1}$. The purpose of transformation is to stabilize the covariance matrix of $\boldsymbol{Z}_n$ so that we may use Theorem 3.6.2 for computing the ADQB and ADQR of the estimators. The relation between $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ is

$$\boldsymbol{Z} = \sigma\boldsymbol{\Omega}^{\frac{1}{2}}\boldsymbol{Z}^* = (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*. \tag{4.1}$$

4.1.1 Asymptotic Distributional Quadratic Bias

Next, the expression of the asymptotic distributional bias (ADB) is presented. From Theorems 3.6.1 and 3.6.2, the ADB expression is derived as the following Theorem.

Theorem 4.1.2. *Under the sequence $\{K_n\}$ and usual regularity condition, as $n \to \infty$. The ADBs of all estimators are*

$$\begin{aligned}
ADB(\widehat{\boldsymbol{\beta}}_1^{UE}) &= \boldsymbol{0}, \\
ADB(\widehat{\boldsymbol{\beta}}_1^{RE}) &= \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}, \\
ADB(\widehat{\boldsymbol{\beta}}_1^{LS}) &= \pi\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}, \\
ADB(\widehat{\boldsymbol{\beta}}_1^{PT}) &= \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta),
\end{aligned}$$

$$ADB(\widehat{\boldsymbol{\beta}}_1^{SP}) = \pi \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta} H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta),$$
$$ADB(\widehat{\boldsymbol{\beta}}_1^{S}) = c\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)],$$
$$ADB(\widehat{\boldsymbol{\beta}}_1^{S+}) = \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\{H_{p_2+2}(c;\Delta) + c\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta) I(\chi^2_{p_2+2}(\Delta) > c)]\}.$$

Here, $c = p_2 - 2$, $p_2 > 2$, $H_\nu(\cdot\,;\Delta)$ is a cumulative distribution function (CDF) of non-central chi-square with ν degrees of freedom and non-centrality parameter Δ, and $\mathbb{E}\left[\chi_\nu^{-2j}(\Delta)\right] = \int_0^\infty x^{-2j} d\phi_\nu(x;\Delta)$.

Proof.

$$\begin{aligned}
\text{ADB}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) &= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \boldsymbol{\beta}_1)] \\
&= \mathbb{E}\left[\lim_{n\to\infty}\boldsymbol{T}_n\right] = \mathbb{E}(\boldsymbol{T}) \\
&= \mathbf{0}.
\end{aligned}$$

$$\begin{aligned}
\text{ADB}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}}) &= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}} - \boldsymbol{\beta}_1)] \\
&= \mathbb{E}\left[\lim_{n\to\infty}\boldsymbol{W}_n\right] = \mathbb{E}(\boldsymbol{W}) \\
&= -\boldsymbol{\omega}\boldsymbol{\delta} = \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}.
\end{aligned}$$

$$\begin{aligned}
\text{ADB}(\widehat{\boldsymbol{\beta}}_1^{\text{LS}}) &= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{LS}} - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\pi\widehat{\boldsymbol{\beta}}_1^{\text{RE}} + (1-\pi)\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\pi\widehat{\boldsymbol{\beta}}_1^{\text{RE}} + \widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \pi\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \boldsymbol{\beta}_1) - \pi\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \widehat{\boldsymbol{\beta}}_1^{\text{RE}})] \\
&= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \boldsymbol{\beta}_1)] - \pi\lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}} - \widehat{\boldsymbol{\beta}}_1^{\text{RE}})] \\
&= \mathbb{E}\left[\lim_{n\to\infty}\boldsymbol{T}_n\right] - \pi\mathbb{E}\left[\lim_{n\to\infty}\boldsymbol{Z}_n\right] \\
&= \mathbb{E}(\boldsymbol{T}) - \pi\mathbb{E}(\boldsymbol{Z}) \\
&= -\pi\boldsymbol{\omega}\boldsymbol{\delta} \\
&= -\pi(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12})\boldsymbol{\delta} \\
&= \pi\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}.
\end{aligned}$$

$$
\begin{aligned}
\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}) &= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}} - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - (\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})I(\Lambda_n \le \lambda_\alpha) - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \boldsymbol{\beta}_1) - \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})I(\Lambda_n \le \lambda_\alpha)] \\
&= \lim_{n\to\infty}\mathbb{E}[\boldsymbol{T}_n - \boldsymbol{Z}_n I(\Lambda_n \le \lambda_\alpha)] \\
&= \mathbb{E}\left[\lim_{n\to\infty}\boldsymbol{T}_n\right] - \mathbb{E}\left[\lim_{n\to\infty}\boldsymbol{Z}_n I(\Lambda_n \le \lambda_\alpha)\right] \\
&= \mathbb{E}(\boldsymbol{T}) - \mathbb{E}[\boldsymbol{Z} I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\
&= -\mathbb{E}[(\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^* I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] && \because \text{ by (4.1)} \\
&= -(\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\mathbb{E}[\boldsymbol{Z}^* I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\
&= -(\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}(\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta}\mathbb{E}[I(\chi^2_{p_2+2}(\Delta) \le \chi^2_{p_2,\alpha})] && \because \text{ by (3.39)} \\
&= \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\mathbb{P}[\chi^2_{p_2+2}(\Delta) \le \chi^2_{p_2,\alpha}] \\
&= \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta).
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) &= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}} - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \pi(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})I(\Lambda_n \le \lambda_\alpha) - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \boldsymbol{\beta}_1) - \pi\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})I(\Lambda_n \le \lambda_\alpha)] \\
&= \lim_{n\to\infty}\mathbb{E}[\boldsymbol{T}_n - \pi\boldsymbol{Z}_n I(\Lambda_n \le \lambda_\alpha)] \\
&= \mathbb{E}\left[\lim_{n\to\infty}\boldsymbol{T}_n\right] - \pi\mathbb{E}\left[\lim_{n\to\infty}\boldsymbol{Z}_n I(\Lambda_n \le \lambda_\alpha)\right] \\
&= \mathbb{E}(\boldsymbol{T}) - \pi\mathbb{E}[\boldsymbol{Z} I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\
&= -\pi\mathbb{E}[(\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^* I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] && \because \text{ by (4.1)} \\
&= -\pi(\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\mathbb{E}[\boldsymbol{Z}^* I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\
&= -\pi(\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}(\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta}\mathbb{E}[I(\chi^2_{p_2+2}(\Delta) \le \chi^2_{p_2,\alpha})] && \because \text{ by (3.39)} \\
&= \pi\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\mathbb{P}[\chi^2_{p_2+2}(\Delta) \le \chi^2_{p_2,\alpha}] \\
&= \pi\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta).
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) &= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}} - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} + (1 - c\Lambda_n^{-1})(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} + \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} - c\Lambda_n^{-1}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \boldsymbol{\beta}_1) - c\Lambda_n^{-1}\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})] \\
&= \lim_{n\to\infty}\mathbb{E}[\boldsymbol{T}_n - c\Lambda_n^{-1}\boldsymbol{Z}_n] \\
&= \mathbb{E}\left[\lim_{n\to\infty}\boldsymbol{T}_n\right] - c\mathbb{E}\left[\lim_{n\to\infty}(\Lambda_n^{-1}\boldsymbol{Z}_n)\right] \\
&= \mathbb{E}(\boldsymbol{T}) - c\mathbb{E}(\boldsymbol{Z}\chi_{p_2}^{-2}(\Delta)) \\
&= -c\mathbb{E}[(\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*\chi_{p_2}^{-2}(\Delta)] \qquad \because \text{ by (4.1)} \\
&= -c(\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}(\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta}\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] \quad \because \text{ by (3.39)} \\
&= c\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)].
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}) &= \lim_{n\to\infty}\mathrm{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}+} - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}} - (1 - c\Lambda_n^{-1})(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})I(\Lambda_n \le c) - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}} - \boldsymbol{\beta}_1) - (1 - c\Lambda_n^{-1})\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})I(\Lambda_n \le c)] \\
&= \mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - \lim_{n\to\infty}\mathbb{E}[(1 - c\Lambda_n^{-1})\boldsymbol{Z}_n I(\Lambda_n \le c)] \\
&= \mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - \mathbb{E}\left[\lim_{n\to\infty}(1 - c\Lambda_n^{-1})\boldsymbol{Z}_n I(\Lambda_n \le c)\right] \\
&= \mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - \mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{Z}I(\chi_{p_2}^{2}(\Delta) \le c)] \\
&= \mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - \mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta))(\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^* I(\chi_{p_2}^{2}(\Delta) \le c)] \quad \because \text{ by (4.1)} \\
&= \mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - \left\{\begin{matrix} (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}(\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta} \\ \mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta) \le c)] \end{matrix}\right\} \qquad \because \text{ by (3.39)} \\
&= \mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\mathrm{E}\,[(1 - c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta) \le c)]. \\
&= c\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta) \le c)] \\
&= \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\left[c\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] + \mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta) \le c)]\right] \\
&= \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\begin{bmatrix} c\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] + \mathbb{E}[I(\chi_{p_2+2}^{2}(\Delta) \le c)] \\ -c\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)I(\chi_{p_2+2}^{2}(\Delta) \le c)] \end{bmatrix} \\
&= \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\{H_{p_2+2}(c;\Delta) + c\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)][1 - I(\chi_{p_2+2}^{2}(\Delta) \le c)]\} \\
&= \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\{H_{p_2+2}(c;\Delta) + c\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)I(\chi_{p_2+2}^{2}(\Delta) > c)]\}.
\end{aligned}
$$

□

For making a possible comparison, we present the asymptotic distributional quadratic bias (ADQB) of all estimators as follows:

Theorem 4.1.3. *Suppose that the conditions of Theorem 4.1.2 hold. The ADQBs of the estimators are*

$$
\begin{aligned}
ADQB(\widehat{\boldsymbol{\beta}}_1^{UE}) &= 0,\\
ADQB(\widehat{\boldsymbol{\beta}}_1^{RE}) &= \Delta^*,\\
ADQB(\widehat{\boldsymbol{\beta}}_1^{LS}) &= \pi^2\Delta^*,\\
ADQB(\widehat{\boldsymbol{\beta}}_1^{PT}) &= \Delta^*[H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)]^2,\\
ADQB(\widehat{\boldsymbol{\beta}}_1^{SP}) &= \pi^2\Delta^*[H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)]^2,\\
ADQB(\widehat{\boldsymbol{\beta}}_1^{S}) &= c^2\Delta^*\{\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)]\}^2,\\
ADQB(\widehat{\boldsymbol{\beta}}_1^{S+}) &= \Delta^*\{H_{p_2+2}(c;\Delta) + c\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)I(\chi^2_{p_2+2}(\Delta) > c)]\}^2,
\end{aligned}
$$

where $c = p_2 - 2$, $p_2 > 2$, $\Delta^* = \sigma^{-2}\boldsymbol{\delta}^\top\boldsymbol{Q}^*\boldsymbol{\delta}$ *and* $\boldsymbol{Q}^* = \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{11.2}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}$.

Proof.

$$
\begin{aligned}
\mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) &= \left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}})\right]^\top \sigma^{-2}\boldsymbol{Q}_{11.2}\left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}})\right]\\
&= 0.
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) &= \left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})\right]^\top \sigma^{-2}\boldsymbol{Q}_{11.2}\left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})\right]\\
&= (\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^\top(\sigma^{-2}\boldsymbol{Q}_{11.2})(\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})\\
&= \sigma^{-2}\boldsymbol{\delta}^\top\boldsymbol{Q}_{12}^\top(\boldsymbol{Q}_{11}^{-1})^\top\boldsymbol{Q}_{11.2}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\\
&= \sigma^{-2}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{11.2}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\\
&= \sigma^{2}\boldsymbol{\delta}^\top\boldsymbol{Q}^*\boldsymbol{\delta} = \Delta^*.
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) &= \left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}})\right]^\top \sigma^{-2}\boldsymbol{Q}_{11.2}\left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}})\right]\\
&= (\pi\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^\top(\sigma^{-2}\boldsymbol{Q}_{11.2})(\pi\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})\\
&= \pi^2(\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^\top(\sigma^{-2}\boldsymbol{Q}_{11.2})(\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})\\
&= \pi^2\Delta^*.
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}) &= \left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}})\right]^{\top} \sigma^{-2}\boldsymbol{Q}_{11.2}\left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}})\right] \\
&= [\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta} H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)]^{\top}(\sigma^{-2}\boldsymbol{Q}_{11.2}) \\
&\quad \times [\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta} H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)] \\
&= (\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\delta)^{\top}(\sigma^{-2}\boldsymbol{Q}_{11.2})(\boldsymbol{Q}_{11}^{-1}\mathrm{Q}_{12}\boldsymbol{\delta})[H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)]^2 \\
&= \Delta^{*}[H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)]^2.
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) &= \left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}})\right]^{\top} \sigma^{-2}\boldsymbol{Q}_{11.2}\left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}})\right] \\
&= [\pi\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta} H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)]^{\top}(\sigma^{-2}\boldsymbol{Q}_{11.2}) \\
&\quad \times [\pi\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta} H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)] \\
&= \pi^2(\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^{\top}(\sigma^{-2}\boldsymbol{Q}_{11.2})(\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})[H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)]^2 \\
&= \pi^2\Delta^{*}[H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)]^2.
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) &= \left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}})\right]^{\top} \sigma^{-2}\boldsymbol{Q}_{11.2}\left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}})\right] \\
&= [(p_2-2)\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)]]^{\top}(\sigma^{-2}\boldsymbol{Q}_{11.2}) \\
&\quad \times [(p_2-2)\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] \\
&= (p_2-2)^2(\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^{\top}(\sigma^{-2}\boldsymbol{Q}_{11.2})(\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})\{\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)]\}^2 \\
&= (p_2-2)^2\Delta^{*}\{\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)]\}^2 \\
&= c^2\Delta^{*}\{\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)]\}^2.
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}) &= \left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}})\right]^{\top} (\sigma^{-2}\boldsymbol{Q}_{11.2})\left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}})\right] \\
&= \left\{\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\left[H_{p_2+2}(c;\Delta)+c\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)I(\chi^2_{p_2+2}(\Delta)>c)]\right]\right\}^{\top} \\
&\quad \times (\sigma^{-2}\boldsymbol{Q}_{11.2}) \\
&\quad \times \left\{\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\left[H_{p_2+2}(c;\Delta)+c\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)I(\chi^2_{p_2+2}(\Delta)>c)]\right]\right\}. \\
&= (\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^{\top}(\sigma^{-2}\boldsymbol{Q}_{11.2})(\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}) \\
&\quad \times \{H_{p_2+2}(c;\Delta)+c\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)I(\chi^2_{p_2+2}(\Delta)>c)]\}^2 \\
&= \Delta^{*}\{H_{p_2+2}(c;\Delta)+c\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)I(\chi^2_{p_2+2}(\Delta)>c)]\}^2.
\end{aligned}
$$

□

The ADQBs of the suggested estimators are different depending on the values of p_2, α, π, and Δ^*. The plots of the ADQBs show this behavior clearly. Figures 4.1 to 4.2 show the graphs of ADQBs of the estimators for different configurations of $p_1 = 3$, $p_2 = 3$ and 7, $\alpha = 0.01$ and 0.05, and $\pi = 0.25, 0.50$, and 0.75. The other values of p_2, α, π, and Δ^* were also studied, however, their results were similar. Hence, we did not reproduce the graph here. The results of ADQB of the estimators can be summarized as follows:

(i) Only $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ is an unbiased estimator for the parameter $\boldsymbol{\beta}_1$.

(ii) All remaining estimators are also unbiased if $\Delta^* = 0$.

(iii) For $\Delta^* > 0$, the ADQBs of both $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ are an unbounded function of Δ^*. The ADQB functions of all remaining estimators are bounded in Δ^*.

(iv) The ADQB of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ is a function of Δ^* and π. The $\mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) = \pi^2 \mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})$ and $\mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) < \mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})$ if $0 < \pi < 1$.

(v) The ADQBs of both $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ increase at first, reach a maximum, and then decrease to become equal with the $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ as $\Delta^* \to \infty$.

(vi) Since $\mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) = \pi^2 \mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}})$, then $\mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) < \mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}})$ when $0 < \pi < 1$.

(vii) For fixed Δ^*, the ADQB of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ is similar to that of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ when $\pi = 0$, and to that of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$ when $\pi = 1$.

(viii) At $\Delta^* = 0$, the ADQBs of both $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ start from zero, increase to a point, and then decrease to zero as $\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)]$ is a decreasing log-convex function of Δ.

(ix) The properties of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ are similar to those of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, however, the ADQB of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ is smaller than or equal to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$; we can derive as follows:

For $p_2 \geq 3$,

$$\begin{aligned}\text{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\text{S}}) &= c^2\Delta^*\{\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)]\}^2 \\ &= \Delta^*\{c\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)]\}^2 \\ &= \Delta^*\left\{\begin{array}{c} c\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)I(\chi^2_{p_2+2}(\Delta) > c)] \\ +c\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)I(\chi^2_{p_2+2}(\Delta) \leq c)] \end{array}\right\}^2.\end{aligned}$$

When $\chi^2_{p_2+2}(\Delta) > c$, we get

$$\text{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\text{S}}) = \Delta^*\{c\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)I(\chi^2_{p_2+2}(\Delta) > c)]\}^2 = \text{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\text{S}^+})$$

due to $H_{p_2+2}(c;\Delta) = 0$ and $\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)I(\chi^2_{p_2+2}(\Delta) \leq c)] = 0$. If $\chi^2_{p_2+2}(\Delta) \leq c$, we obtain

$$\text{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\text{S}}) = \Delta^*\{c\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)I(\chi^2_{p_2+2}(\Delta) \leq c)]\}^2$$

and

$$\text{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\text{S}^+}) = \Delta^*\{H_{p_2+2}(c;\Delta)\}^2 = \Delta^*\{\mathbb{E}[I(\chi^2_{p_2+2}(\Delta) \leq c)]\}^2.$$

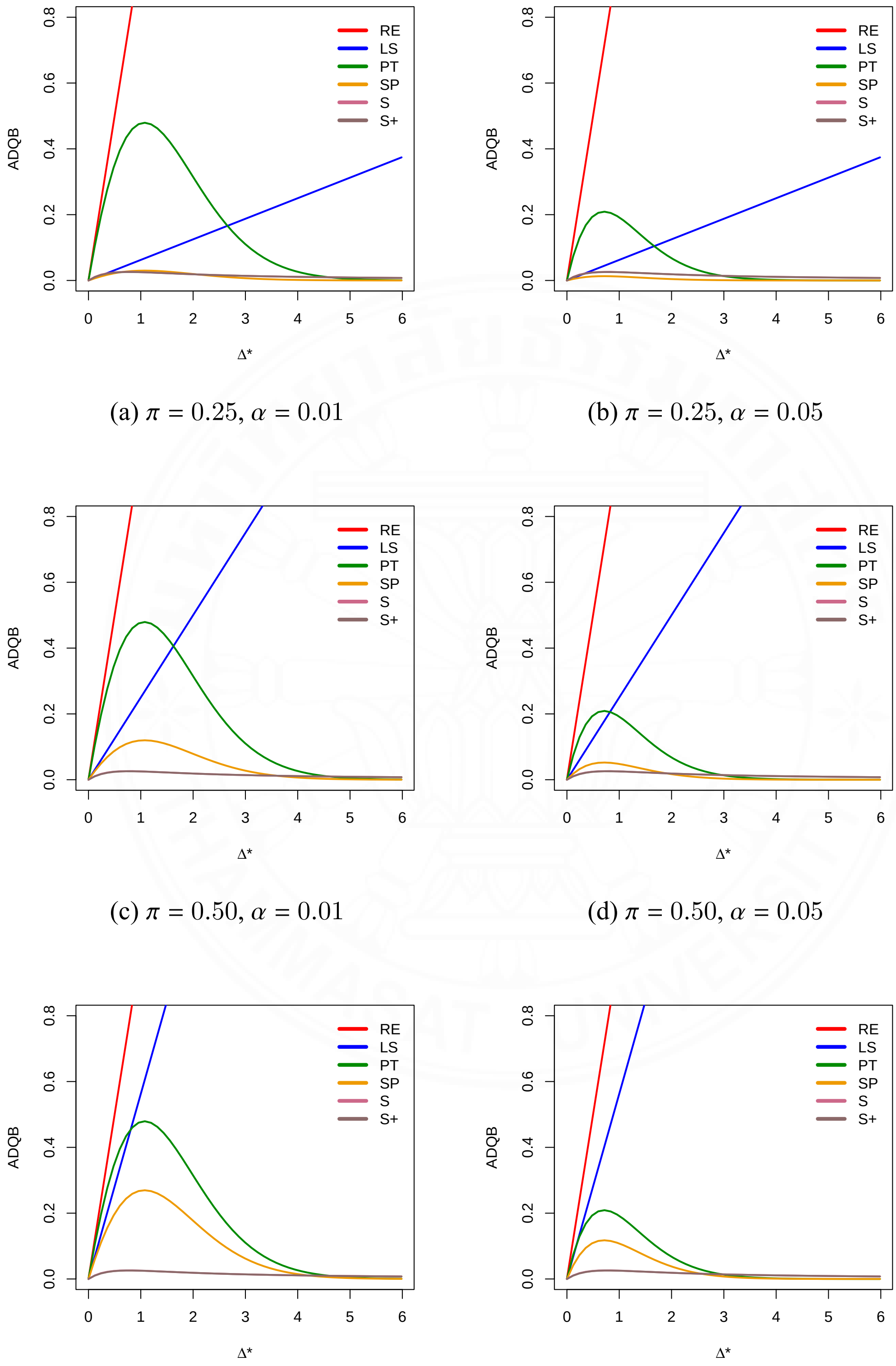
Since $0 < \chi^2_{p_2+2}(\Delta) \leq c$, we therefore get

$$\begin{aligned}1 &\leq \frac{c}{\chi^2_{p_2+2}(\Delta)} \\ \mathbb{E}[I(\chi^2_{p_2+2}(\Delta) \leq c)] &\leq \mathbb{E}\left[\frac{c}{\chi^2_{p_2+2}(\Delta)}I(\chi^2_{p_2+2}(\Delta) \leq c)\right] \\ \Delta^*\{\mathbb{E}[I(\chi^2_{p_2+2}(\Delta) \leq c)]\}^2 &\leq \Delta^*\left\{\mathbb{E}\left[\frac{c}{\chi^2_{p_2+2}(\Delta)}I(\chi^2_{p_2+2}(\Delta) \leq c)\right]\right\}^2 \\ \Delta^*\{H_{p_2+2}(c;\Delta)\}^2 &\leq \Delta^*\{c\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)I(\chi^2_{p_2+2}(\Delta) \leq c)]\}^2.\end{aligned}$$

Hence, $\text{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\text{S}^+}) \leq \text{ADQB}(\widehat{\boldsymbol{\beta}}_1^{\text{S}})$ for all $\Delta^* > 0$.

4.1.2 Asymptotic Distributional Quadratic Risk

To study the asymptotic distributional quadratic risk (ADQR) of an estimator, we display the asymptotic mean squared error matrix (AMSEM) of the proposed estimators which are given by as following Theorem:



(a) $\pi = 0.25$, $\alpha = 0.01$

(b) $\pi = 0.25$, $\alpha = 0.05$

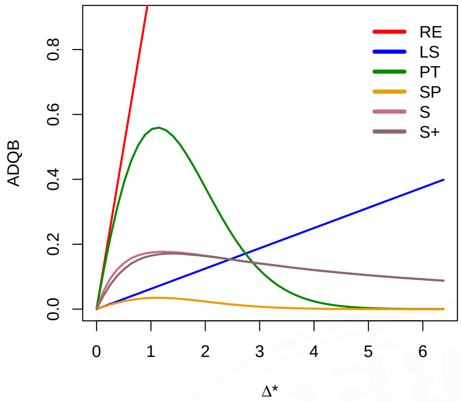
(c) $\pi = 0.50$, $\alpha = 0.01$

(d) $\pi = 0.50$, $\alpha = 0.05$

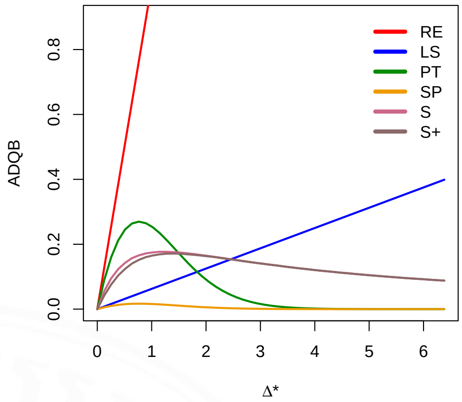
(e) $\pi = 0.75$, $\alpha = 0.01$

(f) $\pi = 0.75$, $\alpha = 0.05$

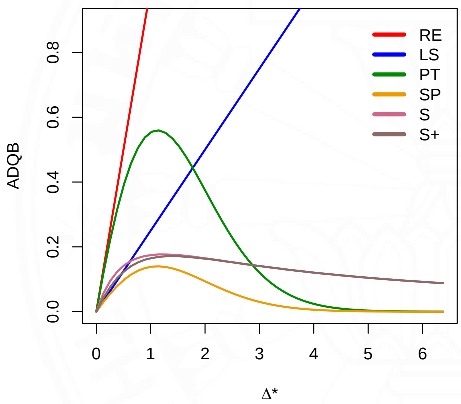
Figure 4.1 ADQB curves of the suggested estimators for nonlinear regression model with $p_1 = 3$ and $p_2 = 3$



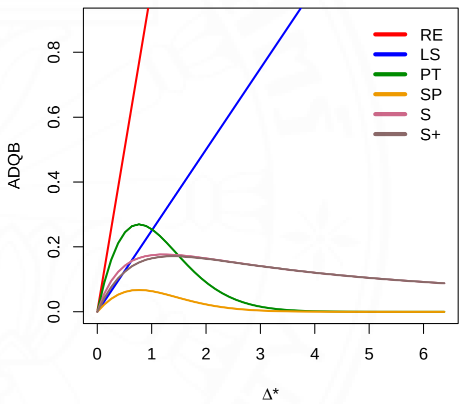
(a) $\pi = 0.25, \alpha = 0.01$



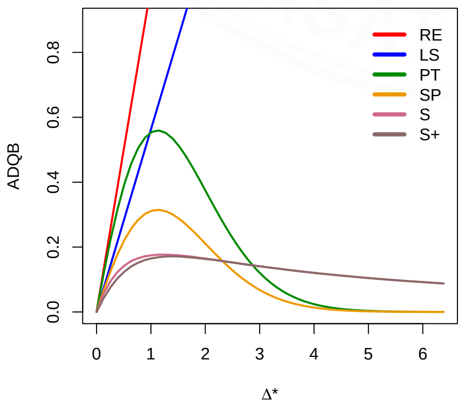
(b) $\pi = 0.25, \alpha = 0.05$



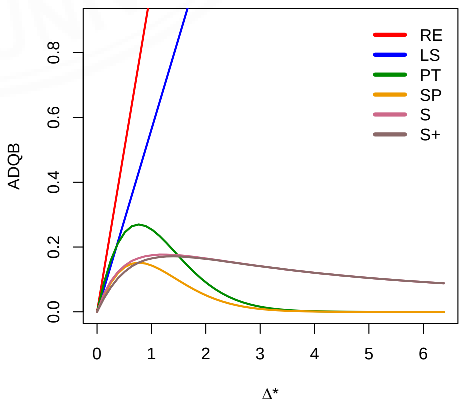
(c) $\pi = 0.50, \alpha = 0.01$



(d) $\pi = 0.50, \alpha = 0.05$



(e) $\pi = 0.75, \alpha = 0.01$



(f) $\pi = 0.75, \alpha = 0.05$

Figure 4.2 ADQB curves of the suggested estimators for nonlinear regression model with $p_1 = 3$ and $p_2 = 7$

Theorem 4.1.4. *Under the sequence of local alternative $\{K_n\}$ and usual regularity conditions, as $n \to \infty$ the AMSEMs of the estimators are*

$$
\begin{aligned}
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{UE}) &= \sigma^2 \boldsymbol{Q}_{11.2}^{-1}, \\
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{RE}) &= \sigma^2 \boldsymbol{Q}_{11}^{-1} + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}, \\
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{LS}) &= \sigma^2 \boldsymbol{Q}_{11.2}^{-1} - \pi(2-\pi)\sigma^2 \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} + \pi^2 \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}, \\
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{PT}) &= \sigma^2 \boldsymbol{Q}_{11.2}^{-1} - \sigma^2 \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}[2H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)], \\
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{SP}) &= \sigma^2 \boldsymbol{Q}_{11.2}^{-1} - \pi(2-\pi)\sigma^2 \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}[2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - \pi(2-\pi) H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)], \\
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{S}) &= \sigma^2 \boldsymbol{Q}_{11.2}^{-1} - c\sigma^2 \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}(2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] - c\mathbb{E}[\chi^{-4}_{p_2+2}(\Delta)]) \\
&\quad + c\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}(2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] - 2\mathbb{E}[\chi^{-2}_{p_2+4}(\Delta)] + c\mathbb{E}[\chi^{-4}_{p_2+4}(\Delta)]), \\
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{S+}) &= \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}_1^{S}) - \sigma^2 \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1 - c\chi^{-2}_{p_2+2}(\Delta))^2 I(\chi^2_{p_2+2}(\Delta) \le c)] \\
&\quad - \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1 - c\chi^{-2}_{p_2+4}(\Delta))^2 I(\chi^2_{p_2+4}(\Delta) \le c)] \\
&\quad + 2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1 - c\chi^{-2}_{p_2+2}(\Delta)) I(\chi^2_{p_2+2}(\Delta) \le c)].
\end{aligned}
$$

Proof. Applying the AMSEM of an estimator $\widehat{\boldsymbol{\beta}}_1^*$ defined as Equation (3.38), we have

$$
\begin{aligned}
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) &= \lim_{n\to\infty} \mathbb{E}\left[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \boldsymbol{\beta}_1)\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \boldsymbol{\beta}_1)^\top\right] \\
&= \mathbb{E}\lim_{n\to\infty}(\boldsymbol{T}_n\boldsymbol{T}_n^\top) = \mathbb{E}(\boldsymbol{T}\boldsymbol{T}^\top) \\
&= \mathbb{V}(\boldsymbol{T}) + \mathbb{E}(\boldsymbol{T})\mathrm{E}(\boldsymbol{T}^\top) \\
&= \sigma^2 \boldsymbol{Q}_{11.2}^{-1}.
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) &= \lim_{n\to\infty} \mathbb{E}\left[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} - \boldsymbol{\beta}_1)\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} - \boldsymbol{\beta}_1)^\top\right] \\
&= \mathbb{E}\lim_{n\to\infty}(\boldsymbol{W}_n\boldsymbol{W}_n^\top) = \mathbb{E}(\boldsymbol{W}\boldsymbol{W}^\top) \\
&= \mathbb{V}(\boldsymbol{W}) + \mathbb{E}(\boldsymbol{W})\mathbb{E}(\boldsymbol{W}^\top) \\
&= \sigma^2 \boldsymbol{Q}_{11}^{-1} + (-\boldsymbol{\omega}\boldsymbol{\delta})(-\boldsymbol{\omega}\boldsymbol{\delta})^\top \\
&= \sigma^2 \boldsymbol{Q}_{11}^{-1} + (\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})(\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^\top \\
&= \sigma^2 \boldsymbol{Q}_{11}^{-1} + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}.
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) &= \lim_{n\to\infty}\mathbb{E}\left[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}-\boldsymbol{\beta}_1)\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}-\boldsymbol{\beta}_1)^\top\right]\\
&= \mathbb{E}\lim_{n\to\infty}[(\boldsymbol{T}_n-\pi\boldsymbol{Z}_n)(\boldsymbol{T}_n-\pi\boldsymbol{Z}_n)^\top]\\
&= \mathbb{E}[(\boldsymbol{T}-\pi\boldsymbol{Z})(\boldsymbol{T}-\pi\boldsymbol{Z})^\top]\\
&= \mathbb{E}[\boldsymbol{T}\boldsymbol{T}^\top-\pi\boldsymbol{T}\boldsymbol{Z}^\top-\pi\boldsymbol{Z}\boldsymbol{T}^\top+\pi^2\boldsymbol{Z}\boldsymbol{Z}^\top]\\
&= \underbrace{\mathbb{E}[\boldsymbol{T}\boldsymbol{T}^\top]}_{\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}})}-2\pi\underbrace{\mathbb{E}[\boldsymbol{T}\boldsymbol{Z}^\top]}_{\mathrm{E}_1}+\pi^2\underbrace{\mathbb{E}[\boldsymbol{Z}\boldsymbol{Z}^\top]}_{\mathrm{E}_2},
\end{aligned}
$$

where

$$
\begin{aligned}
\mathrm{E}_2 &= \mathbb{E}[\boldsymbol{Z}\boldsymbol{Z}^\top]\\
&= \mathbb{V}(\boldsymbol{Z})+\mathbb{E}(\boldsymbol{Z})\mathbb{E}(\boldsymbol{Z}^\top)\\
&= \sigma^2\boldsymbol{\Omega}+(\boldsymbol{\omega}\boldsymbol{\delta})(\boldsymbol{\omega}\boldsymbol{\delta})^\top\\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}+(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^\top\\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}+\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}.
\end{aligned}
$$

By law of total expectation and conditional expectation of a multivariate normal distribution in Theorem 3.6.3, we may rewrite E_1 as

$$
\begin{aligned}
\mathrm{E}_1 &= \mathbb{E}[\boldsymbol{T}\boldsymbol{Z}^\top]\\
&= \mathbb{E}[\mathbb{E}(\boldsymbol{T}\boldsymbol{Z}^\top|\boldsymbol{Z})]\\
&= \mathbb{E}[\mathbb{E}(\boldsymbol{T}|\boldsymbol{Z})\boldsymbol{Z}^\top]\\
&= \mathbb{E}[\{\mathbb{E}(\boldsymbol{T})+Cov(\boldsymbol{T},\boldsymbol{Z})[\mathbb{V}(\boldsymbol{Z})]^{-1}(\boldsymbol{Z}-\mathbb{E}(\boldsymbol{Z})\}\boldsymbol{Z}^\top]\\
&= \mathbb{E}[\{\boldsymbol{0}+(\sigma^2\boldsymbol{\Omega})(\sigma^2\boldsymbol{\Omega})^{-1}(\boldsymbol{Z}-(\boldsymbol{\omega}\boldsymbol{\delta}))\}\boldsymbol{Z}^\top]\\
&= \mathbb{E}[(\boldsymbol{Z}-(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}))\boldsymbol{Z}^\top]\\
&= \mathbb{E}[(\boldsymbol{Z}+\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})\boldsymbol{Z}^\top]\\
&= \mathbb{E}(\boldsymbol{Z}\boldsymbol{Z}^\top)+\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\mathbb{E}(\boldsymbol{Z}^\top)\\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}+\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}+(\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^\top\\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}+\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}.
\end{aligned}
$$

Then, the AMSEM of $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$ becomes

$$\begin{aligned}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\text{LS}}) &= \sigma^2\boldsymbol{Q}_{11.2}^{-1} - 2\pi\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} \\ &\quad + \pi^2(\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}) \\ &= \sigma^2\boldsymbol{Q}_{11.2}^{-1} - \pi(2-\pi)\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} + \pi^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}.\end{aligned}$$

The $\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\text{PT}})$ can be written as

$$\begin{aligned}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\text{PT}}) &= \lim_{n\to\infty}\mathbb{E}\left[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{PT}} - \boldsymbol{\beta}_1)\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\text{PT}} - \boldsymbol{\beta}_1)^\top\right] \\ &= \mathbb{E}\lim_{n\to\infty}[(\boldsymbol{T}_n - \boldsymbol{Z}_n I(\Lambda_n \le \lambda_\alpha))(\boldsymbol{T}_n - \boldsymbol{Z}_n I(\Lambda_n \le \lambda_\alpha))^\top] \\ &= \mathbb{E}[(\boldsymbol{T} - \boldsymbol{Z} I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha}))(\boldsymbol{T} - \boldsymbol{Z} I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha}))^\top] \\ &= \mathbb{E}[\boldsymbol{T}\boldsymbol{T}^\top - 2\boldsymbol{T}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha}) + \boldsymbol{Z}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\ &= \underbrace{\mathbb{E}[\boldsymbol{T}\boldsymbol{T}^\top]}_{\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\text{UE}})} - 2\underbrace{\mathbb{E}[\boldsymbol{T}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})]}_{\text{E}_3} + \underbrace{\mathbb{E}[\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})]}_{\text{E}_4},\end{aligned}$$

Using Equations (4.1) and (3.40), we have

$$\begin{aligned}\text{E}_4 &= \mathbb{E}[\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\ &= \mathbb{E}[(\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*)^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\ &= (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\mathbb{E}[\boldsymbol{Z}^*(\boldsymbol{Z}^*)^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})]((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top \\ &= (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\begin{bmatrix}\mathbb{V}(\boldsymbol{Z}^*)\mathbb{E}[I(\chi^2_{p_2+2}(\Delta) \le \chi^2_{p_2,\alpha})] \\ +\mathbb{E}(\boldsymbol{Z}^*)\mathbb{E}(\boldsymbol{Z}^*)^\top\mathbb{E}[I(\chi^2_{p_2+4}(\Delta) \le \chi^2_{p_2,\alpha})]\end{bmatrix}((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top \\ &= (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\begin{bmatrix}\boldsymbol{I}_{p_2}\mathbb{E}[I(\chi^2_{p_2+2}(\Delta) \le \chi^2_{p_2,\alpha})] \\ +(\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta}((\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta})^\top\mathbb{E}[I(\chi^2_{p_2+4}(\Delta) \le \chi^2_{p_2,\alpha})]\end{bmatrix}((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top \\ &= (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top\mathbb{E}[I(\chi^2_{p_2+2}(\Delta) \le \chi^2_{p_2,\alpha})] \\ &\quad + (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}(\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta}(\boldsymbol{\omega}\boldsymbol{\delta})^\top((\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}})^\top((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top\mathbb{E}[I(\chi^2_{p_2+4}(\Delta) \le \chi^2_{p_2,\alpha})] \\ &= \sigma^2\boldsymbol{\Omega}\mathbb{E}[I(\chi^2_{p_2+2}(\Delta) \le \chi^2_{p_2,\alpha})] + \boldsymbol{\omega}\boldsymbol{\delta}(\boldsymbol{\omega}\boldsymbol{\delta})^\top\mathbb{E}[I(\chi^2_{p_2+4}(\Delta) \le \chi^2_{p_2,\alpha})] \\ &= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{P}(\chi^2_{p_2+2}(\Delta) \le \chi^2_{p_2,\alpha}) \\ &\quad + (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^\top\mathbb{P}(\chi^2_{p_2+4}(\Delta) \le \chi^2_{p_2,\alpha}) \\ &= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta).\end{aligned}$$

Using the rule of conditional expectation in Equation (3.41), E_3 therefore becomes

$$\begin{aligned}
\mathrm{E}_3 &= \mathbb{E}[\boldsymbol{T}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\
&= \mathbb{E}[\mathbb{E}(\boldsymbol{T}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})|\boldsymbol{Z})] \\
&= \mathbb{E}[\mathbb{E}(\boldsymbol{T}|\boldsymbol{Z})\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\
&= \mathbb{E}[\{\mathbb{E}(\boldsymbol{T}) + Cov(\boldsymbol{T},\boldsymbol{Z})[\mathbb{V}(\boldsymbol{Z})]^{-1}(\boldsymbol{Z} - \mathbb{E}(\boldsymbol{Z})\}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\
&= \mathbb{E}[\{\mathbf{0} + (\sigma^2\boldsymbol{\Omega})(\sigma^2\boldsymbol{\Omega})^{-1}(\boldsymbol{Z} - \boldsymbol{\omega}\boldsymbol{\delta})\}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\
&= \mathbb{E}[(\boldsymbol{Z} - (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}))\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\
&= \underbrace{\mathbb{E}[\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})]}_{\mathrm{E}_4} - (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})\underbrace{\mathbb{E}[\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})]}_{\mathrm{E}_5}.
\end{aligned}$$

Applying Equations (4.1) and (3.39), we get

$$\begin{aligned}
\mathrm{E}_5 &= \mathbb{E}[\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\
&= \mathbb{E}[((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*)^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\
&= ((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top((\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta})^\top\mathbb{E}[I(\chi^2_{p_2+2}(\Delta) \le \chi^2_{p_2,\alpha})] \\
&= (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^\top\mathbb{P}(\chi^2_{p_2+2}(\Delta) \le \chi^2_{p_2,\alpha}) \\
&= (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^\top H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta).
\end{aligned}$$

Replacing E_4 and E_5 in E_3, we obtain

$$\begin{aligned}
\mathrm{E}_3 &= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad - (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^\top H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad - \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad - \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}[H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)].
\end{aligned}$$

Substituting E_3 and E_4 into $\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}})$, we then get

$$\begin{aligned}
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}) &= \sigma^2\boldsymbol{Q}_{11.2}^{-1} - 2\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + 2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}[H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)] \\
&\quad + \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta) \\
&= \sigma^2\boldsymbol{Q}_{11.2}^{-1} - \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}[2H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)].
\end{aligned}$$

Next, we consider the AMSEM of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, and we therefore obtain

$$\begin{aligned}
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) &= \lim_{n\to\infty}\mathbb{E}\left[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}} - \boldsymbol{\beta}_1)\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}} - \boldsymbol{\beta}_1)^\top\right] \\
&= \mathbb{E}\lim_{n\to\infty}[(\boldsymbol{T}_n - \pi\boldsymbol{Z}_n I(\Lambda_n \le \lambda_\alpha))(\mathrm{T}_n - \pi\boldsymbol{Z}_n I(\Lambda_n \le \lambda_\alpha))^\top] \\
&= \mathbb{E}[(\boldsymbol{T} - \pi\boldsymbol{Z}I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha}))(\boldsymbol{T} - \pi\boldsymbol{Z}I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha}))^\top] \\
&= \mathbb{E}[\boldsymbol{T}\boldsymbol{T}^\top - \pi\boldsymbol{T}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha}) - \pi\boldsymbol{Z}I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})\boldsymbol{T}^\top \\
&\qquad + \pi^2\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\
&= \mathbb{E}[\boldsymbol{T}\boldsymbol{T}^\top - 2\pi\boldsymbol{T}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha}) + \pi^2\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] \\
&= \underbrace{\mathbb{E}[\boldsymbol{T}\boldsymbol{T}^\top]}_{\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}})} - 2\pi\underbrace{\mathbb{E}[\boldsymbol{T}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})]}_{\mathrm{E}_3} + \pi^2\underbrace{\mathbb{E}[\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})]}_{\mathrm{E}_4} \\
&= \sigma^2\boldsymbol{Q}_{11.2}^{-1} - 2\pi\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + 2\pi\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}[H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)] \\
&\quad + \pi^2\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \pi^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta) \\
&= \sigma^2\boldsymbol{Q}_{11.2}^{-1} - \pi(2-\pi)\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\begin{bmatrix} 2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - 2\pi H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta) \\ +\pi^2 H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta) \end{bmatrix} \\
&= \sigma^2\boldsymbol{Q}_{11.2}^{-1} - \pi(2-\pi)\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}[2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - \pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)].
\end{aligned}$$

Let us regard $\mathbf{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}})$,

$$\begin{aligned}
\mathbf{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) &= \lim_{n\to\infty}\mathbb{E}\left[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}-\boldsymbol{\beta}_1)\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}-\boldsymbol{\beta}_1)^\top\right] \\
&= \mathbb{E}\lim_{n\to\infty}[(\boldsymbol{T}_n - c\boldsymbol{Z}_n\Lambda_n^{-1})(\boldsymbol{T}_n - c\boldsymbol{Z}_n\Lambda_n^{-1})^\top] \\
&= \mathbb{E}[(\boldsymbol{T} - c\boldsymbol{Z}\chi_{p_2}^{-2}(\Delta))(\boldsymbol{T} - c\boldsymbol{Z}\chi_{p_2}^{-2}(\Delta))^\top] \\
&= \mathbb{E}[\boldsymbol{T}\boldsymbol{T}^\top - c\boldsymbol{T}\boldsymbol{Z}^\top\chi_{p_2}^{-2}(\Delta) - c\boldsymbol{Z}\boldsymbol{T}^\top\chi_{p_2}^{-2}(\Delta) + c^2\boldsymbol{Z}\boldsymbol{Z}^\top\chi_{p_2}^{-4}(\Delta)] \\
&= \mathbb{E}[\boldsymbol{T}\boldsymbol{T}^\top - 2c\boldsymbol{T}\boldsymbol{Z}^\top\chi_{p_2}^{-2}(\Delta) + c^2\boldsymbol{Z}\boldsymbol{Z}^\top\chi_{p_2}^{-4}(\Delta)] \\
&= \underbrace{\mathbb{E}[\boldsymbol{T}\boldsymbol{T}^\top]}_{\mathbf{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}})} - 2c\underbrace{\mathbb{E}[\boldsymbol{T}\boldsymbol{Z}^\top\chi_{p_2}^{-2}(\Delta)]}_{\mathrm{E}_6} + c^2\underbrace{\mathbb{E}[\boldsymbol{Z}\boldsymbol{Z}^\top\chi_{p_2}^{-4}(\Delta)]}_{\mathrm{E}_7}.
\end{aligned}$$

Applying Equations (4.1) and (3.40) to E_7, we get

$$\begin{aligned}
\mathrm{E}_7 &= \mathbb{E}[\boldsymbol{Z}\boldsymbol{Z}^\top\chi_{p_2}^{-4}(\Delta)] \\
&= \mathbb{E}[(\sigma^2\mathbf{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*((\sigma^2\mathbf{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*)^\top\chi_{p_2}^{-4}(\Delta)] \\
&= (\sigma^2\mathbf{\Omega})^{\frac{1}{2}}\mathbb{E}[\boldsymbol{Z}^*(\boldsymbol{Z}^*)^\top\chi_{p_2}^{-4}(\Delta)]((\sigma^2\mathbf{\Omega})^{\frac{1}{2}})^\top \\
&= (\sigma^2\mathbf{\Omega})^{\frac{1}{2}}\left[\mathbb{V}(\boldsymbol{Z}^*)\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] + \mathbb{E}(\boldsymbol{Z}^*)\mathbb{E}(\boldsymbol{Z}^*)^\top\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]\right]((\sigma^2\mathbf{\Omega})^{\frac{1}{2}})^\top \\
&= (\sigma^2\mathbf{\Omega})^{\frac{1}{2}}\left[\boldsymbol{I}_{p_2}\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] + (\sigma^2\mathbf{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega\delta}((\sigma^2\mathbf{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega\delta})^\top\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]\right]((\sigma^2\mathbf{\Omega})^{\frac{1}{2}})^\top \\
&= (\sigma^2\mathbf{\Omega})^{\frac{1}{2}}((\sigma^2\mathbf{\Omega})^{\frac{1}{2}})^\top\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] \\
&\quad + (\sigma^2\mathbf{\Omega})^{\frac{1}{2}}(\sigma^2\mathbf{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega\delta}(\boldsymbol{\omega\delta})^\top((\sigma^2\mathbf{\Omega})^{-\frac{1}{2}})^\top((\sigma^2\mathbf{\Omega})^{\frac{1}{2}})^\top\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)] \\
&= \sigma^2\mathbf{\Omega}\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] + \boldsymbol{\omega\delta}(\boldsymbol{\omega\delta})^\top\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)] \\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)].
\end{aligned}$$

Again, using conditional expectation, E_6 becomes,

$$\begin{aligned}
\mathrm{E}_6 &= \mathbb{E}[\boldsymbol{T}\boldsymbol{Z}^\top\chi_{p_2}^{-2}(\Delta)] \\
&= \mathbb{E}[\mathbb{E}(\boldsymbol{T}\boldsymbol{Z}^\top\chi_{p_2}^{-2}(\Delta)|\boldsymbol{Z})] \\
&= \mathbb{E}[\mathbb{E}(\boldsymbol{T}|\boldsymbol{Z})\boldsymbol{Z}^\top\chi_{p_2}^{-2}(\Delta)] \\
&= \mathbb{E}[\{\mathbb{E}(\boldsymbol{T}) + Cov(\boldsymbol{T},\boldsymbol{Z})[\mathbb{V}(\boldsymbol{Z})]^{-1}(\boldsymbol{Z} - \mathbb{E}(\boldsymbol{Z}))\}\boldsymbol{Z}^\top\chi_{p_2}^{-2}(\Delta)] \\
&= \mathbb{E}[\{\mathbf{0} + (\sigma^2\mathbf{\Omega})(\sigma^2\mathbf{\Omega})^{-1}(\boldsymbol{Z} - \boldsymbol{\omega\delta})\}\boldsymbol{Z}^\top\chi_{p_2}^{-2}(\Delta)] \\
&= \mathbb{E}[(\boldsymbol{Z} - (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}))\boldsymbol{Z}^\top\chi_{p_2}^{-2}(\Delta)] \\
&= \underbrace{\mathbb{E}[\boldsymbol{Z}\boldsymbol{Z}^\top\chi_{p_2}^{-2}(\Delta)]}_{\mathrm{E}_8} - (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})\underbrace{\mathbb{E}[\boldsymbol{Z}^\top\chi_{p_2}^{-2}(\Delta)]}_{\mathrm{E}_9}.
\end{aligned}$$

By Equation (4.1) and Theorem 3.6.2, we may rewrite E_8 and E_9 as

$$
\begin{aligned}
\mathrm{E}_8 &= \mathbb{E}[\boldsymbol{Z}\boldsymbol{Z}^\top \chi_{p_2}^{-2}(\Delta)] \\
&= \mathbb{E}[(\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*)^\top \chi_{p_2}^{-2}(\Delta)] \\
&= (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\mathbb{E}[\boldsymbol{Z}^*(\boldsymbol{Z}^*)^\top \chi_{p_2}^{-2}(\Delta)]((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top \\
&= (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\left[\mathbb{V}(\boldsymbol{Z}^*)\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] + \mathbb{E}(\boldsymbol{Z}^*)\mathbb{E}(\boldsymbol{Z}^*)^\top\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)]\right]((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top \\
&= (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\left[\boldsymbol{I}_{p_2}\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] + (\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta}((\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta})^\top\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)]\right]((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top \\
&= (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] \\
&\quad + (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}(\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta}(\boldsymbol{\omega}\boldsymbol{\delta})^\top((\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}})^\top((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)] \\
&= \sigma^2\boldsymbol{\Omega}\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] + \boldsymbol{\omega}\boldsymbol{\delta}(\boldsymbol{\omega}\boldsymbol{\delta})^\top\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)] \\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)]
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{E}_9 &= \mathbb{E}[\boldsymbol{Z}^\top \chi_{p_2}^{-2}(\Delta)] \\
&= \mathbb{E}[((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*)^\top \chi_{p_2}^{-2}(\Delta)] \\
&= ((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top((\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta})^\top\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] \\
&= (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^\top\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)].
\end{aligned}
$$

Thus, E_6 can be rewritten as

$$
\begin{aligned}
\mathrm{E}_6 &= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)] \\
&\quad - (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^\top\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] \\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)] \\
&\quad - \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] \\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] \\
&\quad - \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}(\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - \mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)]).
\end{aligned}
$$

Then, by substituting E_6 and E_7 into $\boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}})$, we obtain

$$\begin{aligned}\boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) &= \sigma^2\boldsymbol{Q}_{11.2}^{-1} - 2c\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] \\ &\quad + 2c\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}(\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - \mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)]) \\ &\quad + c^2\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] \\ &\quad + c^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)] \\ &= \sigma^2\boldsymbol{Q}_{11.2}^{-1} - c\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]) \\ &\quad + c\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - 2\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)] + c\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]).\end{aligned}$$

Finally, we derive the AMSEM of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$, which is

$$\begin{aligned}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}) &= \lim_{n\to\infty}\mathbb{E}\left[\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}+} - \boldsymbol{\beta}_1)\sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}+} - \boldsymbol{\beta}_1)^\top\right] \\ &= \mathbb{E}\lim_{n\to\infty}\begin{bmatrix}\{\boldsymbol{T}_n - c\boldsymbol{Z}_n\Lambda_n^{-1} - \boldsymbol{Z}_nI(\Lambda_n \le c) + c\boldsymbol{Z}_n\Lambda_n^{-1}I(\Lambda_n \le c)\} \\ \{\boldsymbol{T}_n - c\boldsymbol{Z}_n\Lambda_n^{-1} - \boldsymbol{Z}_nI(\Lambda_n \le c) + c\boldsymbol{Z}_n\Lambda_n^{-1}I(\Lambda_n \le c)\}^\top\end{bmatrix} \\ &= \mathbb{E}\lim_{n\to\infty}\begin{bmatrix}\{(\boldsymbol{T}_n - c\boldsymbol{Z}_n\Lambda_n^{-1}) - (1 - c\Lambda_n^{-1})\boldsymbol{Z}_nI(\Lambda_n \le c)\} \\ \{(\boldsymbol{T}_n - c\boldsymbol{Z}_n\Lambda_n^{-1}) - (1 - c\Lambda_n^{-1})\boldsymbol{Z}_nI(\Lambda_n \le c)\}^\top\end{bmatrix} \\ &= \mathbb{E}\begin{bmatrix}\{(\boldsymbol{T} - c\boldsymbol{Z}\chi_{p_2}^{-2}(\Delta)) - (1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{Z}I(\chi_{p_2}^2(\Delta) \le c)\} \\ \{(\boldsymbol{T} - c\boldsymbol{Z}\chi_{p_2}^{-2}(\Delta)) - (1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{Z}I(\chi_{p_2}^2(\Delta) \le c)\}^\top\end{bmatrix} \\ &= \mathbb{E}\begin{bmatrix}(\boldsymbol{T} - c\boldsymbol{Z}\chi_{p_2}^{-2}(\Delta))(\boldsymbol{T} - c\boldsymbol{Z}\chi_{p_2}^{-2}(\Delta))^\top \\ -(\boldsymbol{T} - c\boldsymbol{Z}\chi_{p_2}^{-2}(\Delta))(1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c) \\ -(1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{Z}I(\chi_{p_2}^2(\Delta) \le c)(\boldsymbol{T} - c\boldsymbol{Z}\chi_{p_2}^{-2}(\Delta))^\top \\ +(1 - c\chi_{p_2}^{-2}(\Delta))^2\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)\end{bmatrix} \\ &= \underbrace{\mathbb{E}[(\boldsymbol{T} - c\boldsymbol{Z}\chi_{p_2}^{-2}(\Delta))(\boldsymbol{T} - c\boldsymbol{Z})^\top\chi_{p_2}^{-2}(\Delta)]}_{\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}})} \\ &\quad - 2\mathbb{E}[(\boldsymbol{T} - c\boldsymbol{Z}\chi_{p_2}^{-2}(\Delta))(1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)] \\ &\quad + \mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))^2\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)] \\ &= \boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - 2\mathbb{E}\begin{bmatrix}(1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{T}\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c) \\ -c\chi_{p_2}^{-2}(\Delta)(1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)\end{bmatrix} \\ &\quad + \mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))^2\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)]\end{aligned}$$

$$\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}) = \boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - 2\underbrace{\mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{T}\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)]}_{\mathrm{E}_{10}}$$
$$+ 2\underbrace{\mathbb{E}[c\chi_{p_2}^{-2}(\Delta)(1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)]}_{\mathrm{E}_{11}}$$
$$+ \underbrace{\mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))^2\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)]}_{\mathrm{E}_{12}}.$$

Using conditional expectation, E_{10} becomes

$$\begin{aligned}
\mathrm{E}_{10} &= \mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{T}\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)] \\
&= \mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))\mathbb{E}(\boldsymbol{T}\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)|\boldsymbol{Z})] \\
&= \mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))\mathbb{E}(\boldsymbol{T}|\boldsymbol{Z})\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)] \\
&= \mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))\{\mathbb{E}(\boldsymbol{T}) + Cov(\boldsymbol{T},\boldsymbol{Z})[\mathbb{V}(\boldsymbol{Z})]^{-1}(\boldsymbol{Z} - \mathbb{E}(\boldsymbol{Z}))\}\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)] \\
&= \mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))\{\boldsymbol{0} + (\sigma^2\boldsymbol{\Omega})(\sigma^2\boldsymbol{\Omega})^{-1}(\boldsymbol{Z} - \boldsymbol{\omega}\boldsymbol{\delta})\}\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)] \\
&= \mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))(\boldsymbol{Z} - (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}))\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)] \\
&= \underbrace{\mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)]}_{\mathrm{E}_{13}} \\
&\quad - (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})\underbrace{\mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)]}_{\mathrm{E}_{14}}.
\end{aligned}$$

By Theorem 3.6.2, E_{13} and E_{14} given as

$$\begin{aligned}
\mathrm{E}_{13} &= \mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi_{p_2}^2(\Delta) \le c)] \\
&= \mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))(\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z})^\top I(\chi_{p_2}^2(\Delta) \le c)] \\
&= (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\mathbb{E}[(1 - c\chi_{p_2}^{-2}(\Delta))\boldsymbol{Z}^*(\boldsymbol{Z}^*)^\top I(\chi_{p_2}^2 \le c)]((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top \\
&= (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\begin{bmatrix} \mathbb{V}(\boldsymbol{Z}^*)\mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^2(\Delta) \le c)] \\ +\mathbb{E}(\boldsymbol{Z}^*)\mathbb{E}(\boldsymbol{Z}^*)^\top\mathbb{E}[(1 - c\chi_{p_2+4}^{-2}(\Delta))I(\chi_{p_2+4}^2(\Delta) \le c)] \end{bmatrix}((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top \\
&= (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\begin{bmatrix} \boldsymbol{I}_{p_2}\mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^2(\Delta) \le c)] \\ +\begin{Bmatrix} (\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta}((\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta})^\top \\ \times\mathbb{E}[(1 - c\chi_{p_2+4}^{-2}(\Delta))I(\chi_{p_2+4}^2(\Delta) \le c)] \end{Bmatrix} \end{bmatrix}((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}_{13} &= (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta)\leq c)] \\
&\quad + \begin{bmatrix} (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}(\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega\delta}(\boldsymbol{\omega\delta})^\top((\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}})^\top((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top \\ \times\mathbb{E}[(1-c\chi^{-2}_{p_2+4}(\Delta))I(\chi^2_{p_2+4}(\Delta)\leq c)] \end{bmatrix} \\
&= \sigma^2\boldsymbol{\Omega}\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta)\leq c)] \\
&\quad + \boldsymbol{\omega\delta}(\boldsymbol{\omega\delta})^\top\mathbb{E}[(1-c\chi^{-2}_{p_2+4}(\Delta))I(\chi^2_{p_2+4}(\Delta)\leq c)] \\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta)\leq c)] \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi^{-2}_{p_2+4}(\Delta))I(\chi^2_{p_2+4}(\Delta)\leq c)]
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{E}_{14} &= \mathbb{E}[(1-c\chi^{-2}_{p_2}(\Delta))\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta)\leq c)] \\
&= \mathbb{E}[(1-c\chi^{-2}_{p_2}(\Delta))((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*)^\top I(\chi^2_{p_2}(\Delta)\leq c)] \\
&= ((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top((\sigma^2\boldsymbol{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega\delta})^\top\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta)\leq c)] \\
&= (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^\top\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta)\leq c)].
\end{aligned}$$

Replacing E_{13} and E_{14} in E_{10}, then

$$\begin{aligned}
\mathrm{E}_{10} &= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta)\leq c)] \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi^{-2}_{p_2+4}(\Delta))I(\chi^2_{p_2+4}(\Delta)\leq c)] \\
&\quad - (-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})(-\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta})^\top\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta)\leq c)] \\
&= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta)\leq c)] \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi^{-2}_{p_2+4}(\Delta))I(\chi^2_{p_2+4}(\Delta)\leq c)] \\
&\quad - \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta)\leq c)].
\end{aligned}$$

Using Equation (3.40), we can write E_{11} and E_{12} as

$$\begin{aligned}
\mathrm{E}_{11} &= \mathbb{E}[c\chi^{-2}_{p_2}(\Delta)(1-c\chi^{-2}_{p_2}(\Delta))\boldsymbol{Z}\boldsymbol{Z}^\top I(\chi^2_{p_2}(\Delta)\leq c)] \\
&= \mathbb{E}[c\chi^{-2}_{p_2}(\Delta)(1-c\chi^{-2}_{p_2}(\Delta))(\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*)^\top I(\chi^2_{p_2}(\Delta)\leq c)] \\
&= (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\mathbb{E}[c\chi^{-2}_{p_2}(\Delta)(1-c\chi^{-2}_{p_2}(\Delta))\boldsymbol{Z}^*(\boldsymbol{Z}^*)^\top I(\chi^2_{p_2}\leq c)]((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top \\
&= (\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}}\begin{bmatrix} \mathbb{V}(\boldsymbol{Z}^*)\mathbb{E}[c\chi^{-2}_{p_2+2}(\Delta)(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta)\leq c)] \\ + \left\{ \begin{matrix} \mathbb{E}(\boldsymbol{Z}^*)\mathbb{E}(\boldsymbol{Z}^*)^\top \\ \times\mathbb{E}[c\chi^{-2}_{p_2+4}(\Delta)(1-c\chi^{-2}_{p_2+4}(\Delta))I(\chi^2_{p_2+4}(\Delta)\leq c)] \end{matrix} \right\} \end{bmatrix}((\sigma^2\boldsymbol{\Omega})^{\frac{1}{2}})^\top
\end{aligned}$$

$$\mathrm{E}_{11} = (\sigma^2\mathbf{\Omega})^{\frac{1}{2}} \begin{bmatrix} \boldsymbol{I}_{p_2}\mathbb{E}[c\chi^{-2}_{p_2+2}(\Delta)(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta)\leq c)] \\ +\begin{Bmatrix} (\sigma^2\mathbf{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega\delta}((\sigma^2\mathbf{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega\delta})^\top \\ \times\mathbb{E}[c\chi^{-2}_{p_2+4}(\Delta)(1-c\chi^{-2}_{p_2+4}(\Delta))I(\chi^2_{p_2+4}(\Delta)\leq c)] \end{Bmatrix} \end{bmatrix} ((\sigma^2\mathbf{\Omega})^{\frac{1}{2}})^\top$$

$$= (\sigma^2\mathbf{\Omega})^{\frac{1}{2}}((\sigma^2\mathbf{\Omega})^{\frac{1}{2}})^\top\mathbb{E}[c\chi^{-2}_{p_2+2}(\Delta)(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta)\leq c)]$$

$$+\begin{bmatrix} (\sigma^2\mathbf{\Omega})^{\frac{1}{2}}(\sigma^2\mathbf{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega\delta}(\boldsymbol{\omega\delta})^\top((\sigma^2\mathbf{\Omega})^{-\frac{1}{2}})^\top((\sigma^2\mathbf{\Omega})^{\frac{1}{2}})^\top \\ \times\mathbb{E}[c\chi^{-2}_{p_2+4}(\Delta)(1-c\chi^{-2}_{p_2+4}(\Delta))I(\chi^2_{p_2+4}(\Delta)\leq c)] \end{bmatrix}$$

$$= \sigma^2\mathbf{\Omega}\mathbb{E}[c\chi^{-2}_{p_2+2}(\Delta)(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta)\leq c)]$$

$$+ (\boldsymbol{\omega\delta})(\boldsymbol{\omega\delta})^\top\mathbb{E}[c\chi^{-2}_{p_2+4}(\Delta)(1-c\chi^{-2}_{p_2+4}(\Delta))I(\chi^2_{p_2+4}(\Delta)\leq c)]$$

$$= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[c\chi^{-2}_{p_2+2}(\Delta)(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta)\leq c)]$$

$$+ \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[c\chi^{-2}_{p_2+4}(\Delta)(1-c\chi^{-2}_{p_2+4}(\Delta))I(\chi^2_{p_2+4}(\Delta)\leq c)]$$

and

$$\mathrm{E}_{12} = \mathbb{E}[(1-c\chi^{-2}_{p_2}(\Delta))^2\boldsymbol{ZZ}^\top I(\chi^2_{p_2}\leq c)]$$

$$= \mathbb{E}[(1-c\chi^{-2}_{p_2}(\Delta))^2(\sigma^2\mathbf{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*((\sigma^2\mathbf{\Omega})^{\frac{1}{2}}\boldsymbol{Z}^*)^\top I(\chi^2_{p_2}\leq c)]$$

$$= (\sigma^2\mathbf{\Omega})^{\frac{1}{2}}\mathbb{E}[(1-c\chi^{-2}_{p_2}(\Delta))^2\boldsymbol{Z}^*(\boldsymbol{Z}^*)^\top I(\chi^2_{p_2}\leq c)]((\sigma^2\mathbf{\Omega})^{\frac{1}{2}})^\top$$

$$= (\sigma^2\mathbf{\Omega})^{\frac{1}{2}} \begin{bmatrix} \mathbb{V}(\boldsymbol{Z}^*)\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))^2 I(\chi^2_{p_2+2}(\Delta)\leq c)] \\ +\mathbb{E}(\boldsymbol{Z}^*)\mathbb{E}(\boldsymbol{Z}^*)^\top\mathbb{E}[(1-c\chi^{-2}_{p_2+4}(\Delta))^2 I(\chi^2_{p_2+4}(\Delta)\leq c)] \end{bmatrix} ((\sigma^2\mathbf{\Omega})^{\frac{1}{2}})^\top$$

$$= (\sigma^2\mathbf{\Omega})^{\frac{1}{2}} \begin{bmatrix} \boldsymbol{I}_{p_2}\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))^2 I(\chi^2_{p_2+2}(\Delta)\leq c)] \\ +\begin{Bmatrix} (\sigma^2\mathbf{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega\delta}((\sigma^2\mathbf{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega\delta})^\top \\ \times\mathbb{E}[(1-c\chi^{-2}_{p_2+4}(\Delta))^2 I(\chi^2_{p_2+4}(\Delta)\leq c)] \end{Bmatrix} \end{bmatrix} ((\sigma^2\mathbf{\Omega})^{\frac{1}{2}})^\top$$

$$= (\sigma^2\mathbf{\Omega})^{\frac{1}{2}}((\sigma^2\mathbf{\Omega})^{\frac{1}{2}})^\top\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))^2 I(\chi^2_{p_2+2}(\Delta)\leq c)]$$

$$+\begin{bmatrix} (\sigma^2\mathbf{\Omega})^{\frac{1}{2}}(\sigma^2\mathbf{\Omega})^{-\frac{1}{2}}\boldsymbol{\omega\delta}(\boldsymbol{\omega\delta})^\top((\sigma^2\mathbf{\Omega})^{-\frac{1}{2}})^\top((\sigma^2\mathbf{\Omega})^{\frac{1}{2}})^\top \\ \times\mathbb{E}[(1-c\chi^{-2}_{p_2+4}(\Delta))^2 I(\chi^2_{p_2+4}(\Delta)\leq c)] \end{bmatrix}$$

$$= \sigma^2\mathbf{\Omega}\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))^2 I(\chi^2_{p_2+2}(\Delta)\leq c)]$$

$$+ \boldsymbol{\omega\delta}(\boldsymbol{\omega\delta})^\top\mathbb{E}[(1-c\chi^{-2}_{p_2+4}(\Delta))^2 I(\chi^2_{p_2+4}(\Delta)\leq c)]$$

$$= \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))^2 I(\chi^2_{p_2+2}(\Delta)\leq c)]$$

$$+ \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi^{-2}_{p_2+4}(\Delta))^2 I(\chi^2_{p_2+4}(\Delta)\leq c)].$$

Finally, substituting E_{10}, E_{11}, and E_{12} into $\mathbf{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{S^+})$ and rearranging the terms, we have

$$\begin{aligned}
\mathbf{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{S^+}) &= \mathbf{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{S}) - 2\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta)\le c)] \\
&\quad - 2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))I(\chi_{p_2+4}^{2}(\Delta)\le c)] \\
&\quad + 2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta)\le c)] \\
&\quad + 2\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[c\chi_{p_2+2}^{-2}(\Delta)(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta)\le c)] \\
&\quad + 2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[c\chi_{p_2+4}^{-2}(\Delta)(1-c\chi_{p_2+4}^{-2}(\Delta))I(\chi_{p_2+4}^{2}(\Delta)\le c)] \\
&\quad + \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2I(\chi_{p_2+2}^{2}(\Delta)\le c)] \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2I(\chi_{p_2+4}^{2}(\Delta)\le c)] \\
&= \mathbf{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{S}) \\
&\quad - 2\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} \\
&\qquad \times\mathbb{E}\begin{bmatrix}(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta)\le c)\\ -c\chi_{p_2+2}^{-2}(\Delta)(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta)\le c)\end{bmatrix} \\
&\quad - 2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} \\
&\qquad \times\mathbb{E}\begin{bmatrix}(1-c\chi_{p_2+4}^{-2}(\Delta))I(\chi_{p_2+4}^{2}(\Delta)\le c)\\ -c\chi_{p_2+4}^{-2}(\Delta)(1-c\chi_{p_2+4}^{-2}(\Delta))I(\chi_{p_2+4}^{2}(\Delta)\le c)\end{bmatrix} \\
&\quad + 2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta)\le c)] \\
&\quad + \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2I(\chi_{p_2+2}^{2}(\Delta)\le c)] \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2I(\chi_{p_2+4}^{2}(\Delta)\le c)] \\
&= \mathbf{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{S}) \\
&\quad - 2\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}\begin{bmatrix}(1-c\chi_{p_2+2}^{-2}(\Delta))(1-c\chi_{p_2+2}^{-2}(\Delta))\\ I(\chi_{p_2+2}^{2}(\Delta)\le c)\end{bmatrix} \\
&\quad - 2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}\begin{bmatrix}(1-c\chi_{p_2+4}^{-2}(\Delta))(1-c\chi_{p_2+4}^{-2}(\Delta))\\ I(\chi_{p_2+4}^{2}(\Delta)\le c)\end{bmatrix} \\
&\quad + 2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta)\le c)] \\
&\quad + \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2I(\chi_{p_2+2}^{2}(\Delta)\le c)] \\
&\quad + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2I(\chi_{p_2+4}^{2}(\Delta)\le c)]
\end{aligned}$$

Hence,

$$\begin{aligned}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}) &= \boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - \sigma^2 \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^2(\Delta)\le c)]\\ &\quad - \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^2(\Delta)\le c)]\\ &\quad + 2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^2(\Delta)\le c)].\end{aligned}$$

□

Using Equation (3.37) and the above AMSEMs, the result of the asymptotic distributional quadratic risk (ADQR) of the proposed estimators depend on the following theorem.

Theorem 4.1.5. *Under the assumed regularity condition and local alternative* $\{K_n\}$*, as* $n \to \infty$*, the ADQRs of the estimators are as follows:*

$$\begin{aligned}ADQR(\widehat{\boldsymbol{\beta}}_1^{UE}) &= \sigma^2 tr[\boldsymbol{W}\boldsymbol{Q}_{11.2}^{-1}],\\ ADQR(\widehat{\boldsymbol{\beta}}_1^{RE}) &= ADQR(\widehat{\boldsymbol{\beta}}_1^{UE}) - \sigma^2 tr[\boldsymbol{Q}^\circ\boldsymbol{Q}_{22.1}^{-1}] + \boldsymbol{\delta}^\top\boldsymbol{Q}^\circ\boldsymbol{\delta},\\ ADQR(\widehat{\boldsymbol{\beta}}_1^{LS}) &= ADQR(\widehat{\boldsymbol{\beta}}_1^{UE}) - \pi(2-\pi)\sigma^2 tr[\boldsymbol{Q}^\circ\boldsymbol{Q}_{22.1}^{-1}] + \pi^2\boldsymbol{\delta}^\top\boldsymbol{Q}^\circ\boldsymbol{\delta},\\ ADQR(\widehat{\boldsymbol{\beta}}_1^{PT}) &= ADQR(\widehat{\boldsymbol{\beta}}_1^{UE}) - \sigma^2 tr[\boldsymbol{Q}^\circ\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}(\chi_{p_2,\alpha}^2;\Delta)\\ &\quad + \boldsymbol{\delta}^\top\boldsymbol{Q}^\circ\boldsymbol{\delta}[2H_{p_2+2}(\chi_{p_2,\alpha}^2;\Delta) - H_{p_2+4}(\chi_{p_2,\alpha}^2;\Delta)],\\ ADQR(\widehat{\boldsymbol{\beta}}_1^{SP}) &= ADQR(\widehat{\boldsymbol{\beta}}_1^{UE}) - \pi(2-\pi)\sigma^2 tr[\boldsymbol{Q}^\circ\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}(\chi_{p_2,\alpha}^2;\Delta)\\ &\quad + \boldsymbol{\delta}^\top\boldsymbol{Q}^\circ\boldsymbol{\delta}[2\pi H_{p_2+2}(\chi_{p_2,\alpha}^2;\Delta) - \pi(2-\pi)H_{p_2+4}(\chi_{p_2,\alpha}^2;\Delta)],\\ ADQR(\widehat{\boldsymbol{\beta}}_1^{S}) &= ADQR(\widehat{\boldsymbol{\beta}}_1^{UE}) - c\sigma^2 tr[\boldsymbol{Q}^\circ\boldsymbol{Q}_{22.1}^{-1}](2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)])\\ &\quad + c\boldsymbol{\delta}^\top\boldsymbol{Q}^\circ\boldsymbol{\delta}(2E[\chi_{p_2+2}^{-2}(\Delta)] - 2\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)] + c\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]),\\ ADQR(\widehat{\boldsymbol{\beta}}_1^{S^+}) &= ADQR(\widehat{\boldsymbol{\beta}}_1^{S}) - \sigma^2 tr[\boldsymbol{Q}^\circ\boldsymbol{Q}_{22.1}^{-1}]\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^2(\Delta)\le c)]\\ &\quad - \boldsymbol{\delta}^\top\boldsymbol{Q}^\circ\boldsymbol{\delta}\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^2(\Delta)\le c)]\\ &\quad + 2\boldsymbol{\delta}^\top\boldsymbol{Q}^\circ\boldsymbol{\delta}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^2(\Delta)\le c)],\end{aligned}$$

where $c = p_2 - 2$*,* $p_2 > 2$*, and* $\boldsymbol{Q}^\circ = \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}$.

Proof. The ADQRs of all estimators can be derived as follows:

$$\begin{aligned}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) &= \mathrm{tr}[\boldsymbol{W}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}})]\\ &= \mathrm{tr}[\boldsymbol{W}\sigma^2\boldsymbol{Q}_{11.2}^{-1}]\\ &= \sigma^2\mathrm{tr}[\boldsymbol{W}\boldsymbol{Q}_{11.2}^{-1}].\end{aligned}$$

$$
\begin{aligned}
\mathrm{ADQR}(\widehat{\beta}_1^{\mathrm{RE}}) &= \mathrm{tr}[\boldsymbol{W}\boldsymbol{\Gamma}^*(\widehat{\beta}_1^{\mathrm{RE}})] \\
&= \mathrm{tr}[\boldsymbol{W}(\sigma^2\boldsymbol{Q}_{11}^{-1} + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1})] \\
&= \mathrm{tr}\left[\boldsymbol{W}\begin{pmatrix} \underbrace{\sigma^2\boldsymbol{Q}_{11}^{-1} + \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}}_{\sigma^2\boldsymbol{Q}_{11.2}^{-1}} \\ -\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} \\ +\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} \end{pmatrix}\right] \\
&= \sigma^2\mathrm{tr}[\boldsymbol{W}\boldsymbol{Q}_{11.2}^{-1}] - \sigma^2\mathrm{tr}[\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}] \\
&\quad + \mathrm{tr}[\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}] \\
&= \mathrm{ADQR}(\widehat{\beta}_1^{\mathrm{UE}}) - \sigma^2\mathrm{tr}[\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}] \\
&\quad + \mathrm{tr}[\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}] \\
&= \mathrm{ADQR}(\widehat{\beta}_1^{\mathrm{UE}}) - \sigma^2\mathrm{tr}[\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}] \\
&\quad + \boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta} \\
&= \mathrm{ADQR}(\widehat{\beta}_1^{\mathrm{UE}}) - \sigma^2\mathrm{tr}[\boldsymbol{Q}^\circ\boldsymbol{Q}_{22.1}^{-1}] + \boldsymbol{\delta}^\top\boldsymbol{Q}^\circ\boldsymbol{\delta}.
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{ADQR}(\widehat{\beta}_1^{\mathrm{LS}}) &= \mathrm{tr}[\boldsymbol{W}\boldsymbol{\Gamma}^*(\widehat{\beta}_1^{\mathrm{LS}})] \\
&= \mathrm{tr}\left[\boldsymbol{W}\begin{pmatrix} \sigma^2\boldsymbol{Q}_{11.2}^{-1} - \pi(2-\pi)\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} \\ +\pi^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} \end{pmatrix}\right] \\
&= \sigma^2\mathrm{tr}[\boldsymbol{W}\boldsymbol{Q}_{11.2}^{-1}] - \pi(2-\pi)\sigma^2\mathrm{tr}[\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}] \\
&\quad + \pi^2\mathrm{tr}[\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}] \\
&= \mathrm{ADQR}(\widehat{\beta}_1^{\mathrm{UE}}) - \pi(2-\pi)\sigma^2\mathrm{tr}[\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}] \\
&\quad + \pi^2\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta} \\
&= \mathrm{ADQR}(\widehat{\beta}_1^{\mathrm{UE}}) - \pi(2-\pi)\sigma^2\mathrm{tr}[\boldsymbol{Q}^\circ\boldsymbol{Q}_{22.1}^{-1}] + \pi^2\boldsymbol{\delta}^\top\boldsymbol{Q}^\circ\boldsymbol{\delta}.
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{ADQR}(\widehat{\beta}_1^{\mathrm{PT}}) &= \mathrm{tr}[\boldsymbol{W}\boldsymbol{\Gamma}^*(\widehat{\beta}_1^{\mathrm{PT}})] \\
&= \mathrm{tr}\left[\boldsymbol{W}\begin{pmatrix} \sigma^2\boldsymbol{Q}_{11.2}^{-1} - \sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\ +\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\begin{Bmatrix} 2H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\ -H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta) \end{Bmatrix} \end{pmatrix}\right] \\
&= \sigma^2\mathrm{tr}[\boldsymbol{W}\boldsymbol{Q}_{11.2}^{-1}] - \sigma^2\mathrm{tr}[\,\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \mathrm{tr}[\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}]\,[2H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)]
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}) &= \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) - \sigma^2 \mathrm{tr}[\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}] H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \boldsymbol{\delta}^\top \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}[2H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)] \\
&= \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) - \sigma^2 \mathrm{tr}[\boldsymbol{Q}^\circ \boldsymbol{Q}_{22.1}^{-1}] H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \boldsymbol{\delta}^\top \boldsymbol{Q}^\circ \boldsymbol{\delta}[2H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)].
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) &= \mathrm{tr}[\boldsymbol{W}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}})] \\
&= \mathrm{tr}\left[\boldsymbol{W}\begin{pmatrix} \sigma^2\boldsymbol{Q}_{11.2}^{-1} - \pi(2-\pi)\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\ +\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\left\{\begin{matrix} 2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\ -\pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)\end{matrix}\right\}\end{pmatrix}\right] \\
&= \sigma^2\mathrm{tr}[\boldsymbol{W}\boldsymbol{Q}_{11.1}^{-1}] - \pi(2-\pi)\sigma^2\mathrm{tr}[\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}]H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \mathrm{tr}[\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}]\left\{\begin{matrix} 2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\ -\pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)\end{matrix}\right\} \\
&= \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) - \pi(2-\pi)\sigma^2\mathrm{tr}[\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\left\{\begin{matrix} 2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\ -\pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)\end{matrix}\right\} \\
&= \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) - \pi(2-\pi)\sigma^2\mathrm{tr}[\boldsymbol{Q}^\circ\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \boldsymbol{\delta}^\top\boldsymbol{Q}^\circ\boldsymbol{\delta}[2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - \pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)].
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) &= \mathrm{tr}[\boldsymbol{W}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}})] \\
&= \mathrm{tr}\left[\boldsymbol{W}\begin{pmatrix} \sigma^2\boldsymbol{Q}_{11.2}^{-1} - c\sigma^2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\left\{\begin{matrix} 2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] \\ -c\mathbb{E}[\chi^{-4}_{p_2+2}(\Delta)]\end{matrix}\right\} \\ +c\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\left\{\begin{matrix} 2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] - 2[\chi^{-2}_{p_2+4,\Delta}] \\ +c\mathbb{E}[\chi^{-4}_{p_2+4}(\Delta)]\end{matrix}\right\}\end{pmatrix}\right] \\
&= \sigma^2\mathrm{tr}[\boldsymbol{W}\boldsymbol{Q}_{11.2}^{-1}] - c\sigma^2\mathrm{tr}[\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}]\left\{\begin{matrix} 2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] \\ -c\mathbb{E}[\chi^{-4}_{p_2+2}(\Delta)]\end{matrix}\right\} \\
&\quad + c\,\mathrm{tr}[\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}]\left\{\begin{matrix} 2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] - 2[\chi^{-2}_{p_2+4}(\Delta)] \\ +c\mathbb{E}[\chi^{-4}_{p_2+4}(\Delta)]\end{matrix}\right\}
\end{aligned}
$$

$$\begin{aligned}
\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) &= \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) - c\sigma^2 \mathrm{tr}[\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}] \begin{Bmatrix} 2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] \\ -c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] \end{Bmatrix} \\
&\quad + c\boldsymbol{\delta}^\top \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta} \begin{Bmatrix} 2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - 2[\chi_{p_2+4}^{-2}(\Delta)] \\ +c\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)] \end{Bmatrix} \\
&= \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) - c\sigma^2 \mathrm{tr}[\boldsymbol{Q}^\circ \boldsymbol{Q}_{22.1}^{-1}](2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]) \\
&\quad + c\boldsymbol{\delta}^\top \boldsymbol{Q}^\circ \boldsymbol{\delta}[2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - 2[\chi_{p_2+4}^{-2}(\Delta)] + c\mathbb{E}[\chi_{p_2+4}^{-4}](\Delta)].
\end{aligned}$$

$$\begin{aligned}
\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}) &= \mathrm{tr}[\boldsymbol{W}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}+})] \\
&= \mathrm{tr}\left[\boldsymbol{W}\left(\begin{matrix} \boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) \\ -\sigma^2 \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}\begin{bmatrix} (1-c\chi_{p_2+2}^{-2}(\Delta))^2 \\ I(\chi_{p_2+2}^{2}(\Delta)\le c) \end{bmatrix} \\ +\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} \begin{Bmatrix} 2\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta)\le c)] \\ -\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^{2}(\Delta)\le c)] \end{Bmatrix} \end{matrix}\right)\right] \\
&= \mathrm{tr}[\boldsymbol{W}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}})] \\
&\quad - \sigma^2 \mathrm{tr}[\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}]\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^{2}(\Delta)\le c)] \\
&\quad + \mathrm{tr}[\boldsymbol{\delta}^\top \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}] \begin{Bmatrix} 2\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta)\le c)]) \\ -\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^{2}(\Delta)\le c)] \end{Bmatrix} \\
&= \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) \\
&\quad - \sigma^2 \mathrm{tr}[\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}]\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^{2}(\Delta)\le c)] \\
&\quad + \boldsymbol{\delta}^\top \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta} \begin{Bmatrix} 2\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta)\le c)]) \\ -\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^{2}(\Delta)\le c)] \end{Bmatrix} \\
&= \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - \sigma^2 \mathrm{tr}[\boldsymbol{Q}^\circ \boldsymbol{Q}_{22.1}^{-1}]\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^{2}(\Delta)\le c)] \\
&\quad - \boldsymbol{\delta}^\top \boldsymbol{Q}^\circ \boldsymbol{\delta}\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^{2}(\Delta)\le c)] \\
&\quad + 2\boldsymbol{\delta}^\top \boldsymbol{Q}^\circ \boldsymbol{\delta}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta)\le c)].
\end{aligned}$$

□

In the ADQR analysis, we exclude the case $\boldsymbol{Q}_{12} = \boldsymbol{0}$, since in this situation $\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12} = \boldsymbol{0}$ and $\boldsymbol{Q}_{11.2}^{-1} = \boldsymbol{Q}_{11}^{-1} + \boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{Q}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1} = \boldsymbol{Q}_{11}^{-1}$. The ADQRs of all estimators are then reduced to a common value $\sigma^2 tr[\boldsymbol{W}\boldsymbol{Q}_{11}^{-1}]$, which is the ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$. Hence, all estimators become ADQR equivalent. In the remaining discussion, we assume that $\boldsymbol{Q}_{12}$ is not null ($\boldsymbol{Q}_{12} \neq \boldsymbol{0}$).

The ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ is uncorrelated with the UPI and so does not depend on $\boldsymbol{\delta}$ while the other estimators are functions of $\boldsymbol{\delta}$. In Theorem 4.1.5, the ADQR expressions are explained as a loss function of the Mahalanobis distance. We consider the special case of $\boldsymbol{W} = \sigma^{-2}\boldsymbol{Q}_{11.2}$ and $\sigma^2\mathrm{tr}[\boldsymbol{W}\boldsymbol{Q}_{11.2}^{-1}] = p_1$. We note that ADQR$(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}})$ reduces to p_1, which is constant and independent of $\boldsymbol{\delta}$. The simplified ADQR expressions are provided in the following corollary.

Corollary 2. *For $\boldsymbol{W} = \sigma^{-2}\boldsymbol{Q}_{11.2}$, the ADQR expressions simplify to*

$$
\begin{aligned}
ADQR(\widehat{\boldsymbol{\beta}}_1^{UE}) &= p_1, \\
ADQR(\widehat{\boldsymbol{\beta}}_1^{RE}) &= p_1 - tr[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] + \Delta^*, \\
ADQR(\widehat{\boldsymbol{\beta}}_1^{LS}) &= p_1 - \pi(2-\pi)tr[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] + \pi^2\Delta^*, \\
ADQR(\widehat{\boldsymbol{\beta}}_1^{PT}) &= p_1 - tr[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \Delta^*[2H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)], \\
ADQR(\widehat{\boldsymbol{\beta}}_1^{SP}) &= p_1 - \pi(2-\pi)tr[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \Delta^*[2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - \pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)], \\
ADQR(\widehat{\boldsymbol{\beta}}_1^{S}) &= p_1 - ctr[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}](2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] - c\mathbb{E}[\chi^{-4}_{p_2+2}(\Delta)]) \\
&\quad + c\Delta^*(2E[\chi^{-2}_{p_2+2}(\Delta)] - 2\mathbb{E}[\chi^{-2}_{p_2+4}(\Delta)] + c\mathbb{E}[\chi^{-4}_{p_2+4}(\Delta)]), \\
ADQR(\widehat{\boldsymbol{\beta}}_1^{S+}) &= ADQR(\widehat{\boldsymbol{\beta}}_1^{S}) - tr[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\mathbb{E}[(1 - c\chi^{-2}_{p_2+2}(\Delta))^2 I(\chi^2_{p_2+2}(\Delta) \le c)] \\
&\quad - \Delta^*\mathbb{E}[(1 - c\chi^{-2}_{p_2+4}(\Delta))^2 I(\chi^2_{p_2+4}(\Delta) \le c)] \\
&\quad + 2\Delta^*\mathbb{E}[(1 - c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta) \le c)],
\end{aligned}
$$

where $c = p_2 - 2$, $p_2 > 2$, $\Delta^* = \sigma^{-2}\boldsymbol{\delta}^\top\boldsymbol{Q}^*\boldsymbol{\delta}$ and $\boldsymbol{Q}^* = \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{11.2}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}$.

The RE, LS, PT, SP, S, and S^+ estimators are compared with the UE using ADQR. For comparison, the following definition is very helpful.

Definition 4.1.1. *Let $\boldsymbol{\mathcal{B}}_1$ be the parameter space of $\boldsymbol{\beta}_1$. If two estimators $\widehat{\boldsymbol{\beta}}_1^0$ and $\widehat{\boldsymbol{\beta}}_1^*$ are such that $ADQR(\widehat{\boldsymbol{\beta}}_1^0) \le ADQR(\widehat{\boldsymbol{\beta}}_1^*)$ for all values of $\boldsymbol{\beta}_1 \in \boldsymbol{\mathcal{B}}_1$, with strict inequality for at least one $\boldsymbol{\beta}_1$, we say that $\widehat{\boldsymbol{\beta}}_1^0$ dominates $\widehat{\boldsymbol{\beta}}_1^*$.*

Theorem 4.1.6. *(Courant-Fischer (Gruber, 1998, p. 205)) If $\boldsymbol{B}$ and $\boldsymbol{D}$ are two positive semi-definite matrices, both of order $(m \times m)$, and $\boldsymbol{D}$ is nonsingular, then*

$$ch_{\min}(\boldsymbol{B}\boldsymbol{D}^{-1}) \leq \frac{\boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{D}\boldsymbol{x}} \leq ch_{\max}(\boldsymbol{B}\boldsymbol{D}^{-1}), \tag{4.2}$$

where $ch_{\min}(\cdot)$ and $ch_{\max}(\cdot)$ are the smallest and largest eigenvalues of $(\cdot)$, and $\boldsymbol{x}$ is a column vector of order $(m \times 1)$. We note that the above lower and upper bounds are equal to the infimum and supremum, respectively, of the ratio $\frac{\boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{D}\boldsymbol{x}}$ for $\boldsymbol{x} \neq \boldsymbol{0}$. For $\boldsymbol{D} = \boldsymbol{I}$, the ratio is known as the Rayleigh quotient for matrix $\boldsymbol{B}$.

Further, we consider Theorem 4.1.6 and set $\boldsymbol{W} = \sigma^{-2}\boldsymbol{Q}_{11.2}$, then $\boldsymbol{Q}^\circ = \sigma^{-2}\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{11.2}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12} = \sigma^{-2}\boldsymbol{Q}^*$, so we have

$$\begin{aligned}
ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}) &\leq \frac{\sigma^{-2}\boldsymbol{\delta}^\top\boldsymbol{Q}^*\boldsymbol{\delta}}{\sigma^{-2}\boldsymbol{\delta}^\top\boldsymbol{Q}_{22.1}\boldsymbol{\delta}} \leq ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}) \\
ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}) &\leq \frac{\Delta^*}{\Delta} \leq ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}) \\
\Delta ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}) &\leq \Delta^* \leq \Delta ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}),
\end{aligned} \tag{4.3}$$

where $ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})$ and $ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})$ are the smallest and largest eigenvalues of $(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})$.

The comparison of the ADQR can be regarded as follows:

4.1.2.1 Comparing $\widehat{\beta}_1^{\text{UE}}$ and $\widehat{\beta}_1^{\text{RE}}$

Consider the difference between the ADQRs of $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ and $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, then we get, $\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) - \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}}) = \text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] - \Delta^*$. It can be written in terms of Δ^* as

$$\Delta^* = \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}}) - \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) + \text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}].$$

As a consequence of the Courant-Fischer Theorem using (4.3), we can write

$$\Delta ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}) \leq \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}}) - \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) + \text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] \leq \Delta ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}).$$

Then, we get

$$\overbrace{\underbrace{\begin{array}{c}\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) - \text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] \\ +\Delta ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\end{array}}} \leq \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}}) \leq \overbrace{\underbrace{\begin{array}{c}\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) - \text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] \\ +\Delta ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\end{array}}}. \tag{4.4}$$

If the null hypothesis is true (that is, UPI is correct or $\Delta = 0$), then the lower and upper bounds of $\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}})$ in the above equation are the same, and we have

$$\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}}) = \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) - \text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]$$
$$\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) - \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}}) = \text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] \geq 0,$$

which means that $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$ strictly outperforms $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ when the restriction is correctly specified.

Next, consider the right-hand side of Equation (4.4) for $\Delta \leq \frac{\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]}{ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})}$,

$$\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}}) \leq \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) - \text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] + \left[\frac{\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]}{ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})}\right] ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})$$
$$\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}}) \leq \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}),$$

which shows that $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$ also dominates $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ in the interval $\Delta \in \left[0, \frac{\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]}{ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})}\right)$. We also consider the interval $\Delta \geq \frac{\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]}{ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})}$, the left-hand side of Equation (4.4) is

$$\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) - \text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] + \left[\frac{\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]}{ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})}\right] ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}) \leq \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}})$$
$$\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) \leq \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{RE}}).$$

Hence, $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ outperforms $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$ for $\Delta \in \left[\frac{\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]}{ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})}, \infty\right)$. Clearly, the ADQR of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$ becomes unbounded beyond $\frac{\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]}{ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})}$ when Δ moves away from the null hypothesis.

4.1.2.2 Comparing $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ and $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$

The difference between ADQR of $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ and $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$ is given as $\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) - \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{LS}}) = \pi(2-\pi)\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] - \pi^2\Delta^*$, which gives

$$\Delta^* = \frac{\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{LS}}) - \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) + \pi(2-\pi)\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]}{\pi^2}. \quad (4.5)$$

Replacing Equation (4.5) in Equation (4.3) and solving for ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$), we have

$$\overbrace{\underbrace{\begin{matrix}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) \\ -\pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] \\ +\pi^2\Delta ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\end{matrix}}} \leq \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) \leq \overbrace{\underbrace{\begin{matrix}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) \\ -\pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] \\ +\pi^2\Delta ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\end{matrix}}}. \qquad (4.6)$$

For $\pi \in (0,1)$ and $\Delta = 0$, we get

$$\begin{aligned}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) &= \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) - \pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] \\ \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) &= \pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] \geq 0.\end{aligned}$$

Thus, we conclude that $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ outperforms $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ under the null hypothesis. Moreover, the ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ decreases as π increases, and becomes equal to the ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ when $\pi = 1$. For $\Delta \in \left[0, \frac{\pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]}{\pi^2 ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})}\right)$, the right-hand side of Equation (4.6) reduces to

$$\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) \leq \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}),$$

which means that $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ performs better than $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$. On the other hand, the left-hand side of Equation (4.6) under $\Delta \in \left[\frac{\pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]}{\pi^2 ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})}, \infty\right)$ reduces to

$$\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) \leq \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}),$$

which indicates the superiority of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ in this region.

4.1.2.3 Comparing $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$

We express ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ in terms of ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$) as

$$\begin{aligned}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) = {} & \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) - \pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\ & + \Delta^*[2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - \pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)],\end{aligned}$$

and solving for Δ^*, we get

$$\Delta^* = \frac{\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) + \pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)}{[2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - \pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)]}. \qquad (4.7)$$

A solution of ADQR($\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$) after replacing Equation (4.7) in Equation (4.3) is given below

$$\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{SP}}) \le \overbrace{\underbrace{\begin{array}{c}\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) - \pi(2-\pi)\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}^* \\ +\Delta ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\begin{bmatrix}2\pi H_{p_2+2}^* \\ -\pi(2-\pi)H_{p_2+4}^*\end{bmatrix}\end{array}}}, \tag{4.8}$$

and

$$\overbrace{\underbrace{\begin{array}{c}\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) - \pi(2-\pi)\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}^* \\ +\Delta ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\begin{bmatrix}2\pi H_{p_2+2}^* \\ -\pi(2-\pi)H_{p_2+4}^*\end{bmatrix}\end{array}}} \le \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{SP}}) \tag{4.9}$$

where $H_\nu^* = H_\nu(\chi^2_{p_2,\alpha};\Delta)$. When the hypothesis is true, Equations (4.8) and (4.9) are equal and we have

$$\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) - \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{SP}}) = \pi(2-\pi)\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \ge 0,$$

which indicates that $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$ performs better than $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ at $\Delta^* = 0$. Now, from Equation (4.8) for the interval

$$\Delta \in \left[0, \frac{\pi(2-\pi)\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}^*}{ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})[2\pi H_{p_2+2}^* - \pi(2-\pi)H_{p_2+4}^*]}\right),$$

we have ADQR($\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$) − ADQR($\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$) ≤ 0. This means that $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$ dominates $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$. Similarly, from Equation (4.9) for

$$\Delta \in \left[\frac{\pi(2-\pi)\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}^*}{ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})[2\pi H_{p_2+2}^* - \pi(2-\pi)H_{p_2+4}^*]}, \infty\right)$$

becomes ADQR($\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$) − ADQR($\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$) ≥ 0, which is $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ outperforming $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$.

The preliminary test estimator ($\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$) is a special case of the shrinkage preliminary test estimator ($\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$) for the choice of $\pi = 1$. Hence, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$ performs better than $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ when

$$\Delta \in \left[0, \frac{\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}^*}{ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})[2H_{p_2+2}^* - H_{p_2+4}^*]}\right)$$

and the reverse is true when $\Delta \in \left[\frac{\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}^*}{ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})[2H_{p_2+2}^* - H_{p_2+4}^*]}, \infty\right)$.

4.1.2.4 Comparing $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$

In order to compare the ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$, normally, we write ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$) as

$$\begin{aligned}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) &= \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) - c\,\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}](2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]) \\ &\quad + c\Delta^*(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - 2\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)] + c\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]). \qquad (4.10)\end{aligned}$$

Using a trivial improvement of a result (2.2.13d) and (2.2.13e) cited in Saleh (2006, p. 32) as

$$\begin{aligned}\Delta\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)] &= \mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] \\ \mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] &= c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] + \Delta\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)] \qquad (4.11)\end{aligned}$$

and

$$\begin{aligned}2\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)] &= \mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - \mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)] \\ 4\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)] &= 2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - 2\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)]. \qquad (4.12)\end{aligned}$$

Replacing Equations (4.11) and (4.12) in Equation (4.10), we get

$$\begin{aligned}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) &= -c\,\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\begin{bmatrix}2\left\{c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] + \Delta\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]\right\} \\ -c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]\end{bmatrix} \\ &\quad + c\Delta^*(4\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)] + c\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]) \\ &= -c\,\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\begin{bmatrix}2c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] + 2\Delta\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)] \\ -c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]\end{bmatrix} \\ &\quad + c(4+c)\Delta^*\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)] \\ &= -c\,\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\left[c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] + 2\Delta\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]\right] \\ &\quad + c(4+p_2-2)\Delta^*\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)] \\ &= -c\,\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\begin{bmatrix}c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] + 2\Delta\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)] \\ -\frac{(p_2+2)\Delta^*}{\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]}\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]\end{bmatrix} \\ &= -c\,\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\begin{bmatrix}c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] \\ +2\Delta\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]\left\{1 - \frac{(p_2+2)}{2\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]}\frac{\Delta^*}{\Delta}\right\}\end{bmatrix}.\end{aligned}$$

Since the expectations $\mathbb{E}[\chi^{-4}_{p_2+2}(\Delta)]$ and $\mathbb{E}[\chi^{-4}_{p_2+4}(\Delta)]$ are positive,

$$\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{S}}) - \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}) \geq 0 \text{ if } 1 - \frac{(p_2+2)}{2\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}]}\frac{\Delta^*}{\Delta} \geq 0,$$

and $p_2 \geq 3$. By Equation (4.3), we write

$$1 - \frac{(p_2+2)}{2\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}]}\frac{\Delta^*}{\Delta} \leq 0$$
$$1 - \frac{(p_2+2)}{2\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}]} ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}) \geq 0,$$

which reduces to

$$\frac{\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}]}{ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1})} \geq \frac{(p_2+2)}{2}.$$

Thus, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$ performs better than $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for all Δ, $p_2 \geq 3$, and $\boldsymbol{Q} \in \boldsymbol{\mathcal{Q}}$ where

$$\boldsymbol{\mathcal{Q}} = \left\{ \boldsymbol{Q} : \frac{\text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}]}{ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1})} \geq \frac{(p_2+2)}{2} \right\}.$$

Next, consider the ADQR difference between $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$ and $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ as

$$\begin{aligned}\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{S}}) - \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{S+}}) = {} & \text{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}]\mathbb{E}\left[\left(1 - c\chi^{-2}_{p_2+2}(\Delta)\right)^2 I\left(\chi^2_{p_2+2}(\Delta) \leq c\right)\right] \\ & + \Delta^*\mathbb{E}\left[\left(1 - c\chi^{-2}_{p_2+4}(\Delta)\right)^2 I\left(\chi^2_{p_2+4}(\Delta) \leq c\right)\right] \\ & - 2\Delta^*\mathbb{E}\left[\left(1 - c\chi^{-2}_{p_2+2}(\Delta)\right) I\left(\chi^2_{p_2+2}(\Delta) \leq c\right)\right].\end{aligned}$$

Since,

$$\mathbb{E}\left[\left(1 - c\chi^{-2}_{p_2+2}(\Delta)\right)^2 I\left(\chi^2_{p_2+2}(\Delta) \leq c\right)\right] \geq 0,$$
$$\mathbb{E}\left[\left(1 - c\chi^{-2}_{p_2+4}(\Delta)\right)^2 I\left(\chi^2_{p_2+4}(\Delta) \leq c\right)\right] \geq 0,$$

and

$$\mathbb{E}\left[\left(1 - c\chi^{-2}_{p_2+2}(\Delta)\right) I\left(\chi^2_{p_2+2}(\Delta) \leq c\right)\right] \leq 0.$$

It indicates that $\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{S+}}) \leq \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{S}})$ for all Δ and $\boldsymbol{Q} \in \boldsymbol{\mathcal{Q}}$ with $p_2 \geq 3$. That means the $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ dominates $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$ and hence, $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ also outperforms $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$. Thus,

$$\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{S+}}) \leq \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{S}}) \leq \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}).$$

4.1.2.5 Comparing $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$

Both of ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$) and ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$) are unbound functions of Δ^* and the performance of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ is controlled by the values of $\pi \in (0,1)$. The difference between ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ is

$$\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) = -(1-\pi)^2 \mathrm{tr}[\boldsymbol{Q}^* \boldsymbol{Q}_{22.1}^{-1}] + (1-\pi^2)\Delta^*. \tag{4.13}$$

For $\Delta = 0$, ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$) − ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$) < 0. This means that ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$) holds superiority over ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$) under the null hypothesis. For $\Delta > 0$, Equation (4.13) can be written in the terms

$$\Delta^* = \frac{\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) + (1-\pi)^2 \mathrm{tr}[\boldsymbol{Q}^* \boldsymbol{Q}_{22.1}^{-1}]}{(1-\pi^2)}.$$

Replacing Δ^* in Equation (4.3) gives

$$\overbrace{\underbrace{\begin{array}{c}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) \\ -(1-\pi)^2 \mathrm{tr}[\boldsymbol{Q}^* \boldsymbol{Q}_{22.1}^{-1}] \\ +(1-\pi^2)\Delta ch_{\min}(\boldsymbol{Q}^* \boldsymbol{Q}_{22.1}^{-1})\end{array}}} \leq \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) \leq \overbrace{\underbrace{\begin{array}{c}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) \\ -(1-\pi)^2 \mathrm{tr}[\boldsymbol{Q}^* \boldsymbol{Q}_{22.1}^{-1}] \\ +(1-\pi^2)\Delta ch_{\max}(\boldsymbol{Q}^* \boldsymbol{Q}_{22.1}^{-1})\end{array}}}. \tag{4.14}$$

Next, consider the right-hand side of Equation (4.14) for the interval

$$0 \leq \Delta < \frac{(1-\pi)^2 \mathrm{tr}[\boldsymbol{Q}^* \boldsymbol{Q}_{22.1}^{-1}]}{(1-\pi^2) ch_{\max}(\boldsymbol{Q}^* \boldsymbol{Q}_{22.1}^{-1})},$$

ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$) $\leq$ ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$). It indicates that $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ dominates $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ when the null hypothesis is true or nearly true. On the other hand, consider the left-hand side of Equation (4.14), ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$) $\leq$ ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$), so that the null hypothesis does not hold when

$$\Delta \geq \frac{(1-\pi)^2 \mathrm{tr}[\boldsymbol{Q}^* \boldsymbol{Q}_{22.1}^{-1}]}{(1-\pi^2) ch_{\min}(\boldsymbol{Q}^* \boldsymbol{Q}_{22.1}^{-1})}.$$

4.1.2.6 Comparing $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$

First, we want to compare the ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$. Note that

$$\begin{aligned}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) = &-[1-\pi(2-\pi)H^*_{p_2+2}]\mathrm{tr}[\boldsymbol{Q}^* \boldsymbol{Q}_{22.1}^{-1}] \\ &+ \Delta^*[1-2\pi H^*_{p_2+2} + \pi(2-\pi)H^*_{p_2+4}],\end{aligned}$$

where $H_\nu^* = H_\nu(\chi^2_{p_2,\alpha};\Delta)$ and

$$\Delta^* = \frac{\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) + [1-\pi(2-\pi)H^*_{p_2+2}]\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}]}{[1-2\pi H^*_{p_2+2}+\pi(2-\pi)H^*_{p_2+4}]}.$$

By Equation (4.3), we get

$$\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) \le \overbrace{\underbrace{\begin{array}{c}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) - [1-\pi(2-\pi)H^*_{p_2+2}]\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}] \\ +\Delta ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1})\begin{bmatrix}1-2\pi H^*_{p_2+2}\\ +\pi(2-\pi)H^*_{p_2+4}\end{bmatrix}\end{array}}} \tag{4.15}$$

and

$$\overbrace{\underbrace{\begin{array}{c}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) - [1-\pi(2-\pi)H^*_{p_2+2}]\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}] \\ +\Delta ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1})\begin{bmatrix}1-2\pi H^*_{p_2+2}\\ +\pi(2-\pi)H^*_{p_2+4}\end{bmatrix}\end{array}}} \le \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}). \tag{4.16}$$

When the null hypothesis is true, Equations (4.15) and (4.16) are equal and we have

$$\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) = -[1-\pi(2-\pi)H^*_{p_2+2}]\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}] < 0.$$

This indicates that $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ outperforms $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ when $\Delta = 0$.

For the general case $\Delta > 0$, consider Equation (4.15), we have $\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) \le \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}})$ and then $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ dominates $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ when

$$\Delta \in \left[0, \frac{[1-\pi(2-\pi)H^*_{p_2+2}]\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}]}{ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1})[1-2\pi H^*_{p_2+2}+\pi(2-\pi)H^*_{p_2+4}]}\right).$$

Again, for Equation (4.16), we get $\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) \le \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})$, which means that the dominance situation changes in support of the $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ when

$$\Delta \in \left[\frac{[1-\pi(2-\pi)H^*_{p_2+2}]\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}]}{ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1})[1-2\pi H^*_{p_2+2}+\pi(2-\pi)H^*_{p_2+4}]}, \infty\right).$$

From the above results when $\Delta = 0$ and $\pi = 1$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ performs better than $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$. For $\Delta > 0$ and $\pi = 1$, the use of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ is suggested for

$$\Delta \le \frac{(1-H^*_{p_2+2})\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}]}{ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1})[1-2H^*_{p_2+2}+H^*_{p_2+4}]},$$

but becomes inferior to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$ when

$$\Delta \geq \frac{(1 - H^*_{p_2+2})\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}]}{ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1})[1 - 2H^*_{p_2+2} + H^*_{p_2+4}]}.$$

4.1.2.7 Comparing $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$

The difference in ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$ can be written as

$$\begin{aligned}
\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) &= -\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}] + \Delta^* \\
&\quad + c\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}]\left[2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] - c\mathbb{E}[\chi^{-4}_{p_2+2}(\Delta)]\right] \\
&\quad - c\Delta^*\left[2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] - 2\mathbb{E}[\chi^{-2}_{p_2+4}(\Delta)] + c\mathbb{E}[\chi^{-4}_{p_2+4}(\Delta)]\right] \\
&= -\mathrm{tr}[\mathrm{Q}^*\boldsymbol{Q}^{-1}_{22.1}]\left(1 - c\left\{2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] - c\mathbb{E}[\chi^{-4}_{p_2+2}(\Delta)]\right\}\right) \\
&\quad \Delta^*\left(1 - c\left\{\begin{matrix}2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] - 2\mathbb{E}[\chi^{-2}_{p_2+4}(\Delta)] \\ + c\mathbb{E}[\chi^{-4}_{p_2+4}(\Delta)]\end{matrix}\right\}\right).
\end{aligned}$$

Using Equation (4.12), we can rewrite this as

$$\begin{aligned}
\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) &= -\mathrm{tr}[\mathrm{Q}^*\boldsymbol{Q}^{-1}_{22.1}]\left(1 - c\left\{2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] - c\mathbb{E}[\chi^{-4}_{p_2+2}(\Delta)]\right\}\right) \\
&\quad \Delta^*\left(1 - c(p_2 + 2)\mathbb{E}[\chi^{-4}_{p_2+4}(\Delta)]\right).
\end{aligned}$$

and solving for Δ^*, we get

$$\Delta^* = \frac{\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) + \mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}]\left(1 - c\left\{\begin{matrix}2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] \\ -c\mathbb{E}[\chi^{-4}_{p_2+2}(\Delta)]\end{matrix}\right\}\right)}{1 - c(p_2 + 2)\mathbb{E}[\chi^{-4}_{p_2+4}(\Delta)]}.$$

Using Equation (4.3), we obtain

$$\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) \leq \overbrace{\underbrace{\begin{matrix}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - \mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1}]\left(1 - c\left\{\begin{matrix}2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] \\ -c\mathbb{E}[\chi^{-4}_{p_2+2}(\Delta)]\end{matrix}\right\}\right) \\ +\Delta ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}^{-1}_{22.1})\left(1 - \left\{\begin{matrix}c(p_2 + 2) \\ \mathbb{E}[\chi^{-4}_{p_2+4}(\Delta)]\end{matrix}\right\}\right)\end{matrix}}}\qquad (4.17)$$

and

$$\overbrace{\begin{matrix} \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - \mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\left(1 - c\left\{\begin{matrix} 2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] \\ -c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]\end{matrix}\right\}\right) \\ +\Delta c h_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\left(1 - \left\{\begin{matrix} c(p_2+2) \\ \mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]\end{matrix}\right\}\right)\end{matrix}}_{} \le \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}). \qquad (4.18)$$

From Equation (4.17) for the interval

$$\Delta \in \left[0, \frac{\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\left(1 - c\left[2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]\right]\right)}{c h_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\left(1 - c(p_2+2)\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]\right)}\right),$$

$\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) \le \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}})$ shows that $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ dominates $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$. While from Equation (4.18) for the interval

$$\Delta \in \left[\frac{\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\left(1 - c\left[2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]\right]\right)}{c h_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\left(1 - c(p_2+2)\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]\right)}, \infty\right),$$

$\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) \le \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})$ shows that $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$ starts dominating $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$. Hence, under local alternatives, neither $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ nor $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$ dominates the other asymptotically.

Next, we consider the difference between $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$. This yields

$$\begin{aligned}
\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}+}) &= \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) - \mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] + \Delta^* - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) \\
&\quad + \mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^{2}(\Delta) \le c)] \\
&\quad + \Delta^*\mathbb{E}[(1 - c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^{2}(\Delta) \le c)] \\
&\quad - 2\Delta^*\mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta)) I(\chi_{p_2+2}^{2}(\Delta) \le c)] \\
&= -\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] + \Delta^* \\
&\quad + c\,\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}](2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]) \\
&\quad - c\Delta^*(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - 2\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)] + c\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]) \\
&\quad + \mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^{2}(\Delta) \le c)] \\
&\quad + \Delta^*\mathbb{E}[(1 - c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^{2}(\Delta) \le c)] \\
&\quad - 2\Delta^*\mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta)) I(\chi_{p_2+2}^{2}(\Delta) \le c)].
\end{aligned}$$

Then, we get

$$\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}) = -\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\begin{pmatrix} 1+c(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)]-c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]) \\ -\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^{2}(\Delta)\leq c)] \end{pmatrix} + \Delta^*\begin{pmatrix} 1-c(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)]-2\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)]+c\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]) \\ +\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^{2}(\Delta)\leq c)] \\ -2\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta)) I(\chi_{p_2+2}^{2}(\Delta)\leq c)] \end{pmatrix},$$

and we have

$$\Delta^* = \frac{\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}) + \mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\begin{pmatrix} 1+c(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)]-c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]) \\ -\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^{2}(\Delta)\leq c)] \end{pmatrix}}{\begin{pmatrix} 1-c(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)]-2\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)]+c\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]) \\ +\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^{2}(\Delta)\leq c)] \\ -2\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta)) I(\chi_{p_2+2}^{2}(\Delta)\leq c)] \end{pmatrix}}. \tag{4.19}$$

Substitution of Δ^* into Equation (4.3) obtains:

$$\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) \leq \overbrace{\underbrace{\begin{matrix} \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}})-\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\begin{pmatrix} 1+c(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)]-c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]) \\ -\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^{2}(\Delta)\leq c)] \end{pmatrix} \\ +\Delta ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\begin{pmatrix} 1-c(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)]-2\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)]+c\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]) \\ +\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^{2}(\Delta)\leq c)] \\ -2\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta)) I(\chi_{p_2+2}^{2}(\Delta)\leq c)] \end{pmatrix} \end{matrix}}} \tag{4.20}$$

and

$$\overbrace{\underbrace{\begin{matrix} \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}})-\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\begin{pmatrix} 1+c(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)]-c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]) \\ -\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^{2}(\Delta)\leq c)] \end{pmatrix} \\ +\Delta ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\begin{pmatrix} 1-c(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)]-2\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)]+c\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]) \\ +\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^{2}(\Delta)\leq c)] \\ -2\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta)) I(\chi_{p_2+2}^{2}(\Delta)\leq c)] \end{pmatrix} \end{matrix}}} \leq \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}). \tag{4.21}$$

From Equation (4.20), $\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})$ is smaller than $\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}})$ when

$$\Delta \in \left[0, \frac{\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\begin{pmatrix} 1+c(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)]-c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]) \\ -\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^{2}(\Delta)\leq c)] \end{pmatrix}}{ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\begin{pmatrix} 1-c(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)]-2\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)]+c\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]) \\ +\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^{2}(\Delta)\leq c)] \\ -2\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta)) I(\chi_{p_2+2}^{2}(\Delta)\leq c)] \end{pmatrix}}\right]. \tag{4.22}$$

Also, from Equation (4.21), ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$) is smaller than ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$) when

$$\Delta \in \left[\frac{\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] \begin{pmatrix} 1 + c(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]) \\ -\mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^2(\Delta) \le c)] \end{pmatrix}}{ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}) \begin{pmatrix} 1-c(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)]-2\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)]+c\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]) \\ +\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^2(\Delta)\le c)] \\ -2\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^2(\Delta)\le c)] \end{pmatrix}}, \infty \right]. \tag{4.23}$$

Finally, we compare the ADQRs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$ when $\Delta = 0$, and we obtain

$$\begin{aligned} \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) &= -\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] \left(1 - c\left\{2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]\right\}\right) \\ &\le 0, \end{aligned}$$

$$\begin{aligned} \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}+}) &= -\mathrm{tr}[\mathrm{Q}^*\boldsymbol{Q}_{22.1}^{-1}] \begin{pmatrix} 1 + c(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]) \\ -\mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^2(\Delta) \le c)] \end{pmatrix} \\ &\le 0, \end{aligned}$$

and

$$\begin{aligned} \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) &= -\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\mathbb{E}\left[\left(1 - c\chi_{p_2+2}^{-2}(\Delta)\right)^2 I\left(\chi_{p_2+2}^2(\Delta) \le c\right)\right] \\ &\le 0. \end{aligned}$$

Since ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$) $\le$ ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$), we can conclude that ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$) $\le$ ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$) $\le$ ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$) when the null hypothesis is true and $p_2 \ge 3$.

4.1.2.8 Comparing $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$

Consider the ADQR difference between $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$. It can be written as

$$\begin{aligned} \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) &= -\pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}] + \pi^2\Delta^* \\ &\quad + \pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]H_{p_2+2}(\chi_{p_2,\alpha}^2;\Delta) \\ &\quad - \Delta^*[2\pi H_{p_2+2}(\chi_{p_2,\alpha}^2;\Delta) - \pi(2-\pi)H_{p_2+4}(\chi_{p_2,\alpha}^2;\Delta)] \\ &= -\pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\left(1 - H_{p_2+2}(\chi_{p_2,\alpha}^2;\Delta)\right) \\ &\quad + \Delta^* \begin{pmatrix} \pi^2 - 2\pi H_{p_2+2}(\chi_{p_2,\alpha}^2;\Delta) \\ +\pi(2-\pi)H_{p_2+4}(\chi_{p_2,\alpha}^2;\Delta) \end{pmatrix}. \end{aligned}$$

This give

$$\Delta^* = \frac{\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) + \pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\left(1 - H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)\right)}{\pi^2 - 2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) + \pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)}.$$

Using the Courant-Fisher Theorem, by substituting Δ^* into Equation (4.3), this yields

$$\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) \leq \overbrace{\underbrace{\begin{array}{c}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) - \pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\left(1 - H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)\right) \\ +\Delta ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\begin{pmatrix}\pi^2 - 2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\ +\pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)\end{pmatrix}\end{array}}}, \quad (4.24)$$

and

$$\overbrace{\underbrace{\begin{array}{c}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) - \pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\left(1 - H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)\right) \\ +\Delta ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\begin{pmatrix}\pi^2 - 2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\ +\pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)\end{pmatrix}\end{array}}} \leq \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}). \quad (4.25)$$

From Equation (4.24), we obtain $\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) \leq \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}})$ for the interval

$$0 \leq \Delta < \frac{\pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\left(1 - H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)\right)}{ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\left(\pi^2 - 2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) + \pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)\right)},$$

which implies that $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ is more efficient than $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$. Otherwise, from Equation (4.25), $\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) \leq \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}})$. It can be concluded that $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ performs better than $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ in the region of parameter

$$\Delta \geq \frac{\pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\left(1 - H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)\right)}{ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\left(\pi^2 - 2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) + \pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)\right)}.$$

For the case of $\Delta = 0$, we have

$$\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) = -\pi(2-\pi)\mathrm{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\left(1 - H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)\right),$$

which is always negative for all $\pi \in (0,1)$, and we get $\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) \leq \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}})$. So $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ can be suggested when the null hypothesis holds.

For the special case of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ when $\pi = 1$ is $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, we can conclude that $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ outperforms $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$ in the interval

$$\Delta \in \left[0, \frac{\operatorname{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\left(1 - H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)\right)}{ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\left(-2H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) + H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)\right)}\right).$$

However, when

$$\Delta \in \left[\frac{\operatorname{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}]\left(1 - H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)\right)}{ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})\left(-2H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) + H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)\right)}, \infty\right),$$

$\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ has a higher ADQR than $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$. For $\Delta = 0$, $\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) \leq \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}})$, which means that $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ can be also suggested when the subspace information is true.

4.1.2.9 Comparing $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$

The ADQR difference of the two estimators is given as:

$$\begin{aligned}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) &= -\operatorname{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}](1-\pi)^2 H^*_{p_2+2} \\ &\quad + \Delta^*(1-\pi)[2H^*_{p_2+2} - (1-\pi)H^*_{p_2+4}],\end{aligned}$$

where $H^*_\nu = H_\nu(\chi^2_{p_2,\alpha};\Delta)$ such that

$$\Delta^* = \frac{\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}) - \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) + \operatorname{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}](1-\pi)^2 H^*_{p_2+2}}{(1-\pi)[2H^*_{p_2+2} - (1-\pi)H^*_{p_2+4}]}. \tag{4.26}$$

Replacing Equation (4.26) in Equation (4.3) and solving for $\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}})$, we get

$$\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}) \leq \overbrace{\underbrace{\begin{matrix}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) - \operatorname{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}](1-\pi)^2 H^*_{p_2+2} \\ +\Delta ch_{\max}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})(1-\pi)\begin{bmatrix}2H^*_{p_2+2} \\ -(1-\pi)H^*_{p_2+4}\end{bmatrix}\end{matrix}}}\tag{4.27}$$

and

$$\overbrace{\underbrace{\begin{matrix}\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) - \operatorname{tr}[\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1}](1-\pi)^2 H^*_{p_2+2} \\ +\Delta ch_{\min}(\boldsymbol{Q}^*\boldsymbol{Q}_{22.1}^{-1})(1-\pi)\begin{bmatrix}2H^*_{p_2+2} \\ -(1-\pi)H^*_{p_2+4}\end{bmatrix}\end{matrix}}} \leq \mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}). \tag{4.28}$$

From Equation (4.27) for

$$\Delta \in \left[0, \frac{\text{tr}[\boldsymbol{Q}^{*}\boldsymbol{Q}_{22.1}^{-1}](1-\pi)H_{p_2+2}^{*}}{ch_{\max}(\boldsymbol{Q}^{*}\boldsymbol{Q}_{22.1}^{-1})[2H_{p_2+2}^{*}-(1-\pi)H_{p_2+4}^{*}]}\right)$$

it reduces to $\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{PT}}) \leq \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{SP}})$, showing the dominance of $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$ over $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$. On the other hand, for

$$\Delta \in \left[\frac{\text{tr}[\boldsymbol{Q}^{*}\boldsymbol{Q}_{22.1}^{-1}](1-\pi)H_{p_2+2}^{*}}{ch_{\min}(\boldsymbol{Q}^{*}\boldsymbol{Q}_{22.1}^{-1})[2H_{p_2+2}^{*}-(1-\pi)H_{p_2+4}^{*}]}, \infty\right),$$

the Equation (4.28) as $\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{SP}}) \leq \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{PT}})$, which indicates that $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$ outperforms $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$. Moreover, for a large value of π (close to 1) the dominance region of $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$ is small.

When the null hypothesis is true, the ADQR difference between $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$ and $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$ is as follows:

$$\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{PT}}) - \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{SP}}) = -\text{tr}[\boldsymbol{Q}^{*}\boldsymbol{Q}_{22.1}^{-1}](1-\pi)^2 H_{p_2+2}^{*} \leq 0.$$

Thus, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$ is superior to $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$ for all $\pi \in (0,1)$. Moreover, the $\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{SP}})$ decreases and approaches $\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{PT}})$ when π increases.

4.1.2.10 Comparing $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$ and $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$

Since the ADQRs of $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$ and $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$ intersect each other at some point neither of the two estimators is better. However, in most of the parameter space, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$ performs better than $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$. The difference between $\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{S}})$ and $\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{SP}})$ under the null hypothesis is

$$\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{S}}) - \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{SP}}) = \text{tr}[\boldsymbol{Q}^{*}\boldsymbol{Q}_{22.1}^{-1}]\left[\begin{array}{c}\pi(2-\pi)H_{p_2+2}(\chi_{p_2,\alpha}^2;\Delta) \\ -c\left[2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]\right]\end{array}\right].$$

Since $\mathbb{E}[\chi_{p_2+2}^{-2}] = \frac{1}{p_2}$ and $\mathbb{E}[\chi_{p_2+2}^{-4}] = \frac{1}{p_2(p_2-2)}$, it reduces to

$$\text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{S}}) - \text{ADQR}(\widehat{\boldsymbol{\beta}}_1^{\text{SP}}) = \text{tr}[\boldsymbol{Q}^{*}\boldsymbol{Q}_{22.1}^{-1}]\left[\pi(2-\pi)H_{p_2+2}(\chi_{p_2,\alpha}^2;\Delta) - \frac{(p_2-2)}{p_2}\right] > 0.$$

Thus, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$ has a smaller ADQR than $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$ when $\Delta = 0$, $p_2 \geq 3$, and

$$\pi(2-\pi)H_{p_2+2}(\chi_{p_2,\alpha}^2;0) - \frac{(p_2-2)}{p_2} > 0$$

$$\pi^2 - 2\pi < -\frac{(p_2-2)}{p_2 H_{p_2+2}(\chi_{p_2,\alpha}^2;0)}.$$

Hence,

$$\pi^2 - 2\pi + \frac{(p_2 - 2)}{p_2 H_{p_2+2}(\chi^2_{p_2,\alpha}; 0)} < 0,$$

which is a quadratic equation in π with $a = 1$, $b = -2$, and $c = \frac{(p_2-2)}{p_2 H_{p_2+2}(\chi^2_{p_2,\alpha};0)}$, then the solutions is

$$\pi = \frac{2 \pm \sqrt{4 - \frac{4(p_2-2)}{p_2 H_{p_2+2}(\chi^2_{p_2,\alpha};0)}}}{2} = 1 \pm \sqrt{\left(1 - \frac{(p_2 - 2)}{p_2 H_{p_2+2}(\chi^2_{p_2,\alpha}; 0)}\right)}.$$

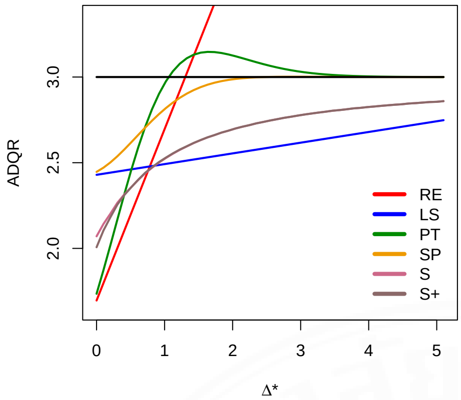
Since $\pi \in (0, 1)$, it must satisfy

$$\pi = 1 - \sqrt{\left(1 - \frac{(p_2 - 2)}{p_2 H_{p_2+2}(\chi^2_{p_2,\alpha}; 0)}\right)} \quad \text{and} \quad \frac{(p_2 - 2)}{p_2} < H_{p_2+2}(\chi^2_{p_2,\alpha}; 0). \tag{4.29}$$

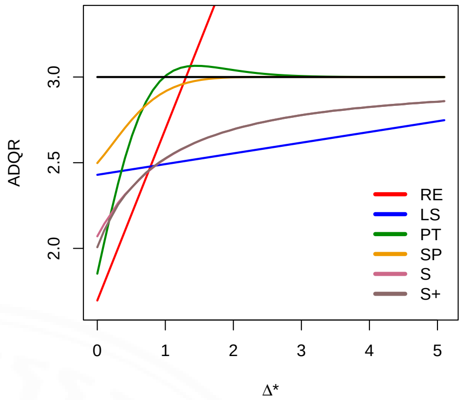
Otherwise, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$ has a smaller ADQR than $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$.

More commonly, we let $\Psi^*_{\pi,\alpha}$ be the point at which $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$ intersect provided the condition in Equation (4.29) is satisfied for fixed π and α. The $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ outperforms $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$ when $\Delta \in [0, \Psi^*_{\pi,\alpha})$. Conversely, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$ dominates $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ in the region $\Delta \in [\Psi^*_{\pi,\alpha}, \infty)$. On the other hand, if Equation (4.29) is not satisfied and there is no intersecting point in the entire parameter space, then $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$ still outperforms $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ for the dimension $p_2 \geq 3$. Therefore, for $p_2 < 3$, the use of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ is suggested for parameter estimation.

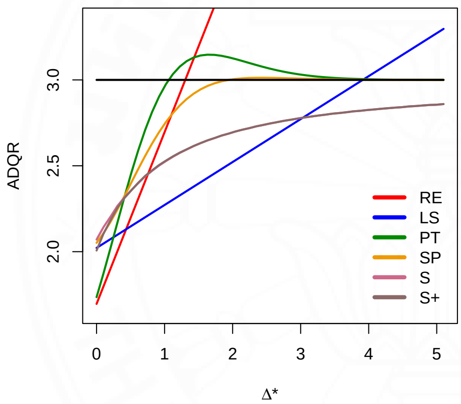
Moreover, we also plotted the ADQRs of the suggested estimators in Corollary 2 to facilitate comparison. A graphical representation of $p_1 = 3$, $p_2 = 7$ and 11, $\alpha = 0.01$ and 0.05, and $\pi = 0.25, 0.50,$ and 0.75 is shown in Figures 4.3 and 4.4. We can see that the ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ is a constant, whereas the ADQRs of all other estimators depend on the values of Δ^*. At $\Delta^* = 0$ or when the null hypothesis is true, all suggested estimators outperform $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ has the lowest ADQR, especially when p_2 increases. When Δ^* increases, the ADQRs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ become unbounded, in which case the ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ increases more slowly since it depends on π. The ADQRs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ increase first, then reach the maximum value. The ADQRs in this part are larger than that of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$. After passing through that point, both ADQRs approach $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$. The ADQRs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ are smaller than that of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$, and are more apparent when p_2 increases.



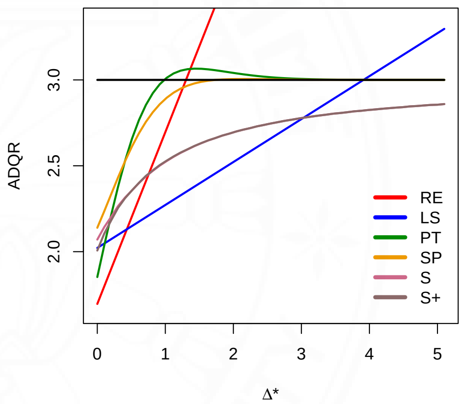
(a) $\pi = 0.25$, $\alpha = 0.01$



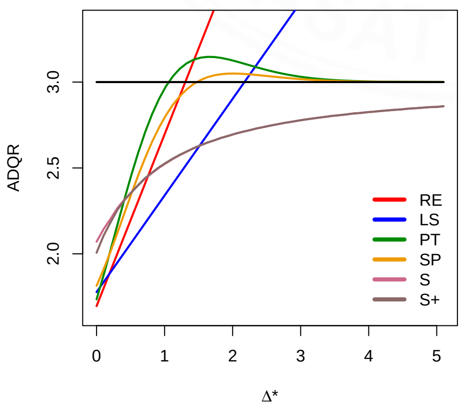
(b) $\pi = 0.25$, $\alpha = 0.05$



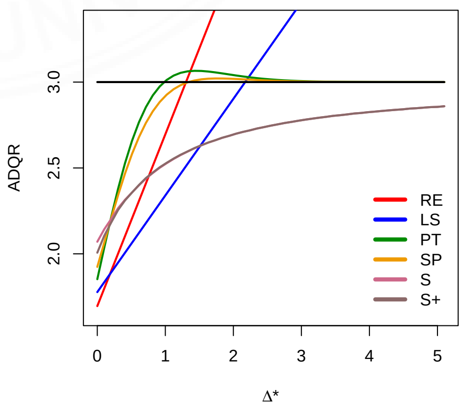
(c) $\pi = 0.50$, $\alpha = 0.01$



(d) $\pi = 0.50$, $\alpha = 0.05$

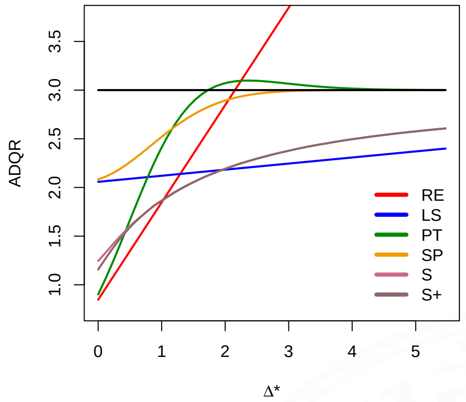


(e) $\pi = 0.75$, $\alpha = 0.01$

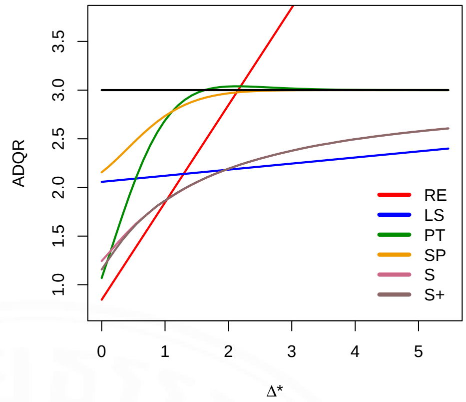


(f) $\pi = 0.75$, $\alpha = 0.05$

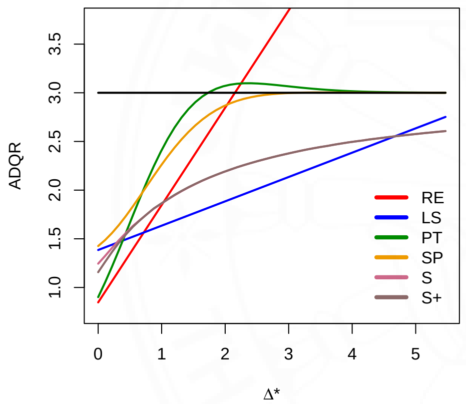
Figure 4.3 ADQR curves of the suggested estimators for nonlinear regression model with $p_1 = 3$ and $p_2 = 7$



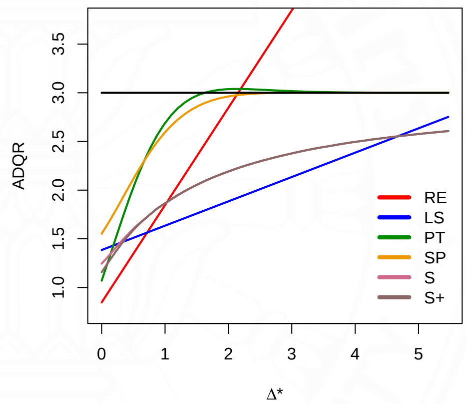
(a) $\pi = 0.25$, $\alpha = 0.01$



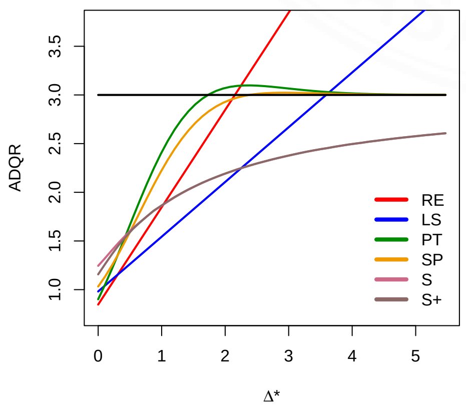
(b) $\pi = 0.25$, $\alpha = 0.05$



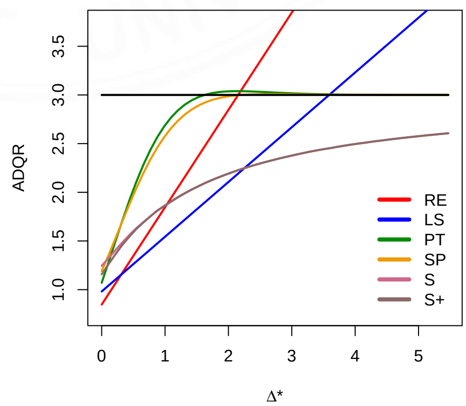
(c) $\pi = 0.50$, $\alpha = 0.01$



(d) $\pi = 0.50$, $\alpha = 0.05$



(e) $\pi = 0.75$, $\alpha = 0.01$



(f) $\pi = 0.75$, $\alpha = 0.05$

Figure 4.4 ADQR curves of the suggested estimators for nonlinear regression model with $p_1 = 3$ and $p_2 = 11$

The comparison of the ADQRs of the proposed estimators shows that none of the proposed estimators outperform the others. Each estimator has a particular region in which its performance is superior. However, the suggested estimators are still more effective than the unrestricted and restricted estimators in cases where the correctness of the UPI is not known.

4.2 Simulation Results

For each model of interest, we assume the model has p parameters and a sample size n. The response values of the Cobb-Douglas model were simulated from

$$y_i = \beta_1(x_{i1}^{\beta_2})(x_{i2}^{\beta_3})\dots(x_{i,p-1}^{\beta_p}) + \varepsilon_i, \tag{4.30}$$

where $x_{ij} \overset{i.i.d}{\sim} \mathcal{N}(8,1)$, $\varepsilon_i \overset{i.i.d}{\sim} \mathcal{N}(0,1)$, and $i = 1, 2, \dots, n$. For the exponential model, the response values were generated from

$$y_i = \beta_1 e^{\beta_2 x_{i1} + \beta_3 x_{i2} + \dots + \beta_p x_{i,p-1}} + \varepsilon_i, \tag{4.31}$$

where $x_{ij} \overset{i.i.d}{\sim} \mathcal{N}(0,1)$, $\varepsilon_i \overset{i.i.d}{\sim} \mathcal{N}(0,1)$, and $i = 1, 2, \dots, n$. Lastly, for $i = 1, 2, \dots, n$, the response values of the monomolecular model were given by

$$y_i = \beta_1 \left(1 - \beta_2 e^{-\beta_3 x_{i1} - \beta_4 x_{i2} - \dots - \beta_p x_{i,p-2}}\right) + \varepsilon_i, \tag{4.32}$$

where $x_{ij} \overset{i.i.d}{\sim} \mathcal{N}(0,1)$ and $\varepsilon_i \overset{i.i.d}{\sim} \mathcal{N}(0,1)$. All models and simulations were written in the R programming language.

4.2.1 Low-Dimensional Data Setting

As our focus was on testing the null hypothesis $\mathrm{H}_0 : \boldsymbol{\beta}_2 = \mathbf{0}$, we partitioned the regression coefficients as $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$ and $\boldsymbol{\beta}_{\mathbf{H}_0} = (\boldsymbol{\beta}_1^\top, \mathbf{0}_{p_2}^\top)^\top$, where $\boldsymbol{\beta}$ is the coefficient vector taking true values in the simulation model, and $\boldsymbol{\beta}_{\mathbf{H}_0}$ is the coefficient vector taking true values under subspace information. To access the effect of uncertainty in the subspace information on the estimators, we determined the divergence between coefficients in the simulation model and the subspace information as follows:

$$\Delta^{\text{sim}} = \left\|\boldsymbol{\beta} - \boldsymbol{\beta}_{\mathbf{H}_0}\right\| = (\boldsymbol{\beta} - \boldsymbol{\beta}_{\mathbf{H}_0})^\top(\boldsymbol{\beta} - \boldsymbol{\beta}_{\mathbf{H}_0}), \tag{4.33}$$

where $\|\cdot\|$ is the Euclidean norm. This is the measure of how far away we go from the null hypothesis. If $\Delta^{\text{sim}} = 0$, the null hypothesis was true and it means that the available subspace information was correct. If $\Delta^{\text{sim}} > 0$, the null hypothesis was not true and it indicates that the available subspace information was incorrect.

In this dissertation, the two situations of the subspace information considered were correct subspace information ($\Delta^{\text{sim}} = 0$) and uncertain subspace information ($\Delta^{\text{sim}} \geq 0$). For $\Delta^{\text{sim}} = 0$, we set the true values of $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$ in the simulations as Equation (3.43). When $\Delta^{\text{sim}} \geq 0$, the simulation model had $\boldsymbol{\beta}_1 = (\beta_1, \beta_2, ..., \beta_{p_1})^\top$ and $\boldsymbol{\beta}_2 = (\Delta^{\text{sim}}, \mathbf{0}_{p_2-1}^\top)^\top$.

To examine the impacts of parameters π and α for the PT and SP estimators, the values of π were set to 0.25, 0.50, and 0.75 and significant levels α were set to 0.01, 0.05, and 0.10. We set the sample size (n) as 100, and the number of trials was $N = 5,000$ as this was sufficient to obtain stable results. Note that the LASSO and aLASSO estimators were excluded when $\Delta^{\text{sim}} > 0$, because this estimator does not make use of subspace information.

We discuss the results of a simulation study for each relevant model in two parts, one for the case when the null hypothesis is assumed to be true and the other when it may not be true.

4.2.1.1 Cobb-Douglas Model

(1) Correct Subspace Information ($\Delta^{\text{sim}} = 0$)

Assuming the subspace information is correct, the true values of parameter $\boldsymbol{\beta}$ in the simulation model were set to $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$, where $\boldsymbol{\beta}_1 = (0.75, 0.75, 0.75)^\top$ and $\boldsymbol{\beta}_2 = \mathbf{0}_{p_2}^\top$. The RMSE was computed using active parameter $p_1 = 3$ and inactive parameters $p_2 = 3, 5, 7, 11, 15$ to cover a small to a large number of the inactive parameters. The RMSEs of all proposed parameters for five pairings of p_1 and p_2 are shown in Tables 4.1. We can summarize these results as follows:

1. For fixed π and α, the RMSEs of all estimators increased as the number of inactive parameters increased. The RE had the highest RMSE, and the UE had the lowest RMSE. It is clear that the RE dominated all other proposed estimators and all estimators were better than UE.

Table 4.1 RMSEs of $\widehat{\beta}_1^{\text{RE}}$, $\widehat{\beta}_1^{\text{LS}}$, $\widehat{\beta}_1^{\text{PT}}$, $\widehat{\beta}_1^{\text{SP}}$, $\widehat{\beta}_1^{\text{S}}$, $\widehat{\beta}_1^{\text{S}^+}$, $\widehat{\beta}_1^{\text{LASSO}}$, and $\widehat{\beta}_1^{\text{aLASSO}}$ with respect to $\widehat{\beta}_1^{\text{UE}}$ for Cobb-Douglas model with $p_1 = 3$ at $\Delta^{\text{sim}} = 0$

Estimator			Number of Inactive Parameters (p_2)				
			3	5	7	11	15
RE			2.128	2.944	4.128	6.118	9.124
LS	$\pi = 0.25$		1.296	1.405	1.489	1.569	1.620
	$\pi = 0.50$		1.647	1.978	2.296	2.651	2.929
	$\pi = 0.75$		1.974	2.621	3.418	4.561	5.810
PT		$\alpha = 0.01$	2.013	2.748	3.307	5.171	6.516
		$\alpha = 0.05$	1.746	2.276	2.515	3.617	4.250
		$\alpha = 0.10$	1.610	1.982	2.177	2.862	3.414
SP	$\pi = 0.25$	$\alpha = 0.01$	1.276	1.384	1.434	1.538	1.572
		$\alpha = 0.05$	1.225	1.325	1.355	1.459	1.490
		$\alpha = 0.10$	1.194	1.277	1.308	1.392	1.437
	$\pi = 0.50$	$\alpha = 0.01$	1.594	1.909	2.082	2.505	2.677
		$\alpha = 0.05$	1.463	1.725	1.818	2.173	2.304
		$\alpha = 0.10$	1.389	1.592	1.677	1.938	2.097
	$\pi = 0.75$	$\alpha = 0.01$	1.880	2.473	2.869	4.049	4.701
		$\alpha = 0.05$	1.660	2.108	2.290	3.088	3.463
		$\alpha = 0.10$	1.543	1.868	2.023	2.545	2.920
S			1.143	1.620	2.087	3.123	4.141
S^+			1.310	1.892	2.397	3.667	5.094
LASSO			1.161	1.537	2.128	2.945	3.924
aLASSO			1.179	1.673	2.142	3.134	4.336

2. For a fixed p_2, the RMSE of the LS estimator increased as π increased. For fixed π, the LS estimator also had larger a RMSE than the SP estimator for all values of α.

3. The PT estimator depended on the value of α and performed well when α decreased. The performance of the SP estimator depended on the choice of π and α, and performance was better as π increased and α decreased.

4. The performance of the PT estimator was superior to that of the SP estimator at the same significant level of α. At $\alpha = 0.01$, the PT estimator also had higher RMSE than LS, S, S^+, LASSO, and aLASSO.

5. For all p_2, the RMSE of the S^+ estimator was greater than that of the shrinkage (S) estimator. The aLASSO estimator also outperformed the LASSO estimator.

(2) Uncertain Subspace Information ($\Delta^{\text{sim}} \geq 0$)

The performance of the estimators was investigated when the correctness of the information was unknown. The simulation model comprised $p_1 = 5$ active parameters and $p_2 - 1$ inactive parameters, with p_2 set to 3, 5, 7, 11, and 15. We set the true values of the active parameter as $\boldsymbol{\beta}_1 = (2, 0.75, 0.75, 0.25, 0.25)^\top$ and the inactive parameter as $\boldsymbol{\beta}_2 = (\beta_6, \mathbf{0})^\top$ where β_6 is a scalar and that can have multiple values. The values of β_6 were set to equal Δ^{sim}, and lay between 0 and 0.1. The RMSEs of the proposed estimators for each p_2 are shown in Tables 4.2 to 4.4 and represented graphically in Figures 4.5 to 4.9. These results can be summarized as follows:

1. As Δ^{sim} moved away from 0, the RMSE of the RE converged on 0 and the curve fell below the horizontal line at RMSE $= 1$. For fixed π, the RMSE of the LS estimator also approached 0 with a slower speed than that of the RE. The performance of the LS estimator was superior to that of all other estimators in some portion of Δ^{sim}.

2. The RMSEs of both the PT and SP estimators first fell below the horizontal line, then increased to 1 as Δ^{sim} moved far away from zero. The performance of the PT and SP estimators was poor when the significant level α was large. For fixed α, the PT estimator performed more efficiently than the SP estimator when Δ^{sim} was close to zero. However, it became inferior to the SP estimator in some space of $\Delta^{\text{sim}} > 0$, then they eventually became equal.

3. When Δ^{sim} was close to zero, the RMSE of the PT estimator exceeded those of the S and S^+ estimators in a small part of the parameter space for a small α. When α was small, π was large, and Δ^{sim} was nearly zero, the SP estimator dominated the shrinkage estimator for all p_2, and it outperformed the S^+ estimator for a small p_2. Conversely, the S and S^+ estimators were superior to the PT and SP estimators when Δ^{sim} was far from zero.

4. The RMSE of the S^+ estimator was initially higher than that of the shrinkage estimator when $\Delta^{\text{sim}} > 0$, but they converged to the same values as Δ^{sim} increased. This suggests that the S^+ estimator is preferred, as its performance remains robust even if the assumed model is grossly wrong.

Table 4.2 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cobb-Douglas model with $p_1 = 5$ and $\pi = 0.25$ at $\Delta^{\mathrm{sim}} \geq 0$

p_2	Δ^{sim}	RE	LS	PT			SP			S	S^+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
3	0.000	1.705	1.222	1.648	1.528	1.416	1.209	1.178	1.148	1.148	1.241
	0.005	1.454	1.209	1.376	1.287	1.233	1.181	1.145	1.118	1.136	1.178
	0.010	1.021	1.180	0.974	0.951	0.946	1.119	1.067	1.044	1.066	1.087
	0.015	0.676	1.135	0.726	0.794	0.849	1.026	0.996	0.991	1.024	1.032
	0.020	0.453	1.077	0.676	0.817	0.888	0.965	0.976	0.984	1.010	1.011
	0.025	0.314	1.008	0.778	0.906	0.946	0.967	0.984	0.991	1.005	1.005
	0.030	0.226	0.933	0.895	0.968	0.990	0.983	0.995	0.998	1.002	1.002
	0.035	0.167	0.854	0.973	0.997	0.998	0.996	0.999	1.000	1.001	1.001
	0.040	0.128	0.776	0.997	1.000	1.000	0.999	1.000	1.000	1.000	1.000
	0.060	0.053	0.501	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.080	0.026	0.315	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.100	0.015	0.200	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.000	2.185	1.318	2.089	1.862	1.664	1.303	1.261	1.217	1.417	1.632
	0.005	1.854	1.305	1.750	1.549	1.425	1.281	1.224	1.185	1.381	1.512
	0.010	1.300	1.277	1.193	1.119	1.100	1.207	1.144	1.109	1.252	1.309
	0.015	0.860	1.234	0.854	0.887	0.906	1.106	1.055	1.033	1.142	1.161
	0.020	0.577	1.178	0.726	0.835	0.886	1.020	0.999	0.996	1.077	1.082
	0.025	0.400	1.112	0.762	0.881	0.926	0.980	0.987	0.990	1.045	1.046
	0.030	0.288	1.039	0.855	0.950	0.975	0.981	0.992	0.996	1.028	1.029
	0.035	0.214	0.961	0.953	0.992	0.996	0.993	0.999	0.999	1.019	1.019
	0.040	0.163	0.882	0.991	0.999	0.999	0.999	1.000	1.000	1.013	1.013
	0.060	0.067	0.591	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
	0.080	0.034	0.382	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001
	0.100	0.019	0.248	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
7	0.000	2.703	1.388	2.555	2.224	1.952	1.368	1.322	1.275	1.786	2.050
	0.005	2.304	1.376	2.145	1.878	1.679	1.345	1.290	1.244	1.680	1.875
	0.010	1.614	1.351	1.473	1.332	1.230	1.282	1.210	1.157	1.452	1.555
	0.015	1.066	1.312	1.005	0.973	0.973	1.178	1.101	1.067	1.278	1.313
	0.020	0.713	1.260	0.792	0.864	0.901	1.065	1.027	1.015	1.171	1.181
	0.025	0.493	1.198	0.776	0.884	0.922	1.006	1.000	0.998	1.108	1.112
	0.030	0.354	1.128	0.846	0.949	0.975	0.993	0.997	0.998	1.073	1.074
	0.035	0.263	1.053	0.947	0.990	0.996	0.995	0.999	0.999	1.051	1.051
	0.040	0.200	0.974	0.991	0.997	1.000	0.999	1.000	1.000	1.037	1.037
	0.060	0.082	0.675	1.000	1.000	1.000	1.000	1.000	1.000	1.012	1.012
	0.080	0.041	0.446	1.000	1.000	1.000	1.000	1.000	1.000	1.004	1.004
	0.100	0.023	0.294	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001

Table 4.2 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for Cobb-Douglas model with $p_1 = 5$ and $\pi = 0.25$ at $\Delta^{\text{sim}} \geq 0$ (Cont.)

				PT			SP				
p_2	Δ^{sim}	RE	LS	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	S	S^+
11	0.000	3.888	1.480	3.483	2.824	2.429	1.450	1.392	1.346	2.411	2.869
	0.005	3.319	1.468	2.937	2.472	2.129	1.430	1.372	1.319	2.216	2.582
	0.010	2.336	1.445	2.023	1.740	1.542	1.375	1.297	1.237	1.869	2.070
	0.015	1.549	1.412	1.337	1.202	1.144	1.274	1.180	1.135	1.576	1.661
	0.020	1.040	1.368	0.969	0.968	0.971	1.150	1.080	1.052	1.382	1.408
	0.025	0.721	1.315	0.846	0.909	0.943	1.054	1.018	1.011	1.259	1.264
	0.030	0.518	1.255	0.859	0.945	0.972	1.006	1.003	1.002	1.180	1.180
	0.035	0.385	1.188	0.931	0.984	0.995	0.999	1.001	1.000	1.127	1.127
	0.040	0.293	1.117	0.982	0.998	1.001	1.000	1.000	1.001	1.092	1.092
	0.060	0.121	0.825	1.000	1.000	1.000	1.000	1.000	1.000	1.029	1.029
	0.080	0.061	0.576	1.000	1.000	1.000	1.000	1.000	1.000	1.010	1.010
	0.100	0.034	0.394	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
15	0.000	5.386	1.551	4.315	3.389	2.762	1.504	1.444	1.385	3.131	3.773
	0.005	4.590	1.541	3.730	2.848	2.394	1.490	1.412	1.355	2.862	3.367
	0.010	3.197	1.523	2.491	2.020	1.731	1.429	1.342	1.270	2.387	2.642
	0.015	2.099	1.495	1.651	1.385	1.286	1.339	1.224	1.173	1.950	2.044
	0.020	1.398	1.458	1.140	1.077	1.039	1.205	1.122	1.081	1.639	1.665
	0.025	0.965	1.413	0.930	0.941	0.951	1.096	1.040	1.023	1.438	1.443
	0.030	0.691	1.360	0.871	0.938	0.965	1.027	1.007	1.003	1.309	1.310
	0.035	0.512	1.300	0.915	0.971	0.987	1.001	0.999	0.999	1.225	1.225
	0.040	0.390	1.235	0.972	0.994	0.997	0.999	1.000	1.000	1.168	1.168
	0.060	0.160	0.954	1.000	1.000	1.000	1.000	1.000	1.000	1.063	1.063
	0.080	0.080	0.693	1.000	1.000	1.000	1.000	1.000	1.000	1.029	1.029
	0.100	0.045	0.488	1.000	1.000	1.000	1.000	1.000	1.000	1.014	1.014

Table 4.3 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cobb-Douglas model with $p_1 = 5$ and $\pi = 0.50$ at $\Delta^{\mathrm{sim}} \geq 0$

				PT			SP				
p_2	Δ^{sim}	RE	LS	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	S	S^+
3	0.000	1.705	1.451	1.648	1.528	1.416	1.420	1.349	1.283	1.148	1.241
	0.005	1.454	1.394	1.376	1.287	1.233	1.334	1.259	1.209	1.136	1.178
	0.010	1.021	1.260	0.974	0.951	0.946	1.158	1.081	1.050	1.066	1.087
	0.015	0.676	1.085	0.726	0.794	0.849	0.972	0.954	0.960	1.024	1.032
	0.020	0.453	0.904	0.676	0.817	0.888	0.887	0.935	0.959	1.010	1.011
	0.025	0.314	0.739	0.778	0.906	0.946	0.915	0.963	0.979	1.005	1.005
	0.030	0.226	0.599	0.895	0.968	0.990	0.959	0.987	0.996	1.002	1.002
	0.035	0.167	0.486	0.973	0.997	0.998	0.990	0.999	0.999	1.001	1.001
	0.040	0.128	0.396	0.997	1.000	1.000	0.999	1.000	1.000	1.000	1.000
	0.060	0.053	0.187	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.080	0.026	0.100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.100	0.015	0.058	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.000	2.185	1.702	2.089	1.862	1.664	1.658	1.546	1.437	1.417	1.632
	0.005	1.854	1.635	1.750	1.549	1.425	1.572	1.435	1.348	1.381	1.512
	0.010	1.300	1.487	1.193	1.119	1.100	1.343	1.227	1.172	1.252	1.309
	0.015	0.860	1.292	0.854	0.887	0.906	1.110	1.051	1.027	1.142	1.161
	0.020	0.577	1.087	0.726	0.835	0.886	0.964	0.967	0.974	1.077	1.082
	0.025	0.400	0.897	0.762	0.881	0.926	0.926	0.962	0.974	1.045	1.046
	0.030	0.288	0.734	0.855	0.950	0.975	0.949	0.981	0.991	1.028	1.029
	0.035	0.214	0.600	0.953	0.992	0.996	0.982	0.997	0.999	1.019	1.019
	0.040	0.163	0.492	0.991	0.999	0.999	0.997	0.999	1.000	1.013	1.013
	0.060	0.067	0.236	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
	0.080	0.034	0.126	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001
	0.100	0.019	0.074	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
7	0.000	2.703	1.914	2.555	2.224	1.952	1.851	1.713	1.585	1.786	2.050
	0.005	2.304	1.847	2.145	1.878	1.679	1.757	1.606	1.490	1.680	1.875
	0.010	1.614	1.694	1.473	1.332	1.230	1.531	1.376	1.270	1.452	1.555
	0.015	1.066	1.486	1.005	0.973	0.973	1.253	1.133	1.086	1.278	1.313
	0.020	0.713	1.263	0.792	0.864	0.901	1.039	1.009	1.002	1.171	1.181
	0.025	0.493	1.052	0.776	0.884	0.922	0.961	0.978	0.983	1.108	1.112
	0.030	0.354	0.868	0.846	0.949	0.975	0.961	0.987	0.993	1.073	1.074
	0.035	0.263	0.714	0.947	0.990	0.996	0.985	0.997	0.998	1.051	1.051
	0.040	0.200	0.588	0.991	0.997	1.000	0.997	0.999	1.000	1.037	1.037
	0.060	0.082	0.286	1.000	1.000	1.000	1.000	1.000	1.000	1.012	1.012
	0.080	0.041	0.153	1.000	1.000	1.000	1.000	1.000	1.000	1.004	1.004
	0.100	0.023	0.090	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001

Table 4.3 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cobb-Douglas model with $p_1 = 5$ and $\pi = 0.50$ at $\Delta^{\mathrm{sim}} \geq 0$ (Cont.)

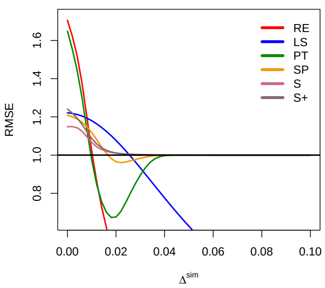
p_2	Δ^{sim}	RE	LS	PT			SP			S	S+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
11	0.000	3.888	2.254	3.483	2.824	2.429	2.141	1.935	1.788	2.411	2.869
	0.005	3.319	2.181	2.937	2.472	2.129	2.044	1.853	1.697	2.216	2.582
	0.010	2.336	2.025	2.023	1.740	1.542	1.812	1.603	1.458	1.869	2.070
	0.015	1.549	1.812	1.337	1.202	1.144	1.490	1.302	1.219	1.576	1.661
	0.020	1.040	1.575	0.969	0.968	0.971	1.203	1.102	1.065	1.382	1.408
	0.025	0.721	1.343	0.846	0.909	0.943	1.039	1.008	1.005	1.259	1.264
	0.030	0.518	1.131	0.859	0.945	0.972	0.981	0.994	0.998	1.180	1.180
	0.035	0.385	0.947	0.931	0.984	0.995	0.987	0.998	1.000	1.127	1.127
	0.040	0.293	0.792	0.982	0.998	1.001	0.997	1.000	1.001	1.092	1.092
	0.060	0.121	0.401	1.000	1.000	1.000	1.000	1.000	1.000	1.029	1.029
	0.080	0.061	0.219	1.000	1.000	1.000	1.000	1.000	1.000	1.010	1.010
	0.100	0.034	0.130	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
15	0.000	5.386	2.559	4.315	3.389	2.762	2.354	2.116	1.913	3.131	3.773
	0.005	4.590	2.491	3.730	2.848	2.394	2.274	1.989	1.806	2.862	3.367
	0.010	3.197	2.337	2.491	2.020	1.731	2.000	1.735	1.547	2.387	2.642
	0.015	2.099	2.121	1.651	1.385	1.286	1.677	1.410	1.305	1.950	2.044
	0.020	1.398	1.874	1.140	1.077	1.039	1.328	1.186	1.119	1.639	1.665
	0.025	0.965	1.622	0.930	0.941	0.951	1.117	1.043	1.022	1.438	1.443
	0.030	0.691	1.385	0.871	0.938	0.965	1.010	0.999	0.998	1.309	1.310
	0.035	0.512	1.175	0.915	0.971	0.987	0.986	0.994	0.997	1.225	1.225
	0.040	0.390	0.993	0.972	0.994	0.997	0.994	0.999	0.999	1.168	1.168
	0.060	0.160	0.515	1.000	1.000	1.000	1.000	1.000	1.000	1.063	1.063
	0.080	0.080	0.286	1.000	1.000	1.000	1.000	1.000	1.000	1.029	1.029
	0.100	0.045	0.170	1.000	1.000	1.000	1.000	1.000	1.000	1.014	1.014

Table 4.4 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for Cobb-Douglas model with $p_1 = 5$ and $\pi = 0.75$ at $\Delta^{\text{sim}} \geq 0$

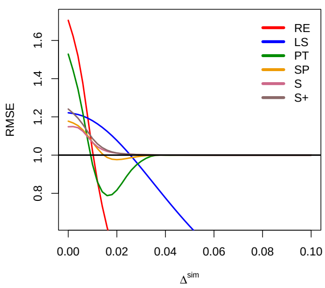
p_2	Δ^{sim}	RE	LS	PT			SP			S	S+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
3	0.000	1.705	1.634	1.648	1.528	1.416	1.585	1.479	1.381	1.148	1.241
	0.005	1.454	1.490	1.376	1.287	1.233	1.409	1.312	1.251	1.136	1.178
	0.010	1.021	1.194	0.974	0.951	0.946	1.102	1.038	1.015	1.066	1.087
	0.015	0.676	0.892	0.726	0.794	0.849	0.860	0.882	0.911	1.024	1.032
	0.020	0.453	0.652	0.676	0.817	0.888	0.784	0.880	0.927	1.010	1.011
	0.025	0.314	0.480	0.778	0.906	0.946	0.850	0.937	0.964	1.005	1.005
	0.030	0.226	0.359	0.895	0.968	0.990	0.930	0.978	0.993	1.002	1.002
	0.035	0.167	0.273	0.973	0.997	0.998	0.982	0.998	0.999	1.001	1.001
	0.040	0.128	0.212	0.997	1.000	1.000	0.998	1.000	1.000	1.000	1.000
	0.060	0.053	0.091	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.080	0.026	0.046	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.100	0.015	0.026	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.000	2.185	2.052	2.089	1.862	1.664	1.973	1.781	1.607	1.417	1.632
	0.005	1.854	1.866	1.750	1.549	1.425	1.765	1.563	1.440	1.381	1.512
	0.010	1.300	1.497	1.193	1.119	1.100	1.337	1.217	1.168	1.252	1.309
	0.015	0.860	1.122	0.854	0.887	0.906	1.010	0.989	0.981	1.142	1.161
	0.020	0.577	0.823	0.726	0.835	0.886	0.855	0.909	0.937	1.077	1.082
	0.025	0.400	0.607	0.762	0.881	0.926	0.850	0.926	0.953	1.045	1.046
	0.030	0.288	0.455	0.855	0.950	0.975	0.906	0.966	0.984	1.028	1.029
	0.035	0.214	0.347	0.953	0.992	0.996	0.969	0.995	0.998	1.019	1.019
	0.040	0.163	0.270	0.991	0.999	0.999	0.994	0.999	0.999	1.013	1.013
	0.060	0.067	0.116	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
	0.080	0.034	0.059	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001
	0.100	0.019	0.034	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
7	0.000	2.703	2.465	2.555	2.224	1.952	2.343	2.077	1.851	1.786	2.050
	0.005	2.304	2.253	2.145	1.878	1.679	2.098	1.846	1.662	1.680	1.875
	0.010	1.614	1.817	1.473	1.332	1.230	1.614	1.425	1.298	1.452	1.555
	0.015	1.066	1.367	1.005	0.973	0.973	1.181	1.084	1.051	1.278	1.313
	0.020	0.713	1.005	0.792	0.864	0.901	0.934	0.951	0.962	1.171	1.181
	0.025	0.493	0.742	0.776	0.884	0.922	0.878	0.938	0.957	1.108	1.112
	0.030	0.354	0.556	0.846	0.949	0.975	0.910	0.971	0.985	1.073	1.074
	0.035	0.263	0.425	0.947	0.990	0.996	0.969	0.994	0.997	1.051	1.051
	0.040	0.200	0.331	0.991	0.997	1.000	0.994	0.998	1.000	1.037	1.037
	0.060	0.082	0.142	1.000	1.000	1.000	1.000	1.000	1.000	1.012	1.012
	0.080	0.041	0.072	1.000	1.000	1.000	1.000	1.000	1.000	1.004	1.004
	0.100	0.023	0.041	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001

Table 4.4 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cobb-Douglas model with $p_1 = 5$ and $\pi = 0.75$ at $\Delta^{\mathrm{sim}} \geq 0$ (Cont.)

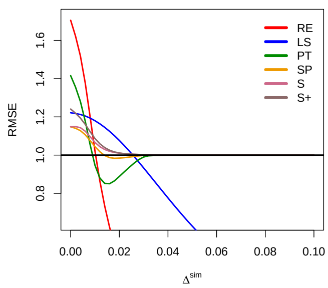
p_2	Δ^{sim}	RE	LS	PT			SP			S	S+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
11	0.000	3.888	3.287	3.483	2.824	2.429	3.003	2.529	2.228	2.411	2.869
	0.005	3.319	3.014	2.937	2.472	2.129	2.705	2.318	2.028	2.216	2.582
	0.010	2.336	2.464	2.023	1.740	1.542	2.108	1.791	1.581	1.869	2.070
	0.015	1.549	1.883	1.337	1.202	1.144	1.518	1.310	1.222	1.576	1.661
	0.020	1.040	1.405	0.969	0.968	0.971	1.129	1.061	1.036	1.382	1.408
	0.025	0.721	1.049	0.846	0.909	0.943	0.962	0.969	0.981	1.259	1.264
	0.030	0.518	0.793	0.859	0.945	0.972	0.930	0.974	0.987	1.180	1.180
	0.035	0.385	0.609	0.931	0.984	0.995	0.964	0.993	0.998	1.127	1.127
	0.040	0.293	0.476	0.982	0.998	1.001	0.991	1.000	1.001	1.092	1.092
	0.060	0.121	0.206	1.000	1.000	1.000	1.000	1.000	1.000	1.029	1.029
	0.080	0.061	0.106	1.000	1.000	1.000	1.000	1.000	1.000	1.010	1.010
	0.100	0.034	0.060	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
15	0.000	5.386	4.206	4.315	3.389	2.762	3.564	2.940	2.482	3.131	3.773
	0.005	4.590	3.879	3.730	2.848	2.394	3.275	2.605	2.238	2.862	3.367
	0.010	3.197	3.193	2.491	2.020	1.731	2.497	2.029	1.735	2.387	2.642
	0.015	2.099	2.456	1.651	1.385	1.286	1.820	1.475	1.349	1.950	2.044
	0.020	1.398	1.840	1.140	1.077	1.039	1.301	1.169	1.104	1.639	1.665
	0.025	0.965	1.377	0.930	0.941	0.951	1.053	1.007	0.997	1.438	1.443
	0.030	0.691	1.043	0.871	0.938	0.965	0.954	0.975	0.985	1.309	1.310
	0.035	0.512	0.802	0.915	0.971	0.987	0.956	0.984	0.993	1.225	1.225
	0.040	0.390	0.628	0.972	0.994	0.997	0.985	0.997	0.998	1.168	1.168
	0.060	0.160	0.272	1.000	1.000	1.000	1.000	1.000	1.000	1.063	1.063
	0.080	0.080	0.139	1.000	1.000	1.000	1.000	1.000	1.000	1.029	1.029
	0.100	0.045	0.080	1.000	1.000	1.000	1.000	1.000	1.000	1.014	1.014



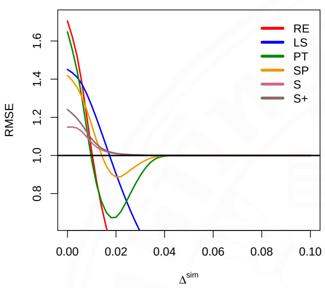
(a) $\pi = 0.25$, $\alpha = 0.01$



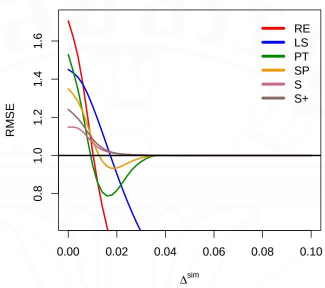
(b) $\pi = 0.25$, $\alpha = 0.05$



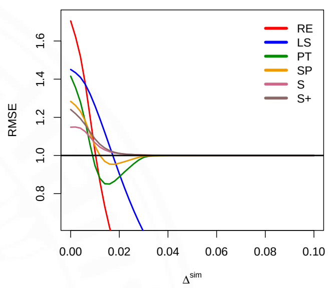
(c) $\pi = 0.25$, $\alpha = 0.10$



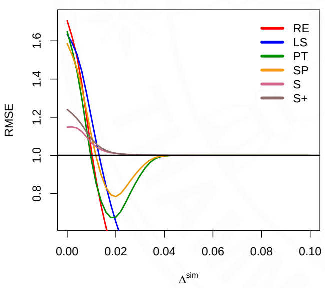
(d) $\pi = 0.50$, $\alpha = 0.01$



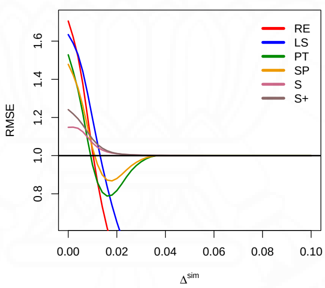
(e) $\pi = 0.50$, $\alpha = 0.05$



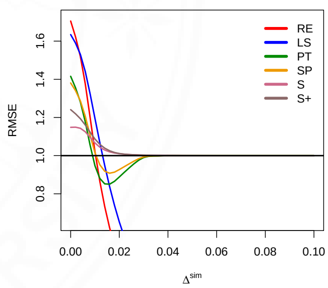
(f) $\pi = 0.50$, $\alpha = 0.10$



(g) $\pi = 0.75$, $\alpha = 0.01$

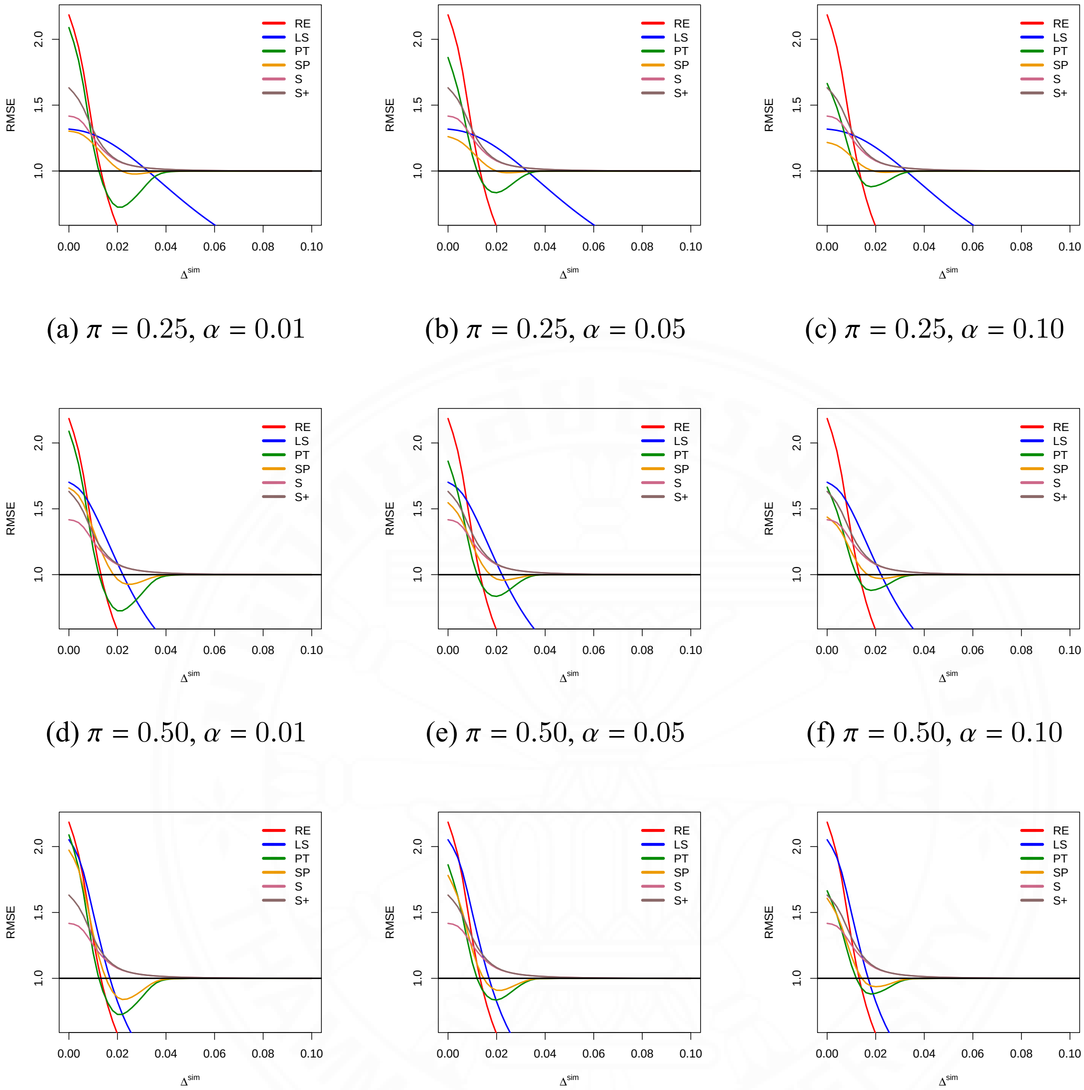


(h) $\pi = 0.75$, $\alpha = 0.05$



(i) $\pi = 0.75$, $\alpha = 0.10$

Figure 4.5 RMSEs of $\widehat{\beta}_1^{\text{RE}}$, $\widehat{\beta}_1^{\text{LS}}$, $\widehat{\beta}_1^{\text{PT}}$, $\widehat{\beta}_1^{\text{SP}}$, $\widehat{\beta}_1^{\text{S}}$, and $\widehat{\beta}_1^{\text{S}^+}$ with respect to $\widehat{\beta}_1^{\text{UE}}$ for Cobb-Douglas model with $p_1 = 5$ and $p_2 - 1 = 2$ at $\Delta^{\text{sim}} \geq 0$

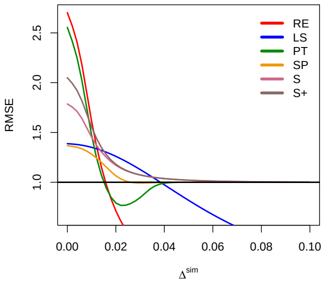


(a) $\pi = 0.25, \alpha = 0.01$ (b) $\pi = 0.25, \alpha = 0.05$ (c) $\pi = 0.25, \alpha = 0.10$

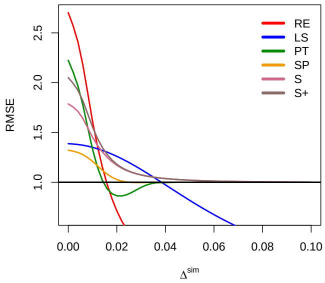
(d) $\pi = 0.50, \alpha = 0.01$ (e) $\pi = 0.50, \alpha = 0.05$ (f) $\pi = 0.50, \alpha = 0.10$

(g) $\pi = 0.75, \alpha = 0.01$ (h) $\pi = 0.75, \alpha = 0.05$ (i) $\pi = 0.75, \alpha = 0.10$

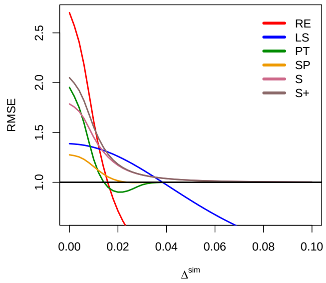
Figure 4.6 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cobb-Douglas model with $p_1 = 5$ and $p_2 - 1 = 4$ at $\Delta^{\mathrm{sim}} \geq 0$



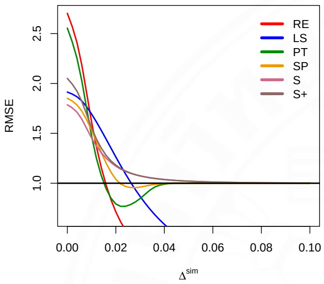
(a) $\pi = 0.25$, $\alpha = 0.01$



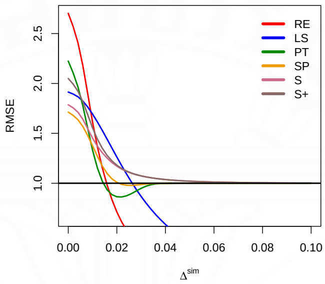
(b) $\pi = 0.25$, $\alpha = 0.05$



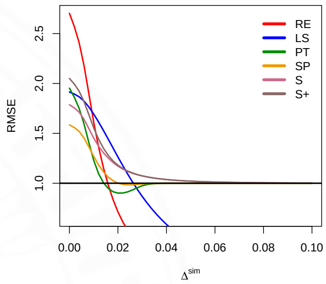
(c) $\pi = 0.25$, $\alpha = 0.10$



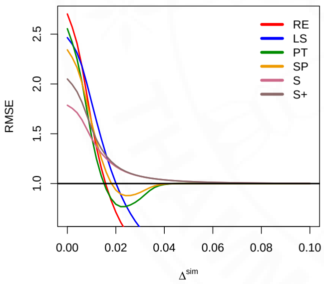
(d) $\pi = 0.50$, $\alpha = 0.01$



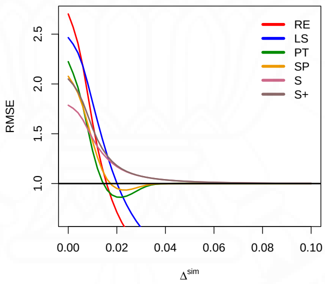
(e) $\pi = 0.50$, $\alpha = 0.05$



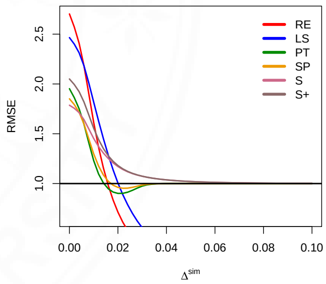
(f) $\pi = 0.50$, $\alpha = 0.10$



(g) $\pi = 0.75$, $\alpha = 0.01$

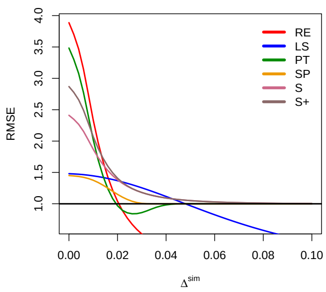


(h) $\pi = 0.75$, $\alpha = 0.05$

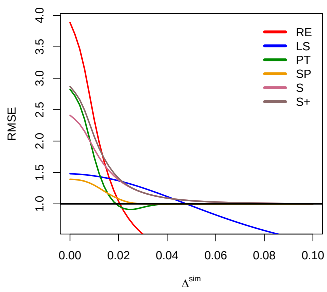


(i) $\pi = 0.75$, $\alpha = 0.10$

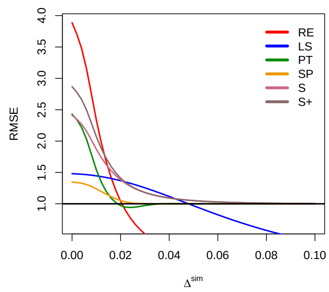
Figure 4.7 RMSEs of $\widehat{\beta}_1^{\text{RE}}$, $\widehat{\beta}_1^{\text{LS}}$, $\widehat{\beta}_1^{\text{PT}}$, $\widehat{\beta}_1^{\text{SP}}$, $\widehat{\beta}_1^{\text{S}}$, and $\widehat{\beta}_1^{\text{S}^+}$ with respect to $\widehat{\beta}_1^{\text{UE}}$ for Cobb-Douglas model with $p_1 = 5$ and $p_2 - 1 = 6$ at $\Delta^{\text{sim}} \geq 0$



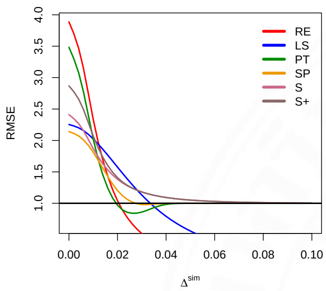
(a) $\pi = 0.25, \alpha = 0.01$



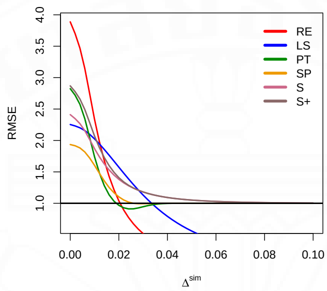
(b) $\pi = 0.25, \alpha = 0.05$



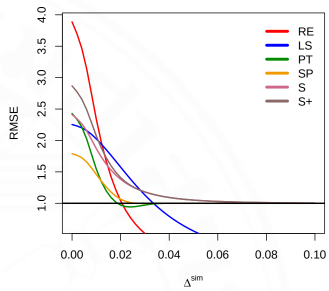
(c) $\pi = 0.25, \alpha = 0.10$



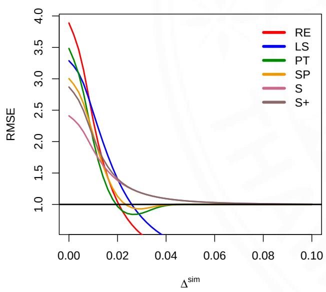
(d) $\pi = 0.50, \alpha = 0.01$



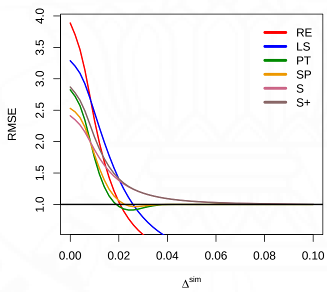
(e) $\pi = 0.50, \alpha = 0.05$



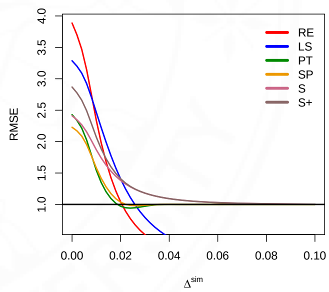
(f) $\pi = 0.50, \alpha = 0.10$



(g) $\pi = 0.75, \alpha = 0.01$

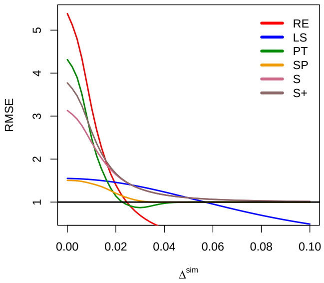


(h) $\pi = 0.75, \alpha = 0.05$

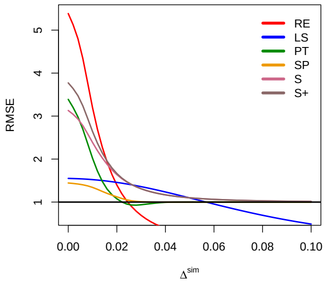


(i) $\pi = 0.75, \alpha = 0.10$

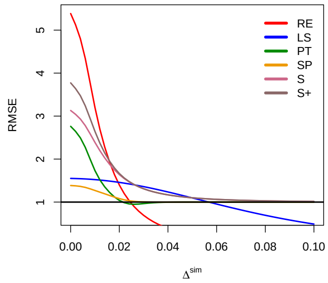
Figure 4.8 RMSEs of $\widehat{\beta}_1^{\mathrm{RE}}$, $\widehat{\beta}_1^{\mathrm{LS}}$, $\widehat{\beta}_1^{\mathrm{PT}}$, $\widehat{\beta}_1^{\mathrm{SP}}$, $\widehat{\beta}_1^{\mathrm{S}}$, and $\widehat{\beta}_1^{\mathrm{S+}}$ with respect to $\widehat{\beta}_1^{\mathrm{UE}}$ for Cobb-Douglas model with $p_1 = 5$ and $p_2 - 1 = 10$ at $\Delta^{\mathrm{sim}} \geq 0$



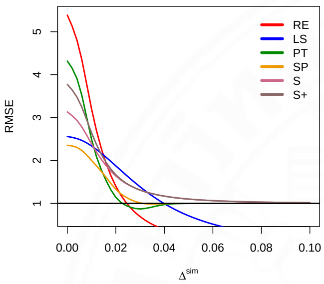
(a) $\pi = 0.25$, $\alpha = 0.01$



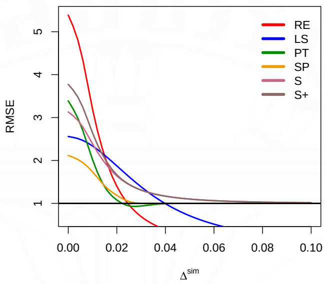
(b) $\pi = 0.25$, $\alpha = 0.05$



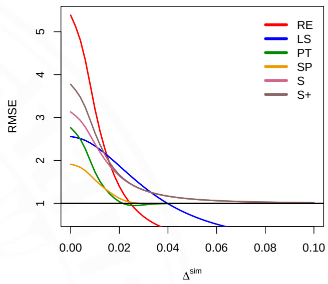
(c) $\pi = 0.25$, $\alpha = 0.10$



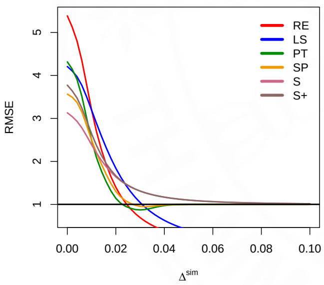
(d) $\pi = 0.50$, $\alpha = 0.01$



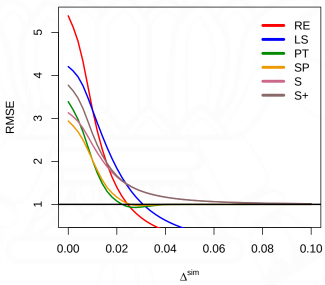
(e) $\pi = 0.50$, $\alpha = 0.05$



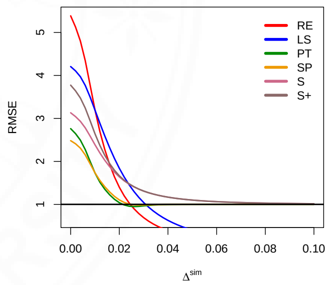
(f) $\pi = 0.50$, $\alpha = 0.10$



(g) $\pi = 0.75$, $\alpha = 0.01$



(h) $\pi = 0.75$, $\alpha = 0.05$



(i) $\pi = 0.75$, $\alpha = 0.10$

Figure 4.9 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S^+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cobb-Douglas model with $p_1 = 5$ and $p_2 - 1 = 14$ at $\Delta^{\mathrm{sim}} \geq 0$

4.2.1.2 Exponential Model

(1) Correct Subspace Information ($\Delta^{\text{sim}} = 0$)

When $\Delta^{\text{sim}} = 0$, we set the true values of $\boldsymbol{\beta}$ in the simulations at $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top = (\boldsymbol{\beta}_1^\top, \mathbf{0}_{p_2}^\top)^\top$. The vector of active parameters ($\boldsymbol{\beta}_1$) was set to 0.27, −1.25, 0.96, and 0.12. Five pairings of active and inactive parameters were chosen so that (p_1, p_2) = (4,3), (4,5), (4,7), (4,11), and (4,15). The RMSE results are shown in Table 4.5. The results can be summarized as follows:

1. The RE performed better than all other estimators as it had the maximum RMSE in all configurations. The RMSEs of all estimators increased as p_2 increased.

2. The RMSE of the LS estimator depended on the choice of π. Its performance was always superior to that of the SP estimator at the same value of π and it performed better than S, S$^+$, LASSO, and aLASSO when π was large. Its RMSE was close to 1 when π was small.

3. The RMSE of the PT estimator decreased as α increased, while that of the SP estimator decreased as π decreased and α increased. For a fixed α, the PT estimator dominated SP, S, S$^+$, LASSO, and aLASSO estimators when α was small. When p_2 was large, the SP estimator outperformed S and LASSO estimators but it was inferior to S$^+$ and aLASSO estimators when π was large and α was small.

4. When π became small, the S$^+$ estimator was more efficient than LS and SP for all p_2 and α. The S$^+$ estimator dominated the shrinkage estimator and the performance of the aLASSO estimator was also superior to the LASSO estimator.

5. These results were similar to the Cobb-Douglas model presented previously, and also consistent with the theoretical results at $\Delta^* = 0$.

(2) Uncertain Subspace Information ($\Delta^{\text{sim}} \geq 0$)

When the submodel is incorrect, $\Delta^{\text{sim}} > 0$. We set the simulation model with $\boldsymbol{\beta}_1 = (0.9, 0.9, 0.9)^\top$ and $\boldsymbol{\beta}_2 = (\beta_4, \mathbf{0}_{p_2-1}^\top)^\top$, where $\beta_4 = \Delta^{\text{sim}}$. To study the behavior of the estimators as correct information changed to incorrect information, the generated values of Δ^{sim} were between 0 and 0.6, and the number of inactive parameters was $p_2 - 1$, where p_2 = 3, 5, 7, 11, 15, and 20. The RMSEs of the proposed

Table 4.5 RMSEs of $\widehat{\beta}_1^{\text{RE}}$, $\widehat{\beta}_1^{\text{LS}}$, $\widehat{\beta}_1^{\text{PT}}$, $\widehat{\beta}_1^{\text{SP}}$, $\widehat{\beta}_1^{\text{S}}$, $\widehat{\beta}_1^{\text{S+}}$, $\widehat{\beta}_1^{\text{LASSO}}$, and $\widehat{\beta}_1^{\text{aLASSO}}$ with respect to $\widehat{\beta}_1^{\text{UE}}$ for exponential model with $p_1 = 4$ at $\Delta^{\text{sim}} = 0$

Estimator			Number of Inactive Parameters (p_2)				
			3	5	7	11	15
RE			1.368	1.615	1.947	2.637	3.778
LS	$\pi = 0.25$		1.131	1.201	1.267	1.374	1.463
	$\pi = 0.50$		1.249	1.402	1.567	1.873	2.197
	$\pi = 0.75$		1.334	1.557	1.831	2.395	3.168
PT		$\alpha = 0.01$	1.337	1.566	1.852	2.425	3.340
		$\alpha = 0.05$	1.290	1.455	1.703	2.130	2.848
		$\alpha = 0.10$	1.227	1.381	1.584	1.909	2.437
SP	$\pi = 0.25$	$\alpha = 0.01$	1.121	1.189	1.250	1.347	1.432
		$\alpha = 0.05$	1.106	1.160	1.219	1.304	1.386
		$\alpha = 0.10$	1.087	1.139	1.191	1.264	1.340
	$\pi = 0.50$	$\alpha = 0.01$	1.229	1.374	1.523	1.790	2.078
		$\alpha = 0.05$	1.198	1.310	1.445	1.664	1.921
		$\alpha = 0.10$	1.160	1.263	1.379	1.558	1.775
	$\pi = 0.75$	$\alpha = 0.01$	1.306	1.514	1.755	2.229	2.872
		$\alpha = 0.05$	1.263	1.418	1.628	1.993	2.519
		$\alpha = 0.10$	1.209	1.351	1.525	1.809	2.215
S			1.086	1.286	1.531	1.992	2.713
S+			1.138	1.370	1.633	2.175	2.983
LASSO			1.096	1.161	1.225	1.391	1.681
aLASSO			1.176	1.336	1.532	1.999	3.126

estimators for each p_2 and π are shown in Tables 4.6 to 4.8. The graphical representations for each p_2 are shown as Figures 4.10 to 4.14. The main findings are as follows:

1. We observed that no single estimator gave the best performance for parameter estimation in all simulations. The RE performed better than all other estimators in an area near the null hypothesis, but as the hypothesis error grew, or Δ^{sim} moved away from zero, its RMSE decreased sharply to zero.

2. Similarly, the LS estimator also decreased to zero, but it was higher than the RE when Δ^{sim} was far from zero. The LS estimator also outperformed other estimators in some areas of Δ^{sim}.

3. The PT estimator performed well at $\Delta^{\text{sim}} = 0$. However, the PT estimator had an estimator error smaller than the SP estimator in a small space of $\Delta^{\text{sim}} > 0$.

4. The RMSE of PT and SP estimators first fell gradually, becoming worse than the UE, then increased to 1. However, both the PT and SP estimators also outperformed all other estimators in some parts of the parameter space Δ^{sim}.

5. The performance of the S^+ estimator was superior to that of the shrinkage estimator and both of them performed better than the other estimators in a large area of Δ^{sim}. The overall results of these simulation studies were in line with the theoretical results.

Table 4.6 RMSEs of $\widehat{\beta}_1^{\text{RE}}$, $\widehat{\beta}_1^{\text{LS}}$, $\widehat{\beta}_1^{\text{PT}}$, $\widehat{\beta}_1^{\text{SP}}$, $\widehat{\beta}_1^{\text{S}}$, and $\widehat{\beta}_1^{\text{S+}}$ with respect to $\widehat{\beta}_1^{\text{UE}}$ for exponential model with $p_1 = 3$ and $\pi = 0.25$ at $\Delta^{\text{sim}} \geq 0$

				PT			SP				
p_2	Δ^{sim}	RE	LS	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	S	S^+
3	0.00	1.234	1.092	1.221	1.191	1.156	1.087	1.076	1.063	1.061	1.096
	0.04	1.117	1.087	1.091	1.064	1.049	1.070	1.050	1.039	1.046	1.063
	0.08	0.859	1.069	0.915	0.933	0.947	1.025	1.011	1.006	1.021	1.027
	0.12	0.614	1.038	0.855	0.920	0.953	0.996	0.996	0.998	1.010	1.011
	0.16	0.433	0.995	0.892	0.955	0.976	0.990	0.996	0.997	1.005	1.005
	0.20	0.309	0.942	0.947	0.977	0.987	0.995	0.997	0.998	1.003	1.003
	0.25	0.208	0.863	0.976	0.992	0.996	0.997	0.999	0.999	1.002	1.002
	0.30	0.116	0.714	0.991	0.999	0.999	0.999	1.000	1.000	1.001	1.001
	0.35	0.072	0.578	0.998	1.000	1.000	1.000	1.000	1.000	1.001	1.001
	0.40	0.049	0.469	0.999	1.000	1.000	1.000	1.000	1.000	1.001	1.001
	0.50	0.008	0.113	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.60	0.003	0.050	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.00	1.430	1.152	1.401	1.338	1.275	1.143	1.125	1.105	1.216	1.269
	0.04	1.288	1.144	1.252	1.190	1.158	1.126	1.098	1.081	1.170	1.205
	0.08	0.993	1.124	1.016	1.010	1.014	1.077	1.046	1.034	1.104	1.116
	0.12	0.713	1.092	0.912	0.945	0.971	1.028	1.012	1.009	1.063	1.065
	0.16	0.506	1.050	0.904	0.951	0.977	1.004	1.000	1.001	1.038	1.039
	0.20	0.362	0.998	0.931	0.974	0.986	0.997	0.998	0.999	1.025	1.025
	0.25	0.244	0.921	0.975	0.993	0.998	0.998	0.999	1.000	1.017	1.016
	0.30	0.131	0.762	0.993	0.999	0.999	0.999	1.000	1.000	1.011	1.011
	0.35	0.085	0.635	0.999	1.000	1.000	1.000	1.000	1.000	1.008	1.008
	0.40	0.050	0.475	1.000	1.000	1.000	1.000	1.000	1.000	1.006	1.006
	0.50	0.011	0.155	1.000	1.000	1.000	1.000	1.000	1.000	1.004	1.004
	0.60	0.005	0.069	1.000	1.000	1.000	1.000	1.000	1.000	1.002	1.002

Table: 4.6 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for exponential model with $p_1 = 3$ and $\pi = 0.25$ at $\Delta^{\mathrm{sim}} \geq 0$ (Cont.)

				PT			SP				
p_2	Δ^{sim}	RE	LS	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	S	S+
7	0.00	1.580	1.194	1.530	1.442	1.365	1.181	1.157	1.135	1.350	1.411
	0.04	1.412	1.186	1.346	1.279	1.215	1.161	1.133	1.107	1.270	1.320
	0.08	1.087	1.166	1.067	1.048	1.032	1.107	1.073	1.052	1.163	1.190
	0.12	0.782	1.136	0.920	0.958	0.963	1.046	1.026	1.015	1.102	1.110
	0.16	0.556	1.095	0.903	0.948	0.963	1.014	1.005	1.002	1.068	1.070
	0.20	0.398	1.044	0.925	0.972	0.976	1.002	1.000	0.998	1.048	1.048
	0.25	0.268	0.966	0.962	0.989	0.992	0.997	1.000	0.999	1.032	1.032
	0.30	0.179	0.868	0.990	0.998	0.998	1.000	1.000	1.000	1.023	1.023
	0.35	0.125	0.768	0.996	0.999	0.999	1.000	1.000	1.000	1.017	1.017
	0.40	0.088	0.661	0.999	1.000	1.000	1.000	1.000	1.000	1.012	1.012
	0.50	0.020	0.250	1.000	1.000	1.000	1.000	1.000	1.000	1.005	1.005
	0.60	0.015	0.197	1.000	1.000	1.000	1.000	1.000	1.000	1.002	1.002
11	0.00	2.056	1.287	1.957	1.800	1.677	1.269	1.239	1.210	1.719	1.819
	0.04	1.877	1.285	1.763	1.603	1.481	1.256	1.214	1.180	1.617	1.699
	0.08	1.456	1.272	1.345	1.246	1.201	1.196	1.140	1.111	1.446	1.482
	0.12	1.045	1.248	1.080	1.073	1.054	1.113	1.075	1.053	1.311	1.322
	0.16	0.737	1.214	0.996	0.994	0.997	1.060	1.030	1.019	1.220	1.225
	0.20	0.525	1.170	0.960	0.983	0.993	1.023	1.011	1.007	1.160	1.162
	0.25	0.352	1.100	0.974	0.991	0.994	1.007	1.003	1.001	1.113	1.113
	0.30	0.227	0.998	0.982	0.995	0.996	1.001	1.000	1.000	1.082	1.082
	0.35	0.156	0.894	0.992	0.997	0.999	1.000	1.000	1.000	1.063	1.063
	0.40	0.098	0.741	0.997	0.998	0.999	1.000	1.000	1.000	1.049	1.049
	0.50	0.011	0.153	1.000	1.000	1.000	1.000	1.000	1.000	1.024	1.024
	0.60	0.004	0.065	1.000	1.000	1.000	1.000	1.000	1.000	1.008	1.008
15	0.00	2.576	1.361	2.355	2.057	1.871	1.334	1.288	1.255	2.088	2.222
	0.04	2.342	1.358	2.086	1.829	1.689	1.316	1.264	1.231	1.960	2.059
	0.08	1.811	1.345	1.588	1.408	1.315	1.259	1.193	1.152	1.702	1.755
	0.12	1.296	1.322	1.216	1.144	1.097	1.171	1.107	1.074	1.491	1.516
	0.16	0.911	1.289	1.054	1.030	1.024	1.091	1.050	1.035	1.352	1.361
	0.20	0.626	1.240	0.985	0.991	0.992	1.042	1.022	1.013	1.262	1.264
	0.25	0.411	1.167	0.972	0.985	0.996	1.016	1.006	1.004	1.188	1.188
	0.30	0.279	1.080	0.986	0.994	0.999	1.006	1.001	1.001	1.140	1.140
	0.35	0.120	0.822	0.992	0.997	0.998	1.001	1.000	1.000	1.106	1.106
	0.40	0.084	0.698	0.997	0.999	0.999	1.000	1.000	1.000	1.083	1.083
	0.50	0.036	0.414	1.000	1.000	1.000	1.000	1.000	1.000	1.036	1.036
	0.60	0.018	0.237	1.000	1.000	1.000	1.000	1.000	1.000	1.019	1.019

Table 4.7 RMSEs of $\widehat{\beta}_1^{\mathrm{RE}}$, $\widehat{\beta}_1^{\mathrm{LS}}$, $\widehat{\beta}_1^{\mathrm{PT}}$, $\widehat{\beta}_1^{\mathrm{SP}}$, $\widehat{\beta}_1^{\mathrm{S}}$, and $\widehat{\beta}_1^{\mathrm{S+}}$ with respect to $\widehat{\beta}_1^{\mathrm{UE}}$ for exponential model with $p_1 = 3$ and $\pi = 0.50$ at $\Delta^{\mathrm{sim}} \geq 0$

p_2	Δ^{sim}	RE	LS	PT			SP			S	S^+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
3	0.00	1.234	1.168	1.221	1.191	1.156	1.159	1.137	1.114	1.061	1.096
	0.04	1.117	1.142	1.091	1.064	1.049	1.112	1.080	1.061	1.046	1.063
	0.08	0.859	1.062	0.915	0.933	0.947	1.017	1.002	0.999	1.021	1.027
	0.12	0.614	0.947	0.855	0.920	0.953	0.967	0.980	0.989	1.010	1.011
	0.16	0.433	0.817	0.892	0.955	0.976	0.968	0.987	0.993	1.005	1.005
	0.20	0.309	0.687	0.947	0.977	0.987	0.984	0.992	0.996	1.003	1.003
	0.25	0.208	0.541	0.976	0.992	0.996	0.992	0.997	0.998	1.002	1.002
	0.30	0.116	0.357	0.991	0.999	0.999	0.997	1.000	1.000	1.001	1.001
	0.35	0.072	0.243	0.998	1.000	1.000	0.999	1.000	1.000	1.001	1.001
	0.40	0.049	0.175	0.999	1.000	1.000	1.000	1.000	1.000	1.001	1.001
	0.50	0.008	0.031	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.60	0.003	0.013	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.00	1.430	1.291	1.401	1.338	1.275	1.274	1.235	1.194	1.216	1.269
	0.04	1.288	1.257	1.252	1.190	1.158	1.224	1.170	1.139	1.170	1.205
	0.08	0.993	1.170	1.016	1.010	1.014	1.109	1.064	1.048	1.104	1.116
	0.12	0.713	1.048	0.912	0.945	0.971	1.021	1.006	1.008	1.063	1.065
	0.16	0.506	0.910	0.904	0.951	0.977	0.988	0.992	0.998	1.038	1.039
	0.20	0.362	0.772	0.931	0.974	0.986	0.984	0.993	0.996	1.025	1.025
	0.25	0.244	0.614	0.975	0.993	0.998	0.992	0.998	1.000	1.017	1.016
	0.30	0.131	0.397	0.993	0.999	0.999	0.998	1.000	1.000	1.011	1.011
	0.35	0.085	0.281	0.999	1.000	1.000	1.000	1.000	1.000	1.008	1.008
	0.40	0.050	0.176	1.000	1.000	1.000	1.000	1.000	1.000	1.006	1.006
	0.50	0.011	0.044	1.000	1.000	1.000	1.000	1.000	1.000	1.004	1.004
	0.60	0.005	0.018	1.000	1.000	1.000	1.000	1.000	1.000	1.002	1.002
7	0.00	1.580	1.384	1.530	1.442	1.365	1.355	1.302	1.255	1.350	1.411
	0.04	1.412	1.346	1.346	1.279	1.215	1.294	1.240	1.189	1.270	1.320
	0.08	1.087	1.255	1.067	1.048	1.032	1.163	1.109	1.077	1.163	1.190
	0.12	0.782	1.127	0.920	0.958	0.963	1.046	1.027	1.014	1.102	1.110
	0.16	0.556	0.983	0.903	0.948	0.963	1.001	0.998	0.997	1.068	1.070
	0.20	0.398	0.836	0.925	0.972	0.976	0.989	0.995	0.993	1.048	1.048
	0.25	0.268	0.666	0.962	0.989	0.992	0.989	0.997	0.998	1.032	1.032
	0.30	0.179	0.508	0.990	0.998	0.998	0.998	1.000	0.999	1.023	1.023
	0.35	0.125	0.390	0.996	0.999	0.999	0.999	1.000	1.000	1.017	1.017
	0.40	0.088	0.293	0.999	1.000	1.000	1.000	1.000	1.000	1.012	1.012
	0.50	0.020	0.075	1.000	1.000	1.000	1.000	1.000	1.000	1.005	1.005
	0.60	0.015	0.057	1.000	1.000	1.000	1.000	1.000	1.000	1.002	1.002

Table 4.7 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for exponential model with $p_1 = 3$ and $\pi = 0.50$ at $\Delta^{\mathrm{sim}} \geq 0$ (Cont.)

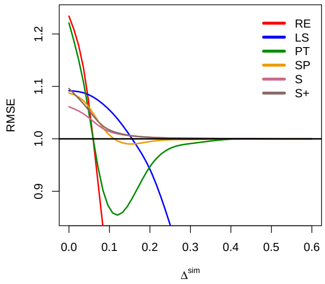
p_2	Δ^{sim}	RE	LS	PT			SP			S	S^{+}
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
11	0.00	2.056	1.621	1.957	1.800	1.677	1.573	1.495	1.427	1.719	1.819
	0.04	1.877	1.596	1.763	1.603	1.481	1.524	1.425	1.347	1.617	1.699
	0.08	1.456	1.508	1.345	1.246	1.201	1.354	1.245	1.192	1.446	1.482
	0.12	1.045	1.374	1.080	1.073	1.054	1.174	1.117	1.081	1.311	1.322
	0.16	0.737	1.213	0.996	0.994	0.997	1.081	1.039	1.026	1.220	1.225
	0.20	0.525	1.044	0.960	0.983	0.993	1.023	1.012	1.009	1.160	1.162
	0.25	0.352	0.844	0.974	0.991	0.994	1.005	1.002	1.001	1.113	1.113
	0.30	0.227	0.637	0.982	0.995	0.996	0.999	0.999	0.999	1.082	1.082
	0.35	0.156	0.486	0.992	0.997	0.999	0.999	0.999	1.000	1.063	1.063
	0.40	0.098	0.333	0.997	0.998	0.999	1.000	1.000	1.000	1.049	1.049
	0.50	0.011	0.042	1.000	1.000	1.000	1.000	1.000	1.000	1.024	1.024
	0.60	0.004	0.017	1.000	1.000	1.000	1.000	1.000	1.000	1.008	1.008
15	0.00	2.576	1.838	2.355	2.057	1.871	1.753	1.624	1.534	2.088	2.222
	0.04	2.342	1.808	2.086	1.829	1.689	1.687	1.549	1.467	1.960	2.059
	0.08	1.811	1.714	1.588	1.408	1.315	1.507	1.359	1.275	1.702	1.755
	0.12	1.296	1.569	1.216	1.144	1.097	1.289	1.176	1.120	1.491	1.516
	0.16	0.911	1.394	1.054	1.030	1.024	1.136	1.073	1.051	1.352	1.361
	0.20	0.626	1.189	0.985	0.991	0.992	1.055	1.028	1.017	1.262	1.264
	0.25	0.411	0.954	0.972	0.985	0.996	1.017	1.006	1.005	1.188	1.188
	0.30	0.279	0.750	0.986	0.994	0.999	1.006	1.001	1.001	1.140	1.140
	0.35	0.120	0.398	0.992	0.997	0.998	1.001	0.999	1.000	1.106	1.106
	0.40	0.084	0.295	0.997	0.999	0.999	0.999	1.000	1.000	1.083	1.083
	0.50	0.036	0.137	1.000	1.000	1.000	1.000	1.000	1.000	1.036	1.036
	0.60	0.018	0.068	1.000	1.000	1.000	1.000	1.000	1.000	1.019	1.019

Table 4.8 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for exponential model with $p_1 = 3$ and $\pi = 0.75$ at $\Delta^{\mathrm{sim}} \geq 0$

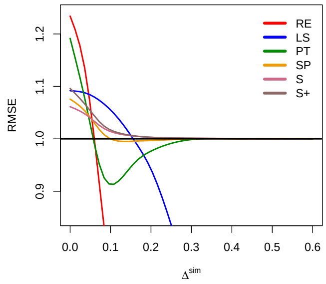
				PT			SP				
p_2	Δ^{sim}	RE	LS	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	S	S^+
3	0.00	1.234	1.218	1.221	1.191	1.156	1.206	1.177	1.146	1.061	1.096
	0.04	1.117	1.152	1.091	1.064	1.049	1.120	1.084	1.065	1.046	1.063
	0.08	0.859	0.983	0.915	0.933	0.947	0.978	0.975	0.979	1.021	1.027
	0.12	0.614	0.783	0.855	0.920	0.953	0.918	0.954	0.974	1.010	1.011
	0.16	0.433	0.603	0.892	0.955	0.976	0.934	0.973	0.985	1.005	1.005
	0.20	0.309	0.459	0.947	0.977	0.987	0.968	0.985	0.992	1.003	1.003
	0.25	0.208	0.326	0.976	0.992	0.996	0.985	0.995	0.997	1.002	1.002
	0.30	0.116	0.192	0.991	0.999	0.999	0.994	0.999	1.000	1.001	1.001
	0.35	0.072	0.123	0.998	1.000	1.000	0.999	1.000	1.000	1.001	1.001
	0.40	0.049	0.085	0.999	1.000	1.000	1.000	1.000	1.000	1.001	1.001
	0.50	0.008	0.014	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.60	0.003	0.006	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.00	1.430	1.392	1.401	1.338	1.275	1.367	1.311	1.254	1.216	1.269
	0.04	1.288	1.311	1.252	1.190	1.158	1.270	1.203	1.166	1.170	1.205
	0.08	0.993	1.119	1.016	1.010	1.014	1.086	1.052	1.041	1.104	1.116
	0.12	0.713	0.896	0.912	0.945	0.971	0.979	0.983	0.995	1.063	1.065
	0.16	0.506	0.694	0.904	0.951	0.977	0.954	0.975	0.989	1.038	1.039
	0.20	0.362	0.531	0.931	0.974	0.986	0.962	0.985	0.992	1.025	1.025
	0.25	0.244	0.380	0.975	0.993	0.998	0.985	0.996	0.999	1.017	1.016
	0.30	0.131	0.216	0.993	0.999	0.999	0.996	1.000	0.999	1.011	1.011
	0.35	0.085	0.143	0.999	1.000	1.000	0.999	1.000	1.000	1.008	1.008
	0.40	0.050	0.085	1.000	1.000	1.000	1.000	1.000	1.000	1.006	1.006
	0.50	0.011	0.020	1.000	1.000	1.000	1.000	1.000	1.000	1.004	1.004
	0.60	0.005	0.008	1.000	1.000	1.000	1.000	1.000	1.000	1.002	1.002
7	0.00	1.580	1.529	1.530	1.442	1.365	1.484	1.407	1.338	1.350	1.411
	0.04	1.412	1.433	1.346	1.279	1.215	1.364	1.293	1.227	1.270	1.320
	0.08	1.087	1.223	1.067	1.048	1.032	1.148	1.100	1.070	1.163	1.190
	0.12	0.782	0.981	0.920	0.958	0.963	1.000	1.004	0.996	1.102	1.110
	0.16	0.556	0.762	0.903	0.948	0.963	0.962	0.978	0.983	1.068	1.070
	0.20	0.398	0.584	0.925	0.972	0.976	0.963	0.986	0.986	1.048	1.048
	0.25	0.268	0.416	0.962	0.989	0.992	0.977	0.994	0.996	1.032	1.032
	0.30	0.179	0.290	0.990	0.998	0.998	0.995	0.999	0.999	1.023	1.023
	0.35	0.125	0.209	0.996	0.999	0.999	0.998	1.000	1.000	1.017	1.017
	0.40	0.088	0.149	0.999	1.000	1.000	1.000	1.000	1.000	1.012	1.012
	0.50	0.020	0.035	1.000	1.000	1.000	1.000	1.000	1.000	1.005	1.005
	0.60	0.015	0.026	1.000	1.000	1.000	1.000	1.000	1.000	1.002	1.002

Table 4.8 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for exponential model with $p_1 = 3$ and $\pi = 0.75$ at $\Delta^{\mathrm{sim}} \geq 0$ (Cont.)

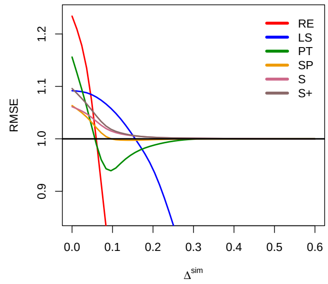
				PT			SP				
p_2	Δ^{sim}	RE	LS	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	S	S+
11	0.00	2.056	1.923	1.957	1.800	1.677	1.840	1.710	1.602	1.719	1.819
	0.04	1.877	1.835	1.763	1.603	1.481	1.725	1.574	1.460	1.617	1.699
	0.08	1.456	1.587	1.345	1.246	1.201	1.413	1.285	1.225	1.446	1.482
	0.12	1.045	1.282	1.080	1.073	1.054	1.161	1.116	1.082	1.311	1.322
	0.16	0.737	0.998	0.996	0.994	0.997	1.058	1.027	1.018	1.220	1.225
	0.20	0.525	0.765	0.960	0.983	0.993	1.001	1.003	1.004	1.160	1.162
	0.25	0.352	0.547	0.974	0.991	0.994	0.993	0.998	0.998	1.113	1.113
	0.30	0.227	0.369	0.982	0.995	0.996	0.992	0.997	0.998	1.082	1.082
	0.35	0.156	0.262	0.992	0.997	0.999	0.996	0.998	0.999	1.063	1.063
	0.40	0.098	0.168	0.997	0.998	0.999	0.999	0.999	0.999	1.049	1.049
	0.50	0.011	0.019	1.000	1.000	1.000	1.000	1.000	1.000	1.024	1.024
	0.60	0.004	0.008	1.000	1.000	1.000	1.000	1.000	1.000	1.008	1.008
15	0.00	2.576	2.333	2.355	2.057	1.871	2.164	1.926	1.772	2.088	2.222
	0.04	2.342	2.222	2.086	1.829	1.689	2.002	1.772	1.644	1.960	2.059
	0.08	1.811	1.924	1.588	1.408	1.315	1.642	1.443	1.336	1.702	1.755
	0.12	1.296	1.557	1.216	1.144	1.097	1.307	1.190	1.128	1.491	1.516
	0.16	0.911	1.212	1.054	1.030	1.024	1.123	1.066	1.048	1.352	1.361
	0.20	0.626	0.905	0.985	0.991	0.992	1.034	1.017	1.009	1.262	1.264
	0.25	0.411	0.635	0.972	0.985	0.996	1.002	0.999	1.002	1.188	1.188
	0.30	0.279	0.451	0.986	0.994	0.999	0.999	0.998	1.000	1.140	1.140
	0.35	0.120	0.205	0.992	0.997	0.998	0.997	0.998	0.999	1.106	1.106
	0.40	0.084	0.146	0.997	0.999	0.999	0.999	1.000	1.000	1.083	1.083
	0.50	0.036	0.064	1.000	1.000	1.000	1.000	1.000	1.000	1.036	1.036
	0.60	0.018	0.031	1.000	1.000	1.000	1.000	1.000	1.000	1.019	1.019



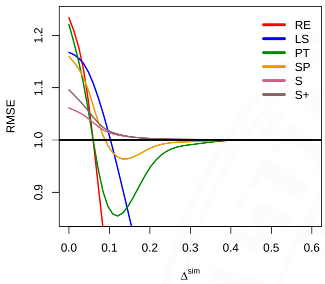
(a) $\pi = 0.25$, $\alpha = 0.01$



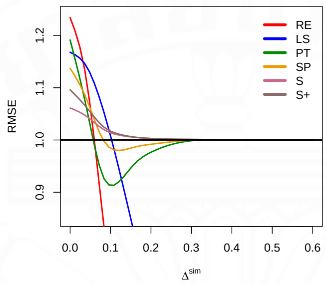
(b) $\pi = 0.25$, $\alpha = 0.05$



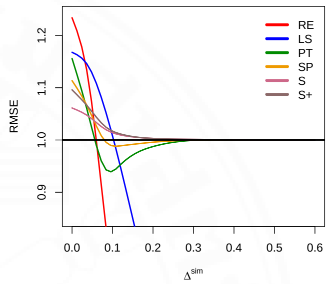
(c) $\pi = 0.25$, $\alpha = 0.10$



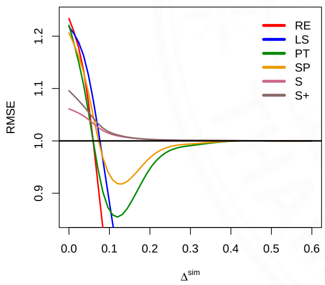
(d) $\pi = 0.50$, $\alpha = 0.01$



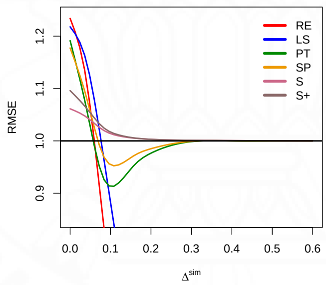
(e) $\pi = 0.50$, $\alpha = 0.05$



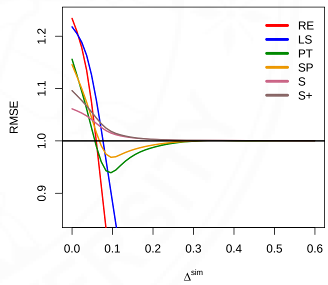
(f) $\pi = 0.50$, $\alpha = 0.10$



(g) $\pi = 0.75$, $\alpha = 0.01$

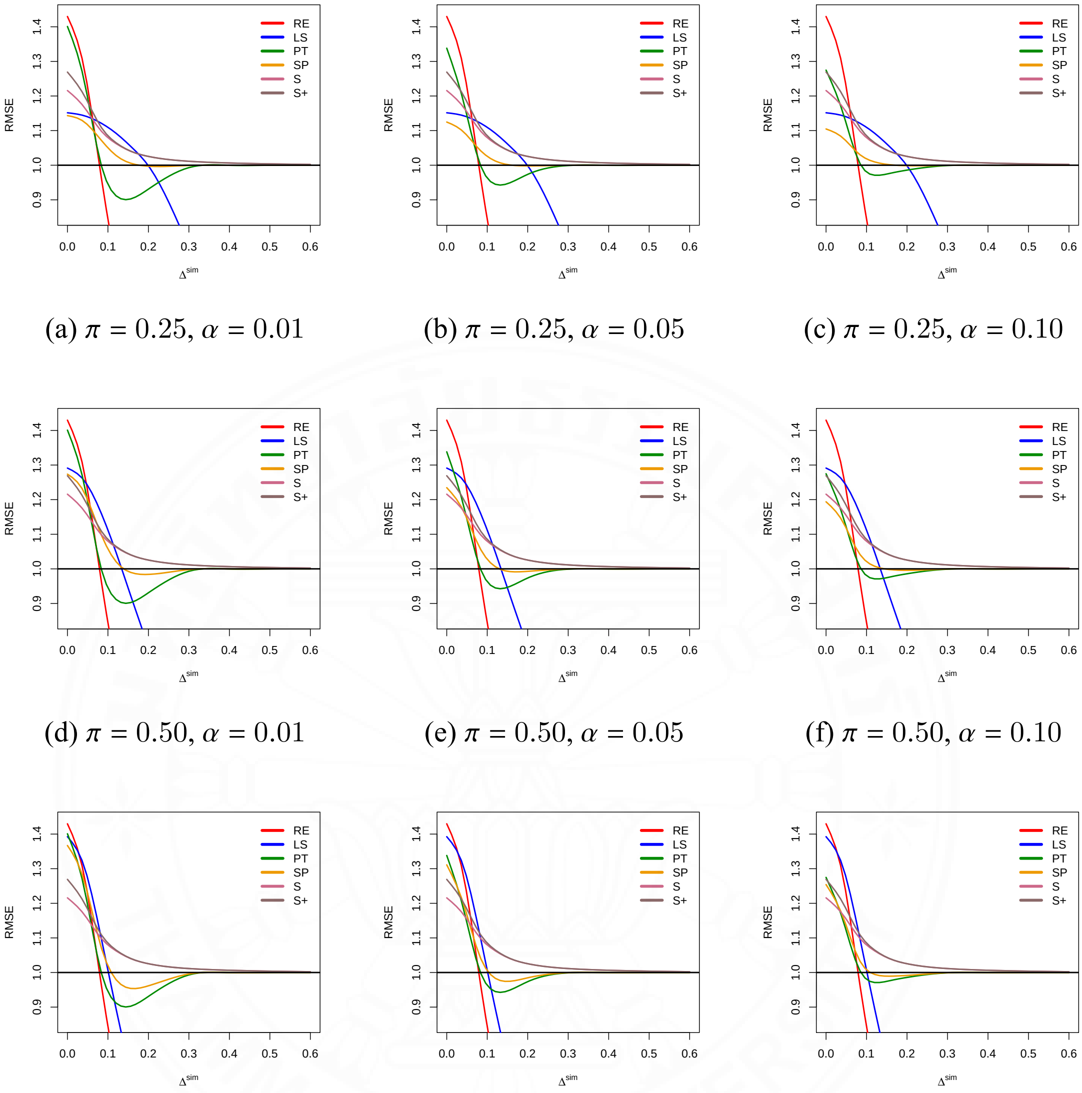


(h) $\pi = 0.75$, $\alpha = 0.05$



(i) $\pi = 0.75$, $\alpha = 0.10$

Figure 4.10 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S^+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for exponential model with $p_1 = 3$ and $p_2 - 1 = 2$ at $\Delta^{\mathrm{sim}} \geq 0$

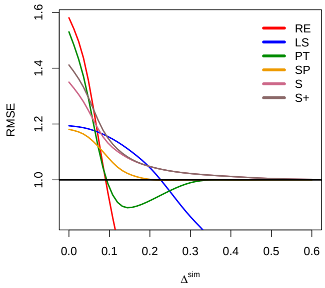


(a) $\pi = 0.25, \alpha = 0.01$ (b) $\pi = 0.25, \alpha = 0.05$ (c) $\pi = 0.25, \alpha = 0.10$

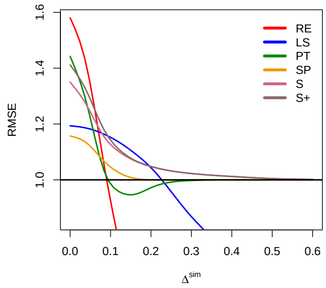
(d) $\pi = 0.50, \alpha = 0.01$ (e) $\pi = 0.50, \alpha = 0.05$ (f) $\pi = 0.50, \alpha = 0.10$

(g) $\pi = 0.75, \alpha = 0.01$ (h) $\pi = 0.75, \alpha = 0.05$ (i) $\pi = 0.75, \alpha = 0.10$

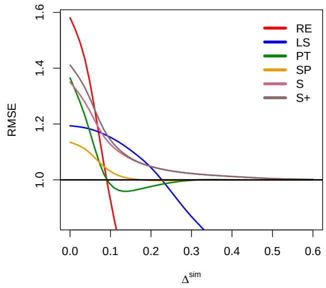
Figure 4.11 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for exponential model with $p_1 = 3$ and $p_2 - 1 = 4$ at $\Delta^{\text{sim}} \geq 0$



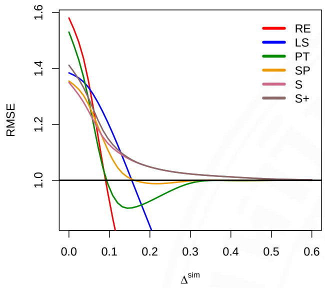
(a) $\pi = 0.25$, $\alpha = 0.01$



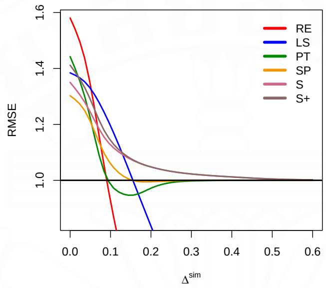
(b) $\pi = 0.25$, $\alpha = 0.05$



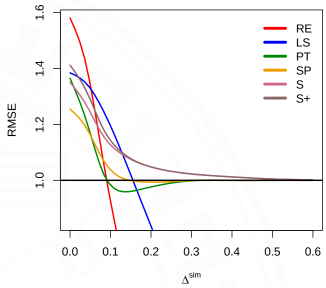
(c) $\pi = 0.25$, $\alpha = 0.10$



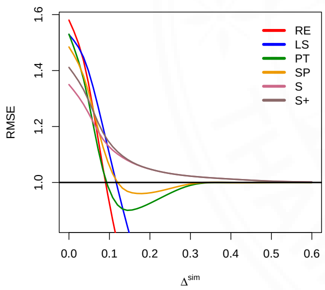
(d) $\pi = 0.50$, $\alpha = 0.01$



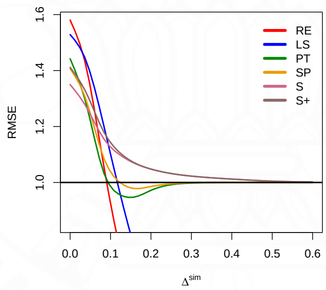
(e) $\pi = 0.50$, $\alpha = 0.05$



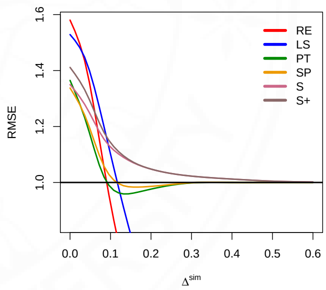
(f) $\pi = 0.50$, $\alpha = 0.10$



(g) $\pi = 0.75$, $\alpha = 0.01$

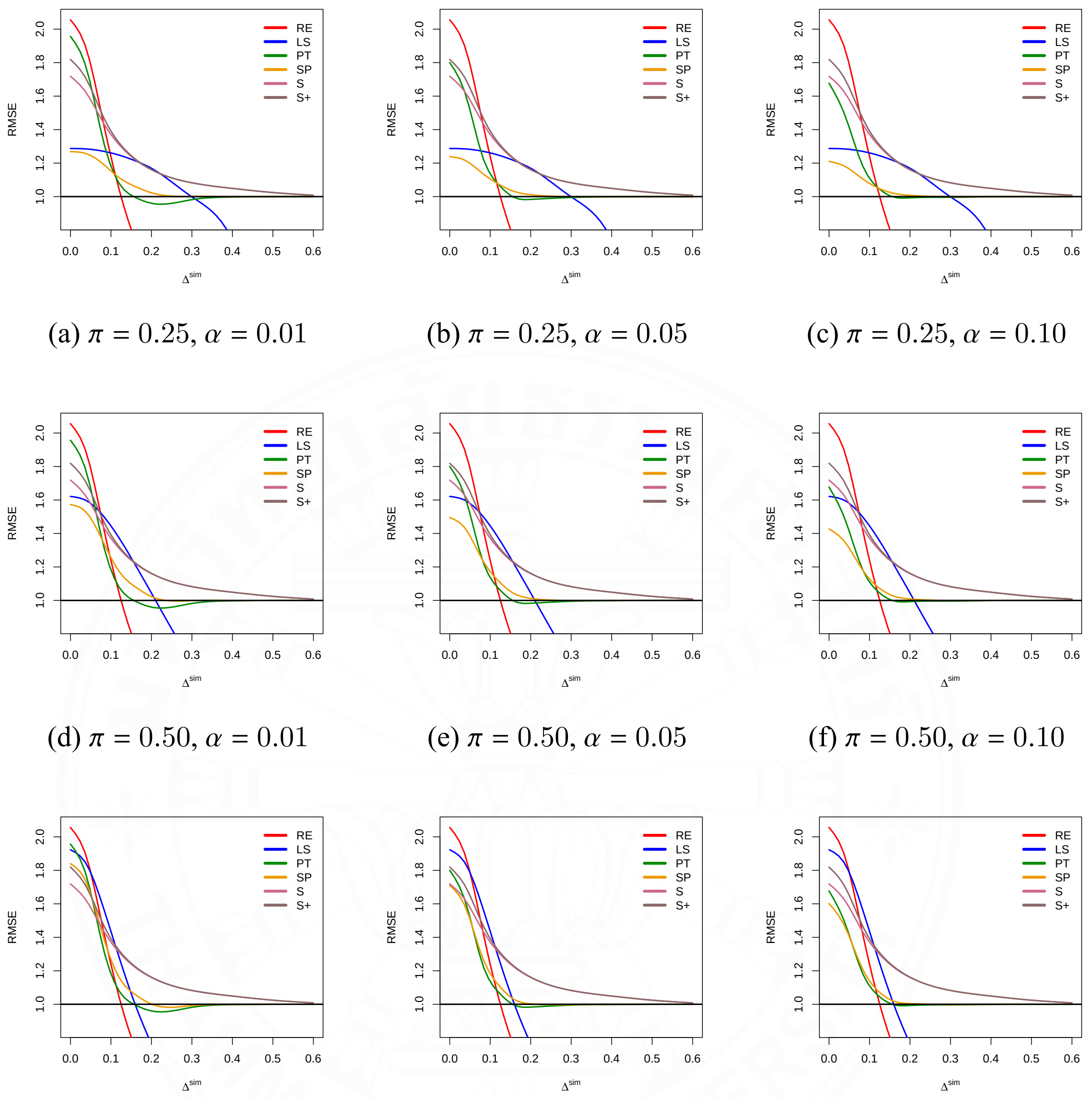


(h) $\pi = 0.75$, $\alpha = 0.05$



(i) $\pi = 0.75$, $\alpha = 0.10$

Figure 4.12 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for exponential model with $p_1 = 3$ and $p_2 - 1 = 6$ at $\Delta^{\mathrm{sim}} \geq 0$

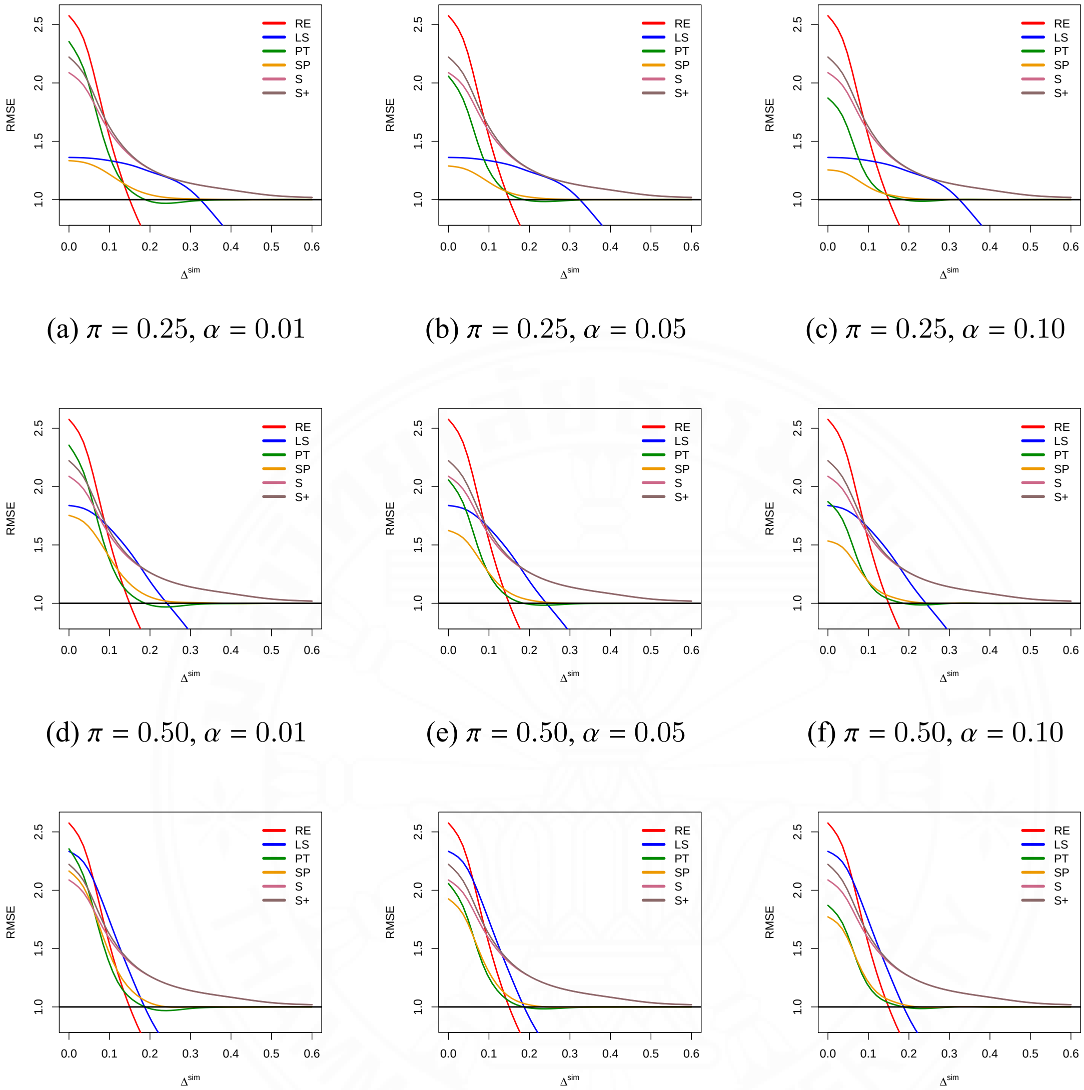


(a) $\pi = 0.25, \alpha = 0.01$ (b) $\pi = 0.25, \alpha = 0.05$ (c) $\pi = 0.25, \alpha = 0.10$

(d) $\pi = 0.50, \alpha = 0.01$ (e) $\pi = 0.50, \alpha = 0.05$ (f) $\pi = 0.50, \alpha = 0.10$

(g) $\pi = 0.75, \alpha = 0.01$ (h) $\pi = 0.75, \alpha = 0.05$ (i) $\pi = 0.75, \alpha = 0.10$

Figure 4.13 RMSEs of $\widehat{\beta}_1^{\text{RE}}$, $\widehat{\beta}_1^{\text{LS}}$, $\widehat{\beta}_1^{\text{PT}}$, $\widehat{\beta}_1^{\text{SP}}$, $\widehat{\beta}_1^{\text{S}}$, and $\widehat{\beta}_1^{\text{S+}}$ with respect to $\widehat{\beta}_1^{\text{UE}}$ for exponential model with $p_1 = 3$ and $p_2 - 1 = 10$ at $\Delta^{\text{sim}} \geq 0$



(a) $\pi = 0.25$, $\alpha = 0.01$ (b) $\pi = 0.25$, $\alpha = 0.05$ (c) $\pi = 0.25$, $\alpha = 0.10$

(d) $\pi = 0.50$, $\alpha = 0.01$ (e) $\pi = 0.50$, $\alpha = 0.05$ (f) $\pi = 0.50$, $\alpha = 0.10$

(g) $\pi = 0.75$, $\alpha = 0.01$ (h) $\pi = 0.75$, $\alpha = 0.05$ (i) $\pi = 0.75$, $\alpha = 0.10$

Figure 4.14 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for exponential model with $p_1 = 3$ and $p_2 - 1 = 14$ at $\Delta^{\text{sim}} \geq 0$

4.2.1.3 Monomolecular Model

(1) Correct Subspace Information ($\Delta^{\text{sim}} = 0$)

In this case, the submodel is assumed to be the true model under the null hypothesis. The true values of $\boldsymbol{\beta}_1$ in this simulation were given as 3.5, 1.9, 1.4, and −0.2. The RMSEs of RE, LS, PT, SP, S, S⁺, LASSO, and aLASSO with respect to UE were computed for $p_1 = 4$ and $p_2 = 3$, 5, 7, 11, and 15. The results are reported in Table 4.9 and summarized as follows:

Table 4.9 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}^+}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LASSO}}$, and $\widehat{\boldsymbol{\beta}}_1^{\text{aLASSO}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for monomolecular model with $p_1 = 4$ at $\Delta^{\text{sim}} = 0$

Estimator			Number of Inactive Parameters (p_2)				
			3	5	7	11	15
RE			1.083	1.127	1.183	1.248	1.402
LS	$\pi = 0.25$		1.035	1.053	1.074	1.101	1.142
	$\pi = 0.50$		1.062	1.094	1.133	1.184	1.272
	$\pi = 0.75$		1.078	1.120	1.171	1.236	1.366
PT		$\alpha = 0.01$	1.080	1.121	1.171	1.233	1.374
		$\alpha = 0.05$	1.067	1.104	1.146	1.218	1.335
		$\alpha = 0.10$	1.057	1.091	1.129	1.198	1.286
SP	$\pi = 0.25$	$\alpha = 0.01$	1.034	1.051	1.070	1.096	1.134
		$\alpha = 0.05$	1.029	1.045	1.061	1.089	1.122
		$\alpha = 0.10$	1.025	1.039	1.054	1.081	1.108
	$\pi = 0.50$	$\alpha = 0.01$	1.059	1.090	1.125	1.174	1.255
		$\alpha = 0.05$	1.051	1.078	1.108	1.162	1.229
		$\alpha = 0.10$	1.044	1.069	1.096	1.147	1.200
	$\pi = 0.75$	$\alpha = 0.01$	1.075	1.114	1.160	1.222	1.341
		$\alpha = 0.05$	1.063	1.099	1.138	1.207	1.305
		$\alpha = 0.10$	1.055	1.086	1.122	1.188	1.263
S			1.023	1.071	1.112	1.185	1.327
S⁺			1.038	1.090	1.141	1.220	1.350
LASSO			1.029	1.073	1.126	1.181	1.317
aLASSO			1.033	1.080	1.128	1.217	1.320

The RMSE behavior of all estimators showed a similar pattern to the result from Cobb-Douglas and exponential models. The performance of all estimators increased as p_2 increased and the RE had the highest RMSE. As π increased, the RMSE of the LS estimator also increased and came close to that of the RE. When α increased, the performance of the PT estimator decreased. The performance of the SP estimator fell as α increased and π decreased. The PT estimator was also superior to the SP estimator at the same α. The performance of the S⁺ estimator was superior to that of shrinkage estimator. Finally, the LASSO estimator was inferior to the aLASSO estimator for all numbers of inactive parameters.

(2) Uncertain Subspace Information ($\Delta^{\text{sim}} \geq 0$)

For cases where the candidate submodel is not true, we set the true values of the active parameter in this simulation as $\boldsymbol{\beta}_1 = (3, 1.2, 1.2, 0.1, 0.1, 0.1)^\top$. The simulation model had $\boldsymbol{\beta}_2 = (\beta_7, \mathbf{0}_{p_2-1}^\top)^\top$, where $\beta_7 = \Delta^{\text{sim}}$ and selected a value of Δ^{sim} between 0 and 0.2. Here, $p_2 - 1$ is the number of inactive parameters and $p_2 = 3$, 5, 7, 11, and 15 were used in the simulations. The RMSEs of the proposed estimator for each p_2 are shown in Tables 4.10 to 4.12 and Figures 4.15 to 4.19. The findings are summarized as follows:

In this case, all estimators showed a similar pattern of RMSE behavior to both previous models. For $\Delta^{\text{sim}} > 0$, the RMSEs of both the RE and LS estimators converged to zero. When Δ^{sim} increased, the RMSEs of the PT and SP estimators dropped below 1, then gradually increased again, converging to 1. They were asymptotically equivalent to the UE when Δ^{sim} was large. The RMSEs of S and S^+ estimator fell gradually to 1. The performance of S^+ was superior to that of the shrinkage estimator when Δ^{sim} was close to zero, but they had the same values when Δ^{sim} was far from zero.

Table 4.10 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for monomolecular model with $p_1 = 6$ and $\pi = 0.25$ at $\Delta^{\text{sim}} \geq 0$

p_2	Δ^{sim}	RE	LS	PT			SP			S	S^+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
3	0.00	1.083	1.034	1.080	1.071	1.062	1.033	1.029	1.025	1.029	1.037
	0.01	1.052	1.031	1.046	1.035	1.028	1.028	1.022	1.017	1.017	1.025
	0.02	0.976	1.026	0.975	0.986	0.987	1.013	1.009	1.006	1.007	1.011
	0.04	0.762	1.005	0.931	0.970	0.983	0.994	0.997	0.999	1.002	1.002
	0.06	0.559	0.973	0.981	0.996	0.997	0.999	1.000	1.000	1.000	1.001
	0.08	0.406	0.933	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.10	0.299	0.885	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.12	0.225	0.832	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.16	0.136	0.718	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.20	0.087	0.603	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 4.10 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for monomolecular model with $p_1 = 6$ and $\pi = 0.25$ at $\Delta^{\text{sim}} \geq 0$ (Cont.)

p_2	Δ^{sim}	RE	LS	PT			SP			S	S^+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
5	0.00	1.147	1.057	1.140	1.125	1.116	1.055	1.049	1.045	1.086	1.101
	0.01	1.115	1.055	1.106	1.089	1.080	1.050	1.043	1.037	1.074	1.084
	0.02	1.034	1.050	1.026	1.016	1.017	1.035	1.024	1.020	1.048	1.054
	0.04	0.804	1.030	0.949	0.965	0.979	1.007	1.001	1.001	1.016	1.016
	0.06	0.586	0.999	0.963	0.986	0.993	0.997	0.999	0.999	1.006	1.006
	0.08	0.423	0.959	0.990	0.995	0.998	0.998	0.999	1.000	1.003	1.003
	0.10	0.310	0.911	0.998	1.000	1.000	1.000	1.000	1.000	1.002	1.002
	0.12	0.232	0.857	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001
	0.16	0.139	0.740	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.20	0.089	0.622	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
7	0.00	1.175	1.073	1.160	1.133	1.120	1.068	1.059	1.053	1.098	1.132
	0.01	1.144	1.072	1.124	1.101	1.092	1.064	1.054	1.048	1.096	1.117
	0.02	1.063	1.067	1.049	1.035	1.037	1.051	1.038	1.032	1.080	1.087
	0.04	0.826	1.048	0.968	0.985	0.987	1.020	1.010	1.006	1.042	1.042
	0.06	0.601	1.018	0.971	0.989	0.994	1.002	1.000	1.000	1.022	1.022
	0.08	0.433	0.978	0.989	0.997	0.998	0.999	1.000	1.000	1.013	1.013
	0.10	0.317	0.929	0.997	1.000	1.000	1.000	1.000	1.000	1.008	1.008
	0.12	0.236	0.874	1.000	1.000	1.000	1.000	1.000	1.000	1.006	1.006
	0.16	0.143	0.760	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
	0.20	0.091	0.639	1.000	1.000	1.000	1.000	1.000	1.000	1.002	1.002
11	0.00	1.306	1.113	1.291	1.259	1.227	1.108	1.098	1.087	1.221	1.258
	0.01	1.277	1.113	1.256	1.215	1.179	1.105	1.091	1.078	1.198	1.231
	0.02	1.190	1.109	1.153	1.115	1.086	1.089	1.068	1.055	1.156	1.177
	0.04	0.930	1.092	0.979	0.984	0.986	1.037	1.020	1.012	1.090	1.094
	0.06	0.678	1.065	0.958	0.989	0.993	1.006	1.004	1.002	1.054	1.054
	0.08	0.489	1.027	0.982	0.996	0.998	1.000	1.000	1.000	1.035	1.035
	0.10	0.358	0.980	0.996	0.998	0.999	1.000	1.000	1.000	1.024	1.024
	0.12	0.268	0.927	1.000	1.000	1.000	1.000	1.000	1.000	1.018	1.018
	0.16	0.155	0.801	1.000	1.000	1.000	1.000	1.000	1.000	1.010	1.010
	0.20	0.101	0.684	1.000	1.000	1.000	1.000	1.000	1.000	1.006	1.006
15	0.00	1.390	1.143	1.358	1.307	1.273	1.135	1.120	1.107	1.307	1.341
	0.01	1.352	1.141	1.315	1.260	1.225	1.130	1.112	1.097	1.286	1.310
	0.02	1.256	1.136	1.203	1.153	1.130	1.112	1.088	1.073	1.235	1.248
	0.04	0.980	1.116	1.008	1.013	1.015	1.052	1.034	1.025	1.139	1.141
	0.06	0.717	1.086	0.970	0.983	0.989	1.013	1.005	1.003	1.082	1.082
	0.08	0.519	1.047	0.979	0.995	0.998	1.001	1.001	1.000	1.051	1.051
	0.10	0.381	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.034	1.034
	0.12	0.285	0.947	1.000	1.001	1.000	1.000	1.000	1.000	1.023	1.023
	0.16	0.170	0.827	1.000	1.000	1.000	1.000	1.000	1.000	1.012	1.012
	0.20	0.110	0.708	1.000	1.000	1.000	1.000	1.000	1.000	1.006	1.006

Table 4.11 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for monomolecular model with $p_1 = 6$ and $\pi = 0.50$ at $\Delta^{\mathrm{sim}} \geq 0$

p_2	Δ^{sim}	RE	LS	PT			SP			S	S$^+$
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
3	0.00	1.083	1.060	1.080	1.071	1.062	1.058	1.051	1.044	1.029	1.037
	0.01	1.052	1.051	1.046	1.035	1.028	1.045	1.035	1.028	1.017	1.025
	0.02	0.976	1.030	0.975	0.986	0.987	1.013	1.010	1.005	1.007	1.011
	0.04	0.762	0.956	0.931	0.970	0.983	0.980	0.991	0.995	1.002	1.002
	0.06	0.559	0.856	0.981	0.996	0.997	0.995	0.999	0.999	1.000	1.001
	0.08	0.406	0.746	0.998	1.000	1.000	0.999	1.000	1.000	1.000	1.000
	0.10	0.299	0.640	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.12	0.225	0.543	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.16	0.136	0.387	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.20	0.087	0.276	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.00	1.147	1.104	1.140	1.125	1.116	1.098	1.088	1.081	1.086	1.101
	0.01	1.115	1.095	1.106	1.089	1.080	1.087	1.073	1.064	1.074	1.084
	0.02	1.034	1.074	1.026	1.016	1.017	1.052	1.035	1.029	1.048	1.054
	0.04	0.804	0.998	0.949	0.965	0.979	1.000	0.996	0.997	1.016	1.016
	0.06	0.586	0.894	0.963	0.986	0.993	0.990	0.996	0.998	1.006	1.006
	0.08	0.423	0.779	0.990	0.995	0.998	0.996	0.998	0.999	1.003	1.003
	0.10	0.310	0.666	0.998	1.000	1.000	0.999	1.000	1.000	1.002	1.002
	0.12	0.232	0.564	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001
	0.16	0.139	0.400	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.20	0.089	0.285	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
7	0.00	1.175	1.131	1.160	1.133	1.120	1.121	1.103	1.093	1.098	1.132
	0.01	1.144	1.124	1.124	1.101	1.092	1.109	1.091	1.081	1.096	1.117
	0.02	1.063	1.103	1.049	1.035	1.037	1.079	1.058	1.050	1.080	1.087
	0.04	0.826	1.027	0.968	0.985	0.987	1.021	1.011	1.006	1.042	1.042
	0.06	0.601	0.920	0.971	0.989	0.994	0.997	0.998	0.999	1.022	1.022
	0.08	0.433	0.801	0.989	0.997	0.998	0.996	0.999	0.999	1.013	1.013
	0.10	0.317	0.685	0.997	1.000	1.000	0.999	1.000	1.000	1.008	1.008
	0.12	0.236	0.577	1.000	1.000	1.000	1.000	1.000	1.000	1.006	1.006
	0.16	0.143	0.414	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
	0.20	0.091	0.293	1.000	1.000	1.000	1.000	1.000	1.000	1.002	1.002
11	0.00	1.306	1.211	1.291	1.259	1.227	1.202	1.181	1.160	1.221	1.258
	0.01	1.277	1.206	1.256	1.215	1.179	1.192	1.163	1.138	1.198	1.231
	0.02	1.190	1.187	1.153	1.115	1.086	1.151	1.114	1.090	1.156	1.177
	0.04	0.930	1.111	0.979	0.984	0.986	1.045	1.023	1.014	1.090	1.094
	0.06	0.678	1.002	0.958	0.989	0.993	1.001	1.004	1.002	1.054	1.054
	0.08	0.489	0.878	0.982	0.996	0.998	0.997	1.000	1.000	1.035	1.035
	0.10	0.358	0.754	0.996	0.998	0.999	0.999	0.999	1.000	1.024	1.024
	0.12	0.268	0.641	1.000	1.000	1.000	1.000	1.000	1.000	1.018	1.018
	0.16	0.155	0.446	1.000	1.000	1.000	1.000	1.000	1.000	1.010	1.010
	0.20	0.101	0.323	1.000	1.000	1.000	1.000	1.000	1.000	1.006	1.006

Table 4.11 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for monomolecular model with $p_1 = 6$ and $\pi = 0.50$ at $\Delta^{\mathrm{sim}} \geq 0$ (Cont.)

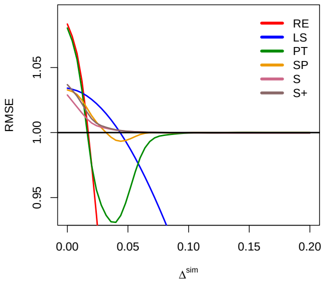
p_2	Δ^{sim}	RE	LS	PT			SP			S	S+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
15	0.00	1.390	1.272	1.358	1.307	1.273	1.253	1.222	1.197	1.307	1.341
	0.01	1.352	1.263	1.315	1.260	1.225	1.239	1.202	1.174	1.286	1.310
	0.02	1.256	1.239	1.203	1.153	1.130	1.194	1.149	1.124	1.235	1.248
	0.04	0.980	1.157	1.008	1.013	1.015	1.072	1.048	1.037	1.139	1.141
	0.06	0.717	1.042	0.970	0.983	0.989	1.013	1.004	1.002	1.082	1.082
	0.08	0.519	0.914	0.979	0.995	0.998	0.997	1.000	1.000	1.051	1.051
	0.10	0.381	0.787	0.993	0.999	1.000	0.998	1.000	1.000	1.034	1.034
	0.12	0.285	0.670	1.000	1.001	1.000	1.000	1.001	1.000	1.023	1.023
	0.16	0.170	0.478	1.000	1.000	1.000	1.000	1.000	1.000	1.012	1.012
	0.20	0.110	0.345	1.000	1.000	1.000	1.000	1.000	1.000	1.006	1.006

Table 4.12 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for monomolecular model with $p_1 = 6$ and $\pi = 0.75$ at $\Delta^{\mathrm{sim}} \geq 0$

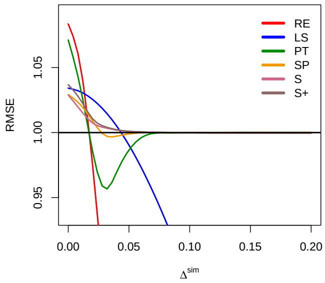
p_2	Δ^{sim}	RE	LS	PT			SP			S	S+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
3	0.00	1.083	1.077	1.080	1.071	1.062	1.074	1.066	1.057	1.029	1.037
	0.01	1.052	1.058	1.046	1.035	1.028	1.051	1.039	1.032	1.017	1.025
	0.02	0.976	1.013	0.975	0.986	0.987	1.000	1.002	0.999	1.007	1.011
	0.04	0.762	0.869	0.931	0.970	0.983	0.959	0.982	0.990	1.002	1.002
	0.06	0.559	0.703	0.981	0.996	0.997	0.989	0.998	0.998	1.000	1.001
	0.08	0.406	0.555	0.998	1.000	1.000	0.999	1.000	1.000	1.000	1.000
	0.10	0.299	0.435	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.12	0.225	0.342	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.16	0.136	0.218	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.20	0.087	0.145	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.00	1.147	1.135	1.140	1.125	1.116	1.128	1.114	1.105	1.086	1.101
	0.01	1.115	1.116	1.106	1.089	1.080	1.106	1.089	1.079	1.074	1.084
	0.02	1.034	1.068	1.026	1.016	1.017	1.049	1.032	1.028	1.048	1.054
	0.04	0.804	0.915	0.949	0.965	0.979	0.980	0.984	0.990	1.016	1.016
	0.06	0.586	0.739	0.963	0.986	0.993	0.979	0.992	0.996	1.006	1.006
	0.08	0.423	0.580	0.990	0.995	0.998	0.993	0.997	0.999	1.003	1.003
	0.10	0.310	0.453	0.998	1.000	1.000	0.999	1.000	1.000	1.002	1.002
	0.12	0.232	0.355	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001
	0.16	0.139	0.225	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.20	0.089	0.149	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 4.12 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for monomolecular model with $p_1 = 6$ and $\pi = 0.75$ at $\Delta^{\mathrm{sim}} \geq 0$ (Cont.)

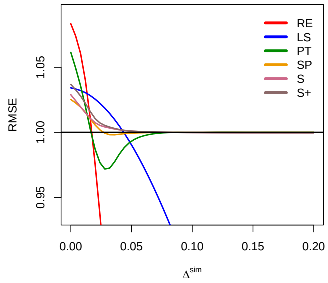
p_2	Δ^{sim}	RE	LS	PT			SP			S	S+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
7	0.00	1.175	1.166	1.160	1.133	1.120	1.153	1.129	1.116	1.098	1.132
	0.01	1.144	1.149	1.124	1.101	1.092	1.130	1.108	1.096	1.096	1.117
	0.02	1.063	1.102	1.049	1.035	1.037	1.078	1.057	1.052	1.080	1.087
	0.04	0.826	0.944	0.968	0.985	0.987	1.003	1.002	0.999	1.042	1.042
	0.06	0.601	0.761	0.971	0.989	0.994	0.987	0.995	0.997	1.022	1.022
	0.08	0.433	0.596	0.989	0.997	0.998	0.993	0.998	0.999	1.013	1.013
	0.10	0.317	0.464	0.997	1.000	1.000	0.998	1.000	1.000	1.008	1.008
	0.12	0.236	0.362	1.000	1.000	1.000	1.000	1.000	1.000	1.006	1.006
	0.16	0.143	0.232	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
	0.20	0.091	0.153	1.000	1.000	1.000	1.000	1.000	1.000	1.002	1.002
11	0.00	1.306	1.280	1.291	1.259	1.227	1.266	1.238	1.209	1.221	1.258
	0.01	1.277	1.265	1.256	1.215	1.179	1.246	1.207	1.174	1.198	1.231
	0.02	1.190	1.216	1.153	1.115	1.086	1.174	1.130	1.101	1.156	1.177
	0.04	0.930	1.049	0.979	0.984	0.986	1.025	1.011	1.005	1.090	1.094
	0.06	0.678	0.850	0.958	0.989	0.993	0.984	0.999	0.999	1.054	1.054
	0.08	0.489	0.669	0.982	0.996	0.998	0.991	0.998	0.999	1.035	1.035
	0.10	0.358	0.522	0.996	0.998	0.999	0.998	0.999	1.000	1.024	1.024
	0.12	0.268	0.409	1.000	1.000	1.000	1.000	1.000	1.000	1.018	1.018
	0.16	0.155	0.251	1.000	1.000	1.000	1.000	1.000	1.000	1.010	1.010
	0.20	0.101	0.170	1.000	1.000	1.000	1.000	1.000	1.000	1.006	1.006
15	0.00	1.390	1.362	1.358	1.307	1.273	1.334	1.289	1.256	1.307	1.341
	0.01	1.352	1.340	1.315	1.260	1.225	1.307	1.256	1.220	1.286	1.310
	0.02	1.256	1.284	1.203	1.153	1.130	1.227	1.172	1.144	1.235	1.248
	0.04	0.980	1.104	1.008	1.013	1.015	1.057	1.041	1.033	1.139	1.141
	0.06	0.717	0.895	0.970	0.983	0.989	0.998	0.996	0.998	1.082	1.082
	0.08	0.519	0.705	0.979	0.995	0.998	0.990	0.998	0.999	1.051	1.051
	0.10	0.381	0.552	0.993	0.999	1.000	0.996	1.000	1.000	1.034	1.034
	0.12	0.285	0.433	1.000	1.001	1.000	1.000	1.001	1.000	1.023	1.023
	0.16	0.170	0.274	1.000	1.000	1.000	1.000	1.000	1.000	1.012	1.012
	0.20	0.110	0.184	1.000	1.000	1.000	1.000	1.000	1.000	1.006	1.006



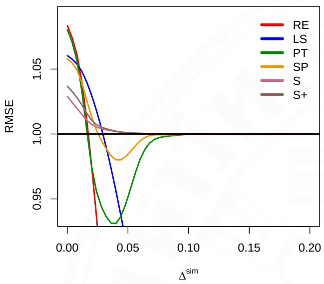
(a) $\pi = 0.25$, $\alpha = 0.01$



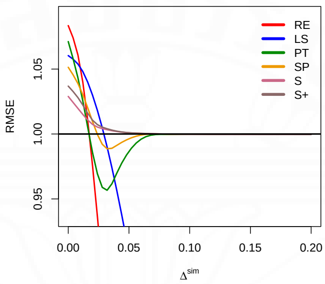
(b) $\pi = 0.25$, $\alpha = 0.05$



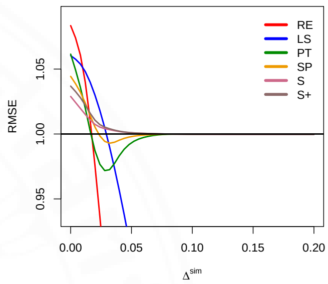
(c) $\pi = 0.25$, $\alpha = 0.10$



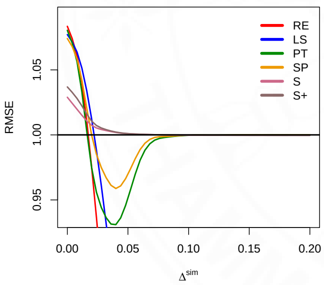
(d) $\pi = 0.50$, $\alpha = 0.01$



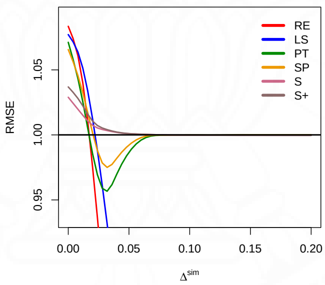
(e) $\pi = 0.50$, $\alpha = 0.05$



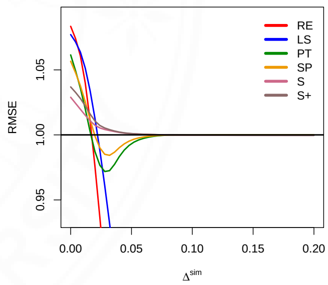
(f) $\pi = 0.50$, $\alpha = 0.10$



(g) $\pi = 0.75$, $\alpha = 0.01$

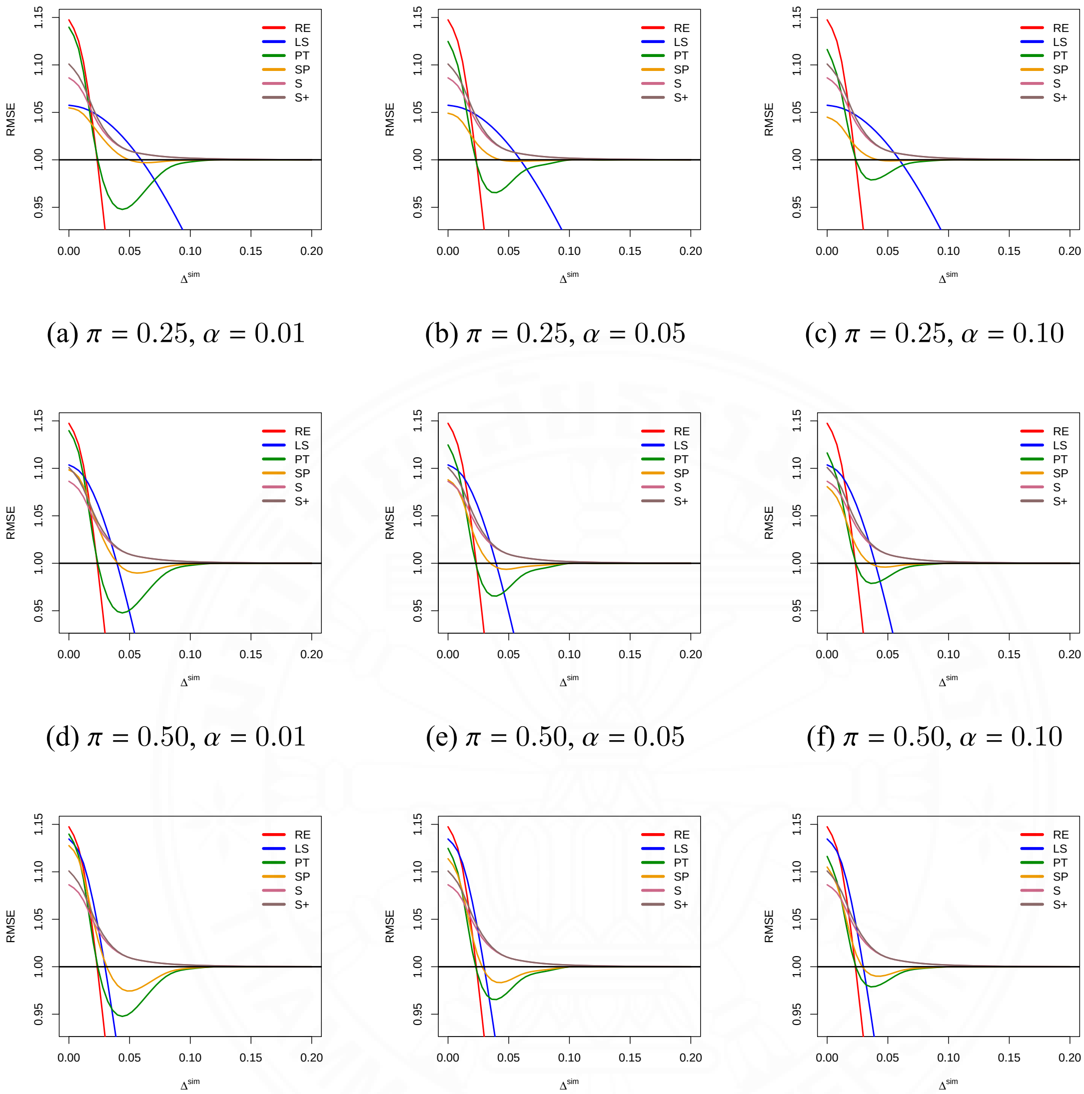


(h) $\pi = 0.75$, $\alpha = 0.05$



(i) $\pi = 0.75$, $\alpha = 0.10$

Figure 4.15 RMSEs of $\widehat{\beta}_1^{\mathrm{RE}}$, $\widehat{\beta}_1^{\mathrm{LS}}$, $\widehat{\beta}_1^{\mathrm{PT}}$, $\widehat{\beta}_1^{\mathrm{SP}}$, $\widehat{\beta}_1^{\mathrm{S}}$, and $\widehat{\beta}_1^{\mathrm{S+}}$ with respect to $\widehat{\beta}_1^{\mathrm{UE}}$ for monomolecular model with $p_1 = 6$ and $p_2 - 1 = 2$ at $\Delta^{\mathrm{sim}} \geq 0$

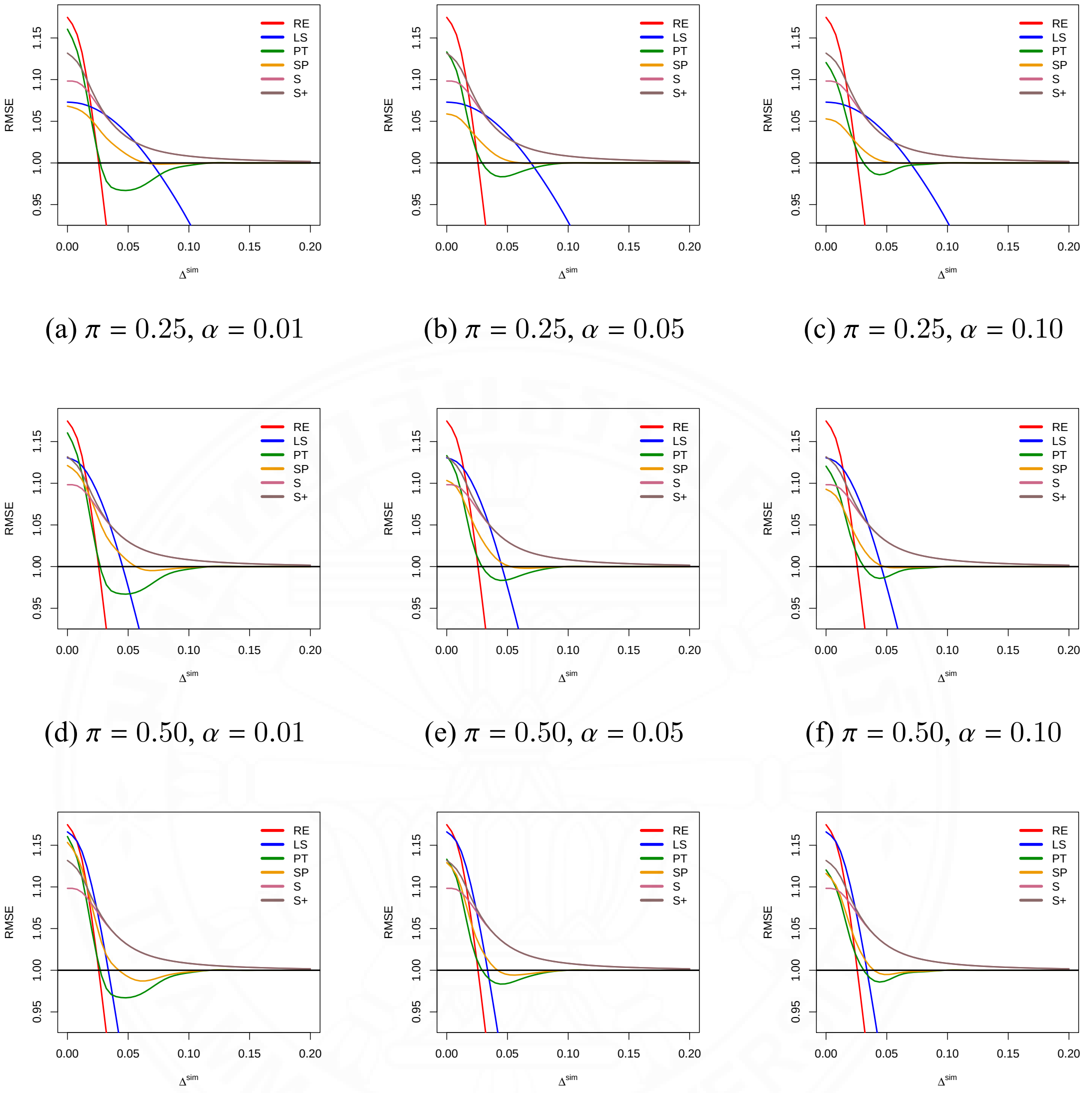


(a) $\pi = 0.25$, $\alpha = 0.01$ (b) $\pi = 0.25$, $\alpha = 0.05$ (c) $\pi = 0.25$, $\alpha = 0.10$

(d) $\pi = 0.50$, $\alpha = 0.01$ (e) $\pi = 0.50$, $\alpha = 0.05$ (f) $\pi = 0.50$, $\alpha = 0.10$

(g) $\pi = 0.75$, $\alpha = 0.01$ (h) $\pi = 0.75$, $\alpha = 0.05$ (i) $\pi = 0.75$, $\alpha = 0.10$

Figure 4.16 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for monomolecular model with $p_1 = 6$ and $p_2 - 1 = 4$ at $\Delta^{\text{sim}} \geq 0$

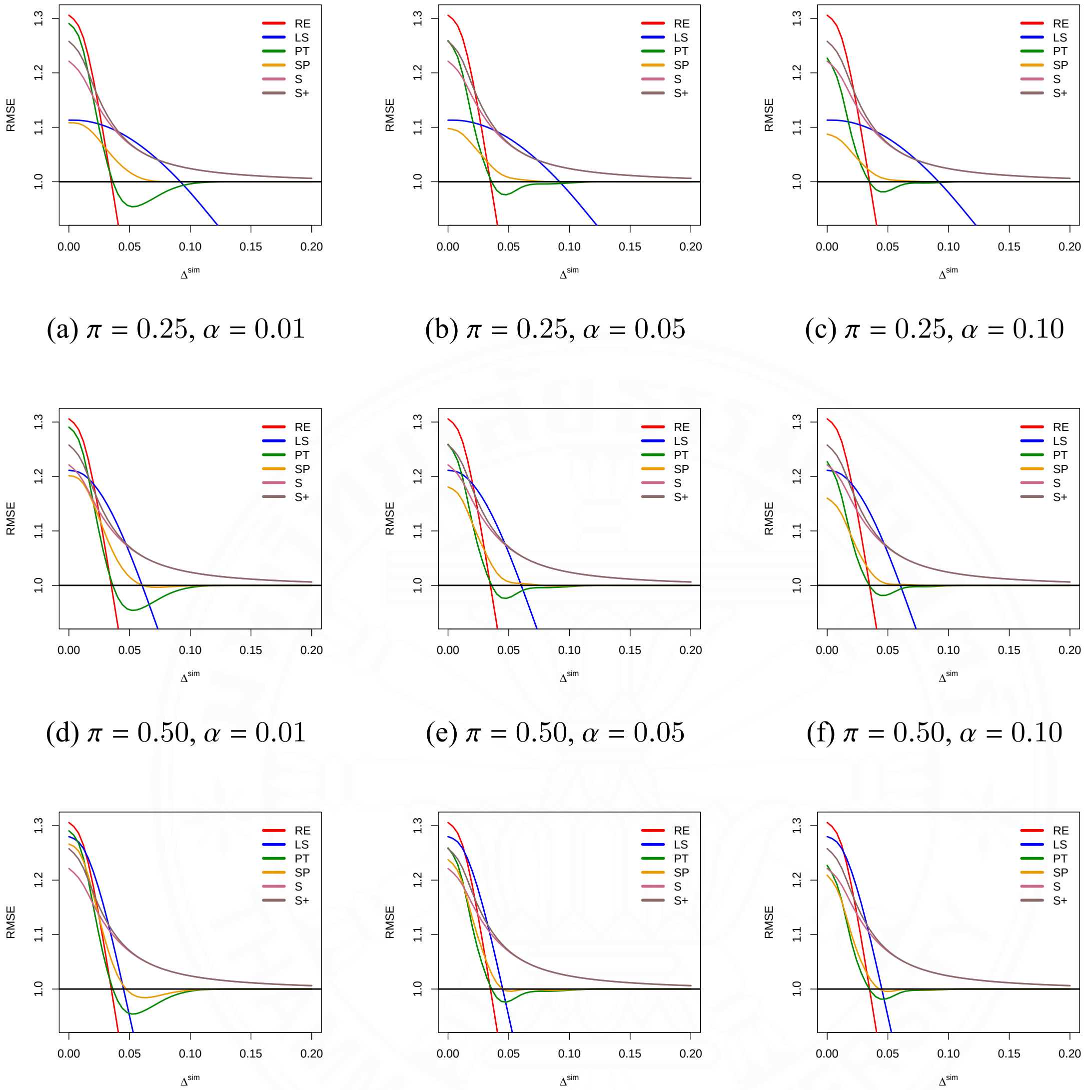


(a) $\pi = 0.25, \alpha = 0.01$ (b) $\pi = 0.25, \alpha = 0.05$ (c) $\pi = 0.25, \alpha = 0.10$

(d) $\pi = 0.50, \alpha = 0.01$ (e) $\pi = 0.50, \alpha = 0.05$ (f) $\pi = 0.50, \alpha = 0.10$

(g) $\pi = 0.75, \alpha = 0.01$ (h) $\pi = 0.75, \alpha = 0.05$ (i) $\pi = 0.75, \alpha = 0.10$

Figure 4.17 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for monomolecular model with $p_1 = 6$ and $p_2 - 1 = 6$ at $\Delta^{\text{sim}} \geq 0$

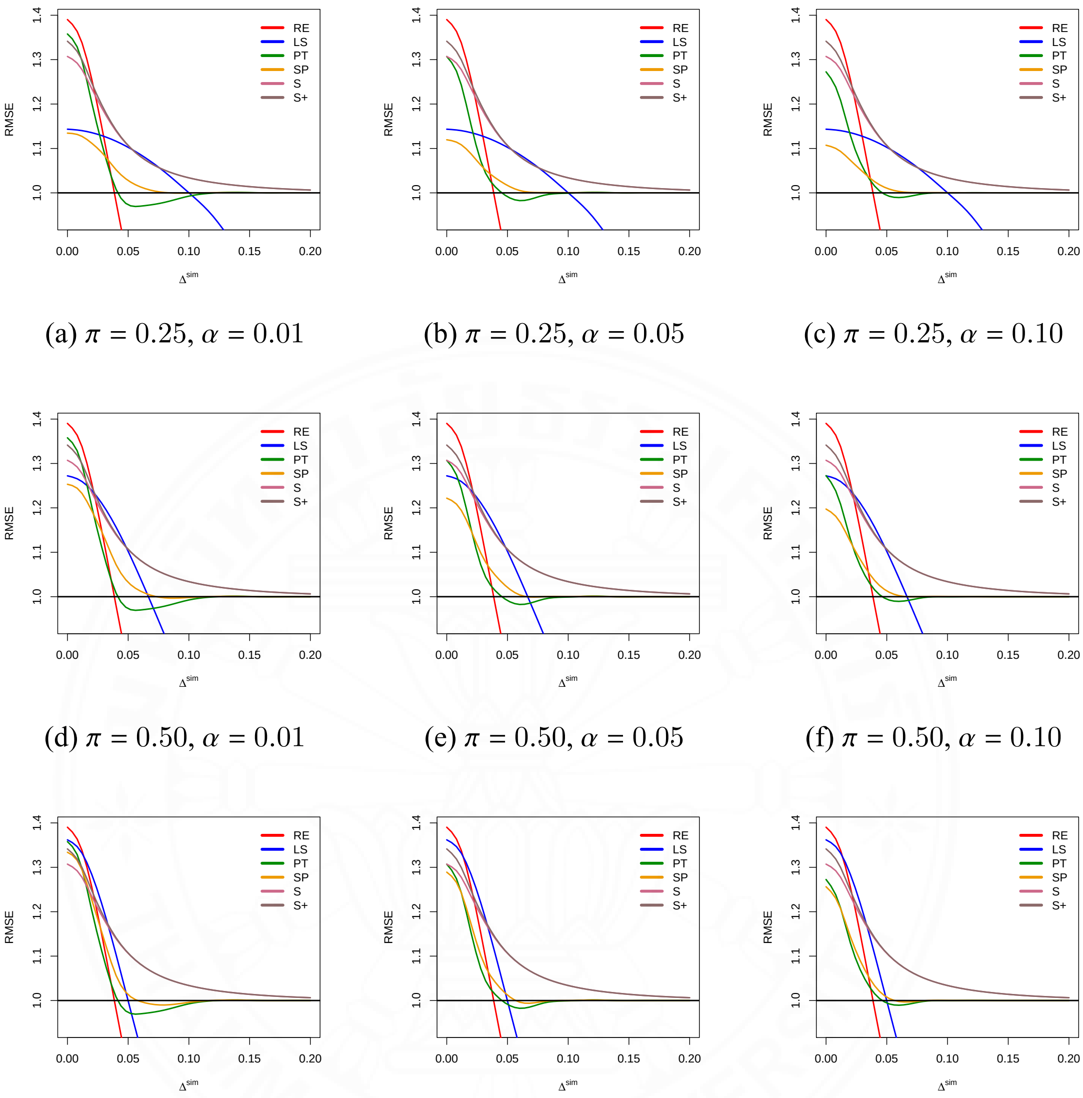


(a) $\pi = 0.25$, $\alpha = 0.01$ (b) $\pi = 0.25$, $\alpha = 0.05$ (c) $\pi = 0.25$, $\alpha = 0.10$

(d) $\pi = 0.50$, $\alpha = 0.01$ (e) $\pi = 0.50$, $\alpha = 0.05$ (f) $\pi = 0.50$, $\alpha = 0.10$

(g) $\pi = 0.75$, $\alpha = 0.01$ (h) $\pi = 0.75$, $\alpha = 0.05$ (i) $\pi = 0.75$, $\alpha = 0.10$

Figure 4.18 RMSEs of $\widehat{\beta}_1^{\mathrm{RE}}$, $\widehat{\beta}_1^{\mathrm{LS}}$, $\widehat{\beta}_1^{\mathrm{PT}}$, $\widehat{\beta}_1^{\mathrm{SP}}$, $\widehat{\beta}_1^{\mathrm{S}}$, and $\widehat{\beta}_1^{\mathrm{S+}}$ with respect to $\widehat{\beta}_1^{\mathrm{UE}}$ for monomolecular model with $p_1 = 6$ and $p_2 - 1 = 10$ at $\Delta^{\mathrm{sim}} \geq 0$



(a) $\pi = 0.25, \alpha = 0.01$ (b) $\pi = 0.25, \alpha = 0.05$ (c) $\pi = 0.25, \alpha = 0.10$

(d) $\pi = 0.50, \alpha = 0.01$ (e) $\pi = 0.50, \alpha = 0.05$ (f) $\pi = 0.50, \alpha = 0.10$

(g) $\pi = 0.75, \alpha = 0.01$ (h) $\pi = 0.75, \alpha = 0.05$ (i) $\pi = 0.75, \alpha = 0.10$

Figure 4.19 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for monomolecular model with $p_1 = 6$ and $p_2 - 1 = 14$ at $\Delta^{\text{sim}} \geq 0$

From the RMSE results for the three nonlinear models of interest, we conclude that the uncertainty of subspace information has severe implications for RE and LS estimator but it had a smaller impact on estimators based on pretest and shrinkage strategies. The use of the S^+ estimator was safe for dealing with overfitting in the presence of subspace information. However, the SP estimator was suggested to apply when $p_2 < 3$.

4.2.2 High-Dimensional Data Setting

For the Cobb-Douglas, exponential, and monomolecular nonlinear regression models, we generated the response variable from Equations (4.30), (4.31), and (4.32), respectively, with regression coefficient vectors $\boldsymbol{\beta} = (\boldsymbol{\beta}_s^\top, \boldsymbol{\beta}_w^\top, \boldsymbol{\beta}_n^\top)^\top$, having strong, weak-to-moderate, and no signals. Each nonzero coefficient with weak-to-moderate signals was randomly assigned to have either positive or negative signs.

We considered the sample sizes and the number of parameters with strong, weak-to-moderate, and no signals were $(n, p_s, p_w, p_n) = (150, 5, 25, 170)$, which satisfied the regular assumptions $p_s \leq p_w < n$ and $p_n > n$ (Ahmed and Yüzbaşı, 2016; Yüzbaşı et al., 2017) with 2,000 simulations. We next consider nonlinear models of interest in the context of the high-dimensional sparse regime in the following subsection.

4.2.2.1 Cobb-Douglas Model

In this part of the study, the regression coefficients of the Cobb-Douglas model were set as $\boldsymbol{\beta} = (\boldsymbol{\beta}_s^\top, \boldsymbol{\beta}_w^\top, \boldsymbol{\beta}_n^\top)^\top = (1, 1, 0.75, 0.75, 0.75, \boldsymbol{\kappa}_{p_w}^\top, \mathbf{0}_{p_n}^\top)^\top$. To study what happens when very weak signal became moderate signals, the size of the weak-to-moderate signals (κ) were set as 0.0100, 0.0125, 0.0150, 0.0175, and 0.0200. In this simulation setting, we simulated 2,000 datasets consisting of $n = 150$, $k_s = 4$, $k_w = 25$, and $k_n = 170$.

We investigated the performance of LASSO and aLASSO as variable selection criteria, and the percentages of predictors selected for each method were displayed. For example, if the percentage of any one predictor was 100, then this predictor was always selected for all simulation steps. Likewise, if the percentage of anyone predictor was zero, then this predictor was never selected in the simulation steps. The findings are summarized in Table 4.13 and Figure 4.20.

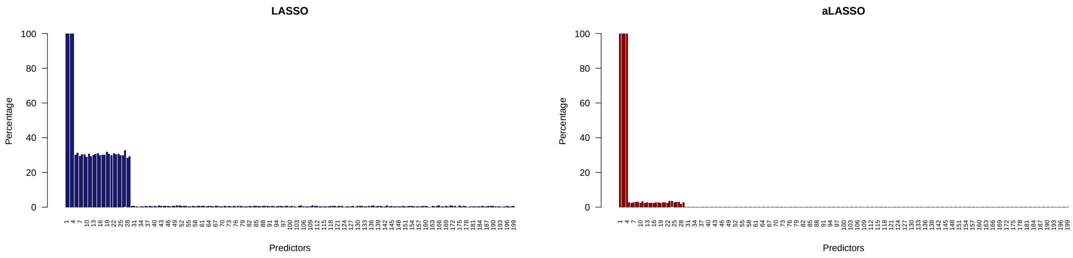
Table 4.13 Selection percentages of predictors using LASSO and aLASSO strategies in Cobb-Douglas model with strong, weak, and no signals for $(n, k_s, k_w, k_n) = (150, 4, 25, 170)$

κ	Strong Signal		Weak Signal		No Signal	
	LASSO	aLASSO	LASSO	aLASSO	LASSO	aLASSO
0.0100	100.00	100.00	30.31	2.77	0.57	0.02
0.0125	100.00	100.00	54.64	8.65	1.58	0.09
0.0150	100.00	100.00	70.20	17.50	2.43	0.25
0.0175	100.00	100.00	80.36	27.40	3.09	0.50
0.0200	100.00	100.00	87.30	37.28	3.57	0.78

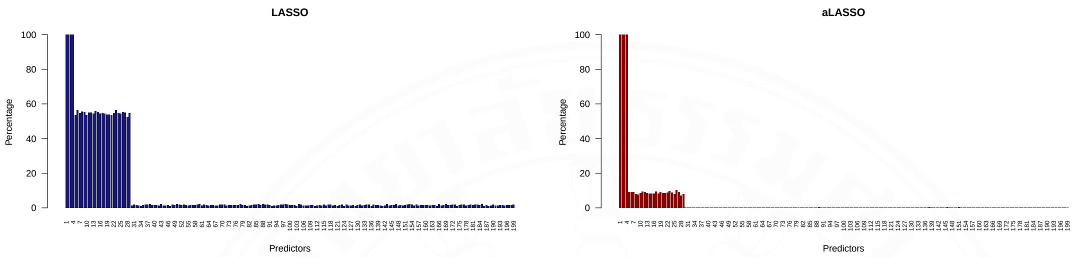
The results show that the LASSO and aLASSO methods produced different structural sparsities in parameter space. LASSO elected a larger number of predictors than aLASSO, though unfortunately, LASSO also retained more predictors with no signal than aLASSO, especially when κ was large. This indicates that variable selection results using the aLASSO strategy presented a lower-dimensional model than when using the LASSO strategy.

When we consider the size of the weak-to-moderate signal (κ), if κ increased, the selection percentage of predictors using LASSO and aLASSO with strong signals remained unchanged while the selection percentage of predictors with weak or no signals increased. However, aLASSO outperformed LASSO well by deleting the predictors with weak and no signals since their selection percentage was tiny. Furthermore, as κ grew, the weak signal became more robust. LASSO showed strong efficiency in picking predictors with both strong and weak signals, although it still stored too many predictors with no signals.

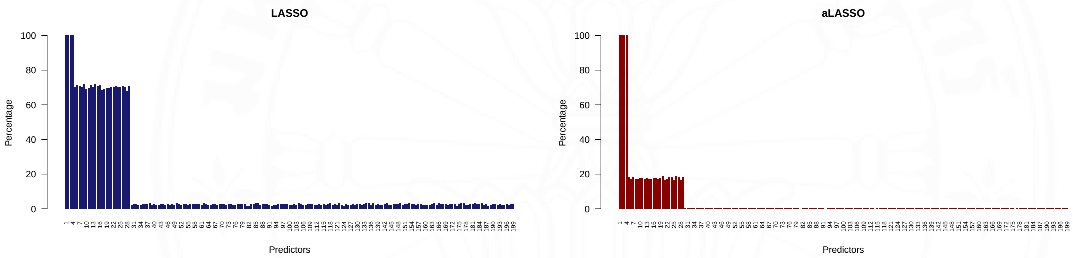
Remarkably, the predictors with weak signals were of little significance for predicting the response variable when κ was very small, and could be removed from the model. In contrast, they become powerful and should be included in the model for large values of κ. We can see that, for small values of κ, LASSO may produce overfitting with too many trivially significant or insignificant predictors being selected. On the other hand, aLASSO may build underfitting because it selects fewer significant predictors as κ becomes large. For this reason, the most suitable variable selection method cannot be identified for all cases, so neither of the two candidate models from the variable selection results using the LASSO and aLASSO strategies may be the best choice.



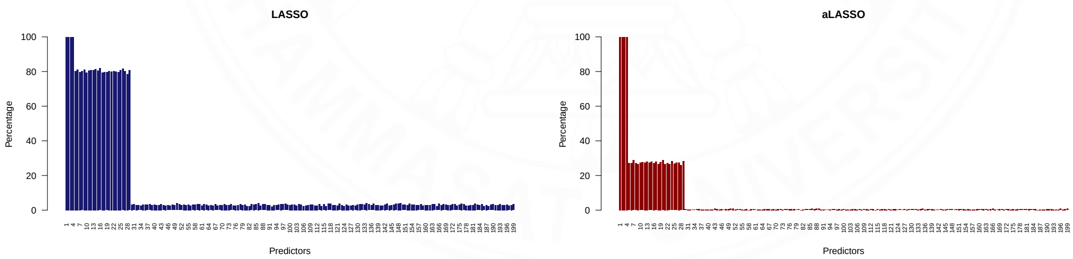
(a) $\kappa = 0.0100$



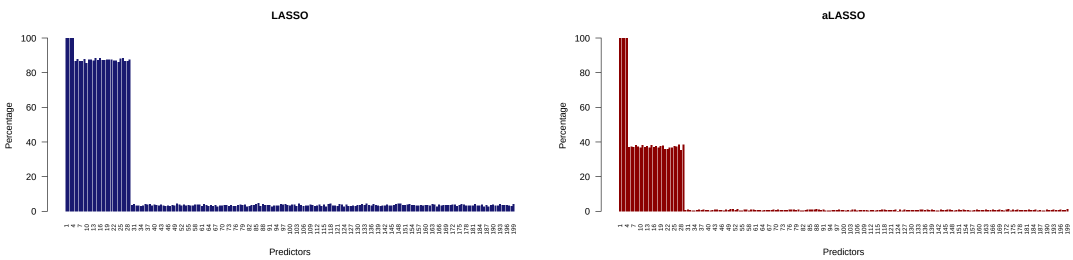
(b) $\kappa = 0.0125$



(c) $\kappa = 0.0150$



(d) $\kappa = 0.0175$



(e) $\kappa = 0.0200$

Figure 4.20 Comparison of percentage of each predictor selected using LASSO and aLASSO strategies in Cobb-Douglas model for each κ and $(n, k_s, k_w, k_n) = (150, 4, 25, 170)$

Finally, the post-selection estimators based on LS, PT, SP, S, and S^+ strategies for parameter estimation were applied. The RMSE results of suggested estimators for $\alpha = 0.05$ and $\pi = 0.5$ are reported in Table 4.14. The findings can be summarized as follows:

1. All estimators had the highest efficiency when $\kappa = 0.0100$, and their RMSEs fell as κ increased.

2. For very small values of κ, the RE confirmed that aLASSO produced a more appropriate model. However, the RE became less effective than the other estimators when κ increased, indicating that the LASSO model was more suitable. At the same time, aLASSO led to underfitting when the null hypothesis was untrue.

3. As κ increased, the LS estimator outperformed SM, which was inferior to other estimators. When κ was small, the performance of the LS estimator was better than all other estimators except the RE.

4. The performance of both PT and SP estimators dropped below one as κ increased, then they increased, becoming equal to the UE. The SP estimator also had a poorer performance than the PT estimator only when κ was small.

5. The RMSE of the S^+ estimator was superior to the shrinkage (S) estimator when κ was small and became equal as κ increased. Both S and S^+ estimators also outperformed RE and LS when the value of κ was large and still dominated PT and SP for small values of κ, which was similar to the simulation results in a low-dimensional setting when the number of inactive parameters was large.

Table 4.14 RMSEs of estimators with respect to the UE in Cobb-Douglas model for a high-dimensional setting where $\alpha = 0.05$ and $\pi = 0.5$

κ	Estimators					
	RE	LS	PT	SP	S	S^+
0.0100	1.1040	1.0506	1.0027	1.0014	1.0168	1.0341
0.0125	1.0116	1.0076	1.0004	1.0002	1.0032	1.0044
0.0150	0.9952	1.0001	1.0000	1.0000	1.0004	1.0005
0.0175	0.9665	0.9861	0.9998	0.9999	0.9955	0.9955
0.0200	0.9478	0.9770	1.0000	1.0000	0.9935	0.9935

These results verified that the proposed post-selection estimators were robust even when the UF and OF models were unworthy, which made the UE and RE impotent. Therefore, the suggested estimators were recommended for post-selection parameter estimation in the high-dimensional sparse Cobb-Douglas model when the accuracy of the results was unknown. These results were also consistent with the theoretical and numerical results in the low-dimensional regime.

4.2.2.2 Exponential Model

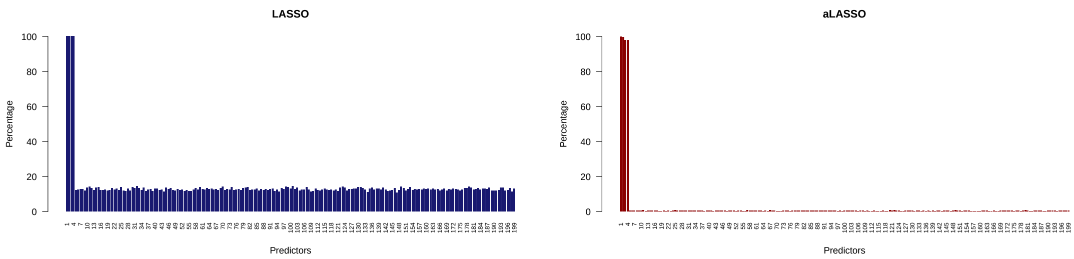
The sparse exponential nonlinear regression was considered and the regression coefficients were chosen to be $\boldsymbol{\beta} = (\boldsymbol{\beta}_s^\top, \boldsymbol{\beta}_w^\top, \boldsymbol{\beta}_n^\top)^\top = (\boldsymbol{\beta}_s^\top, \boldsymbol{\kappa}_{p_w}, \mathbf{0}_{p_n})^\top$, where $\boldsymbol{\beta}_s = (1.2, 1.2, 1.2, 0.9, 0.9)^\top$. In order to investigate the behavior of the estimators from weak to moderate signals, we defined the values of κ as 0.005, 0.050, 0.100, 0.200, 0.300. We generated 2,000 replications of the exponential model with 150 sample sizes (n), and we set the number of predictors with strong (k_s), weak (k_w), and no signals (k_n) as 4, 25, and 170, respectively.

For the variable selection step, the resulting percentages of selected predictors using the LASSO and aLASSO methods with each signal are reported in Table 4.15. In addition, the resulting percentages of each predictor chosen using LASSO and aLASSO are graphically presented in Figure 4.21.

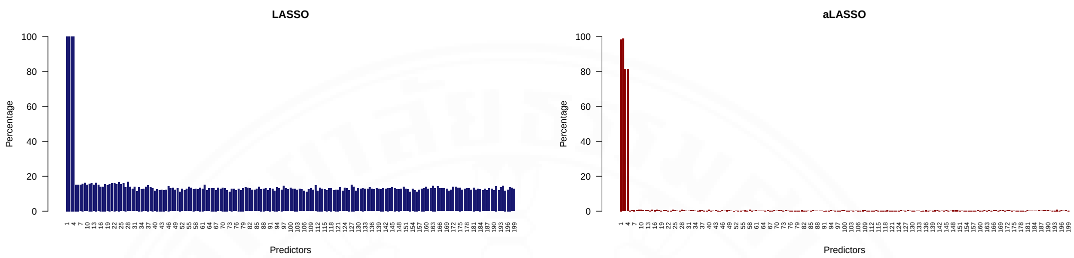
Table 4.15 Selection percentages of predictors using LASSO and aLASSO methods in exponential model with strong, weak, and no signals for $(n, k_s, k_w, k_n) = (150, 4, 25, 170)$

κ	Strong Signal		Weak Signal		No Signal	
	LASSO	aLASSO	LASSO	aLASSO	LASSO	aLASSO
0.005	100.00	98.45	12.64	0.39	12.67	0.35
0.050	100.00	87.18	15.36	0.53	12.86	0.37
0.100	100.00	84.90	23.28	0.96	14.27	0.38
0.200	99.92	81.65	42.71	3.36	17.07	0.56
0.300	99.88	80.50	57.79	8.12	19.80	0.99

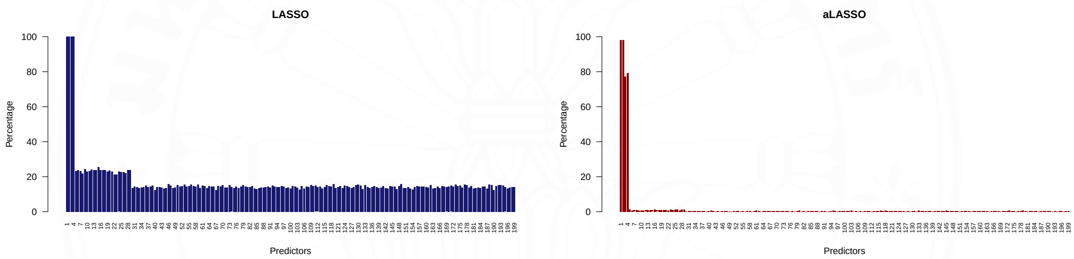
From Table 4.15 and Figure 4.21, we can see that the LASSO strategy performed better in picking predictors with strong and weak signals than the aLASSO strategy for all cases. This confirms that the two different variable selection strategies built different candidate subsets of selected predictors for establishing the model.



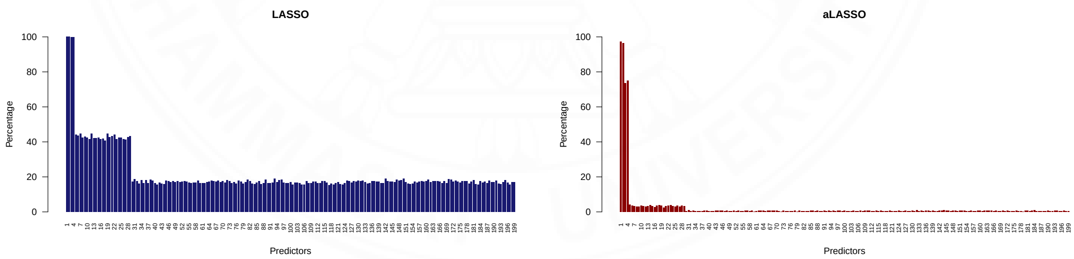
(a) $\kappa = 0.005$



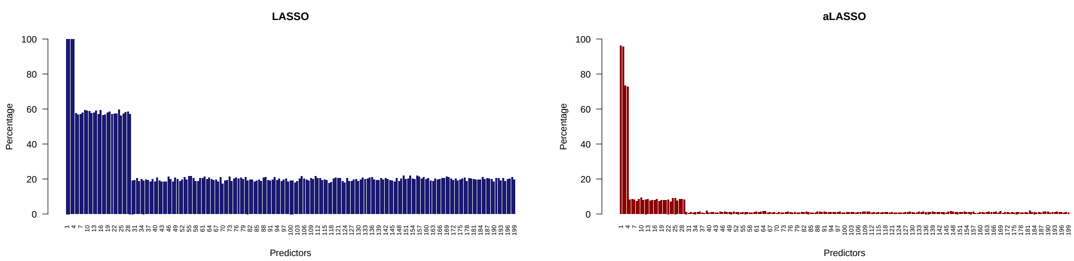
(b) $\kappa = 0.050$



(c) $\kappa = 0.100$



(d) $\kappa = 0.200$



(e) $\kappa = 0.300$

Figure 4.21 Comparison of percentage of each predictor selected using LASSO and aLASSO strategies in exponential model for each κ and $(n, k_s, k_w, k_n) = (150, 4, 25, 170)$

When κ was small, the predictors with weak signals had little or no influence on the response variable for the aLASSO strategy. However, as κ increased, the predictors with these signals were significant, especially in the LASSO strategy, which performed better than aLASSO in selecting predictors.

After using LASSO and aLASSO for variable selection, we applied the proposed estimators such as LS, PT, SP, S, and S^+ to the models from the dimensional reduction step. The RMSE results of the proposed estimators are shown in Tables 4.16.

Table 4.16 RMSEs of estimators with respect to the UE in exponential model for a high-dimensional setting where $\alpha = 0.05$ and $\pi = 0.5$

κ	Estimators					
	RE	LS	PT	SP	S	S^+
0.005	1.0565	1.0402	1.0199	1.0126	1.0365	1.0378
0.050	0.4716	0.8458	1.0000	1.0000	1.0126	1.0126
0.100	0.2614	0.6788	1.0000	1.0000	1.0111	1.0111
0.200	0.0163	0.0645	1.0000	1.0000	1.0107	1.0107
0.300	0.0053	0.0209	1.0000	1.0000	1.0011	1.0011

The results of the post-selection step show that all proposed estimators achieved the maximum performance of $\kappa = 0.005$. Furthermore, the RE was most efficient when $\kappa = 0.005$, but it was less efficient than all other estimators when κ became large. Similarly, the performance of the LS estimator also decreased as κ increased. The RMSEs of both PT and SP estimators fell below one as κ became large. The S^+ estimator showed an improved version of the shrinkage estimator, both of which were more efficient than UE in all cases. They dominated the other estimators when κ was large and the RE was underfitted.

When the correctness of two alternative subsets of selected predictors obtained from LASSO and aLASSO was unknown, our RMSE results were strongly consistent with the theoretical and numerical results in the low-dimensional setting of the exponential model.

4.2.2.3 Monomolecular Model

For a high-dimensional sparse monomolecular model, the true values of parameter $\boldsymbol{\beta}$ in the simulation were set to $\boldsymbol{\beta} = (\boldsymbol{\beta}_s^\top, \boldsymbol{\beta}_w^\top, \boldsymbol{\beta}_n^\top)^\top$, where $\boldsymbol{\beta}_s = (3, 3, 1.2, 1.2, 1.2)^\top$, $\boldsymbol{\beta}_w = \boldsymbol{\kappa}_{p_w}$, and $\boldsymbol{\beta}_n = \mathbf{0}_{p_n}$. We also set κ as 0.001, 0.025, 0.050, 0.075, and 0.100 to study the estimator's behavior from a very weak to a moderate signal.

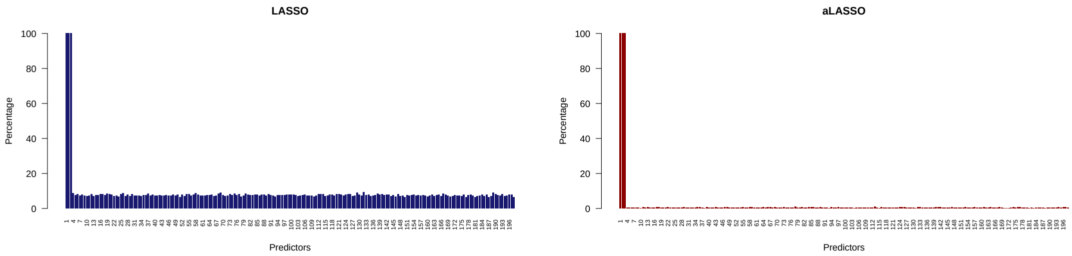
In the dimensional reduction step, the performance of the selecting variable methods was examined only for $(n, k_s, k_w, k_n) = (150, 3, 25, 170)$. Based on 2,000 simulation repetitions, the percentages of predictors selected by the LASSO and aLASSO strategies for each signal level are reported in Table 4.17 and the percentages of selection of each predictor using LASSO and aLASSO strategies are graphically represented in Figure 4.22.

Table 4.17 Selection percentages of predictors using LASSO and aLASSO methods in monomolecular model with strong, weak, and no signals for $(n, k_s, k_w, k_n) = (150, 3, 25, 170)$

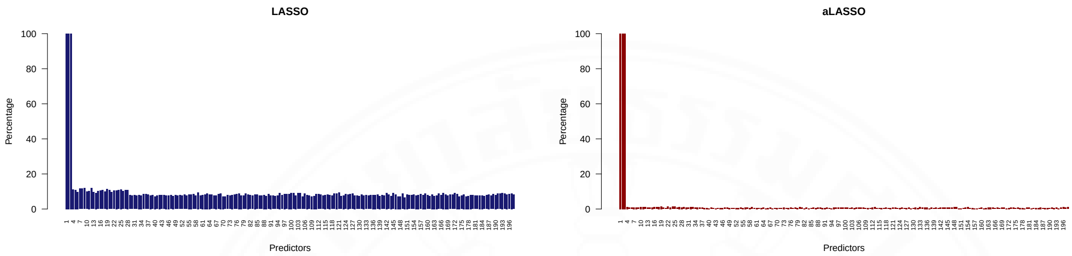
κ	Strong Signal		Weak Signal		No Signal	
	LASSO	aLASSO	LASSO	aLASSO	LASSO	aLASSO
0.001	100.00	100.00	7.71	0.47	7.54	0.49
0.025	100.00	100.00	10.67	1.01	8.08	0.67
0.050	100.00	99.95	22.49	3.37	10.66	0.98
0.075	100.00	99.95	38.02	8.49	13.73	1.57
0.100	100.00	99.93	53.58	16.37	17.54	2.30

The results show the LASSO strategy selected predictors with strong signals for all values of κ, while the performance of the aLASSO strategy decreased as κ increased. As κ increased, the performance in selecting predictors with weak signals of both LASSO and aLASSO increased, but the performance in eliminating predictors with no signals decreased. As we can see in Figure 4.22, the LASSO strategy selected too many predictors when κ was very small, which could produce an overfitted model, whereas aLASSO selected fewer significant predictors for large values of κ, which may produce an underfitted model.

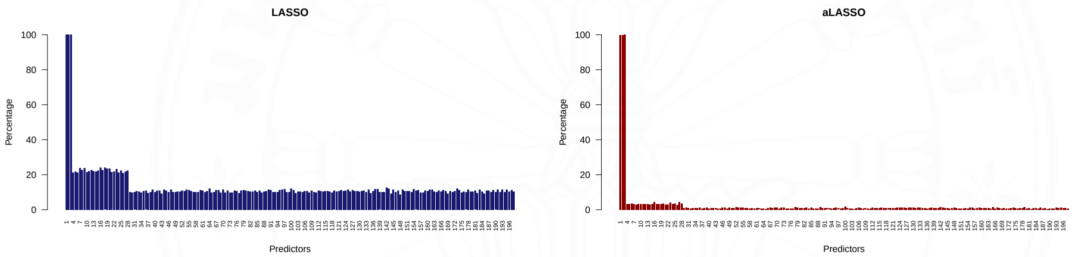
For the post-selection parameter estimation, we suggest parameter estimations based on LS, PT, SP, S, and S^+. To evaluate the RMSEs of the estimators, we determine $\alpha = 0.05$ and $\pi = 0.5$, with the findings reported in Table 4.18.



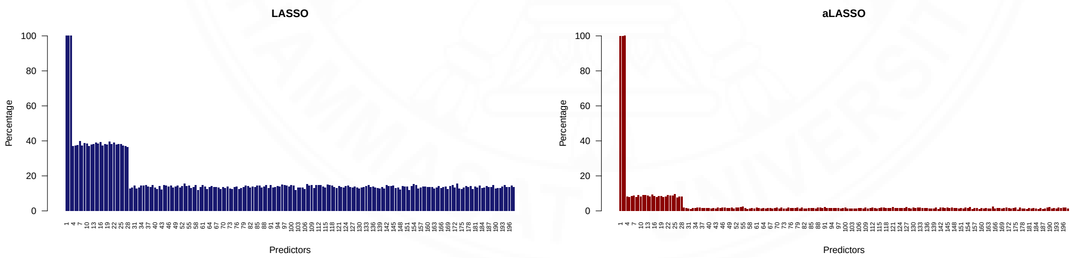
(a) $\kappa = 0.001$



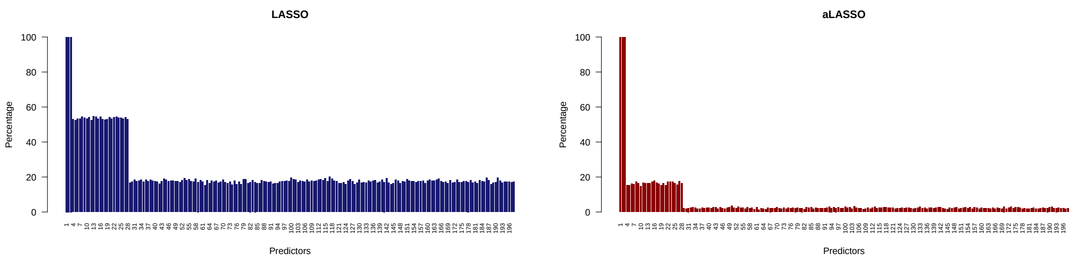
(b) $\kappa = 0.025$



(c) $\kappa = 0.050$



(d) $\kappa = 0.075$



(e) $\kappa = 0.100$

Figure 4.22 Comparison of percentage of each predictor selected using LASSO and aLASSO strategies in monomolecular model for each κ and $(n, k_s, k_w, k_n) = (150, 3, 25, 170)$

Table 4.18 RMSEs of estimators with respect to the UE in monomolecular model for a high-dimensional setting where $\alpha = 0.05$ and $\pi = 0.5$

κ	Estimators					
	RE	LS	PT	SP	S	S^+
0.001	1.7209	1.5584	1.0939	1.0565	1.2975	1.2975
0.025	0.2697	0.8513	1.0510	1.0343	1.1388	1.1388
0.050	0.1811	0.6584	1.0000	1.0000	1.0718	1.0718
0.075	0.0425	0.1705	1.0000	1.0000	1.0658	1.0658
0.100	0.0026	0.0103	1.0000	1.0000	1.0581	1.0581

According to Table 4.18, all estimators had the best performance when $\kappa = 0.001$, but their efficiency then decreased as κ increased. The RE had the highest RMSE at $\kappa = 0.001$. This indicates that the subset of the relevant predictors selected by aLASSO is the right subset, whereas, LASSO identified an overfitted model. On the other hand, aLASSO produced an underfitted model when κ increased, and the efficiency of the RE decreased.

The RMSEs of the post-selection proposed estimators were powerfully consistent with the results in a low-dimensional setting. The efficiency of the LS estimator decreased as κ increased. The performance of PT and SP estimators fell and then became equal to the UE when κ grew. As well, the RMSEs of S and S^+ estimators were the same values.

4.3 Application to Real Data

The suggested and penalized estimators were applied to the analysis of real data examples. In real analysis, the variable selection method helps to evaluate the UPI. We used variable selection procedures based on the Akaike information criterion (AIC), Bayesian information criterion (BIC), and LASSO. We also used the subspace information created by variable selection techniques in only the LS, PT, SP, S, and S^+ estimation processes.

Since the actual parameter values in the real data were unknown, we evaluated the performance of any estimator $\widehat{\boldsymbol{\beta}}_1^*$ of $\widehat{\boldsymbol{\beta}}_1$ with respect to the benchmark estimator $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ by using the simulated relative mean squares prediction error (RMSPE) in Equation (3.45).

4.3.1 Low-dimensional Data Setting

In low-dimensional data, we applied the proposed estimators such as RE, LS, PT, SP, S, S^+, and penalized estimators, i.e., LASSO, and aLASSO to analyze a manufacturing industry dataset, a cost of living and property price index dataset, and a cereal yield dataset for the Cobb-Douglas, exponential, and monomolecular models, respectively.

4.3.1.1 Manufacturing Industry Data

The manufacturing industry dataset was taken from the U.S. National Bureau of Economic Research (2016). This dataset is a joint effort between the National Bureau of Economic Research (NBER) and the U.S. Census Bureau's Center for Economic Studies (CES), and contains annual industry-level data from 1958 to 2011 with various variables. For evaluating the efficiency of the proposed estimators, we applied the Cobb-Douglas model to the manufacturing industry dataset.

The data used in the analysis included a sample of 415 industries from 2009. It contained a dependent variable of 5-factor total factor productivity (TFP) annual growth rate and 8 independent variables as shown in Table 4.19. The AIC, BIC, and LASSO were used to select variables. We considered the submodel generated from the AIC, BIC, or LASSO, only if it satisfied the condition $p_2 \geq 3$.

Table 4.19 List of variables for manufacturing industry data

Variable	Description
Dependent Variable	
Y	5-factor total factor productivity annual growth rate
Independent Variable	
X1	Total employment (1,000s)
X2	End-of-year inventories (millions of dollars)
X3	Cost of electric & fuels (millions of dollars)
X4	Real structures capital stock (millions of dollars)
X5	Deflator for total value of shipments 1997 = 1.0
X6	Deflator for total cost of materials 1997 = 1.0
X7	Deflator for total capital expenditure 1997 = 1.0
X8	Deflator for cost of electric and fuels 1997 = 1.0

To investigate the performance of the estimators, we sampled $m = 100$ bootstrap rows from the complete dataset with replacement $N = 1,000$ iterations and set $\alpha = 0.05$ and $\pi = 0.50$. As can be seen from Table 4.20, six, four, and six predictors were active in the AIC, BIC, and LASSO submodels, respectively. Since AIC and LASSO selection methods eliminated 2 predictors, we therefore considered only the BIC submodel for this case. Table 4.21 shows the RMSPE results for the proposed estimators when $\alpha = 0.05$, and $\pi = 0.50$.

Table 4.20 Variable selection results for manufacturing industry data

Method	p_1	p_2	Active Predictos
AIC	7	2	X2, X3, X4, X5, X6, X8
BIC	5	4	X2, X5, X6, X9
LASSO	7	2	X2, X3, X4, X5, X6, X8

Table 4.21 RMSPEs of estimators with respect to UE for manufacturing industry data

Method	Estimator							
	RE	LS	PT	SP	S	S^+	LASSO	aLASSO
BIC	3.9747	2.2966	2.9771	2.0068	1.8489	2.1838	1.1224	1.1248

The RE estimator had the largest RMSPE since this estimator assumes that the subspace information is accurate. The PT strategy dominated all other estimators, while the performance of LASSO was lower than that of the other estimators. The RMSPE of the RE estimator was superior to that of LASSO and aLASSO estimators, indicating that the LASSO and aLASSO methods eliminated too many significant predictors.

4.3.1.2 Cost of Living and Property Prices Indices Data

Numbeo is the world's largest database for cost of living and worldwide housing (real estate) prices. Numbeo is also a crowd-sourced global database of quality of life information, including perceived crime rates and healthcare quality among many other pieces of information.

We next applied the proposed estimators to a real dataset for an exponential regression model. The data used was a cost of living and property prices index dataset from mid-year 2020 (Numbeo, 2020), and 14 indices and 204 global cities were selected. Table 4.22 displays a dependent variable, the cost of living index, and 13 independent variables.

Table 4.22 List of variables for cost of living and property price indices data

Variable	Description
Dependent Variable	
Y	Cost of Living Index
Independent Variable	
X1	Price to Income Ratio
X2	Gross Rental Yield City Centre
X3	Price to Rent Ratio City Centre
X4	Mortgage as a Percentage of Income
X5	Affordability Index
X6	Purchasing Power Index
X7	Traffic Commute Time Index
X8	Pollution Index
X9	Climate Index
X10	Crime Index
X11	Health Care Index
X12	Inefficiency Index
X13	CO_2 Emission Index

The choice of model and efficient parameter estimation and prediction is important. In real analysis, the variable selection method helps to evaluate the UPI. We applied variable selection procedures based on the AIC, BIC, and LASSO. Table 4.23 shows the independent variables that are significantly influential on the response variable.

Table 4.23 Variable selection results for cost of living and property price indices data

Method	p_1	p_2	Active predictors
AIC	8	6	X1, X3, X4, X6, X7, X8, X9
BIC	7	7	X1, X3, X4, X6, X7, X8
LASSO	13	1	X1, X2, X3, X4, X5, X6, X7, X8, X9, X10, X11, X12

Since the LASSO method gets rid of only one independent variable from the model, in this case, we are also interested only in AIC and BIC submodels, for which $p_2 \geq 3$. We drew $m = 100$ bootstrap rows from the dataset with replacement and $N = 1,000$ replications. Here, α was set to 0.05 and π to 0.50.

The results in Table 4.24 demonstrate that the RE of both the AIC and BIC variable selection methods was the most efficient estimator. All estimators were superior to the UE. The RMSPE of the PT estimator was also greater than the RMSPE of all other estimators. Both S and S^+ estimators outperformed the LS,

Table 4.24 RMSPEs of estimators with respect to UE for cost of living and property price indices data

Method	Estimator							
	RE	LS	PT	SP	S	S^+	LASSO	aLASSO
AIC	3.8004	2.1817	3.2666	2.0408	2.3377	2.7715	1.0003	1.0003
BIC	4.0882	2.2932	3.2476	2.0683	2.4652	2.9044	1.0006	1.0006

SP, LASSO, and aLASSO estimators. The performance of both penalty estimators was equal and was inferior to all other estimators.

4.3.1.3 Cereal Yield Data

The Food and Agriculture Organization (2020) or FAO is a specialized agency of the United Nations (UN) that leads international efforts to defeat hunger. The statistical activities of the FAO cover the areas of agriculture, forestry and fisheries, land and water resources and uses, climate, environment, population, gender, nutrition, poverty, rural development, education, health, and many more. Also, they provide the world's largest database of food and agriculture statistics, and this is publicly available. In this section, we applied the proposed estimators to the real data for the monomolecular model with the cereal yield dataset.

The data that we used comprised 117 countries in 2016. From nine available variables in Table 4.25, the cereal yield was selected as the response variable and the others were regressors. Variable selection procedures based on AIC, BIC, and LASSO were applied, and the variable selection results for establishing candidate submodels are shown in Table 4.26. The RMSPE results with assumed $\alpha = 0.05$ and $\pi = 0.50$ using 100 resampled bootstrap samples that we iterated 1000 times are displayed in Table 4.27.

The results from Table 4.27 show that all of the variable selection methods provided the correct subspace information because the RE had the highest RMSPE. Aside from the RE, all other estimators outperformed the UE, except both penalized estimators in the BIC method. The LASSO and aLASSO estimators performed well in AIC and LASSO variable selections, but poorly in BIC. These results are consistent with simulation results when Δ^{sim} is equal or close to zero.

Table 4.25 List of variables for cereal yield data

Variable	Description
Dependent Variable	
Y	Cereal yield (hg/ha)
Independent Variable	
X1	Area harvested (ha)
X2	Gross Production Index
X3	Import Value (1,000 US$)
X4	Export Value (1,000 US$)
X5	Nitrogen (kg/ha)
X6	Phosphate (kg/ha)
X7	Potash (kg/ha)
X8	Pesticides (kg/ha)

Table 4.26 Variable selection results for cereal yield data

Method	p_1	p_2	Active Predictors
AIC	7	3	X1, X3, X4, X5, X7
BIC	6	4	X1, X3, X4, X5
LASSO	5	5	X4, X5, X7

Table 4.27 RMSPEs of estimators with respect to UE for cereal yield data

Method	Estimator							
	RE	LS	PT	SP	S	S^+	LASSO	aLASSO
AIC	1.0982	1.0676	1.0784	1.0531	1.0311	1.0373	1.0266	1.0274
BIC	1.0621	1.0429	1.0560	1.0387	1.0199	1.0354	0.9894	0.9884
LASSO	1.1517	1.1027	1.1224	1.0817	1.0617	1.0955	1.0725	1.0742

4.3.2 High-dimensional Data Setting

In high-dimensional data examples, we applied the penalized estimators for the variable selection step. In addition, the suggested estimators, such as the RE, LS, PT, SP, S, and S^+ estimators, were applied in practice to examine three datasets: an economic dataset, a dataset on communities and crime, and a dataset on rice yield for the Cobb-Douglas, exponential, and monomolecular models, respectively.

4.3.2.1 Economic Data

The World Bank compiles published statistics on global development, called World Development Indicators (WDI), drawing from officially recognized sources and including national, regional, and global estimates. The WDI that we used in this section is economic data, which contains the topics of growth, economic structure, income and savings, trade, and labor productivity (World Bank, 2018). This study aims to examine the economic factors associated with GDP per capita growth using a Cobb-Douglas sparse regression model in a high-dimensional regime. A total of 64 countries with all measurements on 101 economic factors for the year 2018 are recorded in this dataset.

The numbers of selected variables by LASSO and aLASSO for the economic data set are eleven and four variables, respectively, so that $k_1 = 4$ and $k_2 = 7$. The number of parameters in overfitted and underfitted models was $p_1 + p_2 = 12$ and $p_1 = 5$. After the dimension reduction step, the list of the response variable and explanatory variables selected by the penalized method (LASSO and aLASSO) are shown in Tables 4.28 and 4.29.

Table 4.28 List of variables for economic data

Variable	Description
Dependent Variable	
GDP.PCAP	GDP per capita growth (annual %)
Independent Variable	
ADJ.DRES	Adjusted savings: natural resources depletion (% of GNI)
AGR.TOTL	Agriculture, forestry, and fishing, value added (annual % growth)
CAB.XOKA	Current account balance (% of GDP)
IND.EMPL	Employment in industry, male (% of male employment)
GNP.PCAP	GNI per capita growth (annual %)
CON.PRVT	Households and NPISHs Final consumption per expenditure capita growth (annual %)
IND.TOTL	Industry (including construction), value added (annual % growth)
TLF.CACT	Labor force participation rate, male (% of male population ages 15+)
TDS.DPPG	Public and publicly guaranteed debt service (% of exports of goods, services and primary income)
SRV.TOTL	Services, value added (annual % growth)
EMP.WORK	Wage and salaried workers, female (% of female employment)

Table 4.29 Variable selection results for economic data

Model	Method	Number of Parameters	Selected Predictors as Active
OF	LASSO	12	ADJ.DRES, AGR.TOTL, CAB.XOKA, IND.EMPL, GNP.PCAP, CON.PRVT, IND.TOTL, TLF.CACT, TDS.DPPG, SRV.TOTL, EMP.WORK
UF	aLASSO	5	ADJ.DRES, GNP.PCAP, IND.TOTL, SRV.TOTL

We drew bootstrap samples of size $m = 50$ rows with replacement from the data by $N = 1,000$ times to examine the relative efficiencies of the proposed estimators. We assumed $\alpha = 0.05$ and $\pi = 0.50$. The RMSPE results are reported in Table 4.30.

Table 4.30 RMSPEs of estimators with respect to UE from post-selection for economic data

	Estimators					
	RE	LS	PT	SP	S	S^+
RMSPE	18.9631	3.5600	1.0000	1.0000	1.2195	1.2195

As can be seen, all the estimators dominated the UE except the PT and SP which were equivalent to the UE. The performance of the RE was the best, followed by the LS estimator and both shrinkage estimators, indicating that the information $\boldsymbol{\beta} = \mathbf{0}$ was still correct. This means that aLASSO provided a more accurate subset of selected independent variables, whereas LASSO chose too many no-influence variables.

4.3.2.2 Communities and Crime Data

The communities and crime dataset within the United States consists of socio-economic data from the 1990 US Census, law enforcement data from the 1990 US Law Enforcement Management and Admin Stats (LEMAS) survey, and crime data from the 1995 FBI UCR (UCI Machine Learning Repository, n.d.). After we excluded missing values, the data consisted of 111 instances or crimes reported from across the country and 124 predictive features. Moreover, the potential feature of interest to predict was the total number of violent crimes per 100K population. We then analyzed the data using a high-dimensional exponential sparse regression model.

The LASSO and aLASSO were put to practical use to reduce the predictor dimension. As a result, the LASSO obtained a model that contained 11 ($k_1 + k_2 = 11$) significant predictors and aLASSO selected seven ($k_1 = 7$) significant features into the model (in Table 4.32). Therefore, the overfitted and underfitted models consisted of $p_1 + p_2 = 12$ and $p_1 = 8$ parameters. After variable selection, the list of variables that we considered is shown in Table 4.31.

Table 4.31 List of variables for communities and crime data

Variable	Description
Dependent Variable	
ViolentCrimesPerPop	Total number of violent crimes per 100K popuation
Independent Variable	
pctUrban	Percentage of people living in areas classified as urban
FemalePctDiv	Percentage of females who are divorced
PctKids2Par	Percentage of kids in family housing with two parents
PctTeen2Par	Percentage of kids age 12-17 in two parent households
PctKidsBornNeverMar	Percentage of kids born to never married
PctHousOccup	Percentage of housing occupied
MedRentPctHousInc	Median gross rent as a percentage of household income
PctForeignBorn	Percentage of people foreign born
PctSameHouse85	Percentage of people living in the same house as in 1985
NumKindsDrugsSeiz	Number of different kinds of drugs seized
LemasPctOfficDrugUn	Percentage of officers assigned to drug units

Table 4.32 Variable selection results for communities and crime data

Model	Method	Number of Parameters	Selected Predictors as Active
OF	LASSO	12	pctUrban, FemalePctDiv, PctKids2Par, PctTeen2Par, PctKidsBornNeverMar, PctHousOccup, MedRentPctHousInc, PctForeignBorn, PctSameHouse85, NumKindsDrugsSeiz, LemasPctOfficDrugUn
UF	aLASSO	8	pctUrban, FemalePctDiv, PctKids2Par, PctTeen2Par, PctKidsBornNeverMar, PctHousOccup, MedRentPctHousInc

Table 4.33 RMSPEs of estimators with respect to UE from post-selection for communities and crime data

	Estimators					
	RE	LS	PT	SP	S	S^+
RMSPE	1.2498	5.2741	1.2498	5.2741	6.0502	7.1406

Table 4.33 shows the RMSPEs results where α was set to 0.05 and π was also set to 0.05 with 100 bootstrap rows and 1,000 iterations. All the proposed estimators were superior to the UE, which means that LASSO provided the poorest set or selected too many irrelevant predictors in the OF model. The shrinkage estimators performed best, especially in the positive-part version. The RMSPEs of the RE and PT

estimator had equal values since the information of RE was correct for all iterations. Similarly, the performance of the LS and SP estimators was also equivalent too. This indicates that the test statistic lay in an acceptance region ($H_0 : \boldsymbol{\beta} = \mathbf{0}$) for all iterations.

4.3.2.3 Rice Yield Data

Data on the cultivation of the sample farmer households in the benefit areas of Kwae Noy Dam, Thailand for the 2014/2015 planting year was obtained from the socio-economic situation monitoring and evaluation projects of the Royal Irrigation Department of Thailand, under His Majesty the King of Thailand's initiation, in the 2015 budget year (Chaowagul et al., 2015). Our study aimed to establish a model to predict the average rice yield (kg/0.16 ha). The predictors contained 140 variables and comprised 105 sample households which planted rice twice a year. The monomolecular sparse regression model in a high-dimensional setting was used.

After applying the LASSO and aLASSO strategies for variable selection, five predictors ($k_1 + k_2 = 5$) were selected as significant by LASSO, and two relevant predictors ($k_1 = 2$) by aLASSO, containing $p_1 + p_2 = 7$ and $p_1 = 4$ parameters, respectively. The list of variables and the variable selection results are displayed in Tables 4.34 and 4.35. We report the RMSPEs of the suggested estimators for $\alpha = 0.05$ and $\pi = 0.50$ with $m = 100$ and $N = 1000$ in Table 4.36.

From Table 4.36, we can see that variable selection based on aLASSO provided the correct set of RE as it had the highest RMSPE. The performance of the PT estimator was superior to that of the LS, SP, S, and S$^+$ estimators. The efficiency of the S$^+$ estimator was also better than the shrinkage estimator.

Table 4.34 List of variables for rice yield data

Variable	Description
Dependent Variable	
farmh	Average yield of rice harvested (kg/0.16 ha)
Independent Variable	
chemrg2p	Average price (Thai baht) of herbicides in powder or tablet form
chemrw2q	Average amount of liquid pesticide used (L/0.16 ha)
chemrw2ex	Average cost of liquid pesticides (Thai baht/0.16 ha)
labr13p	Average labor rate for fertilizing rice (Thai baht)
machin2	Number (units) of available rice spray seeding machines in 2014

Table 4.35 Variable selection results for rice yield data

Model	Method	Number of Parameters	Selected Predictors as Active
OF	LASSO	7	chemrg2p, chemrw2q, chemrw2ex, labr13p, machin2
UF	aLASSO	4	chemrw2q, chemrw2ex

Table 4.36 RMSPEs of estimators with respect to UE from post-selection for rice yield data

	Estimators					
	RE	LS	PT	SP	S	S^+
RMSPE	1.0264	1.0175	1.0205	1.0134	1.0066	1.0101

These results of these three real data examples were consistent with the simulation results, which confirmed that the positive-part shrinkage estimator was robust when the correctness of two alternative models obtained from LASSO and adaptive LASSO (aLASSO) variable selection methods was unknown.

CHAPTER 5
PARAMETER ESTIMATION IN THE COX PROPORTIONAL HAZARD REGRESSION MODEL UNDER UNCERTAINTY OF PRIOR INFORMATION

This chapter represents the estimation problem in the Cox proportional hazards regression model when the regression coefficient may be restricted to a subspace in both low- and high-dimensional settings. The classical, preliminary test, and shrinkage estimation strategies are considered, and the large sample properties of these estimators in terms of asymptotic distributional quadratic bias and risk are presented. We also examine two penalized estimators—LASSO and adaptive LASSO—and compare their relative performance numerically with the suggested estimators. The properties of all the estimators are compared through simulated relative mean squared error using a Monte Carlo simulation. Finally, real data examples are applied to illustrate the usefulness of the suggested estimators in practice.

5.1 Introduction

Data that measure lifetime or the length of time until the occurrence of an event are called lifetime, failure time, or survival data. This is an essential topic in many areas, including medicine, engineering, and social sciences. Lifetime data can be defined broadly as the time to the occurrence of a given event. For example, this event can be developing a disease, response to treatment, relapse, or death. Therefore, lifetime data can be the time from the start of treatment to response, time to death, the length of time a person stays on a job, etc.

The statistics used for analyzing lifetime data are called survival analysis. Unfortunately, this analysis is usually a difficult process due to censoring (right censoring is the most common type), which means that participants drop out of the study before the occurrence of the event, leaving incomplete information about the survival time when the study ends. This problem results in a lack of information such that ordinary linear regression models for survival cannot be applied (Baek et al., 2021). Hence, the Cox proportional hazards (PH) regression model, is the most commonly used hazard model in the medical field since it deals with censoring.

In this study, we consider the estimation problem for the Cox PH regression model when there are many potential predictors, some of which may or may not be relevant to the lifetime of patients. If the irrelevant predictors can be removed from the model, the analysis will be more precise. In this case, we can use the advantage of available information derived from the researcher's experience, previous studies, or variable selection approaches to identify the related predictors. Such information is usually called prior information or subspace information.

There are two types of competitors to consider when using available information, which is usually of unknown correctness or uncertain prior information (UPI): the full model (unrestricted model), which includes all predictors, and the candidate submodel (restricted model), which contains the influencing predictors. From this information, the parameter vector can be partitioned into two subvectors, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$, where one parameter subvector corresponds to active predictors and other to inactive predictors. In this study, our point of interest is on the estimation of active parameters ($\boldsymbol{\beta}_1$) when information suggests that the inactive parameters ($\boldsymbol{\beta}_2$) are close to zero.

Under UPI, the full model considers all predictors, which leads to an overfitting problem, as too many inactive parameters are included. If the UPI is correct, the submodel excludes non-significant predictors and contains only active predictors, which can address an overfitting problem. However, we may encounter an underfitting problem if the UPI is incorrect. Therefore, employing either the unrestricted or restricted estimator as the estimator for an active parameter is not a good decision when information accuracy is unknown. Consequently, we introduce preliminary test and shrinkage strategies to address this problem. In addition, a family of penalized approaches for the Cox PH regression model was considered, including the least absolute shrinkage and selection operation (LASSO) and adaptive LASSO (aLASSO) methods.

Many studies have examined the use of preliminary test and shrinkage estimation strategies under UPI in regression models with censored or time-to-event data. The shrinkage and positive-part shrinkage estimators were proposed in Ahmed and Saleh (1999), and a preliminary test estimator was proposed in Khan (2002) for an exponential model with censoring. Hossain and Ahmed (2014) studied the shrinkage and positive-part shrinkage estimation strategies in the Cox proportional hazards regression model. In addition, the two penalized estimators (LASSO and aLASSO) were also considered. Ahmed et al. (2012), Hossain and Howlader (2017), and Hossain and Khan (2020) addressed the problem of estimating parameters in Weibull regression, lognormal regression, and exponentiated Weibull regression models, respectively, for time-to-event data (censored data). The shrinkage and positive-part shrinkage estimation methods were introduced in their study. Moreover, the LASSO estimation strategy was suggested in Ahmed et al. (2012), and both LASSO and aLASSO estimation methods were suggested in Hossain and Howlader (2017).

Therefore, this chapter aims to diagnose the issues related to parameter estimation for the Cox PH regression model when a candidate submodel is available. This model implements preliminary test, shrinkage, and penalized estimation methods for low-dimensional data and extends the work to high-dimensional data.

The rest of the chapter is organized as follows: The Cox proportional hazards regression model is introduced in Section 5.2. The suggested and penalized estimation strategies is displayed in Section 5.3. The asymptotic properties of the suggested estimators and their asymptotic distributional quadratic biases and risks are represented in Section 5.4. The results of a Monte Carlo simulations study are given in Section 5.5. An application to real datasets is displayed in Section 5.6. Finally, the concluding remarks are given in Section 5.7.

5.2 Cox Proportional Hazards Model and Maximum Partial Likelihood Estimation

Models in which covariates have a multiplicative effect on the hazard function play an important role in analyzing lifetime (or survival) data. Therefore, the primary approach to regression modeling for lifetimes typically examines the relationship of the hazard function to covariates. The most common model of this approach is the proportional hazards (PH) model.

Survival analysis is a commonly used method for analyzing failure time, where failure or death is referred to as an event. This analysis tries to model time-to-event data, which is usually censored due to a study's termination. We consider a situation such that lifetimes T_i may be subject to a fixed censoring, and we suppose that there is a lifetime and a censoring time for each individual. Assuming that each individual has a fixed potential censoring time $C_i > 0$ and we have a sample size of n, the data therefore consists of the triple $(t_i, d_i, \boldsymbol{x}_i)$ for $i = 1, 2, ..., n$. Here, $t_i = min(T_i, C_i)$ is the time in the study for the ith individual and d_i is the event indicator of individual i, where $d_i = 1$ if the event has occurred ($T_i \leq C_i$) and $d_i = 0$ if the lifetime is right-censored ($T_i > C_i$). This indicates that t_i is the lifetime or censoring time according to whether $d_i = 1$ or 0, respectively. In addition, $\boldsymbol{x}_i = \{x_{i1}, x_{i2}, ..., x_{ip}\}^\top$ is a $p \times 1$ vector of fixed covariates for the ith individual.

The general form of the hazard function of T depends on the covariate $\boldsymbol{x}_i$ for an individual, is written as:

$$h(t|\boldsymbol{x}_i, \boldsymbol{\beta}) = h_0(t)r(\boldsymbol{x}_i, \boldsymbol{\beta}), \tag{5.1}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_p)^\top$ is a $p \times 1$ vector of regression coefficients. Here, $h_0(t)$ and $r(\boldsymbol{x}_i, \boldsymbol{\beta})$ are positive-value functions. The function $h_0(t)$ is usually called the baseline hazard function and is the hazard function for an individual whose covariate vector $\boldsymbol{x}_i$ is such that $r(\boldsymbol{x}_i, \boldsymbol{\beta}) = 1$. The function $r(\boldsymbol{x}_i, \boldsymbol{\beta})$ characterizes how the hazard function changes with covariates, in which case $h_0(t)$ is the hazard function when $\boldsymbol{x}_i = \mathbf{0}$.

In contrast, Cox (1972) introduced the proportional hazards model which leaves the baseline hazard function $h_0(t)$ unspecified, and a typical specification for $r(\boldsymbol{x}_i, \boldsymbol{\beta})$ is $\exp(\boldsymbol{\beta}^\top \boldsymbol{x}_i)$. Consequently, the hazard function for T given $\boldsymbol{x}_i$ takes the form

$$\begin{aligned} h(t|\boldsymbol{x}_i, \boldsymbol{\beta}) &= h_0(t)\exp(\boldsymbol{\beta}^\top \boldsymbol{x}_i) \\ &= h_0(t)\exp\left(\sum_{j=1}^{p} \beta_j x_{ij}\right) \\ &= h_0(t)\exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}), \end{aligned} \tag{5.2}$$

known as the Cox model or Cox PH model. This model is semi-parametric because the baseline hazard function can take any form and must be estimated nonparametrically. At the same time, the covariates enter the model linearly with no intercept term and estimate $\boldsymbol{\beta}$ parametrically.

The name "proportional hazards" comes from the fact that any two individuals have hazard functions that are constant multiples of one another. Consider two cases

i and i' that differ in their x-values, $\boldsymbol{x}_i$ and $\boldsymbol{x}_{i'}$, with the corresponding linear predictors

$$\boldsymbol{\beta}^\top \boldsymbol{x}_i = \beta_1 \boldsymbol{x}_{i1} + \beta_2 \boldsymbol{x}_{i2} + \cdots + \beta_p \boldsymbol{x}_{ip}$$

and

$$\boldsymbol{\beta}^\top \boldsymbol{x}_{i'} = \beta_1 \boldsymbol{x}_{i'1} + \beta_2 \boldsymbol{x}_{i'2} + \cdots + \beta_p \boldsymbol{x}_{i'p}.$$

The hazard ratio for these two cases is as follows:

$$\begin{aligned}\frac{h(t|\boldsymbol{x}_i, \boldsymbol{\beta})}{h(t|\boldsymbol{x}_{i'}, \boldsymbol{\beta})} &= \frac{h_0(t)\exp(\boldsymbol{\beta}^\top \boldsymbol{x}_i)}{h_0(t)\exp(\boldsymbol{\beta}^\top \boldsymbol{x}_{i'})} \\ &= \frac{\exp(\boldsymbol{\beta}^\top \boldsymbol{x}_i)}{\exp(\boldsymbol{\beta}^\top \boldsymbol{x}_{i'})}, \\ &= \exp\left[\boldsymbol{\beta}^\top (\boldsymbol{x}_i - \boldsymbol{x}_{i'})\right], \end{aligned} \tag{5.3}$$

which is independent of time. Therefore, the Cox model is a proportional hazards model. See Cox and Oakes (1984, p. 91), Klein and Moeschberger (2003, pp. 243–245), Lawless (2003, p. 341), Lee and Wang (2003, pp. 298–301), among others, for more details on Cox proportional hazards model.

5.2.1 Partial Likelihoods

Suppose that a censored random sample $(t_i, d_i), i = 1, 2, ..., n$, yields k distinct observed lifetimes and $n - k$ censoring times. For simplicity, assume that we have an absolutely continuous failure distribution. Let $t_{(1)} < t_{(2)} < \cdots < t_{(k)}$ denote the ordered event times and $\boldsymbol{x}_{(i)}$ be the covariate associated with the individual whose failure time is $t_{(i)}$ or who dies at time $t_{(i)}$. Let $R(t_{(i)})$ be the set of individuals who are still under study at a time (alive) and uncensored just prior to time $t_{(i)}$, and this is referred to as the risk set at $t_{(i)}$; since it consists of those individuals who could be observed to die at $t_{(i)}$, given what has occurred up to that time.

Therefore, Cox's partial likelihood function (Cox, 1972), based on the hazard function as specified by Equation (5.2), for estimating $\boldsymbol{\beta}$ is expressed by

$$\mathcal{L}(\boldsymbol{\beta}) = \prod_{i=1}^{k} \frac{\exp\left(\boldsymbol{\beta}^\top \boldsymbol{x}_{(i)}\right)}{\sum\limits_{j \in R(t_{(i)})} \exp\left(\boldsymbol{\beta}^\top \boldsymbol{x}_j\right)}. \tag{5.4}$$

This is retained as an ordinary likelihood, and inference is carried out by a usual method. Note that the numerator of the likelihood depends only on information from the individual who experiences the event. In contrast, the denominator uses information about all individuals who have not yet experienced the event, including those who will be censored later.

Let $\ell(\boldsymbol{\beta}) = \ln[\mathcal{L}(\boldsymbol{\beta})]$, the log partial likelihood function from Equation (5.4) can be written as follows:

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^{k} \boldsymbol{\beta}^{\top}\boldsymbol{x}_{(i)} - \sum_{i=1}^{k} \ln\left[\sum_{j\in R(t_{(i)})} \exp\left(\boldsymbol{\beta}^{\top}\boldsymbol{x}_j\right)\right]. \tag{5.5}$$

The score equations are found by taking partial derivatives of $\ell(\boldsymbol{\beta})$ with respect to the $\boldsymbol{\beta}$'s as follows:

$$\frac{\partial \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{k} \boldsymbol{x}_{(i)} - \sum_{i=1}^{k} \frac{\sum\limits_{j\in R(t_{(i)})} \boldsymbol{x}_j \exp\left(\boldsymbol{\beta}^{\top}\boldsymbol{x}_j\right)}{\sum\limits_{j\in R(t_{(i)})} \exp\left(\boldsymbol{\beta}^{\top}\boldsymbol{x}_j\right)}. \tag{5.6}$$

The maximum (partial) likelihood estimates are obtained by solving the set of p nonlinear score equations $\frac{\partial \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = 0$. Unfortunately, it cannot be solved since it is nonlinear in $\boldsymbol{\beta}$. Therefore, this can be done numerically using a Newton-Raphson iteration or some other iterative method. Note that Equation (5.5) does not depend upon the baseline hazard ratio, so inferences may be made on the effects of the covariates without knowing the baseline hazard ratio.

The second derivative of the log partial likelihood function in Equation (5.5) with respect to $\boldsymbol{\beta}$ is the following expression:

$$\begin{aligned}\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}^{\top}} = &-\sum_{i=1}^{k} \frac{\sum\limits_{j\in R(t_{(i)})} \boldsymbol{x}_j \mathrm{x}_j^{\top} \exp(\boldsymbol{\beta}^{\top}\boldsymbol{x}_j)}{\sum\limits_{j\in R(t_{(i)})} \exp(\boldsymbol{\beta}^{\top}\boldsymbol{x}_j)} \\ &+ \sum_{i=1}^{k} \frac{\sum\limits_{j\in R(t_{(i)})} \boldsymbol{x}_j \exp(\boldsymbol{\beta}^{\top}\boldsymbol{x}_j)}{\sum\limits_{j\in R(t_{(i)})} \exp(\boldsymbol{\beta}^{\top}\boldsymbol{x}_j)} \frac{\sum\limits_{j\in R(t_{(i)})} \boldsymbol{x}_j^{\top} \exp(\boldsymbol{\beta}^{\top}\boldsymbol{x}_j)}{\sum\limits_{j\in R(t_{(i)})} \exp(\boldsymbol{\beta}^{\top}\boldsymbol{x}_j)}.\end{aligned} \tag{5.7}$$

The negative of the second derivative of the log partial likelihood in Equation (5.7) containing more than one covariate is called the information matrix, and can be denoted as

$$\mathcal{I}(\boldsymbol{\beta}) = -\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}^{\top}}. \tag{5.8}$$

Therefore, the estimator of the variance of the estimated coefficient is the inverse of Equation (5.8) evaluated at $\widehat{\boldsymbol{\beta}}$ and is

$$\mathrm{Var}(\widehat{\boldsymbol{\beta}}) = \mathbb{V}(\widehat{\boldsymbol{\beta}}) = \mathcal{I}(\widehat{\boldsymbol{\beta}})^{-1}. \tag{5.9}$$

More details of the partial likelihood function and Newton-Raphson technique can be found in Hosmer and Lemeshow (1998, pp. 95-97), Klein and Moeschberger (2003, pp. 253–254), Lawless (2003, pp. 342–343), Lee and Wang (2003, pp. 301–302), etc.

5.2.2 Maximum Partial Likelihood Estimation Strategy

In this study, the proportional hazards (PH) regression model was considered when the lifetime of individuals may be related to several potential covariates, some of which may be irrelevant. The Cox PH regression model that we consider is of the form

$$h(t) = h_0(t) \exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}), \qquad t \geq 0. \tag{5.10}$$

The parameter vector $\boldsymbol{\beta}$ can be partitioned into two sub-vectors as $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_1^\top)^\top$, where $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are supposed to have dimensions $p_1 \times 1$ and $p_2 \times 1$, respectively, such that $p = p_1 + p_2$. For inference purposes, when there is censoring, the information matrix, may be partitioned as

$$\mathcal{I}(\boldsymbol{\beta}) = \begin{bmatrix} \mathcal{I}_{11} & \mathcal{I}_{12} \\ \mathcal{I}_{21} & \mathcal{I}_{22} \end{bmatrix}, \tag{5.11}$$

which has a $p \times p$ dimension. We also assume $\boldsymbol{G} = \lim_{n \to \infty} \frac{1}{n} \mathcal{I}(\boldsymbol{\beta})$ as $n \longrightarrow \infty$, then we get a finite positive-definite matrix as

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{G}_{11} & \boldsymbol{G}_{12} \\ \boldsymbol{G}_{21} & \boldsymbol{G}_{22} \end{bmatrix}, \tag{5.12}$$

where $\boldsymbol{G}_{ij} = \lim_{n \to \infty} \frac{1}{n} \mathcal{I}_{ij}$ and $i, j = 1, 2$.

5.2.2.1 Unrestricted Estimator

The unrestricted estimator (UE) of $\boldsymbol{\beta}$, denoted as $\widehat{\boldsymbol{\beta}}^{\mathrm{UE}}$, is the final maximum partial likelihood estimator and is obtained by solving the nonlinear score equation using the Newton-Raphson iterative algorithm or other methods.

Theorem 5.2.1. *Under the usual regularity conditions, as $n \longrightarrow \infty$,*

$$\widehat{\boldsymbol{\beta}}^{UE} \xrightarrow{D} \mathcal{N}_p\left(\boldsymbol{\beta}, \frac{1}{n}\boldsymbol{G}^{-1}\right), \tag{5.13}$$

where $\boldsymbol{G}^{-1}$ *is a* $p \times p$ *asymptotic variance-covariance matrix and* $\xrightarrow{D}$ *implies convergence in distribution. Here,* $\boldsymbol{G} = \lim_{n\to\infty} \frac{1}{n}\mathcal{I}(\boldsymbol{\beta})$*, where* $\mathcal{I}(\boldsymbol{\beta})$ *is the information matrix, which is the negative of the second derivative of the log partial likelihood in Equation* (5.7)*:*

$$\mathcal{I}(\boldsymbol{\beta}) = \sum_{i=1}^{k} \frac{\sum\limits_{j\in R(t_{(i)})} \boldsymbol{x}_j \mathrm{x}_j^\top \exp(\boldsymbol{\beta}^\top \boldsymbol{x}_j)}{\sum\limits_{j\in R(t_{(i)})} \exp(\boldsymbol{\beta}^\top \boldsymbol{x}_j)} - \sum_{i=1}^{k} \frac{\sum\limits_{j\in R(t_{(i)})} \boldsymbol{x}_j \exp(\boldsymbol{\beta}^\top \boldsymbol{x}_j)}{\sum\limits_{j\in R(t_{(i)})} \exp(\boldsymbol{\beta}^\top \boldsymbol{x}_j)} \frac{\sum\limits_{j\in R(t_{(i)})} \boldsymbol{x}_j^\top \exp(\boldsymbol{\beta}^\top \boldsymbol{x}_j)}{\sum\limits_{j\in R(t_{(i)})} \exp(\boldsymbol{\beta}^\top \boldsymbol{x}_j)}. \tag{5.14}$$

Proof. See Andersen and Gill (1982) for detailed proof. □

Since we set the vector of regression coefficients as $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$, then the unrestricted estimator is given by

$$\begin{aligned}\widehat{\boldsymbol{\beta}}^{\mathrm{UE}} &= (\underbrace{\widehat{\beta}_1^{\mathrm{UE}}, \widehat{\beta}_2^{\mathrm{UE}}, \ldots, \widehat{\beta}_{p_1}^{\mathrm{UE}}}_{p_1}, \underbrace{\widehat{\beta}_{p_1+1}^{\mathrm{UE}}, \widehat{\beta}_{p_1+2}^{\mathrm{UE}}, \ldots, \widehat{\beta}_{p_2}^{\mathrm{UE}}}_{p_2}) \\ &= \left((\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}})^\top, (\widehat{\boldsymbol{\beta}}_2^{\mathrm{UE}})^\top\right)^\top, \end{aligned} \tag{5.15}$$

where $p = p_1 + p_2$.

Theorem 5.2.2. *If the usual regularity conditions and Theorem 5.2.1 hold, as* $n \to \infty$*, the marginal distribution of* $\widehat{\boldsymbol{\beta}}_1^{UE} \xrightarrow{D} \mathcal{N}_{p_1}\left(\boldsymbol{\beta}_1, \frac{1}{n}\boldsymbol{G}_{11.2}^{-1}\right)$ *and of* $\widehat{\boldsymbol{\beta}}_2^{UE} \xrightarrow{D} \mathcal{N}_{p_2}\left(\boldsymbol{\beta}_2, \frac{1}{n}\boldsymbol{G}_{22.1}^{-1}\right)$*. Here,* $\boldsymbol{G}_{11.2}^{-1} = \left(\boldsymbol{Q}_{11} - \boldsymbol{G}_{12}\boldsymbol{G}_{22}^{-1}\boldsymbol{G}_{21}\right)^{-1}$*,* $\boldsymbol{G}_{22.1}^{-1} = \left(\boldsymbol{G}_{22} - \boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\right)^{-1}$*, and* $\xrightarrow{D}$ *means convergence in distribution.*

Proof. See Ravishanker and Dey (2001, p. 155) for detailed proof. □

5.2.2.2 Restricted Estimator

The restricted estimator (RE) of $\boldsymbol{\beta}$, denoted by $\widehat{\boldsymbol{\beta}}^{\mathrm{RE}}$, can be obtained by maximizing the log partial likelihood function in (5.5) under the linear restriction $\boldsymbol{R\beta} - \boldsymbol{r} = \boldsymbol{0}$. Using the Lagrange multiplier technique, the RE would be

$$\widehat{\boldsymbol{\beta}}^{\mathrm{RE}} = \widehat{\boldsymbol{\beta}}^{\mathrm{UE}} - \mathcal{I}(\boldsymbol{\beta})^{-1}\boldsymbol{R}^\top\left(\boldsymbol{R}\mathcal{I}(\boldsymbol{\beta})^{-1}\boldsymbol{R}^\top\right)^{-1}(\boldsymbol{R}\widehat{\boldsymbol{\beta}}^{\mathrm{UE}} - \boldsymbol{r}). \tag{5.16}$$

If there is reason to believe that $\boldsymbol{\beta}_2$ is close to zero, as a special case we can consider $R = [\boldsymbol{0}_{p_2\times p_1}, \boldsymbol{I}_{p_2}]$ and $\boldsymbol{r} = \boldsymbol{0}_{p_2\times 1}$. Thus, the null hypotheses

$\boldsymbol{R\beta} = \mathbf{0}$ become $\boldsymbol{\beta}_2 = \mathbf{0}$. From Lawless & Singhal (1978), it can be shown that the RE of $\boldsymbol{\beta}_1$ or $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ is as follows:

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} &= \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - (-\mathcal{I}_{11}^{-1}\mathcal{I}_{12}\widehat{\boldsymbol{\beta}}_2^{\mathrm{UE}}) \\ &= \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \boldsymbol{\omega}_n\widehat{\boldsymbol{\beta}}_2^{\mathrm{UE}}.\end{aligned} \tag{5.17}$$

Here, $\boldsymbol{\omega}_n = -\mathcal{I}_{11}^{-1}\mathcal{I}_{12}$ and we assume that $\boldsymbol{\omega}_n \xrightarrow{P} \boldsymbol{\omega} = -\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}$ as $n \to \infty$, where $\xrightarrow{P}$ indicates convergence in probability. In the simulation, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ can also be obtained using the Newton-Raphson iterative or other method under the restriction $\boldsymbol{\beta}_2 = \mathbf{0}$.

5.2.2.3 Large Sample Test Statistic

In order to test $\mathrm{H}_0 : \boldsymbol{\beta}_2 = \mathbf{0}$ versus $\mathrm{H}_1 : \boldsymbol{\beta}_2 \neq \mathbf{0}$, the likelihood ratio test statistic is defined by

$$\mathscr{L}_n = -2\ln\left[\frac{\mathcal{L}(\widehat{\boldsymbol{\beta}}^{\mathrm{RE}})}{\mathcal{L}(\widehat{\boldsymbol{\beta}}^{\mathrm{UE}})}\right] = 2\left[\ell(\widehat{\boldsymbol{\beta}}^{\mathrm{UE}}) - \ell(\widehat{\boldsymbol{\beta}}^{\mathrm{RE}})\right], \tag{5.18}$$

where $\ell(\widehat{\boldsymbol{\beta}}^{\mathrm{UE}})$ and $\ell(\widehat{\boldsymbol{\beta}}^{\mathrm{RE}})$ are the values of the log-likelihood of the UE and RE, respectively. Here, $\mathscr{L}_n$ is approximately $\chi^2_{p_2}$ under the null hypothesis for large n.

5.3 Various Estimation Strategies

In this section a variety of estimation strategies for parameter vector $\boldsymbol{\beta}_1$ are explored.

5.3.1 Strategy 1: Linear Shrinkage Estimator

The linear shrinkage (LS) estimator derived by taking a linear combination of the competing UE and RE is defined by

$$\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \pi(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}), \tag{5.19}$$

where π is a coefficient reflecting the degree of distrust in the prior information and $\pi \in (0, 1)$.

5.3.2 Strategy 2: Preliminary Test Estimator

The preliminary test (PT) estimator is obtained by substituting π with $I(\mathscr{L}_n \leq l_\alpha)$ in the LS estimator to give a random weighting. Therefore, the result is

$$\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - (\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) I(\mathscr{L}_n \leq \lambda_\alpha), \tag{5.20}$$

where λ_α is the α-level critical value for the test statistic $\mathscr{L}_n$ and $I(\cdot)$ is an indicator function.

5.3.3 Strategy 3: Shrinkage Preliminary Test Estimator

To improve the PT estimator in terms of α, replacing the RE with the LS estimator in the PT estimator is called the shrinkage preliminary test (SP) estimator:

$$\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - (\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) I(\mathscr{L}_n \leq \lambda_\alpha). \tag{5.21}$$

The performance of this strategy is much better than that of the PT strategy in a large portion of the parameter space.

5.3.4 Strategy 4: Stein-Type Shrinkage Estimator

The Stein-type shrinkage or shrinkage (S) estimator optimally combines the UE and RE to outperform the UE, which is given as

$$\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} + \left(1 - \frac{c}{\mathscr{L}_n}\right) (\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}), \tag{5.22}$$

where $c = p_2 - 2$ is the shrinkage constant and $p_2 \geq 3$. The Stein-type shrinkage estimator tends to over-shrink the UE towards the RE when the test statistic $\mathscr{L}_n$ is very small compared to c. Therefore, the truncated version suggested below is applied to avoid this behavior.

5.3.5 Strategy 5: Positive-Part Stein-Type Shrinkage Estimator

The positive-part Stein-type shrinkage or positive-part shrinkage (S^+) estimator is obtained from the shrinkage estimator by changing the factor $1 - c\mathscr{L}_n^{-1}$ to 0 whenever $\mathscr{L}_n \leq c$, which is

$$\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+} = \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} + \left(1 - \frac{c}{\mathscr{L}_n}\right)^+ (\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}), \tag{5.23}$$

where $a^+ = \max(0, a)$ is a positive-part function. Thus, the S^+ is peculiarly essential to manage the over-shrinking inherent in the shrinkage estimator.

5.3.6 Strategy 6: Least Absolute Shrinkage and Selection Operator Estimator

The least absolute shrinkage and selection operator (LASSO) method minimizes the negative partial log-likelihood under the L_1 constraint. It can be assigned as follows:

$$\widehat{\boldsymbol{\beta}}^{\text{LASSO}} = \arg\min_{\boldsymbol{\beta}} \left\{ -\ell(\boldsymbol{\beta}) + \tau \sum_{j=1}^{p} |\beta_j| \right\}, \tag{5.24}$$

where τ is a tuning parameter that controls the amount of shrinkage and is greater than zero. The LASSO shrinks some coefficients to zero, which means that it simultaneously performs variable selection and parameter estimation.

5.3.7 Strategy 7: Adaptive Least Absolute Shrinkage and Selection Operator Estimator

The adaptive least absolute shrinkage and selection operator (aLASSO) method is based on a penalized partial likelihood with adaptively weighted L_1 penalties on regression coefficients. Therefore, the aLASSO objective function becomes

$$\widehat{\boldsymbol{\beta}}^{\text{aLASSO}} = \arg\min_{\boldsymbol{\beta}} \left\{ -\ell(\boldsymbol{\beta}) + \tau \sum_{j=1}^{p} \frac{|\beta_j|}{|\widehat{\beta}_j|^{\gamma}} \right\}, \tag{5.25}$$

where $\widehat{\beta}_j$ is an initial maximum partial likelihood estimator. The tuning parameter τ produces a balance between the goodness of fit and sparsity, and the weights $1/|\widehat{\beta}_j|^{\gamma}$, $\gamma > 0$, allowing for a precise tuning parameter of the penalization.

5.4 Asymptotic Properties and Results

The main asymptotic results of this chapter are derived from thinking along the same lines as described in the previous chapter. The asymptotic distribution of the UE and RE and their joint distribution to facilitate a derivation of the asymptotic properties of the proposed estimators are first presented. We next derive their asymptotic distributional quadratic bias (ADQB) and asymptotic distributional quadratic risk (ADQR) and compare the asymptotic results of the proposed estimators. However, the penalty estimators are not derived since these do not consider the subspace $\boldsymbol{\beta}_2 \neq 0$.

To obtain the expressions for the asymptotic properties of the suggested estimators, we first let $\boldsymbol{\delta} = (\delta_1, \delta_2, ..., \delta_{p_2})^\top \in \mathbb{R}^{p_2}$ and consider the following sequence of local alternatives:

$$\{K_n\} : \boldsymbol{\beta}_2 = \frac{\boldsymbol{\delta}}{\sqrt{n}}. \tag{5.26}$$

Next, we present the following lemmas about the asymptotic distributions under $\{K_n\}$.

Lemma 1. *Under the sequence of local alternatives, the usual regularity conditions of maximum partial likelihood estimation, and as $n \to \infty$, we obtain*

$$\begin{aligned}
\boldsymbol{J}_n &= \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{UE} - \boldsymbol{\beta}_1) \xrightarrow{D} \boldsymbol{J} \sim \mathcal{N}_{p_1}(\mathbf{0}, \boldsymbol{G}_{11.2}^{-1}),\\
\boldsymbol{L}_n &= \sqrt{n}(\widehat{\boldsymbol{\beta}}_2^{UE} - \boldsymbol{\beta}_2) \xrightarrow{D} \boldsymbol{L} \sim \mathcal{N}_{p_2}(\mathbf{0}, \boldsymbol{G}_{22.1}^{-1}),\\
\boldsymbol{M}_n &= \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{RE} - \boldsymbol{\beta}_1) \xrightarrow{D} \boldsymbol{M} \sim \mathcal{N}_{p_1}(-\boldsymbol{\omega\delta}, \boldsymbol{G}_{11}^{-1}),\\
\boldsymbol{O}_n &= \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^{UE} - \widehat{\boldsymbol{\beta}}_1^{RE}) \xrightarrow{D} \boldsymbol{O} \sim \mathcal{N}_{p_1}(\boldsymbol{\omega\delta}, \boldsymbol{\Omega}),\\
\begin{bmatrix} \boldsymbol{J}_n \\ \boldsymbol{O}_n \end{bmatrix} &\xrightarrow{D} \begin{bmatrix} \boldsymbol{J} \\ \boldsymbol{O} \end{bmatrix} \sim \mathcal{N}_{2p_1}\left(\begin{bmatrix} \mathbf{0} \\ \boldsymbol{\omega\delta} \end{bmatrix}, \begin{bmatrix} \boldsymbol{G}_{11.2}^{-1} & \boldsymbol{\Phi} \\ \boldsymbol{\Phi} & \boldsymbol{\Phi} \end{bmatrix}\right),\\
\begin{bmatrix} \boldsymbol{M}_n \\ \boldsymbol{O}_n \end{bmatrix} &\xrightarrow{D} \begin{bmatrix} \boldsymbol{M} \\ \boldsymbol{O} \end{bmatrix} \sim \mathcal{N}_{2p_1}\left(\begin{bmatrix} -\boldsymbol{\omega\delta} \\ \boldsymbol{\omega\delta} \end{bmatrix}, \begin{bmatrix} \boldsymbol{G}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Phi} \end{bmatrix}\right),
\end{aligned}$$

where $\boldsymbol{\omega} = -\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}$, $\boldsymbol{\Phi} = \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}$, and $\xrightarrow{D}$ implies converge in distribution.

Lemma 2. *Under usual regularity conditions and the sequence of local alternatives $\{K_n\}$, as $n \to \infty$,*

$$\boldsymbol{O}_n^* = \sqrt{n}\boldsymbol{\Phi}_n^{-\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{UE} - \widehat{\boldsymbol{\beta}}_1^{RE}) \xrightarrow{D} \boldsymbol{O}^* \sim \mathcal{N}_{p_1}(\boldsymbol{\Phi}^{-\frac{1}{2}}\boldsymbol{\omega\delta}, \boldsymbol{I}_{p_1}),$$

where $\boldsymbol{\Phi}_n = \mathcal{I}_{11}^{-1}\mathcal{I}_{12}\mathcal{I}_{22.1}^{-1}\mathcal{I}_{21}\mathcal{I}_{11}^{-1}$ and $\boldsymbol{\Phi}_n \xrightarrow{P} \boldsymbol{\Phi}$.

The asymptotic distribution of $\boldsymbol{O}_n$ has covariance matrix $\boldsymbol{\Phi}$, while $\boldsymbol{O}_n^*$ also has covariance matrix $\boldsymbol{I}_{p_1}$. The relation between $\boldsymbol{O}$ and $\boldsymbol{O}^*$ is as follows:

$$\boldsymbol{O} = \boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{O}^* = \boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{O}^*. \tag{5.27}$$

The following lemmas facilitate the computation of ADQB and ADQR under local alternatives $\{K_n\}$.

Lemma 3. *Under local alternatives and usual regularity conditions, as $n \to \infty$, the test statistic $\mathscr{L}_n$ converges to a non-central chi-squared distribution with p_2 degrees of freedom and non-centrality parameter $\Delta = \boldsymbol{\delta}^\top \boldsymbol{G}_{22.1}\boldsymbol{\delta}$.*

Lemma 4. *Let* $\boldsymbol{x} = (x_1, x_2, ..., x_k)^\top$ *be a k-dimensional normal vector distributed as* $\mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{x}}, \boldsymbol{\Sigma}_{\boldsymbol{x}})$*. Then, for any measurable function* ϕ*, we have*

$$\mathbb{E}\left[\boldsymbol{x}\phi\left(\boldsymbol{x}^\top\boldsymbol{x}\right)\right] = \boldsymbol{\mu}_{\boldsymbol{x}}\mathbb{E}\left[\phi\left(\chi^2_{k+2}(\Delta)\right)\right], \tag{5.28}$$

$$\mathbb{E}\left[\boldsymbol{x}\boldsymbol{x}^\top\phi\left(\boldsymbol{x}^\top\boldsymbol{x}\right)\right] = \boldsymbol{\Sigma}_{\boldsymbol{x}}\mathbb{E}\left[\phi\left(\chi^2_{k+2}(\Delta)\right)\right] + \boldsymbol{\mu}_{\boldsymbol{x}}\boldsymbol{\mu}_{\boldsymbol{x}}^\top\mathbb{E}\left[\phi\left(\chi^2_{k+4}(\Delta)\right)\right]. \tag{5.29}$$

5.4.1 Asymptotic Distributional Bias

To compare the suggested estimator's performance in terms of estimation bias, we calculated the asymptotic distributional bias (ADB) in the following Equation and used Lemmas 3 and 4.

$$\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^*) = \lim_{n\to\infty}\mathbb{E}\left[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^* - \boldsymbol{\beta}_1)\right], \tag{5.30}$$

where $\widehat{\boldsymbol{\beta}}_1^*$ can be any estimator of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$. Therefore, the ADBs of the estimators are given in the following Theorem.

Theorem 5.4.1. *Under the sequence* $\{K_n\}$ *and usual regularity condition, as* $n \to \infty$*, the ADBs of the suggested estimators are*

$$\begin{aligned}
ADB(\widehat{\boldsymbol{\beta}}_1^{UE}) &= \mathbf{0},\\
ADB(\widehat{\boldsymbol{\beta}}_1^{RE}) &= \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta},\\
ADB(\widehat{\boldsymbol{\beta}}_1^{LS}) &= \pi\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta},\\
ADB(\widehat{\boldsymbol{\beta}}_1^{PT}) &= \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta),\\
ADB(\widehat{\boldsymbol{\beta}}_1^{SP}) &= \pi\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta),\\
ADB(\widehat{\boldsymbol{\beta}}_1^{S}) &= c\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)],\\
ADB(\widehat{\boldsymbol{\beta}}_1^{S+}) &= \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\{H_{p_2+2}(c;\Delta) + c\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)I(\chi^2_{p_2+2}(\Delta) > c)]\},
\end{aligned}$$

where $c = p_2 - 2$, $p_2 > 2$, $H_\nu(\cdot\,;\Delta)$ *is a CDF of non-central chi-square with* ν *degrees of freedom and non-centrality parameter* Δ*, and* $\mathbb{E}\left[\chi_\nu^{-2j}(\Delta)\right] = \int_0^\infty x^{-2j}\,d\phi_\nu(x;\Delta)$.

Proof. Under the usual regularity conditions, a sequence of local alternatives, and using

Lemma 1, 2, and 4, the ADB of the estimators is obtained as follows:

$$\begin{aligned}\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}) &= \lim_{n\to\infty}\mathbb{E}[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \boldsymbol{\beta}_1)]\\ &= \mathbb{E}(\boldsymbol{J}) = \mathbf{0}.\end{aligned}$$

$$\begin{aligned}\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) &= \lim_{n\to\infty}\mathbb{E}[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} - \boldsymbol{\beta}_1)]\\ &= \mathbb{E}(\boldsymbol{M}) = -\boldsymbol{\omega\delta} = \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}.\end{aligned}$$

$$\begin{aligned}\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) &= \lim_{n\to\infty}\mathbb{E}[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}} - \boldsymbol{\beta}_1)]\\ &= \lim_{n\to\infty}\mathbb{E}[n^{\frac{1}{2}}(\pi\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} + (1-\pi)\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \boldsymbol{\beta}_1)]\\ &= \mathbb{E}(\boldsymbol{J}) - \pi\mathbb{E}(\boldsymbol{O}) = -\pi\boldsymbol{\omega\delta} = \pi\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}.\end{aligned}$$

$$\begin{aligned}\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}) &= \lim_{n\to\infty}\mathbb{E}[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}} - \boldsymbol{\beta}_1)]\\ &= \lim_{n\to\infty}\mathbb{E}[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \pi(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})I(\mathscr{L}_n \le \lambda_\alpha) - \boldsymbol{\beta}_1)]\\ &= \mathbb{E}(\boldsymbol{J}) - \pi\mathbb{E}[\boldsymbol{O}I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})]\\ &= -\pi\mathbb{E}[\boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{O}^* I(\chi^2_{p_2}(\Delta) \le \chi^2_{p_2,\alpha})] && \because \text{ by (5.27)}\\ &= -\pi\boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{\Phi}^{-\frac{1}{2}}\boldsymbol{\omega\delta}\mathbb{E}[I(\chi^2_{p_2+2}(\Delta) \le \chi^2_{p_2,\alpha})] && \because \text{ by (5.28)}\\ &= \pi\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta).\end{aligned}$$

For $\pi = 1$, we obtain $\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}) = \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)$. Next, we consider the ADBs of the shrinkage and positive-part shrinkage estimators.

$$\begin{aligned}\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) &= \lim_{n\to\infty}\mathbb{E}[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}} - \boldsymbol{\beta}_1)]\\ &= \lim_{n\to\infty}\mathbb{E}[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} + (1 - c\mathscr{L}_n^{-1})(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \boldsymbol{\beta}_1)]\\ &= \lim_{n\to\infty}\mathbb{E}[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \boldsymbol{\beta}_1) - c\mathscr{L}_n^{-1}n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})]\\ &= \mathbb{E}(\boldsymbol{J}) - c\mathbb{E}(\boldsymbol{O}\chi^{-2}_{p_2}(\Delta))\\ &= -c\mathbb{E}[\boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{O}^*\chi^{-2}_{p_2}(\Delta)] && \because \text{ by (5.27)}\\ &= -c\boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{\Phi}^{-\frac{1}{2}}\boldsymbol{\omega\delta}\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] && \because \text{ by (5.28)}\\ &= c\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)].\end{aligned}$$

$$
\begin{aligned}
\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}) &= \lim_{n\to\infty} \mathrm{E}[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}} - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty} \mathbb{E}[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} + (1 - c\mathscr{L}_n^{-1})^{+}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty} \mathbb{E}\left[n^{\frac{1}{2}}\begin{pmatrix} \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}} + (1 - c\mathscr{L}_n^{-1})(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) \\ -(1 - c\mathscr{L}_n^{-1})(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})I(\mathscr{L}_n \le c) - \boldsymbol{\beta}_1 \end{pmatrix}\right] \\
&= \lim_{n\to\infty} \mathbb{E}[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}} - (1 - c\mathscr{L}_n^{-1})(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})I(\mathscr{L}_n \le c) - \boldsymbol{\beta}_1)] \\
&= \lim_{n\to\infty} \mathbb{E}[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}} - \boldsymbol{\beta}_1) - (1 - c\mathscr{L}_n^{-1})n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}} - \widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}})I(\mathscr{L}_n \le c)] \\
&= \mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - \mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta))\boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{O}^{*}I(\chi_{p_2}^{2}(\Delta) \le c)] \quad \because \text{ by (5.27)} \\
&= \mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - \begin{Bmatrix} \boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{\Phi}^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta} \\ \mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta) \le c)] \end{Bmatrix} \quad \because \text{ by (5.28)} \\
&= \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\left[c\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] + \mathbb{E}[(1 - c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^{2}(\Delta) \le c)]\right] \\
&= \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\{H_{p_2+2}(c;\Delta) + c\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)I(\chi_{p_2+2}^{2}(\Delta) > c)]\}.
\end{aligned}
$$

□

We can see that all the estimators are unbiased in the particular case that $\boldsymbol{\delta} = 0$ or under the null hypothesis. Therefore, the ADBs of all estimators are equivalent. We limit ourselves to the case in which $\boldsymbol{\delta} \neq 0$. In this case, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ is the only unbiased estimator of $\boldsymbol{\beta}_1$ because it does not rely on UPI. The ADB of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$ converges to that of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, and the ADB of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ converges to that of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, as $\alpha \to 0$ and $H_{p_2+2}(\chi_{p_2,\alpha}^{2};\Delta) \to 1$. When $\alpha \to 1$, the ADBs of both $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ converge to that of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$.

We transform the ADB to the scalar form of the asymptotic distributional quadratic bias (ADQB) to make the comparison clearer and more meaningful. Therefore, the ADQB of the estimator $\widehat{\boldsymbol{\beta}}_1^{*}$ is defined as

$$
\mathrm{ADQB}(\widehat{\boldsymbol{\beta}}_1^{*}) = \left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{*})\right]^{\top}\boldsymbol{G}_{11.2}\left[\mathrm{ADB}(\widehat{\boldsymbol{\beta}}_1^{*})\right]. \tag{5.31}
$$

The following theorem expresses the ADQBs of the suggested estimators:

Theorem 5.4.2. *Suppose that the conditions of Theorem 5.4.1 hold. The ADQBs of the suggested estimators are*

$$ADQB(\widehat{\boldsymbol{\beta}}_1^{UE}) = 0,$$
$$ADQB(\widehat{\boldsymbol{\beta}}_1^{RE}) = \Delta^*,$$
$$ADQB(\widehat{\boldsymbol{\beta}}_1^{LS}) = \pi^2\Delta^*,$$
$$ADQB(\widehat{\boldsymbol{\beta}}_1^{PT}) = \Delta^*[H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)]^2,$$
$$ADQB(\widehat{\boldsymbol{\beta}}_1^{SP}) = \pi^2\Delta^*[H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)]^2,$$
$$ADQB(\widehat{\boldsymbol{\beta}}_1^{S}) = c^2\Delta^*\{\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)]\}^2,$$
$$ADQB(\widehat{\boldsymbol{\beta}}_1^{S+}) = \Delta^*\{H_{p_2+2}(c;\Delta) + c\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)I(\chi^2_{p_2+2}(\Delta) > c)]\}^2,$$

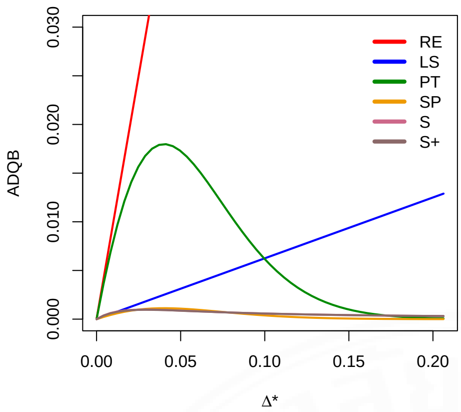
where $c = p_2 - 2$, $p_2 > 2$, $\Delta^* = \boldsymbol{\delta}^\top \boldsymbol{G}^*\boldsymbol{\delta}$ *and* $\boldsymbol{G}^* = \boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{11.2}\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}$.

Proof. The proof of this theorem is straightforward and can be easily derived by using the above ADB and Equation (5.31). □

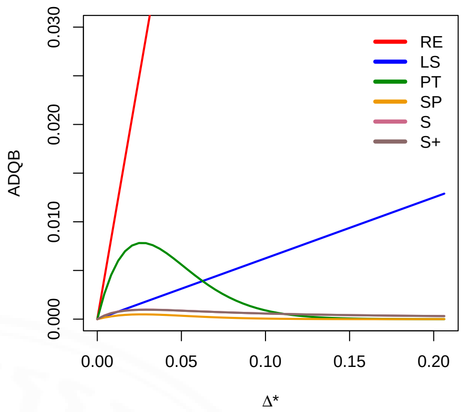
For the choice of $\alpha = 0.01$ and 0.05, and $\pi = 0.25$, 0.50, and 0.75, Figures 5.1 and 5.2 shows the ADQBs behaviour of the suggested estimators for $p_1 = 3$, and $p_2 = 3$ and 7. We can see that the ADQBs of all estimators except the UE are functions of Δ^*. Therefore, we investigate the behavior of the ADQB of the suggested estimators in terms of Δ^*. The ADQB($\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$) and ADQB($\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$) are unbounded functions of Δ^* and tend to ∞ when $\Delta^* \to \infty$. The ADQBs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ are functions of Δ^* and α. The ADQBs of both estimators start from the initial value 0, increase to a certain point, then gradually decrease to zero. Moreover, ADQB($\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$) $= \pi$ADQB($\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$) $<$ ADQB($\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$) for $\pi \in [0, 1)$, so $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ has asymptotically less bias than $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$ depending upon the value of π. Thus, one can think of π as a bias reduction factor in the pretest estimation. The ADQBs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$ start from 0 at $\Delta^* = 0$ and elevate to a point, and after that go to zero, since $\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)]$ is a decreasing log-convex function of Δ. However, the curve of ADQB($\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$) stays below the curve of ADQB($\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$) for all values of Δ^*.

5.4.2 Asymptotic Distributional Risk

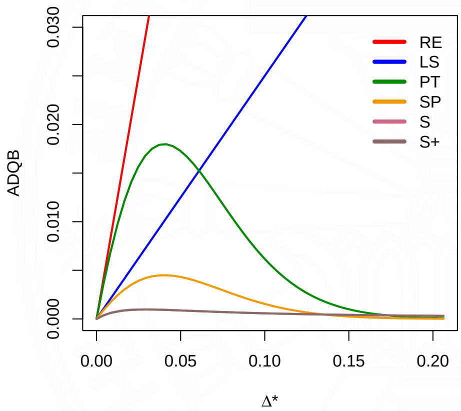
For the local alternatives $\{K_n\}$, the suggested estimators may not be asymptotically unbiased estimators of $\boldsymbol{\beta}_1$. With that in mind, we introduce the following



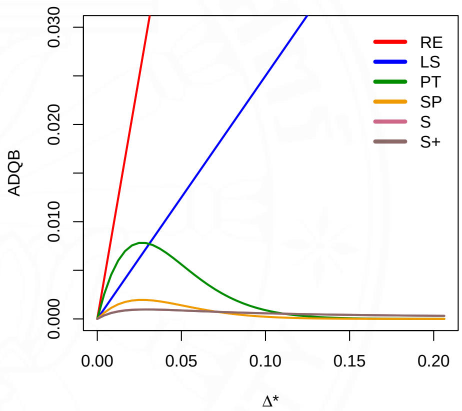
(a) $\pi = 0.25$, $\alpha = 0.01$



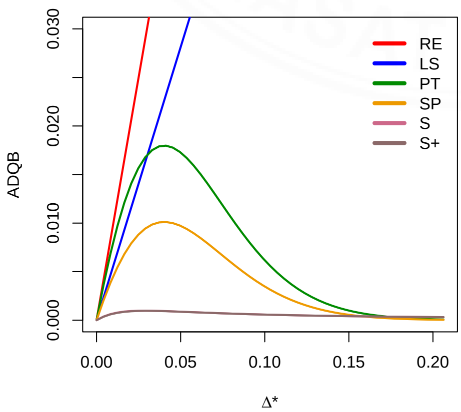
(b) $\pi = 0.25$, $\alpha = 0.05$



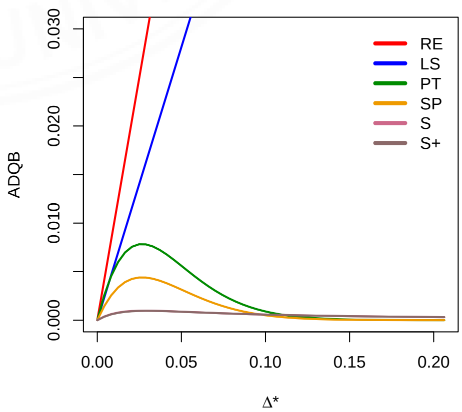
(c) $\pi = 0.50$, $\alpha = 0.01$



(d) $\pi = 0.50$, $\alpha = 0.01$

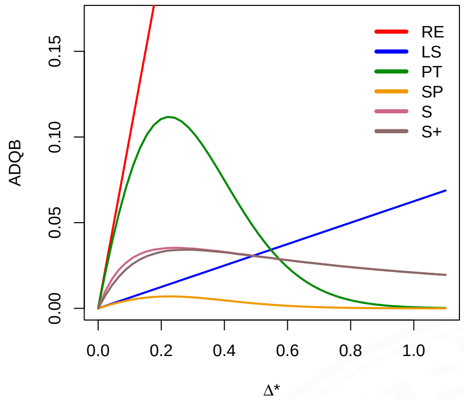


(e) $\pi = 0.75$, $\alpha = 0.01$

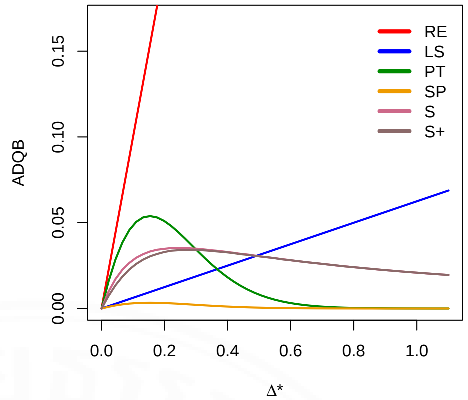


(f) $\pi = 0.75$, $\alpha = 0.05$

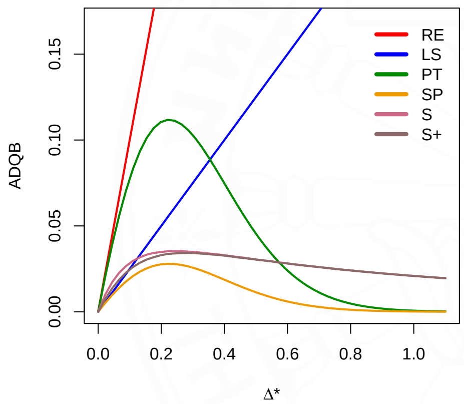
Figure 5.1 ADQB curves of the suggested estimators for Cox PH regression model with $p_1 = 3$ and $p_2 = 3$



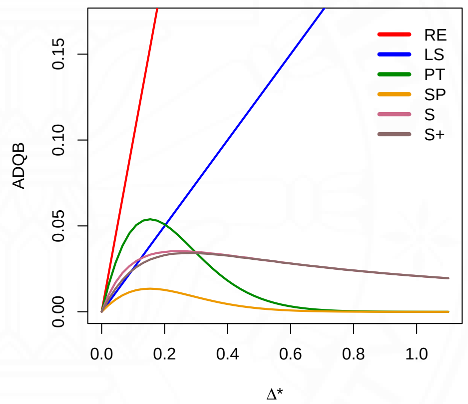
(a) $\pi = 0.25$, $\alpha = 0.01$



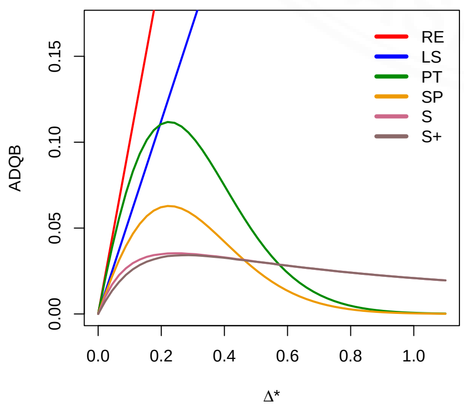
(b) $\pi = 0.25$, $\alpha = 0.05$



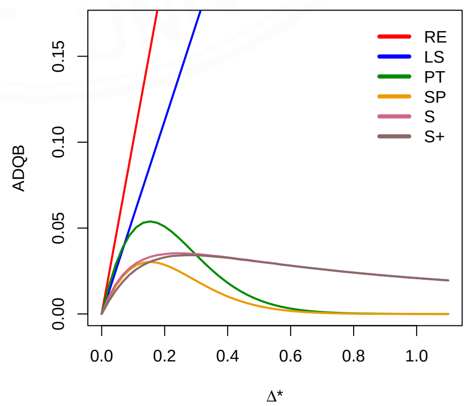
(c) $\pi = 0.50$, $\alpha = 0.01$



(d) $\pi = 0.50$, $\alpha = 0.01$



(e) $\pi = 0.75$, $\alpha = 0.01$



(f) $\pi = 0.75$, $\alpha = 0.05$

Figure 5.2 ADQB curves of the suggested estimators for Cox PH regression model with $p_1 = 3$ and $p_2 = 7$

loss function:

$$\mathcal{L}(\widehat{\boldsymbol{\beta}}_1^*, \boldsymbol{\beta}_1; \boldsymbol{W}) = \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^* - \boldsymbol{\beta}_1)^\top \boldsymbol{W} \sqrt{n}(\widehat{\boldsymbol{\beta}}_1^* - \boldsymbol{\beta}_1), \tag{5.32}$$

where $\boldsymbol{W}$ is a positive semi-definite weight matrix. A generic choice for $\boldsymbol{W}$ is the identity matrix so that $\mathcal{L}(\widehat{\boldsymbol{\beta}}_1^*, \boldsymbol{\beta}_1; \boldsymbol{W})$ is the un-weighted quadratic loss. The asymptotic mean squared error matrix (AMSEM) formula of an estimator under the quadratic loss function would be as follows:

$$\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^*) = \lim_{n\to\infty} \mathbb{E}\left[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^* - \boldsymbol{\beta}_1) n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^* - \boldsymbol{\beta}_1)^\top \right]. \tag{5.33}$$

We first derive the AMSEM of the suggested estimators and then use them to compute the asymptotic distributional quadratic risk (ADQR). The results of $\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^*)$ are given in the following Theorem.

Theorem 5.4.3. *Under the sequence of local alternative $\{K_n\}$ and usual regularity conditions, as $n \to \infty$ the AMSEMs of the suggested estimators are given as follows:*

$$\begin{aligned}
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{UE}) &= \boldsymbol{G}_{11.2}^{-1}, \\
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{RE}) &= \boldsymbol{G}_{11}^{-1} + \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}, \\
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{LS}) &= \boldsymbol{G}_{11.2}^{-1} - \pi(2-\pi)\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1} + \pi^2\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}, \\
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{PT}) &= \boldsymbol{G}_{11.2}^{-1} - \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{Q}_{21}\boldsymbol{G}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}[2H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)], \\
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{SP}) &= \boldsymbol{G}_{11.2}^{-1} - \pi(2-\pi)\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad + \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}[2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - \pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)], \\
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{S}) &= \boldsymbol{G}_{11.2}^{-1} - c\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}(2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] - c\mathbb{E}[\chi^{-4}_{p_2+2}(\Delta)]) \\
&\quad + c\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}(2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] - 2\mathbb{E}[\chi^{-2}_{p_2+4}(\Delta)] + c\mathbb{E}[\chi^{-4}_{p_2+4}(\Delta)]),
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{S+}) &= \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}_1^{S}) - \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[(1 - c\chi^{-2}_{p_2+2}(\Delta))^2 I(\chi^2_{p_2+2}(\Delta) \le c)] \\
&\quad - \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[(1 - c\chi^{-2}_{p_2+4}(\Delta))^2 I(\chi^2_{p_2+4}(\Delta) \le c)] \\
&\quad + 2\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\mathbb{E}[(1 - c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta) \le c)].
\end{aligned}$$

where $c = p_2 - 2$ and $p_2 > 2$.

Proof. By the above Lemmas, the asymptotic mean squared error matrix of an estimator $\boldsymbol{\beta}_1^*$ is derived as follows:

$$\begin{aligned}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) &= \lim_{n\to\infty}\mathbb{E}\left[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}-\boldsymbol{\beta}_1)n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}-\boldsymbol{\beta}_1)^\top\right]\\ &= \mathbb{E}[(\boldsymbol{J}-\pi\boldsymbol{O})(\boldsymbol{J}-\pi\boldsymbol{O})^\top]\\ &= \underbrace{\mathbb{E}[\boldsymbol{J}\boldsymbol{J}^\top]}_{\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}})}-2\pi\underbrace{\mathbb{E}[\boldsymbol{J}\boldsymbol{O}^\top]}_{\mathrm{A}_1}+\pi^2\underbrace{\mathbb{E}[\boldsymbol{O}\boldsymbol{O}^\top]}_{\mathrm{A}_2},\end{aligned} \tag{5.34}$$

where

$$\mathrm{A}_2=\mathbb{V}(\boldsymbol{O})+\mathbb{E}(\boldsymbol{O})\mathbb{E}(\boldsymbol{Z}^\top)=\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}+\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}.$$

By the law of conditional expectation of a multivariate normal distribution, we may write A_1 as

$$\begin{aligned}\mathrm{A}_1 &= \mathbb{E}[\mathbb{E}(\boldsymbol{J}\boldsymbol{O}^\top|\boldsymbol{O})]=\mathbb{E}[\mathbb{E}(\boldsymbol{J}|\boldsymbol{O})\boldsymbol{O}^\top]\\ &= \mathbb{E}[\{\mathbb{E}(\boldsymbol{J})+Cov(\boldsymbol{J},\boldsymbol{O})[\mathbb{V}(\boldsymbol{O})]^{-1}(\boldsymbol{O}-\mathbb{E}(\boldsymbol{O})\}\boldsymbol{O}^\top]\\ &= \underbrace{\mathbb{E}(\boldsymbol{O}\boldsymbol{O}^\top)}_{\mathrm{A}_2}+\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\mathbb{E}(\boldsymbol{O}^\top)=\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}.\end{aligned}$$

Therefore, Equation (5.34) becomes

$$\begin{aligned}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}) &= \boldsymbol{G}_{11.2}^{-1}-\pi(2-\pi)\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\\ &\quad+\pi^2\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}.\end{aligned} \tag{5.35}$$

For $\pi=0$, Equation (5.35) reduces to

$$\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}})=\boldsymbol{G}_{11.2}^{-1}.$$

When $\pi=1$, Equation (5.35) becomes

$$\begin{aligned}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}) &= \boldsymbol{G}_{11.2}^{-1}-\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}+\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\\ &= \boldsymbol{G}_{11}^{-1}+\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}.\end{aligned}$$

Next, the AMSEM of $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$ can be written as

$$\begin{aligned}
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\text{SP}}) &= \lim_{n\to\infty}\mathbb{E}\left[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\text{SP}}-\boldsymbol{\beta}_1)n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\text{SP}}-\boldsymbol{\beta}_1)^\top\right] \\
&= \mathbb{E}\lim_{n\to\infty}[(\boldsymbol{J}_n-\pi\boldsymbol{O}_n I(\mathscr{L}_n\le\lambda_\alpha))(\text{J}_n-\pi\boldsymbol{O}_n I(\mathscr{L}_n\le\lambda_\alpha))^\top] \\
&= \mathbb{E}[(\boldsymbol{J}-\pi\boldsymbol{O}I(\chi^2_{p_2}(\Delta)\le\chi^2_{p_2,\alpha}))(\boldsymbol{J}-\pi\boldsymbol{O}I(\chi^2_{p_2}(\Delta)\le\chi^2_{p_2,\alpha}))^\top] \\
&= \underbrace{\mathbb{E}[\boldsymbol{J}\boldsymbol{J}^\top]}_{\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\text{UE}})} -2\pi\underbrace{\mathbb{E}[\boldsymbol{J}\boldsymbol{O}^\top I(\chi^2_{p_2}(\Delta)\le\chi^2_{p_2,\alpha})]}_{\text{A}_3} +\pi^2\underbrace{\mathbb{E}[\boldsymbol{O}\boldsymbol{O}^\top I(\chi^2_{p_2}(\Delta)\le\chi^2_{p_2,\alpha})]}_{\text{A}_4}.
\end{aligned}$$

Using Equations (5.27) and (5.29), we have

$$\begin{aligned}
\text{A}_4 &= \mathbb{E}[\boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{O}^*(\boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{O}^*)^\top I(\chi^2_{p_2}(\Delta)\le\chi^2_{p_2,\alpha})] \\
&= \boldsymbol{\Phi}^{\frac{1}{2}}\begin{bmatrix}\boldsymbol{I}_{p_2}\mathbb{E}[I(\chi^2_{p_2+2}(\Delta)\le\chi^2_{p_2,\alpha})] \\ +\boldsymbol{\Phi}^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta}(\boldsymbol{\Phi}^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta})^\top\mathbb{E}[I(\chi^2_{p_2+4}(\Delta)\le\chi^2_{p_2,\alpha})]\end{bmatrix}(\boldsymbol{\Phi}^{\frac{1}{2}})^\top \\
&= \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)+\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta).
\end{aligned}$$

Using the rule of conditional expectation along with Equation (5.28), A_3 becomes

$$\begin{aligned}
\text{A}_3 &= \mathbb{E}[\mathbb{E}(\boldsymbol{J}\boldsymbol{O}^\top I(\chi^2_{p_2}(\Delta)\le\chi^2_{p_2,\alpha})|\boldsymbol{O})] = \mathbb{E}[\mathbb{E}(\boldsymbol{J}|\boldsymbol{O})\boldsymbol{O}^\top I(\chi^2_{p_2}(\Delta)\le\chi^2_{p_2,\alpha})] \\
&= \mathbb{E}[\{\mathbb{E}(\boldsymbol{J})+Cov(\boldsymbol{J},\boldsymbol{O})[\mathbb{V}(\boldsymbol{O})]^{-1}(\boldsymbol{O}-\mathbb{E}(\boldsymbol{O})\}\boldsymbol{O}^\top I(\chi^2_{p_2}(\Delta)\le\chi^2_{p_2,\alpha})] \\
&= \underbrace{\mathbb{E}[\boldsymbol{O}\boldsymbol{O}^\top I(\chi^2_{p_2}(\Delta)\le\chi^2_{p_2,\alpha})]}_{\text{A}_4} -(-\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta})\mathbb{E}[\boldsymbol{O}^\top I(\chi^2_{p_2}(\Delta)\le\chi^2_{p_2,\alpha})] \\
&= \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)+\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad -(-\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta})(-\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta})^\top H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&= \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad -\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}[H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)-H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)].
\end{aligned}$$

Substitute A_3 and A_4 into $\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\text{SP}})$, then we obtain

$$\begin{aligned}
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\text{SP}}) &= \boldsymbol{G}_{11.2}^{-1}-2\pi\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad +2\pi\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}[H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)-H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)] \\
&\quad +\pi^2\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad +\pi^2\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta) \\
&= \boldsymbol{G}_{11.2}^{-1}-\pi(2-\pi)\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) \\
&\quad +\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}[2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)-\pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)].
\end{aligned}$$

For $\pi = 1$, $\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\text{SP}})$ reduces to

$$\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\text{PT}}) = \boldsymbol{G}_{11.2}^{-1} - \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) + \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}[2H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)].$$

Let us consider $\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\text{S}})$. This yields

$$\begin{aligned}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\text{S}}) &= \lim_{n\to\infty}\mathbb{E}\left[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\text{S}} - \boldsymbol{\beta}_1)n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\text{S}} - \boldsymbol{\beta}_1)^\top\right] \\ &= \mathbb{E}\lim_{n\to\infty}[(\boldsymbol{J}_n - c\boldsymbol{O}_n\mathscr{L}_n^{-1})(\boldsymbol{J}_n - c\boldsymbol{O}_n\mathscr{L}_n^{-1})^\top] \\ &= \underbrace{\mathbb{E}[\boldsymbol{J}\boldsymbol{J}^\top]}_{\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\text{UE}})} - 2c\underbrace{\mathbb{E}[\boldsymbol{J}\boldsymbol{O}^\top\chi_{p_2}^{-2}(\Delta)]}_{\text{A}_5} + c^2\underbrace{\mathbb{E}[\boldsymbol{O}\boldsymbol{O}^\top\chi_{p_2}^{-4}(\Delta)]}_{\text{A}_6}.\end{aligned}$$

Applying Equations (5.27) and (5.29) to A_6, we get

$$\begin{aligned}\text{A}_6 &= \mathbb{E}[\boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{O}^*(\boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{O}^*)^\top\chi_{p_2}^{-4}(\Delta)] \\ &= \boldsymbol{\Phi}^{\frac{1}{2}}\left[\boldsymbol{I}_{p_2}\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] + \boldsymbol{\Phi}^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta}(\boldsymbol{\Phi}^{-\frac{1}{2}}\boldsymbol{\omega}\boldsymbol{\delta})^\top\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]\right](\boldsymbol{\Phi}^{\frac{1}{2}})^\top \\ &= \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] + \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)].\end{aligned}$$

Using conditional expectation and Equations (5.28) and (5.27), therefore, A_5 becomes,

$$\begin{aligned}\text{A}_5 &= \mathbb{E}[\mathbb{E}(\boldsymbol{J}\boldsymbol{O}^\top\chi_{p_2}^{-2}(\Delta)|\boldsymbol{O})] = \mathbb{E}[\mathbb{E}(\boldsymbol{J}|\boldsymbol{O})\boldsymbol{O}^\top\chi_{p_2}^{-2}(\Delta)] \\ &= \mathbb{E}[(\boldsymbol{O} - (-\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}))\boldsymbol{O}^\top\chi_{p_2}^{-2}(\Delta)] \\ &= \mathbb{E}[\boldsymbol{O}\boldsymbol{O}^\top\chi_{p_2}^{-2}(\Delta)] - (-\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta})\mathbb{E}[\boldsymbol{O}^\top\chi_{p_2}^{-2}(\Delta)] \\ &= \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] + \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)] \\ &\quad - \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)].\end{aligned}$$

Then, the AMSEM of $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$ is given by

$$\begin{aligned}\boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}_1^{\text{S}}) &= \boldsymbol{G}_{11.2}^{-1} - 2c\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] \\ &\quad + 2c\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}(\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - \mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)]) \\ &\quad + c^2\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)] + c^2\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)] \\ &= \boldsymbol{G}_{11.2}^{-1} - c\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]) \\ &\quad + c\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}(2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - 2\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)] + c\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]).\end{aligned}$$

Finally, we derive the AMSEM of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$, which is

$$\begin{aligned}
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}) &= \lim_{n\to\infty}\mathbb{E}\left[n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}+}-\boldsymbol{\beta}_1)n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}+}-\boldsymbol{\beta}_1)^\top\right] \\
&= \mathbb{E}\lim_{n\to\infty}\begin{bmatrix}\{\boldsymbol{J}_n - c\boldsymbol{O}_n\Lambda_n^{-1} - \boldsymbol{O}_n I(\Lambda_n \le c) + c\boldsymbol{O}_n\Lambda_n^{-1}I(\Lambda_n\le c)\} \\ \{\boldsymbol{J}_n - c\boldsymbol{O}_n\Lambda_n^{-1} - \boldsymbol{O}_n I(\Lambda_n \le c) + c\boldsymbol{O}_n\Lambda_n^{-1}I(\Lambda_n\le c)\}^\top\end{bmatrix} \\
&= \underbrace{\mathbb{E}[(\boldsymbol{J}-c\boldsymbol{O}\chi_{p_2}^{-2}(\Delta))(\boldsymbol{J}-c\boldsymbol{O})^\top\chi_{p_2}^{-2}(\Delta)]}_{\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}})} \\
&\quad -2\mathbb{E}[(\boldsymbol{J}-c\boldsymbol{O}\chi_{p_2}^{-2}(\Delta))(1-c\chi_{p_2}^{-2}(\Delta))\boldsymbol{O}^\top I(\chi_{p_2}^2(\Delta)\le c)] \\
&\quad +\mathbb{E}[(1-c\chi_{p_2}^{-2}(\Delta))^2\boldsymbol{O}\boldsymbol{O}^\top I(\chi_{p_2}^2(\Delta)\le c)] \\
&= \boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - 2\underbrace{\mathbb{E}[(1-c\chi_{p_2}^{-2}(\Delta))\boldsymbol{J}\boldsymbol{O}^\top I(\chi_{p_2}^2(\Delta)\le c)]}_{\mathrm{A}_7} \\
&\quad + 2\underbrace{\mathbb{E}[c\chi_{p_2}^{-2}(\Delta)(1-c\chi_{p_2}^{-2}(\Delta))\boldsymbol{O}\boldsymbol{O}^\top I(\chi_{p_2}^2(\Delta)\le c)]}_{\mathrm{A}_8} \\
&\quad + \underbrace{\mathbb{E}[(1-c\chi_{p_2}^{-2}(\Delta))^2\boldsymbol{O}\boldsymbol{O}^\top I(\chi_{p_2}^2(\Delta)\le c)]}_{\mathrm{A}_9}.
\end{aligned}$$

Using conditional expectation and Lemma 4, A_7 becomes

$$\begin{aligned}
\mathrm{A}_7 &= \mathbb{E}[(1-c\chi_{p_2}^{-2}(\Delta))\mathbb{E}(\boldsymbol{J}\boldsymbol{O}^\top I(\chi_{p_2}^2(\Delta)\le c)|\boldsymbol{O})] \\
&= \mathbb{E}[(1-c\chi_{p_2}^{-2}(\Delta))\boldsymbol{O}\boldsymbol{O}^\top I(\chi_{p_2}^2(\Delta)\le c)] \\
&\quad -(-\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta})\mathbb{E}[(1-c\chi_{p_2}^{-2}(\Delta))\boldsymbol{O}^\top I(\chi_{p_2}^2(\Delta)\le c)] \\
&= \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^2(\Delta)\le c)] \\
&\quad +\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))I(\chi_{p_2+4}^2(\Delta)\le c)] \\
&\quad -\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^2(\Delta)\le c)].
\end{aligned}$$

Using Equation (5.29), we can write A_8 and A_9 as

$$\begin{aligned}
\mathrm{A}_8 &= \mathbb{E}[c\chi_{p_2}^{-2}(\Delta)(1-c\chi_{p_2}^{-2}(\Delta))\boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{O}^*(\boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{O}^*)^\top I(\chi_{p_2}^2(\Delta)\le c)] \\
&= \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[c\chi_{p_2+2}^{-2}(\Delta)(1-c\chi_{p_2+2}^{-2}(\Delta))I(\chi_{p_2+2}^2(\Delta)\le c)] \\
&\quad +\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[c\chi_{p_2+4}^{-2}(\Delta)(1-c\chi_{p_2+4}^{-2}(\Delta))I(\chi_{p_2+4}^2(\Delta)\le c)],
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{A}_9 &= \mathbb{E}[(1-c\chi_{p_2}^{-2}(\Delta))^2\boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{O}^*(\boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{O}^*)^\top I(\chi_{p_2}^2 \le c)] \\
&= \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^2(\Delta) \le c)] \\
&\quad + \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^2(\Delta) \le c)].
\end{aligned}$$

Substituting A_7, A_8, and A_9 into $\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+})$ and rearranging the terms, so we have

$$\begin{aligned}
\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}) &= \boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}) - \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{G}_{22.1}^{-1}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^2(\Delta) \le c)] \\
&\quad - \boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^2(\Delta) \le c)] \\
&\quad + 2\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta)) I(\chi_{p_2+2}^2(\Delta) \le c)].
\end{aligned}$$

□

The results for ADQR of the suggested estimators are derived by using the AMSEM results in Theorem 5.4.3 and the following Equation.

$$\mathrm{ADQR}(\widehat{\boldsymbol{\beta}}_1^*) = \mathrm{tr}[\boldsymbol{W}\boldsymbol{\Gamma}^*(\widehat{\boldsymbol{\beta}}_1^*)]. \tag{5.36}$$

The ADQR expressions are given as follows:

Theorem 5.4.4. *Under the assumed regularity condition and local alternative $\{K_n\}$, as $n \to \infty$, the ADQRs of the estimators are given by*

$$\begin{aligned}
ADQR(\widehat{\boldsymbol{\beta}}_1^{UE}) &= tr[\boldsymbol{W}\boldsymbol{G}_{11.2}^{-1}], \\
ADQR(\widehat{\boldsymbol{\beta}}_1^{RE}) &= ADQR(\widehat{\boldsymbol{\beta}}_1^{UE}) - tr[\boldsymbol{G}^\circ\boldsymbol{G}_{22.1}^{-1}] + \boldsymbol{\delta}^\top\boldsymbol{G}^\circ\boldsymbol{\delta}, \\
ADQR(\widehat{\boldsymbol{\beta}}_1^{LS}) &= ADQR(\widehat{\boldsymbol{\beta}}_1^{UE}) - \pi(2-\pi)tr[\boldsymbol{G}^\circ\boldsymbol{G}_{22.1}^{-1}] + \pi^2\boldsymbol{\delta}^\top\boldsymbol{G}^\circ\boldsymbol{\delta}, \\
ADQR(\widehat{\boldsymbol{\beta}}_1^{PT}) &= ADQR(\widehat{\boldsymbol{\beta}}_1^{UE}) - tr[\boldsymbol{G}^\circ\boldsymbol{G}_{22.1}^{-1}]H_{p_2+2}(\chi_{p_2,\alpha}^2;\Delta) \\
&\quad + \boldsymbol{\delta}^\top\boldsymbol{G}^\circ\boldsymbol{\delta}[2H_{p_2+2}(\chi_{p_2,\alpha}^2;\Delta) - H_{p_2+4}(\chi_{p_2,\alpha}^2;\Delta)], \\
ADQR(\widehat{\boldsymbol{\beta}}_1^{SP}) &= ADQR(\widehat{\boldsymbol{\beta}}_1^{UE}) - \pi(2-\pi)tr[\boldsymbol{G}^\circ\boldsymbol{G}_{22.1}^{-1}]H_{p_2+2}(\chi_{p_2,\alpha}^2;\Delta) \\
&\quad + \boldsymbol{\delta}^\top\boldsymbol{G}^\circ\boldsymbol{\delta}[2\pi H_{p_2+2}(\chi_{p_2,\alpha}^2;\Delta) - \pi(2-\pi)H_{p_2+4}(\chi_{p_2,\alpha}^2;\Delta)], \\
ADQR(\widehat{\boldsymbol{\beta}}_1^{S}) &= ADQR(\widehat{\boldsymbol{\beta}}_1^{UE}) - ctr[\boldsymbol{G}^\circ\boldsymbol{G}_{22.1}^{-1}](2\mathbb{E}[\chi_{p_2+2}^{-2}(\Delta)] - c\mathbb{E}[\chi_{p_2+2}^{-4}(\Delta)]) \\
&\quad + c\boldsymbol{\delta}^\top\boldsymbol{G}^\circ\boldsymbol{\delta}(2E[\chi_{p_2+2}^{-2}(\Delta)] - 2\mathbb{E}[\chi_{p_2+4}^{-2}(\Delta)] + c\mathbb{E}[\chi_{p_2+4}^{-4}(\Delta)]), \\
ADQR(\widehat{\boldsymbol{\beta}}_1^{S^+}) &= ADQR(\widehat{\boldsymbol{\beta}}_1^{S}) - tr[\boldsymbol{G}^\circ\boldsymbol{G}_{22.1}^{-1}]\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta))^2 I(\chi_{p_2+2}^2(\Delta) \le c)] \\
&\quad - \boldsymbol{\delta}^\top\boldsymbol{G}^\circ\boldsymbol{\delta}\mathbb{E}[(1-c\chi_{p_2+4}^{-2}(\Delta))^2 I(\chi_{p_2+4}^2(\Delta) \le c)] \\
&\quad + 2\boldsymbol{\delta}^\top\boldsymbol{G}^\circ\boldsymbol{\delta}\mathbb{E}[(1-c\chi_{p_2+2}^{-2}(\Delta)) I(\chi_{p_2+2}^2(\Delta) \le c)],
\end{aligned}$$

where $c = p_2 - 2$, $p_2 > 2$, *and* $\boldsymbol{G}^{\circ} = \boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\boldsymbol{W}\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}$.

Proof. The proof of ADQR of the estimators can be derived from Equation (5.36) using the above asymptotic covariance matrix. □

The ADQR expressions given above depend on the weight matrix $\boldsymbol{W}$ and if we select a suitable choice of $\boldsymbol{W}$, then these ADQR simplify, and the results are presented in the following corollary.

Corollary 3. *When* $\boldsymbol{W} = \boldsymbol{G}_{11.2}$, *the ADQRs of the estimators simplify to*

$$\begin{aligned}
ADQR(\widehat{\boldsymbol{\beta}}_1^{UE}) &= p_1,\\
ADQR(\widehat{\boldsymbol{\beta}}_1^{RE}) &= p_1 - tr[\boldsymbol{G}^*\boldsymbol{G}_{22.1}^{-1}] + \Delta^*,\\
ADQR(\widehat{\boldsymbol{\beta}}_1^{LS}) &= p_1 - \pi(2-\pi)tr[\boldsymbol{G}^*\boldsymbol{G}_{22.1}^{-1}] + \pi^2\Delta^*,\\
ADQR(\widehat{\boldsymbol{\beta}}_1^{PT}) &= p_1 - tr[\boldsymbol{G}^*\boldsymbol{G}_{22.1}^{-1}]H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)\\
&\quad + \Delta^*[2H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)],\\
ADQR(\widehat{\boldsymbol{\beta}}_1^{SP}) &= p_1 - \pi(2-\pi)tr[\boldsymbol{G}^*\boldsymbol{G}_{22.1}^{-1}]H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta)\\
&\quad + \Delta^*[2\pi H_{p_2+2}(\chi^2_{p_2,\alpha};\Delta) - \pi(2-\pi)H_{p_2+4}(\chi^2_{p_2,\alpha};\Delta)],\\
ADQR(\widehat{\boldsymbol{\beta}}_1^{S}) &= p_1 - c\,tr[\boldsymbol{G}^*\boldsymbol{G}_{22.1}^{-1}](2\mathbb{E}[\chi^{-2}_{p_2+2}(\Delta)] - c\mathbb{E}[\chi^{-4}_{p_2+2}(\Delta)])\\
&\quad + c\Delta^*(2E[\chi^{-2}_{p_2+2}(\Delta)] - 2\mathbb{E}[\chi^{-2}_{p_2+4}(\Delta)] + c\mathbb{E}[\chi^{-4}_{p_2+4}(\Delta)]),\\
ADQR(\widehat{\boldsymbol{\beta}}_1^{S+}) &= ADQR(\widehat{\boldsymbol{\beta}}_1^{S}) - tr[\boldsymbol{G}^*\boldsymbol{G}_{22.1}^{-1}]\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))^2 I(\chi^2_{p_2+2}(\Delta) \le c)]\\
&\quad - \Delta^*\mathbb{E}[(1-c\chi^{-2}_{p_2+4}(\Delta))^2 I(\chi^2_{p_2+4}(\Delta) \le c)]\\
&\quad + 2\Delta^*\mathbb{E}[(1-c\chi^{-2}_{p_2+2}(\Delta))I(\chi^2_{p_2+2}(\Delta) \le c)].
\end{aligned}$$

where $c = p_2 - 2$, $p_2 > 2$, $\Delta^* = \boldsymbol{\delta}^\top\boldsymbol{G}^*\boldsymbol{\delta}$ *and* $\boldsymbol{G}^* = \boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{11.2}\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{12}$.

There were similar ADQR analyses for comparison of the suggested estimators in the previous chapter, and as such they are not reported here. However, we have plotted the ADQRs of the suggested estimators in Corollary 3 for $p_1 = 3$ and $p_2 = 5$ and 11 here. Figures 5.3 and 5.4 represent the ADQR curves with $\alpha = 0.01$ and 0.05, and $\pi = 0.25$, 0.50, and 0.75.

The results show that the ADQRs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$ are an unbounded function of Δ^*. When Δ^* was equal or close to zero, the ADQRs of all suggested estimators were smaller than ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$), which means they were superior in performance to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$. For all α, the ADQRs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$ and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ approached the ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ when $\Delta^* \to \infty$. The ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$ depends on α and decreased as α increased, while the ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$ depended on α and π. For fixed α, there was an increase in π associated with larger variation in ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$). In contrast, for fixed π, the small α tended to have larger variation in ADQR($\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$). Moreover, the ADQR of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ was always greater or equivalent to that of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$.

5.5 Simulation Results

We assumed that the Cox PH model has p available predictors with a sample size of n. The survival times were generated from the following:

$$h(t) = h_0(t)\exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}). \tag{5.37}$$

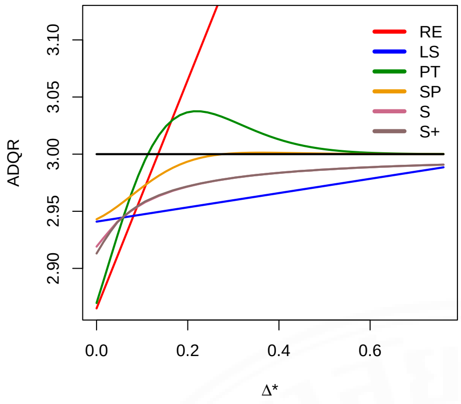
The values of independent variable x_{ij} were generated from a standard normal distribution for $i = 1, 2, ...n$. Here $x_{i1}, x_{i2}, ..., x_{ip}$ were independent and identically distributed. The proportion of observations specified by the censor was randomly and uniformly selected to be right-censored. The censoring times were generated from a uniform distribution in the interval $(0, c)$, where c was chosen to obtain the desired censoring rate. We used two different types of censoring percentages or proportions of censoring (pc), i.e., 20% and 30%. The baseline hazard function was generated using the flexible-hazard method described in Harden and Kropko (2019). We used the `sim.survdata` function in the `coxed` package for generating data in the Cox PH model in Equation (5.37) and ran all the simulated data and calculations on the statistical software `R`.

In this section, the Monte Carlo simulation was used to examine the performance of the suggested estimators by comparing them with the UE via the simulated relative mean square error (RMSE). The RMSE is a ratio of the simulated mean square error (MSE) of the suggested estimators and the simulated MSE of the UE. Therefore, the RMSE of any estimator $\widehat{\boldsymbol{\beta}}_1^{*}$ is

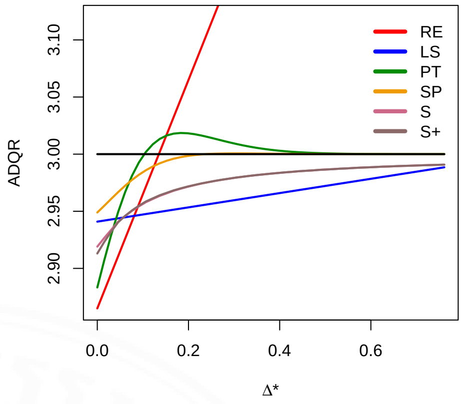
$$\mathrm{RMSE}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}, \widehat{\boldsymbol{\beta}}_1^{*}) = \frac{\mathrm{MSE}(\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}})}{\mathrm{MSE}(\widehat{\boldsymbol{\beta}}_1^{*})},$$

where

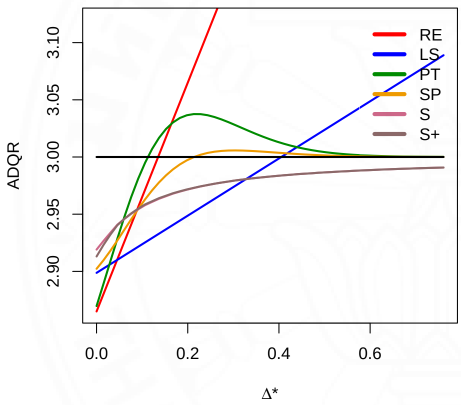
$$\mathrm{MSE}(\widehat{\boldsymbol{\beta}}_1^{*}) = \frac{1}{n}(\boldsymbol{\beta}_1 - \widehat{\boldsymbol{\beta}}_1^{*})^\top(\boldsymbol{\beta}_1 - \widehat{\boldsymbol{\beta}}_1^{*}).$$



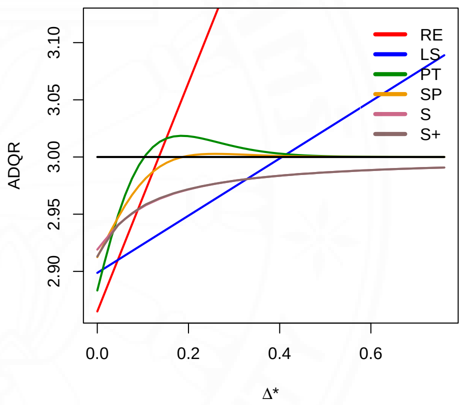
(a) $\pi = 0.25$, $\alpha = 0.01$



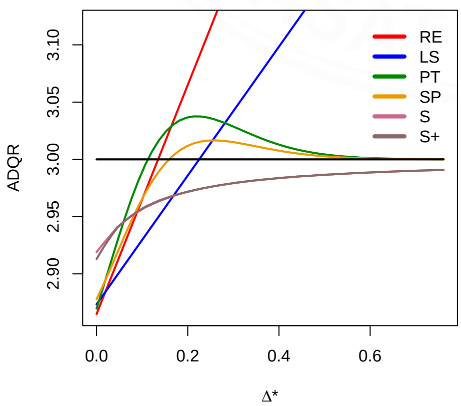
(b) $\pi = 0.25$, $\alpha = 0.05$



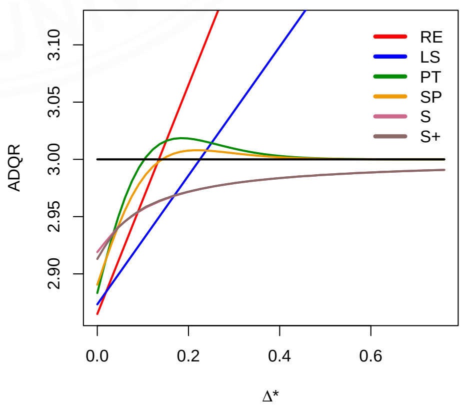
(c) $\pi = 0.50$, $\alpha = 0.01$



(d) $\pi = 0.50$, $\alpha = 0.05$

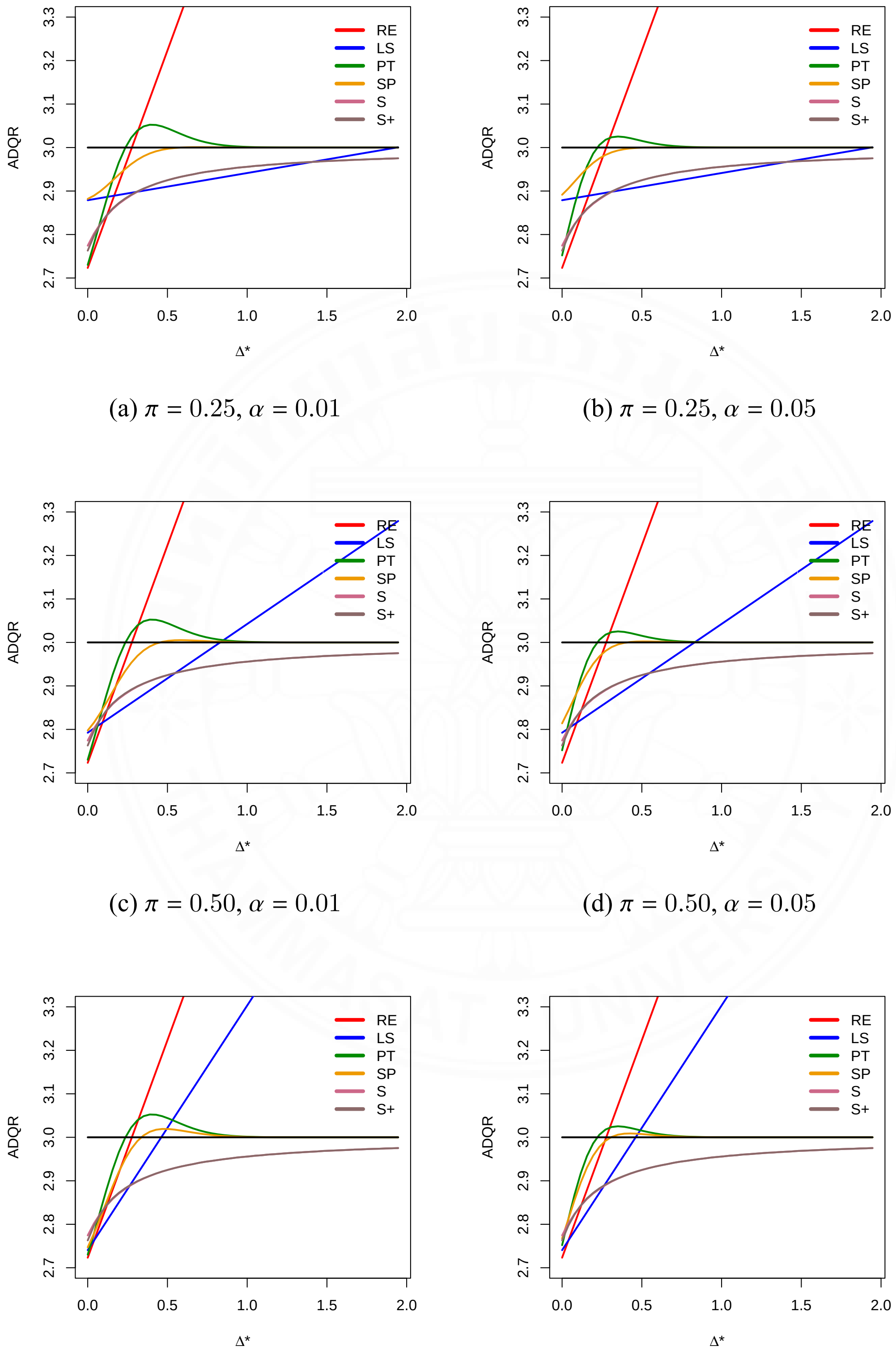


(e) $\pi = 0.75$, $\alpha = 0.01$



(f) $\pi = 0.75$, $\alpha = 0.05$

Figure 5.3 ADQR curves of the suggested estimators for Cox PH regression model with $p_1 = 3$ and $p_2 = 5$



(a) $\pi = 0.25$, $\alpha = 0.01$ (b) $\pi = 0.25$, $\alpha = 0.05$

(c) $\pi = 0.50$, $\alpha = 0.01$ (d) $\pi = 0.50$, $\alpha = 0.05$

(e) $\pi = 0.75$, $\alpha = 0.01$ (f) $\pi = 0.75$, $\alpha = 0.05$

Figure 5.4 ADQR curves of the suggested estimators for Cox PH regression model with $p_1 = 3$ and $p_2 = 11$

A RMSE larger than one indicates the degree of superiority of the estimator $\widehat{\boldsymbol{\beta}}_1^*$ over $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$.

5.5.1 Low-Dimensional Data

In the Cox PH regression model under a low-dimensional setting, the number of covariates is less than the sample size. We assumed that the p parameters contained p_1 active parameters and p_2 inactive parameters, such that $p_1 + p_2 = p$. To assess the behavior of the suggested estimators, we defined the degrees of model misspecification Δ^{sim} for representing the divergence between the simulation model and the restricted model under the null hypothesis $\mathrm{H}_0 : \boldsymbol{\beta}_2 = \mathbf{0}$ by

$$\Delta^{\mathrm{sim}} = \left\| \boldsymbol{\beta} - \boldsymbol{\beta}_{\mathbf{H}_0} \right\|.$$

Here, $\boldsymbol{\beta}$ is the coefficient vector of the true parameter for the simulation model and $\boldsymbol{\beta}_{\mathbf{H}_0}$ is the coefficient vector for the restricted model under the null hypothesis.

For a realistic situation, the regression coefficients $\boldsymbol{\beta}$ were assumed to scatter and we rearranged them to be active ($\boldsymbol{\beta}_1$) and inactive ($\boldsymbol{\beta}_2$) groups. Therefore, the coefficient vectors for the simulation model are in the following forms:

$$\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top = (\underbrace{\beta_1, \beta_2, ..., \beta_{p_1}}_{p_1}, \underbrace{\Delta^{\mathrm{sim}}, 0, ..., 0}_{p_2})^\top,$$

and the coefficient vectors of the restricted model are $\boldsymbol{\beta}_{\mathbf{H}_0} = (\boldsymbol{\beta}_1^\top, \mathbf{0}_{p_2}^\top)^\top$. If $\Delta^{\mathrm{sim}} = 0$, $\boldsymbol{\beta} = \boldsymbol{\beta}_{\mathbf{H}_0}$, it indicates that the restricted model was correct. If $\Delta^{\mathrm{sim}} > 0$, it means that the restricted model was incorrect.

Moreover, the choice of significance level (α) was fixed to 0.01, 0.05, and 0.1, while the shrinkage intensity (π) was set to 0.25, 0.50, and 0.75. In this study, we set the sample size (n) as 250, and the number of simulations (N) was 5,000 iterations, which was adequate to produce stable results.

For this simulation, we considered two cases for assessing the behavior of the proposed estimators when the size of Δ^{sim} is changed; one for the case when the null hypothesis was assumed to be true and other for when it may not be true.

5.5.1.1 Correct Subspace Information ($\Delta^{\text{sim}} = 0$)

In this case, we set the true value of the regression coefficients for the simulation model as $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top = ((0.16, -0.54, 0.23, 0.75, -0.39)^\top, \mathbf{0}_{p_2}^\top)^\top$, with $p_1 = 5$ and $p_2 = 3, 5, 7, 11$, and 15. For comparing the performance of the suggested estimators and two penalized estimators, the RMSEs results of the estimators for each censoring percentage are presented in Tables 5.1 to 5.2.

Table 5.1 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LASSO}}$, and $\widehat{\boldsymbol{\beta}}_1^{\text{aLASSO}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for Cox PH model with $pc = 20\%$ and $p_1 = 5$ at $\Delta^{\text{sim}} = 0$

Estimator			Number of Inactive Parameters (p_2)				
			3	5	7	11	15
RE			1.059	1.105	1.165	1.313	1.491
LS	$\pi = 0.25$		1.021	1.037	1.056	1.097	1.140
	$\pi = 0.50$		1.038	1.068	1.104	1.187	1.279
	$\pi = 0.75$		1.050	1.091	1.141	1.262	1.403
PT		$\alpha = 0.01$	1.054	1.097	1.154	1.279	1.446
		$\alpha = 0.05$	1.046	1.075	1.124	1.216	1.332
		$\alpha = 0.10$	1.037	1.062	1.103	1.175	1.257
SP	$\pi = 0.25$	$\alpha = 0.01$	1.019	1.034	1.052	1.088	1.128
		$\alpha = 0.05$	1.016	1.027	1.042	1.070	1.099
		$\alpha = 0.10$	1.013	1.022	1.035	1.057	1.079
	$\pi = 0.50$	$\alpha = 0.01$	1.035	1.062	1.096	1.168	1.254
		$\alpha = 0.05$	1.029	1.049	1.078	1.132	1.193
		$\alpha = 0.10$	1.023	1.040	1.065	1.108	1.152
	$\pi = 0.75$	$\alpha = 0.01$	1.047	1.084	1.131	1.234	1.365
		$\alpha = 0.05$	1.040	1.065	1.106	1.182	1.274
		$\alpha = 0.10$	1.032	1.054	1.087	1.148	1.213
S			1.012	1.046	1.092	1.201	1.343
S^+			1.017	1.053	1.100	1.209	1.348
LASSO			0.877	0.880	0.889	0.941	1.022
aLASSO			0.943	0.986	1.033	1.153	1.293

From Tables 5.1 and 5.2, we can summarize the results under the null hypothesis as follows:

1. All suggested estimators dominated the UE for all p_2 and censoring percentages, while the penalized estimators were inferior to the UE when p_2 was small for all pc.

2. The RMSEs of all estimators increased when p_2 increased and their RMSEs also increased when pc increased.

Table 5.2 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LASSO}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{aLASSO}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cox PH model with $pc = 30\%$ and $p_1 = 5$ at $\Delta^{\mathrm{sim}} = 0$

Estimator			Number of Inactive Parameters (p_2)				
			3	5	7	11	15
RE			1.065	1.112	1.177	1.320	1.513
LS	$\pi = 0.25$		1.022	1.039	1.059	1.099	1.144
	$\pi = 0.50$		1.041	1.071	1.110	1.192	1.289
	$\pi = 0.75$		1.055	1.096	1.150	1.269	1.419
PT		$\alpha = 0.01$	1.061	1.106	1.164	1.295	1.458
		$\alpha = 0.05$	1.049	1.083	1.135	1.233	1.347
		$\alpha = 0.10$	1.039	1.067	1.112	1.181	1.264
SP	$\pi = 0.25$	$\alpha = 0.01$	1.021	1.037	1.054	1.091	1.130
		$\alpha = 0.05$	1.017	1.029	1.045	1.074	1.102
		$\alpha = 0.10$	1.013	1.024	1.038	1.059	1.081
	$\pi = 0.50$	$\alpha = 0.01$	1.038	1.067	1.102	1.176	1.259
		$\alpha = 0.05$	1.031	1.053	1.084	1.140	1.200
		$\alpha = 0.10$	1.025	1.043	1.070	1.110	1.155
	$\pi = 0.75$	$\alpha = 0.01$	1.052	1.091	1.139	1.246	1.374
		$\alpha = 0.05$	1.042	1.071	1.114	1.195	1.284
		$\alpha = 0.10$	1.033	1.058	1.095	1.152	1.218
S			1.013	1.048	1.098	1.211	1.353
S+			1.019	1.057	1.107	1.217	1.357
LASSO			0.921	0.928	0.952	1.020	1.111
aLASSO			0.953	1.005	1.052	1.180	1.327

3. As expected, the RMSEs of the RE were highest for all cases of pc, which indicates that the RE outperformed all the other estimators.

4. For fixed pc and p_2, the RMSEs of the LS estimator increased to be equal to that of the RE when π increased to 1. Moreover, the RMSEs of the PT estimator increased as α decreased.

5. The SP estimator combines the LS and PT estimators, so it depends on π and α, and its performance was similar to the LS and PT estimators. For fixed pc, the RMSEs of the SP estimator increased when π increased but decreased when α increased.

6. At the same level of α, the performance of the PT estimator was always superior to that of the SP estimator because the PT estimator is a particular case of the SP estimator when $\pi = 1$.

7. When we fixed pc, the RMSEs of the PT estimator at $\alpha = 0.01$ was superior to that of the S and S$^+$ estimators for all p_2. While the performance of the PT estimator at $\alpha = 0.05$ and 0.10 outperformed the S and S$^+$ estimators when p_2 was small.

8. The S$^+$ estimator performed better than the shrinkage estimator for all p_2 for fixed pc.

9. Furthermore, the two penalized estimators were comparable to the suggested estimators when p_2 was large.

5.5.1.2 Uncertain Subspace Information ($\Delta^{\text{sim}} \geq 0$)

In this part of the study, the behavior of the proposed estimators was examined when ($\Delta^{\text{sim}} \geq 0$). The regression coefficient vector for the simulation model was set as $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top = ((-0.85, 0.27, 0.52)^\top, (\Delta^{\text{sim}}, \mathbf{0}_{p_2-1})^\top)^\top$, where $\Delta^{\text{sim}} \in [0, 0.8]$ with $(p_1, p_2) = (3,3), (3,5), (3,7), (3,11)$, and $(3,15)$.

In this case, the both LASSO and aLASSO estimators were not considered here because they do not take advantage of the fact that $\boldsymbol{\beta}$ is partitioned into active parameter vector $\boldsymbol{\beta}_1$ and inactive parameter vector $\boldsymbol{\beta}_2$, in which $\boldsymbol{\beta}_2 = \mathbf{0}$. The RMSE results for $pc = 20\%$ are reported in Tables 5.3 to 5.5 and for $pc = 30\%$ are displayed in Tables 5.6 to 5.8. The RMSEs are represented in Figures 5.5 to 5.9 for $pc = 20\%$ and in Figures 5.10 to 5.14 for $pc = 30\%$.

Table 5.3 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for Cox PH model with $pc = 20\%$, $p_1 = 3$, and $\pi = 0.25$ at $\Delta^{\text{sim}} \geq 0$

p_2	Δ^{sim}	RE	LS	PT			SP			S	S$^+$
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
3	0.0	1.037	1.014	1.035	1.027	1.024	1.013	1.010	1.009	1.012	1.014
	0.1	1.029	1.016	1.022	1.016	1.014	1.012	1.009	1.007	1.010	1.011
	0.2	0.988	1.023	0.988	0.992	0.994	1.005	1.001	1.000	1.007	1.007
	0.3	0.868	1.027	0.977	0.993	0.997	0.998	0.999	0.999	1.006	1.006
	0.4	0.684	1.020	0.996	1.000	1.000	0.999	1.000	1.000	1.005	1.005
	0.5	0.507	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.004	1.004
	0.6	0.371	0.962	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
	0.7	0.277	0.912	1.000	1.000	1.000	1.000	1.000	1.000	1.002	1.002
	0.8	0.217	0.862	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001

Table 5.3 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cox PH model with $pc = 20\%$, $p_1 = 3$, and $\pi = 0.25$ at $\Delta^{\mathrm{sim}} \geq 0$ (Cont.)

				PT			SP				
p_2	Δ^{sim}	RE	LS	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	S	S+
5	0.0	1.068	1.025	1.064	1.054	1.043	1.024	1.020	1.016	1.038	1.040
	0.1	1.058	1.028	1.046	1.032	1.025	1.023	1.016	1.012	1.031	1.034
	0.2	1.004	1.035	0.988	0.991	0.991	1.011	1.005	1.002	1.024	1.025
	0.3	0.869	1.040	0.964	0.987	0.994	0.998	0.999	0.999	1.020	1.020
	0.4	0.679	1.033	0.991	0.998	0.999	0.999	1.000	1.000	1.016	1.016
	0.5	0.500	1.012	1.000	1.000	1.000	1.000	1.000	1.000	1.013	1.013
	0.6	0.366	0.975	1.000	1.000	1.000	1.000	1.000	1.000	1.010	1.010
	0.7	0.274	0.924	1.000	1.000	1.000	1.000	1.000	1.000	1.007	1.007
	0.8	0.215	0.878	1.000	1.000	1.000	1.000	1.000	1.000	1.005	1.005
7	0.0	1.112	1.040	1.105	1.092	1.077	1.038	1.033	1.028	1.075	1.077
	0.1	1.102	1.045	1.082	1.060	1.046	1.036	1.027	1.021	1.064	1.068
	0.2	1.054	1.056	1.017	1.006	1.003	1.024	1.012	1.007	1.055	1.056
	0.3	0.915	1.066	0.963	0.983	0.993	1.002	0.999	1.000	1.045	1.045
	0.4	0.716	1.067	0.987	0.997	1.000	0.999	1.000	1.000	1.038	1.038
	0.5	0.526	1.052	0.999	1.000	1.000	1.000	1.000	1.000	1.032	1.032
	0.6	0.382	1.019	1.000	1.000	1.000	1.000	1.000	1.000	1.026	1.026
	0.7	0.285	0.971	1.000	1.000	1.000	1.000	1.000	1.000	1.021	1.021
	0.8	0.220	0.913	1.000	1.000	1.000	1.000	1.000	1.000	1.015	1.015
11	0.0	1.202	1.070	1.186	1.155	1.131	1.065	1.056	1.048	1.153	1.155
	0.1	1.185	1.073	1.159	1.120	1.096	1.064	1.049	1.039	1.137	1.140
	0.2	1.136	1.089	1.070	1.036	1.022	1.050	1.027	1.018	1.122	1.123
	0.3	0.992	1.106	0.974	0.988	0.993	1.016	1.005	1.002	1.104	1.104
	0.4	0.779	1.116	0.977	0.992	0.997	1.000	0.999	1.000	1.090	1.090
	0.5	0.574	1.110	0.995	1.000	1.000	0.999	1.000	1.000	1.077	1.077
	0.6	0.417	1.084	1.000	1.000	1.000	1.000	1.000	1.000	1.065	1.065
	0.7	0.309	1.040	1.000	1.000	1.000	1.000	1.000	1.000	1.053	1.053
	0.8	0.238	0.985	1.000	1.000	1.000	1.000	1.000	1.000	1.042	1.042
15	0.0	1.315	1.105	1.290	1.241	1.203	1.098	1.083	1.071	1.251	1.255
	0.1	1.307	1.113	1.261	1.193	1.151	1.098	1.075	1.060	1.236	1.241
	0.2	1.240	1.132	1.137	1.065	1.038	1.080	1.043	1.028	1.209	1.212
	0.3	1.073	1.155	0.986	0.989	0.993	1.029	1.011	1.006	1.182	1.182
	0.4	0.837	1.174	0.969	0.990	0.995	1.003	1.001	1.000	1.161	1.161
	0.5	0.613	1.179	0.990	0.998	0.998	0.999	1.000	1.000	1.142	1.142
	0.6	0.445	1.165	0.998	1.000	1.000	1.000	1.000	1.000	1.124	1.124
	0.7	0.329	1.128	1.000	1.000	1.000	1.000	1.000	1.000	1.107	1.107
	0.8	0.252	1.073	1.000	1.000	1.000	1.000	1.000	1.000	1.089	1.089

Table 5.4 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for Cox PH model with $pc = 20\%$, $p_1 = 3$, and $\pi = 0.50$ at $\Delta^{\text{sim}} \geq 0$

				PT			SP				
p_2	Δ^{sim}	RE	LS	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	S	S+
3	0.0	1.037	1.025	1.035	1.027	1.024	1.023	1.019	1.016	1.012	1.014
	0.1	1.029	1.027	1.022	1.016	1.014	1.020	1.014	1.011	1.010	1.011
	0.2	0.988	1.029	0.988	0.992	0.994	1.005	1.000	1.000	1.007	1.007
	0.3	0.868	1.010	0.977	0.993	0.997	0.993	0.998	0.999	1.006	1.006
	0.4	0.684	0.952	0.996	1.000	1.000	0.998	1.000	1.000	1.005	1.005
	0.5	0.507	0.860	1.000	1.000	1.000	1.000	1.000	1.000	1.004	1.004
	0.6	0.371	0.749	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
	0.7	0.277	0.639	1.000	1.000	1.000	1.000	1.000	1.000	1.002	1.002
	0.8	0.217	0.551	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001
5	0.0	1.068	1.045	1.064	1.054	1.043	1.043	1.036	1.029	1.038	1.040
	0.1	1.058	1.048	1.046	1.032	1.025	1.039	1.027	1.021	1.031	1.034
	0.2	1.004	1.048	0.988	0.991	0.991	1.013	1.005	1.001	1.024	1.025
	0.3	0.869	1.027	0.964	0.987	0.994	0.991	0.996	0.998	1.020	1.020
	0.4	0.679	0.966	0.991	0.998	0.999	0.997	0.999	1.000	1.016	1.016
	0.5	0.500	0.869	1.000	1.000	1.000	1.000	1.000	1.000	1.013	1.013
	0.6	0.366	0.756	1.000	1.000	1.000	1.000	1.000	1.000	1.010	1.010
	0.7	0.274	0.644	1.000	1.000	1.000	1.000	1.000	1.000	1.007	1.007
	0.8	0.215	0.558	1.000	1.000	1.000	1.000	1.000	1.000	1.005	1.005
7	0.0	1.112	1.073	1.105	1.092	1.077	1.069	1.060	1.050	1.075	1.077
	0.1	1.102	1.078	1.082	1.060	1.046	1.063	1.046	1.035	1.064	1.068
	0.2	1.054	1.085	1.017	1.006	1.003	1.035	1.017	1.010	1.055	1.056
	0.3	0.915	1.071	0.963	0.983	0.993	0.996	0.996	0.999	1.045	1.045
	0.4	0.716	1.018	0.987	0.997	1.000	0.997	0.999	1.000	1.038	1.038
	0.5	0.526	0.922	0.999	1.000	1.000	1.000	1.000	1.000	1.032	1.032
	0.6	0.382	0.803	1.000	1.000	1.000	1.000	1.000	1.000	1.026	1.026
	0.7	0.285	0.686	1.000	1.000	1.000	1.000	1.000	1.000	1.021	1.021
	0.8	0.220	0.582	1.000	1.000	1.000	1.000	1.000	1.000	1.015	1.015
11	0.0	1.202	1.131	1.186	1.155	1.131	1.121	1.102	1.087	1.153	1.155
	0.1	1.185	1.132	1.159	1.120	1.096	1.114	1.087	1.070	1.137	1.140
	0.2	1.136	1.149	1.070	1.036	1.022	1.080	1.043	1.028	1.122	1.123
	0.3	0.992	1.145	0.974	0.988	0.993	1.017	1.004	1.002	1.104	1.104
	0.4	0.779	1.100	0.977	0.992	0.997	0.996	0.998	0.999	1.090	1.090
	0.5	0.574	1.008	0.995	1.000	1.000	0.998	1.000	1.000	1.077	1.077
	0.6	0.417	0.886	1.000	1.000	1.000	1.000	1.000	1.000	1.065	1.065
	0.7	0.309	0.757	1.000	1.000	1.000	1.000	1.000	1.000	1.053	1.053
	0.8	0.238	0.644	1.000	1.000	1.000	1.000	1.000	1.000	1.042	1.042

Table 5.4 RMSEs of $\widehat{\beta}_1^{\mathrm{RE}}$, $\widehat{\beta}_1^{\mathrm{LS}}$, $\widehat{\beta}_1^{\mathrm{PT}}$, $\widehat{\beta}_1^{\mathrm{SP}}$, $\widehat{\beta}_1^{\mathrm{S}}$, and $\widehat{\beta}_1^{\mathrm{S+}}$ with respect to $\widehat{\beta}_1^{\mathrm{UE}}$ for Cox PH model with pc = 20%, p_1 = 3, and π = 0.50 at $\Delta^{\mathrm{sim}} \geq 0$ (Cont.)

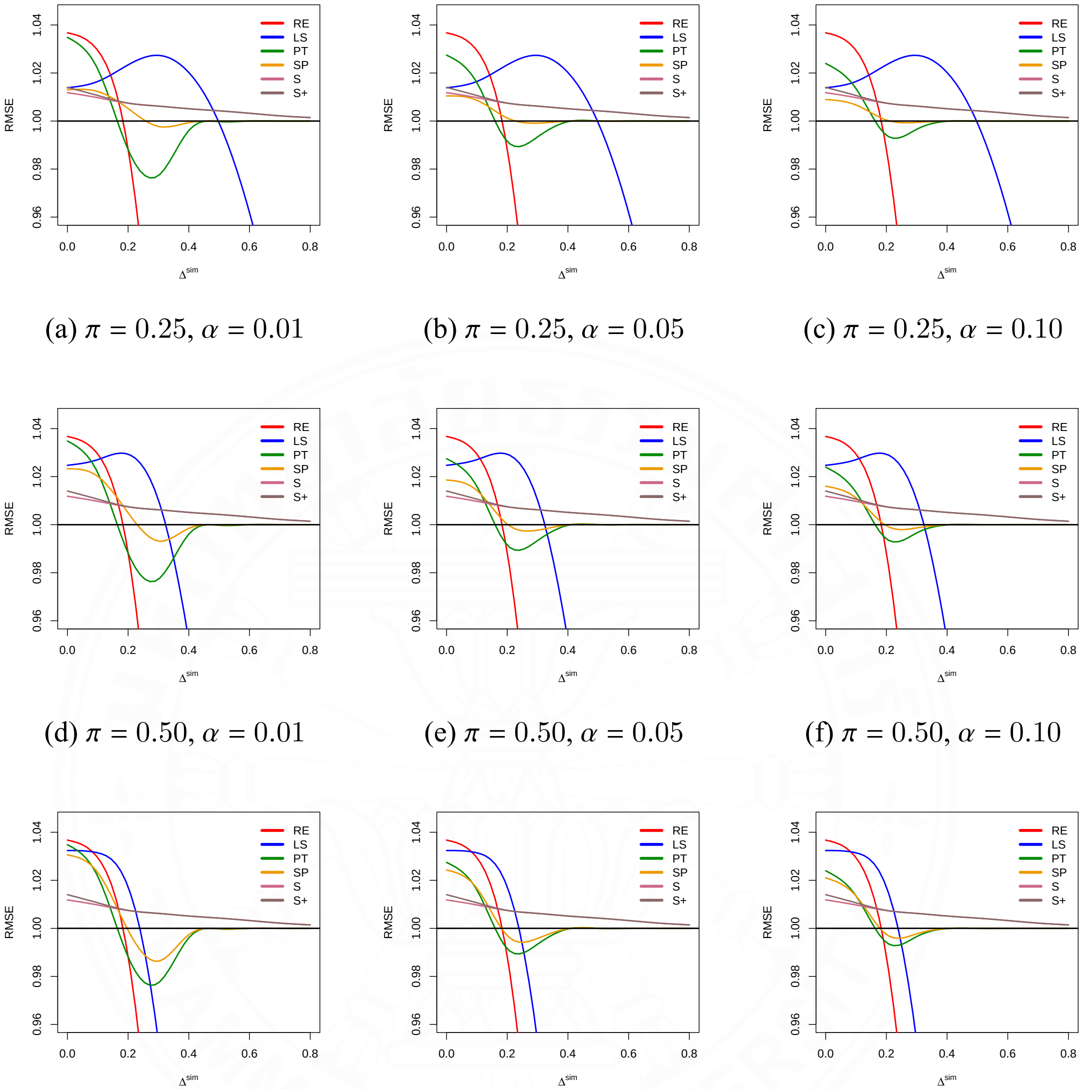
p_2	Δ^{sim}	RE	LS	PT α = 0.01	PT α = 0.05	PT α = 0.10	SP α = 0.01	SP α = 0.05	SP α = 0.10	S	S+
15	0.0	1.315	1.200	1.290	1.241	1.203	1.185	1.156	1.132	1.251	1.255
	0.1	1.307	1.211	1.261	1.193	1.151	1.181	1.137	1.108	1.236	1.241
	0.2	1.240	1.230	1.137	1.065	1.038	1.135	1.070	1.044	1.209	1.212
	0.3	1.073	1.234	0.986	0.989	0.993	1.037	1.013	1.007	1.182	1.182
	0.4	0.837	1.198	0.969	0.990	0.995	0.999	0.999	0.999	1.161	1.161
	0.5	0.613	1.108	0.990	0.998	0.998	0.997	0.999	0.999	1.142	1.142
	0.6	0.445	0.981	0.998	1.000	1.000	0.999	1.000	1.000	1.124	1.124
	0.7	0.329	0.841	1.000	1.000	1.000	1.000	1.000	1.000	1.107	1.107
	0.8	0.252	0.712	1.000	1.000	1.000	1.000	1.000	1.000	1.089	1.089

Table 5.5 RMSEs of $\widehat{\beta}_1^{\mathrm{RE}}$, $\widehat{\beta}_1^{\mathrm{LS}}$, $\widehat{\beta}_1^{\mathrm{PT}}$, $\widehat{\beta}_1^{\mathrm{SP}}$, $\widehat{\beta}_1^{\mathrm{S}}$, and $\widehat{\beta}_1^{\mathrm{S+}}$ with respect to $\widehat{\beta}_1^{\mathrm{UE}}$ for Cox PH model with pc = 20%, p_1 = 3, and π = 0.75 at $\Delta^{\mathrm{sim}} \geq 0$

p_2	Δ^{sim}	RE	LS	PT α = 0.01	PT α = 0.05	PT α = 0.10	SP α = 0.01	SP α = 0.05	SP α = 0.10	S	S+
3	0.0	1.037	1.032	1.035	1.027	1.024	1.031	1.024	1.021	1.012	1.014
	0.1	1.029	1.031	1.022	1.016	1.014	1.023	1.017	1.013	1.010	1.011
	0.2	0.988	1.017	0.988	0.992	0.994	0.999	0.997	0.997	1.007	1.007
	0.3	0.868	0.953	0.977	0.993	0.997	0.986	0.996	0.998	1.006	1.006
	0.4	0.684	0.826	0.996	1.000	1.000	0.997	1.000	1.000	1.005	1.005
	0.5	0.507	0.672	1.000	1.000	1.000	1.000	1.000	1.000	1.004	1.004
	0.6	0.371	0.530	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
	0.7	0.277	0.416	1.000	1.000	1.000	1.000	1.000	1.000	1.002	1.002
	0.8	0.217	0.338	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001
5	0.0	1.068	1.059	1.064	1.054	1.043	1.056	1.048	1.038	1.038	1.040
	0.1	1.058	1.059	1.046	1.032	1.025	1.046	1.033	1.025	1.031	1.034
	0.2	1.004	1.037	0.988	0.991	0.991	1.005	1.000	0.997	1.024	1.025
	0.3	0.869	0.964	0.964	0.987	0.994	0.980	0.993	0.996	1.020	1.020
	0.4	0.679	0.831	0.991	0.998	0.999	0.994	0.998	0.999	1.016	1.016
	0.5	0.500	0.672	1.000	1.000	1.000	1.000	1.000	1.000	1.013	1.013
	0.6	0.366	0.529	1.000	1.000	1.000	1.000	1.000	1.000	1.010	1.010
	0.7	0.274	0.415	1.000	1.000	1.000	1.000	1.000	1.000	1.007	1.007
	0.8	0.215	0.338	1.000	1.000	1.000	1.000	1.000	1.000	1.005	1.005

Table 5.5 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\text{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\text{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\text{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\text{UE}}$ for Cox PH model with $pc = 20\%$, $p_1 = 3$, and $\pi = 0.75$ at $\Delta^{\text{sim}} \geq 0$ (Cont.)

p_2	Δ^{sim}	RE	LS	PT			SP			S	S^+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
7	0.0	1.112	1.097	1.105	1.092	1.077	1.091	1.080	1.067	1.075	1.077
	0.1	1.102	1.098	1.082	1.060	1.046	1.078	1.057	1.044	1.064	1.068
	0.2	1.054	1.085	1.017	1.006	1.003	1.033	1.015	1.008	1.055	1.056
	0.3	0.915	1.016	0.963	0.983	0.993	0.983	0.991	0.996	1.045	1.045
	0.4	0.716	0.881	0.987	0.997	1.000	0.993	0.998	1.000	1.038	1.038
	0.5	0.526	0.714	0.999	1.000	1.000	0.999	1.000	1.000	1.032	1.032
	0.6	0.382	0.559	1.000	1.000	1.000	1.000	1.000	1.000	1.026	1.026
	0.7	0.285	0.438	1.000	1.000	1.000	1.000	1.000	1.000	1.021	1.021
	0.8	0.220	0.349	1.000	1.000	1.000	1.000	1.000	1.000	1.015	1.015
11	0.0	1.202	1.176	1.186	1.155	1.131	1.162	1.136	1.116	1.153	1.155
	0.1	1.185	1.171	1.159	1.120	1.096	1.147	1.111	1.089	1.137	1.140
	0.2	1.136	1.165	1.070	1.036	1.022	1.087	1.046	1.029	1.122	1.123
	0.3	0.992	1.102	0.974	0.988	0.993	1.002	0.999	0.999	1.104	1.104
	0.4	0.779	0.963	0.977	0.992	0.997	0.988	0.996	0.998	1.090	1.090
	0.5	0.574	0.786	0.995	1.000	1.000	0.997	1.000	1.000	1.077	1.077
	0.6	0.417	0.617	1.000	1.000	1.000	1.000	1.000	1.000	1.065	1.065
	0.7	0.309	0.481	1.000	1.000	1.000	1.000	1.000	1.000	1.053	1.053
	0.8	0.238	0.382	1.000	1.000	1.000	1.000	1.000	1.000	1.042	1.042
15	0.0	1.315	1.273	1.290	1.241	1.203	1.252	1.211	1.177	1.251	1.255
	0.1	1.307	1.280	1.261	1.193	1.151	1.239	1.178	1.140	1.236	1.241
	0.2	1.240	1.271	1.137	1.065	1.038	1.155	1.077	1.047	1.209	1.212
	0.3	1.073	1.202	0.986	0.989	0.993	1.022	1.005	1.002	1.182	1.182
	0.4	0.837	1.054	0.969	0.990	0.995	0.987	0.996	0.998	1.161	1.161
	0.5	0.613	0.860	0.990	0.998	0.998	0.994	0.999	0.999	1.142	1.142
	0.6	0.445	0.675	0.998	1.000	1.000	0.998	1.000	1.000	1.124	1.124
	0.7	0.329	0.524	1.000	1.000	1.000	1.000	1.000	1.000	1.107	1.107
	0.8	0.252	0.414	1.000	1.000	1.000	1.000	1.000	1.000	1.089	1.089

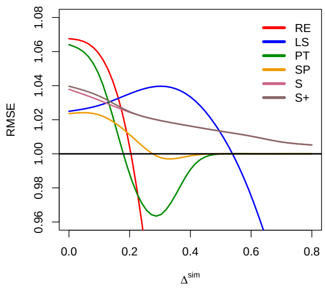


(a) $\pi = 0.25, \alpha = 0.01$ (b) $\pi = 0.25, \alpha = 0.05$ (c) $\pi = 0.25, \alpha = 0.10$

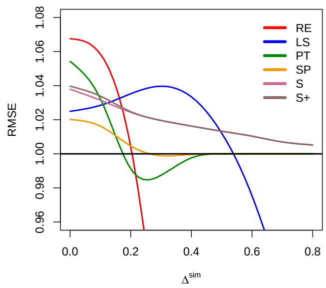
(d) $\pi = 0.50, \alpha = 0.01$ (e) $\pi = 0.50, \alpha = 0.05$ (f) $\pi = 0.50, \alpha = 0.10$

(g) $\pi = 0.75, \alpha = 0.01$ (h) $\pi = 0.75, \alpha = 0.05$ (i) $\pi = 0.75, \alpha = 0.10$

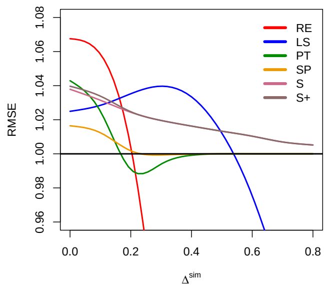
Figure 5.5 RMSEs of $\widehat{\beta}_1^{\text{RE}}$, $\widehat{\beta}_1^{\text{LS}}$, $\widehat{\beta}_1^{\text{PT}}$, $\widehat{\beta}_1^{\text{SP}}$, $\widehat{\beta}_1^{\text{S}}$, and $\widehat{\beta}_1^{\text{S+}}$ with respect to $\widehat{\beta}_1^{\text{UE}}$ for Cox PH model with $pc = 20\%$, $p_1 = 3$, and $p_2 - 1 = 2$ at $\Delta^{\text{sim}} \geq 0$



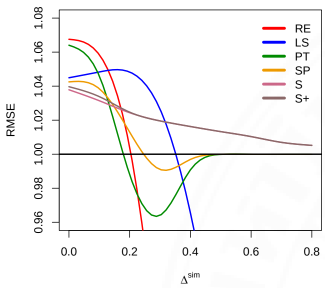
(a) $\pi = 0.25, \alpha = 0.01$



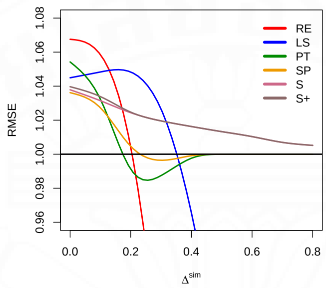
(b) $\pi = 0.25, \alpha = 0.05$



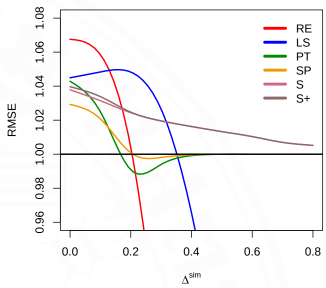
(c) $\pi = 0.25, \alpha = 0.10$



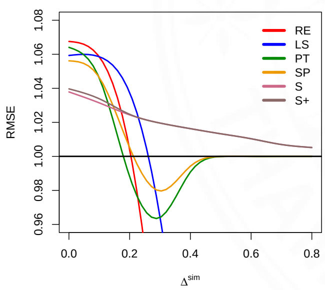
(d) $\pi = 0.50, \alpha = 0.01$



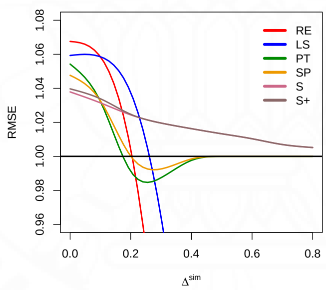
(e) $\pi = 0.50, \alpha = 0.05$



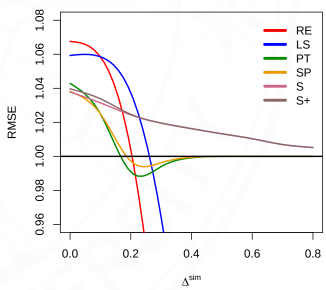
(f) $\pi = 0.50, \alpha = 0.10$



(g) $\pi = 0.75, \alpha = 0.01$

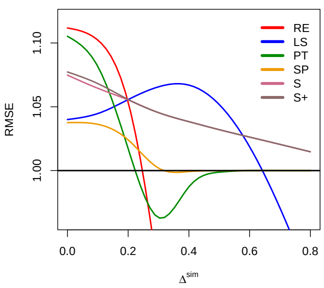


(h) $\pi = 0.75, \alpha = 0.05$

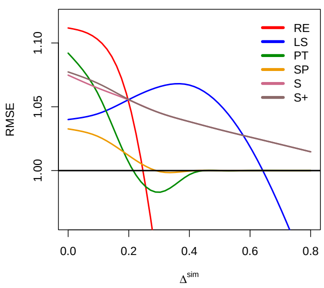


(i) $\pi = 0.75, \alpha = 0.10$

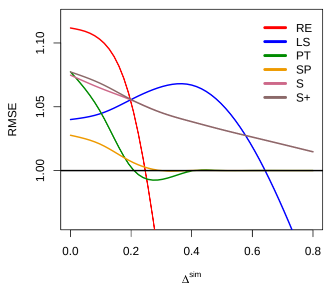
Figure 5.6 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cox PH model with $pc = 20\%$, $p_1 = 3$, and $p_2 - 1 = 4$ at $\Delta^{\mathrm{sim}} \geq 0$



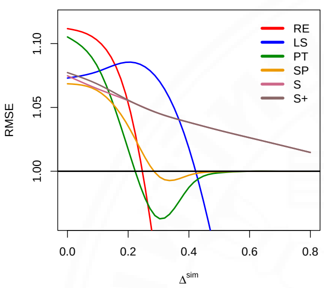
(a) $\pi = 0.25$, $\alpha = 0.01$



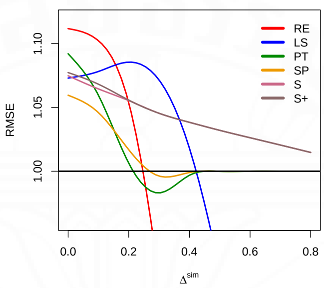
(b) $\pi = 0.25$, $\alpha = 0.05$



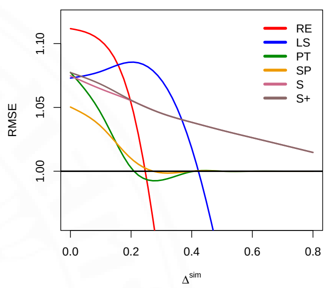
(c) $\pi = 0.25$, $\alpha = 0.10$



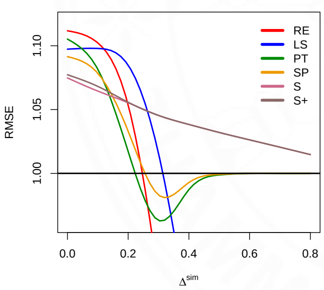
(d) $\pi = 0.50$, $\alpha = 0.01$



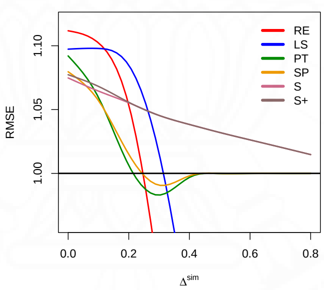
(e) $\pi = 0.50$, $\alpha = 0.05$



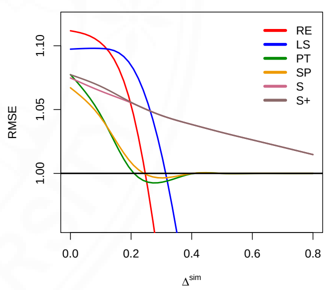
(f) $\pi = 0.50$, $\alpha = 0.10$



(g) $\pi = 0.75$, $\alpha = 0.01$

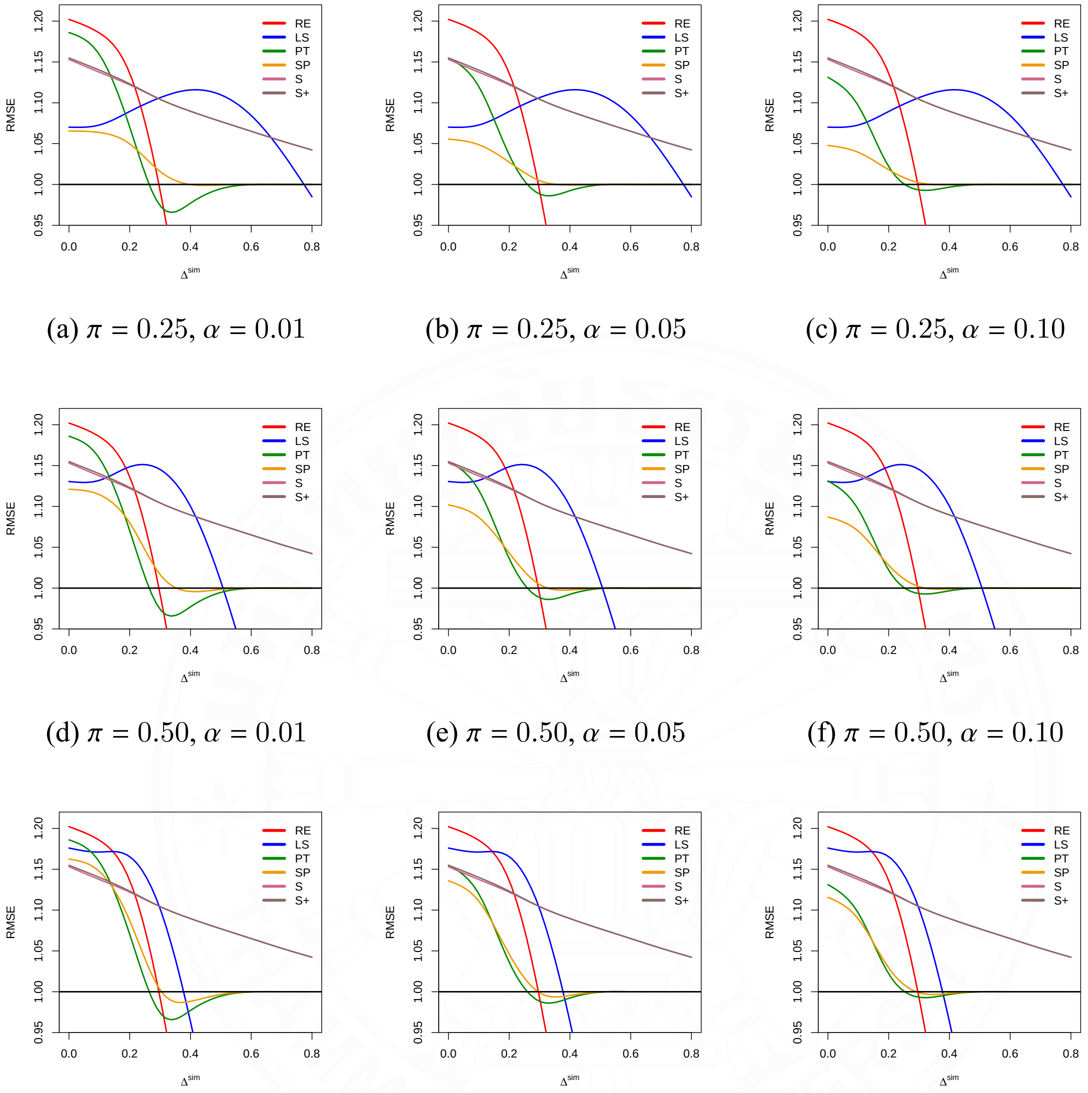


(h) $\pi = 0.75$, $\alpha = 0.05$



(i) $\pi = 0.75$, $\alpha = 0.10$

Figure 5.7 RMSEs of $\widehat{\beta}_1^{\text{RE}}$, $\widehat{\beta}_1^{\text{LS}}$, $\widehat{\beta}_1^{\text{PT}}$, $\widehat{\beta}_1^{\text{SP}}$, $\widehat{\beta}_1^{\text{S}}$, and $\widehat{\beta}_1^{\text{S+}}$ with respect to $\widehat{\beta}_1^{\text{UE}}$ for Cox PH model with $pc = 20\%$, $p_1 = 3$, and $p_2 - 1 = 6$ at $\Delta^{\text{sim}} \geq 0$

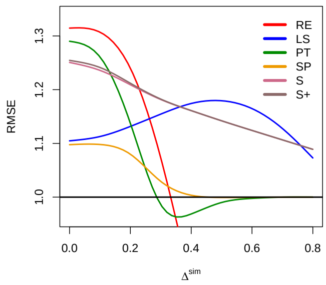


(a) $\pi = 0.25$, $\alpha = 0.01$ (b) $\pi = 0.25$, $\alpha = 0.05$ (c) $\pi = 0.25$, $\alpha = 0.10$

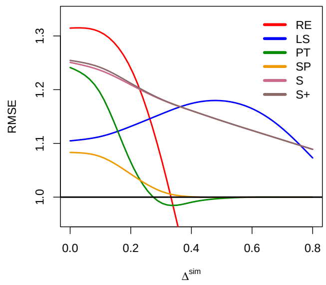
(d) $\pi = 0.50$, $\alpha = 0.01$ (e) $\pi = 0.50$, $\alpha = 0.05$ (f) $\pi = 0.50$, $\alpha = 0.10$

(g) $\pi = 0.75$, $\alpha = 0.01$ (h) $\pi = 0.75$, $\alpha = 0.05$ (i) $\pi = 0.75$, $\alpha = 0.10$

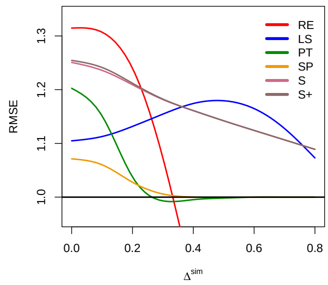
Figure 5.8 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cox PH model with $pc = 20\%$, $p_1 = 3$, and $p_2 - 1 = 10$ at $\Delta^{\mathrm{sim}} \geq 0$



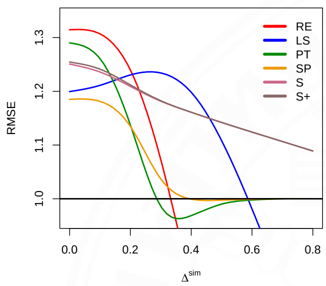
(a) $\pi = 0.25$, $\alpha = 0.01$



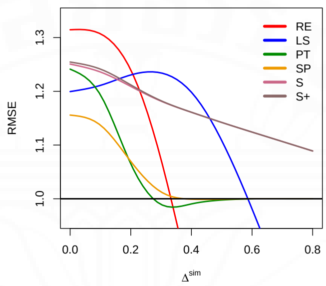
(b) $\pi = 0.25$, $\alpha = 0.05$



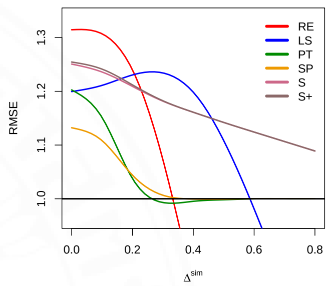
(c) $\pi = 0.25$, $\alpha = 0.10$



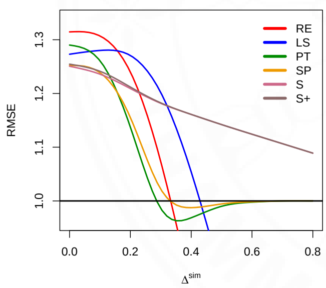
(d) $\pi = 0.50$, $\alpha = 0.01$



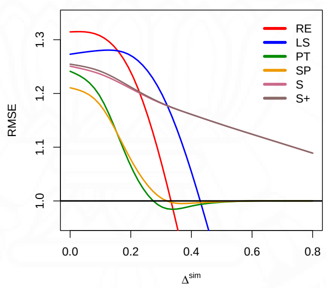
(e) $\pi = 0.50$, $\alpha = 0.05$



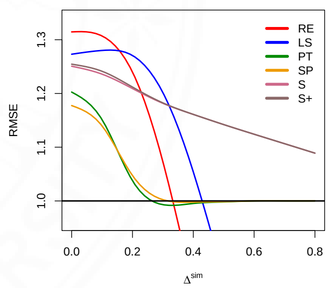
(f) $\pi = 0.50$, $\alpha = 0.10$



(g) $\pi = 0.75$, $\alpha = 0.01$



(h) $\pi = 0.75$, $\alpha = 0.05$



(i) $\pi = 0.75$, $\alpha = 0.10$

Figure 5.9 RMSEs of $\widehat{\beta}_1^{\text{RE}}$, $\widehat{\beta}_1^{\text{LS}}$, $\widehat{\beta}_1^{\text{PT}}$, $\widehat{\beta}_1^{\text{SP}}$, $\widehat{\beta}_1^{\text{S}}$, and $\widehat{\beta}_1^{\text{S+}}$ with respect to $\widehat{\beta}_1^{\text{UE}}$ for Cox PH model with $pc = 20\%$, $p_1 = 3$, and $p_2 - 1 = 14$ at $\Delta^{\text{sim}} \geq 0$

Table 5.6 RMSEs of $\widehat{\beta}_1^{\mathrm{RE}}$, $\widehat{\beta}_1^{\mathrm{LS}}$, $\widehat{\beta}_1^{\mathrm{PT}}$, $\widehat{\beta}_1^{\mathrm{SP}}$, $\widehat{\beta}_1^{\mathrm{S}}$, and $\widehat{\beta}_1^{\mathrm{S+}}$ with respect to $\widehat{\beta}_1^{\mathrm{UE}}$ for Cox PH model with $pc = 30\%$, $p_1 = 3$, and $\pi = 0.25$ at $\Delta^{\mathrm{sim}} \geq 0$

p_2	Δ^{sim}	RE	LS	PT			SP			S	S^+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
3	0.0	1.042	1.015	1.039	1.032	1.026	1.014	1.011	1.009	1.014	1.015
	0.1	1.038	1.018	1.029	1.019	1.013	1.014	1.009	1.006	1.010	1.012
	0.2	1.005	1.025	0.987	0.995	0.995	1.005	1.003	1.001	1.008	1.008
	0.3	0.900	1.030	0.977	0.990	0.994	0.999	0.999	0.999	1.007	1.007
	0.4	0.730	1.026	0.989	0.998	0.999	0.998	1.000	1.000	1.006	1.006
	0.5	0.553	1.008	0.998	1.000	1.000	1.000	1.000	1.000	1.005	1.005
	0.6	0.412	0.977	1.000	1.000	1.000	1.000	1.000	1.000	1.004	1.004
	0.7	0.311	0.934	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
	0.8	0.243	0.883	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001
5	0.0	1.077	1.027	1.071	1.060	1.051	1.025	1.021	1.018	1.045	1.046
	0.1	1.066	1.029	1.055	1.040	1.029	1.024	1.018	1.013	1.035	1.037
	0.2	1.028	1.038	1.004	1.000	0.996	1.015	1.007	1.003	1.029	1.029
	0.3	0.915	1.046	0.967	0.986	0.993	1.001	0.999	1.000	1.024	1.024
	0.4	0.738	1.046	0.983	0.997	0.999	0.998	1.000	1.000	1.021	1.021
	0.5	0.555	1.031	0.998	1.000	1.000	1.000	1.000	1.000	1.017	1.017
	0.6	0.412	1.002	1.000	1.000	1.000	1.000	1.000	1.000	1.014	1.014
	0.7	0.312	0.961	1.000	1.000	1.000	1.000	1.000	1.000	1.011	1.011
	0.8	0.242	0.908	1.000	1.000	1.000	1.000	1.000	1.000	1.007	1.007
7	0.0	1.112	1.040	1.105	1.092	1.077	1.038	1.033	1.028	1.075	1.077
	0.1	1.101	1.042	1.086	1.069	1.054	1.036	1.028	1.022	1.068	1.070
	0.2	1.058	1.052	1.024	1.010	1.008	1.027	1.014	1.009	1.055	1.056
	0.3	0.937	1.062	0.965	0.982	0.989	1.005	1.000	1.000	1.045	1.045
	0.4	0.752	1.064	0.974	0.995	0.997	0.998	1.000	1.000	1.037	1.037
	0.5	0.566	1.051	0.995	0.999	1.000	0.999	1.000	1.000	1.031	1.031
	0.6	0.419	1.023	1.000	1.000	1.000	1.000	1.000	1.000	1.025	1.025
	0.7	0.316	0.981	1.000	1.000	1.000	1.000	1.000	1.000	1.019	1.019
	0.8	0.246	0.931	1.000	1.000	1.000	1.000	1.000	1.000	1.013	1.013
11	0.0	1.205	1.072	1.188	1.157	1.134	1.066	1.056	1.048	1.155	1.158
	0.1	1.192	1.073	1.169	1.128	1.100	1.064	1.050	1.040	1.144	1.146
	0.2	1.145	1.086	1.085	1.046	1.033	1.053	1.031	1.022	1.125	1.126
	0.3	1.017	1.102	0.986	0.988	0.995	1.022	1.008	1.005	1.106	1.106
	0.4	0.818	1.112	0.966	0.989	0.995	1.002	1.000	1.000	1.090	1.090
	0.5	0.616	1.108	0.987	0.998	1.000	0.999	1.000	1.000	1.077	1.077
	0.6	0.456	1.086	0.999	1.000	1.000	1.000	1.000	1.000	1.065	1.065
	0.7	0.342	1.048	1.000	1.000	1.000	1.000	1.000	1.000	1.053	1.053
	0.8	0.265	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.041	1.041

Table 5.6 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cox PH model with $pc = 30\%$, $p_1 = 3$, and $\pi = 0.25$ at $\Delta^{\mathrm{sim}} \geq 0$ (Cont.)

p_2	Δ^{sim}	RE	LS	PT			SP			S	S^{+}
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
15	0.0	1.322	1.106	1.293	1.252	1.206	1.098	1.085	1.071	1.254	1.257
	0.1	1.312	1.112	1.268	1.205	1.153	1.098	1.077	1.060	1.242	1.245
	0.2	1.251	1.127	1.163	1.090	1.061	1.085	1.050	1.035	1.212	1.216
	0.3	1.104	1.148	1.009	0.996	0.995	1.042	1.016	1.009	1.184	1.185
	0.4	0.885	1.166	0.958	0.982	0.993	1.006	1.001	1.000	1.161	1.161
	0.5	0.664	1.172	0.979	0.995	0.998	0.999	1.000	1.000	1.141	1.141
	0.6	0.491	1.161	0.995	0.999	0.999	0.999	1.000	1.000	1.122	1.122
	0.7	0.367	1.130	0.999	1.000	1.000	1.000	1.000	1.000	1.103	1.103
	0.8	0.284	1.083	1.000	1.000	1.000	1.000	1.000	1.000	1.086	1.086

Table 5.7 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cox PH model with $pc = 30\%$, $p_1 = 3$, and $\pi = 0.50$ at $\Delta^{\mathrm{sim}} \geq 0$

p_2	Δ^{sim}	RE	LS	PT			SP			S	S^{+}
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
3	0.0	1.042	1.027	1.039	1.032	1.026	1.025	1.021	1.017	1.014	1.015
	0.1	1.038	1.030	1.029	1.019	1.013	1.023	1.015	1.011	1.010	1.012
	0.2	1.005	1.035	0.987	0.995	0.995	1.005	1.003	1.000	1.008	1.008
	0.3	0.900	1.020	0.977	0.990	0.994	0.994	0.996	0.998	1.007	1.007
	0.4	0.730	0.973	0.989	0.998	0.999	0.996	0.999	1.000	1.006	1.006
	0.5	0.553	0.890	0.998	1.000	1.000	0.999	1.000	1.000	1.005	1.005
	0.6	0.412	0.788	1.000	1.000	1.000	1.000	1.000	1.000	1.004	1.004
	0.7	0.311	0.683	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
	0.8	0.243	0.589	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001
5	0.0	1.077	1.049	1.071	1.060	1.051	1.045	1.038	1.033	1.045	1.046
	0.1	1.066	1.050	1.055	1.040	1.029	1.042	1.031	1.022	1.035	1.037
	0.2	1.028	1.056	1.004	1.000	0.996	1.021	1.010	1.004	1.029	1.029
	0.3	0.915	1.045	0.967	0.986	0.993	0.996	0.997	0.998	1.024	1.024
	0.4	0.738	0.999	0.983	0.997	0.999	0.995	0.999	1.000	1.021	1.021
	0.5	0.555	0.914	0.998	1.000	1.000	0.999	1.000	1.000	1.017	1.017
	0.6	0.412	0.810	1.000	1.000	1.000	1.000	1.000	1.000	1.014	1.014
	0.7	0.312	0.703	1.000	1.000	1.000	1.000	1.000	1.000	1.011	1.011
	0.8	0.242	0.603	1.000	1.000	1.000	1.000	1.000	1.000	1.007	1.007
7	0.0	1.112	1.073	1.105	1.092	1.077	1.069	1.060	1.050	1.075	1.077
	0.1	1.101	1.075	1.086	1.069	1.054	1.063	1.050	1.039	1.068	1.070
	0.2	1.058	1.081	1.024	1.010	1.008	1.040	1.020	1.014	1.055	1.056
	0.3	0.937	1.071	0.965	0.982	0.989	1.001	0.997	0.998	1.045	1.045
	0.4	0.752	1.026	0.974	0.995	0.997	0.992	0.999	0.999	1.037	1.037

Table 5.7 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cox PH model with $pc = 30\%$, $p_1 = 3$, and $\pi = 0.50$ at $\Delta^{\mathrm{sim}} \geq 0$ (Cont.)

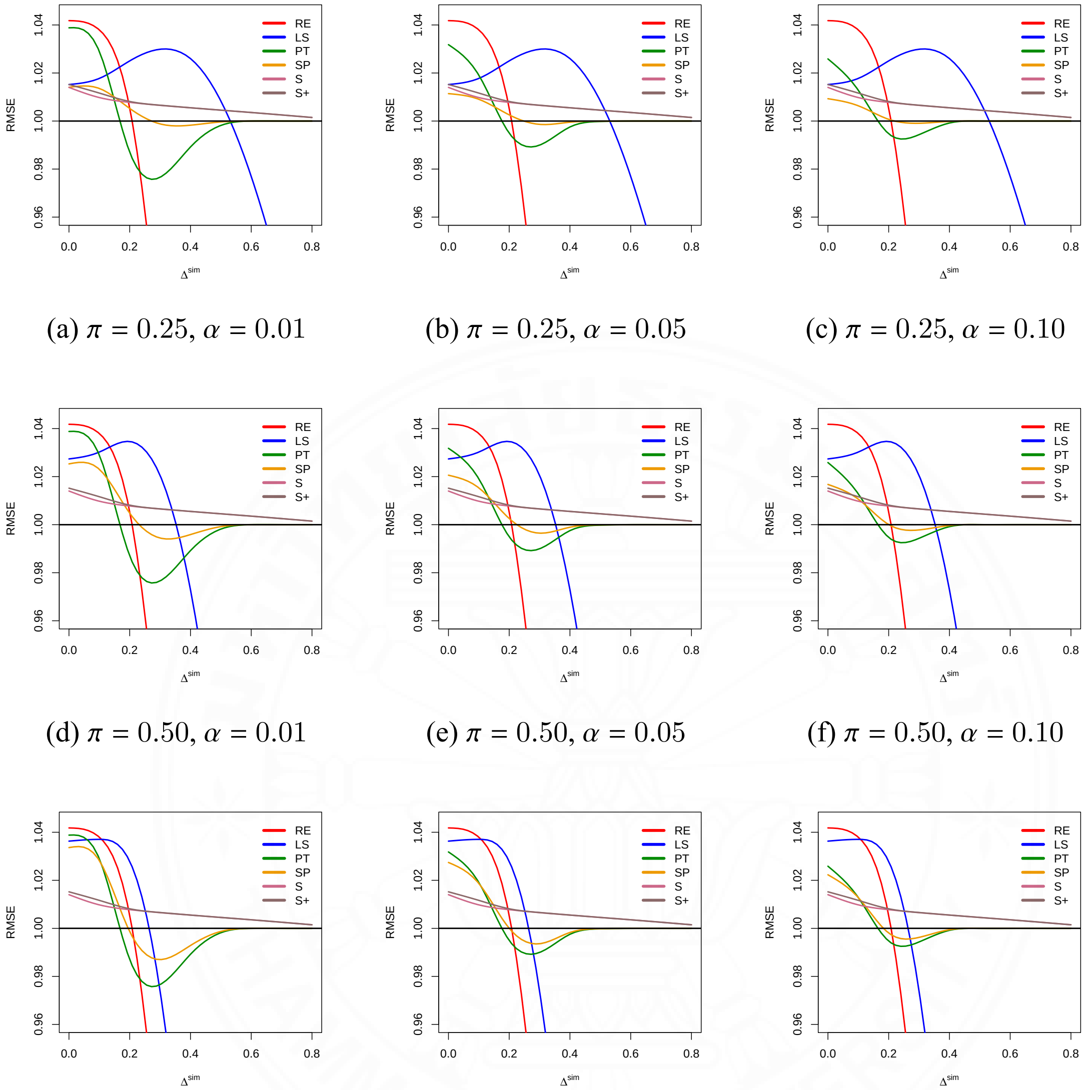
p_2	Δ^{sim}	RE	LS	PT			SP			S	S^+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
7	0.5	0.566	0.941	0.995	0.999	1.000	0.998	1.000	1.000	1.031	1.031
	0.6	0.419	0.832	1.000	1.000	1.000	1.000	1.000	1.000	1.025	1.025
	0.7	0.316	0.722	1.000	1.000	1.000	1.000	1.000	1.000	1.019	1.019
	0.8	0.246	0.622	1.000	1.000	1.000	1.000	1.000	1.000	1.013	1.013
11	0.0	1.205	1.133	1.188	1.157	1.134	1.122	1.103	1.089	1.155	1.158
	0.1	1.192	1.133	1.169	1.128	1.100	1.117	1.090	1.071	1.144	1.146
	0.2	1.145	1.146	1.085	1.046	1.033	1.087	1.049	1.035	1.125	1.126
	0.3	1.017	1.144	0.986	0.988	0.995	1.028	1.008	1.005	1.106	1.106
	0.4	0.818	1.107	0.966	0.989	0.995	0.997	0.998	0.999	1.090	1.090
	0.5	0.616	1.026	0.987	0.998	1.000	0.996	0.999	1.000	1.077	1.077
	0.6	0.456	0.915	0.999	1.000	1.000	1.000	1.000	1.000	1.065	1.065
	0.7	0.342	0.794	1.000	1.000	1.000	1.000	1.000	1.000	1.053	1.053
	0.8	0.265	0.684	1.000	1.000	1.000	1.000	1.000	1.000	1.041	1.041
15	0.0	1.322	1.202	1.293	1.252	1.206	1.185	1.160	1.133	1.254	1.257
	0.1	1.312	1.210	1.268	1.205	1.153	1.182	1.141	1.108	1.242	1.245
	0.2	1.251	1.225	1.163	1.090	1.061	1.147	1.084	1.058	1.212	1.216
	0.3	1.104	1.231	1.009	0.996	0.995	1.058	1.020	1.011	1.184	1.185
	0.4	0.885	1.202	0.958	0.982	0.993	1.001	0.998	0.999	1.161	1.161
	0.5	0.664	1.125	0.979	0.995	0.998	0.995	0.998	0.999	1.141	1.141
	0.6	0.491	1.012	0.995	0.999	0.999	0.999	1.000	1.000	1.122	1.122
	0.7	0.367	0.882	0.999	1.000	1.000	1.000	1.000	1.000	1.103	1.103
	0.8	0.284	0.759	1.000	1.000	1.000	1.000	1.000	1.000	1.086	1.086

Table 5.8 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}^+}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cox PH model with $pc = 30\%$, $p_1 = 3$, and $\pi = 0.75$ at $\Delta^{\mathrm{sim}} \geq 0$

p_2	Δ^{sim}	RE	LS	PT			SP			S	S^+
				$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$		
3	0.0	1.042	1.036	1.039	1.032	1.026	1.034	1.027	1.022	1.014	1.015
	0.1	1.038	1.037	1.029	1.019	1.013	1.028	1.019	1.013	1.010	1.012
	0.2	1.005	1.028	0.987	0.995	0.995	0.999	1.000	0.998	1.008	1.008
	0.3	0.900	0.974	0.977	0.990	0.994	0.987	0.994	0.996	1.007	1.007
	0.4	0.730	0.862	0.989	0.998	0.999	0.993	0.998	0.999	1.006	1.006
	0.5	0.553	0.716	0.998	1.000	1.000	0.999	1.000	1.000	1.005	1.005
	0.6	0.412	0.576	1.000	1.000	1.000	1.000	1.000	1.000	1.004	1.004
	0.7	0.311	0.459	1.000	1.000	1.000	1.000	1.000	1.000	1.003	1.003
	0.8	0.243	0.371	1.000	1.000	1.000	1.000	1.000	1.000	1.001	1.001

Table 5.8 RMSEs of $\widehat{\boldsymbol{\beta}}_1^{\mathrm{RE}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{LS}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{PT}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{SP}}$, $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S}}$, and $\widehat{\boldsymbol{\beta}}_1^{\mathrm{S+}}$ with respect to $\widehat{\boldsymbol{\beta}}_1^{\mathrm{UE}}$ for Cox PH model with $pc = 30\%$, $p_1 = 3$, and $\pi = 0.75$ at $\Delta^{\mathrm{sim}} \geq 0$ (Cont.)

p_2	Δ^{sim}	RE	LS	PT $\alpha = 0.01$	PT $\alpha = 0.05$	PT $\alpha = 0.10$	SP $\alpha = 0.01$	SP $\alpha = 0.05$	SP $\alpha = 0.10$	S	S+
5	0.0	1.077	1.066	1.071	1.060	1.051	1.061	1.051	1.044	1.045	1.046
	0.1	1.066	1.063	1.055	1.040	1.029	1.053	1.038	1.028	1.035	1.037
	0.2	1.028	1.053	1.004	1.000	0.996	1.017	1.008	1.001	1.029	1.029
	0.3	0.915	0.998	0.967	0.986	0.993	0.984	0.993	0.996	1.024	1.024
	0.4	0.738	0.882	0.983	0.997	0.999	0.990	0.998	1.000	1.021	1.021
	0.5	0.555	0.729	0.998	1.000	1.000	0.998	1.000	1.000	1.017	1.017
	0.6	0.412	0.585	1.000	1.000	1.000	1.000	1.000	1.000	1.014	1.014
	0.7	0.312	0.466	1.000	1.000	1.000	1.000	1.000	1.000	1.011	1.011
	0.8	0.242	0.375	1.000	1.000	1.000	1.000	1.000	1.000	1.007	1.007
7	0.0	1.112	1.097	1.105	1.092	1.077	1.091	1.080	1.067	1.075	1.077
	0.1	1.101	1.095	1.086	1.069	1.054	1.080	1.063	1.050	1.068	1.070
	0.2	1.058	1.083	1.024	1.010	1.008	1.039	1.019	1.014	1.055	1.056
	0.3	0.937	1.025	0.965	0.982	0.989	0.987	0.991	0.994	1.045	1.045
	0.4	0.752	0.905	0.974	0.995	0.997	0.984	0.997	0.998	1.037	1.037
	0.5	0.566	0.749	0.995	0.999	1.000	0.997	0.999	1.000	1.031	1.031
	0.6	0.419	0.598	1.000	1.000	1.000	1.000	1.000	1.000	1.025	1.025
	0.7	0.316	0.476	1.000	1.000	1.000	1.000	1.000	1.000	1.019	1.019
	0.8	0.246	0.384	1.000	1.000	1.000	1.000	1.000	1.000	1.013	1.013
11	0.0	1.205	1.179	1.188	1.157	1.134	1.164	1.138	1.118	1.155	1.158
	0.1	1.192	1.175	1.169	1.128	1.100	1.153	1.117	1.092	1.144	1.146
	0.2	1.145	1.167	1.085	1.046	1.033	1.098	1.054	1.039	1.125	1.126
	0.3	1.017	1.112	0.986	0.988	0.995	1.015	1.002	1.002	1.106	1.106
	0.4	0.818	0.989	0.966	0.989	0.995	0.984	0.995	0.997	1.090	1.090
	0.5	0.616	0.823	0.987	0.998	1.000	0.992	0.999	1.000	1.077	1.077
	0.6	0.456	0.659	0.999	1.000	1.000	0.999	1.000	1.000	1.065	1.065
	0.7	0.342	0.522	1.000	1.000	1.000	1.000	1.000	1.000	1.053	1.053
	0.8	0.265	0.419	1.000	1.000	1.000	1.000	1.000	1.000	1.041	1.041
15	0.0	1.322	1.278	1.293	1.252	1.206	1.253	1.218	1.180	1.254	1.257
	0.1	1.312	1.281	1.268	1.205	1.153	1.242	1.186	1.140	1.242	1.245
	0.2	1.251	1.271	1.163	1.090	1.061	1.174	1.098	1.066	1.212	1.216
	0.3	1.104	1.214	1.009	0.996	0.995	1.047	1.014	1.006	1.184	1.185
	0.4	0.885	1.084	0.958	0.982	0.993	0.985	0.992	0.997	1.161	1.161
	0.5	0.664	0.904	0.979	0.995	0.998	0.988	0.997	0.999	1.141	1.141
	0.6	0.491	0.725	0.995	0.999	0.999	0.997	0.999	1.000	1.122	1.122
	0.7	0.367	0.573	0.999	1.000	1.000	0.999	1.000	1.000	1.103	1.103
	0.8	0.284	0.458	1.000	1.000	1.000	1.000	1.000	1.000	1.086	1.086

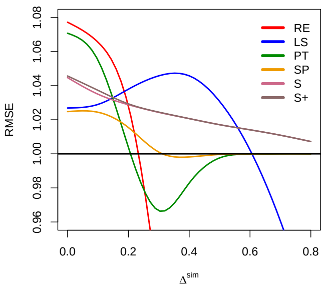


(a) $\pi = 0.25$, $\alpha = 0.01$ (b) $\pi = 0.25$, $\alpha = 0.05$ (c) $\pi = 0.25$, $\alpha = 0.10$

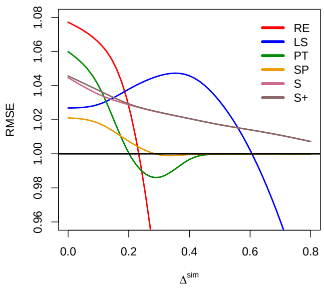
(d) $\pi = 0.50$, $\alpha = 0.01$ (e) $\pi = 0.50$, $\alpha = 0.05$ (f) $\pi = 0.50$, $\alpha = 0.10$

(g) $\pi = 0.75$, $\alpha = 0.01$ (h) $\pi = 0.75$, $\alpha = 0.05$ (i) $\pi = 0.75$, $\alpha = 0.10$

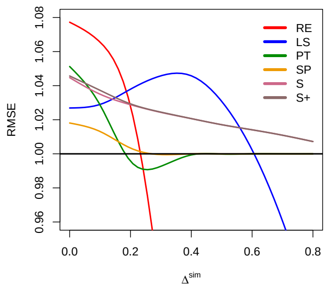
Figure 5.10 RMSEs of $\widehat{\beta}_1^{\mathrm{RE}}$, $\widehat{\beta}_1^{\mathrm{LS}}$, $\widehat{\beta}_1^{\mathrm{PT}}$, $\widehat{\beta}_1^{\mathrm{SP}}$, $\widehat{\beta}_1^{\mathrm{S}}$, and $\widehat{\beta}_1^{\mathrm{S+}}$ with respect to $\widehat{\beta}_1^{\mathrm{UE}}$ for Cox PH model with $pc = 30\%$, $p_1 = 3$, and $p_2 - 1 = 2$ at $\Delta^{\mathrm{sim}} \geq 0$



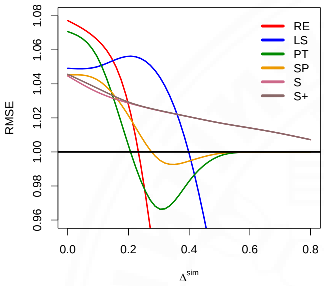
(a) $\pi = 0.25$, $\alpha = 0.01$



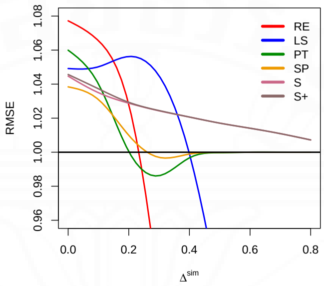
(b) $\pi = 0.25$, $\alpha = 0.05$



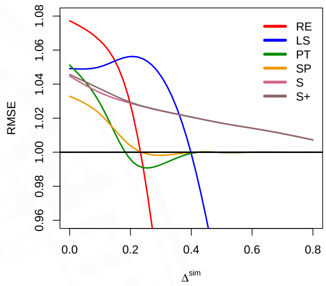
(c) $\pi = 0.25$, $\alpha = 0.10$



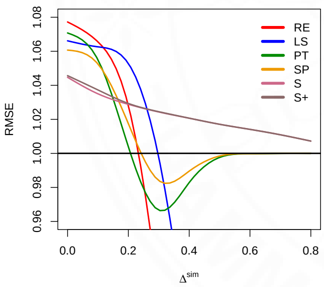
(d) $\pi = 0.50$, $\alpha = 0.01$



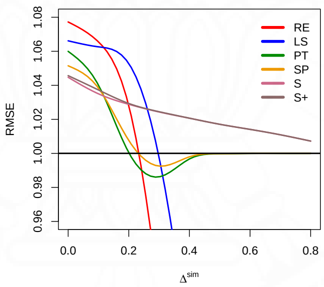
(e) $\pi = 0.50$, $\alpha = 0.05$



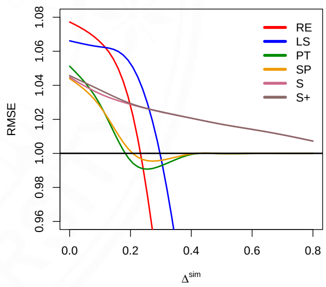
(f) $\pi = 0.50$, $\alpha = 0.10$



(g) $\pi = 0.75$, $\alpha = 0.01$

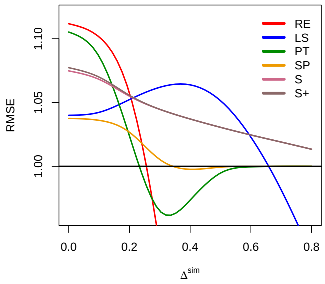


(h) $\pi = 0.75$, $\alpha = 0.05$

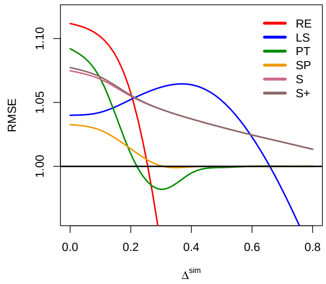


(i) $\pi = 0.75$, $\alpha = 0.10$

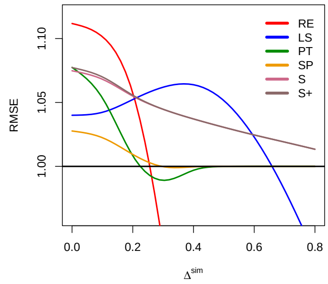
Figure 5.11 RMSEs of $\widehat{\beta}_1^{\text{RE}}$, $\widehat{\beta}_1^{\text{LS}}$, $\widehat{\beta}_1^{\text{PT}}$, $\widehat{\beta}_1^{\text{SP}}$, $\widehat{\beta}_1^{\text{S}}$, and $\widehat{\beta}_1^{\text{S}^+}$ with respect to $\widehat{\beta}_1^{\text{UE}}$ for Cox PH model with $pc = 30\%$, $p_1 = 3$, and $p_2 - 1 = 4$ at $\Delta^{\text{sim}} \geq 0$



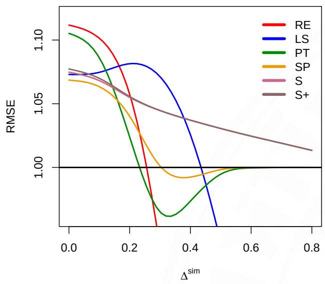
(a) $\pi = 0.25$, $\alpha = 0.01$



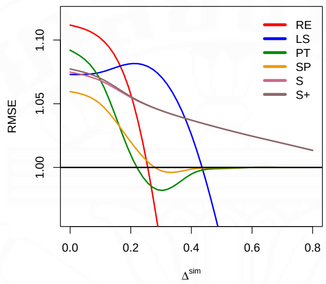
(b) $\pi = 0.25$, $\alpha = 0.05$



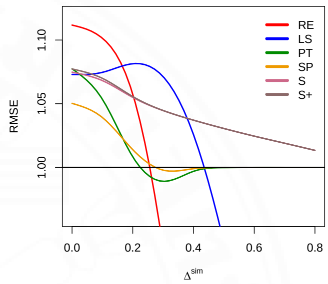
(c) $\pi = 0.25$, $\alpha = 0.10$



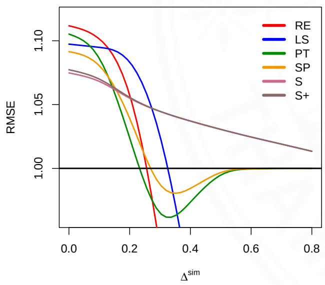
(d) $\pi = 0.50$, $\alpha = 0.01$



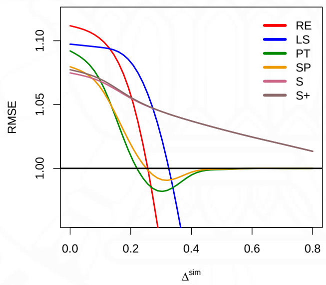
(e) $\pi = 0.50$, $\alpha = 0.05$



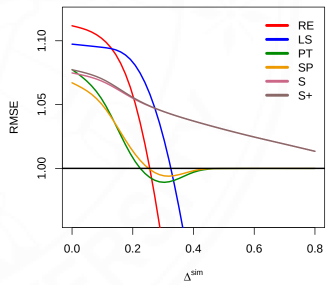
(f) $\pi = 0.50$, $\alpha = 0.10$



(g) $\pi = 0.75$, $\alpha = 0.01$

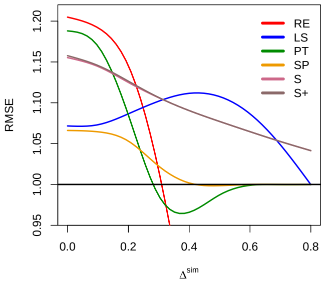


(h) $\pi = 0.75$, $\alpha = 0.05$

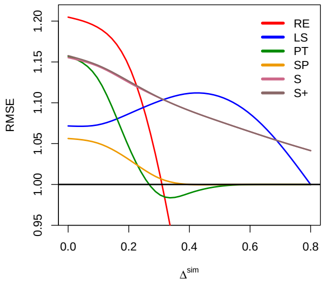


(i) $\pi = 0.75$, $\alpha = 0.10$

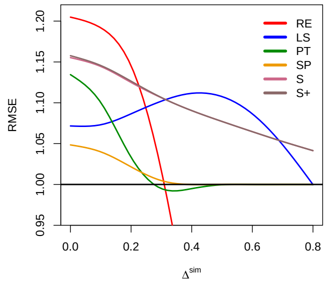
Figure 5.12 RMSEs of $\widehat{\beta}_1^{\mathrm{RE}}$, $\widehat{\beta}_1^{\mathrm{LS}}$, $\widehat{\beta}_1^{\mathrm{PT}}$, $\widehat{\beta}_1^{\mathrm{SP}}$, $\widehat{\beta}_1^{\mathrm{S}}$, and $\widehat{\beta}_1^{\mathrm{S+}}$ with respect to $\widehat{\beta}_1^{\mathrm{UE}}$ for Cox PH model with $pc = 30\%$, $p_1 = 3$, and $p_2 - 1 = 6$ at $\Delta^{\mathrm{sim}} \geq 0$



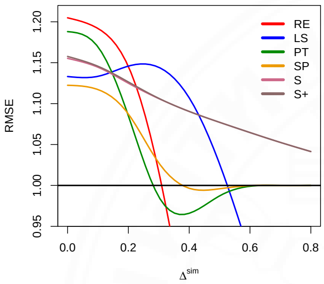
(a) $\pi = 0.25$, $\alpha = 0.01$



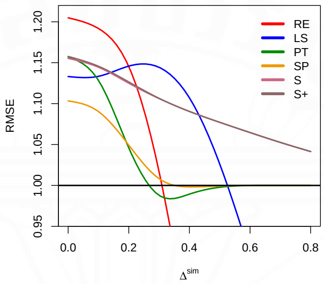
(b) $\pi = 0.25$, $\alpha = 0.05$



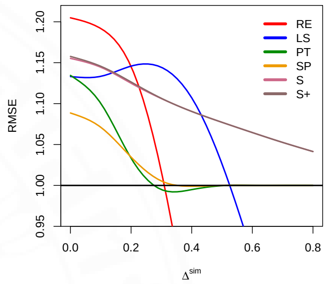
(c) $\pi = 0.25$, $\alpha = 0.10$



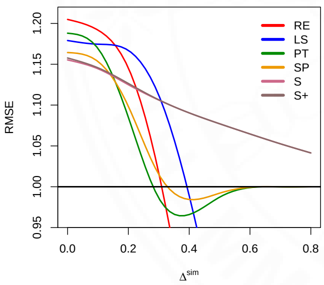
(d) $\pi = 0.50$, $\alpha = 0.01$



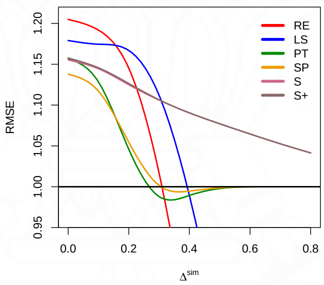
(e) $\pi = 0.50$, $\alpha = 0.05$



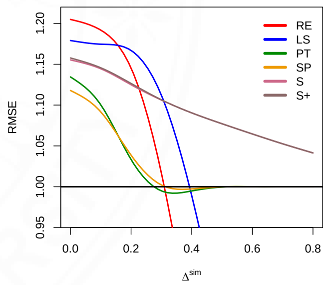
(f) $\pi = 0.50$, $\alpha = 0.10$



(g) $\pi = 0.75$, $\alpha = 0.01$

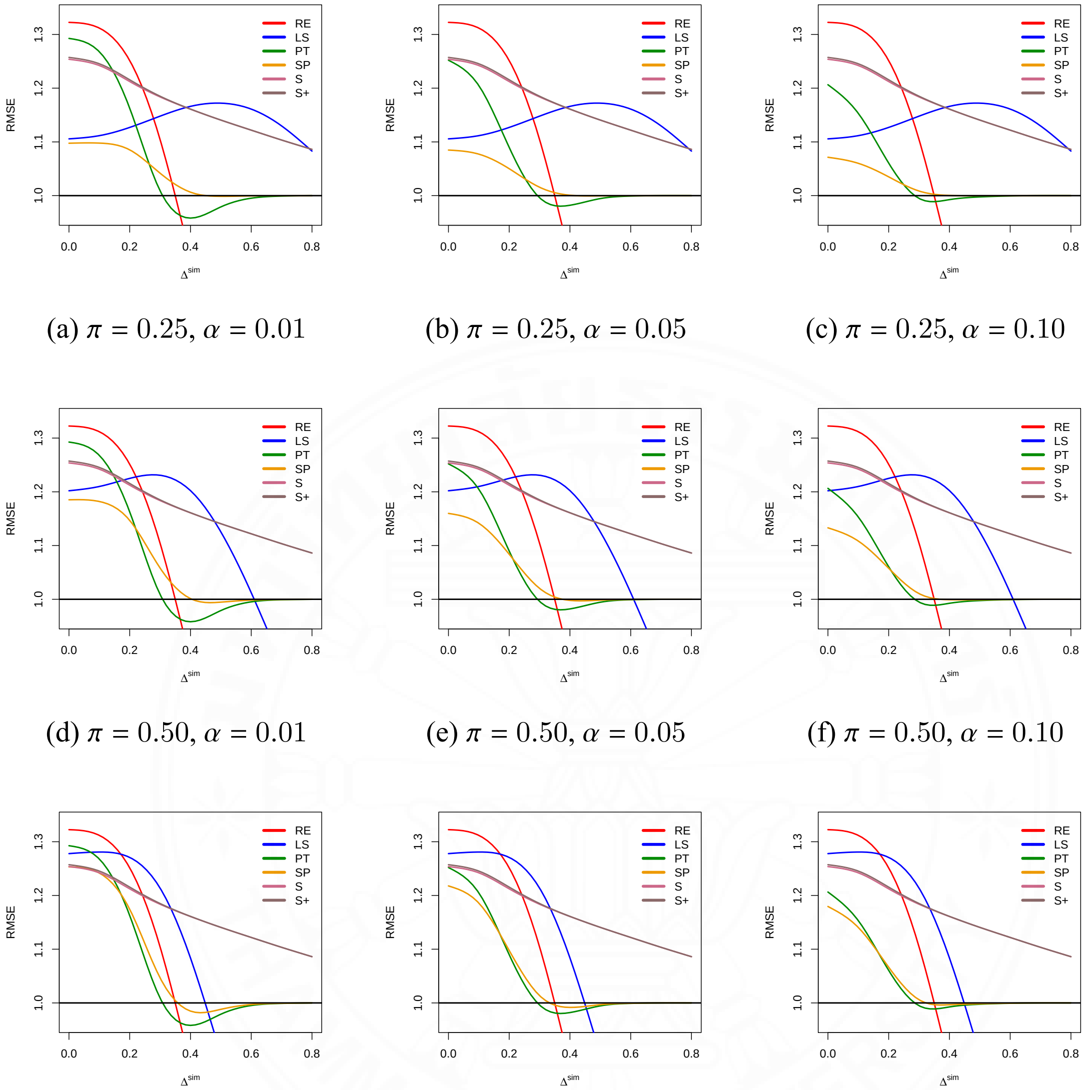


(h) $\pi = 0.75$, $\alpha = 0.05$



(i) $\pi = 0.75$, $\alpha = 0.10$

Figure 5.13 RMSEs of $\widehat{\beta}_1^{\text{RE}}$, $\widehat{\beta}_1^{\text{LS}}$, $\widehat{\beta}_1^{\text{PT}}$, $\widehat{\beta}_1^{\text{SP}}$, $\widehat{\beta}_1^{\text{S}}$, and $\widehat{\beta}_1^{\text{S}^+}$ with respect to $\widehat{\beta}_1^{\text{UE}}$ for Cox PH model with $pc = 30\%$, $p_1 = 3$, and $p_2 - 1 = 10$ at $\Delta^{\text{sim}} \geq 0$



(a) $\pi = 0.25$, $\alpha = 0.01$ (b) $\pi = 0.25$, $\alpha = 0.05$ (c) $\pi = 0.25$, $\alpha = 0.10$

(d) $\pi = 0.50$, $\alpha = 0.01$ (e) $\pi = 0.50$, $\alpha = 0.05$ (f) $\pi = 0.50$, $\alpha = 0.10$

(g) $\pi = 0.75$, $\alpha = 0.01$ (h) $\pi = 0.75$, $\alpha = 0.05$ (i) $\pi = 0.75$, $\alpha = 0.10$

Figure 5.14 RMSEs of $\widehat{\beta}_1^{\mathrm{RE}}$, $\widehat{\beta}_1^{\mathrm{LS}}$, $\widehat{\beta}_1^{\mathrm{PT}}$, $\widehat{\beta}_1^{\mathrm{SP}}$, $\widehat{\beta}_1^{\mathrm{S}}$, and $\widehat{\beta}_1^{\mathrm{S+}}$ with respect to $\widehat{\beta}_1^{\mathrm{UE}}$ for Cox PH model with $pc = 30\%$, $p_1 = 3$, and $p_2 - 1 = 14$ at $\Delta^{\mathrm{sim}} \geq 0$

For each censoring percentage, we obtained similar results for the suggested estimators. Therefore, as can be seen in Tables 5.3 to 5.8 and Figures 5.5 to 5.14, the simulation analysis findings can be summarized as follows:

1. The RE was superior to all the other estimators when the subspace information was correct or nearly correct, which means Δ^{sim} was at zero or near zero. However, when Δ^{sim} moved away from zero, the RMSE of the RE decreased and converged on zero.

2. The LS estimator is a linear combination of the UE and RE, which depends on π. Its performance was similar to that of the RE in that the RMSEs decreased slowly, converging on zero as Δ^{sim} increased. However, the LS estimator outperformed all the other estimators in some portion of $\Delta^{\text{sim}} \geq 0$.

3. When Δ^{sim} increased slightly from zero, the RMSEs of the PT and SP estimators first fell. In this phase, the PT estimator outperformed the SP estimator. However, as Δ^{sim} increased further, the SP estimator dominated the PT estimator. Finally, if Δ^{sim} increased far from zero, the RMSEs of both estimators converged to one.

4. When Δ^{sim} was at or near zero, the performance of the PT estimator was superior to that of the S and S^+ estimators at $\alpha = 0.01$. Still, its performance was inferior to that of the S and S^+ estimators at $\alpha = 0.05, 0.10$, and large p_2. However, the performance of the PT estimator was lower than the S and S^+ estimators in some areas of $\Delta^{\text{sim}} \geq 0$.

5. The SP estimator depended on values of α and π, and its performance was similar to the PT estimator. Under the null hypothesis, the SP estimator dominated the S and S^+ estimators when α was small and π was large. In contrast, for fixed p_2, the SP estimator was inferior to the S and S^+ estimators when π was small.

6. For all p_2, the S^+ estimator dominated the shrinkage estimator when Δ^{sim} was at or near zero, but their performance was equivalent when Δ^{sim} was far from zero. Furthermore, the performance of both these estimators was also superior to all other estimators in some parts of $\Delta^{\text{sim}} > 0$.

5.5.2 High-Dimensional Data

We extended the estimation problem to the high-dimensional Cox PH regression model in which $p > n$. The response variable was generated using the formula in Equation (5.37) with the following regression coefficient vector under the three effect size such as strong, weak-to-moderate, and no effect,

$$\boldsymbol{\beta} = (\underbrace{2.45, 1.51, -0.34, -1.38, 0.75}_{p_s}, \underbrace{\kappa, \kappa, ..., \kappa}_{p_w}, \underbrace{0, 0, ..., 0}_{p_n})^\top, \tag{5.38}$$

where κ denotes weak to moderate signals. We generated the dataset with values of κ as $0.01, 0.05, 0.10$, and 0.15 to examine the performance of the estimators in the presence of the weak signals. We randomly assigned all nonzero coefficients with weak to moderate signals to have either positive or negative signs.

To satisfy the usual assumptions $p_s \leq p_w < n$ and $p_n \geq n$, we considered the case $(n, p_s, p_w, p_n) = (250, 5, 25, 270)$ and $(250, 5, 45, 300)$. Here, the significance level (α) was set as 0.05, the value of shrinkage intensity (π) was set as 0.50, and the censoring percentages (pc) were set as 20% and 30%. To assess the performance of the post-selection estimators, we used the RMSE criterion with each design repeated 500 times.

Moreover, we also applied two steps, i.e., dimensional reduction and post-selection parameter estimation steps, to improve estimation for the high-dimensional sparse Cox PH regression model. This is demonstrated next.

5.5.2.1 Dimensional Reduction Step

To effectively eliminate irrelevant or select influential predictors in high-dimensional data, the two most widely used penalized estimations, namely LASSO and aLASSO, are used to produce two models with different subsets of relevant predictors.

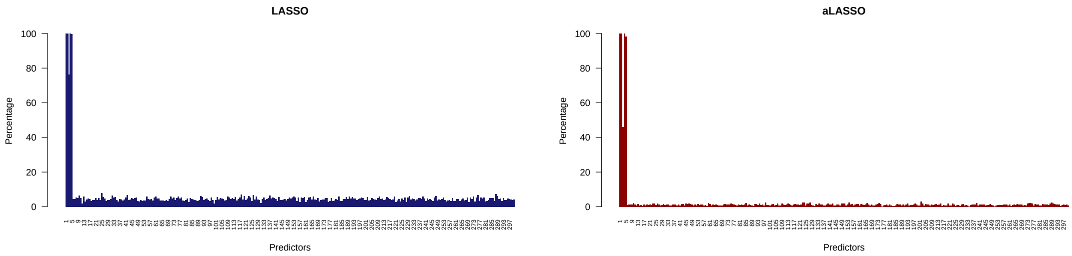
For checking the behavior and comparing the subset selection of the LASSO and aLASSO, Figures 5.15 to 5.18 display the percentage of each predictor variable selected by LASSO and aLASSO for $pc = 20\%$ and $pc = 30\%$ with $(n, p_s, p_w, p_n) = (250, 5, 25, 270)$ and $(250, 5, 45, 300)$. Table 5.9 reports the selection percentage of the predictors for each signal, censoring percentage, and (n, p_s, p_w, p_n).

Table 5.9 Selection percentage of predictors using LASSO and aLASSO methods in Cox PH regression model with strong, weak, and no signals

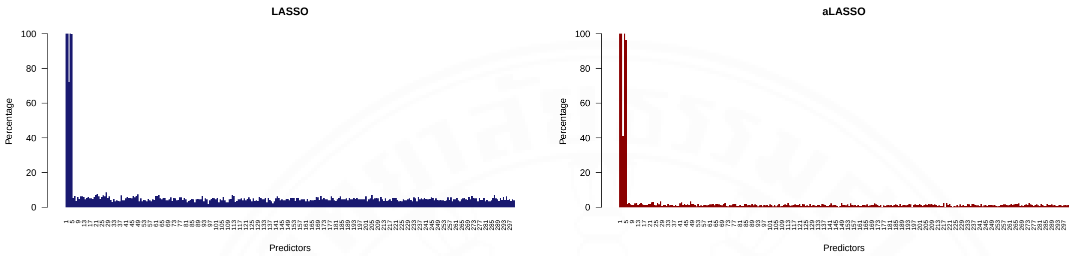
(n, p_s, p_w, p_n)	pc	κ	Strong signal		Weak signal		No signal	
			LASSO	aLASSO	LASSO	aLASSO	LASSO	aLASSO
(250,5,25,270)	20%	0.01	95.16	88.84	4.54	1.23	4.48	1.21
		0.05	94.36	87.44	5.80	1.82	4.57	1.25
		0.10	93.12	86.68	11.07	3.75	4.62	1.29
		0.15	91.84	85.80	17.65	7.83	5.18	1.75
	30%	0.01	92.52	86.00	4.55	1.16	4.47	1.25
		0.05	91.92	85.24	6.09	1.85	4.60	1.29
		0.10	91.76	84.72	9.79	3.23	4.63	1.30
		0.15	91.20	83.92	15.63	5.80	5.46	1.69
(250,5,45,300)	20%	0.01	95.32	88.20	4.02	1.00	4.10	1.09
		0.05	94.28	87.20	5.78	1.50	4.27	1.11
		0.10	93.28	86.16	10.93	3.63	5.22	1.40
		0.15	91.04	84.68	17.24	7.24	5.80	2.08
	30%	0.01	93.52	85.88	4.55	1.26	4.80	1.32
		0.05	93.12	84.56	6.36	1.85	5.18	1.43
		0.10	91.76	83.84	10.67	3.39	5.60	1.65
		0.15	89.92	81.84	14.87	6.02	5.78	2.05

Table 5.9 shows that the LASSO method was more capable than the aLASSO method in choosing predictors that had strong and weak signals. However, it also stored many nuisance (no signal) predictors. Therefore, two different submodels were obtained from LASSO and aLASSO strategies, and LASSO picked more predictors than aLASSO. Moreover, we can see that the performance in choosing the predictors with strong signals and in eliminating the predictors with no influence of LASSO and aLASSO decreased when κ increased. In contrast, the performance in selecting the predictors with weak signals increased.

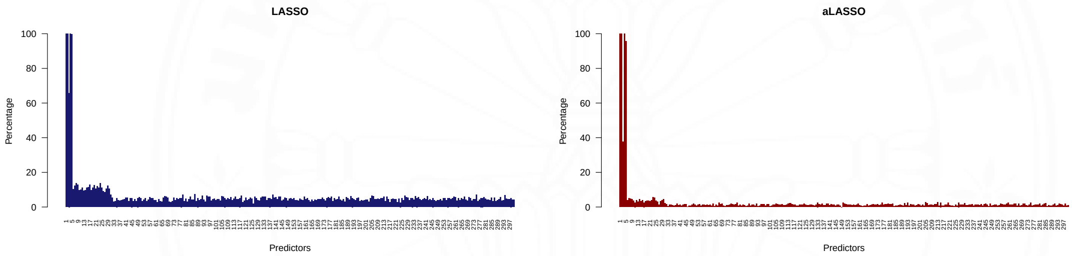
For small κ, predictors with weak signals may have little or no effect on predicting the response variable, and they should be removed from the model. On the other hand, the predictors with weak signals were essential and were selected in the model when κ was large. Furthermore, the LASSO performed better than aLASSO in choosing predictors with weak signals. However, the LASSO could not remove all the predictors with no signal, unlike aLASSO, which could eliminate most nuisance predictors. In addition, when the censoring percentage (pc) was at 30%, the performance in selecting predictors with strong signals worsened.



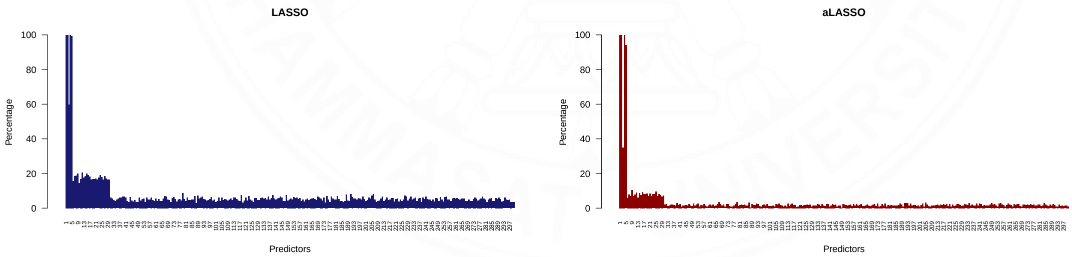
(a) $\kappa = 0.01$



(b) $\kappa = 0.05$

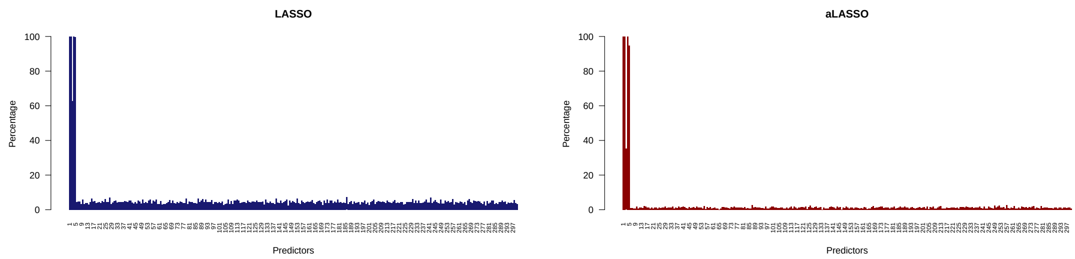


(c) $\kappa = 0.10$

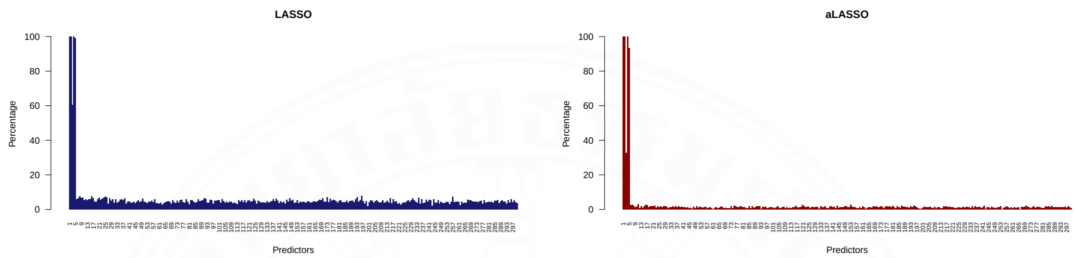


(d) $\kappa = 0.15$

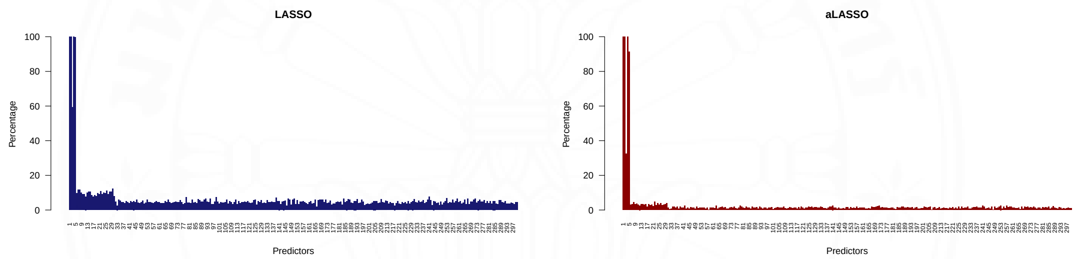
Figure 5.15 Selection percentage of predictors using LASSO and aLASSO methods in Cox PH model with $pc = 20\%$ for strong, weak, and no signals and $(n, p_s, p_w, p_n) = (250, 5, 25, 270)$



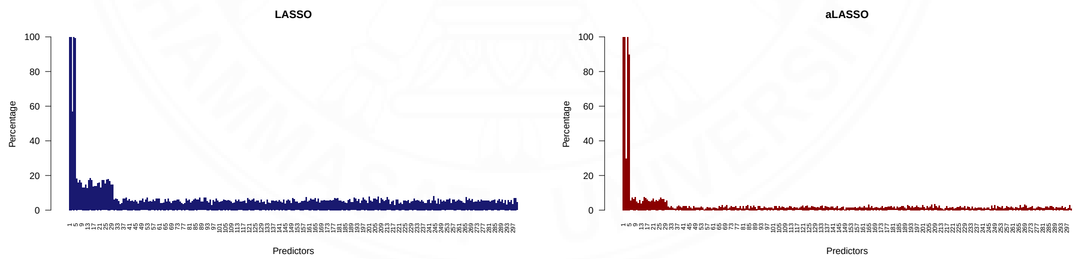
(a) $\kappa = 0.01$



(b) $\kappa = 0.05$

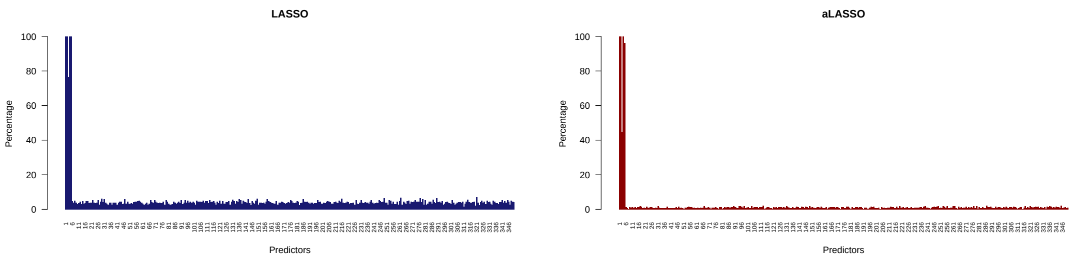


(c) $\kappa = 0.10$

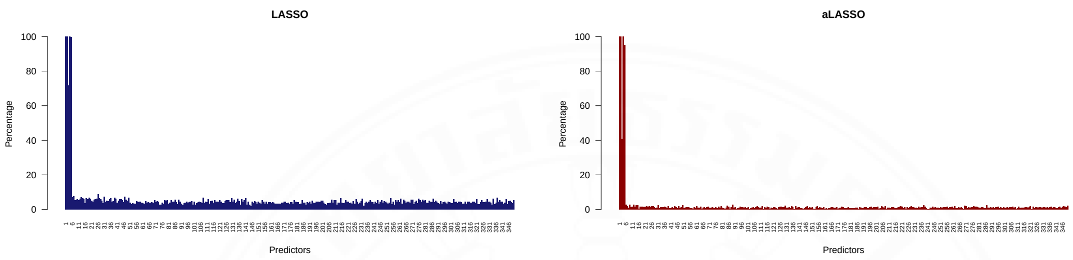


(d) $\kappa = 0.15$

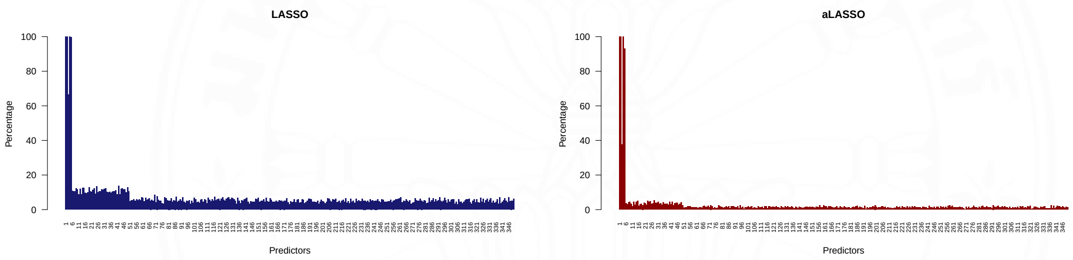
Figure 5.16 Selection percentage of predictors using LASSO and aLASSO methods in Cox PH model with $pc = 30\%$ for strong, weak, and no signals and $(n, p_s, p_w, p_n) = (250, 5, 25, 270)$



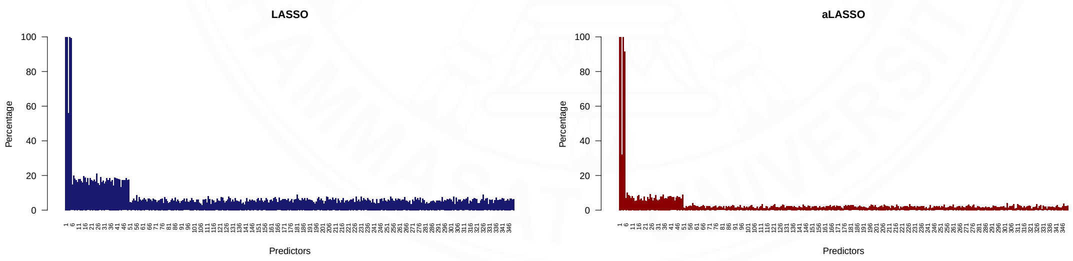
(a) $\kappa = 0.01$



(b) $\kappa = 0.05$

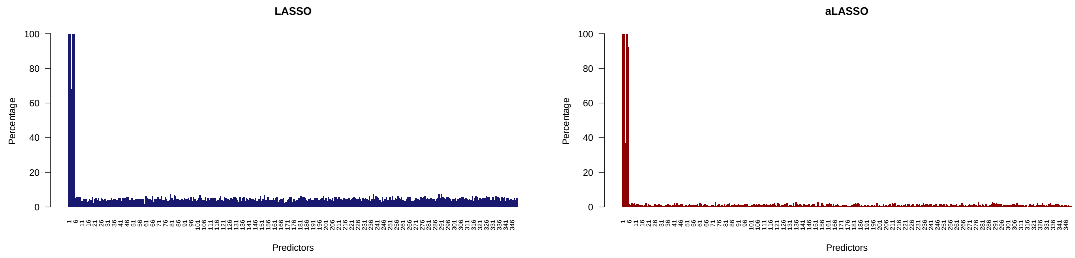


(c) $\kappa = 0.10$

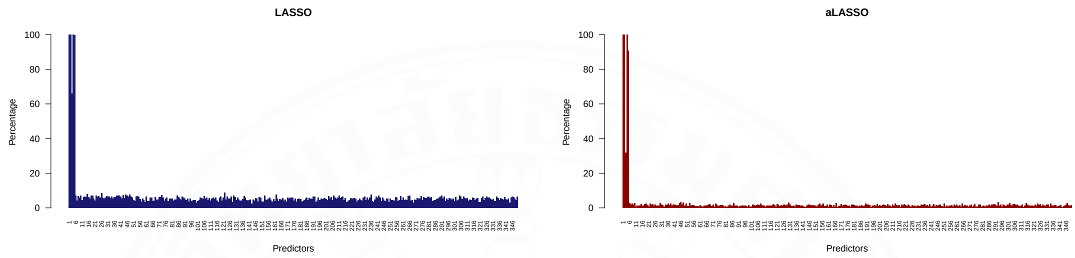


(d) $\kappa = 0.15$

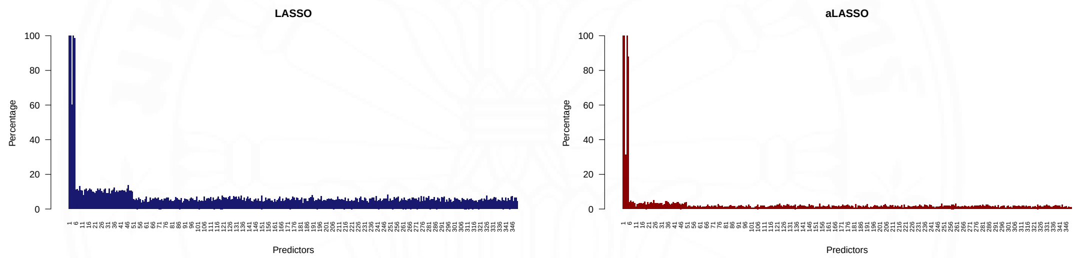
Figure 5.17 Selection percentage of predictors using LASSO and aLASSO methods in Cox PH model with $pc = 20\%$ for strong, weak, and no signals and $(n, p_s, p_w, p_n) = (250, 5, 45, 300)$



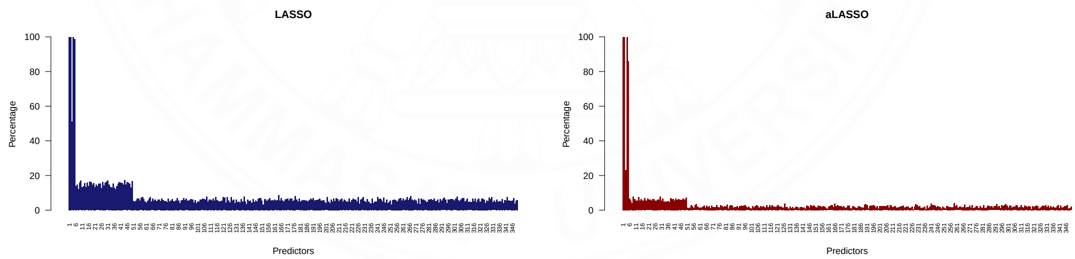
(a) $\kappa = 0.01$



(b) $\kappa = 0.05$



(c) $\kappa = 0.10$



(d) $\kappa = 0.15$

Figure 5.18 Selection percentage of predictors using LASSO and aLASSO methods in Cox PH model with $pc = 30\%$ for strong, weak, and no signals and $(n, p_s, p_w, p_n) = (250, 5, 45, 300)$

From the above results, when κ was small, the LASSO selected too many predictors with weak and no signals. Therefore, it may build an overfitted (OF) model. In contrast, when κ was large, the aLASSO may produce an underfitted (UF) model, which gives a poor performance in picking the predictors with strong and weak signals and selects fewer significant predictors than LASSO. These two models led to the consideration of the overfitted and underfitted problems. Therefore, we applied post-selection suggested estimation strategies such as the LS, PT, SP, S, and S^+ to manage this issue and compared them with the penalized estimators (LASSO and aLASSO).

5.5.2.2 Post-Selection Parameter Estimation Step

From the previous step, the LASSO and aLASSO methods reduced the variable dimension to a low-dimensional (LD) model. So, we assumed that the subset of predictors using the LASSO and aLASSO strategies contains $p_1 + p_2$ and p_1 relevant predictors, where $p_1 + p_2 < p$ in this step.

After dimensional reduction of the variables, we applied the estimation strategies in Sections 5.3.1 to 5.3.5 to parameter estimation $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$ when it was plausible that $\boldsymbol{\beta}_2$ was a zero vector. This means that the regression coefficients were divided into $\boldsymbol{\beta}_1 = (\beta_1, \beta_2, ..., \beta_{p_1})^\top$ and $\boldsymbol{\beta}_2 = (\beta_{p_1+1}, \beta_{p_1+2}, ..., \beta_{p_1+p_2})^\top$ subsets, which are coefficients from the OF with $p_1 + p_2$ parameters and the UF with p_1 parameters, respectively.

To provide parameter estimation after performing variable selection, the UE and RE were the maximum partial likelihood estimators from the OF (LASSO) and UF (aLASSO) models, respectively. For testing $H_0 : \boldsymbol{\beta}_2 = \mathbf{0}_{p_2}$, we set $\boldsymbol{\beta}_2$ as the coefficient of the p_2 predictors as existing in the OF model, but not in the UF model. The RMSE criterion was used to compare the performance of the post-selection parameter estimation, and the results of the estimators for each pc and (n, p_s, p_w, p_n) are displayed in Table 5.10.

The simulation results show that when κ increased, the RE was inferior to all other estimators since the aLASSO method was poorer than the LASSO method in selecting the predictors with strong and weak signals. This means that the LASSO method may be produced an appropriate model, in contrast to the aLASSO method, which built an underfitted model.

Table 5.10 RMSEs of estimators with respect to the UE in Cox PH model for a high-dimensional setting where $\alpha = 0.05$ and $\pi = 0.5$

(n, p_s, p_w, p_n)	pc	κ	Estimators							
			RE	LS	PT	SP	S	S+	LASSO	aLASSO
(250,5,25,270)	20%	0.01	1.3007	1.2570	1.0009	1.0005	1.1322	1.1556	0.2872	0.3627
		0.05	1.0241	1.1120	1.0000	1.0000	1.0833	1.1081	0.2911	0.3709
		0.10	0.9332	1.0647	1.0000	1.0000	1.0716	1.0760	0.3280	0.4004
		0.15	0.9095	1.0404	1.0000	1.0000	1.0663	1.0733	0.4117	0.4823
	30%	0.01	1.1150	1.1562	1.0003	1.0002	1.1047	1.1165	0.2640	0.3400
		0.05	0.9882	1.0825	1.0000	1.0000	1.0738	1.0848	0.3207	0.4066
		0.10	0.8488	0.9969	1.0000	1.0000	1.0424	1.0460	0.3824	0.4620
		0.15	0.8220	0.9765	1.0000	1.0000	1.0382	1.0382	0.5058	0.5786
(250,5,45,300)	20%	0.01	1.3376	1.2917	1.0013	1.0007	1.1495	1.2426	0.3274	0.4048
		0.05	1.1265	1.1991	1.0000	1.0000	1.1175	1.1828	0.3811	0.4723
		0.10	0.9153	1.0914	1.0000	1.0000	1.0991	1.1262	0.4114	0.5064
		0.15	0.8931	1.0357	1.0000	1.0000	1.0681	1.0694	0.4756	0.5590
	30%	0.01	1.2795	1.2771	1.0001	1.0001	1.1507	1.1941	0.3370	0.4219
		0.05	1.1466	1.2161	1.0000	1.0000	1.1366	1.1788	0.3757	0.4635
		0.10	0.9064	1.0561	1.0000	1.0000	1.0801	1.0993	0.4510	0.5436
		0.15	0.8594	1.0011	1.0000	1.0000	1.0476	1.0589	0.5614	0.6372

The behavior of the RE fell below one when κ increased. The behavior of the LS estimator was similar to the RE, but it was still more outstanding than the RE. The RMSEs of the PT and SP estimators converged to one when κ increased. The S^+ estimator outperformed the shrinkage (S) estimator, and the RMSEs of both estimators were close to the same value when κ was large. Moreover, when κ increased, the performance of the S^+ was still superior to all estimators, especially two penalized estimators.

Since the coefficients with weak and no signals in the post-selection model grew as κ increased, the penalized estimators were estimated and shrank the coefficients with no signal to zero in the post-selection model. This may explain why the behavior of both post-selection penalized estimators increased as κ increased. The aLASSO estimator outperformed the LASSO estimator for all κ values. However, both penalized estimators performed worse than all other estimators.

Based on these simulation results, we can conclude that high-dimensional sparse Cox PH regression model results were consistent with the low-dimensional regime.

5.6 Application to Real Data

To examine our approach's practical use, the proposed and penalized estimators were applied to real datasets. We do not know the subspace information on which the predictors affect the response variable in real situations. Still, we can use the variable selection methods to identify the predictors that influence the response variable. We selected significant variables in low- and high-dimensional data examples using the Bayesian information criterion (BIC) and penalized approaches (LASSO and aLASSO), respectively. Moreover, we applied the resampling bootstrap approach to estimate the regression coefficients.

Since the Cox model is a type of hazard model, it cannot directly predict the survival time. To do this, the survival analysis in the Cox model converts the hazard ratio to survival times through distributions. However, since the user must manually pre-select the distribution in the Cox model, it is problematic that a specific distribution is chosen to generate survival time (Baek et al., 2021). As this result, we evaluated the model's performance for the Cox PH model by the predictive risk score (Chen et al., 2012; Moncada-Torres et al., 2021), which indicates the chance of premature death, instead of using the prediction error of the survival time. Therefore, the suggested estimators' performance can be assessed using the MSE of risk scores (MSER) for each bootstrap replication. To facilitate the comparison, we also calculated the RMSE of risk scores (RMSER) of the estimators with respect to the UE, defined as follows:

$$\text{RMSER}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}}, \widehat{\boldsymbol{\beta}}_1^{*}) = \frac{\text{MSER}(\widehat{\boldsymbol{\beta}}_1^{\text{UE}})}{\text{MSER}(\widehat{\boldsymbol{\beta}}_1^{*})}, \tag{5.39}$$

where

$$\text{MSER}(\widehat{\boldsymbol{\beta}}_1^{*}) = \frac{1}{m}\left(e^{\boldsymbol{x}\widehat{\boldsymbol{\beta}}^{*}_{1\,(\text{true})}} - e^{\boldsymbol{x}\widehat{\boldsymbol{\beta}}^{*}_{1\,(\text{simulated})}}\right)^{\top}\left(e^{\boldsymbol{x}\widehat{\boldsymbol{\beta}}^{*}_{1\,(\text{true})}} - e^{\boldsymbol{x}\widehat{\boldsymbol{\beta}}^{*}_{1\,(\text{simulated})}}\right).$$

Here $\widehat{\boldsymbol{\beta}}_1^{*}$ is the suggested estimators and $e^{\boldsymbol{x}\widehat{\boldsymbol{\beta}}^{*}_{1\,(\text{true})}}$ and $e^{\boldsymbol{x}\widehat{\boldsymbol{\beta}}^{*}_{1\,(\text{simulated})}}$ are the predictive risk scores from the suggested estimators of the true model and of the simulated model, respectively.

In the Cox PH model, we applied the estimators used in this study, i.e., RE, LS, PT, SP, S, S^{+}, LASSO, and aLASSO estimators, to analyze real datasets. Furthermore, we applied them to a breast cancer dataset in the low-dimensional setting and a diffuse large-B-cell lymphoma dataset in the high-dimensional regime. The data analysis and results are described in the following sections.

5.6.1 Breast Cancer Data

The breast cancer data set utilized in Royston and Altman (2013) includes patient records from 720 patients with node-positive breast cancer who participated in a study undertaken by the German Breast Cancer Study Group (GBSG) between 1984 and 1989.

This dataset contains 686 patients with complete data for the prognostic variables. The median of follow-up was 1084 days overall (1,443 days for lives and 646 days for deaths). Approximately 44% (387) of patients died during this time, and the censoring rate was around 56%. The dataset was available in the `survival` package in the `R` program. The list of covariates related with this dataset is given in Table 5.11.

Table 5.11 List of variables for breast cancer data

Variable	Description
Time	Days to first of recurrence, death, or last follow-up
Status	0 = alive without recurrence (censored), 1 = recurrence or death
X1	Age (years)
X2	Menopausal status (0 = premenopausal, 1 = postmenopausal)
X3	Tumor size (mm)
X4	Tumor grade
X5	Number of positive lymph nodes
X6	Progesterone receptors (fmol/l)
X7	Estrogen receptors (fmol/l)
X8	Hormonal therapy (0 = no, 1 = yes)

Since subspace information was unavailable, we applied variable selection technique via the BIC method to identify the influence predictors. The variable selection result for establishing the submodel are given in Table 5.12.

Table 5.12 Variable selection results for breast cancer data

Method	Number of active parameters (p_1)	Number of inactive parameters (p_2)	Active predictors
BIC	4	4	X4, X5, X6, X8

We used resampling bootstrap simulations of size $m = 250$ from the dataset with replacement and $N = 1,000$ times. The RMSER results with $\alpha = 0.05$ and $\pi = 0.50$ are presented in Table 5.13.

Table 5.13 RMSERs of estimators with respect to UE for breast cancer data

Estimator							
RE	LS	PT	SP	S	S^+	LASSO	aLASSO
8.2542	3.5072	8.2542	3.5072	2.1615	2.1615	1.0763	1.7895

According to the results in Table 5.13, all estimators were more efficient than the UE. The RE estimator performed better than all other estimators. The performance of the RE and PT estimators was equivalent since their RMSERs were equal. Likewise, the RMSERs of the LS and SP also had equal values. This indicates that the test statistic lay in an acceptance region ($H_0 : \boldsymbol{\beta} = \mathbf{0}$) for all iterations. The S and S^+ estimators had the same performance, and the aLASSO estimator was superior to the LASSO estimator. In addition, the performance of the suggested estimators was superior to both the LASSO and aLASSO estimators.

5.6.2 Diffuse Large-B-Cell Lymphoma Data

The diffuse large-B-cell lymphoma (DLBCL) dataset in Rosenwald et al. (2002) was acquired from tumor-biopsy specimens and clinical data of retrospective patients to examine gene expression with the use of DNA microarrays and to analyze genomic abnormalities.

This dataset includes 240 patients with untreated diffuse large-B-cell lymphoma who had no previous history of lymphoma, but only 235 patients with follow-up times were non-zero. From 235 patients, 133 (57%) patient deaths were found during the follow-up period, the median of follow-up was 2.8 years overall (7.3 years for survivors and 1.1 years for deaths), and the censoring percentage was approximate 43%. From the available gene expression measurements data, there were 5,674 genes for analysis. However, there were many missing gene expression values in the dataset, and only 398 genes had no missing values, so we used only this complete data in the analysis.

Applying the LASSO and aLASSO strategies for variable selection in the high-dimensional Cox PH regression model, the LASSO selected 11 influential predictors ($p_1+p_2 = 11$). In addition, the aLASSO picked 6 relevant predictors ($p_1 = 6$)

Table 5.14 List of variables for diffuse large-B-cell lymphoma data

Variable	Description
Time	Follow-up (years)
Status	Status at follow-up (0 = alive, 1 = death)
GenBank ID	
LC_29447	
M27364	Eukaryotic translation elongation factor 1 alpha 1
LC_24432	
V00568	V-myc myelocytomatosis viral oncogene homolog (avian)
D32050	Alanyl-tRNA synthetase
U72511	Repressor of estrogen receptor activity
X90858	Uridine phosphorylase
X00452	Major histocompatibility complex, class II, DQ alpha 1
M20430	Major histocompatibility complex, class II, DR beta 5
X89984	B-cell CLL/lymphoma 7A
X62055	Protein tyrosine phosphatase, non-receptor type 6

Table 5.15 Variable selection results for diffuse large-B-cell lymphoma data

Method	Number of parameters	Selected predictors as active
LASSO	11	LC_29447, M27364, LC_24432, V00568, D32050, U72511, X90858, X00452, M20430, X89984, X62055,
aLASSO	6	M27364, LC_24432, V00568, U72511, X00452, M20430,

that were a subset of LASSO. The set of selected predictors using LASSO and aLASSO methods is shown in Table 5.14, and no description was provided for genes LC_29447 and LC_24432 in Rosenwald et al. (2002). The variable selection results that provided overfitted and underfitted models are represented in Table 5.15.

Later, we applied the post-selection suggested estimators to two different models with the subspace information $\boldsymbol{\beta}_2 = (\beta_{LC29447}, \beta_{D32050}, \beta_{X90858}, \beta_{X89984}, \beta_{X62055}) = (0, 0, 0, 0, 0)$, in which its correctness was unknown. Finally, to examine the performance of the proposed estimators for $\alpha = 0.05$ and $\pi = 0.50$, the RMSERs were computed using $m = 175$ bootstrap samples with data replacement $N = 1,000$ times. The findings are reported in Table 5.16.

The results in Table 5.16 show that the RE was inferior to the UE. This result indicated that the submodel provided by aLASSO was unreliable since the performance of the RE was poor. The LS, S, and S^+ were superior to the UE, except for

Table 5.16 RMSERs of estimators with respect to UE from post-selection for diffuse large-B-cell lymphoma data

Estimators							
RE	LS	PT	SP	S	S^+	LASSO	aLASSO
0.9473	1.0638	1.0000	1.0000	1.0834	1.0834	0.6165	0.6211

the PT and SP estimators, which were equivalent to the UE. The RMSER of the S^+ estimator was equal to that of the shrinkage estimator, and both estimators also performed better than all other estimators, especially with the LASSO and aLASSO estimators. Moreover, the aLASSO estimator outperformed the LASSO estimator, and the two penalized estimators were inferior to all post-selection suggested estimators.

5.7 Concluding Remarks

In this chapter, we have regarded the partial maximum likelihood, and the suggested and penalized estimators in the Cox proportional hazards regression model under the restriction of parameters in the context of low- and high-dimensional settings. We established the asymptotic properties of the suggested estimators via asymptotic distributional quadratic bias and risk. We also evaluated the performance of the suggested and penalized estimators in terms of the Monte Carlo simulations. Furthermore, the proposed estimators' performance was studied using real data examples.

The results from the variable selection step showed that the LASSO strategy was able to screen a greater number of significant predictors than the aLASSO strategy. When there were many predictors with very weak signals, LASSO caused an overfitting problem, while aLASSO introduced an underfitting problem when very weak signals became moderate. As a result, the partial maximum likelihood estimators may produce overfitting when the model was LASSO-based and underfitting when the model was aLASSO-based. They both performed poorly when their subsets were incorrect.

For the post-selection parameter estimation step, the use of the suggested estimators was suitable when the accuracy of the variable selection results was unknown because each estimator has a particular region in which its performance is superior. Furthermore, in a large portion of the parameter space, the positive-part shrinkage estimator gave better estimates than other estimators.

CHAPTER 6
CONCLUSIONS AND FUTURE RESEARCH WORK

For parameter estimation in nonlinear and Cox PH regression models using preliminary test and shrinkage strategies, this summary is divided into two parts: conclusions and future research work.

6.1 Conclusions

In this dissertation, the regression parameter estimation problem for nonlinear regression models (Cobb-Douglas, exponential, and monomolecular models) and the Cox proportional hazards regression model (special chapter) was regarded in the presence of overfitting and uncertain prior information (UPI). Estimation based on linear shrinkage (LS), preliminary test (PT), shrinkage preliminary test (SP), and two shrinkage strategies, i.e., shrinkage (S) and positive-part shrinkage (S^{+}), were proposed which efficiently combine the unrestricted and restricted estimators. We derived the asymptotic properties of the proposed estimators and mathematically compared them using asymptotic distributional quadratic bias and risk under local alternatives. We used Monte Carlo simulations to examine the risk performance of the estimators and compare them with the unrestricted estimator (UE) when subspace information misspecification existed, which was measured in terms of relative mean square error (RMSE) in low-dimensional ($k < n$) and high-dimensional ($k > n$) settings. Two penalized estimators, i.e., the least absolute shrinkage and selection operator (LASSO) and adaptive LASSO (aLASSO), were also compared numerically in a low-dimensional regime and both estimators were used for dimensional reduction in a high-dimensional case.

For statistical models in regression parameter estimation in low-dimensional data, the full model will produce overfitting when the prior information is uncertain or some of the variables in the full model have a strong influence on the response variable and some have no influence. This can be directly addressed by using the restricted model or submodel if it is *a priori* known that some regressors do not significantly contribute to the prediction of the dependent variable or that the UPI is correct. However, if the UPI is incorrect, the submodel may encounter underfitting. This indicates that the cause of the poor performance of the unrestricted and restricted model is the uncertainty of the information.

For dimensional reduction in the high-dimensional data, the results confirmed that the LASSO and aLASSO methods may not be diagnosing the optimal subset of significant predictors in all situations. The LASSO strategy contained many predictors with very weak signals that provided overfitting, and the model with the picked predictors from LASSO was set as the overfitted model. However, the aLASSO strategy eliminated many significant predictors when very weak signals became moderate signals, resulting in underfitting. Since aLASSO selected a smaller number of predictors than LASSO, the model with the chosen predictors from aLASSO was determined as the underfitted model. Consequently, inappropriate screening variables resulted in the classical estimators (unrestricted and restricted) based on these screening results becoming unreliable and ineffective.

For parameter estimation and considering the behaviour of the suggested estimators, these estimators were applied to estimate the regression coefficients to solve the overfitting and underfitting problems caused by information uncertainty in both low- and high-dimensional settings. We found that the RE dominated all other estimators when the subspace information was true or nearly true. Then its risk rapidly increased and became greater than the risk of all other estimators when the information misspecification increased. The LS estimator was also impacted by incorrect UPI similarly to but less than the RE, as the LS estimator was controlled by the degree of trust in the null hypothesis. For the PT strategy, at the same level, the performance of the SP estimator was better than the PT estimator when the subspace information was untrue. The performance of the PT and SP estimators was superior to that of the UE and the two shrinkage estimators only when the subspace information was true or nearly true. The two shrinkage estimators had higher efficiency than the UE estimator, especially when the number of inactive parameters was large, and the S^+ estimator performed uniformly better than other estimators in the largest parameter space. Moreover, the aLASSO estimator was superior to the LASSO estimator when the number of inactive parameters increased. Both penalty estimators were comparable to the proposed estimators when the number of inactive parameters was large.

To assert the benefit of the proposed estimators, these estimators were applied to real datasets and their performance was compared. When the restricted model or submodel was dependable, the RE, LS, PT, and SP estimators performed well. When the submodel was unreliable, their performance was worse. The two shrinkage estimators consistently outperformed the FM estimator and were regarded as robust irrespective of the uncertainty of prior information. These results strongly asserted the theoretical and simulation findings.

We can conclude that when there exists overfitting and UPI is untrue, the S^+ estimator can be safely used for estimating regression parameters of models of interest of low- and high-dimensional data. Still, its use is limited to the number of the inactive parameters (p_2) which must be greater than or equal to three. The use of the SP estimator is recommended for $p_2 < 3$.

6.2 Future Research Work

The work represented in this dissertation can be extended in following areas for future research.

6.2.1 In this work, we considered two penalty estimation strategies, i.e. LASSO and adaptive LASSO. There are many other penalties that exist, including smoothly clipped absolute deviation (SCAD), and elastic net (ENET), which are interesting estimators for the future study of nonlinear models. These may be used as variable selection methods in high-dimensional sparse data analysis.

6.2.2 We suggest extending the ridge estimation strategy for constructing the preliminary test and shrinkage strategies in nonlinear models to the problem of parameter estimation when models are assumed to be sparse and multicollinear.

6.2.3 We suggest extending high-dimensional data analysis in censored data (time to event data) with parametric models (i.e., Weibull and lognormal, etc.) and applying pretest and shrinkage strategies.

REFERENCES

Books and Book Articles

Bates, D. M., & Watts, D. G. (1988). *Nonlinear regression analysis and its applications* (Vol. 2). Wiley New York.

Cox, D. R., & Oakes, D. (1984). *Analysis of survival data*. Chapman and Hall/CRC.

Draper, N. R., & Smith, H. (1998). *Applied regression analysis*. John Wiley & Sons.

Durlauf, S. N., & Blume, L. E. (2008). *The new palgrave dictionary of economics* (2nd ed., Vol. 1). Palgrave Macmillan: Basingstoke.

Gowariker, V., Krishnamurthy, V. N., Gowariker, S., Dhanorkar, M., & Paranjape, K. (2009). *The fertilizer encyclopedia*. John Wiley & Sons.

Gruber, M. H. J. (1998). *Improving efficiency by shrinkage: The James-Stein and ridge regression estimators*. Marcel Dekker: New York.

Harrell, F. E. (2015). *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis* (2nd ed.). Springer.

Hosmer, D. W., & Lemeshow, S. (1998). *Applied survival analysis: regression modeling of time-to-event data*. Wiley-Interscience.

Judge, G. G., & Bock, M. E. (1978). *The statistical implications of pre-test and steinrule estimators in econometrics* (Vol. 25). North-Holland.

Klein, J. P., & Moeschberger, M. L. (2003). *Survival analysis: techniques for censored and truncated data*. Springer.

Lawless, J. F. (2003). *Statistical models and methods for lifetime data* (2nd ed.). John Wiley & Sons.

Lee, E. T., & Wang, J. (2003). *Statistical methods for survival data analysis* (3rd ed.). John Wiley & Sons.

Madden, L. V., & Campbell, C. L. (1990). Nonlinear disease progress curves. In J. Kranz (Ed.), *Epidemics of plant diseases* (pp. 181–229). Springer.

Myers, R. H., et al. (2010). *Generalized linear models.: with applications in engineering and the sciences* (2nd ed.). John Wiley & Sons.

Pal, R. (2016). *Predictive modeling of drug sensitivity*. Academic Press.

Ravishanker, N., & Dey, D. K. (2001). *A first course in linear model theory*. CRC Press.

Reangsephet, O., Lisawadi, S., & Ahmed, S. E. (2021). Weak signals in highdimensional poisson regression models (Vol. 7). In Y. Dimotikalis, A. Karagrigoriou, C. Parpoula, & C. Skiadas (Eds.), *Applied Modeling Techniques and Data Analysis 1: Computational Data Analysis Methods and Tools* (pp. 203–219). Wiley. https://doi.org/10.1002/9781119821588.ch11

Saleh, A. M. E. (2006). *Theory of preliminary test and stein-type estimation with applications* (Vol. 517). John Wiley & Sons.

Seber, G. A. F., & Wild, C. J. (2003). *Nonlinear regression*. New Jersey: John Wiley & Sons.

Articles

Abichou, T., Clark, J., Chanton, J., Hater, G., Green, R., Goldsmith, D., Barlaz, M., & Swan, N. (2012). A new approach to characterize emission contributions from area sources during optical remote sensing technique testing. *Journal of the Air & Waste Management Association*, *62*(12), 1403–1410.

Ahmed, S. E. (1992). Shrinkage preliminary test estimation in multivariate normal distributions. *Journal of Statistical Computation and Simulation*, *43*(3-4), 177–195.

Ahmed, S. E. (1997). Asymptotic shrinkage estimation: the regression case. *Applied Statistical Science II*, 113–139.

Ahmed, S. E., Hossain, S., & Doksum, K. A. (2012). Lasso and shrinkage estimation in weibull censored regression models. *Journal of Statistical Planning and Inference*, *142*(6), 1273–1284.

Ahmed, S. E., & Nicol, C. J. (2012). An application of shrinkage estimation to the nonlinear regression model. *Computational Statistics & Data Analysis*, *56*(11), 3309–3321.

Ahmed, S. E., & Raheem, S. E. (2012). Shrinkage and absolute penalty estimation in linear regression models. *Wiley Interdisciplinary Reviews: Computational Statistics*, *4*(6), 541–553.

Ahmed, S. E., & Saleh, A. K. M. (1999). Estimation of regression coefficients in an exponential regression model with censored observations. *Journal of the Japan Statistical Society*, *29*(1), 55–64.

Ahmed, S. E., & Yüzbaşı, B. (2016). Big data analytics: integrating penalty strategies. *International Journal of Management Science and Engineering Management*, *11*(2), 105–115.

Andersen, P. K., & Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *The Annals of Statistics*, 1100–1120.

Annamalai, C. (2010). Applications of exponential decay and geometric series in effective medicine dosage. *Advances in Bioscience and Biotechnology*, *1*(1), 51.

Archontoulis, S. V., & Miguez, F. E. (2015). Nonlinear regression models and applications in agricultural research. *Agronomy Journal*, *107*(2), 786–798.

Baek, E. T., Yang, H. J., Kim, S. H., Lee, G. S., Oh, I. J., Kang, S. R., & Min, J. J. (2021). Survival time prediction by integrating cox proportional hazards network and distribution function network. *BMC bioinformatics*, *22*(1), 1–15.

Bancroft, T. A. (1944). On biases in estimation due to the use of preliminary tests of significance. *The Annals of Mathematical Statistics*, *15*(2), 190–204.

Chen, H. C., Kodell, R. L., Cheng, K. F., & Chen, J. J. (2012). Assessment of performance of survival prediction models for cancer prognosis. *BMC medical research methodology*, *12*(1), 1–11.

Cheng, M. L., & Han, Y. (2013). A modified Cobb-Douglas production function model and its application. *IMA Journal of Management Mathematics*, *25*(3), 353–365.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, *34*(2), 187–202.

Davidson, R. R., & Lever, W. E. (1970). The limiting distribution of the likelihood ratio statistic under a class of local alternatives. *Sankhyaā: The Indian Journal of Statistics, Series A*, 209–224.

Dotson, W. G. (1966). Linearization transformations for least squares problems. *Mathematics Magazine*, *39*(3), 178–183.

Fekedulegn, D., Mac Siurtain, M. P., & Colbert, J. J. (1999). Parameter estimation of nonlinear growth models in forestry. *Silva Fennica*, *33*(4), 327–336.

Gao, X., Ahmed, S. E., & Feng, Y. (2017). Post selection shrinkage estimation for highdimensional data analysis. *Applied Stochastic Models in Business and Industry*, *33*(2), 97–120.

Harden, J. J., & Kropko, J. (2019). Simulating duration data for the cox model. *Political Science Research and Methods*, *7*(4), 921–928.

Hossain, M. M., Majumder, A. K., & Basak, T. (2012). An application of nonlinear Cobb-Douglas production function to selected manufacturing industries in Bangladesh. *Open Journal of Statistics*, *2*(4), 460–468.

Hossain, M. Z., & Al-Amri, K. S. (2010). Use of Cobb-Douglas production model on some selected manufacturing industries in Oman. *Education, Business and Society: Contemporary Middle Eastern Issues*, *3*(2), 78–85.

Hossain, S., & Ahmed, S. E. (2014). Penalized and shrinkage estimation in the cox proportional hazards model. *Communications in Statistics-Theory and Methods*, *43*(5), 1026–1040.

Hossain, S., Ahmed, S. E., & Howlader, H. A. (2014). Model selection and parameter estimation of a multinomial logistic regression model. *Journal of Statistical Computation and Simulation*, *84*(7), 1412–1426.

Hossain, S., & Howlader, H. A. (2017). Shrinkage estimation in lognormal regression model for censored data. *Journal of Applied Statistics*, *44*(1), 162–180.

Hossain, S., & Khan, S. A. (2020). Shrinkage estimation of the exponentiated weibull regression model for time-to-event data. *Statistica Neerlandica*, *74*(4), 592–610.

Khamis, A., Ismail, Z., Khalid, H., & Mohammed, A. T. (2005). Nonlinear growth models for modeling oil palm yield growth. *Journal of Mathematics and Statistics*, *1*(3), 225–233.

Khan, S. M. (2002). Preliminary test estimation of regression coefficients in an exponential regression model with censored observations. *Pakistan Journal of Statistics*, *18*(2), 265–274.

Lawless, J., & Singhal, K. (1978). Efficient screening of nonnormal regression models. Biometrics, 318–327.

Lisawadi, S., Ahmed, S., & Reangsephet, O. (2021). Post estimation and prediction strategies in negative binomial regression model. *International Journal of Modelling and Simulation*, *41*(6), 463-477.

Lisawadi, S., Shah, M. K. A., & Ahmed, S. E. (2016). Model selection and post estimation based on a pretest for logistic regression models. *Journal of Statistical Computation and Simulation*, *86*(17), 3495–3511.

Moncada-Torres, A., van Maaren, M. C., Hendriks, M. P., Siesling, S., & Geleijnse, G. (2021). Explainable machine learning can outperform cox regression predictions and provide insights in breast cancer survival. *Scientific reports*, *11*(1), 1–13.

Phukongtong, S., Lisawadi, S., & Ahmed, S. E. (2020). Penalty, post pretest and shrinkage strategies in a partially linear model. *Communications in StatisticsSimulation and Computation*. https://doi.org/10.1080/03610918.2020.1788589

Raheem, S. E., & Ahmed, S. E. (2011). Positive-shrinkage and pretest estimation in multiple regression: a monte carlo study with applications. *Journal of the Iranian Statistical Society*, *10*(2), 267–289.

Raheem, S. E., Ahmed, S. E., & Doksum, K. A. (2012). Absolute penalty and shrinkage estimation in partially linear models. *Computational Statistics & Data Analysis*, *56*(4), 874–891.

Rosenwald, A., Wright, G., Chan, W. C., Connors, J. M., Campo, E., Fisher, R. I., Gascoyne, R. D., Muller-Hermelink, H. K., Smeland, E. B., Giltnane, J. M., Hurt, E. M., Zhao, H., Averett, L., Yang, L., Wilson, W. H., Jaffe, E. S., Simon, R., Klausner, R. D., Powell, J., . . . Staudt, L. M. (2002). The use of molecular profiling to predict survival after chemotherapy for diffuse large-b-cell lymphoma. *New England Journal of Medicine*, *346*(25), 1937–1947.

Royston, P., & Altman, D. G. (2013). External validation of a cox prognostic model: principles and methods. *BMC medical research methodology*, *13*(1), 1–15.

Sapra, S. (2003). Pre-test estimation in poisson regression model. *Applied Economics Letters*, *10*(9), 541–543.

Thompson, J. R. (1968). Some shrinkage techniques for estimating the mean. *Journal of the American Statistical Association*, *63*(321), 113–122.

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

Wang, Y., Wang, F., & Wang, H. (2017). Influencing factors regression analysis of heating energy consumption of rural buildings in china. *Procedia Engineering*, *205*, 3585–3592.

Ware, G., Ohki, K., & Moon, L. (1982). The mitscherlich plant growth model for determining critical nutrient deficiency levels 1. *Agronomy Journal*, *74*(1), 88–91.

Yang, Z., Wang, Z., Liu, H., Eldar, Y. C., & Zhang, T. (2015). Sparse nonlinear regression: Parameter estimation and asymptotic inference. *arXiv preprint arXiv:1511.04514*.

Yasar, A., Bilgili, M., & Simsek, E. (2012). Water demand forecasting based on stepwise multiple nonlinear regression analysis. *Arabian Journal for Science and Engineering*, *37*(8), 2333–2341.

Yüzbaşı, B., Arashi, M., & Ahmed, S. E. (2017). Big data analysis using shrinkage strategies. *arXiv preprint arXiv:1704.05074*.

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

Proceedings

James, W., & Stein, C. (1961). Estimation with quadratic loss. In J. Neyman (Ed.), *Proceedings of the fourth berkeley symposium on mathematical statistics and probability: Vol. 1. Contributions to the theory of statistics* (pp. 361–379). University of California Press.

Reangsephet, O., Lisawadi, S., & Ahmed, S. E. (2018). A comparison of pretest, stein-type and penalty estimators in logistic regression model. In J. Xu, M. Gen, A. Hajiyev, & F. Cooke (Eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management* (pp. 19–34). Springer, Cham. https://doi.org/10.1007/978-3-319-59280-0_2

Reangsephet, O., Lisawadi, S., & Ahmed, S. E. (2019). Improving estimation of regression parameters in negative binomial regression model. In J. Xu, F. Cooke, M. Gen, & S. Ahmed (Eds.), *Proceedings of the Twelfth International Conference on Management Science and Engineering Management* (pp. 265–275). Springer, Cham. https://doi.org/10.1007/978-3-319-93351-1_22

Reangsephet, O., Lisawadi, S., & Ahmed, S. E. (2020). Weak signals in highdimensional logistic regression models. In J. Xu, S. Ahmed, F. Cooke, & G. Duca (Eds.), *Proceedings of the Thirteenth International Conference on Management Science and Engineering Management* (pp. 121–133). Springer, Cham. https://doi.org/10.1007/978-3-030-21248-3_9

Technical and Research Reports

Stein, C. (1956). *Inadmissibility of the usual estimator for the mean of a multivariate normal distribution* (Technical Report). Stanford University Stanford United States.

Chaowagul, M., Phahuyuth, D., & Pooncharoen, N. (2015). *Report on the socioeconomic situation monitoring and evaluation, kwae noy dam under his majesty the king's initiation, budjet year 2015* (Research Report). Naresuan University, Phitsanulok.

Dissertation

Rajarshi, V. M. (1995). *Verification of mitscherlich-bray equation as modified by bangar for efficient N and P fertilizer use for winter sorghum on vertisol under dryland condition* [Unpublished doctoral dissertation]. Mahatma Phule Krishi Vidyapeeth, Rahuri.

Webpages and Websites

Food and Agriculture Organization. (2020). *Food and agriculture data*. Retrieved June 15, 2020, from http://www.fao.org/faostat/en/#data

National Bureau of Economic Research. (2016). *NBER-CES manufacturing industry database*. Retrieved September 2, 2020, from https://www.nber.org/research/data/nber-ces-manufacturing-industrydatabase

Numbeo. (2020). *Cost of living and property prices indices*. Retrieved October 8, 2020, from https://www.numbeo.com/

UCI Machine Learning Repository. (n.d.). *Communities and crime unnormalized data set*. Retrieved November 30, 2021, from https://archive.ics.uci.edu/ml/datasets/Communities+and+Crime+Unnormalized

World Bank. (2018). *World development indicators data set*. Retrieved December 12, 2021, from https://databank.worldbank.org/source/world-development-indicators

BIOGRAPHY

Name	Janjira Piladaeng
Educational Attainment	Academic Year 2015: Master of Science (Applied Statistics), Thammasat University, Thailand Academic Year 2012: Bachelor of Science (Statistics), Khon Kaen University, Thailand
Scholarship	2018 – 2019: Canada-ASEAN Scholarships and Educational Exchanges for Development (SEED) 2009 – 2018: Human Resource Development in Science Project (Science Achievement Scholarship of Thailand, SAST)

Publications

Piladaeng, J., Lisawadi, S., & Ahmed, S. E. (2020). Improving the performance of least squares estimator in a nonlinear regression model. In J. Xu, G. Duca, S. Ahmed, F. García Márquez, & A. Hajiyev (Eds.), *Proceedings of the Fourteenth International Conference on Management Science and Engineering Management* (pp. 492–502). Springer, Cham. https://doi.org/10.1007/978-3-030-49829-0_37

Piladaeng, J., Ahmed, S. E., & Lisawadi, S. (in press). Penalised and post pretest and shrinkage strategies in nonlinear growth models. *Australian & New Zealand Journal of Statistics*.