UNIVERSITY OF SOUTHERN QUEENSLAND

IMPROVED ESTIMATION FOR LINEAR MODELS UNDER DIFFERENT LOSS FUNCTIONS

A Dissertation submitted by

Zahirul Hoque
B.Sc. Honours, M.Sc.

For the award of
Doctor of Philosophy

2004

Abstract

This thesis investigates improved estimators of the parameters of the linear regression models with normal errors, under sample and non-sample prior information about the value of the parameters. The estimators considered are the unrestricted estimator (UE), restricted estimator (RE), shrinkage restricted estimator (SRE), preliminary test estimator (PTE), shrinkage preliminary test estimator (SPTE), and shrinkage estimator (SE). The performances of the estimators are investigated with respect to bias, squared error and linex loss. For the analyses of the risk functions of the estimators, analytical, graphical and numerical procedures are adopted.

In Part I the SRE, SPTE and SE of the slope and intercept parameters of the simple linear regression model are considered. The performances of the estimators are investigated with respect to their biases and mean square errors. The efficiencies of the SRE, SPTE and SE relative to the UE are obtained. It is revealed that under certain conditions, SE outperforms the other estimators considered in this thesis.

In Part II in addition to the likelihood ratio (LR) test, the Wald (W) and Lagrange multiplier (LM) tests are used to define the SPTE and SE of the parameter vector of the multiple linear regression model with normal errors. Moreover, the modified and size-corrected W, LR and LM tests are used in the definition of SPTE. It is revealed that a great deal of conflict exists among the quadratic biases (QB) and quadratic risks (QR) of the SPTEs under the three original tests. The use of the modified tests reduces the conflict among the QRs, but not among the QBs. However, the use of the size-corrected

tests in the definition of the SPTE almost eliminates the conflict among both QBs and QRs. It is also revealed that there is a great deal of conflict among the performances of the SEs when the three original tests are used as the preliminary test statistics. With respect to quadratic bias, the W test statistic based SE outperforms that based on the LR and LM test statistics. However, with respect to the QR criterion, the LM test statistic based SE outperforms the W and LM test statistics based SEs, under certain conditions.

In Part III the performance of the PTE of the slope parameter of the simple linear regression model is investigated under the linex loss function. This is motivated by increasing criticism of the squared error loss function for its inappropriateness in many real life situations where underestimation of a parameter is more serious than its overestimation or vice-versa. It is revealed that under the linex loss function the PTE outperforms the UE if the non-sample prior information about the value of the parameter is not too far from its true value. Like the linex loss function, the risk function of the PTE is also asymmetric. However, if the magnitude of the scale parameter of the linex loss is very small, the risk of the PTE is nearly symmetric.

Certification of Dissertation

I certify that the ideas, mathematical derivation of the formulas, results, analysis and conclusions reported in this dissertation are entirely the outcome of my own effort, except where otherwise acknowledged. I also certify that the work is original and has not been submitted for any other award, except where otherwise acknowledged.

Signature of Candidate | Date

ENDORSEMENT

Signature of Supervisor | Date

Acknowledgement

First of all, I wish to express my deepest gratitude to almighty Allah, whose divine support helped my to complete this dissertation, *Alhamdulillah.*

To my supervisor, Associate Professor Dr Shahjahan Khan— You have provided me with unbelievable support in all respects throughout my candidature, especially in my introduction to the research area, about which I had absolutely no idea before I met you. I have never seen you unhappy to answer any of my questions, even to the most stupid ones. I could not have finished this without your help. *I find no words to express my gratitude to you.*

To my associate supervisor, Dr Ashley Plank— I gratefully acknowledge your intellectual criticism and valuable time, especially during the preparation of my dissertation proposal and the writing up of this dissertation. *Thanking you may not be enough, no matter how big the thank you.*

I would like to thank USQ for providing me with an USQ Postgraduate Research Scholarship and a Faculty of Sciences Postgraduate Scholarship. My special thanks to the Department of Mathematics and Computing for providing me with a part-time teaching position at the Department, travel grants to attend a number of conferences and visiting different universities. I thankfully acknowledge the excellent support of the Office of Research and Higher Degrees.

I would like to express my gratitude to the University of Chittagong, Bangladesh for granting me study leave and a special grant towards my study at USQ.

To my colleagues— Prof. Roberts, Asso. Prof. Harman, Dr Billah, Ms

Lennox, and Mr Farrell, to mention a few. *I appreciate your help.*

To my parents— It would not have been possible for me to reach this stage if you did not sacrifice so much. *Thanks God, you blessed me with such lovely parents.*

Finally, to Nipa and Irfan— Nipa, it could not be possible to complete this dissertation in due time if you were not that caring wife. Irfan, when you were only three years old you knew that in the morning your Dad goes to the uni to do his PhD. You did not know what it was!! *Your imagination encouraged me to complete it.*

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Linear regression models are used to represent the linear relationships between the response or dependent variable and a set of explanatory variables or predictors. If appropriate, such a model can be used to predict the value of a response variable for a set of known values of the predictors. For any such prediction, estimation of the regression parameters, is essential. Estimation of parameters is also essential for performing statistical tests on any individual or set of regression parameters. In practice, there are many variables, the relationship among which can be explained by a linear regression model, and hence it is one of the most popular models used in data analysis. However, fitting any model to a set of data, involves the estimation of the parameters of the model.

The commonly-used classical estimators of the unknown parameters of the liner regression models are based exclusively on the sample information. In real life situations, researchers may have prior information on the parameters available either in the form of a prior distribution or as a value of a parameter. The source of such prior information can be previous studies or expert knowledge.

The prior distribution of a parameter is used in the Bayesian approach to statistical analysis. However, if the prior information about the parameter is available as a constant value rather than as a distribution, the Bayesian approach cannot be pursued. There are however estimation methods that use prior information about the value of a parameter in addition to the sample information. The expectation is that the inclusion of such additional information in the estimation process would result in a better estimator than using sample information alone. In some cases this may be true, but in many other cases the risk of worse consequences cannot be ruled out.

This thesis deals with the improved estimation strategies of the parameters of the simple and multiple linear regression models with normal errors, where sample as well as non-sample prior information about the value of the parameter are used. The performances of the estimators are investigated under various loss functions.

There are three main parts of this thesis. Part I consisting of Chapters 2 and 3 studies four different alternative estimators for each of the slope and intercept parameters of the simple linear regression model under the squared error loss function. Part II consisting of Chapters 4 and 5 studies the impact of using three alternative tests in the definition of the same estimator of the coefficient vector of multiple linear regression model. Finally, Part III consisting of Chapter 6 studies the performance of an improved estimator of the slope parameter of the simple linear regression model under the linex loss function.

1.1 Improved Estimation Under Squared Error Loss Function

The credit for discovery of the method of least squares, generally, is given to Carl Friedrich Gauss, who used the procedure in the early part of the nineteenth century. The exclusive sample information based least-squares estimator (LSE) or equivalent maximum likelihood estimator (MLE) of the parameters of the linear models with normal error are unbiased and uniformly minimum variance. Such an estimator is known as the unrestricted estimator (UE) as it is obtained from the sample information alone, without any restriction. However, with respect to some other statistical criteria, the UE is not appropriate and indeed can be improved upon by using additional information such as non-sample prior information about the value of the parameter.

Credible non-sample prior information about the value of a parameter is known as uncertain non-sample prior information (UNSPI) as there is doubt about the accuracy of such information. According to Fisher, the UNSPI about the value of the parameter can be expressed in the form of a null hypothesis and the uncertainty can be removed by performing an appropriate statistical test on that hypothesis (cf. Khan *et al.*, 2002). Under the null hypothesis, the suspected value of the parameter is known as the restricted estimator (RE). The RE outperforms the UE when the null hypothesis holds; otherwise the UE outperforms the RE. Therefore, it is a natural expectation to combine the sample and non-sample prior information to define an estimator that may outperform both UE and RE, under certain conditions.

Khan and Saleh (2001) defined the SRE (they called it RE) of the univariate normal mean, as a convex combination of the UE and RE, with a coefficient of distrust d $(0 \leq d \leq 1)$ representing the measure of distrust in the UNSPI. The value of d is determined by the experimenter according to his/her belief on the null hypothesis. Bancroft (1944), and later Han and Bancroft (1968), developed the preliminary test estimator (PTE) that uses the sample as well as uncertain non-sample prior information about the value of the parameter, in its definition. Some authors call PTE testimator for obvious reasons (cf. Pandey and Rai, 1996). If the UNSPI about the value of the parameter is not too far from its true value with respect to the squared error loss, the PTE dominates the UE (cf. Ahsanullah and Saleh, 1972). Khan and Saleh (2001) introduced the coefficient of distrust d to the PTE of the univariate normal mean, and called the new estimator the shrinkage PTE (SPTE). Unlike the PTE, the SPTE is a continuous function of the UE and RE. For $d = 0$, the SPTE becomes the PTE. Therefore, the PTE is a special case of the SPTE.

Stein (1956) surprised the statistical world by declaring that with respect to the squared error loss function, the sample mean of a p-dimensional $(p \geq 3)$ population is an inadmissible estimator of the population mean, as one can find another estimator that dominates the sample mean. Later, James and Stein (1961) introduced the Stein-type or James-Stein shrinkage estimator (SE) for multivariate normal population that dominates the usual maximum likelihood estimator, the sample mean, under the squared error loss criterion, if the dimension of the population is three or more. The seminal work of Bancroft (1944), Stein (1956), and James and Stein (1961) generated a large volume of research on improved estimators of parameters.

Ahmed and Saleh (1989) provided a comparison among the UE, RE, PTE and SE, for two multivariate normal populations with a common covariance structure. Their study showed that under certain conditions, the SE outperforms the other three estimators. Later, Khan and Hoque (2002) extended the study by proposing the positive-rule SE (PRSE). They showed that under the squared error loss criterion, the PRSE is a better choice than the SE. Khan and Saleh (2001) defined the SE of the univariate normal mean with a slightly different approach from that for the multivariate set-up. They showed that with respect to the mean square error criterion, the SE outperforms the UE, SRE and SPTE, under certain conditions. In a series of papers, Saleh and Sen (1984, 1985, 1986) introduced the preliminary test approach to Stein's approach in the non-parametric set-up.

The simple linear regression model is one of the most widely used models in many disciplines, and hence improvement in the estimation of its parameters is desirable. Ahsanullah and Saleh (1972) defined the preliminary test estimator of the intercept parameter of the simple linear regression model with normal error, assuming the value of the slope parameter is zero. They derived the bias and mean square error functions of the PTE and compared them with those of the maximum likelihood estimator. As the value of the slope parameter is not necessarily zero, it is of interest to define the PTE/SPTE of the intercept parameter under the suspicion that the value of the slope parameter is some constant that may or may not be zero. Bhoj and Ahsanullah (1993) considered two linear regression models with normal errors and studied the preliminary test estimator of the conditional mean of the dependent variable in the first model under the suspicion that the values of the slope parameters for both models are equal. Later, Bhoj and Ahsanullah (1994) extended the problem of the preliminary test estimation of the conditional mean of the first model under the suspicion that the values of both the slope and intercept parameters of one model are the same as those of the other model. Saleh and Sen (1978, 1979) considered the non-parametric estimation strategies for the intercept parameter after a preliminary test on regression, for univariate and multivariate cases.

Khan and Saleh (1995, 1997) investigated the improved estimation problem for a family of Student's t populations. Khan and Saleh (1998) discussed different estimators of the location parameter for a location-scale model based on samples from p multivariate Student's t populations. Many authors have contributed to this area, notably Sclove *et al.* (1972), Judge and Bock (1978), Stein (1981), Matta and Casella (1990), and Khan (1998) to mention a few. However, the relative performances of the SRE, SPTE and SE of the intercept and slope parameters of the simple linear regression model have not been previously investigated.

In this thesis we investigate the alternative estimators of the slope and intercept parameters those are biased but possess superior statistical property in terms of a popular statistical criterion, namely the mean square error (mse). The estimators of slope and intercept parameters considered in this thesis are the three biased estimators: the SRE, SPTE and SE. The bias and mean square error functions of the estimators are derived. To compare the performances of the estimators, the bias and mean square error functions have been analysed both analytically and graphically. The efficiencies of the estimators relative to the unrestricted estimator are also investigated.

1.2 The W, LR and LM Tests in Improved Estimation

Until recent years, the likelihood ratio test based on t or F statistic, was used to define preliminary test based estimators. In the literature, there are alternative tests to the LR test, namely, the Wald (W) and Lagrange multiplier (LM) tests. The W test was introduced by Wald (1943), and the LM test by Aitchison and Silvey (1958) and Silvey (1959). Among others, Breusch (1979) pointed out that the LM test is the same as the score test of Rao (1947). Engle (1984) distinguished the three tests by stating that "the LM approach starts at the null and asks whether movement toward the alternative would be an improvement, the W approach starts at the alternative and considers movement toward the null, and the LR approach compares the two hypotheses directly on an equal basis." Therefore, the three tests based on different test statistics measure the difference between the null and alternative hypotheses, but in different fashions. For a geometrical interpretation of these differences readers may see Engle (1984). Among others, Savin (1976), and Berndt and Savin (1977) pointed out that a systematic inequality relation exists among the values of the three test statistics.

The exact sampling distributions of the W, LR and LM test statistics are complicated (cf. Rothenberg, 1977). In practice, the critical regions of the tests are determined based on their asymptotic approximations. Under the null hypothesis, the three test statistics are asymptotically equivalent and distributed as chi-square with the same degrees of freedom (cf. Engle, 1984). Evans and Savin (1982) showed that the tests based on the approximate chi-square critical value differ with respect to their size and power, particularly for small samples, and there may be conflict among their conclusions. The probability of conflict is substantial when the three tests are based on the same asymptotic chi-square critical value. It may not be surprising that the use of conflicting tests in the definitions of the SPTE and SE will affect the statistical properties of the estimators.

Billah (1997) and Billah and Saleh (1998) introduced the three tests in the formation of the PTE and SPTE for multiple linear regression models with normal errors. Their studies showed that the performances of the PTEs and SPTEs are different for different tests. Later, Billah and Saleh (2002a,b) extended their earlier studies to the regression model with Student's t errors revealing the same properties of the estimators as those for the model with normal errors. Recently, Kibria (2002) and Khan and Hoque (2003) introduce the three W, LR and LM tests in the formation of the shrinkage preliminary test maximum likelihood estimator (SPTMLE) and PTEs, respectively for the multivariate normal mean. Kibria (2002) considered the p-dimensional multivariate normal model with mean vector $\boldsymbol{\mu}$ and a special covariance structure $\Sigma = \sigma^2 I_p$ and defined the SPTMLE of $\boldsymbol{\mu}$ under the suspicion that the values of the p components of the population mean vector are equal. Khan and Hoque (2003) considered the same model but defined the PTE of $\boldsymbol{\mu}$ under the suspicion that $\boldsymbol{\mu} = \boldsymbol{\mu}_0$, a given vector of the same dimension. The studies of Kibria (2002) and Khan and Hoque (2003) revealed that the use of the asymptotically equivalent tests in improved estimation of the parameters results in conflicting performances of the estimators.

With a view to dealing with the conflict among the three tests, Evans and Savin (1982) studied the properties of the three tests after the introduction of some correction factors to the test statistics. They used degrees of freedom corrections to the W and LM test statistics and Edgeworth correction to the LR test statistic. The degrees of freedom correction was originally introduced by Gallant (1975) for non-linear regression model. It corrects the bias of the estimate of the error variance. The Edgeworth correction was derived by Rothenberg (1977) from the second order Edgeworth approximation to the exact distribution of the test statistic. The tests with these corrections are known as the modified W, LR and LM tests. Evans and Savin (1982) calculated the powers of the modified tests and the probability of conflict and showed that the modification results in a better approximation to the power function of the exact tests. However, the conflict remains still substantial. The inequality relation that holds for the original test statistics does not hold for the modified test statistics. Khan and Hoque (2003) also introduced the modified W, LR and LM tests in the formation of the SPTE of the multivariate normal mean. Their study showed that the use of the modified tests reduces the conflict among the properties of the SPTEs to some extent but remains considerable.

The conflict among the W, LR and LM tests with or without modification is due to the fact that they use the same chi-square critical value despite the fact that the values of the test statistics are not the same, in general. Further, Evans and Savin (1982), and Rothenberg (1977) suggested the Edgeworth correction to the chi-square critical values of the W and LM test statistics, in addition to that of the LR test statistics. They showed that the Edgeworth corrections to the critical values of the tests almost eliminate the conflict among the powers of the tests. Therefore, it is of interest to investigate the performances of the improved estimators under the three alternative tests and under their different versions.

In this thesis, we define the SPTEs of the parameter vector of the multiple linear regression model under the original, modified and size-corrected W, LR and LM tests. The quadratic risk functions of the estimators are derived. The efficiencies of the SPTEs (with respect to the quadratic risk) relative to the UE are obtained. The conflict among the relative efficiencies is calculated. Here conflict is defined as the difference between the maximum and minimum relative efficiencies of the SPTEs under different tests and for any particular value of the non-centrality parameter of the non-central F distribution. It is seen that the use of the original tests in the formation of the SPTE results in a great deal of conflict among the statistical properties of the SPTEs. Though the use of the modified tests reduces this conflict to some extent, it remains considerable. However, the use of the size-corrected tests almost eliminates the conflict among the statistical properties of the SPTEs.

By definition, unlike the SPTE, the SE does not depend on the level of significance of the preliminary test of the null hypothesis. On the other hand, the modified LR and the size-corrected W, LR and LM tests are obtained by using the size correction of the tests. Therefore, the SE is defined under the original W, LR and LM test statistics only. The performances of the SEs with respect to the quadratic bias and quadratic risk are investigated. Both graphical and numerical analyses are pursued. It is revealed that under the quadratic bias criterion the performance of the W test statistic based SE is the best having smallest QB, followed by the LR and LM test statistics based SEs, respectively. Under the quadratic risk criterion there is no uniform domination of one SE over the others. However, under certain conditions, the LM test statistic based SE dominates the other two SEs.

1.3 Improved Estimation Under Linex Loss Function

The popularity of the squared error loss function is due to its mathematical and interpretational convenience. In spite of the wide popularity of this symmetric loss function, many authors have recognised it as inappropriate in various problems (see, for instance, Ferguson, 1967; Zellner and Geisel, 1968; Aitchison and Dunsmore, 1975; Varian, 1975; Berger, 1980). As pointed out by Zellner (1986), the admissibility of an estimator may depend quite sensitively on features of the loss function, such as symmetry, is not generally appreciated. Due to the symmetric nature of the squared error loss function it cannot differentiate between the overestimation and underestimation of any parameter. In real life situations there are numerous cases where underestimation of a parameter leads to more or less severe consequences than overestimation. In dam construction, for example, underestimation of the peak water level is more serious than overestimation. On the other hand, for a manufacturing company, overestimation of the mean life of the product for the purposes of customers warranty is more serious than underestimation. As the squared error loss function is unable to assign appropriate unequal weights for underestimation and overestimation of any parameter, the use of this loss function is inappropriate and hence not useful.

In an applied study of real estate assessment, Varian (1975) introduced a very useful non-symmetric loss function called the linex loss function, that has both linear and exponential components and is appropriate to represent asymmetric losses. This loss function grows approximately exponentially on one side of zero, the value of the estimation error, and approximately linearly on the other side. The linex loss function assigns unequal weights to the underestimation and overestimation by using a shape parameter. For small values of the shape parameter, the linex loss function is approximately symmetric and not too far from the quadratic loss function (cf. Zellner, 1986). The linex loss function is more general than the squared error loss function as the later is a special case of the former.

Zellner (1986) studied the properties of estimation and prediction procedures under the linex loss function. He showed that some usual estimators that are admissible relative to the squared error loss function, are inadmissible relative to the linex loss function. For example, Zellner (1986) proved that the UE $\bar{X}$ of the univariate normal mean is inadmissible relative to the linex loss function, as the risk of the estimator $\bar{X} - a\sigma^2/2n$ is less than that of the UE, where a is the shape parameter of the linex loss function, σ^2 is the population variance and n is the size of the sample. In the case of unknown σ^2, he suggested using $S^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2/n$. Later, Rojo (1987) generalized Zellner's result and showed that under the linex loss function any estimator of the form $c\bar{X} + d$, of θ is admissible if either $0 \leq c < 1$, or $c = 1$ and $d = -a\sigma^2/2n$. Otherwise, $c\bar{X} + d$ is inadmissible. Pandey and Rai (1996) investigated the properties of the so called testimator, a choice between $\bar{X}$ and $\bar{X} - a\sigma^2/2n$, of the univariate normal mean under the linex loss function. They showed that the testimator dominates the admissible estimator in terms of the linex risk in certain region of the parametric space. Further contributions to this area include Bhattacharaya *et al.* (2002), Parsian and Farispour (1993), Parsian *et al.* (1993), and Pandey and Rai (1992), to mention a few.

In this thesis we have derived the risk functions of the UE and PTE of the slope parameter of the simple linear regression model under the linex loss function. Using the moment generating function of the PTE, the bias and mse functions of the PTE of the slope parameter are derived. With respect to the linex loss function, the performance of the PTE relative to that of the UE is investigated. It is found that with respect to the linex loss function, the PTE dominates the UE in the neighbourhood of $\Delta = 0$, where Δ is the non-centrality parameter of non-central Student's t distribution. For very large value of Δ, the performance of the PTE is the same as that of the UE. Like the form of the linex loss function, the form of the risk function of the PTE is also asymmetric. However, for very small value of the shape parameter of the linex loss function, the form of the risk function of the PTE is almost symmetric.

Part I

Estimation Under Squared Error Loss Function

Chapter 2

Estimation of the Slope Parameter of Simple Linear Regression Model

2.1 Introduction

Consider a set of n random sample observations $(x_i,\ y_i)$ for $i = 1, 2, \ldots, n$ from the simple linear regression model

$$y = \beta_0 + \beta_1 x + \varepsilon \tag{2.1.1}$$

where y is the response variable, β_0 is the intercept parameter, β_1 is the slope parameter, x is the predictor and ε is the error component associated with the response variable. Assume that the errors are independently and identically distributed as a normal variable with mean 0 and variance σ^2. Then, in conventional notation we write ε iid $N(0,\ \sigma^2)$.

The exclusive sample information based maximum likelihood estimator of the slope parameter β_1 is known as the unrestricted estimator (UE). Assume that uncertain non-sample prior information on the value of the slope parameter is available either from previous study or from practical experience of researchers or experts. Such non-sample prior information about the value of β_1 can be expressed in the form of the null hypothesis

$$H_0 : \beta_1 = \beta_{10} \tag{2.1.2}$$

which may be true, but there is doubt. The estimator of β_1 under the null hypothesis in (2.1.2) is known as the restricted estimator (RE). We wish to combine the sample and uncertain non-sample prior information in estimating the slope β_1.

Following Khan and Saleh (2001), we assign a coefficient of distrust d, $0 \leq d \leq 1$, for the non-sample prior information, as a measure of the degree of distrust in the null hypothesis. First we obtain the unrestricted and restricted estimators of the unknown slope β_1 and the common variance σ^2 from the likelihood function of the sample. Based on the UE and RE of σ^2, we select the likelihood ratio test for testing H_0 in (2.1.2) against the alternative hypothesis

$$H_a : \beta_1 \neq \beta_{10}. \tag{2.1.3}$$

We then use the test statistic and coefficient of distrust, as well as the sample and non-sample prior information to define some alternative estimators of the unknown slope β_1.

Using the above methods we define a number of improved estimators of the slope parameter β_1, namely the shrinkage restricted estimator (SRE), shrinkage preliminary test estimator (SPTE), and the shrinkage estimator (SE). To compare the performances of the estimators we investigate their bias and mean square error (mse) functions, both analytically and graphically. The relative efficiencies of the estimators are also studied to search for a better estimator in some sense. Extensive computations have been undertaken to check the properties of the estimators. Analytical and graphical analyses reveal that although none of the estimators has uniformly superior statistical properties, the SE dominates the other estimators considered in this study, provided the non-sample prior information regarding the value of β_1 is not too far from its true value. As the prior information is obtained from previous studies or expert's knowledge, it is expected that such an information will not be too far from the true value of the parameter.

The layout of this chapter is as follows. Section 2.2 deals with the specification of the model and definition of the unrestricted maximum likelihood estimators of β_1 and σ^2 as well as the derivation of the likelihood ratio test statistic to test the null hypothesis in (2.1.2). Three alternative estimators of the slope parameter are defined in Section 2.3. Expressions for the bias and mse of the estimators are derived in Section 2.4. A comparative study of the relative efficiencies of the estimators is provided in Section 2.5. Some concluding remarks are presented in Section 2.6. Selected MATLAB codes, used for producing graphs, are presented in Appendix 2.A.

2.2 Some Preliminaries

Let the n sample responses from the linear regression model in (2.1.1) be expressed in the following convenient form

$$\boldsymbol{y} = \beta_0 \mathbf{1}_n + \beta_1 \boldsymbol{x} + \boldsymbol{\varepsilon} \tag{2.2.1}$$

where $\boldsymbol{y} = (y_1, \ldots, y_n)'$ is an $n \times 1$ vector of responses, $\mathbf{1}_n = (1, \ldots, 1)'$ is a vector of one's, $\boldsymbol{x}$ is an $n \times 1$ vector of predictors, β_0 and β_1 are the unknown intercept and slope parameters respectively, and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)'$ is a vector of errors with independent components which is distributed as $N_n(\mathbf{0},\ \sigma^2 I_n)$. Hence,

$$\mathrm{E}[\boldsymbol{\varepsilon}] = \mathbf{0} \tag{2.2.2}$$

and

$$\mathrm{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] = \sigma^2 I_n. \tag{2.2.3}$$

Here, σ^2 stands for the variance of each of the error components in $\boldsymbol{\varepsilon}$ and I_n is the identity matrix of order n.

The exclusive sample information based UE of the slope parameter β_1 is given by

$$\tilde{\beta}_1 = S_{xx}^{-1} S_{xy} \tag{2.2.4}$$

where $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2$, the sum of squares of x and $S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$, the sum of products of x and y .

It is well known that the UE of σ^2 is

$$S_n^{*2} = \frac{1}{n}(\boldsymbol{y} - \hat{\boldsymbol{y}})'(\boldsymbol{y} - \hat{\boldsymbol{y}}) \tag{2.2.5}$$

where $\hat{\boldsymbol{y}} = \tilde{\beta}_0 \mathbf{1}_n + \tilde{\beta}_1 \boldsymbol{x}$ in which $\tilde{\beta}_0$ is the UE of β_0. This estimator of σ^2 is biased. However, an unbiased estimator of σ^2 is

$$S_n^2 = \frac{1}{n-2}(\boldsymbol{y} - \hat{\boldsymbol{y}})'(\boldsymbol{y} - \hat{\boldsymbol{y}}). \tag{2.2.6}$$

The unbiased estimator of σ^2 has a scaled χ^2 distribution with shape parameter $\nu = (n-2)$. The estimated standard error of $\tilde{\beta}_1$ is $S_n S_{xx}^{-1/2}$.

To be able to use the uncertain non-sample prior information in the estimation of the slope, it is essential to remove the element of uncertainty concerning its value. To remove the uncertainty in the non-sample prior information Fisher suggested (cf. Khan and Saleh, 2001) conducting an appropriate statistical test on the null-hypothesis. For the problem under study, an appropriate test is the likelihood ratio test, and the test statistic is given by

$$\mathcal{L}_\nu = \frac{\sqrt{S_{xx}}(\tilde{\beta}_1 - \beta_{10})}{S_n}. \tag{2.2.7}$$

Under H_a, the above test statistic $\mathcal{L}_\nu$ follows the non-central Student's t distribution with ν degrees of freedom (d.f.) and non-centrality parameter Δ, given by

$$\Delta = \frac{\sqrt{S_{xx}}(\beta_1 - \beta_{10})}{\sigma}. \tag{2.2.8}$$

Equivalently, under H_a, $\mathcal{L}_\nu^2$ follows the non-central F distribution with $(1, \nu)$ degrees of freedom and non-centrality parameter Δ^2. Under the null-hypothesis $\mathcal{L}_\nu$ and $\mathcal{L}_\nu^2$ follow central Student's t and F distributions respectively with appropriate degrees of freedom. This test statistic was used by Bancroft (1944) to define the preliminary test estimator (PTE). In this study, the same test statistic is used to define the shrinkage preliminary test and shrinkage estimators by following the preliminary test approach.

For the model in (2.1.1) the sampling distribution of the UE of β_1 is normal with mean and variance given by

$$\mathrm{E}[\tilde{\beta}_1] = \beta_1 \tag{2.2.9}$$

and

$$\mathrm{Var}\left[\tilde{\beta}_1\right] = \sigma^2 S_{xx}^{-1} \tag{2.2.10}$$

respectively. Therefore, $\tilde{\beta}_1$ is unbiased for β_1, and hence its mse is the same as its variance. Evidently, the bias and the mse of $\tilde{\beta}_1$ are given by

$$\mathrm{B}_1\left[\tilde{\beta}_1\right] = 0 \tag{2.2.11}$$

and

$$\mathrm{M}_1\left[\tilde{\beta}_1\right] = \sigma^2 S_{xx}^{-1} \tag{2.2.12}$$

respectively. In this study, the above bias and mse functions are compared with those of the SRE, SPTE and SE of β_1.

2.3 Proposed Improved Estimators of the Slope

As part of incorporating the uncertain non-sample prior information into the estimation process, first we combine the exclusive sample information based UE $\tilde{\beta}_1$ with the non-sample prior information presented in the form of the null hypothesis in (2.1.2) in some reasonable way. Consider a simple convex combination of $\tilde{\beta}_1$ and $\hat{\beta}_1$ as

$$\hat{\beta}_1^{\mathrm{SRE}} = d\tilde{\beta}_1 + (1-d)\hat{\beta}_1 \tag{2.3.1}$$

where $\hat{\beta}_1 = \beta_{10}$ and $0 \leq d \leq 1$. This estimator of β_1 is called the shrinkage restricted estimator (SRE). Here, $d = 0$ means that there is *no distrust* in the H_0, and then we get $\hat{\beta}_1^{\mathrm{SRE}} = \hat{\beta}_1$, the RE, while $d = 1$ means that there is *complete distrust* in the H_0 and then we get $\hat{\beta}_1^{\mathrm{SRE}} = \tilde{\beta}_1$, the UE . If $0 < d < 1$ (that is, the degree of distrust is an intermediate value) then the SRE of β_1 takes an interpolated value between $\hat{\beta}_1$ and $\tilde{\beta}_1$, given by (2.3.1). The shrinkage restricted estimator, as defined above, is normally distributed with mean and mean square error given by

$$\mathrm{E}\left[\hat{\beta}_1^{\mathrm{SRE}}\right] = d\beta_1 + (1-d)\beta_{10} \tag{2.3.2}$$

and

$$\mathrm{M}_2\left[\hat{\beta}_1^{\mathrm{SRE}}\right] = \frac{\sigma^2}{S_{xx}}\left[d^2 + (1-d)^2\Delta^2\right] \tag{2.3.3}$$

respectively.

Following Khan and Saleh (2001), the shrinkage preliminary test estimator of the slope parameter β_1 is defined as

$$\begin{aligned}
\hat{\beta}_1^{\mathrm{SPTE}} &= \hat{\beta}_1^{\mathrm{SRE}}\, I\big(|t_\nu| < t_{\alpha/2}\big) + \tilde{\beta}_1\, I\big(|t_\nu| \geq t_{\alpha/2}\big) \\
&= \hat{\beta}_1^{\mathrm{SRE}}\, I\big(|t_\nu| < t_{\alpha/2}\big) + \tilde{\beta}_1 \left\{1 - I\big(|t_\nu| < t_{\alpha/2}\big)\right\} \\
&= \tilde{\beta}_1 - \left[\tilde{\beta}_1 - \hat{\beta}_1^{\mathrm{SRE}}\right] I\big(|t_\nu| < t_{\alpha/2}\big) \\
&= \tilde{\beta}_1 - \left[\tilde{\beta}_1 - d\tilde{\beta}_1 - (1-d)\hat{\beta}_1\right] I\big(|t_\nu| < t_{\alpha/2}\big) \\
&= \tilde{\beta}_1 - (1-d)(\tilde{\beta}_1 - \hat{\beta}_1)\, I(|t_\nu| < t_{\alpha/2})
\end{aligned} \tag{2.3.4}$$

where $I(A)$ is an indicator function of the set A and $t_{\alpha/2}$ is the critical value chosen for the two-sided α-level test based on the Student's t distribution with ν degrees of freedom. A simplified form of the above shrinkage preliminary test estimator is given by

$$\hat{\beta}_1^{\mathrm{SPTE}} = \hat{\beta}_1\, I\big(|t_\nu| < t_{\alpha/2}\big) + \tilde{\beta}_1\, I\big(|t_\nu| \geq t_{\alpha/2}\big) \tag{2.3.5}$$

which is a special case of (2.3.4) when $d = 0$. Note that, $\hat{\beta}_1^{\mathrm{SPTE}}(d \neq 0)$ is a combination of $\hat{\beta}_1^{\mathrm{SRE}}$ and $\tilde{\beta}_1$, and $\hat{\beta}_1^{\mathrm{SPTE}}$ is a choice between $\hat{\beta}_1$ and $\tilde{\beta}_1$. For the convenience of the derivation of the bias and mean square error function of the SPTE, (2.3.4) may be rewritten as

$$\hat{\beta}_1^{\mathrm{SPTE}} = \tilde{\beta}_1 - (1-d)(\tilde{\beta}_1 - \hat{\beta}_1)\, I(F < F_\alpha) \tag{2.3.6}$$

where F_α is the $(1-\alpha)^{\text{th}}$ quantile of the central F distribution with $(1, \nu)$ degrees of freedom. For an equivalent expression of the SPTE see Khan and Saleh (2001). If $d = 0$, (2.3.6) is

$$\hat{\beta}_1^{\text{SPTE}}(d = 0) = \tilde{\beta}_1 - (\tilde{\beta}_1 - \hat{\beta}_1)\, I(F < F_\alpha) \tag{2.3.7}$$

which is the PTE of β_1.

The performance of the SPTE depends on the pre-selected level of significance, α of the test. To overcome this limitation, the shrinkage estimator (SE) of β_1 is considered and is defined as

$$\begin{aligned}\hat{\beta}_1^{\text{SE}} &= \hat{\beta}_1 + (1 - c|t_\nu|^{-1})(\tilde{\beta}_1 - \hat{\beta}_1) \\ &= \hat{\beta}_1 + \left\{1 - \frac{cS_n}{|\sqrt{S_{xx}}(\tilde{\beta}_1 - \hat{\beta}_1)|}\right\}(\tilde{\beta}_1 - \hat{\beta}_1) \\ &= \tilde{\beta}_1 - \frac{cS_n(\tilde{\beta}_1 - \hat{\beta}_1)}{\left|\sqrt{S_{xx}}(\tilde{\beta}_1 - \hat{\beta}_1)\right|}\end{aligned} \tag{2.3.8}$$

where c is the shrinkage constant, a function of n. If $|t_\nu| = \left|\frac{\sqrt{S_{xx}}(\tilde{\beta}_1 - \hat{\beta}_1)}{S_n}\right|$ is large, $\hat{\beta}_1^{\text{SE}}$ tends towards $\tilde{\beta}_1$, while if $|t_\nu|$ is small, equal to c, $\hat{\beta}_1^{\text{SE}}$ tends towards β_{10}. Unlike the shrinkage preliminary test estimator, the shrinkage estimator does not depend on the level of significance of the test. Though the dimension of the population considered in this study, is less than three, unlike the Stein-type shrinkage estimator, the SE in (2.3.8) is admissible over the UE with respect to the squared error loss criterion.

2.4 Some Statistical Properties

In this section, the bias and mean square error functions of the SRE, SPTE and SE of the slope parameter β_1 are derived. Also, some of the important features of these functions are discussed.

2.4.1 The Bias and MSE of the SRE

Theorem 2.41 *The bias function of the shrinkage restricted estimator of the slope parameter* β_1 *is given by*

$$\mathrm{B}_2\left[\hat{\beta}_1^{\mathrm{SRE}}\right] = -(1-d)\frac{\sigma}{\sqrt{S_{xx}}}\Delta \tag{2.4.1}$$

where Δ *is the non-centrality parameter of non-central Student's* t *distribution.*

Proof. By definition, the bias function of the SRE of β_1 is

$$\begin{aligned}\mathrm{B}_2\left[\hat{\beta}_1^{\mathrm{SRE}}\right] &= \mathrm{E}\left[\hat{\beta}_1^{\mathrm{SRE}} - \beta_1\right] \\ &= \mathrm{E}\left[d\tilde{\beta}_1 - \beta_1 + (1-d)\beta_{10}\right] \\ &= d\beta_1 - \beta_1 + (1-d)\beta_{10} \\ &= -(1-d)(\beta_1 - \beta_{10}) \\ &= -(1-d)\frac{\sigma}{\sqrt{S_{xx}}}\Delta. \end{aligned} \tag{2.4.2}$$

This completes the proof of the theorem.

Theorem 2.42 *The mean square error function of the shrinkage restricted estimator of the slope parameter* β_1 *is given by*

$$\mathrm{M}_2\left[\hat{\beta}_1^{\mathrm{SRE}}\right] = \frac{\sigma^2}{S_{xx}}\left[d^2 + (1-d)^2\Delta^2\right] \tag{2.4.3}$$

where Δ^2 *is the non-centrality parameter of non-central* F *distribution.*

Proof. By definition, the mse function of the SRE of β_1 is

$$\begin{aligned}\mathrm{M}_2\left[\hat{\beta}_1^{\mathrm{SRE}}\right] &= \mathrm{E}\left[\hat{\beta}_1^{\mathrm{SRE}} - \beta_1\right]^2 \\ &= \mathrm{E}\left[d\tilde{\beta}_1 - \beta_1 + (1-d)\beta_{10}\right]^2\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{E}\Big[d(\tilde{\beta}_1 - \beta_1) - (1-d)(\beta_1 - \beta_{10})\Big]^2 \\
&= d^2 \mathrm{E}\Big[\tilde{\beta}_1 - \beta_1\Big]^2 + (1-d)^2(\beta_1 - \beta_{10})^2 \\
&= \frac{d^2\sigma^2}{S_{xx}} + (1-d)^2(\beta_1 - \beta_{10})^2 \\
&= \frac{\sigma^2}{S_{xx}}\left[d^2 + (1-d)^2\Delta^2\right]. \qquad (2.4.4)
\end{aligned}
$$

This completes the proof of the theorem.

If the null hypothesis in (2.1.2) is true, the value of the parameter Δ^2 is 0; otherwise, it is always positive. The statistical properties of the SRE, SPTE and SE depend on the value of this parameter. This feature is investigated in greater detail in the forthcoming sections.

2.4.2 The Bias and MSE of the SPTE

Theorem 2.43 *The bias function of the shrinkage preliminary test estimator of the slope parameter β_1 is given by*

$$
\mathrm{B}_3\left[\hat{\beta}_1^{\mathrm{SPTE}}\right] = -(1-d)\frac{\sigma}{\sqrt{S_{xx}}}\Delta\, G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right) \qquad (2.4.5)
$$

where $G_{a,b}(\cdot;\, \Delta^2)$ is the c.d.f. of the non-central F distribution with $(a,\, b)$ degrees of freedom and non-centrality parameter Δ^2.

Proof. By definition, the bias function of the SPTE of β_1 is

$$
\begin{aligned}
\mathrm{B}_3\left[\hat{\beta}_1^{\mathrm{SPTE}}\right] &= \mathrm{E}\left[\hat{\beta}_1^{\mathrm{SPTE}} - \beta_1\right] \\
&= \mathrm{E}\Big[\tilde{\beta}_1 - \beta_1 - (1-d)\,(\tilde{\beta}_1 - \beta_{10})\, I\big(F < F_\alpha\big)\Big] \\
&= -(1-d)\frac{\sigma}{\sqrt{S_{xx}}}\mathrm{E}\left[\frac{\sqrt{S_{xx}}(\tilde{\beta}_1 - \beta_{10})}{\sigma}\, I\left(\frac{S_{xx}(\tilde{\beta}_1 - \beta_{10})^2}{S_n^2} < F_\alpha\right)\right]. \qquad (2.4.6)
\end{aligned}
$$

Note $Z = \sqrt{S_{xx}}(\tilde{\beta}_1 - \beta_{10})\sigma^{-1}$ is distributed as $N(\Delta, 1)$, and $\nu\sigma^{-2}S_n^2$ is distributed as a central chi-square variable with ν degrees of freedom. Therefore,

$$\mathrm{B}_3\left[\hat{\beta}_1^{\text{SPTE}}\right] = -(1-d)\frac{\sigma}{\sqrt{S_{xx}}}\,\mathrm{E}\left[Z\ I\left(\frac{\nu Z^2}{\chi_\nu^2} < F_\alpha\right)\right]. \tag{2.4.7}$$

Applying Theorem 1, Appendix B2, Judge and Bock (1978) to (2.4.7), the bias function of the SPTE of β_1 can be written as

$$\begin{aligned}\mathrm{B}_3\left[\hat{\beta}_1^{\text{SPTE}}\right] &= -(1-d)(\beta_1 - \beta_{10})\,G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right)\\ &= -(1-d)\frac{\sigma}{\sqrt{S_{xx}}}\Delta\,G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right).\end{aligned} \tag{2.4.8}$$

This completes the proof of the theorem.

Theorem 2.44 *The mean square error function of the shrinkage preliminary test estimator of the slope parameter β_1 is given by*

$$\begin{aligned}\mathrm{M}_3\left[\hat{\beta}_1^{\text{SPTE}}\right] = \frac{\sigma^2}{S_{xx}}\Bigg[1 - (1-d^2)G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right) + (1-d)\Delta^2\\ \times\left\{2G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right) - (1+d)G_{5,\nu}\left(\frac{1}{5}F_\alpha; \Delta^2\right)\right\}\Bigg]\end{aligned} \tag{2.4.9}$$

where $G_{a,b}(\cdot; \Delta^2)$ is the c.d.f. of the non-central F distribution with (a, b) degrees of freedom and non-centrality parameter Δ^2.

Proof. By definition, the mse function of the SPTE of β_1 is

$$\begin{aligned}\mathrm{M}_3\left[\hat{\beta}_1^{\text{SPTE}}\right] &= \mathrm{E}\left[\hat{\beta}_1^{\text{SPTE}} - \beta_1\right]^2\\ &= \mathrm{E}\left[(\tilde{\beta}_1 - \beta_1) - (1-d)(\tilde{\beta}_1 - \beta_{10})\,I(F < F_\alpha)\right]^2\\ &= \mathrm{E}\left[\tilde{\beta}_1 - \beta_1\right]^2 + (1-d)^2\mathrm{E}\left[(\tilde{\beta}_1 - \beta_{10})^2\,I(F < F_\alpha)\right]\\ &\quad - 2(1-d)\,\mathrm{E}\left[(\tilde{\beta}_1 - \beta_1)(\beta_1 - \beta_{10})\,I(F < F_\alpha)\right]\\ &= \frac{\sigma^2}{S_{xx}} + (1-d)^2\mathrm{E}\left[(\tilde{\beta}_1 - \beta_{10})^2\,I(F < F_\alpha)\right]\\ &\quad - 2(1-d)\,\mathrm{E}\left[(\tilde{\beta}_1 - \beta_1)(\beta_1 - \beta_{10})\,I(F < F_\alpha)\right].\end{aligned} \tag{2.4.10}$$

The second term of the right hand side of (2.4.10) is

$$(1-d)^2 \mathrm{E}\Big[(\tilde{\beta}_1 - \beta_{10})^2\, I(F < F_\alpha)\Big] = (1-d)^2 \frac{\sigma^2}{S_{xx}}\, \mathrm{E}\left[Z^2\ I\left(\frac{\nu Z^2}{\chi^2_\nu} < F_\alpha\right)\right]. \tag{2.4.11}$$

Applying Theorem 3, Appendix B2, Judge and Bock (1978) to (2.4.11), we get

$$\begin{aligned}&(1-d)^2 \mathrm{E}\Big[(\tilde{\beta}_1 - \beta_{10})^2\, I(F < F_\alpha)\Big]\\ &= (1-d)^2 \frac{\sigma^2}{S_{xx}}\, G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right) + (1-d)^2 \frac{\sigma}{S_{xx}} \Delta^2\, G_{5,\nu}\left(\frac{1}{5}F_\alpha; \Delta^2\right).\end{aligned} \tag{2.4.12}$$

Now,

$$\begin{aligned}&\mathrm{E}\Big[(\tilde{\beta}_1 - \beta_1)(\tilde{\beta}_1 - \beta_{10})\, I(F < F_\alpha)\Big]\\ &= \mathrm{E}\Big[\Big\{(\tilde{\beta}_1 - \beta_{10}) - (\beta_1 - \beta_{10})\Big\}\, (\tilde{\beta}_1 - \beta_{10})\, I(F < F_\alpha)\Big]\\ &= \mathrm{E}\Big[(\tilde{\beta}_1 - \beta_{10})^2\, I(F < F_\alpha)\Big] - \frac{\sigma}{\sqrt{S_{xx}}} \Delta \mathrm{E}\Big[(\tilde{\beta}_1 - \beta_{10})\, I(F < F_\alpha)\Big]\\ &= \frac{\sigma^2}{S_{xx}}\, \mathrm{E}\left[Z^2\ I\left(\frac{\nu Z^2}{\chi^2_\nu} < F_\alpha\right)\right] - \frac{\sigma^2}{S_{xx}} \Delta\, \mathrm{E}\left[Z\ I\left(\frac{\nu Z^2}{\chi^2_\nu} < F_\alpha\right)\right].\end{aligned} \tag{2.4.13}$$

Applying Theorems 1 and 3, Appendix B2, Judge and Bock (1978) to (2.4.13), the last term of (2.4.10) becomes

$$\begin{aligned}&2(1-d)\, \mathrm{E}\Big[(\tilde{\beta}_1 - \beta_1)(\tilde{\beta}_1 - \beta_{10})\Big]\ I(F < F_\alpha)\\ &= 2(1-d) \frac{\sigma^2}{S_{xx}} \left[\Delta^2 G_{5,\nu}\left(\frac{1}{5}F_\alpha; \Delta^2\right) + \Big\{1 - \Delta^2\Big\} G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right)\right].\end{aligned} \tag{2.4.14}$$

Collecting all terms, the mse function of the SPTE of β_1 can be expressed as

$$\begin{aligned}\mathrm{M}_3\Big[\hat{\beta}_1^{\mathrm{SPTE}}\Big] = \frac{\sigma^2}{S_{xx}}\bigg[&1 - (1-d^2)\, G_{3,\nu}\left(\frac{1}{3}F_\alpha;\ \Delta^2\right) + (1-d)\Delta^2\\ &\times \left\{2G_{3,\nu}\left(\frac{1}{3}F_\alpha;\ \Delta^2\right) - (1+d)\, G_{5,\nu}\left(\frac{1}{5}F_\alpha;\ \Delta^2\right)\right\}\bigg].\end{aligned} \tag{2.4.15}$$

This completes the proof of the theorem.

Figure 2.1 displays the behaviour of the mse function for a range of values of Δ^2 and selected values of α and d.

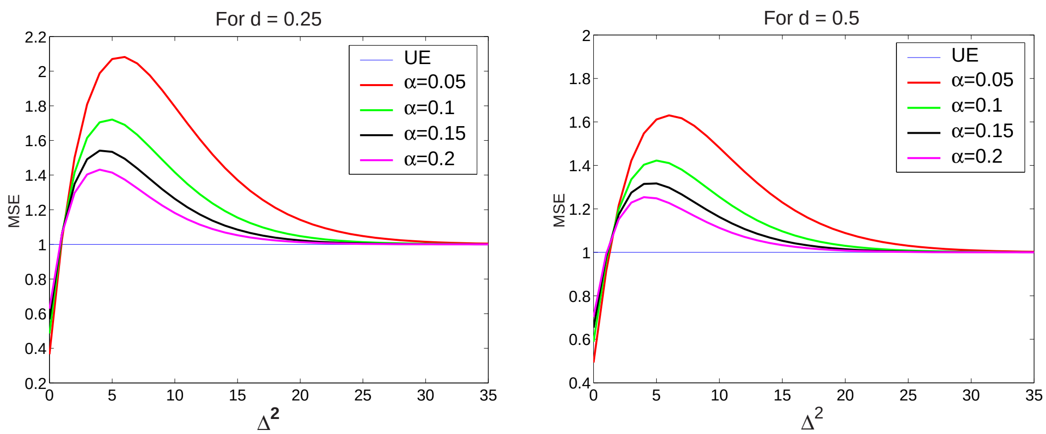

Figure 2.1: The mean square error of the SPTE of the Slope.

2.4.2.1 Some Properties of the MSE of the SPTE

Here we discuss some important features of the mse function of the SPTE of β_1.

- Under the null hypothesis, $\Delta^2 = 0$ and hence the mse of $\hat{\beta}_1^{\text{SPTE}}$ is

$$\frac{\sigma^2}{S_{xx}}\left[1-(1-d^2)\,G_{3,\nu}\left(\frac{1}{3}F_\alpha;0\right)\right] < \frac{\sigma^2}{S_{xx}}, \qquad \text{if } d<1. \tag{2.4.16}$$

Thus, when $\Delta^2 = 0$ the SPTE of β_1 performs better than $\tilde{\beta}_1$, the UE. As $\alpha \to 0$, $G_{3,\nu}\left(\frac{1}{3}F_\alpha;0\right) \to 1$, and hence

$$\frac{\sigma^2}{S_{xx}}\left[1-(1-d^2)\,G_{3,\nu}\left(\frac{1}{3}F_\alpha;0\right)\right] \to \frac{d^2\sigma^2}{S_{xx}}, \tag{2.4.17}$$

which is the mse of $\hat{\beta}_1^{\text{SRE}}$. On the other hand, if $F_\alpha \to 0$, $G_{3,\nu}\left(\frac{1}{3}F_\alpha;0\right) \to$

0, and hence

$$\frac{\sigma^2}{S_{xx}}\left[1-(1-d^2)\,G_{3,\nu}\left(\frac{1}{3}F_\alpha;0\right)\right]\to\frac{\sigma^2}{S_{xx}} \tag{2.4.18}$$

which is the mse of $\tilde{\beta}_1$.

- As $\Delta^2 \to \infty$, $G_{m,\nu}\left(\frac{1}{m}F_\alpha;\Delta^2\right) \to 0$, and $\mathrm{M}_3\left[\hat{\beta}_1^{\text{SPTE}}\right]$ in (2.4.15) tends towards $\frac{\sigma^2}{S_{xx}}$, the mse of $\tilde{\beta}_1$.

- Since $G_{3,\nu}\left(\frac{1}{3}F_\alpha;\Delta^2\right)$ is always greater than $G_{5,\nu}\left(\frac{1}{5}F_\alpha;\Delta^2\right)$ for any value of α, replacing $G_{5,\nu}\left(\frac{1}{5}F_\alpha;\Delta^2\right)$ by $G_{3,\nu}\left(\frac{1}{3}F_\alpha;\Delta^2\right)$ in (2.4.15) implies

$$\begin{aligned}\mathrm{M}_3\left[\hat{\beta}_1^{\text{SPTE}}\right] &\geq \frac{\sigma^2}{S_{xx}}\left[1+(1-d^2)G_{3,\nu}\left(\frac{1}{3}F_\alpha;\Delta^2\right)\left\{(1-d)\Delta^2-(1+d)\right\}\right]\\ &\geq \frac{\sigma^2}{S_{xx}} \quad \text{whenever } \Delta^2 > \frac{1+d}{1-d}.\end{aligned}$$

  On the other hand, (2.4.15) may be rewritten as

$$\begin{aligned}&\frac{\sigma^2}{S_{xx}}\left[1+(1-d)G_{3,\nu}\left(\frac{1}{3}F_\alpha;\Delta^2\right)\left\{2\Delta^2-(1+d)\right\}-(1-d^2)G_{5,\nu}\left(\frac{1}{5}F_\alpha;\Delta^2\right)\right]\\ &\leq \frac{\sigma^2}{S_{xx}} \quad \text{whenever } \Delta^2 < \frac{1+d}{2}.\end{aligned}$$

  This means that the mse of $\hat{\beta}_1^{\text{SPTE}}$ as a function of Δ^2 crosses the constant line of $\mathrm{M}_1\left[\tilde{\beta}_1\right] = \frac{\sigma^2}{S_{xx}}$ in the interval $\left(\frac{1+d}{2}, \frac{1+d}{1-d}\right)$.

A general picture of the mse function of the SPTE of β_1 can be described as follows:

The mse function begins with the smallest value $\frac{\sigma^2}{S_{xx}}\left[1-(1-d^2)\times G_{3,\nu}\left(\frac{1}{3}F_\alpha\,;0\right)\right]$ at $\Delta^2=0$. As Δ^2 grows larger, the function increases monotonically, crossing the constant line $\sigma^2 S_{xx}^{-1}$ in the interval $\left(\frac{1+d}{2}, \frac{1+d}{1-d}\right)$ and reaching its maximum in the interval $\left(\frac{1+d}{1-d}, \infty\right)$. Finally, as $\Delta^2 \to \infty$, the mse of the SPTE of β_1 monotonically decreases and approaches $\sigma^2 S_{xx}^{-1}$, the mse of the UE of β_1.

2.4.2.2 Determination of Optimum α for the SPTE

Clearly, the mse, and hence the relative efficiency of the shrinkage preliminary test estimator relative to the unrestricted estimator, depends on the level of significance α of the test of the null-hypothesis and the value of Δ^2.

Let the efficiency of the SPTE relative to the UE of β_1 be denoted by $\text{Eff}(\alpha; \Delta^2)$. Then

$$\begin{aligned}\text{Eff}(\alpha;\, \Delta^2) &= \mathrm{M}_1\left[\tilde{\beta}_1\right] \Big/ \mathrm{M}_3\left[\hat{\beta}_1^{\text{SPTE}}\right] \\ &= [1 + g(\Delta^2)]^{-1} \qquad (2.4.19)\end{aligned}$$

where $$\begin{aligned}g(\Delta^2) &= (1-d)\Delta^2\Big\{2G_{3,\nu}\Big(\frac{1}{3}F_\alpha; \Delta^2\Big)\Big\} - (1+d)\, G_{5,\nu}(\frac{1}{5}F_\alpha; \Delta^2\Big) \\ &\quad - (1-d^2)G_{3,\nu}\Big(\frac{1}{3}F_\alpha; \Delta^2\Big). \qquad (2.4.20)\end{aligned}$$

The relative efficiency function $\text{Eff}(\alpha;\, \Delta^2)$ attains its maximum at $\Delta^2 = 0$ for all α, and is given by

$$\text{Eff}(\alpha;\, \Delta^2) = \left[1 - (1-d^2)G_{3,\nu}\Big(\frac{1}{3}F_\alpha; 0\Big)\right]^{-1} \geq 1. \qquad (2.4.21)$$

As Δ^2 departs from the origin, $\text{Eff}(\alpha;\, \Delta^2)$ decreases monotonically, crossing the line $\text{Eff}(\alpha;\, \Delta^2) = 1$ in the interval $\left(\frac{1+d}{2}, \frac{1+d}{1-d}\right)$ and reaching a minimum at $\Delta^2 = \Delta^2_{\text{min}}$. From that point on it increases monotonically towards one as $\Delta^2 \to \infty$. For $\Delta^2 = 0$ and varying significance level, we have

$$\max_{\alpha} \text{Eff}(\alpha;\, 0) = \text{Eff}(0; 0) = d^{-2}. \qquad (2.4.22)$$

As a function of α, $\text{Eff}(\alpha; 0)$ decreases as α increases. On the other hand, $\text{Eff}(\alpha; \Delta^2)$ as a function of Δ^2 is decreasing, and the curves $\text{Eff}(0; \Delta^2)$ and

$\text{Eff}(1/2;\Delta^2) = 1$ intersect at $\Delta^2 = \frac{1+d}{1-d}$. The value of Δ^2 at the intersection decreases as α increases. Therefore, for two different levels of significance say, α_1 and α_2, $\text{Eff}(\alpha_1;\Delta^2)$ and $\text{Eff}(\alpha_2;\Delta^2)$ intersects below one. In order to choose an optimum level of significance with maximum relative efficiency, the following rule is adopted:

If it is known that $0 \leq \Delta^2 \leq \frac{1+d}{1-d}$, $\hat{\beta}_1$ is always chosen since $\text{Eff}(0,\Delta^2)$ is maximum for all Δ^2 in this interval. Generally, Δ^2 is unknown. In this case there is no way of choosing the uniformly best estimator of β_1.

Let us pre-assign a tolerable relative efficiency, say, Eff_0, and consider the set

$$A_\alpha = \left\{\alpha \mid \text{Eff}(\alpha;\Delta^2) \geq \text{Eff}_0\right\}. \tag{2.4.23}$$

An estimator $\hat{\beta}_1^{\text{SPTE}}$ is chosen which maximizes $\text{Eff}(\alpha;\Delta^2)$ over all $\alpha \in A_\alpha$ and Δ^2. Thus, for given Eff_0 the solution $\alpha = \alpha^*$

$$\max_{\alpha} \min_{\Delta^2} \text{Eff}(\alpha;\Delta^2) = \text{Eff}_0 \tag{2.4.24}$$

provides an optimal choice of α, and the procedure is known as the *maximin rule* of the optimum level of significance of the preliminary test. A numerical procedure along with practical illustration of selecting an optimal α is provided in Khan and Saleh (2001).

2.4.3 The Bias and MSE of the SE

Following Bolfarine and Zacks (1992), the bias and mean square error functions of the shrinkage estimator of the slope parameter β_1 are derived and presented in the following theorems.

Theorem 2.45 *The bias function of the shrinkage estimator of* β_1 *is given by*

$$\mathrm{B}_4\left[\hat{\beta}_1^{\mathrm{SE}}\right] = \frac{c\bar{x}\sigma}{\sqrt{S_{xx}}} K_\nu \Big\{ 1 - 2\Phi(-\Delta) \Big\} \tag{2.4.25}$$

where $K_n = \sqrt{\frac{2}{n-2} \frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-2}{2})}}$ *and* $\Phi(\cdot)$ *is the c.d.f. of the standard normal distribution.*

Proof. By definition, the bias function of the SE of β_1 is

$$\begin{aligned}
\mathrm{B}_4\left[\hat{\beta}_1^{\mathrm{SE}}\right] &= \mathrm{E}\left[\hat{\beta}_1^{\mathrm{SE}} - \beta_1\right] \\
&= \mathrm{E}\left[\tilde{\beta}_1 - \beta_1 - \frac{cS_n(\tilde{\beta}_1 - \beta_{10})}{\left|\sqrt{S_{xx}}(\tilde{\beta}_1 - \beta_{10})\right|}\right] \\
&= -c\,\mathrm{E}\left[\frac{S_n(\tilde{\beta}_1 - \beta_{10})}{\left|\sqrt{S_{xx}}(\tilde{\beta}_1 - \beta_{10})\right|}\right] \\
&= -\frac{c}{\sqrt{S_{xx}}}\,\mathrm{E}[S_n]\,\mathrm{E}\left[\frac{Z}{|Z|}\right] \qquad (2.4.26)
\end{aligned}$$

where $Z = \frac{\sqrt{S_{xx}}(\tilde{\beta}_1 - \beta_{10})}{\sigma} \sim \mathcal{N}(\Delta, 1)$.

Now, we evaluate $\mathrm{E}[S_n]$ and $\mathrm{E}\left[\frac{Z}{|Z|}\right]$.

By definition, $\frac{(n-2)S_n^2}{\sigma^2} \sim \chi^2_{n-2}$. Therefore,

$$f_{S_n}(y) = \frac{2}{2^{\frac{n-2}{2}}\Gamma\left(\frac{n-2}{2}\right)} e^{\frac{(n-2)y^2}{2\sigma^2}} \left\{\frac{(n-2)y^2}{\sigma^2}\right\}^{\frac{n-2}{2}-1} \frac{(n-2)y}{\sigma^2}. \tag{2.4.27}$$

Then, using the definition of expectation,

$$\mathrm{E}_{S_n}[y] = \frac{2}{2^{\frac{n-2}{2}}\Gamma(\frac{n-2}{2})} \int_0^\infty e^{-\frac{(n-2)y^2}{2\sigma^2}} \left\{\frac{(n-2)y^2}{\sigma^2}\right\}^{\frac{n-2}{2}} dy. \tag{2.4.28}$$

Consider $\frac{(n-2)y^2}{2\sigma^2} = y_1$. The Jacobian of the transformation is $|J| = \left(\frac{\sigma^2}{2(n-2)y_1}\right)^{1/2}$. Therefore,

$$\mathrm{E}_{S_n}[y] = \frac{2}{2^{\frac{n-2}{2}}\Gamma\left(\frac{n-2}{2}\right)} \int_0^\infty e^{-y_1} \left(2\,y_1\right)^{\frac{n-2}{2}} \left(\frac{\sigma^2}{2\,y_1(n-2)}\right)^{1/2} dy_1$$

$$= \frac{2^{1/2}\,\sigma}{(n-2)^{1/2}\,\Gamma(\frac{n-2}{2})} \int_0^\infty e^{-y_1}\, y_1^{\frac{n-1}{2}-1}\, dy_1$$
$$= \frac{(2\sigma^2)^{1/2}}{(n-2)^{1/2}} \frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-2}{2})}. \tag{2.4.29}$$

By definition,

$$\begin{aligned} \mathrm{E}\left[\frac{Z}{|Z|}\right] &= E\left[\frac{Z}{|Z|}\Big| Z > \Delta\right] + \mathrm{E}\left[\frac{Z}{|Z|}\Big| Z < -\Delta\right] \\ &= \int_\Delta^\infty \frac{z}{z} f(z)dz + \int_{-\infty}^{-\Delta} \frac{z}{-z} f(z)dz \\ &= 1 - \int_{-\infty}^{-\Delta} \frac{z}{-z} f(z)dz - \int_{-\infty}^{-\Delta} \frac{z}{-z} f(z)dz \\ &= 1 - P(z < -\Delta) - P(z < \Delta) \\ &= 1 - 2\,\Phi(-\Delta). \end{aligned} \tag{2.4.30}$$

Therefore, the bias function of the SE of β_1 is obtained as

$$\mathrm{B}_4\left[\hat{\beta}_1^{\mathrm{SE}}\right] = \frac{c\bar{x}\sigma}{\sqrt{S_{xx}}} K_n \Big\{1 - 2\Phi(-\Delta)\Big\}. \tag{2.4.31}$$

This completes the proof of the theorem.

From the expression of the above bias function, the quadratic bias function of the SE of β_1 is obtained as

$$\mathrm{QB}_4\left[\hat{\beta}_1^{\mathrm{SE}}\right] = \frac{\sigma^2\bar{x}^2}{S_{xx}} c^2 K_n^2 \Big\{2\Phi(\Delta) - 1\Big\}^2 \tag{2.4.32}$$

where $K_n = \sqrt{\frac{2}{n-2}} \frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-2}{2})}$.

Theorem 2.46 *The mean square error function of the shrinkage estimator of the slope parameter β_1 is given by*

$$\mathrm{M}_4\left[\hat{\beta}_1^{\mathrm{SE}}\right] = \frac{\sigma^2}{S_{xx}} \left\{1 + c^2 - 2cK_n\sqrt{\frac{2}{\pi}} e^{-\Delta^2/2}\right\} \tag{2.4.33}$$

where K_n is as defined in Theorem 2.45

Proof. By definition, the mse function of the SE of β_1 is

$$\begin{aligned}
\mathrm{M}_4\left[\hat{\beta}_1^{\mathrm{SE}}\right] &= \mathrm{E}\left[\hat{\beta}_1^{\mathrm{SE}} - \beta_1\right]^2 \\
&= \mathrm{E}\left[\tilde{\beta}_1 - \beta_1 - \frac{cS_n(\tilde{\beta}_1 - \beta_{10})}{|\sqrt{S_{xx}}(\tilde{\beta}_1 - \beta_{10})|}\right]^2 \\
&= \mathrm{E}\left[\tilde{\beta}_1 - \beta_1\right]^2 + c^2\,\mathrm{E}\big[S_n^2\big]\,\mathrm{E}\left[\frac{(\tilde{\beta}_1 - \beta_{10})^2}{[\sqrt{S_{xx}}(\tilde{\beta}_1 - \beta_{10})]^2}\right] \\
&\quad - 2c\,\mathrm{E}\left[\frac{(\tilde{\beta}_1 - \beta_1)(\tilde{\beta}_1 - \beta_{10})}{|\sqrt{S_{xx}}(\tilde{\beta}_1 - \beta_{10})|}\right]\mathrm{E}[S_n] \\
&= \frac{\sigma^2}{S_{xx}} + \frac{c^2\sigma^2}{S_{xx}} - 2c\frac{\sigma^2 K_n}{S_{xx}}\left\{E[|Z|] - \Delta\mathrm{E}\left[\frac{Z}{|Z|}\right]\right\}. \qquad (2.4.34)
\end{aligned}$$

where $Z \sim \mathcal{N}(\Delta, 1)$.

As $Z \sim \mathcal{N}(\Delta, 1)$, we write

$$f_{|Z|}(z) = \phi(z - \Delta) + \phi(z + \Delta). \qquad (2.4.35)$$

Now, by definition

$$\begin{aligned}
\mathrm{E}[|Z|] &= \int_0^\infty z\,\phi(z - \Delta)\,dz + \int_0^\infty z\,\phi(z + \Delta)\,dz \\
&= \int_\Delta^\infty z\,\phi(z)\,dz + \int_{-\Delta}^\infty z\,\phi(z)\,dz + \Delta\left\{\int_{-\Delta}^\infty \phi(z)\,dz - \int_\Delta^\infty \phi(z)\,dz\right\} \\
&= \int_\Delta^\infty z\,\phi(z)\,dz + \int_{-\Delta}^\infty z\,\phi(z)\,dz + \Delta\Big\{2\,\Phi(\Delta) - 1\Big\} \\
&= \sqrt{\frac{2}{\pi}}e^{-\Delta^2/2} + \Delta\Big\{2\Phi(\Delta) - 1\Big\} \qquad (2.4.36)
\end{aligned}$$

where $\Phi(\cdot)$ is the c.d.f. of the standard normal variable.

Therefore, the mse of the SE of β_1 is obtained as

$$\mathrm{M}_4\left[\hat{\beta}_1^{\mathrm{SE}}\right] = \frac{\sigma^2}{S_{xx}}\left\{1 + c^2 - 2cK_n\sqrt{\frac{2}{\pi}}e^{-\Delta^2/2}\right\}. \qquad (2.4.37)$$

2.4.3.1 Determination of the Optimum Value of c

A stationary point of $\mathrm{M}_4\left[\hat{\beta}_1^{\mathrm{SE}}\right]$ with respect to c occurs when the first derivative (with respect to c)

$$\mathrm{M}_4'\left[\hat{\beta}_1^{\mathrm{SE}}\right] = \frac{\sigma^2}{S_{xx}}\left\{2c - 2K_n\sqrt{\frac{2}{\pi}}e^{-\Delta^2/2}\right\} = 0\,, \tag{2.4.38}$$

from which

$$c = c^* = K_n\sqrt{\frac{2}{\pi}}e^{-\Delta^2/2}. \tag{2.4.39}$$

The second derivative of $\mathrm{M}_4\left[\hat{\beta}_1^{\mathrm{SE}}\right]$ with respect to c is

$$\mathrm{M}_4''\left[\hat{\beta}_1^{\mathrm{SE}}\right] = \frac{2\sigma^2}{S_{xx}} > 0\,. \tag{2.4.40}$$

Therefore, c^* is the value of c that minimizes (2.4.37). It depends on Δ^2 as shown in (2.4.39).

To make c^* independent of Δ^2, we choose $c^0 = \sqrt{\frac{2}{\pi}}K_n$. Thus, optimum $\mathrm{M}_4\left[\hat{\beta}_1^{\mathrm{SE}}\right]$ becomes

$$\mathrm{M}_4\left[\hat{\beta}_1^{\mathrm{SE}}\right] = \frac{\sigma^2}{S_{xx}}\left\{1 - \frac{2}{\pi}K_n^2\left[2e^{-\Delta^2/2} - 1\right]\right\}. \tag{2.4.41}$$

The above mse function of the SE of β_1 is compared with those of the other estimators of β_1 in the next section.

2.5 Performances Comparison of the Estimators

In this section, the quadratic bias functions and relative efficiencies of the SRE, SPTE and SE are compared with those of the UE of the slope parameter.

2.5.1 Comparison of the Quadratic Bias Functions

The quadratic bias functions of the SRE, SPTE and SE of the slope parameter β_1 are given by

$$\mathrm{QB}_2\left[\hat{\beta}_1^{\mathrm{SRE}}\right] = (1-d)^2\frac{\sigma^2}{S_{xx}}\Delta^2, \tag{2.5.1}$$

$$\mathrm{QB}_3\left[\hat{\beta}_1^{\mathrm{SPTE}}\right] = (1-d)^2\frac{\sigma^2}{S_{xx}}\Delta^2 G_{3,\nu}^2\left(\frac{1}{3}F_\alpha;\Delta^2\right) \tag{2.5.2}$$

and

$$\mathrm{QB}_4\left[\hat{\beta}_1^{\mathrm{SE}}\right] = \bar{x}^2c^2\frac{\sigma^2}{S_{xx}}K_n^2\Big\{2\Phi(\Delta)-1\Big\}^2 \tag{2.5.3}$$

respectively.

Clearly, under the null-hypothesis, $\mathrm{QB}_2\left[\hat{\beta}_1^{\mathrm{SRE}}\right] = \mathrm{QB}_3\left[\hat{\beta}_1^{\mathrm{SPTE}}\right] = \mathrm{QB}_4\left[\hat{\beta}_1^{\mathrm{SE}}\right]$ $= 0$ for all d and α. When $\Delta^2 \to \infty$, $\mathrm{QB}_2\left[\hat{\beta}_1^{\mathrm{SRE}}\right] \to \infty$ except at $d = 1$; $\mathrm{QB}_3\left[\hat{\beta}_1^{\mathrm{SPTE}}\right]$ tends to 0 for all α and d; and $\mathrm{QB}_4\left[\hat{\beta}_1^{\mathrm{SE}}\right] \to \frac{\sigma^2}{S_{xx}}c^2K_n^2$, a constant that does not depend on d. Therefore, in terms of quadratic bias, the SRE is uniformly dominated by both the SPTE and SE.

For very large values of Δ^2, the SE is dominated by the SPTE regardless of the value of α. From a small to moderate values of Δ^2, there is no uniform domination of one estimator over the others. In this case, domination depends on the level of significance, α. For small values of α, the SPTE is dominated by the SE, and for larger values of α, the SE is dominated by the SPTE. However, Chiou and Saleh (2002) suggest the value of α to be between 20% and 25%. In this interval of α, the quadratic bias of the SPTE is relatively small for not too small values of Δ^2. However, in practice, the non-centrality parameter is unlikely to be very large (otherwise the credibility of prior information is in serious question) and α is usually preferred to be small.

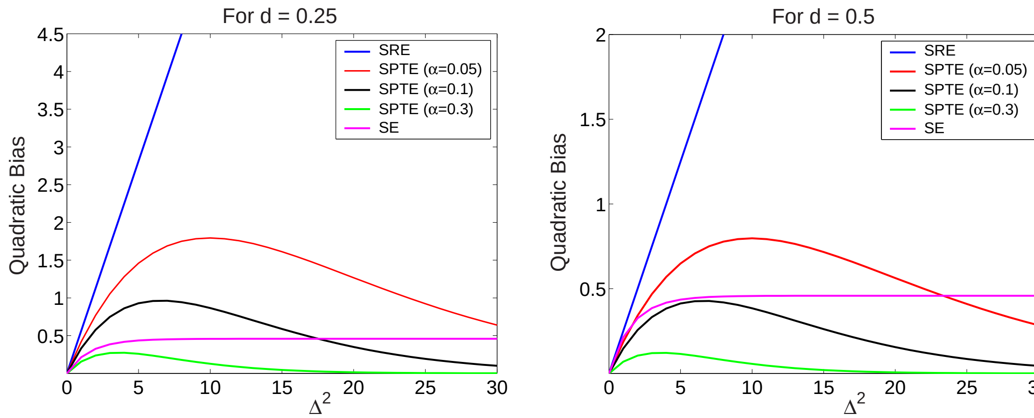

Figure 2.2: The quadratic bias of the SRE, SPTE and SE of the Slope.

The quadratic bias of the SE is relatively stable and is essentially a constant starting from some moderate value of Δ^2. It is unaffected by the choice of the values of d and α. Therefore, the SE may be a better choice among the three biased estimators considered in this study. The behaviour of the quadratic bias functions of the SRE, SPTE and SE of β_1 is displayed in Figure 2.2.

2.5.2 Comparison of the Relative Efficiencies

The relative efficiency (Eff) of an estimator is defined as the ratio of the reciprocal of the mse function. The performances of the estimators are compared on the basis of their relative efficiencies.

2.5.2.1 Comparing SRE with UE

The efficiency of $\hat{\beta}_1^{\text{SRE}}$ relative to $\tilde{\beta}_1$ is denoted by $\text{Eff}[\hat{\beta}_1^{\text{SRE}} : \tilde{\beta}_1]$, and is given by

$$\text{Eff}\left[\hat{\beta}_1^{\text{SRE}} : \tilde{\beta}_1\right] = \left[d^2 + (1-d)^2\Delta^2\right]^{-1}. \tag{2.5.4}$$

Based on (2.5.4), the following properties of the SRE are observed.

- If the non-sampling prior information is correct, that is, $\Delta^2 = 0$, $\text{Eff}\left[\hat{\beta}_1^{\text{SRE}} : \tilde{\beta}_1\right] = d^{-2} > 1$, and hence $\hat{\beta}_1^{\text{SRE}}$ is more efficient than $\tilde{\beta}_1$. Thus, under the null hypothesis, the SRE of β_1 performs better than the UE of β_1.

- If the non-sampling prior information is incorrect (that is, $\Delta^2 > 0$) we study the expression in (2.5.4) as a function of Δ^2 for a fixed value of d. As a function of Δ^2, (2.5.4) is a decreasing function with its maximum value d^{-2} (> 1) at $\Delta^2 = 0$ and minimum value 0 for very large values of Δ^2. The relative efficiency equals 1 at $\Delta^2 = \frac{1+d}{1-d}$. Thus, if $\Delta^2 \in [0, \frac{1+d}{1-d})$, $\hat{\beta}_1^{\text{SRE}}$ is more efficient than $\tilde{\beta}_1$ and outside this interval $\tilde{\beta}_1$ is more efficient than $\hat{\beta}_1^{\text{SRE}}$. For example, if $d = 0.5$, $\hat{\beta}_1^{\text{SRE}}$ is more efficient than $\tilde{\beta}_1$ in $[0, 3)$, while $\tilde{\beta}_1$ is more efficient than $\hat{\beta}_1^{\text{SRE}}$ in $[3, \infty)$. Also, for $d = 0.5$ the maximum efficiency of $\hat{\beta}_1^{\text{SRE}}$ relative to $\tilde{\beta}_1$ is 4.

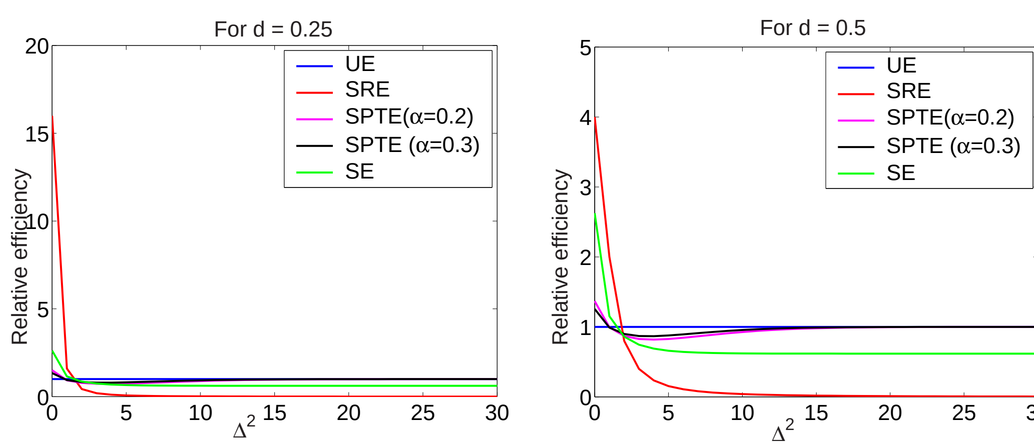

Figure 2.3: The efficiency of the SRE, SPTE and SE relative to the UE for selected values of d and α.

2.5.2.2 Comparing SPTE with UE

The efficiency of the SPTE relative to the UE is given by

$$\begin{aligned} \text{Eff}\left[\hat{\beta}_1^{\text{SPTE}} : \tilde{\beta}_1\right] = \Bigg[1 - (1-d^2)G_{3,\nu}\Big(\frac{1}{3}F_\alpha;\Delta^2\Big) + (1-d)\Delta^2 \\ \times \left\{2G_{3,\nu}\Big(\frac{1}{3}F_\alpha;\Delta^2\Big) - (1+d)G_{5,\nu}\Big(\frac{1}{5}F_\alpha;\Delta^2\Big)\right\}\Bigg]^{-1} \end{aligned} \tag{2.5.5}$$

for any fixed d $(0 \le d \le 1)$ and at a fixed level of significance α. As $F_\alpha \to \infty$, $\text{Eff}\left[\hat{\beta}_1^{\text{SPTE}} : \tilde{\beta}_1\right] \to [1-(1-d^2)+(1-d)^2\Delta^2]^{-1} = [d^2+(1-d)^2\Delta^2]^{-1}$ which is the relative efficiency of $\hat{\beta}_1^{\text{SRE}}$ relative to $\tilde{\beta}_1$. On the other hand, as $F_\alpha \to 0$, $\text{Eff}\left[\hat{\beta}_1^{\text{SPTE}} : \tilde{\beta}_1\right] \to 1$. This means that the efficiency of the SPTE is the same as that of the UE, $\tilde{\beta}_1$.

For varying Δ^2, the following properties of the efficiency of the SPTE relative to the UE are observed.

- Under the null hypothesis, $\Delta^2 = 0$, at which the relative efficiency in (2.5.5) attains its maximum,

$$\left[1-(1-d^2)G_{3,\nu}\Big(\frac{1}{3}F_\alpha;0\Big)\right]^{-1} \ge 1. \tag{2.5.6}$$

- As Δ^2 grows larger than zero, the relative efficiency function monotonically decreases, crossing the 1-line for a Δ^2-value between $\frac{1+d}{2}$ and $\frac{1+d}{1-d}$, and reaching a minimum for some $\Delta^2 = \Delta^2_{\text{min}}$. It then monotonically increases and approaches unity from below. The relative efficiency of the shrinkage preliminary test estimator equals unity whenever

$$\Delta^2_* = \frac{(1+d)}{2-(1+d)\frac{G_{5,\nu}(\frac{1}{5}F_\alpha;\Delta^2)}{G_{3,\nu}(\frac{1}{3}F_\alpha;\Delta^2)}} \tag{2.5.7}$$

where Δ_*^2 lies in the interval $\left(\frac{1+d}{2}, \frac{1+d}{1-d}\right)$. This means that

$$\text{Eff}\left[\hat{\beta}_1^{\text{SPTE}} : \tilde{\beta}_1\right] \begin{matrix}\le\\=\\>\end{matrix} 1 \text{ according as } \Delta_*^2 \begin{matrix}\le\\=\\>\end{matrix} \Delta^2. \tag{2.5.8}$$

- Finally, as $\Delta^2 \to \infty$, $\text{Eff}[\hat{\beta}_1^{\text{SPTE}} : \tilde{\beta}_1] \to 1$.

In conclusion, the shrinkage preliminary test estimator is more efficient than the unrestricted estimator whenever $\Delta^2 < \Delta_*^2$. Otherwise $\tilde{\beta}_1$ is more efficient than SPTE up to some moderate value of Δ^2. For very large values of Δ^2, the efficiency of the SPTE are the same as that of the UE. Figures 2.3 and 2.4 display the efficiency of the SPTE relative to the UE.

2.5.2.3 Comparing SPTE with SRE

As for the efficiency of $\hat{\beta}_1^{\text{SPTE}}$ relative to $\hat{\beta}_1^{\text{SRE}}$ we have

$$\text{Eff}\left[\hat{\beta}_1^{\text{SPTE}} : \hat{\beta}_1^{\text{SRE}}\right] = \left[d^2 + (1-d)^2\Delta^2\right]\left[1 + g(\Delta^2)\right]^{-1} \tag{2.5.9}$$

where

$$g(\Delta^2) = (1-d)\Delta^2\left\{2G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right) - (1+d)G_{5,\nu}\left(\frac{1}{5}F_\alpha; \Delta^2\right)\right\} - (1+d^2)G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right). \tag{2.5.10}$$

- Under the null-hypothesis, $\Delta^2 = 0$, and hence

$$\text{Eff}\left[\hat{\beta}_1^{\text{SPTE}} : \hat{\beta}_1^{\text{SRE}}\right] = d^2\left[1 - (1-d^2)G_{3,\nu}\left(\frac{1}{3}F_\alpha; 0\right)\right]^{-1} \ge d^2. \tag{2.5.11}$$

Combining this result with (2.5.6), we obtain

$$d^2 \le \text{Eff}\left[\hat{\beta}_1^{\text{SPTE}} : \hat{\beta}_1^{\text{SRE}}\right] \le 1 \le \text{Eff}\left[\hat{\beta}_1^{\text{SPTE}} : \tilde{\beta}_1\right]. \tag{2.5.12}$$

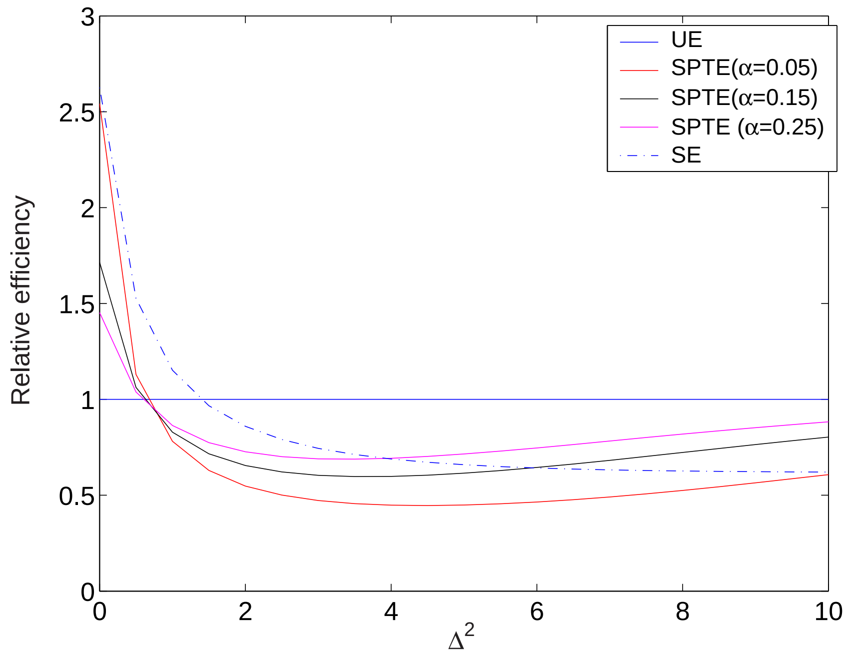

Figure 2.4: The efficiency of the SE and SPTE relative to the UE for $d = 0$ and selected values of α.

- For $\Delta^2 > 0$, we have

$$\text{Eff}\left[\hat{\beta}_1^{\text{SPTE}} : \hat{\beta}_1^{\text{SRE}}\right] \lesseqgtr 1 \ \text{ according as} \tag{2.5.13}$$

$$\Delta^2 \lesseqgtr \frac{1+d}{1-d} \frac{\left\{1 - G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right)\right\}}{\left\{1 - 2G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right) - (1+d)G_{5,\nu}\left(\frac{1}{5}F_\alpha; \Delta^2\right)\right\}}. \tag{2.5.14}$$

- Finally, as $\Delta^2 \to \infty$, $\text{Eff}\left[\hat{\beta}_1^{\text{SPTE}}; \hat{\beta}_1^{\text{SRE}}\right] \to 0$. Thus, except for a small interval around 0, $\hat{\beta}_1^{\text{SPTE}}$ is more efficient than $\hat{\beta}_1^{\text{SRE}}$.

2.5.2.4 Comparing SE with UE

The efficiency of $\hat{\beta}_1^{\text{SE}}$ relative to $\tilde{\beta}_1$ is given by

$$\text{Eff}\left[\hat{\beta}_1^{\text{SE}} : \tilde{\beta}_1\right] = \left[1 - \frac{2}{\pi} K_n^2 \left\{2e^{-\Delta^2/2} - 1\right\}\right]^{-1}. \tag{2.5.15}$$

- Under the null-hypothesis $\Delta^2 = 0$, and hence

$$\text{Eff}\left[\hat{\beta}_1^{\text{SE}} : \tilde{\beta}_1\right] = \left[1 - \frac{2}{\pi}K_n^2\right]^{-1} \geq 1. \tag{2.5.16}$$

- As Δ^2 grows larger than zero, $\text{Eff}\left[\hat{\beta}_1^{\text{SE}} : \tilde{\beta}_1\right]$ decreases monotonically from $\left[1 - \frac{2}{\pi}K_n^2\right]^{-1}$ at $\Delta^2 = 0$, crossing unity at $\Delta^2 = \ln 4$, and approaching the minimum value $\left[1 + \frac{2}{\pi}K_n^2\right]^{-1}$ as $\Delta^2 \to \infty$. Thus, the loss of efficiency of $\hat{\beta}_1^{\text{SE}}$ relative to $\tilde{\beta}_1$ is $1 - \left[1 + \frac{2}{\pi}K_n^2\right]^{-1}$, while the gain in efficiency is $\left[1 - \frac{2}{\pi}K_n^2\right]^{-1}$, which is achieved at $\Delta^2 = 0$. Hence, for $\Delta^2 < \ln 4$, $\hat{\beta}_1^{\text{SE}}$ performs better than $\tilde{\beta}_1$. Otherwise $\tilde{\beta}_1$ performs better $\hat{\beta}_1^{\text{SE}}$.

- Finally, as $\Delta^2 \to \infty$ the efficiency of the SPTE relative to the UE approaches one and that of the SE relative to the UE approaches $\left[1 + \frac{2}{\pi}K_n^2\right]^{-1}$. Figures 2.3 and 2.4 display the efficiency of the SE relative to the UE.

2.5.2.5 Comparing SE with SRE

The efficiency of $\hat{\beta}_1^{\text{SE}}$ relative to $\hat{\beta}_1^{\text{SRE}}$ is given by

$$\text{Eff}\left[\hat{\beta}_1^{\text{SE}} : \hat{\beta}_1^{\text{SRE}}\right] = \left[d^2 + (1-d)^2\Delta^2\right]\left[1 - \frac{2}{\pi}K_n^2\left\{2e^{-\Delta^2/2} - 1\right\}\right]^{-1}. \tag{2.5.17}$$

- Under the null-hypothesis $\Delta^2 = 0$, and hence

$$\text{Eff}\left[\hat{\beta}_1^{\text{SE}} : \hat{\beta}_1^{\text{SRE}}\right] = d^2\left[1 - \frac{2}{\pi}K_n^2\right]^{-1} \lesseqqgtr 1 \tag{2.5.18}$$

depending on

$$d^2 \lesseqqgtr \left(1 - \frac{2}{\pi}K_n^2\right)^{-1}. \tag{2.5.19}$$

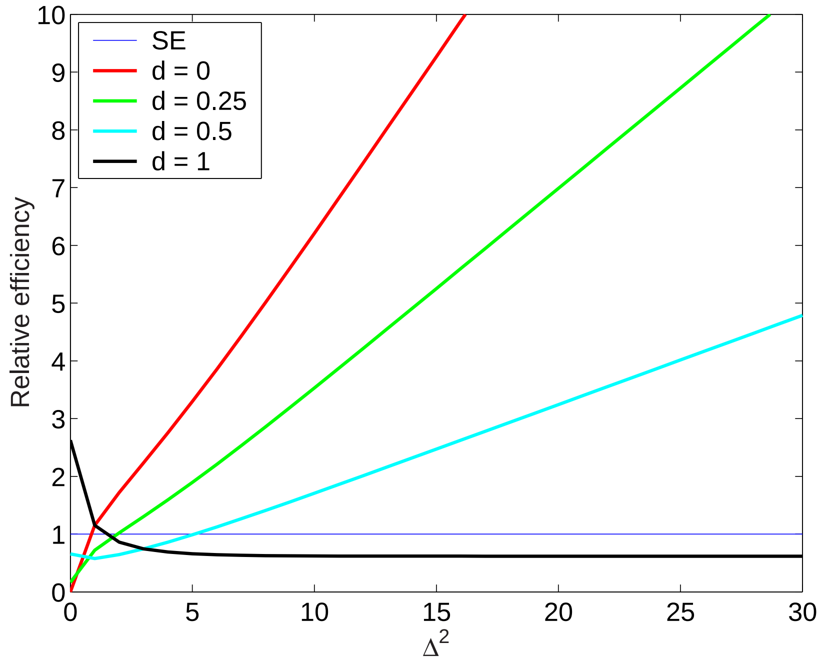

Figure 2.5: The efficiency of the SE relative to the SRE for $n = 20$ and selected values of d.

- As Δ^2 grows larger than zero, the efficiency of the SE relative to the SRE increases or decreases depending on the values of d and n.

- Finally, as Δ^2 approaches a very large value, the relative efficiency increases unboundedly, except for $d = 1$. For $d = 1$ and very large value of Δ^2, the relative efficiency of the SE relative to the RE is some constant, less than one.

Figure 2.5 displays the efficiency of the SE relative to the RE for a range of values of Δ^2.

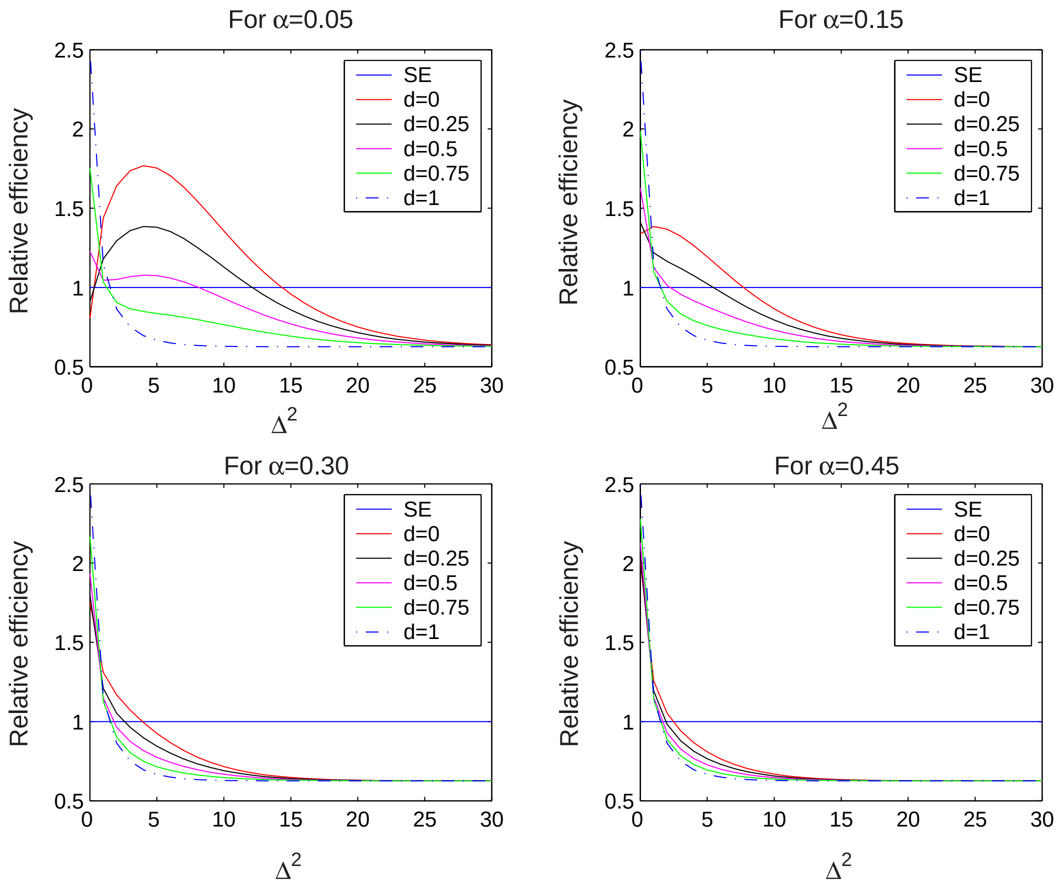

Figure 2.6: The efficiency of the SE relative to the SPTE for selected values of d.

2.5.2.6 Comparing SE with SPTE

In Figure 2.6, the maximum efficiency of the SE relative to the SPTE is attained for $\Delta^2 = 0$ and $d = 1$, regardless of the value of α. At $\Delta^2 = 0$, as d decreases, the relative efficiency of the SE also decreases, and it moves down to one for very small values of d. Starting from some moderate value of Δ^2, the relative efficiency of the SE becomes less than one and converges to a stable value, below one, as $\Delta^2 \to \infty$. Except for $\Delta^2 = 0$ and near 0, the relative efficiency of SE is always higher for smaller values of d than for larger values of

d. The difference between the relative efficiencies of the SE for different values of d is higher for lower value of α than for higher values of α. As α increases this difference decreases. Moreover, as α increases the relative efficiency of the SE also increases for $\Delta^2 = 0$ or near 0.

2.6 Concluding Remarks

The UE is based on the sample data alone, and it is the only unbiased estimator among the four estimators considered in this chapter. The introduction of non-sample prior information in the estimation process causes the estimators to be biased. However, the biased estimators perform better than the unbiased estimator when they are judged based on the mse criterion. The performance of the biased estimators depend on the value of Δ. In the case of the SPTE, the performance also depends on the value of the level of significance. Under the null hypothesis, the departure parameter is 0 and the SE dominates all other estimators if α is not too high. As α increases, the performance of the SPTE improves when Δ is not too close to zero. At a lower level of significance, the SE outperforms the SPTE over a wide range of values of Δ. When the value of Δ is not far from 0, the SE always outperforms the SPTE and SRE. Therefore, in practice if a researcher postulates a value of β_1 from prior knowledge or experience that is not too far from its true value, the SE would be the best choice as an *improved estimator* of the slope.

2.A Appendix

- The following MATLAB codes are used for producing Figure 2.2.

```
d=0.25; n=5; v=n-2; D=0:1:30; x=(1-d).^2; B2=x.*D;
plot(D,B2); hold on
G3=ncfcdf(finv(.95,3,v)./3,3,v, D); B3=x.*D.*G3;
plot(D,B3,'r')
G3=ncfcdf(finv(.9,3,v)./3,3,v, D); B4=x.*D.*G3;
plot(D,B4,'k')
G3=ncfcdf(finv(.7,3,v)./3,3,v, D); B6=x.*D.*G3;
plot(D,B6,'m')
K=sqrt(2./(n-2)).*gamma((n-1)./2)./gamma((n-2)./2);
c=sqrt(2).*1./sqrt(pi).*K.*exp(-D./2);
F=normcdf(sqrt(D), 0, 1);
B7=2./pi.*K.^4.*(2.*F-1).^2;
plot(D,B7,'g')
legend('SRE', 'SPTE (\alpha=0.05)', 'SPTE (\alpha=0.1)',
'SPTE(\alpha=0.3)','SE',1)
xlabel('\Delta^2');
ylabel('Quadratic Bias');
title('For d = 0.25')
```

- The following MATLAB codes are used for producing Figure 2.4.

```
d=0; D=0:0.5:10; q=ones(1,length(D));
plot(D,q)
```

```
hold on
n=20; v=n-2;
G3=ncfcdf(finv(.95,3,v)./3,3,v, D);
G5=ncfcdf(finv(.95,5,v)./5,5,v, D); x=1-d.^2; y=1-d; z=1+d;
R2=1./(1 - x.*G3 + y.*D.*(2.*G3 - z.*G5));
plot(D, R2, 'r')
G3=ncfcdf(finv(.85,3,v)./3,3,v, D);
G5=ncfcdf(finv(.85,5,v)./5,5,v, D);
R3=1./(1-x.*G3 +y.*D.*(2.*G3-z.*G5));
plot(D, R3, 'k')
G3=ncfcdf(finv(.75,3,v)./3,3,v, D);
G5=ncfcdf(finv(.75,5,v)./5,5,v, D);
R4=1./(1 - x.*G3 +y.*D.*(2.*G3- z.*G5));
plot(D, R4, 'm')
k=sqrt(2./(n-2)).*gamma((n-1)./2)./gamma((n-2)./2);
R4=1./(1-2.*(1./pi).*k.^2.*(2.*exp(-D./2)-1));
plot(D, R4, 'b-.')
legend ('UE', 'PTE(\alpha=0.05)', 'PTE(\alpha=0.15)','PTE
(\alpha=0.25)', 'SE',1) xlabel('\Delta^2');
ylabel('Relative efficiency')
```

Chapter 3

Estimation of the Intercept Parameter of Simple Linear Regression Model

3.1 Introduction

In Chapter 2 we studied the performances of the three improved estimators, SRE, SPTE and SE, of the slope parameter β_1 of the simple linear regression model (2.1.1). The focus of this chapter is on the estimation of the intercept parameter β_0 assuming that *uncertain non-sample prior information* on the value of the slope parameter β_1 is available, either from previous study or from practical experience of a researcher or expert. It is well known that the estimation of the intercept parameter involves that of the slope parameter. As a result, an estimator of β_1 is required in the definition of the estimator of β_0.

Let the non-sample prior information about the value of β_1 be expressed in the form of the null hypothesis in (2.1.2). We wish to combine both the sample data and the uncertain non-sample prior information on the value of β_1 in estimating the intercept parameter β_0. Similar to the estimation of the slope parameter, we introduce a coefficient of distrust d $(0 \leq d \leq 1)$ for the non-sample prior information that represents the degree of distrust in the null hypothesis in (2.1.2). First we define the UE of the unknown intercept β_0 from the likelihood function of the parameter of the model based on a random sample. Based on the unrestricted and restricted (by the null hypothesis) estimators of σ^2, we define the likelihood ratio test for testing the null hypothesis. Then we use the coefficient of distrust, as well as the sample and non-sample prior information, to define the shrinkage restricted estimator. Using the preliminary test approach we define the shrinkage preliminary test and shrinkage estimators of the unknown population intercept β_0. The coefficient of distrust d is introduced to both the SPTE and SE of β_0.

To compare the performances of the estimators of the intercept parameter, their bias, mean square error and relative efficiency functions have been analysed both analytically and graphically. The analyses reveal that although there is no uniformly superior estimator with respect to both unbiasedness and mse criteria, the shrinkage estimator dominates the other two biased estimators if the non-sample prior information regarding the value of the slope parameter is not too far from its true value. In practice, the non-sample prior information is usually obtained from past experience or expert knowledge and hence it is expected that such information will not be too far from the true value.

The layout of this chapter is as follows. Some preliminaries are provided in Section 3.2. The three alternative estimators of the intercept parameter are defined in Section 3.3. Expressions of the bias and mse functions of the estimators are derived in Section 3.4. A comparative study of the quadratic biases and relative efficiencies of the estimators are included in Section 3.5. Some concluding remarks are presented in Section 3.6. Selected MATLAB codes, used for producing graphs, are presented in Appendix 3.A.

3.2 Some Preliminaries

Based on the random sample observations (x_i, y_i) for $i = 1, 2, \ldots, n$, the UE of the intercept β_0 is given by

$$\tilde{\beta}_0 = \bar{y} - \tilde{\beta}_1 \bar{x} \tag{3.2.1}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $\tilde{\beta}_1$ is the UE of the slope β_1, given by the expression (2.2.4).

It is well known that for the normal model, the sampling distribution of the UE of β_0 is normal with mean and variance given by

$$\mathrm{E}\left[\tilde{\beta}_0\right] = \beta_0 \tag{3.2.2}$$

and

$$\mathrm{Var}\left[\tilde{\beta}_0\right] = \sigma^2 H \tag{3.2.3}$$

respectively, in which $H = \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)$ and $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2$. Therefore, the bias and mse (variance) functions of the unrestricted estimator of β_0 are given by

$$\mathrm{B}_1\left[\tilde{\beta}_0\right] = \mathrm{E}\left[\tilde{\beta}_0 - \beta_0\right] = 0 \tag{3.2.4}$$

and

$$\mathrm{M}_1\left[\tilde{\beta}_0\right] = \sigma^2 H \tag{3.2.5}$$

respectively.

The above bias and mean square error functions are compared with those of the SRE, SPTE and SE to investigate the relative performance of the estimators under various conditions.

3.3 Proposed Improved Estimators of the Intercept

Consider a convex combination of $\tilde{\beta}_0 = \bar{y} - \tilde{\beta}_1\bar{x}$ (the mle of β_0) and $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$ (the mle of β_0 under the null hypothesis in (2.1.2)), as

$$\hat{\beta}_0^{\text{SRE}} = d\tilde{\beta}_0 + (1-d)\hat{\beta}_0. \tag{3.3.1}$$

The estimator $\hat{\beta}_0^{\text{SRE}}$ is called the *shrinkage restricted estimator* (SRE) of the intercept parameter β_0, where d is the *degree of distrust* on the null hypothesis. Here, $d = 0$ means that there is *no distrust* in H_0, and then we get $\hat{\beta}_0^{\text{SRE}} = \hat{\beta}_0$ (complete reliance on the prior information), while $d = 1$ means that there is *complete distrust* on H_0, and we then get $\hat{\beta}_0^{\text{SRE}} = \tilde{\beta}_0$ (exclusive sample information based estimator). If $0 < d < 1$ (that is, the degree of distrust is an intermediate value between 0 and 1) then the SRE of β_0 yields an interpolated value between $\hat{\beta}_0$ and $\tilde{\beta}_0$ given by (3.3.1).

Following Khan and Saleh (2001), the shrinkage preliminary test estimator of the intercept parameter β_0 is defined as

$$\begin{aligned}
\hat{\beta}_0^{\text{SPTE}} &= \hat{\beta}_0^{\text{SRE}}\, I(F < F_\alpha) + \tilde{\beta}_0\, I\big(F \geq F_\alpha\big) \\
&= \hat{\beta}_0^{\text{SRE}} I(F < F_\alpha) + \tilde{\beta}_0\{1 - I(F < F_\alpha)\} \\
&= \tilde{\beta}_0 + \hat{\beta}_0^{\text{SRE}} I(F < F_\alpha) - \tilde{\beta}_0\, I(F < F_\alpha)
\end{aligned}$$

$$
\begin{aligned}
&= \tilde{\beta}_0 + \{d\tilde{\beta}_0 + (1-d)\hat{\beta}_0\}\, I(F < F_\alpha) - \tilde{\beta}_0\, I(F < F_\alpha) \\
&= \tilde{\beta}_0 + (1-d)\hat{\beta}_0\, I(F < F_\alpha) - (1-d)\tilde{\beta}_0\, I(F < F_\alpha) \\
&= \tilde{\beta}_0 - (1-d)(\tilde{\beta}_0 - \hat{\beta}_0)\, I(F < F_\alpha) \\
&= \tilde{\beta}_0 - (1-d)\,(\tilde{\beta}_0 - \hat{\beta}_0)\, I(F < F_\alpha) \qquad (3.3.2)
\end{aligned}
$$

where $I(A)$ is an indicator function of the set A, and F_α is the $(1-\alpha)^{\text{th}}$ upper quantile of the central F distribution with $(1,\, \nu)$ degrees of freedom. For $d = 0$, the shrinkage preliminary test estimator becomes

$$
\hat{\beta}_0^{\text{SPTE}} = \tilde{\beta}_0 - (\tilde{\beta}_0 - \hat{\beta}_0)\, I(F < F_\alpha), \qquad (3.3.3)
$$

the preliminary test estimator of β_0.

As shown in Chapter 2, the performance of the SPTE depends on the choice of the level of significance α of the preliminary test. Therefore, we define a shrinkage estimator (SE) that does not depend on α. The SE of β_0 is

$$
\begin{aligned}
\hat{\beta}_0^{\text{SE}} &= \hat{\beta}_0^{\text{SRE}} + (1 - c|t|^{-1})(\tilde{\beta}_0 - \hat{\beta}_0^{\text{SRE}}) \\
&= \tilde{\beta}_0 - c|t|^{-1}(\tilde{\beta}_0 - \hat{\beta}_0^{\text{SRE}}) \\
&= \tilde{\beta}_0 - c|t|^{-1}(1-d)(\tilde{\beta}_0 - \hat{\beta}_0) \\
&= \tilde{\beta}_0 + (1-d)c|t|^{-1}\bar{x}(\tilde{\beta}_1 - \hat{\beta}_1) \\
&= \tilde{\beta}_0 + (1-d)\frac{cS_n\bar{x}}{\sqrt{S_{xx}}|\tilde{\beta}_1 - \beta_{10}|}(\tilde{\beta}_1 - \hat{\beta}_1) \qquad (3.3.4)
\end{aligned}
$$

where t is the test statistic to test the null hypothesis and c is the shrinkage constant, a function of n. Unlike the Stein-type SE, the SE in (3.3.4) is admissible over the UE for one-dimensional populations (see sub-subsection 3.5.2.3).

3.4 Some Statistical Properties

In this section, the bias and mean square error (mse) functions of the SRE, SPTE and SE are derived. Also, some important features of these functions are discussed.

3.4.1 The Bias and MSE of the SRE

Theorem 3.47 *The bias function of the shrinkage restricted estimator of the intercept parameter β_0 of the simple linear regression model is given by*

$$\mathrm{B}_2\left[\hat{\beta}_0^{\mathrm{SRE}}\right] = (1-d)\frac{\bar{x}\sigma}{\sqrt{S_{xx}}}\Delta \tag{3.4.1}$$

where $\Delta = S_{xx}^{1/2}(\beta_1 - \beta_{10})\sigma^{-1}$.

Proof. By definition, the bias function of the SRE of β_0 is

$$\begin{aligned}
\mathrm{B}_2\left[\hat{\beta}_0^{\mathrm{SRE}}\right] &= \mathrm{E}\left[\hat{\beta}_0^{\mathrm{SRE}} - \beta_0\right] \\
&= \mathrm{E}\left[d\tilde{\beta}_0 - \beta_0 + (1-d)\hat{\beta}_0\right] \\
&= d\beta_0 - \beta_0 + (1-d)\hat{\beta}_0 \\
&= -(1-d)\beta_0 + (1-d)\hat{\beta}_0 \\
&= -(1-d)(\bar{y} - \beta_1\bar{x}) + (1-d)(\bar{y} - \beta_{10}\bar{x}) \\
&= (1-d)\bar{x}(\beta_1 - \beta_{10}) \\
&= (1-d)\frac{\bar{x}\sigma}{\sqrt{S_{xx}}}\Delta\,.
\end{aligned} \tag{3.4.2}$$

This completes the proof of the theorem.

Theorem 3.48 *The mean square error function of the shrinkage restricted estimator of the intercept parameter β_0 of the simple linear regression model* *is given by*

$$\mathrm{M}_2\left[\hat{\beta}_0^{\mathrm{SRE}}\right] = \sigma^2\left[d^2 H + \frac{(1-d)^2\bar{x}^2\Delta^2}{S_{xx}}\right] \tag{3.4.3}$$

where Δ^2 *is the non-centrality parameter of non-central* F *distribution, a function of the distance between the true value of* β_1 *and that under the null hypothesis.*

Proof. By definition, the mse function of the SRE of β_0 is

$$\begin{aligned}
\mathrm{M}_2\left[\hat{\beta}_0^{\mathrm{SRE}}\right] &= \mathrm{E}\left[\hat{\beta}_0^{\mathrm{SRE}} - \beta_0\right]^2 \\
&= \mathrm{E}\left[d\tilde{\beta}_0 - \beta_0 + (1-d)\hat{\beta}_0\right]^2 \\
&= \mathrm{E}\left[d\tilde{\beta}_0 - d\beta_0 - (1-d)\beta_0 + (1-d)\hat{\beta}_0\right]^2 \\
&= \mathrm{E}\left[d(\tilde{\beta}_0 - \beta_0) - (1-d)(\beta_0 - \hat{\beta}_0)\right]^2 \\
&= d^2\,\mathrm{E}\left[\tilde{\beta}_0 - \beta_0\right]^2 - (1-d)^2\,\mathrm{E}\left[\beta_0 - \hat{\beta}_0\right]^2 \\
&= d^2\sigma^2 H + (1-d)^2\left(\bar{y} - \beta_1\bar{x} - \bar{y} + \beta_{10}\bar{x}\right)^2 \\
&= d^2\sigma^2 H + (1-d)^2\bar{x}^2\left(\beta_1 - \beta_{10}\right)^2 \\
&= d^2\sigma^2 H + \frac{(1-d)^2\bar{x}^2\Delta^2}{S_{xx}} \\
&= \sigma^2\left[d^2 H + \frac{(1-d)^2\bar{x}^2\Delta^2}{S_{xx}}\right].
\end{aligned} \tag{3.4.4}$$

This completes the proof of the theorem.

Under the null hypothesis, the value of the parameter Δ^2 is 0, while under the alternative hypothesis it takes a positive value. The performances of the estimators SRE, SPTE and SE change with the change in the value of Δ^2. We investigate this feature in a greater detail in the forthcoming sections.

3.4.2 The Bias and MSE of the SPTE

Theorem 3.49 *The bias function of the shrinkage preliminary test estimator of the intercept parameter β_0 of the simple linear regression model is given by*

$$\mathrm{B}_3\left[\hat{\beta}_0^{\mathrm{SPTE}}\right] = (1-d)\frac{\bar{x}\sigma}{\sqrt{S_{xx}}}\Delta\, G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right) \tag{3.4.5}$$

where $G_{a,b}(\cdot;\Delta^2)$ is the cumulative distribution function of the non-central F distribution with (a,b) degrees of freedom and non-centrality parameter Δ^2.

Proof. By definition, the bias function of the SPTE of β_0 is

$$\begin{aligned}
\mathrm{B}_3\left[\hat{\beta}_0^{\mathrm{SPTE}}\right] &= \mathrm{E}\left[\hat{\beta}_0^{\mathrm{SPTE}} - \beta_0\right] \\
&= \mathrm{E}\left[\tilde{\beta}_0 - \beta_0 - (1-d)\,(\tilde{\beta}_0 - \hat{\beta}_0)\, I(F < F_\alpha)\right] \\
&= -(1-d)\,\mathrm{E}\left[(\tilde{\beta}_0 - \hat{\beta}_0)\, I(F < F_\alpha)\right] \\
&= -(1-d)\,\mathrm{E}\left[(\bar{y} - \tilde{\beta}_1\bar{x} - \bar{y} + \beta_{10}\bar{x})\, I(F < F_\alpha)\right] \\
&= (1-d)\bar{x}\,\mathrm{E}\left[(\tilde{\beta}_1 - \beta_{10})\, I(F < F_\alpha)\right] \\
&= (1-d)\frac{\bar{x}\sigma}{\sqrt{S_{xx}}}\,\mathrm{E}\left[\frac{\sqrt{S_{xx}}(\tilde{\beta}_1 - \beta_{10})}{\sigma}\, I\left(\frac{\sqrt{S_{xx}}(\tilde{\beta}_1 - \beta_{10})^2}{S_n^2} < F_\alpha\right)\right].
\end{aligned} \tag{3.4.6}$$

Note $Z = \sqrt{S_{xx}}(\tilde{\beta}_1 - \beta_{10})\sigma^{-1}$ is distributed as $N(\Delta,\, 1)$, and $\nu\sigma^{-2}S_n^2$ is distributed as a central chi-square variable with ν degrees of freedom.

Applying Theorem 1, Appendix B2, Judge and Bock (1978), to the right hand side of (3.4.6), the bias function of the SPTE of β_0 is obtained as

$$\mathrm{B}_3\left[\hat{\beta}_0^{\mathrm{SPTE}}\right] = (1-d)\frac{\bar{x}\sigma}{\sqrt{S_{xx}}}\Delta\, G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right). \tag{3.4.7}$$

This completes the proof of the theorem.

Theorem 3.410 *The mean square error function of the shrinkage preliminary test estimator of the intercept parameter* β_0 *of the simple linear regression model is given by*

$$\mathrm{M}_3\left[\hat{\beta}_0^{\mathrm{SPTE}}\right] = \sigma^2 H + \frac{\bar{x}^2\sigma^2}{S_{xx}}\left[\Delta^2\left\{2(1-d)\,G_{3,v}\left(\frac{1}{3}F_\alpha;\Delta^2\right) - (1-d^2)\right.\right.$$
$$\left.\left.\times G_{5,v}\left(\frac{1}{5}F_\alpha;\Delta^2\right)\right\} - (1-d^2)\,G_{3,v}\left(\frac{1}{3}F_\alpha;\Delta^2\right)\right] \qquad (3.4.8)$$

where $G_{a,b}(\cdot;\,\Delta^2)$ *is the c.d.f. of the non-central* F *distribution with* $(a,\,b)$ *degrees of freedom and non-centrality parameter* Δ^2*.*

Proof. By definition, the mse function of the SPTE of β_0 is

$$\begin{aligned}
\mathrm{M}_3\left[\hat{\beta}_0^{\mathrm{SPTE}}\right] &= \mathrm{E}\left[\hat{\beta}_0^{\mathrm{SPTE}} - \beta_0\right]^2 \\
&= \mathrm{E}\left[(\tilde{\beta}_0 - \beta_0) - (1-d)\,(\tilde{\beta}_0 - \hat{\beta}_0)\,I(F < F_\alpha)\right] \\
&= \mathrm{E}\left[\tilde{\beta}_0 - \beta_0\right]^2 + (1-d)^2\mathrm{E}\left[(\tilde{\beta}_0 - \hat{\beta}_0)^2\,I(F < F_\alpha)\right] \\
&\quad - 2(1-d)\,\mathrm{E}\left[(\tilde{\beta}_0 - \beta_0)(\tilde{\beta}_0 - \hat{\beta}_0)\,I(F < F_\alpha)\right] \\
&= \sigma^2 H + (1-d)^2\mathrm{E}\left[(\tilde{\beta}_0 - \hat{\beta}_0)^2\,I(F < F_\alpha)\right] \\
&\quad - 2(1-d)\,\mathrm{E}\left[(\tilde{\beta}_0 - \beta_0)(\tilde{\beta}_0 - \hat{\beta}_0)\,I(F < F_\alpha)\right]. \qquad (3.4.9)
\end{aligned}$$

The second term of the right hand side of (3.4.9) is

$$\begin{aligned}
(1-d)^2\mathrm{E}\left[(\tilde{\beta}_0 - \hat{\beta}_0)^2\,I(F < F_\alpha)\right] &= (1-d)^2\bar{x}^2\mathrm{E}\left[(\tilde{\beta}_1 - \beta_{10})^2\,I(F < F_\alpha)\right] \\
&= (1-d)^2\frac{\bar{x}^2\sigma^2}{S_{xx}}\mathrm{E}\left[\frac{S_{xx}(\tilde{\beta}_1 - \beta_{10})^2}{\sigma^2}\right. \\
&\quad \left.\times I\left(\frac{S_{xx}(\tilde{\beta}_1 - \beta_{10})^2}{S_n^2} < F_\alpha\right)\right]. \qquad (3.4.10)
\end{aligned}$$

Applying Theorem 3, Appendix B2, Judge and Bock (1978), to the right hand side of (3.4.10), we get

$$(1-d)^2\mathrm{E}\Big[(\tilde{\beta}_0-\hat{\beta}_0)^2\, I(F<F_\alpha)\Big] = (1-d)^2\frac{\bar{x}^2\sigma^2}{S_{xx}}\, G_{3,\nu}\left(\frac{1}{3}F_\alpha;\Delta^2\right) + (1-d)^2\frac{\bar{x}^2\sigma^2}{S_{xx}}\Delta^2\, G_{5,\nu}\left(\frac{1}{5}F_\alpha;\Delta^2\right). \tag{3.4.11}$$

The last term of the right hand side of (3.4.9) is

$$\begin{aligned}
&-2(1-d)\,\mathrm{E}\Big[(\tilde{\beta}_0-\beta_0)(\tilde{\beta}_0-\hat{\beta}_0)\, I(F<F_\alpha)\Big]\\
&= -2(1-d)\bar{x}^2\,\mathrm{E}\Big[(\tilde{\beta}_1-\beta_1)(\tilde{\beta}_1-\beta_{10})\, I(F<F_\alpha)\Big]\\
&= -2(1-d)\bar{x}^2\,\mathrm{E}\Big[\{(\tilde{\beta}_1-\beta_{10})-(\beta_1-\beta_{10})\}(\tilde{\beta}_1-\beta_{10})\, I(F<F_\alpha)\Big]\\
&= -2(1-d)\bar{x}^2\,\Big\{\mathrm{E}\Big[(\tilde{\beta}_1-\beta_{10})^2\, I(F<F_\alpha)\Big]\\
&\quad -\frac{\sigma}{\sqrt{S_{xx}}}\Delta\,\mathrm{E}\Big[(\tilde{\beta}_1-\beta_{10})\, I(F<F_\alpha)\Big]\Big\}.
\end{aligned} \tag{3.4.12}$$

Applying Theorems 1 and 3, Appendix B2, Judge and Bock (1978), to the right hand side of (3.4.12), we get

$$\begin{aligned}
&2(1-d)\,\mathrm{E}\Big[(\tilde{\beta}_0-\beta_0)(\tilde{\beta}_0-\hat{\beta}_0)\, I(F<F_\alpha)\Big]\\
&= -2(1-d)\frac{\bar{x}^2\sigma^2}{S_{xx}}\left[G_{3,\nu}\left(\frac{1}{3}F_\alpha;\Delta^2\right)+\Delta^2\, G_{5,\nu}\left(\frac{1}{5}F_\alpha;\Delta^2\right)\right.\\
&\quad \left. -\Delta^2\, G_{3,\nu}\left(\frac{1}{3}F_\alpha;\Delta^2\right)\right].
\end{aligned} \tag{3.4.13}$$

Collecting all terms on the right hand side of (3.4.9), the mse function of the SPTE of β_0 is obtained as

$$\begin{aligned}
\mathrm{M}_3\Big[\hat{\beta}_0^{\mathrm{SPTE}}\Big] &= \sigma^2 H + \frac{\bar{x}^2\sigma^2}{S_{xx}}\left[\Delta^2\left\{2(1-d)\, G_{3,v}\left(\frac{1}{3}F_\alpha;\Delta^2\right)-(1-d^2)\right.\right.\\
&\quad \left.\left.\times G_{5,v}\left(\frac{1}{5}F_\alpha;\Delta^2\right)\right\}-(1-d^2)\, G_{3,v}\left(\frac{1}{3}F_\alpha;\Delta^2\right)\right].
\end{aligned} \tag{3.4.14}$$

This completes the proof of the theorem.

3.4.2.1 Some Properties of the MSE of the SPTE

Here we provide some analytical discussion of the mean square error function of the SPTE for varying values of Δ^2.

- Under the null hypothesis, $\Delta^2 = 0$, in which case the mse of $\hat{\beta}_0^{\text{SPTE}}$,

$$\mathrm{M}_3\left[\hat{\beta}_0^{\text{SPTE}}\right] = \sigma^2 H - \frac{\sigma^2\bar{x}^2}{S_{xx}}(1-d^2)G_{3,v}\left(\frac{1}{3}F_\alpha;\Delta\right) < \sigma^2 H, \qquad (3.4.15)$$

the mse of $\tilde{\beta}_0$. Thus, when $\Delta^2 = 0$ the SPTE of β_0 performs better than the UE.

As $\alpha \to 0$, $G_{3,\nu}\left(\frac{1}{3}F_\alpha;0\right) \to 1$. Therefore,

$$\mathrm{M}_3\left[\hat{\beta}_0^{\text{SPTE}}\right] = \sigma^2 H - \frac{\bar{x}^2\sigma^2}{S_{xx}}(1-d^2)\, G_{3,v}\left(\frac{1}{3}F_\alpha;\Delta^2\right) \to \sigma^2 H - \frac{\bar{x}^2\sigma^2}{S_{xx}}(1-d^2)$$

which is the mse of $\hat{\beta}_0^{\text{SRE}}$ at $\Delta = 0$. On the other hand, as $F_\alpha \to 0$, $G_{3,\nu}\left(\frac{1}{3}F_\alpha;0\right) \to 0$. Then

$$\mathrm{M}_3\left[\hat{\beta}_0^{\text{SPTE}}\right] = \sigma^2 H - \frac{\bar{x}^2\sigma^2}{S_{xx}}(1-d^2)\, G_{3,v}\left(\frac{1}{3}F_\alpha;\Delta^2\right) \to \sigma^2 H, \qquad (3.4.17)$$

the mse of $\tilde{\beta}_0$.

- As $\Delta^2 \to \infty$, $G_{m,\nu}\left(\frac{1}{m}F_\alpha;\Delta^2\right) \to 0$, and hence from (3.4.8), the mse of the SPTE tends towards that of the UE.

- Since $G_{3,\nu}\left(\frac{1}{3}F_\alpha;\Delta^2\right)$ is always greater than $G_{5,\nu}\left(\frac{1}{5}F_\alpha;\Delta^2\right)$ for any value of α, replacing $G_{5,\nu}\left(\frac{1}{5}F_\alpha;\Delta^2\right)$ by $G_{3,\nu}\left(\frac{1}{3}F_\alpha;\Delta^2\right)$, the expression of the mse function of the SPTE of β_0 in (3.4.8) yields

$$\begin{aligned}\mathrm{M}_3\left[\hat{\beta}_0^{\text{SPTE}}\right] &\geq \sigma^2 H + \frac{\bar{x}^2\sigma^2}{S_{xx}}\left[(1-d)\, G_{3,\nu}\left(\frac{1}{3}F_\alpha;\Delta^2\right)\left\{\Delta^2(1-d)-(1+d)\right\}\right] \\ &\geq \sigma^2 H, \quad \text{whenever } \Delta^2 > \frac{1+d}{1-d}.\end{aligned}$$

On the other hand, (3.4.8) may be rewritten as

$$\mathrm{M}_3\left[\hat{\beta}_0^{\mathrm{SPTE}}\right] = \sigma^2 H + \frac{\bar{x}^2\sigma^2}{S_{xx}}\Big[1 + (1-d)G_{3,\nu}\Big(\frac{1}{3}F_\alpha;\Delta^2\Big)\big\{2\Delta^2 - (1+d)\big\} - (1-d^2)G_{5,\nu}\Big(\frac{1}{5}F_\alpha;\Delta^2\Big)\Big] \le \sigma^2 H \qquad (3.4.18)$$

whenever $\Delta^2 < \frac{1+d}{2}$. Therefore, the mse of $\hat{\beta}_0^{\mathrm{SPTE}}$ as a function of Δ^2 crosses the constant line $\mathrm{M}\left[\tilde{\beta}_0\right] = \sigma^2 H$ in the interval $\left(\frac{1+d}{2}, \frac{1+d}{1-d}\right)$.

A general picture of the mse graph may be described as follows:

The mse function of the SPTE has minimum value at $\Delta^2 = 0$. As Δ^2 grows larger, the function increases monotonically, crossing the constant line $\sigma^2 H$ in the interval $\left(\frac{1+d}{2}, \frac{1+d}{1-d}\right)$, and reaching the maximum in the interval $\left(\frac{1+d}{1-d}, \infty\right)$ before monotonically decreasing towards $\sigma^2 H$ as $\Delta^2 \to \infty$.

3.4.2.2 Determination of Optimum α for the SPTE

Clearly, the mse function, and hence the efficiency of the shrinkage preliminary test estimator relative to the unrestricted estimator, depends on the level of significance α of the test and the non-centrality parameter Δ^2.

Let the efficiency of the SPTE relative to the UE be denoted by $\mathrm{Eff}(\alpha;\Delta^2)$ which is given by

$$\mathrm{Eff}(\alpha;\Delta^2) = [1 + g_1(\Delta^2)]^{-1} \qquad (3.4.19)$$

where

$$g_1(\Delta^2) = \frac{\bar{x}^2}{HS_{xx}}\Big[(1-d)\Delta^2\Big\{2\,G_{3,\nu}\Big(\frac{1}{3}F_\alpha;\Delta^2\Big) - (1+d) \times G_{5,\nu}(\frac{1}{5}F_\alpha;\Delta^2\Big)\Big\} - (1-d^2)\,G_{3,\nu}\Big(\frac{1}{3}F_\alpha;\Delta^2\Big)\Big]. \qquad (3.4.20)$$

The efficiency function attains its maximum at $\Delta^2 = 0$ for all α, given by

$$\text{Eff}(\alpha; 0) = \left[1 - (1 - d^2)\frac{\bar{x}^2}{HS_{xx}}\, G_{3,\nu}\left(\frac{1}{3}F_\alpha; 0\right)\right]^{-1} \geq 1. \qquad (3.4.21)$$

As Δ^2 departs from the origin, $\text{Eff}(\alpha; \Delta^2)$ decreases monotonically crossing the line $\text{Eff}(\alpha; \Delta^2) = 1$ in the interval $\left(\frac{1+d}{2}, \frac{1+d}{1-d}\right)$ to a minimum at $\Delta^2 = \Delta^2_{\text{min}}$. Then from that point increases monotonically towards 1 as $\Delta^2 \to \infty$ from below. Now, for $\Delta^2 = 0$ and varying significance level, we have

$$\max_\alpha \text{Eff}(\alpha, 0) = \text{Eff}(0, 0) = \left[1 - (1 - d^2)\frac{\bar{x}^2}{HS_{xx}}\right]^{-1}. \qquad (3.4.22)$$

As a function of α, $\text{Eff}(\alpha; 0)$ decreases as α increases. On the other hand, $\text{Eff}(\alpha; \Delta^2)$ as a function of Δ^2 is decreasing, and the curves $\text{Eff}(0; \Delta^2)$ and $\text{Eff}(1/2; \Delta^2) = 1$ intersect at $\Delta^2 = \frac{1+d}{1-d}$. The value of Δ^2 at the intersection decreases as α increases. Therefore, for two different levels of significance say, α_1 and α_2, $\text{Eff}(\alpha_1; \Delta^2)$ and $\text{Eff}(\alpha_2; \Delta^2)$ intersect below 1. In order to choose an optimum level of significance with maximum relative efficiency we adopt the following rule.

If it is known that $0 \leq \Delta \leq \frac{1+d}{1-d}$, $\hat{\beta}_0$ is always chosen since $\text{Eff}(0, \Delta^2)$ is maximum for all Δ^2 in this interval. Generally, Δ^2 is unknown. In this case there is no way of choosing a uniformly best estimator of β_0. Thus, we pre-assign a tolerable relative efficiency, say Eff_0. Then, consider the set

$$A_\alpha = \left\{\alpha | \text{Eff}(\alpha; \Delta^2) \geq \text{Eff}_0\right\}. \qquad (3.4.23)$$

An estimator $\hat{\beta}_0^{\text{SPTE}}$ is chosen which maximizes $\text{Eff}(\alpha; \Delta^2)$ over all $\alpha \in A_\alpha$ and Δ^2. For any given Eff_o, solving the equation

$$\max_\alpha \min_{\Delta^2} \text{Eff}(\alpha; \Delta^2) = \text{Eff}_0 \qquad (3.4.24)$$

Table 3.1: Maximum and minimum efficiencies of the SPTE of the intercept parameter relative to the UE for $d = 0.2$.

		Sample size, n						
α		10	15	20	25	30	35	40
0.05	Eff*	2.4395	2.3064	2.2457	2.2114	2.1894	2.1742	2.1630
	Eff_o	0.4810	0.5182	0.5347	0.5441	0.5501	0.5542	0.5573
	Δ_o	5.5700	4.9600	4.6300	4.5300	4.3400	4.3300	4.2600
0.10	Eff*	1.9192	1.8503	1.8201	1.8033	1.7926	1.7852	1.7798
	Eff_o	0.5822	0.6091	0.6209	0.6274	0.6316	0.6345	0.6366
	Δ_o	4.6900	4.2800	4.0600	4.0600	3.9600	3.8800	3.8700
0.15	Eff*	1.6755	1.6336	1.6156	1.6057	1.5994	1.5951	1.5920
	Eff_o	0.6492	0.6695	0.6694	0.6782	0.6830	0.6861	0.6897
	Δ_o	4.3200	4.0300	3.9500	3.7700	3.7800	3.6700	3.6400
0.20	Eff*	1.5270	1.4997	1.4881	1.4818	1.4779	1.4752	1.4732
	Eff_o	0.7000	0.7155	0.7220	0.7257	0.7280	0.7295	0.7307
	Δ_o	3.9900	3.6800	3.6400	3.6500	3.4900	3.5400	3.4600
0.25	Eff*	1.4245	1.4061	1.3985	1.3944	1.3918	1.3900	1.3887
	Eff_o	0.7411	0.7529	0.7579	0.7606	0.7624	0.7635	0.7644
	Δ_o	3.8100	3.5700	3.4900	3.4500	3.4300	3.4100	3.4100
0.35	Eff*	1.2895	1.2809	1.2775	1.2756	1.2745	1.2738	1.2733
	Eff_o	0.8053	0.8120	0.8148	0.8163	0.8173	0.8179	0.8183
	Δ_o	3.5200	3.3400	3.2800	3.2300	3.1900	3.1800	3.2200
0.50	Eff*	1.1696	1.1672	1.1664	1.1660	1.1658	1.1657	1.1656
	Eff_o	0.8747	0.8772	0.8781	0.8786	0.8789	0.8791	0.8792
	Δ_o	3.2100	3.0800	3.0900	3.0900	3.0300	2.9900	3.0100

for α, the solution α^* provides an optimal choice of α, and the procedure is known as *maximin rule* of the optimum level of significance of the preliminary test. Table 3.1 provides the maximum and minimum relative efficiencies Eff* and Eff_o respectively, of the SPTE relative to the UE, of β_0 for selected values of α. It also provides the value $\Delta = \Delta_o$ at which the minimum relative efficiency occurs. For example, if a practitioner has a sample of size 20, chooses $d = 0.2$ and wishes to achieve the minimum relative efficiency 0.7220 of the SPTE, the recommended value of α is 0.20.

3.4.3 The Bias and MSE of the SE

Following Bolfarine and Zacks (1992), the bias and mse functions of the SE of β_0 are derived and presented in the following theorems.

Theorem 3.411 *The bias function of the shrinkage estimator of the intercept parameter β_0 of the simple linear regression model is given by*

$$\mathrm{B}_4\left[\hat{\beta}_0^{\mathrm{SE}}\right] = (1-d)\frac{c\bar{x}\sigma}{\sqrt{S_{xx}}}K_\nu\{1 - 2\,\Phi(-\Delta)\} \tag{3.4.25}$$

where $K_\nu = \sqrt{\frac{2}{n-2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-2}{2}\right)}$ and $\Phi(-\Delta)$ is the c.d.f. of the standard normal distribution evaluated at $-\Delta$.

Proof. By definition, the bias function of the shrinkage estimator of the intercept parameter β_0 is

$$\begin{aligned}
\mathrm{B}_4\left[\hat{\beta}_0^{\mathrm{SE}}\right] &= \mathrm{E}\left[\hat{\beta}_0^{\mathrm{SE}} - \beta_0\right] \\
&= \mathrm{E}\left[\tilde{\beta}_0 - \beta_0 + (1-d)\frac{cS_n\bar{x}}{\sqrt{S_{xx}}|\tilde{\beta}_1 - \beta_{10}|}(\tilde{\beta}_1 - \hat{\beta}_1)\right] \\
&= (1-d)\,\mathrm{E}\left[\frac{cS_n\bar{x}}{\sqrt{S_{xx}}|\tilde{\beta}_1 - \beta_{10}|}(\tilde{\beta}_1 - \hat{\beta}_1)\right] \\
&= (1-d)\frac{c\bar{x}}{\sqrt{S_{xx}}}\,\mathrm{E}[S_n]\,\mathrm{E}\left[\frac{Z}{|Z|}\right]
\end{aligned} \tag{3.4.26}$$

where $Z = \frac{\sqrt{S_{xx}}(\tilde{\beta}_1 - \beta_{10})}{\sigma} \sim \mathcal{N}(\Delta, 1)$.

Recollecting the expressions for $\mathrm{E}[S_n]$ and $\mathrm{E}\left[\frac{Z}{|Z|}\right]$ from (2.4.29) and (2.4.30) respectively, and substituting them into (3.4.26), the bias function of the SE of β_0 is obtained as

$$\mathrm{B}_4\left[\hat{\beta}_0^{\mathrm{SE}}\right] = (1-d)\frac{c\bar{x}\sigma}{\sqrt{S_{xx}}}K_\nu\{1 - 2\Phi(-\Delta)\} \tag{3.4.27}$$

where $K_\nu = \sqrt{\frac{2}{n-2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-2}{2}\right)}$ and $\Phi(-\Delta)$ is the c.d.f. of the standard normal distribution evaluated at $-\Delta$. This completes the proof of the theorem.

From (3.4.27), the quadratic bias function of the SE of β_0 is obtained as

$$\mathrm{QB}_4\left[\hat{\beta}_0^{\mathrm{SE}}\right] = (1-d)^2\frac{c^2\bar{x}^2\sigma^2}{S_{xx}}K_\nu^2\{1-2\,\Phi(-\Delta)\}^2$$
$$= (1-d)^2\frac{c^2\bar{x}^2\sigma^2}{S_{xx}}K_\nu^2\{2\,\Phi(\Delta)-1\}^2. \qquad (3.4.28)$$

Theorem 3.412 *The mean square error function of the shrinkage estimator of the intercept parameter β_0 of the simple linear regression model is given by*

$$\mathrm{M}_4\left[\hat{\beta}_0^{\mathrm{SE}}\right] = \sigma^2\left[\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}\left\{1+(1-d)^2c^2-2(1-d)cK_\nu\sqrt{\frac{2}{\pi}}e^{-\Delta^2/2}\right\}\right]. \qquad (3.4.29)$$

Proof. By definition, the mse function of the SE of β_0 is

$$\mathrm{M}_4\left[\hat{\beta}_0^{\mathrm{SE}}\right] = \mathrm{E}\left[\hat{\beta}_0^{\mathrm{SE}}-\beta_0\right]^2$$
$$= \mathrm{E}\left[\tilde{\beta}_0-\beta_0\right]^2 + (1-d)^2\frac{c^2\bar{x}^2}{S_{xx}}\mathrm{E}\left[\frac{S_n^2(\tilde{\beta}_1-\beta_{10})^2}{|\tilde{\beta}_1-\beta_{10}|^2}\right]$$
$$+\frac{2(1-d)c\bar{x}}{\sqrt{S_{xx}}}\,\mathrm{E}\left[(\tilde{\beta}_0-\beta_0)\frac{S_n(\tilde{\beta}_1-\beta_{10})}{|\tilde{\beta}_1-\beta_{10}|}\right]$$

$$= \sigma^2 H + (1-d)^2\frac{c^2\bar{x}^2\sigma^2}{S_{xx}} - 2c(1-d)\frac{\bar{x}^2\sigma^2K_\nu}{S_{xx}}$$
$$\times\left\{\mathrm{E}\left[|Z|\right]-\Delta\,\mathrm{E}\left[\frac{Z}{|Z|}\right]\right\} \qquad (3.4.30)$$

where $Z\sim\mathcal{N}(\Delta,1)$.

Recollecting the expressions for $\mathrm{E}[|Z|]$ and $\mathrm{E}[Z/|Z|]$ from (2.4.36) and (2.4.30) respectively, and substituting them into (3.4.30), the mse function of the shrinkage estimator of β_0 is obtained as

$$\mathrm{M}_4\left[\hat{\beta}_0^{\mathrm{SE}}\right] = \sigma^2\left[\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}\left\{1+(1-d)^2c^2-2(1-d)cK_\nu\sqrt{\frac{2}{\pi}}e^{-\Delta^2/2}\right\}\right].$$

3.4.3.1 Determination of the Optimum Value of c

A stationary point of $\mathrm{M}_4\left[\hat{\beta}_0^{\mathrm{SE}}\right]$ with respect to c occurs when the first derivative (with respect to c)

$$\mathrm{M}_4'\left[\hat{\beta}_0^{\mathrm{SE}}\right] = \frac{\bar{x}^2\sigma^2}{S_{xx}}\left[2(1-d)^2c - 2(1-d)K_\nu\sqrt{\frac{2}{\pi}}e^{-\Delta^2/2}\right] = 0$$

or

$$2(1-d)^2c = 2(1-d)K_\nu\sqrt{\frac{2}{\pi}}e^{-\Delta^2/2},$$

from which

$$c = (1-d)^{-1}K_\nu\sqrt{\frac{2}{\pi}}e^{-\Delta^2/2} = c^* \quad \text{(say)}. \tag{3.4.32}$$

The second derivative of $\mathrm{M}_4\left[\hat{\beta}_0^{\mathrm{SE}}\right]$ with respect to c is

$$\mathrm{M}_4''\left[\hat{\beta}_0^{\mathrm{SE}}\right] = \frac{2\bar{x}^2\sigma^2(1-d)^2}{S_{xx}} > 0. \tag{3.4.33}$$

Therefore, c^* is the value of c which minimizes (3.4.29). Clearly, the optimum value of c depends on Δ^2 as shown in (3.4.32).

To make c^* independent of Δ^2 we choose $c^0 = (1-d)^{-1}\sqrt{\frac{2}{\pi}}K_\nu$. Hence, the optimum $\mathrm{M}_4\left[\hat{\beta}_0^{\mathrm{SE}}\right]$, say $\mathrm{M}_4^*\left[\hat{\beta}_0^{\mathrm{SE}}\right]$ becomes

$$\mathrm{M}_4^*\left[\hat{\beta}_0^{\mathrm{SE}}\right] = \sigma^2\left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\left\{1 + \frac{2}{\pi}K_\nu^2\left(1 - 2e^{-\Delta^2/2}\right)\right\}\right]. \tag{3.4.34}$$

3.5 Performance Comparison of the Estimators

The quadratic bias and relative efficiency functions of the SRE, SPTE and SE relative to the UE are analysed in this section.

3.5.1 Comparison of the Quadratic Bias Functions

Here the bias functions of the three biased estimators are analysed by analyzing their quadratic bias functions. Also, a best-performed estimator is proposed, under certain conditions.

The quadratic bias functions of the SRE, SPTE and SE of the intercept parameter β_0 are given by

$$\mathrm{QB}_2\left[\hat{\beta}_0^{\mathrm{SRE}}\right] = \frac{\bar{x}^2\sigma^2}{S_{xx}}(1-d)^2\Delta^2, \tag{3.5.1}$$

$$\mathrm{QB}_3\left[\hat{\beta}_0^{\mathrm{SPTE}}\right] = \frac{\bar{x}^2\sigma^2}{S_{xx}}(1-d)^2\Delta^2 G_{3,\nu}^2\left(\frac{1}{3}F_\alpha;\Delta^2\right) \tag{3.5.2}$$

and

$$\mathrm{QB}_4\left[\hat{\beta}_0^{\mathrm{SE}}\right] = \frac{\sigma^2\bar{x}^2}{S_{xx}}K_\nu^2\{2\,\Phi(\Delta)-1\}^2 \tag{3.5.3}$$

respectively. Note that in the derivation of $\mathrm{QB}_4\left[\hat{\beta}_0^{\mathrm{SE}}\right]$ the optimal value of the shrinkage constant has been used, and as a result the factor involving d in $\mathrm{QB}_4\left[\hat{\beta}_0^{\mathrm{SE}}\right]$ cancels out. Figure 3.1 displays the graphs of the quadratic bias functions of the SRE, SPTE and SE of β_0.

Under the null-hypothesis, $\Delta^2 = 0$ and hence the quadratic biases of the SRE, SPTE and SE are all 0 for all d and α. It is observed that as $\Delta^2 \to \infty$, $\mathrm{QB}_2\left[\hat{\beta}_0^{\mathrm{SRE}}\right] \to \infty$ except for $d = 1$; $\mathrm{QB}_3\left[\hat{\beta}_0^{\mathrm{SPTE}}\right] \to 0$ for all α and d; and $\mathrm{QB}_4\left[\hat{\beta}_0^{\mathrm{SE}}\right] \to \frac{\bar{x}^2\sigma^2}{S_{xx}}K_\nu^2$, a constant not depending on d. Therefore, in terms of the quadratic bias, the SRE is uniformly dominated by both the SPTE and SE except for $d = 1$. Also, for very large values of Δ^2 the SE is dominated by the SPTE regardless of the value of α. From some small to moderate values of Δ^2, there is no uniform domination of one estimator over the others. In this case, domination depends on the level of significance α and the degree of distrust d.

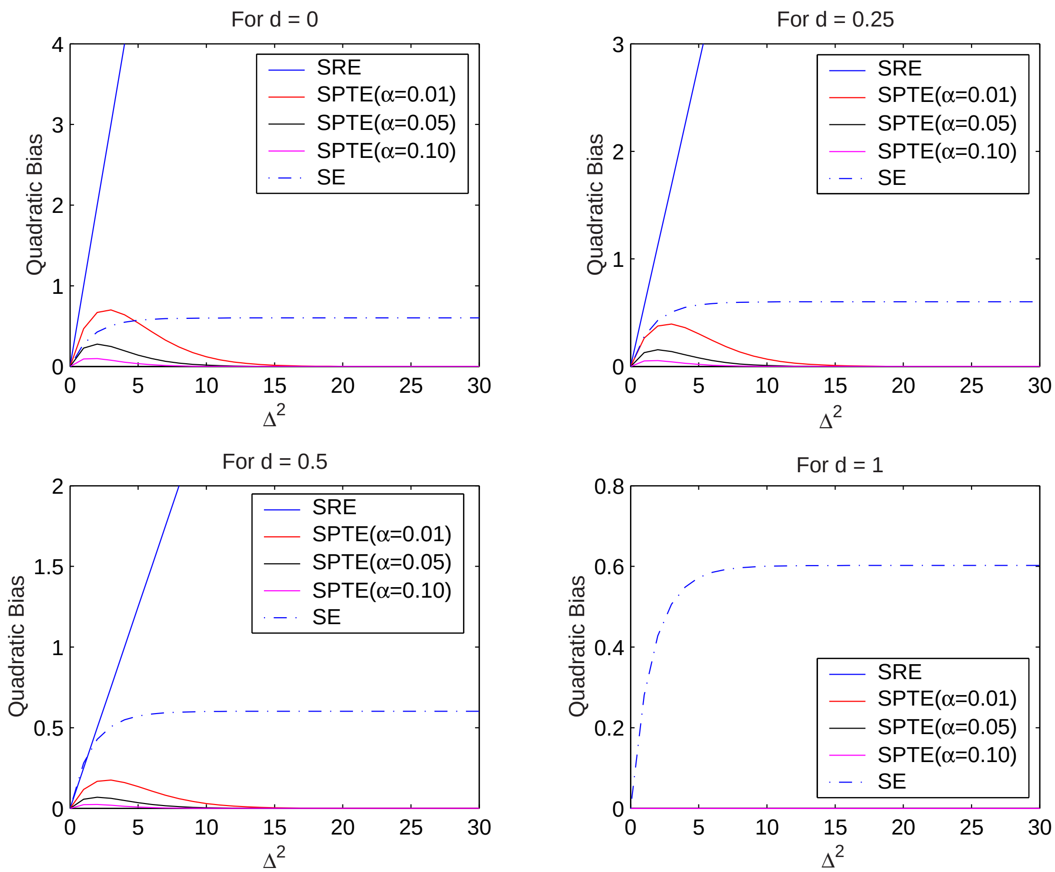

Figure 3.1: The QBs of the SRE, SPTE and SE for selected values of d.

However, Chiou and Saleh (2002) suggested the value of α to be between 20% and 25%. In this interval for α, the quadratic bias of the SPTE approaches zero for some relatively large value of Δ^2. If there is complete distrust of the null hypothesis, the quadratic biases of the RE and SPTE become 0 for any α and Δ^2, while that of the SE remains greater than 0 except for $\Delta^2 = 0$. As the prior information is usually obtained from previous studies or expert knowledge, in practice, the chance of the non-centrality parameter being very large is very slim and α is usually preferred to be reasonably small. Also, the quadratic bias of the SE is relatively stable and approaches a constant value starting from some moderate value of Δ^2 and is unaffected by the choice of d

and α. Although the quadratic bias of the SE stabilizes to a constant for some moderate value of Δ^2, it does not outperform the SPTE except for very small α and d near 0.

3.5.2 Comparison of the Relative Efficiencies

3.5.2.1 Comparing SRE with UE

The efficiency function of the SRE relative to the UE is

$$\text{Eff}\left[\hat{\beta}_0^{\text{SRE}} : \tilde{\beta}_0\right] = H\left[d^2 H + (1-d)^2 \frac{\bar{x}^2}{S_{xx}} \Delta^2\right]^{-1}. \qquad (3.5.4)$$

The efficiency function of the SRE relative to the UE takes its highest possible value at $\Delta^2 = 0$ for $d = 0$. As Δ^2 increases, the efficiency function decreases for all d. It crosses the 1-line at some value of Δ^2 near zero and for some moderate to large value of Δ^2 approaches 0. For $d = 1$, the SRE and UE are equally efficient regardless of the value of Δ^2.

From (3.5.4), the following conclusions are drawn:

- Under H_0, $\Delta^2 = 0$ and hence $\text{Eff}\left[\hat{\beta}_0^{\text{SRE}} : \tilde{\beta}_0\right] = d^{-2} \geq 1$. When $d = 0$, the efficiency function of the SRE grows unboundedly large. As d grows larger from 0, the efficiency decreases, and finally reaches 1-line for $d = 1$. Therefore, under H_0, the SRE is a better choice than the UE.

- As Δ^2 grows larger, the efficiency function grows smaller, and finally as $\Delta^2 \to \infty$, $\text{Eff}\left[\hat{\beta}_0^{\text{SRE}}; \tilde{\beta}_0\right] \to 0$ except for $d = 1$. As $d \to 1$, $\text{Eff}\left[\hat{\beta}_0^{\text{SRE}}; \tilde{\beta}_0\right] \to 1$ from below regardless of the value of Δ^2. Therefore, for any large value of Δ^2, the UE is a better choice than the SRE.

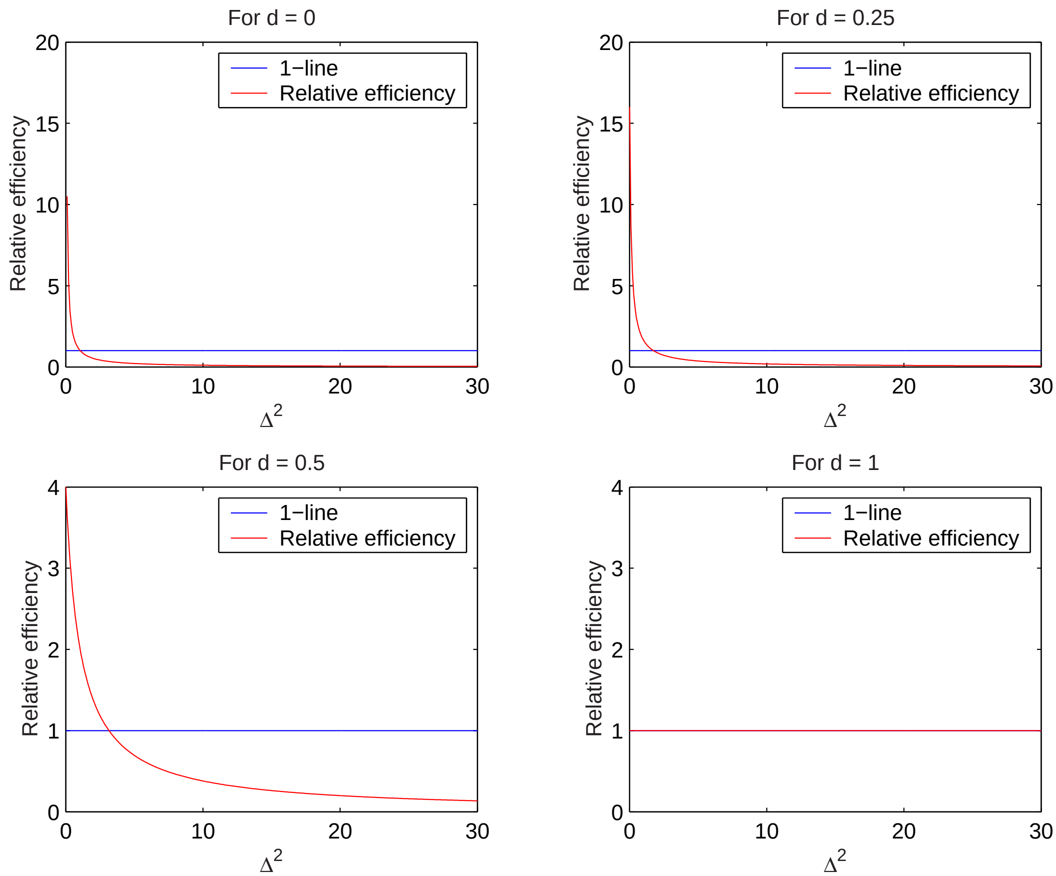

Figure 3.2: The efficiency of the SRE relative to the UE for selected d-values.

In general, the efficiency of the SRE relative to the UE is a decreasing function of Δ^2 with it's maximum value $d^{-2}(\geq 1)$ at $\Delta^2 = 0$ and minimum value 0 at $\Delta^2 = \infty$, unless $d = 1$. The efficiency of the SRE equals 1 at $\Delta^2 = \frac{H(1+d)S_{xx}}{(1-d)\bar{x}^2}$. Thus, if $\Delta^2 \in \left[0, \frac{H(1+d)S_{xx}}{(1-d)\bar{x}^2}\right]$, the SRE is more efficient than the UE, otherwise, the reverse is true. However, in practice the non-sample prior information is usually obtained from some previous experience or expert knowledge and hence it is very likely that Δ^2 would be close to 0. Therefore, for $\Delta^2 = 0$ or near 0, the restricted estimator is a better choice than the unrestricted estimator. Figure 3.2 displays the change in the efficiency of the SRE relative to the UE for change in the value of Δ^2.

3.5.2.2 Comparing SPTE with UE and SRE

The efficiency of the SPTE relative to the UE and SRE are respectively

$$\text{Eff}\left[\hat{\beta}_0^{\text{SPTE}} : \tilde{\beta}_0\right] = H\left[H + \frac{\bar{x}^2\sigma^2}{S_{xx}} g_1(\Delta^2)\right]^{-1} \tag{3.5.5}$$

and

$$\text{Eff}\left[\hat{\beta}_0^{\text{SPTE}} : \hat{\beta}_0^{\text{SRE}}\right] = \left[d^2 H + (1-d)^2\Delta^2\frac{\bar{x}^2}{S_{xx}}\right]\left[H + \frac{\bar{x}^2}{S_{xx}} g_1(\Delta^2)\right]^{-1} \tag{3.5.6}$$

where

$$\begin{aligned} g_1(\Delta^2) = \Delta^2\Big\{&2(1-d)\, G_{3,v}\Big(\frac{1}{3}F_\alpha;\Delta^2\Big) - (1-d^2) \\ &\times G_{5,v}\Big(\frac{1}{5}F_\alpha;\Delta^2\Big)\Big\} - (1-d^2)\, G_{3,v}\Big(\frac{1}{3}F_\alpha;\Delta^2\Big). \end{aligned} \tag{3.5.7}$$

From the expressions in (3.5.5) and (3.5.6), the following conclusions are drawn.

- Under H_0, $\Delta^2 = 0$ and the relative efficiency functions become

$$\text{Eff}\left[\hat{\beta}_0^{\text{SPTE}} : \tilde{\beta}_0\right] = H\left[H - (1-d^2)\frac{\bar{x}^2\sigma^2}{S_{xx}} G_{3,v}\Big(\frac{1}{3}F_\alpha;0\Big)\right]^{-1} \tag{3.5.8}$$

and

$$\text{Eff}\left[\hat{\beta}_0^{\text{SPTE}} : \hat{\beta}_0^{\text{SRE}}\right] = d^2 H\left[H - (1-d^2)\frac{\bar{x}^2\sigma^2}{S_{xx}} G_{3,\nu}\Big(\frac{1}{3}F_\alpha;0\Big)\right]^{-1}. \tag{3.5.9}$$

Therefore, for any fixed d (< 1) the maximum efficiency of the SPTE relative to the UE occurs at $\Delta^2 = 0$ while the minimum relative efficiency of the SPTE relative to the SRE occurs at $\Delta^2 = 0$. As d grows larger, the maximum efficiency of the SPTE relative to the UE decreases while the minimum efficiency of the SPTE relative to the SRE increases. For $d = 1$, the efficiencies of the SPTE, SRE and UE are all equal regardless of the values of α and Δ^2.

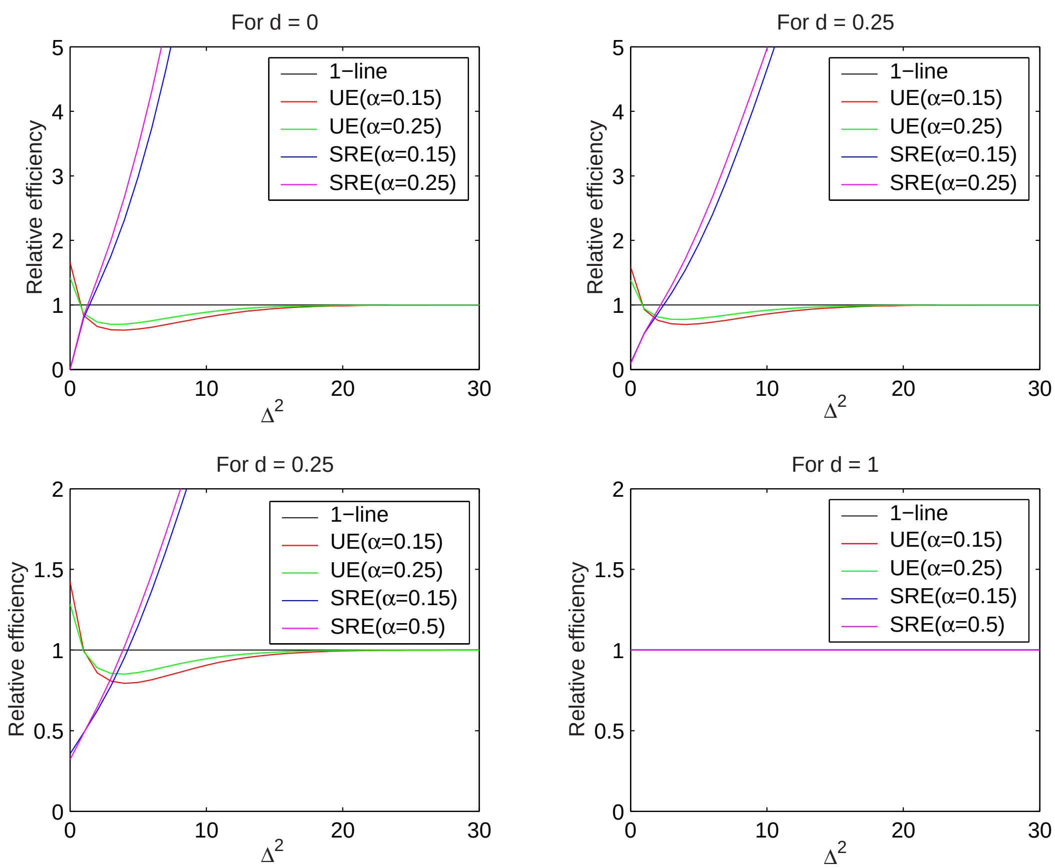

Figure 3.3: The efficiency of the SPTE relative to the UE and SRE for selected values of d.

- As Δ^2 grows larger, the efficiency of the SPTE relative to the UE goes down and crosses the 1-line at

$$\Delta_*^2 = \frac{(1+d)G_{3,\nu}(\frac{1}{3}F_\alpha;\Delta^2)}{2G_{3,\nu}(\frac{1}{3}F_\alpha;\Delta^2) - (1+d)G_{5,\nu}(\frac{1}{5}F_\alpha;\Delta^2)}$$

while the efficiency of the SPTE compared to the SRE goes up and crosses the 1-line at

$$\Delta_{**}^2 = \frac{(1+d)\left\{1 - G_{3,\nu}\left(\frac{1}{3}F_\alpha;\Delta^2\right)\right\}}{(1-d)\left\{1 - 2\,G_{3,\nu}\left(\frac{1}{3}F_\alpha;\Delta^2\right) - (1+d)\,G_{5,\nu}\left(\frac{1}{5}F_\alpha;\Delta^2\right)\right\}}. \tag{3.5.11}$$

- Finally, as $\Delta^2 \to \infty$, $\text{Eff}\left[\hat{\beta}_0^{\text{SPTE}} : \tilde{\beta}_0\right] \to 1$ regardless of the value of

d and α, while the efficiency of the SPTE relative to the SRE grows unboundedly large regardless of the value of α, unless $d = 1$.

In general, the SPTE is more efficient than the UE if $0 \le \Delta^2 < \Delta_*^2$. Starting from some $\Delta^2 > \Delta_*^2$, the UE is more efficient than the SPTE up to some moderate value of Δ^2, and then slowly approaches one. On the other hand, for general $\Delta^2 > 0$, $\text{Eff}\left[\hat{\beta}_0^{\text{SPTE}} : \hat{\beta}_0^{\text{SRE}}\right] \lesseqqgtr 1$ according as $\Delta^2 \lesseqqgtr \Delta_{**}^2$. Figure 3.3 displays the change in the efficiency of the SPTE relative to the UE and SRE, for change in the value of Δ^2.

3.5.2.3 Comparing SE with UE, SRE and SPTE

The relative efficiencies of the SE relative to the UE, SRE and SPTE are respectively

$$\text{Eff}\left[\hat{\beta}_0^{\text{SE}} : \tilde{\beta}_0\right] = \left[1 + \frac{2\bar{x}^2 K_\nu^2 \varphi}{\pi H S_{xx}}\right]^{-1}, \tag{3.5.12}$$

$$\text{Eff}\left[\hat{\beta}_0^{\text{SE}} : \hat{\beta}_0^{\text{SRE}}\right] = \left[d^2 H + (1-d)^2 \frac{\bar{x}^2 \Delta^2}{S_{xx}}\right] \left[H + \frac{2\bar{x}^2 K_\nu^2 \varphi}{\pi S_{xx}}\right]^{-1}, \tag{3.5.13}$$

and

$$\text{Eff}\left[\hat{\beta}_0^{\text{SE}} : \hat{\beta}_0^{\text{SPTE}}\right] = \left[H + \frac{\bar{x}^2 \sigma^2}{S_{xx}} g_1(\Delta^2)\right] \left[H + \frac{2\bar{x}^2 K_\nu^2 \varphi}{\pi S_{xx}}\right]^{-1} \tag{3.5.14}$$

where $\varphi = \left(1 - 2\, e^{-\Delta^2/2}\right)$ and $g_1(\Delta^2)$ is defined in (3.4.20).

The efficiency of the SE relative to the UE is a decreasing function of Δ^2 which takes its maximum value at $\Delta^2 = 0$. It falls sharply as Δ^2 moves away from 0, and approaches a constant value for some moderate value of Δ^2.

The efficiency of the SE relative to the SRE is an increasing function of Δ^2 which takes its minimum value at $\Delta^2 = 0$. It grows unboundedly large as Δ^2 increases.

The efficiency of the SE relative to the SPTE is neither an increasing nor a decreasing function of Δ^2. Moreover, it depends on the choice of the level of significance. For moderate to large value of Δ^2 it approaches a constant value regardless of the choice of α.

From the expressions in (3.5.12) – (3.5.14), the following conclusions are drawn.

- Under H_0, $\Delta^2 = 0$ and hence

$$\text{Eff}\left[\hat{\beta}_0^{\text{SE}} : \tilde{\beta}_0\right] = \left[1 - 2H^{-1}VK_\nu^2\right]^{-1}, \quad (3.5.15)$$

$$\text{Eff}\left[\hat{\beta}_0^{\text{SE}} : \hat{\beta}_0^{\text{SRE}}\right] = d^2\left[1 - 2H^{-1}VK_\nu^2\right]^{-1}, \quad (3.5.16)$$

and

$$\text{Eff}\left[\hat{\beta}_0^{\text{SE}} : \hat{\beta}_0^{\text{SPTE}}\right] = \frac{1 - H^{-1}V(1-d^2)G_{3,\nu}\left(\frac{1}{3}F_\alpha; 0\right)}{1 - 2H^{-1}VK_\nu^2} \quad (3.5.17)$$

where $V = \frac{H^{-1}\bar{x}^2}{\pi S_{xx}}$. The second term on the right hand side of (3.5.15) is always positive, so the maximum efficiency of the SE relative to the UE is attained at $\Delta^2 = 0$, and it is greater than 1 for all values of Δ^2 near 0. The relative efficiencies of the SE relative to the SRE and SPTE depend on d. When $d = 0$, the minimum efficiency of the SE relative to the SRE is 0 at $\Delta^2 = 0$. No such minimum or maximum efficiency of the SE relative to the SPTE exists at $\Delta^2 = 0$. The value of $G_{3,v}\left(\frac{1}{3}F_\alpha; 0\right)$ is smaller for larger α. Therefore, at $\Delta^2 = 0$ the efficiency of the SE relative to the SPTE is higher for larger value of α. As $\alpha \to 1$, $G_{3,v}\left(\frac{1}{3}F_\alpha; 0\right) \to 0$, and hence the efficiency of the SE relative to the SPTE tends to that of the SE relative to the UE. However, for a fixed α, as d increases, the efficiency function of the SE relative to the SPTE also increases. When there is complete distrust of the null hypothesis, the relative efficiencies of the SE relative to all other estimators become the same regardless of the choice of Δ^2 and α.

- As Δ^2 moves away from 0, the efficiency of the SE relative to the UE falls sharply, the efficiency relative to the SRE quickly increases, and the efficiency relative to the SPTE goes up or down according as $2\, G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right) \lessgtr (1+d)\, G_{5,\nu}\left(\frac{1}{5}F_\alpha; \Delta^2\right)$.
- $\Delta^2 \to \infty$, $\text{Eff}\left[\hat{\beta}_0^{\text{SE}} : \tilde{\beta}_0\right] \to \left\{1 + (1+\frac{S_{xx}}{n\bar{x}^2})^{-1}\frac{2}{\pi}K_\nu^2\right]^{-1} < 1$; $\text{Eff}\left[\hat{\beta}_0^{\text{SE}} : \hat{\beta}_0^{\text{SRE}}\right] \to \infty$, except for $d = 1$; and $\text{Eff}\left[\hat{\beta}_0^{\text{SE}} : \hat{\beta}_0^{\text{SPTE}}\right]$ approaches the constant value $\left[1 + \frac{2H^{-1}\bar{x}^2K_\nu^2}{\pi S_{xx}}\right]^{-1}$, which does not depend on d and α.

In general, the efficiency of the SE relative to the UE decreases from $\left\{1 - (1+\frac{S_{xx}}{n\bar{x}^2})^{-1}\frac{2}{\pi}K_\nu^2\right]^{-1}$ at $\Delta^2 = 0$, crosses the 1-line at $\Delta^2 = \ln 4$, and approaches a constant value as $\Delta^2 \to \infty$. Therefore, for $\Delta^2 < \ln 4$, the SE performs better than the UE; otherwise, the UE performs better than the SE. On the other hand, $\text{Eff}\left[\hat{\beta}_0^{\text{SE}} : \hat{\beta}_0^{\text{SRE}}\right]$ increases as Δ^2 moves away from 0. It grows unboundedly large as $\Delta^2 \to \infty$. The general picture of the efficiency of the SE compare to the SPTE can be described as follows. The relative efficiency function has the value in (3.5.17) at $\Delta^2 = 0$ and crosses the 1-line at

$$\Delta^2 = \frac{\frac{2}{\pi}K_\nu^2\varphi - (1-d^2)G_{5,\nu}\left(\frac{1}{5}F_\alpha; \Delta^2\right)}{\sigma^2\left[2(1-d)\, G_{3,\nu}\left(\frac{1}{3}F_\alpha; \Delta^2\right) - (1+d)^2\, G_{5,\nu}\left(\frac{1}{5}F_\alpha; \Delta^2\right)\right]}. \qquad (3.5.18)$$

As $\Delta^2 \to \infty$, the relative efficiency function approaches the constant value $\left[1 + \frac{2H^{-1}\bar{x}^2K_\nu^2}{\pi S_{xx}}\right]^{-1}$. Figure 3.4 displays the change in the efficiency of the SE relative to the UE, SRE and SPTE for change in the value of Δ^2.

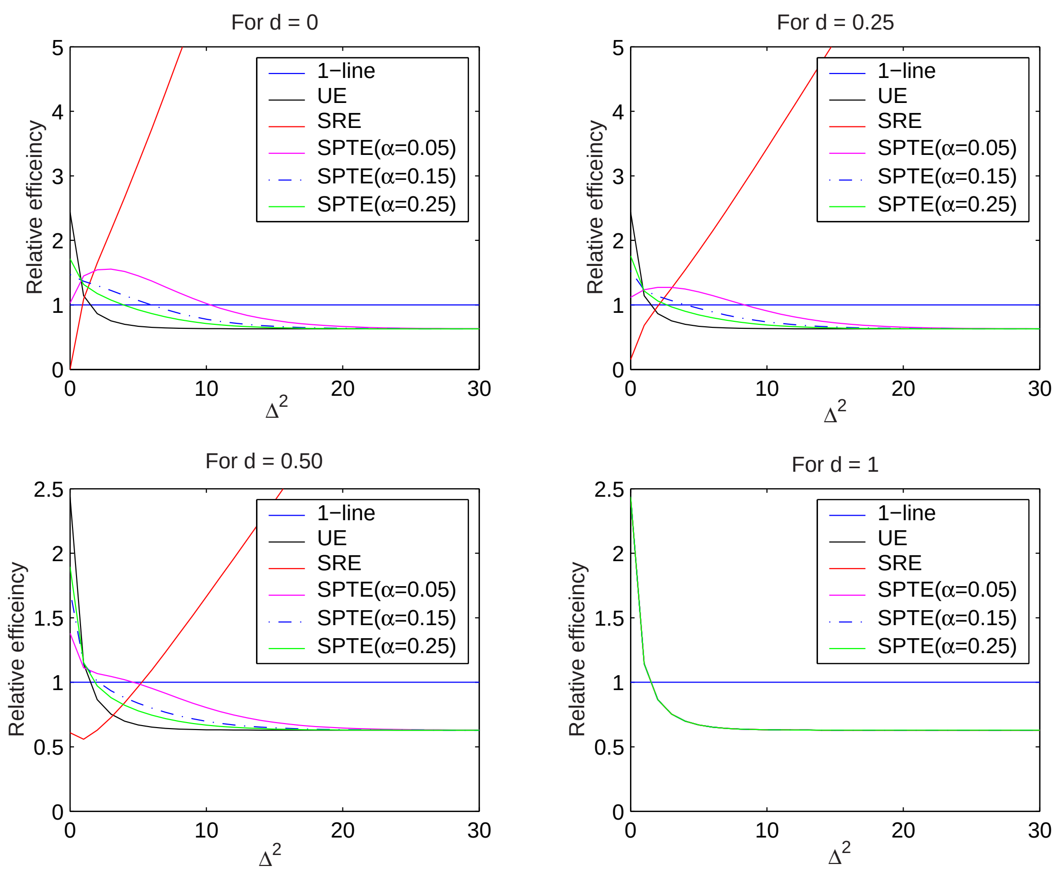

Figure 3.4: The efficiency of the SE relative to the UE, SRE and SPTE for selected values of d.

3.6 Concluding Remarks

Among the four estimators considered in this study, the UE is the only unbiased estimator and is based exclusively on the sample information. The estimators based on both the sample and non-sample prior information are biased. However, the inclusion of non-sample prior information increases the efficiencies of the estimators. The relative efficiencies of the biased estimators depend on the departure parameter Δ^2 and the degree of distrust d. From

0 to some moderate value of Δ^2, the SE dominates the UE for all values of d. Starting from some moderate value of Δ^2 the SE is dominated by the UE. From 0 to some moderate value of Δ^2 the SE is dominated by the SRE. But starting from that moderate value of Δ^2 the SE dominates the SRE. However, the increasing rate of the efficiency of the SE relative to the SRE decreases as the value of the coefficient of distrust increases. Under the null hypothesis the SE dominates the SPTE unless α or d is not too small. From some small to moderate values of Δ^2, the SE dominates the SPTE if α is not too large. Starting from some moderate value of Δ^2, the SE is dominated by the SPTE. In practice, the non-sample prior information is obtained from expert knowledge or previous studies, and hence the value of the degree of distrust on the null hypothesis is very unlikely to be close to 1. Also, the level of significance is preferred to be small. Therefore, under such situation the shrinkage estimator would be the best choice as an improved estimator of the intercept parameter among all the estimators considered in this study.

3.A Appendix

- The following MATLAB codes are used for producing Figure 3.2.

```
d=0.25; D=0:1:30; n=20; x=1; x2=(1-d).^2;
M2=(d.^2.*(1./n+x))+x2.*D.*x; x1=1-d.^2; q2=x2.*D.*x;
plot(D,q2)
hold on
v=n-2; G39=ncfcdf(finv(.95,3,v)./3,3,v, D);
q39=x2.*D.*x.*G39.^2;
plot(D,q39,'r')
G35=ncfcdf(finv(.85,3,v)./3,3,v, D); q35=x2.*D.*x.*G35.^2;
plot(D,q35,'k')
G30=ncfcdf(finv(.75,3,v)./3,3,v, D);
q30=x2.*D.*x.*G30.^2;
plot(D,q30,'m')
k=sqrt(2./(n-2)).*gamma((n-1)./2)./gamma((n-2)./2);
O=sqrt(D);F=normcdf(O, 0, 1);
q4=2./pi.*k.^4.*x.*(2.*F- 1).^2;
plot(D,q4,'-.')
legend('RE', 'PTE(\alpha=0.01)',
'PTE(\alpha=0.05)', 'PTE(\alpha=0.15)', 'SE', 1)
xlabel('\Delta^2'); ylabel('Quadratic bias');
title('For d = 0.25');
```

- The following MATLAB codes are used for producing Figure 3.4.

```
d=0.25; D=0:1:30; n=20; v=n-2; x=1-d.^2;
```

```
q=ones(1,length(D));
plot(D,q)
hold on
y=1-d;z=1+d;
k=sqrt(2./(n-2)).*gamma((n-1)./2)./gamma((n-2)./2);
t=1;M1=1./n+t;
M4= 1./n + t.*(1+2./pi.*k.^2.*(1-2.*exp(-D./2)));
R1=M1./M4;plot(D, R1, 'k')
M2= d.^2.*(1./n+t)+(1-d).^2.*t.*D;
R2=M2./M4; plot(D, R2, 'r')
a=0.95;G3=ncfcdf(finv(a,3,v)./3,3,v,D);
G5=ncfcdf(finv(a,5,v)./5,5,v, D);
M3=1./n+t+t.*(D.*(2.*y.*G3-x.*G5)-x.*G3);
R3=M3./M4; plot(D,R3,'m')
a=0.85; G3=ncfcdf(finv(a,3,v)./3,3,v,D);
G5=ncfcdf(finv(a,5,v)./5,5,v, D);
M3=1./n + t + t.*(D.*(2.*y.*G3-x.*G5)-x.*G3); R4=M3./M4;
plot(D, R4, 'b-.')
a=0.75; G3=ncfcdf(finv(a,3,v)./3,3,v, D);
G5=ncfcdf(finv(a,5,v)./5,5,v,D);
M3=1./n + t + t.*(D.*(2.*y.*G3-x.*G5)-x.*G3);
R5=M3./M4; plot(D, R5, 'g')
legend ('1-line', 'UE','SRE','SPTE(\alpha=0.05)',
'SPTE(\alpha=0.15)','SPTE(\alpha=0.25)', 1)
xlabel('\Delta^2');ylabel('Relative efficiency');
title('For d =0.25')
```

Part II

Estimation Based on Three Tests

Chapter 4

Shrinkage Preliminary Test Estimator of Multiple Linear Regression Model Based on Three Tests

4.1 Introduction

Consider the multiple linear regression model

$$\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{4.1.1}$$

where $\boldsymbol{Y}$ is an $n \times 1$ vector of response variables, X is an $n \times p$ matrix of non-stochastic predictors, $\boldsymbol{\beta}$ is a $p \times 1$ vector of regression coefficients and $\boldsymbol{\varepsilon}$ is an error vector having the same dimension as $\boldsymbol{Y}$. It is assumed that X is of full rank, and $n \geq p$. Also assume that the error vector follows the n-dimensional normal distribution with mean vector $\mathbf{0}$ and covariance matrix $\sigma^2 I_n$, where I_n is an identity matrix of order n.

The exclusive sample information based maximum likelihood estimator or the least square estimator of $\boldsymbol{\beta}$ is known as the unrestricted estimator (UE).

Suppose uncertain non-sample prior information about the value of $\boldsymbol{\beta}$ is available and is expressed in the form of the null hypothesis

$$H_0 : H\boldsymbol{\beta} = \boldsymbol{h} \tag{4.1.2}$$

where H is a $q \times p$ matrix of full row rank and $\boldsymbol{h}$ is a known $q \times 1$ vector. The estimators of $\boldsymbol{\beta}$ and σ^2 under the null hypothesis in (4.1.2) are called restricted estimators (RE).

With respect to the quadratic loss function, the RE of $\boldsymbol{\beta}$ performs better than the UE when the null hypothesis holds. Otherwise, the RE may be considerably biased, inefficient and inconsistent while the performance of the UE remains steady over any such departures (cf. Billah and Saleh, 1998). Therefore, we define a SPTE of $\boldsymbol{\beta}$ that combines both sample and uncertain non-sample prior information and performs better than both UE and RE, under certain conditions, as shown in Chapters 2 and 3.

To remove the uncertainty in the non-sample prior information, we use the Wald (W), likelihood ratio (LR) and Lagrange multiplier (LM) tests as the preliminary tests on the null hypothesis. Moreover, we use the modified and Edgeworth size-corrected W, LR and LM tests. We investigate whether the corrections to the tests reduce the conflict among the statistical properties of the SPTEs of $\boldsymbol{\beta}$. Conflict is defined as the difference between the largest and smallest quadratic biases, and efficiencies with respect to the quadratic risk relative to the UE. The conflict among the relative efficiencies of the SPTEs under both the original and modified W, LR and LM tests are calculated. Also, we calculate the conflict among the relative efficiencies of the estimators based on the Edgeworth size-corrected tests. The conflicts among the properties of the SPTEs are studied to make a recommendation for the choice of the preliminary test.

The layout of this chapter is as follows. Some preliminaries are outlined in Section 4.2. The three tests are briefly discussed in Section 4.3. The bias and quadratic risk (QR) functions of the SPTEs under the original W, LR and LM tests are derived in Section 4.4. The bias and QR functions of the SPTEs under the modified and size-corrected W, LR and LM tests are derived in Section 4.5 and 4.6 respectively. The bias and efficiencies of the SPTEs relative to the UE are analysed in Section 4.7. An example, that illustrates the definitions of the test statistics, the tables of conflict among the efficiencies of the SPTEs relative to the UE and selected MATLAB codes, used for producing graphs, are presented in Appendix 4.A.

4.2 Some Preliminaries

By definition, the UE of $\boldsymbol{\beta}$ is

$$\begin{aligned}\tilde{\boldsymbol{\beta}} &= (X'X)^{-1}X'\boldsymbol{Y} \\ &= C^{-1}X'\boldsymbol{Y} \qquad (4.2.3)\end{aligned}$$

where $C = X'X$. Furthermore, the UE of σ^2 is

$$\tilde{\sigma}^2 = \frac{1}{n}(\boldsymbol{Y} - X\tilde{\boldsymbol{\beta}})'(\boldsymbol{Y} - X\tilde{\boldsymbol{\beta}}). \qquad (4.2.4)$$

Under the null hypothesis in (4.1.2), the restricted estimators of $\boldsymbol{\beta}$ and σ^2 are

$$\hat{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}} - C^{-1}H'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h}) \qquad (4.2.5)$$

and

$$\hat{\sigma}^2 = \frac{1}{n}(\boldsymbol{Y} - X\hat{\boldsymbol{\beta}})'(\boldsymbol{Y} - X\hat{\boldsymbol{\beta}}) \tag{4.2.6}$$

respectively.

If ξ is any appropriate test statistic for testing the null hypothesis in (4.1.2), then a simple form of the SPTE of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}}_{\xi}^{\text{SPTE}} = \tilde{\boldsymbol{\beta}} - (1-d)(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\, I(\xi < \xi_{\alpha}), \tag{4.2.7}$$

where ξ_α is the α-level critical value of the test statistic ξ , d $(0 \leq d \leq 1)$ is the coefficient of distrust on the null hypothesis, and $I(\cdot)$ is an indicator function which assumes value unity when the inequality in the argument holds, and 0 otherwise.

For the computation of the bias functions, mse matrices and quadratic risk functions of the estimators of $\boldsymbol{\beta}$, the following definitions are used.

Let $\boldsymbol{\beta}^*$ be an estimator of $\boldsymbol{\beta}$. Then the bias and quadratic bias (QB) functions of $\boldsymbol{\beta}^*$ are defined as

$$\text{B}[\boldsymbol{\beta}^*] = \text{E}[\boldsymbol{\beta}^* - \boldsymbol{\beta}] = \mathbf{b} \text{ (say)} \tag{4.2.8}$$

and

$$\text{QB}[\boldsymbol{\beta}^*] = \mathbf{b}'\mathbf{b} \tag{4.2.9}$$

respectively. Now, consider the quadratic loss function

$$\text{L}[\boldsymbol{\beta}^*] = (\boldsymbol{\beta}^* - \boldsymbol{\beta})'\mathcal{W}(\boldsymbol{\beta}^* - \boldsymbol{\beta}) \tag{4.2.10}$$

for a given positive definite matrix $\mathcal{W}$ of appropriate order. Then the quadratic risk for estimating $\boldsymbol{\beta}$ by $\boldsymbol{\beta}^*$ is

$$\text{R}[\boldsymbol{\beta}^*] = \text{E}[(\boldsymbol{\beta}^* - \boldsymbol{\beta})'\mathcal{W}(\boldsymbol{\beta}^* - \boldsymbol{\beta})] = \text{tr}[\mathcal{W}M] \tag{4.2.11}$$

where 'tr' denotes the trace operator and M is the mean square error (mse) matrix of the estimator $\boldsymbol{\beta}^*$ given by

$$M = \mathrm{M}[\boldsymbol{\beta}^*] = \mathrm{E}[(\boldsymbol{\beta}^* - \boldsymbol{\beta})(\boldsymbol{\beta}^* - \boldsymbol{\beta})'] \,. \tag{4.2.12}$$

To facilitate the computation of the quadratic risk functions of the estimators of $\boldsymbol{\beta}$, we assume $\mathcal{W} = \sigma^{-2}C$. Therefore, the bias function, mse matrix and quadratic risk of the UE of $\boldsymbol{\beta}$ are $\mathbf{0}$, $\sigma^2 C^{-1}$ and p, respectively. The bias and mse matrix of the RE of $\boldsymbol{\beta}$ are

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}\right] = -C^{-1}H'(HC^{-1}H')^{-1}(H\boldsymbol{\beta} - \boldsymbol{h}) = \boldsymbol{\eta} \text{ (say)} \tag{4.2.13}$$

and

$$\mathrm{M}\left[\hat{\boldsymbol{\beta}}\right] = \sigma^2 C^{-1} - \sigma^2 \Lambda + \boldsymbol{\eta}\boldsymbol{\eta}' \tag{4.2.14}$$

respectively, where $\Lambda = C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}$.

4.3 The W, LR and LM Tests

Suppose there are n independent responses $y_1, \ldots, y_n$, with identical density function $f(y;\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a $p \times 1$ parameter vector. Also, let $l_i(\boldsymbol{\theta}) = \ln f(y_i;\boldsymbol{\theta})$ be the log-density function of the i^{th} response. Then the log-likelihood function, score vector and information matrix are

$$l(\boldsymbol{\theta}) = \sum_{i=1}^{n} l_i(\boldsymbol{\theta}), \tag{4.3.1}$$

$$d(\boldsymbol{\theta}) = \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \tag{4.3.2}$$

and

$$I(\boldsymbol{\theta}) = -\mathrm{E}\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right] \tag{4.3.3}$$

respectively.

Suppose we are interested in testing the null hypothesis

$$H_0 : h(\boldsymbol{\theta}) = \mathbf{0} \tag{4.3.4}$$

where $h(\boldsymbol{\theta})$ is a $r \times 1$ vector function of $\boldsymbol{\theta}$ with $r \leq p$.

Let $H(\boldsymbol{\theta}) = \frac{\partial h(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}$ be the $r \times p$ matrix of derivatives. Based on the unrestricted estimator $\tilde{\boldsymbol{\theta}}$ and restricted (by H_0) estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$, the W, LR and LM test statistics for testing the null hypothesis in (4.3.4) are

$$\xi_{\mathrm{W}} = h(\tilde{\boldsymbol{\theta}})' \left[H(\tilde{\boldsymbol{\theta}}) I(\tilde{\boldsymbol{\theta}})^{-1} H(\tilde{\boldsymbol{\theta}})' \right]^{-1} h(\tilde{\boldsymbol{\theta}}), \tag{4.3.5}$$

$$\xi_{\mathrm{LR}} = 2 \left[l(\tilde{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}) \right] \tag{4.3.6}$$

and

$$\xi_{\mathrm{LM}} = d(\hat{\boldsymbol{\theta}})' I(\hat{\boldsymbol{\theta}})^{-1} d(\hat{\boldsymbol{\theta}}) \tag{4.3.7}$$

respectively (see Evans and Savin, 1980). Under the null hypothesis, the asymptotic distributions of the test statistics in (4.3.5) – (4.3.7) are the same as χ^2_r (see Engle, 1984). To illustrate the above definitions of the test statistics, an example is provided in Appendix4.A.

4.4 The Original W, LR and LM Tests and the SPTE of $\boldsymbol{\beta}$

To test the null hypothesis in (4.1.2) the usual F statistic is

$$F = \frac{(RRSS - URSS)}{URSS} \frac{m}{q} \tag{4.4.1}$$

where $m = n - p$, $URSS = (\boldsymbol{Y} - X\tilde{\boldsymbol{\beta}})'(\boldsymbol{Y} - X\tilde{\boldsymbol{\beta}})$ is the unrestricted residual sum of squares and $RRSS = (\boldsymbol{Y} - X\hat{\boldsymbol{\beta}})'(\boldsymbol{Y} - X\hat{\boldsymbol{\beta}})$ is the restricted residual sum of squares. The above test statistic can be written in the following form.

$$F = \frac{(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})}{qS^2} \tag{4.4.2}$$

where $S^2 = \frac{1}{m}(\boldsymbol{y} - X\tilde{\boldsymbol{\beta}})'(\boldsymbol{y} - X\tilde{\boldsymbol{\beta}})$ is an unbiased estimator of σ^2.

To find the sampling distribution of the test statistic in (4.4.2) consider the transformation,

$$\boldsymbol{Z} = \sigma^{-1}(HC^{-1}H')^{-1/2}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h}). \tag{4.4.3}$$

Here, $\boldsymbol{Z} \sim N_q\left(\sigma^{-1}(HC^{-1}H')^{-1/2}(H\boldsymbol{\beta} - \boldsymbol{h}),\ \boldsymbol{I}_q\right)$.

Therefore, under the alternative hypothesis,

$$\boldsymbol{Z}'\boldsymbol{Z} = \frac{(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})}{\sigma^2} \tag{4.4.4}$$

is distributed as a non-central chi-square variable with q d.f. and non centrality parameter

$$\Delta = \frac{(H\boldsymbol{\beta} - \boldsymbol{h})'(HC^{-1}H')^{-1}(H\boldsymbol{\beta} - \boldsymbol{h})}{\sigma^2}. \tag{4.4.5}$$

Also,

$$\frac{S^2}{\sigma^2} = \frac{(\boldsymbol{y} - X\tilde{\boldsymbol{\beta}})'(\boldsymbol{y} - X\tilde{\boldsymbol{\beta}})}{m\sigma^2} = \frac{\chi^2_m}{m}. \tag{4.4.6}$$

Therefore, under the alternative hypothesis, the distribution of the test statistic in (4.4.2) is the non-central F distribution with $(q,\ m)$ d.f. and non-centrality parameter Δ. An equivalent derivation of the distribution of F can is given in Ashish and Srivastava (1997), p. 64 – 65.

As alternative to the F test, the W, LR and LM tests can also be used to test the same hypothesis. The respective test statistics are

$$\xi_{\mathrm{W}} = (H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'(\tilde{\sigma}^2 HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h}), \tag{4.4.7}$$

$$\xi_{\mathrm{LR}} = n\,[\ln\hat{\sigma}^2 - \ln\tilde{\sigma}^2] \tag{4.4.8}$$

and

$$\xi_{\mathrm{LM}} = (H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'(\hat{\sigma}^2 HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h}) \tag{4.4.9}$$

(cf. Evans and Savin, 1982). The above test statistics can be written as functions of the F statistic as follows (cf. Ullah and Zinde-Walsh, 1984).

$$\xi_{\mathrm{W}} = \frac{nq}{m}F, \tag{4.4.10}$$

$$\xi_{\mathrm{LR}} = n\,\ln\!\left(1 + \frac{q}{m}F\right) \tag{4.4.11}$$

and

$$\xi_{\mathrm{LM}} = \frac{nqF}{m + qF}. \tag{4.4.12}$$

Under the null hypothesis, the asymptotic distributions of the test statistics in (4.4.10) - (4.4.12) are the same as χ^2_q (cf. Evans and Savin, 1980).

By definition, the SPTE involves an appropriate test statistic for testing the null hypothesis that is formed by the non-sample prior information. Since the above tests statistics are equally appropriate, we may use each one of them. Therefore, the SPTEs of $\boldsymbol{\beta}$ under the W, LR and LM test statistics are

$$\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{W}} = \tilde{\boldsymbol{\beta}} - \bar{d}\,(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\,I(\xi_{\mathrm{W}} < \chi^2_\alpha), \tag{4.4.13}$$

$$\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LR}} = \tilde{\boldsymbol{\beta}} - \bar{d}\,(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\,I(\xi_{\mathrm{LR}} < \chi^2_\alpha) \tag{4.4.14}$$

and

$$\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LM}} = \tilde{\boldsymbol{\beta}} - \bar{d}\,(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\,I(\xi_{\mathrm{LM}} < \chi^2_\alpha) \tag{4.4.15}$$

respectively, where $\bar{d} = (1 - d)$, and χ^2_α is the chi-square critical value at significance level α and d.f. q.

The bias, quadratic bias and quadratic risk functions of the SPTEs in (4.4.13) – (4.4.15) are presented in the following theorems.

Theorem 4.13. *The bias functions of the SPTEs of $\boldsymbol{\beta}$ under the original W, LR and LM tests are respectively*

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{W}}\right] = \boldsymbol{\eta}\, \bar{d}\, G_{q+2,\, m}\, (r^{\mathrm{W}}_2;\, \Delta)\,, \tag{4.4.16}$$

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LR}}\right] = \boldsymbol{\eta}\, \bar{d}\, G_{q+2,\, m}\, (r^{\mathrm{LR}}_2;\, \Delta) \tag{4.4.17}$$

and

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LM}}\right] = \boldsymbol{\eta}\, \bar{d}\, G_{q+2,\, m}\, (r^{\mathrm{LM}}_2;\, \Delta) \tag{4.4.18}$$

where $r^{\mathrm{W}}_2 = \frac{m\chi^2_\alpha}{n(q+2)}$, $r^{\mathrm{LR}}_2 = \frac{m}{(q+2)}\left(e^{\frac{\chi^2_\alpha}{n}} - 1\right)$, $r^{\mathrm{LM}}_2 = \frac{m\chi^2_\alpha}{(q+2)\left(n-\chi^2_\alpha\right)}$; *and* $G_{a,b}(r;\, \Delta)$ *is the cumulative distribution function of the non-central* F *distribution with* $(a,\, b)$ *d.f., non-centrality parameter* Δ *and is evaluated at* r.

Proof. By definition, the bias function of the SPTE of $\boldsymbol{\beta}$ under the W test is

$$\begin{aligned}
\mathrm{B}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{W}}\right] &= \mathrm{E}\left[\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta} - \bar{d}\,(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\, I\left(\xi_{\mathrm{W}} < \chi^2_\alpha\right)\right] \\
&= -\bar{d}\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\, I\left(\xi_{\mathrm{W}} < \chi^2_\alpha\right)\right] \\
&= -\bar{d}\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\, I\left(\frac{nq}{m}F < \chi^2_\alpha\right)\right] \\
&= -\bar{d}\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\, I\left(F < \frac{m\chi^2_\alpha}{nq}\right)\right] \\
&= -\bar{d}\, C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\Bigg[(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h}) \\
&\qquad \times I\left(\frac{(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})}{qS^2} < \frac{m\chi^2_\alpha}{nq}\right)\Bigg]
\end{aligned}$$

$$= -\bar{d}\, C^{-1}H'(HC^{-1}H')^{-1}\sigma(HC^{-1}H')^{\frac{1}{2}}\, \mathrm{E}\left[\boldsymbol{Z}\, I\left(\frac{m\boldsymbol{Z}'\boldsymbol{Z}}{q\chi^2_m} < \frac{m\chi^2_\alpha}{nq}\right)\right].$$

Applying Theorem 1, Appendix B2, Judge and Bock (1978), to the above expression, the bias function of the SPTE of $\boldsymbol{\beta}$ under the W test is obtained as

$$\begin{aligned}\mathrm{B}\left[\hat{\boldsymbol{\beta}}^{\text{SPTE}}_{\text{W}}\right] &= -\bar{d}C^{-1}H'(HC^{-1}H')^{-1}(H\boldsymbol{\beta} - \boldsymbol{h})\, \mathrm{E}\left[I\left(\frac{m\chi^2_{q+2}(\Delta)}{(q+2)\chi^2_m} < \frac{m\chi^2_\alpha}{n(q+2)}\right)\right] \\ &= \boldsymbol{\eta}\bar{d}\, G_{q+2,\,m}\,(r^{\text{W}}_2;\Delta)\,. \end{aligned} \tag{4.4.19}$$

This completes the proof of (4.4.16) of the theorem.

Now, we prove (4.4.17) for the bias function of the SPTEs of $\boldsymbol{\beta}$ under the LR test.

By definition, the bias function of the SPTE of $\boldsymbol{\beta}$ under the LR test is

$$\begin{aligned}\mathrm{B}\left[\hat{\boldsymbol{\beta}}^{\text{SPTE}}_{\text{LR}}\right] &= \mathrm{E}\left[\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta} - \bar{d}\,(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\, I\big(\xi_{\text{LR}} < \chi^2_\alpha\big)\right] \\ &= -\bar{d}\, \mathrm{E}\left[(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\, I\big(\xi_{\text{LR}} < \chi^2_\alpha\big)\right] \\ &= -\bar{d}\, \mathrm{E}\left[(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\, I\left(n\, \ln\left\{1 + \frac{qF}{m}\right\} < \chi^2_\alpha\right)\right] \\ &= -\bar{d}\, \mathrm{E}\left[(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\, I\left(F < \frac{m}{q}\{e^{\chi^2_\alpha/n} - 1\}\right)\right] \\ &= -\bar{d}C^{-1}H'(HC^{-1}H')^{-1}\, \mathrm{E}\Bigg[(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h}) \\ &\qquad \times I\left(\frac{(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})}{q\hat{\sigma}^2} < \frac{m}{q}\left\{e^{\chi^2_\alpha/n} - 1\right\}\right)\Bigg] \\ &= -\bar{d}\,\sigma C^{-1}H'(HC^{-1}H')^{-\frac{1}{2}}\mathrm{E}\left[\boldsymbol{Z}\, I\left(\frac{m\boldsymbol{Z}'\boldsymbol{Z}}{q\chi^2_m} < \frac{m}{q}\left\{e^{\chi^2_\alpha/n} - 1\right\}\right)\right]. \end{aligned} \tag{4.4.20}$$

Applying Theorem 1, Appendix B2, Judge and Bock (1978), to the right hand side of (4.4.20), the bias function of the SPTE of $\boldsymbol{\beta}$ under the LR test is obtained as

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}_{\mathrm{LR}}^{\mathrm{SPTE}}\right] = \boldsymbol{\eta}\,\bar{d}\,G_{q+2,\,m}\left(r_2^{\mathrm{LR}};\Delta\right). \tag{4.4.21}$$

This completes the proof of (4.4.17) of the theorem.

Finally, we prove (4.4.18) for the bias function of the SPTE of $\boldsymbol{\beta}$ under the LM test.

By definition the bias function of the SPTE of $\boldsymbol{\beta}$ under the LM test is

$$\begin{aligned}
\mathrm{B}\left[\hat{\boldsymbol{\beta}}_{\mathrm{LM}}^{\mathrm{SPTE}}\right] &= \mathrm{E}\left[\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta} - \bar{d}\,(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\,I\left(\xi_{\mathrm{LM}} < \chi_\alpha^2\right)\right] \\
&= -\bar{d}\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\,I\left(\xi_{\mathrm{LM}} < \chi_\alpha^2\right)\right] \\
&= -\bar{d}\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\,I\left(\frac{nqF}{m+qF} < \chi_\alpha^2\right)\right] \\
&= -\bar{d}\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\,I\left(F < \frac{m\chi_\alpha^2}{q(n-\chi_\alpha^2)}\right)\right] \\
&= -\bar{d}\,C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\Bigg[(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h}) \\
&\qquad \times I\left(\frac{(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})}{q\hat{\sigma}^2} < \frac{m\chi_\alpha^2}{q(n-\chi_\alpha^2)}\right)\Bigg] \\
&= -\bar{d}\sigma C^{-1}H'(HC^{-1}H')^{-\frac{1}{2}}\mathrm{E}\left[\boldsymbol{Z}\,I\left(\frac{m\boldsymbol{Z}'\boldsymbol{Z}}{q\chi_m^2} < \frac{m\chi_\alpha^2}{q(n-\chi_\alpha^2)}\right)\right].
\end{aligned} \tag{4.4.22}$$

Applying Theorem 1, Appendix B2, Judge and Bock (1978), to the right hand side of (4.4.20), the bias function of the SPTE of $\boldsymbol{\beta}$ under the LM test is obtained as

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}_{\mathrm{LM}}^{\mathrm{SPTE}}\right] = \boldsymbol{\eta}\,\bar{d}\,G_{q+2,\,m}\left(r_2^{\mathrm{LM}};\Delta\right). \tag{4.4.23}$$

This completes the proof of the statement of the theorem.

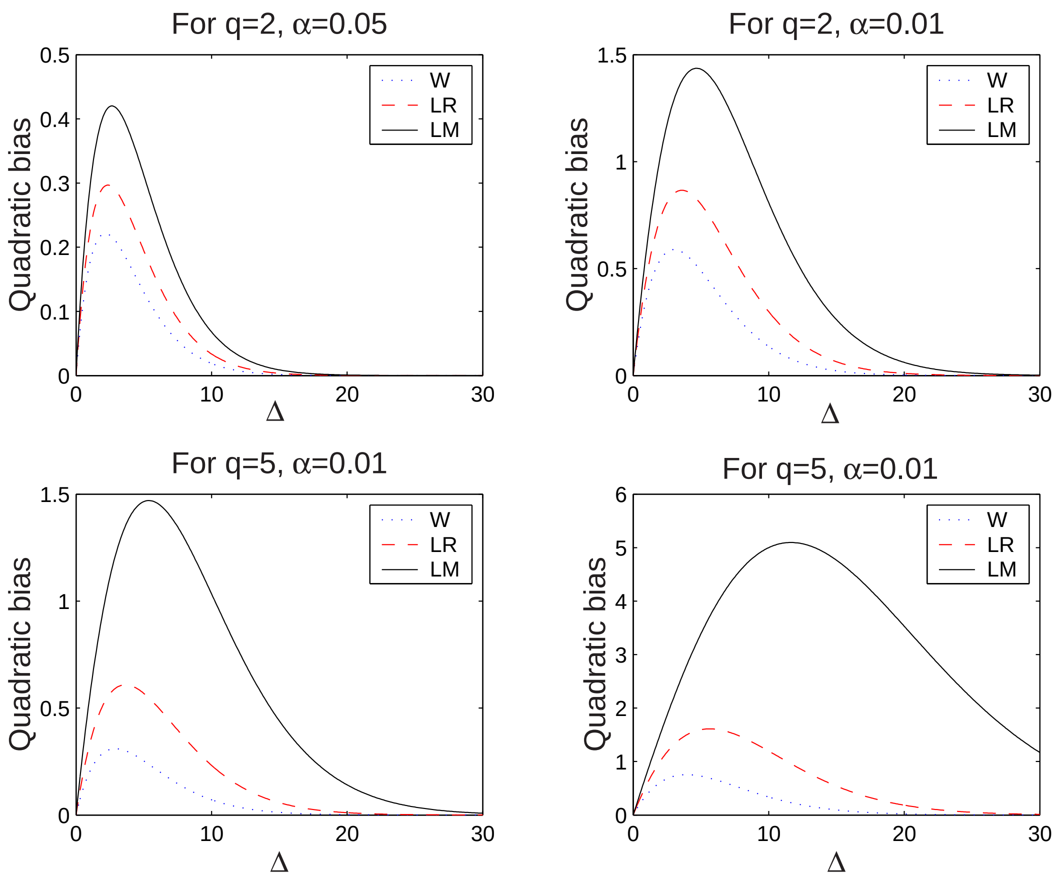

Figure 4.1: The QBs of the SPTEs under the original W, LR and LM tests for $n = 25$, $d = 0.1$, $p = 8$, and selected values of q and α.

Theorem 4.14. *The quadratic bias functions of the SPTEs of* $\boldsymbol{\beta}$ *under the original W, LR and LM tests are*

$$\text{QB}\left[\hat{\boldsymbol{\beta}}_{\text{W}}^{\text{SPTE}}\right] = \boldsymbol{\eta}'\boldsymbol{\eta}\,\bar{d}^2\,G^2_{q+2,m}\,(r_2^{\text{W}};\,\Delta)\,, \tag{4.4.24}$$

$$\text{QB}\left[\hat{\boldsymbol{\beta}}_{\text{LR}}^{\text{SPTE}}\right] = \boldsymbol{\eta}'\boldsymbol{\eta}\,\bar{d}^2\,G^2_{q+2,m}\,(r_2^{\text{LR}};\,\Delta) \tag{4.4.25}$$

and

$$\text{QB}\left[\hat{\boldsymbol{\beta}}_{\text{LM}}^{\text{SPTE}}\right] = \boldsymbol{\eta}'\boldsymbol{\eta}\,\bar{d}^2\,G^2_{q+2,m}\,(r_2^{\text{LM}};\,\Delta) \tag{4.4.26}$$

respectively.

Proof. Using the definition of the quadratic bias in (4.2.9), the proof of the theorem is straightforward.

Figure 4.1 displays the graph of the QB functions of the SPTEs of $\boldsymbol{\beta}$ under the original W, LR and LM tests against Δ.

Theorem 4.15. *The quadratic risk functions of the SPTEs of* $\boldsymbol{\beta}$ *under the original W, LR and LM tests are*

$$\begin{aligned} \mathrm{R}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SPTE}}\Big] &= p - qd^* G_{q+2,\,m}\left(r_2^{\mathrm{W}};\Delta\right) + 2\bar{d}\Delta\, G_{q+2,\,m}\left(r_2^{\mathrm{W}};\Delta\right) \\ &\quad - d^*\Delta G_{q+4,\,m}\left(r_4^{\mathrm{W}};\Delta\right), \qquad (4.4.27) \\ \mathrm{R}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{LR}}^{\mathrm{SPTE}}\Big] &= p - qd^* G_{q+2,m}\left(r_2^{\mathrm{LR}};\Delta\right) + 2\bar{d}\Delta G_{q+2,\,m}\left(r_2^{\mathrm{LR}};\Delta\right) \\ &\quad - d^*\Delta G_{q+4,\,m}\left(r_4^{\mathrm{LR}};\Delta\right) \qquad (4.4.28) \end{aligned}$$

and

$$\begin{aligned} \mathrm{R}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{LM}}^{\mathrm{SPTE}}\Big] &= p - qd^* G_{q+2,\,m}\left(r_2^{\mathrm{LM}};\Delta\right) + 2\bar{d}\Delta\, G_{q+2,m}\left(r_2^{\mathrm{LM}};\Delta\right) \\ &\quad - d^*\Delta G_{q+4,\,m}\left(r_4^{\mathrm{LM}};\Delta\right) \qquad (4.4.29) \end{aligned}$$

respectively, where $d^* = (1-d^2)$; $r_i^{\mathrm{W}} = \frac{m\chi_\alpha^2}{n(q+i)}$, $r_i^{\mathrm{LR}} = \frac{m}{(q+i)}\Big(e^{\frac{\chi_\alpha^2}{n}} - 1\Big)$, $r_i^{\mathrm{LM}} = \frac{m\chi_\alpha^2}{(q+i)\left(n-\chi_\alpha^2\right)}$, $i = 2,\, 4$; *and* $G_{a,b}(r;\,\Delta)$ *is the cumulative distribution function of the non-central* F *distribution with* $(a,\, b)$ *d.f., non-centrality parameter* Δ *and is evaluated at* r*.*

Proof. By definition, the mse matrix of the SPTE of $\boldsymbol{\beta}$ under the W test is

$$\begin{aligned} \mathrm{M}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SPTE}}\Big] &= \mathrm{E}\Big[\Big(\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SPTE}} - \boldsymbol{\beta}\Big)\Big(\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SPTE}} - \boldsymbol{\beta}\Big)'\Big] \\ &= \mathrm{E}\Big[\Big(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\Big) - \bar{d}C^{-1}H'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})\, I(\xi_{\mathrm{W}} < \chi_\alpha^2)\Big] \\ &\quad \times \Big[\Big(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\Big) - \bar{d}C^{-1}H'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})\, I(\xi_{\mathrm{W}} < \chi_\alpha^2)\Big]' \end{aligned}$$

$$= \mathrm{E}\Big[\left(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)\left(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)'\Big] + \bar{d}^2\, \mathrm{E}\Big[C^{-1}H'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})$$
$$\times(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\, I(\xi_{\mathrm{W}} < \chi^2_\alpha)\Big]$$
$$- 2\bar{d}\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\, I(\xi_{\mathrm{W}} < \chi^2_\alpha)\Big]. \tag{4.4.30}$$

The first term in the right hand side of (4.4.30) is

$$\mathrm{E}\Big[\left(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)\left(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)'\Big] = \sigma^2 C^{-1}, \tag{4.4.31}$$

the mse matrix of the unrestricted estimator of $\boldsymbol{\beta}$.

The second term in the right hand side of (4.4.30) is

$$\bar{d}^2 C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'\, I(\xi_{\mathrm{W}} < \chi^2_\alpha)\Big]$$
$$\times (HC^{-1}H')^{-1}HC^{-1}$$
$$= \bar{d}^2\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\left[\boldsymbol{Z}\boldsymbol{Z}'\, I\left(\frac{m\boldsymbol{Z}'\boldsymbol{Z}}{q\chi^2_m} < \frac{m\chi^2_\alpha}{nq}\right)\right] HC^{-1}. \tag{4.4.32}$$

Applying Theorem 3, Appendix B2, Judge and Bock (1978), to (4.4.32), we get

$$\bar{d}^2 C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'\, I(\xi_{\mathrm{W}} < \chi^2_\alpha)\Big]$$
$$\times (HC^{-1}H')^{-1}HC^{-1}$$
$$= \bar{d}^2\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\Big[G_{q+2,\,m}(r_2^{\mathrm{W}};\, \Delta) + \sigma^{-2}(HC^{-1}H')^{-1/2}$$
$$\times (H\boldsymbol{\beta} - \boldsymbol{h})(H\boldsymbol{\beta} - \boldsymbol{h})'(HC^{-1}H')^{-1/2}G_{q+4,\,m}(r_4^{\mathrm{W}};\, \Delta)\Big]HC^{-1}$$
$$= \bar{d}^2\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\,G_{q+2,\,m}(r_2^{\mathrm{W}};\, \Delta) + \bar{d}^2 C^{-1}H'(HC^{-1}H')^{-1}$$
$$\times (H\boldsymbol{\beta} - \boldsymbol{h})(H\boldsymbol{\beta} - \boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\,G_{q+4,\,m}(r_4^{\mathrm{W}};\, \Delta). \tag{4.4.33}$$

Therefore, the second term in the right hand side of (4.4.30) becomes

$$\bar{d}^2\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\,G_{q+2,\,m}(r_2^{\mathrm{W}};\, \Delta) + \bar{d}^2\boldsymbol{\eta}\boldsymbol{\eta}'\,G_{q+4,\,m}(r_4^{\mathrm{W}};\, \Delta). \tag{4.4.34}$$

The third term in the right hand side of (4.4.30) is

$$
\begin{aligned}
&2\bar{d}\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\,I(\xi_{\mathrm{W}}<\chi^2_\alpha)\Big]\\
&=2\bar{d}\,\mathrm{E}\Big[E\left\{(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})/(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})\right\}(H\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})'\,I(\xi_{\mathrm{W}}<\chi^2_\alpha)\Big]\,(HC^{-1}H')^{-1}HC^{-1}\\
&=2\bar{d}\,\mathrm{E}\Big[\sigma^2C^{-1}H'(\sigma^2HC^{-1}H')^{-1}\left\{(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})-(H\boldsymbol{\beta}-\boldsymbol{h})\right\}\\
&\quad\times(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'\,I(\xi_{\mathrm{W}}<\chi^2_\alpha)\Big]\,(HC^{-1}H')^{-1}HC^{-1}\\
&=2\bar{d}C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'\,I(\xi_{\mathrm{W}}<\chi^2_\alpha)\Big]\\
&\quad\times(HC^{-1}H')^{-1}HC^{-1}-2dC^{-1}H'(HC^{-1}H')^{-1}(H\boldsymbol{\beta}-\boldsymbol{h})\\
&\quad\times\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'\,I(\xi_{\mathrm{W}}<\chi^2_\alpha)\Big]\,(HC^{-1}H')^{-1}HC^{-1}.
\end{aligned}
\tag{4.4.35}
$$

Using the expressions (4.4.34) and (4.4.19) in (4.4.35), the last term of (4.4.30) is obtained as

$$
\begin{aligned}
&2d\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\,I(\xi_{\mathrm{W}}<\chi^2_\alpha)\Big]\\
&=2\bar{d}\sigma^2C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\,G_{q+2,\,m}(r_2^{\mathrm{W}};\,\Delta)+2\bar{d}\boldsymbol{\eta}\boldsymbol{\eta}'\,G_{q+4,\,m}(r_4^{\mathrm{W}};\,\Delta)\\
&\quad-2\bar{d}c^{-1}H'(HC^{-1}H')^{-1}(H\boldsymbol{\beta}-\boldsymbol{h})(H\boldsymbol{\beta}-\boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\\
&\quad\quad\times G_{q+2,\,m}(r_2^{\mathrm{W}};\,\Delta)\\
&=2\bar{d}\sigma^2C^{-1}H(HC^{-1}H')^{-1}HC^{-1}\,G_{q+2,\,m}(r_2;\Delta)+2\bar{d}\boldsymbol{\eta}\boldsymbol{\eta}'\,G_{q+4,\,m}(r_2^{\mathrm{W}};\,\Delta)\\
&\quad\quad-2\bar{d}\boldsymbol{\eta}\boldsymbol{\eta}'\,G_{q+2,\,m}(r_2^{\mathrm{W}};\,\Delta).
\end{aligned}
\tag{4.4.36}
$$

Collecting the expressions of the three terms of (4.4.30), the mse matrix of the SPTE of $\boldsymbol{\beta}$ under the W test is obtained as

$$
\begin{aligned}
\mathrm{M}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SPTE}}\Big]&=\sigma^2C^{-1}-d^*\sigma^2C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\,G_{q+2,\,m}(r_2^{\mathrm{W}};\,\Delta)\\
&\quad+2\bar{d}\boldsymbol{\eta}\boldsymbol{\eta}'\,G_{q+2,\,m}(r_2^{\mathrm{W}};\,\Delta)-d^*\boldsymbol{\eta}\boldsymbol{\eta}'G_{q+4,\,m}(r_4^{\mathrm{W}};\,\Delta)
\end{aligned}
\tag{4.4.37}
$$

where $d^*=1-d^2$.

By definition, the quadratic risk function of the SPTE of $\boldsymbol{\beta}$ under the W test is

$$\begin{aligned}
\mathrm{R}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SPTE}}\Big] &= \mathrm{tr}\Big[\mathcal{W} \times \mathrm{M}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SPTE}}\Big]\Big] \\
&= \mathrm{tr}\Big[\sigma^{-2}C \times \mathrm{M}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SPTE}}\Big]\Big] \\
&= \mathrm{tr}\Big[\sigma^{-2}C\Big\{\sigma^2C^{-1} - d^*\sigma^2C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\,G_{q+2,\,m}(r_2^{\mathrm{W}};\,\Delta) \\
&\quad - 2\bar{d}\boldsymbol{\eta}\boldsymbol{\eta}'\,G_{q+2,\,m}(r_2^{\mathrm{W}};\,\Delta) - d^*\boldsymbol{\eta}\boldsymbol{\eta}'G_{q+4,\,m}(r_4^{\mathrm{W}};\,\Delta)\Big\}\Big] \\
&= p - qd^*G_{q+2,\,m}\,(r_2^{\mathrm{W}};\Delta) + 2\bar{d}\Delta\,G_{q+2,\,m}\,(r_2^{\mathrm{W}};\Delta) - d^*\Delta G_{q+4,\,m}\,(r_4^{\mathrm{W}};\Delta)\,.
\end{aligned} \tag{4.4.38}$$

This completes the proof of (4.A.17) of the theorem.

By definition, the mse matrix of the SPTE of $\boldsymbol{\beta}$ under the LR test is

$$\begin{aligned}
\mathrm{M}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{LR}}^{\mathrm{SPTE}}\Big] &= \Big[\Big(\hat{\boldsymbol{\beta}}_{\mathrm{LR}}^{\mathrm{SPTE}} - \boldsymbol{\beta}\Big)\Big(\hat{\boldsymbol{\beta}}_{\mathrm{LR}}^{\mathrm{SPTE}} - \boldsymbol{\beta}\Big)'\Big] \\
&= \mathrm{E}\Big[\Big(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\Big) - \bar{d}C^{-1}H'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})\,I(\xi_{\mathrm{LR}} < \chi^2_\alpha)\Big] \\
&\quad \times \Big[\Big(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\Big) - \bar{d}C^{-1}H'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})\,I(\xi_{\mathrm{LR}} < \chi^2_\alpha)\Big]' \\
&= \sigma^2C^{-1} + d^2\,\mathrm{E}\Big[C^{-1}H'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})' \\
&\quad \times (HC^{-1}H')^{-1}HC^{-1}\,I(\xi_{\mathrm{LR}} < \chi^2_\alpha)\Big] - 2\bar{d}\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})' \\
&\quad \times (HC^{-1}H')^{-1}HC^{-1}\,I(\xi_{\mathrm{LR}} < \chi^2_\alpha)\Big]\,.
\end{aligned} \tag{4.4.39}$$

The second term in the right hand side of (4.4.39) is

$$\begin{aligned}
&\bar{d}^2C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'\,I(\xi_{\mathrm{LR}} < \chi^2_\alpha)\Big] \\
&\qquad \times (HC^{-1}H')^{-1}HC^{-1} \\
&= \bar{d}^2\sigma^2C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\left[\boldsymbol{Z}\boldsymbol{Z}'\,I\left(\frac{m\boldsymbol{Z}'\boldsymbol{Z}}{q\chi^2_m} < \frac{m\{e^{\chi^2_\alpha/n} - 1\}}{q}\right)\right]HC^{-1}.
\end{aligned} \tag{4.4.40}$$

Applying Theorem 3, Appendix B2, Judge and Bock (1978), to (4.4.40) we get

$$
\begin{aligned}
&\bar{d}^2 C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'\,I(\xi_{\mathrm{LR}}<\chi^2_\alpha)\Big]\,(HC^{-1}H')^{-1}HC^{-1}\\
&=\bar{d}^2\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\Big[G_{q+2,\,m}(r_2^{\mathrm{LR}};\Delta)+\sigma^{-2}(HC^{-1}H')^{-1/2}\\
&\quad\times(H\boldsymbol{\beta}-\boldsymbol{h})(H\boldsymbol{\beta}-\boldsymbol{h})'(HC^{-1}H')^{-1/2}G_{q+4,\,m}(r_4^{\mathrm{LR}};\Delta)\Big]HC^{-1}\\
&=\bar{d}^2\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\,G_{q+2,\,m}(r_2^{\mathrm{LR}};\Delta)+\bar{d}^2C^{-1}H'(HC^{-1}H')^{-1}\\
&\quad\times(H\boldsymbol{\beta}-\boldsymbol{h})(H\boldsymbol{\beta}-\boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\,G_{q+4,\,m}(r_4^{\mathrm{LR}};\Delta)\\
&=\bar{d}^2\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\,G_{q+2,\,m}(r_2^{\mathrm{LR}};\Delta)+d^2\boldsymbol{\eta}\boldsymbol{\eta}'\,G_{q+4,\,m}(r_4^{\mathrm{LR}};\Delta).
\end{aligned}
\tag{4.4.41}
$$

The third term in the right hand side of (4.4.39) is

$$
\begin{aligned}
&2\bar{d}\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\,I(\xi_{\mathrm{LR}}<\chi^2_\alpha)\Big]\\
&=2\bar{d}\,\mathrm{E}\Big[\mathrm{E}\Big\{(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})/(\mathrm{H}\tilde{\boldsymbol{\beta}}-\boldsymbol{h})\Big\}\,(\mathrm{H}\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})'\,\mathrm{I}(\xi_{\mathrm{LR}}<\chi^2_\alpha)\Big]\,(HC^{-1}H')^{-1}HC^{-1}\\
&=2\bar{d}\,\mathrm{E}\Big[\sigma^2C^{-1}H'(\sigma^2HC^{-1}H')^{-1}\Big\{(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})-(H\boldsymbol{\beta}-\boldsymbol{h})\Big\}\,(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'\\
&\quad\times I(\xi_{\mathrm{LR}}<\chi^2_\alpha)\Big](HC^{-1}H')^{-1}HC^{-1}\\
&=2\bar{d}C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'\,I(\xi_{\mathrm{W}}<\chi^2_\alpha)\Big]\,(HC^{-1}H')^{-1}\\
&\quad\times HC^{-1}-2\bar{d}C^{-1}H'(HC^{-1}H')^{-1}(H\boldsymbol{\beta}-\boldsymbol{h})\,\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'\,I(\xi_{\mathrm{LR}}<\chi^2_\alpha)\Big]\\
&\qquad\times(HC^{-1}H')^{-1}HC^{-1}.
\end{aligned}
\tag{4.4.42}
$$

Collecting the expressions of the two terms of the right hand side of (4.4.42) from (4.4.41) and (4.4.20), the last term of (4.4.39) is obtained as

$$
\begin{aligned}
&2\bar{d}\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\,I(\xi_{\mathrm{LR}}<\chi^2_\alpha)\Big]\\
&=2\bar{d}\sigma^2C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\,G_{q+2,\,m}(r_2^{\mathrm{LR}};\Delta)+2\bar{d}\boldsymbol{\eta}\boldsymbol{\eta}'\,G_{q+4,\,m}(r_4^{\mathrm{LR}};\Delta)\\
&\quad-2\bar{d}c^{-1}H'(HC^{-1}H')^{-1}(H\boldsymbol{\beta}-\boldsymbol{h})(H\boldsymbol{\beta}-\boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\\
&\quad\times G_{q+2,\,m}(r_2^{\mathrm{LR}};\Delta)
\end{aligned}
$$

$$= 2\bar{d}\sigma^2 C^{-1} H(HC^{-1}H')^{-1} HC^{-1}\, G_{q+2,\,m}(r_2^{\text{LR}};\Delta) + 2\bar{d}\boldsymbol{\eta}\boldsymbol{\eta}'\, G_{q+4,\,m}(r_2^{\text{LR}};\Delta)$$
$$- 2\bar{d}\boldsymbol{\eta}\boldsymbol{\eta}'\, G_{q+2,\,m}(r_2^{\text{LR}};\Delta). \qquad (4.4.43)$$

Collecting the expressions of the three terms of (4.4.39), the mse matrix of the SPTE of $\boldsymbol{\beta}$ under the LR test is obtained as

$$\text{M}\Big[\hat{\boldsymbol{\beta}}_{\text{LR}}^{\text{SPTE}}\Big] = \sigma^2 C^{-1} - d^*\sigma^2 C^{-1} H'(HC^{-1}H')^{-1} HC^{-1}\, G_{q+2,\,m}(r_2^{\text{LR}};\Delta)$$
$$+ 2\bar{d}\boldsymbol{\eta}\boldsymbol{\eta}'\, G_{q+2,\,m}(r_2^{\text{LR}};\Delta) - d^*\boldsymbol{\eta}\boldsymbol{\eta}' G_{q+4,\,m}(r_4^{\text{LR}};\Delta). \qquad (4.4.44)$$

By definition, the quadratic risk function of the SPTE of $\boldsymbol{\beta}$ under the LR test is

$$\begin{aligned}
\text{R}\Big[\hat{\boldsymbol{\beta}}_{\text{LR}}^{\text{SPTE}}\Big] &= \text{tr}\Big[\mathcal{W} \times \text{M}\Big[\hat{\boldsymbol{\beta}}_{\text{LR}}^{\text{SPTE}}\Big]\Big] \\
&= \text{tr}\Big[\sigma^{-2} C \times \text{M}\Big[\hat{\boldsymbol{\beta}}_{\text{LR}}^{\text{SPTE}}\Big]\Big] \\
&= \text{tr}\Big[\sigma^{-2} C \Big\{\sigma^2 C^{-1} - d^*\sigma^2 C^{-1} H'(HC^{-1}H')^{-1} HC^{-1}\, G_{q+2,\,m}(r_2^{\text{LR}};\ \Delta) \\
&\quad - 2d\boldsymbol{\eta}\boldsymbol{\eta}'\, G_{q+2,\,m}(r_2^{\text{LR}};\ \Delta) - d^*\boldsymbol{\eta}\boldsymbol{\eta}' G_{q+4,\,m}(r_4^{\text{LR}};\ \Delta)\Big\}\Big] \\
&= p - qd^* G_{q+2,\,m}\,(r_2^{\text{LR}};\Delta) + 2\bar{d}\Delta\, G_{q+2,\,m}\,(r_2^{\text{LR}};\Delta) \\
&\quad - d^*\Delta G_{q+4,\,m}\,(r_4^{\text{LR}};\Delta)\,. \qquad (4.4.45)
\end{aligned}$$

This completes the proof of (4.4.28) of the theorem.

Finally, we derive the quadratic risk function of the SPTE of $\boldsymbol{\beta}$ under the LM test.

By definition, the mse matrix of the SPTE of $\boldsymbol{\beta}$ under the LM test is

$$\begin{aligned}
\text{M}\Big[\hat{\boldsymbol{\beta}}_{\text{LM}}^{\text{SPTE}}\Big] &= \text{E}\Big[\Big(\hat{\boldsymbol{\beta}}_{\text{LM}}^{\text{SPTE}} - \boldsymbol{\beta}\Big)\Big(\hat{\boldsymbol{\beta}}_{\text{LM}}^{\text{SPTE}} - \boldsymbol{\beta}\Big)'\Big] \\
&= \text{E}\Big[\Big(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\Big) - \bar{d}C^{-1}H'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})\, I(\xi_{\text{LM}} < \chi_\alpha^2)\Big] \\
&\quad \times \Big[\Big(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\Big) - \bar{d}C^{-1}H'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})\, I(\xi_{\text{LM}} < \chi_\alpha^2)\Big]'
\end{aligned}$$

$$= \sigma^2 C^{-1} + d^2\,\mathrm{E}\Big[C^{-1}H'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h}) \times (H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\,I(\xi_{\mathrm{LM}} < \chi^2_\alpha)\Big] - 2\bar{d}\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\,I(\xi_{\mathrm{LM}} < \chi^2_\alpha)\Big]. \tag{4.4.46}$$

The second term of the right hand side of (4.4.46) is

$$\bar{d}^2 C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'\,I(\xi_{\mathrm{LM}} < \chi^2_\alpha)\Big]\,(HC^{-1}H')^{-1}HC^{-1} = \bar{d}^2\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\left[\boldsymbol{Z}\boldsymbol{Z}'\,I\left(\frac{m\boldsymbol{Z}'\boldsymbol{Z}}{q\chi^2_m} < \frac{m\{e^{\chi^2_\alpha/n} - 1\}}{q}\right)\right] HC^{-1}. \tag{4.4.47}$$

Applying Theorem 3, Appendix B2, Judge and Bock (1978), to (4.4.47), we get

$$\begin{aligned}
&\bar{d}^2 C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'\,I(\xi_{\mathrm{LM}} < \chi^2_\alpha)\Big]\,(HC^{-1}H')^{-1}HC^{-1} \\
&= \bar{d}^2\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\Big[G_{q+2,\,m}(r_2^{\mathrm{LM}};\Delta) + \sigma^{-2}(HC^{-1}H')^{-1/2} \\
&\quad \times (H\boldsymbol{\beta} - \boldsymbol{h})(H\boldsymbol{\beta} - \boldsymbol{h})'(HC^{-1}H')^{-1/2}G_{q+4,\,m}(r_4^{\mathrm{LM}};\Delta)\Big]HC^{-1} \\
&= \bar{d}^2\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\,G_{q+2,\,m}(r_2^{\mathrm{LM}};\Delta) + d^2 C^{-1}H'(HC^{-1}H')^{-1} \\
&\quad \times (H\boldsymbol{\beta} - \boldsymbol{h})(H\boldsymbol{\beta} - \boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\,G_{q+4,\,m}(r_4^{\mathrm{LM}};\Delta) \\
&= \bar{d}^2\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\,G_{q+2,\,m}(r_2^{\mathrm{LM}};\Delta) + \bar{d}^2\boldsymbol{\eta}\boldsymbol{\eta}'\,G_{q+4,\,m}(r_4^{\mathrm{LM}};\Delta).
\end{aligned} \tag{4.4.48}$$

The third term in the right hand side of (4.4.46) is

$$\begin{aligned}
&2\bar{d}\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\,I(\xi_{\mathrm{LM}} < \chi^2_\alpha)\Big] \\
&= 2\bar{d}\,\mathrm{E}\Big[\mathrm{E}\Big\{(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})/(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})\Big\}\,(H\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})'\,I(\xi_{\mathrm{LM}} < \chi^2_\alpha)\Big]\,(HC^{-1}H')^{-1}HC^{-1} \\
&= 2\bar{d}\,\mathrm{E}\Big[\sigma^2 C^{-1}H'(\sigma^2 HC^{-1}H')^{-1}\Big\{(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h}) - (H\boldsymbol{\beta} - \boldsymbol{h})\Big\}\,(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})' \\
&\quad \times I(\xi_{\mathrm{LM}} < \chi^2_\alpha)\Big](HC^{-1}H')^{-1}HC^{-1}
\end{aligned}$$

$$= 2\bar{d}C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'\, I(\xi_{\text{LM}} < \chi^2_\alpha)\Big]\,(HC^{-1}H')^{-1}$$
$$\times\, HC^{-1} - 2\bar{d}C^{-1}H'(HC^{-1}H')^{-1}(H\boldsymbol{\beta} - \boldsymbol{h})\,\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'\, I(\xi_{\text{LM}} < \chi^2_\alpha)\Big]$$
$$\times\,(HC^{-1}H')^{-1}HC^{-1}. \tag{4.4.49}$$

Collecting the expressions of the two terms of the right hand side of (4.4.49) from (4.4.48) and (4.4.23), the last term of (4.4.46) is obtained as

$$2\bar{d}\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\, I(\xi_{\text{LM}} < \chi^2_\alpha)\Big]$$
$$= 2\bar{d}\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\, G_{q+2,\,m}(r_2^{\text{LM}};\Delta) + 2\bar{d}\boldsymbol{\eta}\boldsymbol{\eta}'\, G_{q+4,\,m}(r_4^{\text{LM}};\Delta)$$
$$- 2\bar{d}c^{-1}H'(HC^{-1}H')^{-1}(H\boldsymbol{\beta} - \boldsymbol{h})(H\boldsymbol{\beta} - \boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}$$
$$\times\, G_{q+2,\,m}(r_2^{\text{LM}};\Delta)$$
$$= 2\bar{d}\sigma^2 C^{-1}H(HC^{-1}H')^{-1}HC^{-1}\, G_{q+2,\,m}(r_2^{\text{LM}};\Delta) + 2\bar{d}\boldsymbol{\eta}\boldsymbol{\eta}'\, G_{q+4,\,m}(r_2^{\text{LM}};\Delta)$$
$$- 2\bar{d}\boldsymbol{\eta}\boldsymbol{\eta}'\, G_{q+2,\,m}(r_2^{\text{LM}};\Delta). \tag{4.4.50}$$

Collecting the expressions of the three terms of (4.4.46), the mse matrix of the SPTE of $\boldsymbol{\beta}$ under the LM test is obtained as

$$\mathrm{M}\Big[\hat{\boldsymbol{\beta}}^{\text{SPTE}}_{\text{LM}}\Big] = \sigma^2 C^{-1} - d^*\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\, G_{q+2,\,m}(r_2^{\text{LM}};\Delta)$$
$$+ 2\bar{d}\boldsymbol{\eta}\boldsymbol{\eta}'\, G_{q+2,\,m}(r_2^{\text{LM}};\Delta) - d^*\boldsymbol{\eta}\boldsymbol{\eta}' G_{q+4,\,m}(r_4^{\text{LM}};\Delta). \tag{4.4.51}$$

By definition, the quadratic risk function of the SPTE of $\boldsymbol{\beta}$ under the LM test is

$$\mathrm{R}\Big[\hat{\boldsymbol{\beta}}^{\text{SPTE}}_{\text{LM}}\Big] = \mathrm{tr}\Big[\mathcal{W} \times \mathrm{M}\Big[\hat{\boldsymbol{\beta}}^{\text{SPTE}}_{\text{LM}}\Big]\Big]$$
$$= \mathrm{tr}\Big[\sigma^{-2}C \times \mathrm{M}\Big[\hat{\boldsymbol{\beta}}^{\text{SPTE}}_{\text{LM}}\Big]\Big]$$
$$= \mathrm{tr}\Big[\sigma^{-2}C\Big\{\sigma^2 C^{-1} - d^*\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\, G_{q+2,\,m}(r_2^{\text{LM}};\,\Delta)$$
$$+ 2\bar{d}\boldsymbol{\eta}\boldsymbol{\eta}'\, G_{q+2,\,m}(r_2^{\text{LM}};\,\Delta) - d^*\boldsymbol{\eta}\boldsymbol{\eta}' G_{q+4,\,m}(r_4^{\text{LM}};\,\Delta)\Big\}\Big]$$

$$= p - qd^* G_{q+2,\,m}\left(r_2^{\text{LM}}; \Delta\right) + 2\bar{d}\Delta\, G_{q+2,\,m}\left(r_2^{\text{LM}}; \Delta\right)$$
$$- d^* \Delta G_{q+4,\,m}\left(r_4^{\text{LM}}; \Delta\right). \tag{4.4.52}$$

This completes the proof of the theorem.

Figure 4.2 displays the efficiencies of the SPTEs of $\boldsymbol{\beta}$ relative to the UE, against Δ.

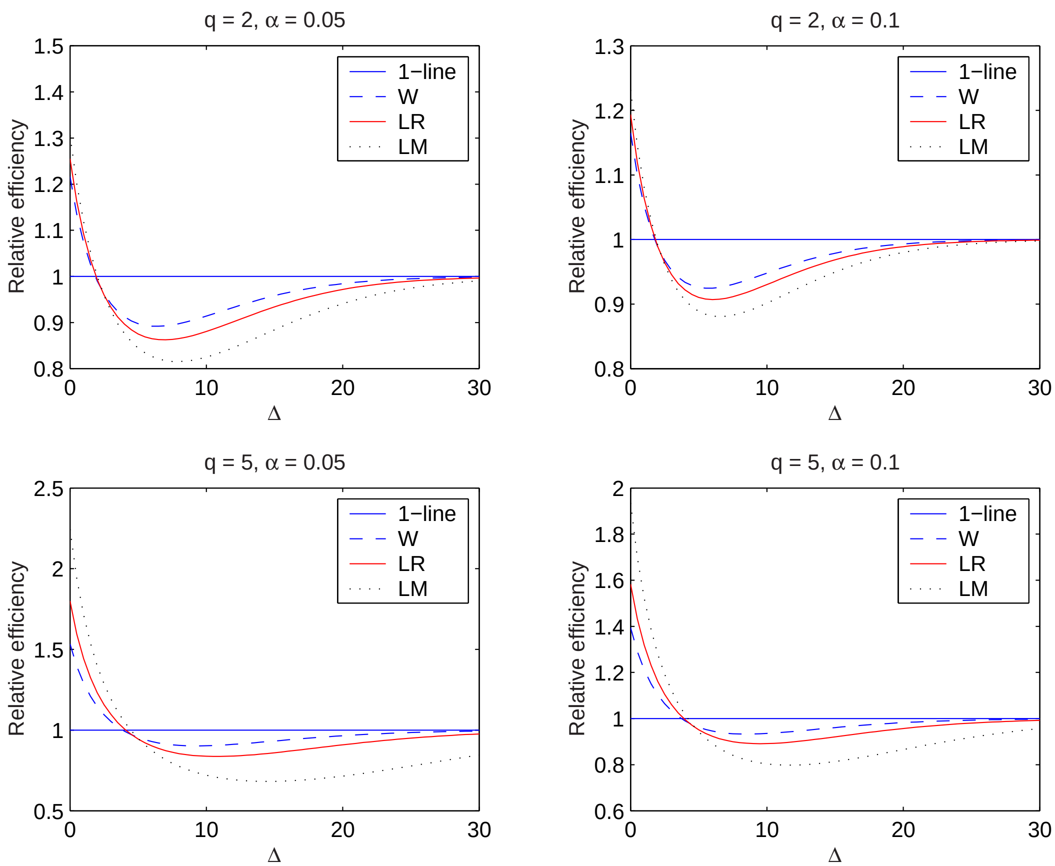

Figure 4.2: Efficiencies of the SPTEs under the original W, LR and LM tests relative to the UE for $n = 25$, $d = 0.1$, $p = 8$, and selected values of q and α.

4.5 The Modified W, LR and LM Tests and the SPTE of $\boldsymbol{\beta}$

To test the null hypothesis in (4.1.2), the modified W, LR and LM test statistics are

$$\xi^*_{\mathrm{W}} = qF, \tag{4.5.1}$$

$$\xi^*_{\mathrm{LR}} = (m + \frac{q}{2} - 1)\,\ln\left(1 + \frac{qF}{m}\right) \tag{4.5.2}$$

and

$$\xi^*_{\mathrm{LM}} = \frac{(m+q)qF}{m+qF} \tag{4.5.3}$$

respectively. The modified W statistic is obtained by replacing n by m in (4.4.10) and the modified LM statistic by replacing n by $(m+q)$ in (4.4.12). This degrees of freedom correction corrects the bias of the estimator of the error variance σ^2. The modified LR test statistic is obtained by replacing n by $m + (q/2) - 1$. This correction to the LR statistic is the Edgeworth-size correction which ensures that the LR test has the correct significance level to order $1/m$ (cf. Evans and Savin, 1982). The inequality relation, $\xi_{\mathrm{W}} \geq \xi_{\mathrm{LR}} \geq \xi_{\mathrm{LM}}$, that holds for the original test statistics (see Appendix for the proof), does not hold for the modified test statistics for all m and q (cf. Evans and Savin, 1982).

By definition, the SPTEs of $\boldsymbol{\beta}$ under the modified tests will differ from those of $\boldsymbol{\beta}$ under the original tests only with respect to the corresponding indicator function $I(\cdot)$. The indicator functions involved with the definitions of SPTEs of $\boldsymbol{\beta}$ under the three modified tests are as follows.

$$I_{\mathrm{W}^*} = I\big(qF < \chi^2_\alpha\big) = I\left(F < \frac{\chi^2_\alpha}{q}\right), \tag{4.5.4}$$

$$I_{\mathrm{LR}^*} = I\left(\left\{\left(m + \frac{q}{2} - 1\right) \ln\left(1 + \frac{qF}{m}\right)\right\} < \chi^2_\alpha\right) = I\left(F < \left\{\frac{m}{q}(e^{\chi^2_\alpha/(m+\frac{q}{2}-1)} - 1)\right\}\right) \tag{4.5.5}$$

and

$$I_{\mathrm{LM}^*} = I\left(\frac{(m+q)qF}{m+qF} < \chi^2_\alpha\right) = I\left(F < \frac{m\chi^2_\alpha}{q(m+q-\chi^2_\alpha)}\right). \tag{4.5.6}$$

The derivation of the bias and quadratic risk functions of the SPTEs of $\boldsymbol{\beta}$ under the modified tests is straightforward. The bias, quadratic bias and quadratic risk functions of the SPTEs of $\boldsymbol{\beta}$ under the modified tests are stated in Theorems 4.16. – 4.18. respectively.

Theorem 4.16. *The bias functions of the SPTEs of $\boldsymbol{\beta}$ under the modified W, LR and LM tests are respectively*

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{W}^*}\right] = -\boldsymbol{\eta}\, d\, G_{q+2,\,m}\left(r_2^{\mathrm{W}^*}; \Delta\right), \tag{4.5.7}$$

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LR}^*}\right] = -\boldsymbol{\eta}\, d\, G_{q+2,\,m}\left(r_2^{\mathrm{LR}^*}; \Delta\right) \tag{4.5.8}$$

and

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LM}^*}\right] = -\boldsymbol{\eta}\, d\, G_{q+2,\,m}\left(r_2^{\mathrm{LM}^*}; \Delta\right) \tag{4.5.9}$$

where $r_2^{\mathrm{W}^*} = \frac{\chi^2_\alpha}{(q+2)}$, $r_2^{\mathrm{LR}^*} = \frac{m}{(q+2)}\left(e^{\chi^2_\alpha/(m+\frac{q}{2}-1)} - 1\right)$, $r_2^{\mathrm{LM}^*} = \frac{m\chi^2_\alpha}{(q+2)\left(m+q-\chi^2_\alpha\right)}$; *and* $G_{a,b}(r;\Delta)$ *is the cumulative distribution function of the non-central F distribution $(a,\, b)$ d.f., with non-centrality parameter Δ and evaluated at r.*

Theorem 4.17. *The quadratic bias functions of the SPTEs of $\boldsymbol{\beta}$ under the modified W, LR and LM tests are*

$$\mathrm{QB}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{W}^*}\right] = -\boldsymbol{\eta}'\boldsymbol{\eta}\,\bar{d}^2\,G^2_{q+2,\,m}\left(r_2^{\mathrm{W}^*};\Delta\right), \tag{4.5.10}$$

$$\mathrm{QB}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LR}^*}\right] = -\boldsymbol{\eta}'\boldsymbol{\eta}\,\bar{d}^2\,G^2_{q+2,\,m}\left(r_2^{\mathrm{LR}^*};\Delta\right) \tag{4.5.11}$$

and

$$\mathrm{QB}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LM}^*}\right] = -\boldsymbol{\eta}'\boldsymbol{\eta}\,\bar{d}^2\,G^2_{q+2,\,m}\left(r_2^{\mathrm{LM}^*};\,\Delta\right) \tag{4.5.12}$$

respectively.

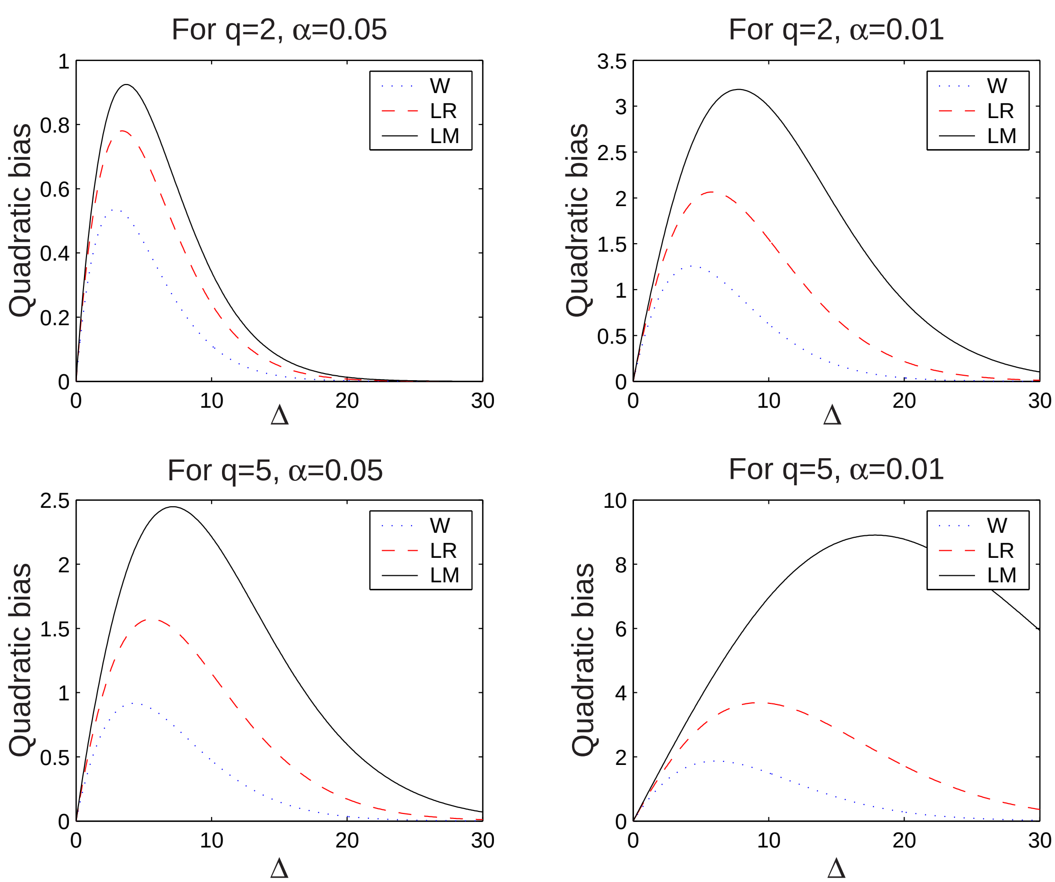

Figure 4.3: The QBs of the SPTEs under the modified W, LR and LM tests for $n = 25$, $d = 0.1$, $p = 8$ and selected values of q and α.

Theorem 4.18. *The quadratic risk functions of the SPTEs of $\boldsymbol{\beta}$ under the modified W, LR and LM tests, are*

$$\mathrm{R}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{W}^*}\Big] = p - qd^*\, G_{q+2,\,m}\Big(r_2^{\mathrm{W}^*};\Delta\Big) + 2d\Delta\, G_{q+2,\,m}\Big(r_2^{\mathrm{W}^*};\Delta\Big) - d^*\Delta\, G_{q+4,\,m}\Big(r_4^{\mathrm{W}^*};\Delta\Big), \tag{4.5.13}$$

$$\mathrm{R}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LR}^*}\Big] = p - qd^*\, G_{q+2,\,m}\Big(r_2^{\mathrm{LR}^*};\Delta\Big) + 2d\Delta\, G_{q+2,\,m}\Big(r_2^{\mathrm{LR}^*};\Delta\Big) - d^*\Delta\, G_{q+4,\,m}\Big(r_4^{\mathrm{LR}^*};\Delta\Big) \tag{4.5.14}$$

and

$$\mathrm{R}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LM}^*}\Big] = p - qd^*\, G_{q+2,\,m}\Big(r_2^{\mathrm{LM}^*};\Delta\Big) + 2d\Delta\, G_{q+2,\,m}\Big(r_2^{\mathrm{LM}^*};\Delta\Big) - d^*\Delta\, G_{q+4,\,m}\Big(r_4^{\mathrm{LM}^*};\Delta\Big) \tag{4.5.15}$$

respectively, where $r_i^{\mathrm{W}^*} = \frac{\chi^2_\alpha}{(q+i)}$, $r_i^{\mathrm{LR}^*} = \frac{m}{(q+i)}\Big(e^{\chi^2_\alpha/(m+\frac{q}{2}-1)} - 1\Big)$, $r_i^{\mathrm{LM}^*} = \frac{m\chi^2_\alpha}{(q+i)\big(m+q-\chi^2_\alpha\big)}$, $i = 2, 4$; *and* $G_{a,\,b}(r;\,\Delta)$ *is the cumulative distribution function of the non-central F distribution with $(a,\,b)$ d.f., non-centrality parameter Δ and evaluated at r.*

Figures 4.3 and 4.4 display the QBs and efficiencies of the SPTEs under the modified tests relative to the UE of $\boldsymbol{\beta}$, against Δ. It is observed that the use of the modified tests in the definition of the SPTE reduces the conflict among the relative efficiencies of the SPTEs as compared to the use of the original tests, but does not reduce the conflict among the QBs. Also, the reduced conflict among the relative efficiencies of the SPTEs under the modified test is considerable.

As the conflict among the size and power properties of the three Edgeworth size-corrected tests is almost negligible, we conjecture that the use of the size-corrected tests in the definition of the SPTE will eliminate the conflict among the statistical properties of the estimators. Therefore, we study the performance of the SPTEs under the Edgeworth size-corrected tests.

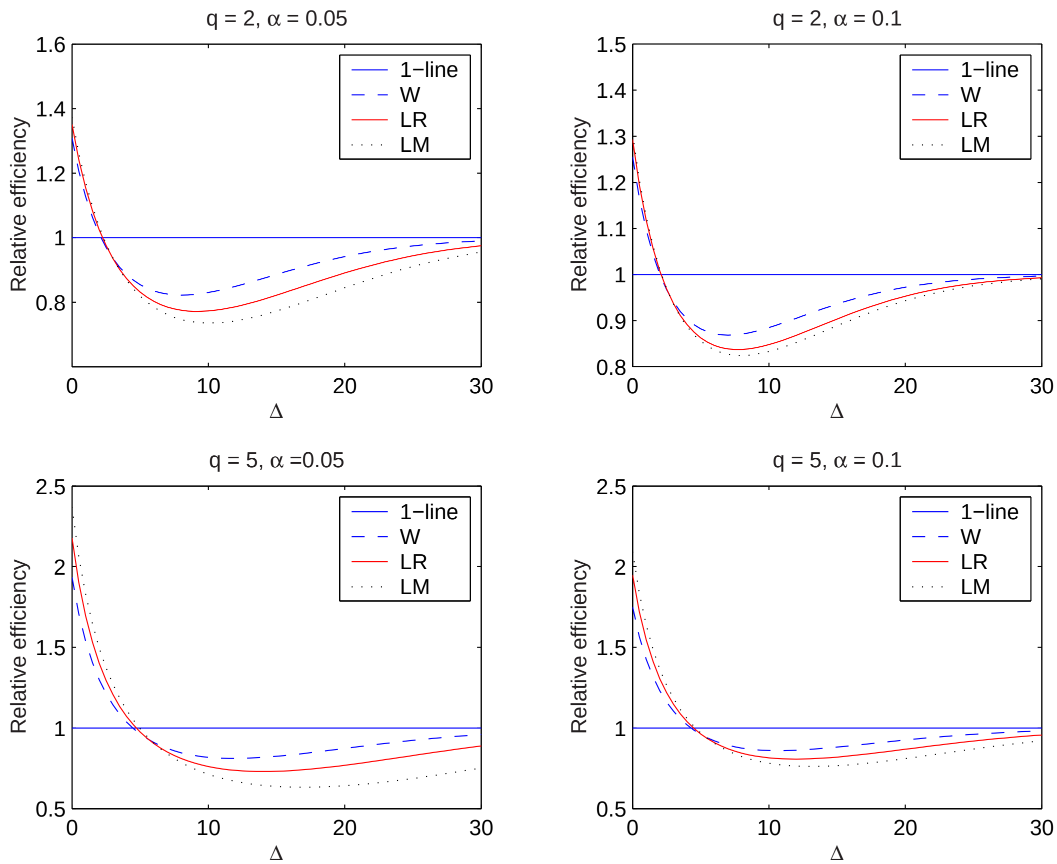

Figure 4.4: Efficiencies of the SPTEs under the modified W, LR and LM tests relative to the UE for $n = 25$, $d = 0.1$, $p = 8$ and selected values of q and α.

4.6 The Size-Corrected W, LR and LM Tests and the SPTE of $\boldsymbol{\beta}$

Following Evans and Savin (1982), we now consider the correction factors for the chi-square critical values of the W and LM test statistics. As the modified

LR test is based on the Edgeworth approximation to the exact distribution of the test statistic, it has the correct significance level. The Edgeworth size-corrected critical values (to order $1/m$) of the ξ_{W^*} and ξ_{LM^*} test statistics are

$$\zeta_{\mathrm{W}^c} = \chi^2_\alpha \left\{ 1 + \frac{\chi^2_\alpha - q + 2}{2m} \right\} \tag{4.6.1}$$

and

$$\zeta_{\mathrm{LM}^c} = \chi^2_\alpha \left\{ 1 + \frac{\chi^2_\alpha - q - 2}{2m} \right\} \tag{4.6.2}$$

respectively (cf. Evans and Savin, 1982). The tests with the adjusted critical values are known as the Edgeworth size-corrected tests.

The associated indicator functions with the SPTEs of $\boldsymbol{\beta}$ under the Edgeworth size-corrected W, LR and LM tests are given by

$$\begin{aligned} I_{\mathrm{W}^c} &= I\left(qF < \chi^2_\alpha \left\{ 1 + \frac{\chi^2_\alpha - q + 2}{2m} \right\} \right) \\ &= I\left(F < \frac{\chi^2_\alpha}{q} \left\{ 1 + \frac{\chi^2_\alpha - q + 2}{2m} \right\} \right), \end{aligned} \tag{4.6.3}$$

$$I_{\mathrm{LR}^c} = I\left(F < \left\{ \frac{m}{q} \left(e^{\chi^2_\alpha / (m + \frac{q}{2} - 1)} - 1 \right) \right\} \right) \tag{4.6.4}$$

and

$$\begin{aligned} I_{\mathrm{LM}^c} &= I\left(\frac{(m+q)qF}{m+qF} < \chi^2_\alpha \left\{ 1 - \frac{\chi^2_\alpha - q - 2}{2m} \right\} \right) \\ &= I\left(F < \frac{m\chi^2_\alpha (2m - \chi^2_\alpha + q + 2)}{q\left\{ 2m^2 + 2mq - \chi^2_\alpha (2m - \chi^2_\alpha + q + 2) \right\}} \right) \end{aligned} \tag{4.6.5}$$

respectively. The bias, QB and quadratic risk functions of the SPTEs of $\boldsymbol{\beta}$ under the Edgeworth size-corrected W, LR and LM tests are stated in the following theorems.

Theorem 4.19. *The bias functions of the SPTE of $\boldsymbol{\beta}$ under the Edgeworth size-corrected W, LR and LM tests are respectively*

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{W}^c}\right] = -\boldsymbol{\eta}\, d\, G_{q+2,\,m}\left(r^{\mathrm{W}^c}_{E2};\, \Delta\right), \tag{4.6.6}$$

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LR}^c}\right] = -\boldsymbol{\eta}\, d\, G_{q+2,\,m}\left(r^{\mathrm{LR}^c}_{E2};\, \Delta\right) \tag{4.6.7}$$

and

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LM}^c}\right] = -\boldsymbol{\eta}\, d\, G_{q+2,\,m}\left(r^{\mathrm{LM}^c}_{E2};\, \Delta\right) \tag{4.6.8}$$

where $r^{\mathrm{W}^c}_{E2} = \frac{\chi^2_\alpha}{(q+2)}\left(1 + \frac{\chi^2_\alpha - q + 2}{2m}\right)$, $r^{\mathrm{LM}^c}_{E2} = \frac{m\chi^2_\alpha\left(2m-\chi^2_\alpha+q+2\right)}{(q+2)\left(2m^2+2mq-\chi^2_\alpha(2m-\chi^2_\alpha+q+2)\right)}$, $r^{\mathrm{LR}^c}_{E2}$ *is defined in Theorem 4.16.; $G_{a,\,b}(r;\,\Delta)$ is defined as before.*

Theorem 4.20. *The quadratic bias functions of the SPTE of $\boldsymbol{\beta}$ under the Edgeworth size-corrected W, LR and LM tests are*

$$\mathrm{QB}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{W}^c}\right] = -\boldsymbol{\eta}'\boldsymbol{\eta}\, \bar{d}^2\, G^2_{q+2,\,m}\left(r^{\mathrm{W}^c}_{E2};\, \Delta\right) \tag{4.6.9}$$

$$\mathrm{QB}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LR}^c}\right] = -\boldsymbol{\eta}'\boldsymbol{\eta}\, \bar{d}^2\, G^2_{q+2,\,m}\left(r^{\mathrm{LR}^c}_{E2};\, \Delta\right) \tag{4.6.10}$$

and

$$\mathrm{QB}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LM}^c}\right] = -\boldsymbol{\eta}'\boldsymbol{\eta}\, \bar{d}^2\, G^2_{q+2,\,m}\left(r^{\mathrm{LM}^c}_{E2};\, \Delta\right) \tag{4.6.11}$$

respectively.

Theorem 4.21. *The quadratic risk functions of the SPTEs of $\boldsymbol{\beta}$ under the Edgeworth size-corrected W, LR and LM tests are respectively*

$$\mathrm{R}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{W}^c}\right] = p - qd^c\, G_{q+2,\,m}\left(r^{\mathrm{W}^c}_{E2};\, \Delta\right) + 2d\Delta\, G_{q+2,\,m}\left(r^{\mathrm{W}^c}_{E2};\, \Delta\right) - d^*\Delta\, G_{q+4,\,m}\left(r^{\mathrm{W}^c}_{E4};\, \Delta\right), \tag{4.6.12}$$

$$\mathrm{R}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SPTE}}_{\mathrm{LR}^c}\right] = p - qd^*\, G_{q+2,\,m}\left(r^{\mathrm{LR}^c}_{E2};\, \Delta\right) + 2d\Delta\, G_{q+2,\,m}\left(r^{\mathrm{LR}^c}_{E2};\, \Delta\right) - d^*\Delta\, G_{q+4,\,m}\left(r^{\mathrm{LR}^c}_{E4};\, \Delta\right) \tag{4.6.13}$$

and

$$\mathrm{R}\left[\hat{\boldsymbol{\beta}}_{\mathrm{LM}^c}^{\mathrm{SPTE}}\right] = p - qd^* G_{q+2,m}\left(r_{E2}^{\mathrm{LM}^c};\, \Delta\right) + 2d\Delta\, G_{q+2,m}\left(r_{E2}^{\mathrm{LM}^c};\, \Delta\right) - d^*\Delta\, G_{q+4,m}\left(r_{E4}^{\mathrm{LM}^c}; \Delta\right) \quad (4.6.14)$$

where $r_{Ei}^{\mathrm{W}^c} = \frac{\chi_\alpha^2}{(q+i)}\left(1 + \frac{\chi_\alpha^2 - q + 2}{2m}\right)$, $r_{Ei}^{\mathrm{LM}^c} = \frac{m\chi_\alpha^2\left(2m - \chi_\alpha^2 + q + 2\right)}{(q+i)\left(2m^2 + 2mq - \chi_\alpha^2(2m - \chi_\alpha^2 + q + 2)\right)}$, $r_{Ei}^{\mathrm{LR}^c}$ *is defined in Theorem 4.16.*, $i = 2, 4$; $G_{a,b}(r;\, \Delta)$ *is defined as before.*

The proofs of the above theorems are straightforward.

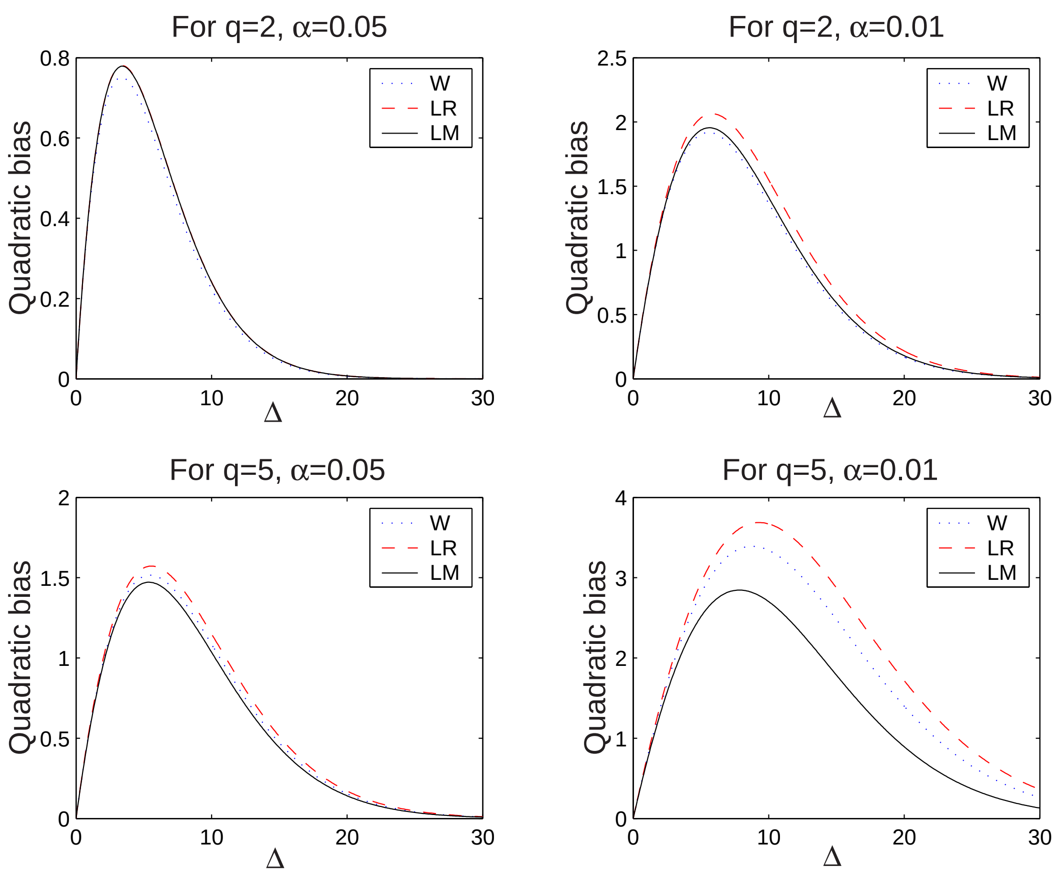

Figure 4.5: The QBs of the SPTEs under the size-corrected W, LR and LM tests for $n = 25$, $d = 0.1$, $p = 8$, and selected values of q and α.

Figures 4.5 and 4.6 display the QBs and efficiencies of the SPTEs under the size-corrected tests relative to the UE of $\boldsymbol{\beta}$, against Δ.

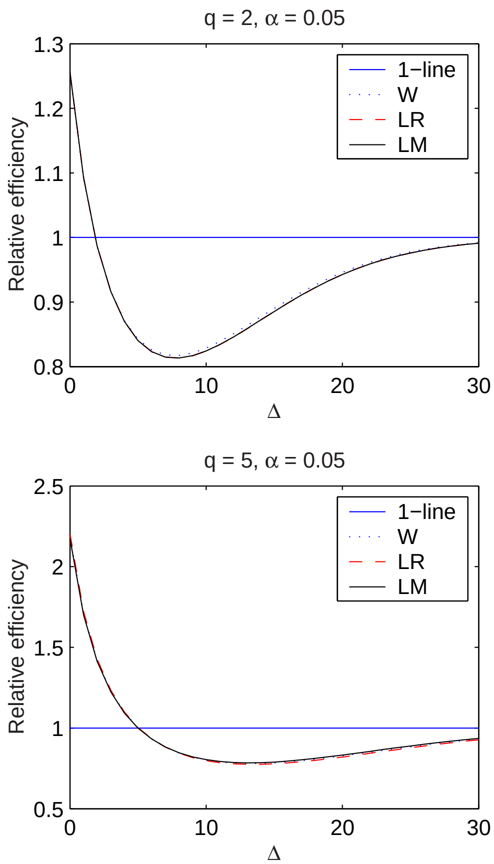

Figure 4.6: Efficiencies of the SPTEs under the size-corrected W, LR and LM tests relative to UE for $n = 25$, $d = 0.1$, $p = 8$ and selected values of q and α.

4.7 The QB and Relative Efficiency Analysis

The QBs and relative efficiencies of the SPTEs based on the original, modified as well as size-corrected W, LR and LM tests are calculated. The calculations were carried out for selected values of n, p, q, d and α. As SPTE is a biased estimator, the focus of this study is on the efficiency of the SPTEs relative to the UE with respect to the quadratic risk criterion. From Figures 4.1, 4.3 and 4.6 it is observed that there is a great deal of conflict among the quadratic biases of the SPTEs under both original and modified W, LR and LM tests. The use of the size-corrected tests almost eliminates the conflict among the QBs of the SPTEs under the three different tests.The QBs of the SPTEs under the original, modified and Edgeworth size-corrected tests are zero at $\Delta = 0$. The inequality relation

$$\mathrm{QB}\left[\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SPTE}}\right] \leq \mathrm{QB}\left[\hat{\boldsymbol{\beta}}_{\mathrm{LR}}^{\mathrm{SPTE}}\right] \leq \mathrm{QB}\left[\hat{\boldsymbol{\beta}}_{\mathrm{LM}}^{\mathrm{SPTE}}\right] \tag{4.7.1}$$

holds for the SPTEs under the original, modified as well as the size-corrected tests.

Table 4.1 and Figure 4.2 display the relative efficiencies of the estimators and the conflict among them under the three original tests. The results show that from 0 to some moderate value of Δ (say Δ_0), the SPTE based on the LM test performs the best followed by those based on the LR and W tests. For $\Delta > \Delta_0$ they perform in reverse order. The conflict among the estimators is considerably large and increases as q increases. The conflict is as large as 2.02 for $n = 25$, $q = 8$, $\Delta = 0$ and $\alpha = 0.1$. This is due to the fact that the original tests do not have the correct significance level.

Selected results for the modified tests based SPTEs are presented in Table 4.2 and Figure 4.4. It is revealed that the performance patterns of the SPTEs based on the modified tests are very similar to those based on the original tests. As compare to the original tests, the modified tests reduce the conflict among the relative efficiencies of the SPTEs. However, the conflict is still substantial. This is not unexpected, because the corrections to the original W and LM tests are based only on the bias correction in the estimates of error variance σ^2.

As stated earlier, the Edgeworth size-corrected tests have the correct significance level to order $1/m$. Table 4.3 and Figure 4.6 present the selected results for the estimators under the Edgeworth size-corrected tests. Clearly, the size-corrected tests reduce the conflict significantly. For example, when $n = 25$, $q = 8$, $\Delta = 1.5$ and $\alpha = 0.1$, the conflict is 0.038 as compared to 0.789 and 0.495 for the original and modified tests, respectively. The conflict among the SPTEs based on the size-corrected tests is negligible compared to those among the SPTEs based on the original and modified tests. For example, if $n = 25$, $q = 5$, $\Delta = 1.5$ and $\alpha = 0.1$, the conflicts are 0.541, 0.342 and 0.009 respectively for the original, modified and Edgeworth size-corrected tests. As q decreases, the conflict reduces.

If a test does not have the correct significance level, the quadratic bias and relative efficiencie of the SPTE under this test may be artificial and hence one should not rely on the performance of the estimator based on such test. The results illustrate the importance of using the tests with correct significance level in the formation of SPTE.

4.8 Concluding Remarks

The SPTE of the coefficient vector of the multiple linear regression model has been defined based on the original, modified and size-corrected W, LR and LM tests. There is a great deal of conflict among the statistical properties of the three SPTEs under the original W, LR and LM tests. The use of three modified tests in the definition of the SPTE reduces the conflict among the statistical properties of the estimators to some extent, but remains considerable. However, the use of the size-corrected tests in the definition of the SPTE almost eliminate the conflict among the statistical properties of the SPTEs. Therefore, the practitioners of the SPTE can use any of the three size-corrected W, LR and LM tests in the definition of the estimator without risking of conflicting performance of the SPTE.

4.A Appendix

4.A.1 Example

Let $y_1, y_2, \ldots, y_n$ be a random sample of size n drawn from a univariate normal distribution with unknown mean μ and variance σ^2. We want to test the null hypothesis

$$H_0 : \mu = \mu_0 \tag{4.A.1}$$

against the alternative hypothesis

$$H_1 : \mu \neq \mu_0. \tag{4.A.2}$$

The log-likelihood function is

$$l(\boldsymbol{\theta}) = \frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2 \tag{4.A.3}$$

where $\boldsymbol{\theta} = (\mu, \sigma^2)'$. From this we get

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \mu) = \frac{n(\bar{y} - \mu)}{\sigma^2} \tag{4.A.4}$$

and

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(y_i - \mu)^2. \tag{4.A.5}$$

The unrestricted and restricted estimators are

$$\tilde{\mu} = \bar{y}, \quad \tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 \tag{4.A.6}$$

and

$$\hat{\mu} = \mu_0, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu_0)^2 \tag{4.A.7}$$

respectively. The information matrix is

$$I(\boldsymbol{\theta}) = -E\left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right) = \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{n}{2\sigma^4} \end{bmatrix}. \tag{4.A.8}$$

Here, $h(\boldsymbol{\theta}) = \mu - \mu_0$, and hence $H(\boldsymbol{\theta}) = \frac{\partial h(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} = \frac{\partial}{\partial \boldsymbol{\theta}'}(\mu - \mu_0) = [1 \quad 0]$.

Therefore, from (4.3.5) the Wald statistic is obtained as

$$\begin{aligned} \xi_{\mathrm{W}} &= h(\tilde{\boldsymbol{\theta}})' \left[H(\tilde{\boldsymbol{\theta}}) I(\tilde{\boldsymbol{\theta}})^{-1} H(\tilde{\boldsymbol{\theta}})'\right]^{-1} h(\tilde{\boldsymbol{\theta}}) \\ &= (\tilde{\mu} - \mu_0)^2 \left[\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\tilde{\sigma}^2}{n} & 0 \\ 0 & \frac{2\tilde{\sigma}^4}{n} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right]^{-1} \\ &= \frac{(\bar{y} - \mu_0)^2}{\tilde{\sigma}^2/n}. \end{aligned} \tag{4.A.9}$$

The traditional t statistic for testing the null hypothesis in (4.A.1) is

$$t = \frac{(\bar{y} - \mu_0)}{\tilde{\sigma}/\sqrt{n-1}}, \tag{4.A.10}$$

and is distributed as Student's t distribution with $(n-1)$ d.f.

Using (4.A.10) in (4.A.9) we get

$$\xi_{\mathrm{W}} = \frac{n}{n-1} t^2. \tag{4.A.11}$$

Now, we derive the LR test statistic for testing the null hypothesis in (4.A.1) against the alternative hypothesis in (4.A.2).

By definition, the LR test statistic is

$$\begin{aligned}
\xi_{\mathrm{LR}} &= 2[l(\tilde{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}})] = 2\left[\frac{n}{2}\ln\left(\frac{\hat{\sigma}^2}{\tilde{\sigma}^2}\right)\right] \\
&= n\ \ln\frac{\sum_{i=1}^n (y_i-\mu_0)^2}{\sum_{i=1}^n (y_i-\bar{y})^2} \\
&= n\ \ln\left(1+\frac{(\bar{y}-\mu_0)^2}{\frac{1}{n}\sum_{i=1}^n (y_i-\bar{y})^2}\right)
\end{aligned}$$

$$\begin{aligned}
&= n\ \ln\left(1+\frac{\xi_{\mathrm{W}}}{n}\right) \\
&= n\ \ln\left(1+\frac{t^2}{n-1}\right).
\end{aligned} \tag{4.A.12}$$

Finally, the LM test statistic for testing the same hypotheses is

$$\begin{aligned}
\xi_{\mathrm{LM}} &= d(\hat{\boldsymbol{\theta}})' I(\hat{\boldsymbol{\theta}})^{-1} d(\hat{\boldsymbol{\theta}}) \\
&= \begin{bmatrix} \frac{n(\bar{y}-\mu_0)}{\hat{\sigma}^2} & -\frac{n}{2\hat{\sigma}^2}+\frac{1}{2\hat{\sigma}^4}\sum (y_i-\mu_0)^2 \end{bmatrix} \begin{bmatrix} \frac{\hat{\sigma}^2}{n} & 0 \\ 0 & \frac{2\hat{\sigma}^4}{n} \end{bmatrix} \\
&\quad \times \begin{bmatrix} \frac{n(\bar{y}-\mu_0)}{\hat{\sigma}^2} \\ -\frac{n}{2\hat{\sigma}^2}+\sum_{i=1}^n (y_i-\mu_0)^2/2\hat{\sigma}^4 \end{bmatrix} \\
&= \frac{(\bar{y}-\mu_0)^2}{\hat{\sigma}^2/n} \\
&= \frac{n(\bar{y}-\mu_0)^2}{\frac{1}{n}\sum_{i=1}^n (y_i-\mu_0)^2} \\
&= \frac{n(\bar{y}-\mu_0)^2}{\frac{1}{n}\sum_{i=1}^n (y_i-\bar{y})^2+(\bar{y}-\mu_0)^2} \\
&= \frac{n(\bar{y}-\mu_0)^2/\hat{\sigma}^2}{1+(\bar{y}-\mu_0)^2/\hat{\sigma}^2} \\
&= \frac{\xi_{\mathrm{W}}}{1+\frac{\xi_{\mathrm{W}}}{n}} \\
&= \frac{nt^2}{n-1+t^2}.
\end{aligned} \tag{4.A.13}$$

Under the null hypothesis, the distribution of t^2 is the central F with 1 and $(n-1)$ d.f., and under the alternative hypothesis, a non-central F with the same d.f. and non-centrality parameter $\Delta = \frac{(\mu-\mu_0)^2}{\sigma^2/n}$.

The three test statistics in (4.A.11) – (4.A.13) are asymptotically equivalent to χ^2_1 (see Engle, 1984). When the critical regions are calculated from the limiting distributions, the conflict among the inferences based on the tests is evident. This is illustrated by the following numerical inequality among the test statistics.

Theorem 4.22. *For any given set of data,*

$$\xi_{\text{LM}} \leq \xi_{\text{LR}} \leq \xi_{\text{W}} \tag{4.A.14}$$

where ξ_{W}, ξ_{LR} and ξ_{LM} are the respective test statistics for Wald, likelihood ratio and Lagrangian multiplier tests.

Proof. By definition,

$$\begin{aligned}
\xi_{\text{LR}} &= n \ \ln\left(1 + \frac{\xi_{\text{W}}}{n}\right) \\
1 + \frac{\xi_{\text{W}}}{n} &= e^{\xi_{\text{LR}}/n} \\
1 + \frac{\xi_{\text{W}}}{n} &= 1 + \frac{\xi_{\text{LR}}}{n} + \frac{\xi_{\text{LR}}^2}{2n^2} + \cdots \\
\xi_{\text{W}} &= \xi_{\text{LR}} + \frac{\xi_{\text{LR}}^2}{2n} + \cdots \\
\xi_{\text{LR}} &\leq \xi_{\text{W}}.
\end{aligned} \tag{4.A.15}$$

Again,

$$\begin{aligned}
\xi_{\text{LM}} &= \xi_{\text{W}} \left(1 + \frac{\xi_{\text{W}}}{n}\right)^{-1} \\
\frac{1}{\xi_{\text{LM}}} &= \frac{1}{\xi_{\text{W}}} + \frac{1}{n}
\end{aligned}$$

$$\frac{1}{\xi_{\text{LM}}} \geq \frac{1}{\xi_{\text{W}}}$$
$$\xi_{\text{LM}} \leq \xi_{\text{W}}. \tag{4.A.16}$$

Also, we know

$$\xi_{\text{LR}} = n \ \ln\left(1 + \frac{\xi_{\text{W}}}{n}\right)$$
$$\frac{\xi_{\text{LR}}}{n} = \ln\left(1 + \frac{\xi_{\text{W}}}{n}\right)$$
$$e^{\xi_{\text{LR}}/n} = 1 + \xi_{\text{W}}/n$$
$$e^{\xi_{\text{LR}}/n} \geq 1 + \frac{\xi_{\text{LM}}}{n}$$
$$1 + \frac{\xi_{\text{LR}}}{n} + \frac{\xi_{\text{LR}}^2}{2n^2} + \cdots \geq 1 + \frac{\xi_{\text{LM}}}{n}$$
$$\xi_{\text{LM}} \leq \xi_{\text{LR}}. \tag{4.A.17}$$

Combining (4.A.15) and (4.A.17), we get

$$\xi_{\text{LM}} \leq \xi_{\text{LR}} \leq \xi_{\text{W}}. \tag{4.A.18}$$

This completes the proof of the statement of the theorem.

For the general case, the above inequality was established by Breusch (1979).

4.A.2 Tables of Conflict Among the Efficiencies

Table 4.1: Conflict (Conf) among the SPTEs under the original W, LR and LM tests for $n = 25$, $p = 8$, and selected values of q and α.

q	Δ	$\alpha = 0.05$			Conf	$\alpha = 0.1$			Conf	$\alpha = 0.2$			Conf
		W	LR	LM		W	LR	LM		W	LR	LM	
	0	1.216	1.255	1.306	0.090	1.165	1.193	1.231	0.066	1.108	1.124	1.145	0.037
2	0.5	1.135	1.162	1.200	0.065	1.100	1.119	1.146	0.046	1.063	1.074	1.087	0.024
	1.0	1.073	1.091	1.118	0.045	1.052	1.063	1.080	0.028	1.030	1.036	1.044	0.014
	1.5	1.026	1.036	1.053	0.027	1.015	1.021	1.030	0.015	1.006	1.009	1.012	0.006
	2.0	0.990	0.993	1.000	0.010	0.988	0.989	0.991	0.003	0.989	0.988	0.988	0.001
	2.5	0.962	0.959	0.958	0.004	0.967	0.964	0.961	0.006	0.976	0.973	0.970	0.006
	3.0	0.940	0.933	0.925	0.015	0.952	0.945	0.937	0.015	0.967	0.963	0.957	0.010
	3.5	0.924	0.912	0.897	0.027	0.941	0.931	0.919	0.022	0.961	0.955	0.948	0.013
	4.0	0.912	0.896	0.875	0.037	0.933	0.921	0.905	0.028	0.957	0.950	0.941	0.016
	5.0	0.897	0.875	0.844	0.053	0.925	0.910	0.888	0.037	0.955	0.947	0.936	0.019
	10.0	0.913	0.880	0.825	0.088	0.948	0.930	0.901	0.047	0.975	0.968	0.958	0.017
	15.0	0.958	0.933	0.884	0.074	0.978	0.968	0.949	0.029	0.991	0.988	0.984	0.007
	30.0	0.998	0.996	0.990	0.008	0.999	0.999	0.997	0.002	0.999	0.999	0.999	0.000
	0	1.533	1.799	2.245	0.712	1.394	1.582	1.934	0.540	1.254	1.363	1.564	0.310
5	0.5	1.395	1.595	1.941	0.546	1.291	1.432	1.699	0.408	1.186	1.268	1.418	0.232
	1.0	1.291	1.442	1.713	0.422	1.212	1.318	1.522	0.310	1.134	1.195	1.308	0.174
	1.5	1.209	1.324	1.537	0.328	1.151	1.230	1.386	0.235	1.093	1.138	1.222	0.129
	2.0	1.145	1.231	1.398	0.253	1.102	1.161	1.278	0.176	1.062	1.093	1.155	0.093
	2.5	1.094	1.157	1.285	0.191	1.064	1.105	1.192	0.128	1.037	1.058	1.101	0.064
	3.0	1.053	1.096	1.192	0.139	1.033	1.060	1.122	0.089	1.017	1.030	1.058	0.041
	3.5	1.020	1.047	1.115	0.095	1.009	1.024	1.064	0.055	1.001	1.007	1.022	0.021
	4.0	0.993	1.006	1.050	0.057	0.990	0.995	1.016	0.026	0.989	0.989	0.994	0.005
	5.0	0.953	0.944	0.948	0.009	0.962	0.951	0.942	0.020	0.973	0.964	0.952	0.021
	10.0	0.903	0.837	0.720	0.183	0.935	0.891	0.803	0.132	0.964	0.942	0.895	0.069
	15.0	0.932	0.859	0.682	0.250	0.961	0.920	0.813	0.148	0.982	0.966	0.925	0.057
	30.0	0.994	0.976	0.845	0.149	0.997	0.992	0.956	0.041	0.999	0.998	0.992	0.007
	0	1.820	2.622	4.652	2.832	1.580	2.092	3.603	2.023	1.362	1.633	2.374	1.012
8	0.5	1.623	2.198	3.556	1.933	1.444	1.822	2.864	1.420	1.278	1.484	2.024	0.746
	1.0	1.477	1.904	2.879	1.402	1.341	1.626	2.385	1.044	1.213	1.371	1.776	0.563
	1.5	1.365	1.689	2.420	1.055	1.260	1.479	2.049	0.789	1.162	1.284	1.593	0.431
	2.0	1.277	1.526	2.089	0.812	1.197	1.365	1.802	0.605	1.121	1.215	1.453	0.332
	2.5	1.207	1.399	1.840	0.633	1.145	1.275	1.613	0.468	1.088	1.159	1.342	0.254
	3.0	1.151	1.298	1.645	0.494	1.104	1.202	1.464	0.360	1.061	1.114	1.254	0.193
	3.5	1.105	1.216	1.489	0.384	1.070	1.143	1.345	0.275	1.040	1.078	1.183	0.143
	4.0	1.067	1.149	1.362	0.295	1.043	1.095	1.247	0.204	1.022	1.048	1.124	0.102
	4.5	1.036	1.093	1.256	0.220	1.020	1.055	1.166	0.146	1.008	1.024	1.075	0.067
	5.0	1.010	1.046	1.168	0.158	1.002	1.021	1.099	0.097	0.997	1.004	1.034	0.037
	5.5	0.989	1.007	1.092	0.103	0.987	0.993	1.041	0.054	0.988	0.987	1.000	0.013
	6.0	0.971	0.974	1.027	0.056	0.975	0.970	0.992	0.022	0.981	0.974	0.971	0.010
	7.0	0.945	0.922	0.921	0.024	0.957	0.935	0.913	0.044	0.972	0.954	0.927	0.045
	10.0	0.912	0.839	0.725	0.187	0.939	0.885	0.776	0.163	0.965	0.933	0.859	0.106
	15.0	0.924	0.821	0.586	0.338	0.954	0.889	0.704	0.250	0.977	0.948	0.852	0.125
	30.0	0.989	0.945	0.587	0.402	0.995	0.979	0.825	0.170	0.998	0.994	0.963	0.035

Table 4.2: Conflict (Conf) among the SPTEs under the modified W, LR and LM tests for $n = 25$, $p = 8$, and selected values of q and α.

q	Δ	$\alpha = 0.05$			Conf	$\alpha = 0.1$			Conf	$\alpha = 0.2$			Conf
		W	LR	LM		W	LR	LM		W	LR	LM	
	0	1.308	1.350	1.373	0.065	1.255	1.292	1.306	0.051	1.184	1.209	1.208	0.025
2	0.5	1.207	1.242	1.262	0.055	1.166	1.194	1.205	0.039	1.115	1.132	1.131	0.017
	1.0	1.127	1.155	1.171	0.044	1.097	1.117	1.125	0.028	1.062	1.074	1.073	0.012
	1.5	1.063	1.083	1.096	0.033	1.043	1.056	1.062	0.019	1.023	1.029	1.029	0.006
	2.0	1.011	1.024	1.033	0.022	1.000	1.007	1.010	0.010	0.992	0.994	0.994	0.002
	2.5	0.969	0.975	0.981	0.012	0.967	0.968	0.969	0.002	0.969	0.968	0.968	0.001
	3.0	0.935	0.935	0.937	0.002	0.940	0.936	0.935	0.005	0.952	0.947	0.947	0.005
	3.5	0.908	0.901	0.899	0.009	0.919	0.911	0.908	0.011	0.939	0.932	0.932	0.007
	4.0	0.886	0.873	0.868	0.018	0.903	0.891	0.886	0.017	0.929	0.920	0.920	0.009
	5.0	0.854	0.831	0.819	0.035	0.882	0.862	0.855	0.027	0.919	0.906	0.907	0.013
	10.0	0.830	0.772	0.735	0.095	0.885	0.848	0.832	0.053	0.938	0.922	0.922	0.016
	15.0	0.886	0.821	0.772	0.114	0.935	0.903	0.889	0.046	0.973	0.962	0.963	0.011
	30.0	0.990	0.974	0.955	0.035	0.996	0.993	0.990	0.006	0.999	0.998	0.998	0.001
	0	1.933	2.177	2.370	0.437	1.751	1.948	2.093	0.342	1.533	1.651	1.693	0.160
5	0.5	1.708	1.903	2.064	0.356	1.566	1.720	1.835	0.269	1.399	1.489	1.521	0.122
	1.0	1.537	1.695	1.830	0.293	1.426	1.547	1.639	0.213	1.297	1.366	1.391	0.094
	1.5	1.404	1.531	1.645	0.241	1.317	1.412	1.486	0.169	1.217	1.270	1.290	0.073
	2.0	1.299	1.401	1.496	0.197	1.230	1.304	1.364	0.134	1.154	1.194	1.209	0.055
	2.5	1.213	1.294	1.373	0.160	1.160	1.218	1.264	0.104	1.103	1.133	1.144	0.041
	3.0	1.143	1.206	1.271	0.128	1.103	1.147	1.183	0.080	1.062	1.084	1.092	0.030
	3.5	1.085	1.133	1.185	0.100	1.057	1.088	1.115	0.058	1.029	1.043	1.049	0.020
	4.0	1.037	1.071	1.112	0.075	1.018	1.039	1.058	0.040	1.002	1.010	1.013	0.011
	5.0	0.962	0.974	0.995	0.033	0.960	0.963	0.969	0.009	0.962	0.960	0.960	0.002
	10.0	0.818	0.761	0.712	0.106	0.860	0.814	0.781	0.079	0.909	0.883	0.873	0.036
	15.0	0.824	0.731	0.637	0.187	0.881	0.819	0.766	0.115	0.936	0.908	0.897	0.039
	30.0	0.959	0.888	0.751	0.208	0.981	0.956	0.920	0.061	0.994	0.989	0.986	0.008
	0	2.809	3.673	4.652	1.843	2.362	2.945	3.603	1.241	1.907	2.190	2.374	0.467
8	0.5	2.349	2.951	3.624	1.275	2.028	2.444	2.902	0.874	1.694	1.903	2.037	0.343
	1.0	2.028	2.472	2.968	0.940	1.789	2.099	2.436	0.647	1.536	1.694	1.796	0.260
	1.5	1.793	2.132	2.514	0.721	1.610	1.847	2.105	0.495	1.414	1.537	1.615	0.201
	2.0	1.615	1.879	2.182	0.567	1.472	1.657	1.858	0.386	1.319	1.415	1.476	0.157
	2.5	1.475	1.684	1.929	0.454	1.363	1.508	1.667	0.304	1.243	1.318	1.366	0.123
	3.0	1.364	1.530	1.730	0.366	1.275	1.390	1.517	0.242	1.181	1.240	1.277	0.096
	3.5	1.273	1.406	1.569	0.296	1.203	1.294	1.395	0.192	1.131	1.176	1.205	0.074
	4.0	1.198	1.304	1.438	0.240	1.144	1.215	1.295	0.151	1.089	1.123	1.146	0.057
	4.5	1.136	1.219	1.328	0.192	1.095	1.149	1.212	0.117	1.055	1.080	1.096	0.041
	5.0	1.084	1.147	1.235	0.151	1.054	1.093	1.142	0.088	1.026	1.043	1.055	0.029
	5.5	1.040	1.086	1.156	0.116	1.019	1.046	1.082	0.063	1.002	1.012	1.020	0.018
	6.0	1.002	1.034	1.087	0.085	0.990	1.007	1.031	0.041	0.982	0.987	0.990	0.008
	7.0	0.943	0.950	0.976	0.033	0.945	0.943	0.949	0.006	0.952	0.947	0.944	0.008
	10.0	0.845	0.802	0.768	0.077	0.875	0.837	0.805	0.070	0.912	0.888	0.874	0.038
	15.0	0.812	0.719	0.618	0.194	0.864	0.796	0.726	0.138	0.919	0.885	0.862	0.057
	30.0	0.932	0.824	0.608	0.324	0.966	0.918	0.835	0.131	0.987	0.975	0.965	0.022

Table 4.3: Conflict (Conf) among the SPTEs under the size-corrected W, LR and LM tests for $n = 25$, $p = 8$, and selected values of q and α.

q	Δ	$\alpha = 0.05$			Conf	$\alpha = 0.1$			Conf	$\alpha = 0.2$			Conf
		W	LR	LM		W	LR	LM		W	LR	LM	
	0	1.347	1.350	1.349	0.003	1.290	1.292	1.293	0.003	1.208	1.209	1.211	0.003
2	0.5	1.239	1.242	1.241	0.003	1.192	1.194	1.195	0.003	1.131	1.132	1.134	0.003
	1.0	1.152	1.155	1.154	0.003	1.116	1.117	1.118	0.002	1.073	1.074	1.075	0.002
	1.5	1.081	1.083	1.082	0.002	1.055	1.056	1.057	0.002	1.029	1.029	1.029	0.000
	2.0	1.023	1.024	1.024	0.001	1.007	1.007	1.008	0.001	0.994	0.994	0.995	0.001
	2.5	0.975	0.975	0.975	0.000	0.968	0.968	0.968	0.000	0.968	0.968	0.968	0.000
	3.0	0.935	0.935	0.935	0.000	0.937	0.936	0.936	0.001	0.947	0.947	0.947	0.000
	3.5	0.902	0.901	0.901	0.001	0.912	0.911	0.911	0.001	0.932	0.932	0.931	0.001
	4.0	0.874	0.873	0.874	0.001	0.892	0.891	0.890	0.002	0.920	0.920	0.919	0.001
	5.0	0.833	0.831	0.832	0.002	0.864	0.862	0.862	0.002	0.907	0.906	0.905	0.002
	10.0	0.778	0.772	0.775	0.006	0.850	0.848	0.847	0.003	0.922	0.922	0.920	0.002
	15.0	0.828	0.821	0.824	0.007	0.906	0.903	0.902	0.004	0.963	0.962	0.961	0.002
	30.0	0.977	0.974	0.975	0.003	0.993	0.993	0.992	0.001	0.998	0.998	0.998	0.000
	0	2.165	2.177	2.136	0.041	1.945	1.948	1.939	0.009	1.654	1.651	1.666	0.015
5	0.5	1.893	1.903	1.869	0.034	1.717	1.720	1.713	0.007	1.492	1.489	1.501	0.012
	1.0	1.686	1.695	1.667	0.028	1.545	1.547	1.541	0.006	1.368	1.366	1.375	0.009
	1.5	1.525	1.531	1.508	0.023	1.410	1.412	1.407	0.005	1.272	1.270	1.277	0.007
	2.0	1.395	1.401	1.382	0.019	1.303	1.304	1.301	0.003	1.196	1.194	1.200	0.006
	2.5	1.290	1.294	1.279	0.015	1.217	1.218	1.215	0.003	1.134	1.133	1.137	0.004
	3.0	1.203	1.206	1.194	0.012	1.146	1.147	1.144	0.003	1.084	1.084	1.086	0.002
	3.5	1.130	1.133	1.124	0.009	1.087	1.088	1.086	0.002	1.044	1.043	1.045	0.002
	4.0	1.069	1.071	1.065	0.006	1.038	1.039	1.037	0.002	1.010	1.010	1.011	0.001
	5.0	0.974	0.974	0.972	0.002	0.963	0.963	0.963	0.000	0.960	0.960	0.960	0.000
	10.0	0.763	0.761	0.771	0.010	0.815	0.814	0.817	0.003	0.882	0.883	0.879	0.004
	15.0	0.736	0.731	0.749	0.018	0.820	0.819	0.822	0.003	0.907	0.908	0.904	0.004
	30.0	0.893	0.888	0.905	0.017	0.957	0.956	0.958	0.002	0.988	0.989	0.988	0.001
	0	3.661	3.673	3.396	0.277	2.965	2.945	2.869	0.096	2.220	2.190	2.230	0.040
8	0.5	2.942	2.951	2.759	0.192	2.458	2.444	2.391	0.067	1.925	1.903	1.932	0.029
	1.0	2.465	2.472	2.331	0.141	2.109	2.099	2.059	0.050	1.711	1.694	1.717	0.023
	1.5	2.127	2.132	2.024	0.108	1.855	1.847	1.817	0.038	1.550	1.537	1.554	0.017
	2.0	1.875	1.879	1.795	0.084	1.663	1.657	1.633	0.030	1.425	1.415	1.428	0.013
	2.5	1.681	1.684	1.618	0.066	1.513	1.508	1.490	0.023	1.326	1.318	1.328	0.010
	3.0	1.528	1.530	1.477	0.053	1.394	1.390	1.375	0.019	1.246	1.240	1.248	0.008
	3.5	1.404	1.406	1.363	0.043	1.297	1.294	1.282	0.015	1.181	1.176	1.182	0.006
	4.0	1.302	1.304	1.269	0.035	1.217	1.215	1.206	0.011	1.127	1.123	1.128	0.005
	4.5	1.218	1.219	1.192	0.027	1.151	1.149	1.142	0.009	1.082	1.080	1.083	0.003
	5.0	1.146	1.147	1.126	0.021	1.095	1.093	1.088	0.007	1.045	1.043	1.046	0.003
	5.5	1.086	1.086	1.070	0.016	1.047	1.046	1.043	0.004	1.014	1.012	1.014	0.002
	6.0	1.034	1.034	1.023	0.011	1.007	1.007	1.004	0.003	0.987	0.987	0.987	0.000
	10.0	0.802	0.802	0.814	0.012	0.836	0.837	0.841	0.005	0.885	0.888	0.885	0.003
	15.0	0.720	0.719	0.747	0.028	0.794	0.796	0.805	0.011	0.881	0.885	0.880	0.005
	30.0	0.826	0.824	0.865	0.041	0.916	0.918	0.926	0.010	0.974	0.975	0.973	0.002

4.A.3 MATLAB Codes

- The following MATLAB codes are used for producing Figure 4.6.

```
n=25; D=0:1:30; p=8; q=5; m=n-p; s=ones(1,length(D));
plot(D,s)
hold on
A=0.95;d=0.1; GW2=ncfcdf(chi2inv(A, q).*(1 + (chi2inv(A,
q)-q+2)./(2.*m))./(q+2), q+2, m, D);
GW4=ncfcdf(chi2inv(A, q).*(1
+(chi2inv(A, q)-q+2)./(2.*m))./(q+4), q+4, m, D);
REW=p./(p-(1-d.^2).*q.*GW2+2.*(1-d).*D.*GW2
-(1-d.^2).*D.*GW4);
plot(D, REW, ':')
GLR2=ncfcdf((m.*(exp(chi2inv(A,
q)./(m+q./2-1))-1)./(q+2)), q+2, m, D);
GLR4=ncfcdf((m.*(exp(chi2inv(A, q)./(m+q./2-1))-1)./(q+4)),
q+4,m, D);
RELR=p./(p-(1-d.^2).*q.*GLR2+2.*(1-d).*D.*GLR2
-(1-d.^2).*D.*GLR4);plot(D, RELR, 'r--')
GLM2=ncfcdf((m.*(chi2inv(A,q)).*(2.*m-chi2inv(A,
q)+q+2))./((q+2).*(2.*m.^2+2.*m.*q
-chi2inv(A, q).*(2.*m-chi2inv(A,
q)+q+2))),q+2,m,D);
legend('1-line', 'W', 'LR','LM', 1);
xlabel('\Delta'); ylabel('Relative efficiency');
title('Size-corrected tests with d=0.2')
GLM4=ncfcdf((m.*(chi2inv(A, q)).*(2.*m-chi2inv(A,
q)+q+2))./((q+4).*(2.*m.^2+2.*m.*q
-chi2inv(A, q).*(2.*m-chi2inv(A,q)+q+2))),q+4,m,D);
RELM=p./(p-(1-d.^2).*q.*GLM2+2.*(1-d).*D.*GLM2
-(1-d.^2).*D.*GLM4);plot(D, RELM, 'k')
legend('1-line', 'W', 'LR','LM', 1);
xlabel('\Delta'); ylabel('Relative efficiency');
title('q = 5 and d = 0.1')
```

Chapter 5

Shrinkage Estimator of Multiple Linear Regression Model Based on Three Test Statistics

5.1 Introduction

In Chapter 4 we studied the conflict among the properties of the shrinkage preliminary test estimator (SPTE) of the coefficient vector $\boldsymbol{\beta}$ of the multiple linear regression model (4.1.1). The performances of both the PTE and SPTE depend on the choice of the level of significance of the preliminary test. We now define a shrinkage estimator (SE) of $\boldsymbol{\beta}$ that involves an appropriate test statistic for testing the null hypothesis but does not depend on the level of significance of the test. A simple form of the SE of $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}}^{\text{SE}} = \tilde{\boldsymbol{\beta}} - k\,\xi^{-1}(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}) \tag{5.1.1}$$

where $\tilde{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}$ are the unrestricted and restricted estimators of $\boldsymbol{\beta}$, ξ is any appropriate test statistic to test the null hypothesis in (4.1.2), and k is a positive constant, generally known as the shrinkage constant. An optimal k, in the sense of minimizing the quadratic risk of the SE, is given by $k = \frac{(q-2)(n-p)}{q(n-p+2)}$

for $q \geq 3$.

Traditionally, the LR test statistic has been used to define the SE of parameters in linear models (see, for instance Khan and Saleh, 2001). As there are alternative tests, namely the Wald (W) and Lagrange multiplier (LM) tests, the SE can be defined based on the alternative test statistics in a similar way to the SPTE of $\boldsymbol{\beta}$. However, due to the conflicting size and power properties of the W, LR and LM tests, conflicting performances of the SEs based on these test statistics may not be avoidable. Since there is no recommendation for the choice of any particular test statistic in the definition of the SE, the practitioners may face the challenge of selecting one from the three available alternative test statistics. The main objective of this study is to investigate the performances of the SEs of $\boldsymbol{\beta}$ under the W, LR and LM test statistics with a view to suggesting an optimal test statistic among the three test statistics.

The layout of this chapter is as follows. Some preliminaries are outlined in Section 5.2. The bias functions of the SEs under the W, LR and LM test statistics are derived and analysed in Section 5.3. Section 5.4 is devoted to the derivation and analysis of the quadratic risk functions. Finally, some concluding remarks are presented in Section 5.5.

5.2 Some Preliminaries

For the multiple linear regression model in (4.1.1), the UE of $\boldsymbol{\beta}$ is

$$\tilde{\boldsymbol{\beta}} = C^{-1}X'\boldsymbol{Y} \tag{5.2.1}$$

and that of σ^2 is

$$\tilde{\sigma}^2 = \frac{1}{n}(\boldsymbol{Y} - X\tilde{\boldsymbol{\beta}})'(\boldsymbol{Y} - X\tilde{\boldsymbol{\beta}}) \tag{5.2.2}$$

where $C = X'X$. Under the null hypothesis in (4.1.2), the restricted estimators of $\boldsymbol{\beta}$ and σ^2 are

$$\hat{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}} - C^{-1}H'(HC^{-1}H')^{-1}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h}) \tag{5.2.3}$$

and

$$\hat{\sigma}^2 = \frac{1}{n}(\boldsymbol{Y} - X\hat{\boldsymbol{\beta}})'(\boldsymbol{Y} - X\hat{\boldsymbol{\beta}}) \tag{5.2.4}$$

respectively. Following the definition in (5.1.1), the shrinkage estimators of $\boldsymbol{\beta}$ under the W, LR and LM test statistics are

$$\hat{\boldsymbol{\beta}}_{\text{W}}^{\text{SE}} = \tilde{\boldsymbol{\beta}} - k\,\xi_{\text{W}}^{-1}(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}), \tag{5.2.5}$$

$$\hat{\boldsymbol{\beta}}_{\text{LR}}^{\text{SE}} = \tilde{\boldsymbol{\beta}} - k\,\xi_{\text{LR}}^{-1}(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}) \tag{5.2.6}$$

and

$$\hat{\boldsymbol{\beta}}_{\text{LM}}^{\text{SE}} = \tilde{\boldsymbol{\beta}} - k\,\xi_{\text{LM}}^{-1}(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}) \tag{5.2.7}$$

respectively, where ξ_{W}, ξ_{LR} and ξ_{LM} are the respective test statistics, given in (4.4.10) — (4.4.12). With respect to the unbiasedness and quadratic risk criteria, the performances of the above SEs of $\boldsymbol{\beta}$ are studied in the remainder of this chapter.

5.3 The Bias Functions

In this section the bias functions of the shrinkage estimators of $\boldsymbol{\beta}$ in (5.2.5) – (5.2.7) are derived and analysed. Following Judge and Bock (1978), direct computation leads to the following theorems.

Theorem 5.23. *The bias functions of the shrinkage estimators of $\boldsymbol{\beta}$ for the multiple linear regression model with iid normal error under the W, LR, and LM test statistics are given by*

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SE}};\,\boldsymbol{\beta}\right] = \frac{mk}{n}\boldsymbol{\eta}\,\mathrm{E}\left[\chi_{q+2}^{-2}(\Delta)\right], \tag{5.3.1}$$

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}_{\mathrm{LR}}^{\mathrm{SE}};\,\boldsymbol{\beta}\right] = \frac{k}{n}\boldsymbol{\eta}\,\mathrm{E}\left[\ln\left(1+\frac{q+2}{m}F_{q+2,\,m}(\Delta)\right)\right]^{-1} \tag{5.3.2}$$

and

$$\mathrm{B}\left[\hat{\boldsymbol{\beta}}_{\mathrm{LM}}^{\mathrm{SE}};\,\boldsymbol{\beta}\right] = \frac{k}{n}\boldsymbol{\eta}\,\left\{1+m\mathrm{E}\left[\chi_{q+2}^{-2}(\Delta)\right]\right\} \tag{5.3.3}$$

respectively, where $\chi_{q+2}^{2}(\Delta)$ is the noncentral chi-square random variable with $(q+2)$ d.f. and non-centrality parameter Δ given by (4.4.5), and $F_{q+2,\,m}(\Delta)$ is the noncentral F variable with $(q+2,\,m)$ d.f. and the same non-centrality parameter Δ.

Proof. By definition, the bias function of the SE of $\boldsymbol{\beta}$ under the W test statistic is

$$\begin{aligned}
\mathrm{B}\left[\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SE}};\,\boldsymbol{\beta}\right] &= \mathrm{E}\left[\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SE}} - \boldsymbol{\beta}\right] \\
&= -k\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})\,\xi_{\mathrm{W}}^{-1}\right] \\
&= -k\,C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\left[(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})\left(\frac{nq}{m}F\right)^{-1}\right]
\end{aligned}$$

$$= -\frac{mk\sigma}{nq}\, C^{-1}H'(HC^{-1}H')^{-\frac{1}{2}}\, \mathrm{E}\left[\boldsymbol{Z}\, \left(\frac{m\chi_q^2(\Delta)}{q\chi_m^2}\right)^{-1}\right]$$

$$= -\frac{km\sigma}{n}\, C^{-1}H'(HC^{-1}H')^{-\frac{1}{2}}\, \mathrm{E}\big[\boldsymbol{Z}\, (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\big]\,. \tag{5.3.4}$$

Applying Theorem 1, Appendix B2, Judge and Bock (1978), to the right hand side of (5.3.4), we get

$$\mathrm{B}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SE}};\, \boldsymbol{\beta}\Big] = -\frac{km}{n}\, C^{-1}H'(HC^{-1}H')^{-1}(H\boldsymbol{\beta} - \boldsymbol{h})\, \mathrm{E}\big[\chi_{q+2}^{-2}(\Delta)\big]$$

$$= \frac{mk}{n}\boldsymbol{\eta}\, \mathrm{E}\big[\chi_{q+2}^{-2}(\Delta)\big] \tag{5.3.5}$$

where $\mathrm{E}\big[\chi_{q+2}^{-2}(\Delta)\big]$ is the expected value of the inverted non-central chi-square random variable with $(q+2)$ d.f. and non-centrality parameter Δ.

Now, we derive the bias function of the SE of $\boldsymbol{\beta}$ under the LR test statistic. By definition, the bias function of the SE of $\boldsymbol{\beta}$ under the LR test statistic is

$$\mathrm{B}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{LR}}^{\mathrm{SE}};\, \boldsymbol{\beta}\Big] = \mathrm{E}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{LR}}^{\mathrm{SE}} - \boldsymbol{\beta}\Big]$$

$$= -k\, \mathrm{E}\left[(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\left\{n \ln\left(1 + \frac{q}{m}F\right)\right\}^{-1}\right]$$

$$= -\frac{k}{n}C^{-1}H'(HC^{-1}H')^{-1/2}\sigma\, \mathrm{E}\left[\boldsymbol{Z}\left\{\ln\left(1 + \frac{\boldsymbol{Z}'\boldsymbol{Z}}{\chi_m^2}\right)\right\}^{-1}\right]. \tag{5.3.6}$$

Applying Theorem 1, Appendix B2, Judge and Bock (1978) to the right hand side of (5.3.6), we get

$$\mathrm{B}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{LR}}^{\mathrm{SE}};\, \boldsymbol{\beta}\Big] = -\frac{k}{n}C^{-1}H'(HC^{-1}H')^{-1}(H\boldsymbol{\beta} - \boldsymbol{h})\, \mathrm{E}\left[\ln\left(1 + \frac{q+2}{m}F_{q+2,\, m}(\Delta)\right)\right]^{-1}$$

$$= \frac{k}{n}\boldsymbol{\eta}\, \mathrm{E}\left[\ln\left(1 + \frac{q+2}{m}F_{q+2,\, m}(\Delta)\right)\right]^{-1} \tag{5.3.7}$$

where $F_{q+2,\, m}(\Delta)$ is the noncentral F with $(q+2,\, m)$ d.f. and non-centrality parameter Δ.

Finally, we derive the bias function of the SE of $\boldsymbol{\beta}$ under the LM test statistic. By definition, the bias function of the SE of $\boldsymbol{\beta}$ under the LM test statistic is

$$
\begin{aligned}
\mathrm{B}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LM}};\, \boldsymbol{\beta}\Big] &= \mathrm{E}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LM}} - \boldsymbol{\beta}\Big] \\
&= -k\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\left(\frac{m+qF}{nqF}\right)\right] \\
&= -\frac{k}{n}\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\left(1+\frac{m}{qF}\right)\right] \\
&= -\frac{k}{n}\,C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\left[(H\tilde{\boldsymbol{\beta}} - \boldsymbol{h})\left(1+\frac{\chi^2_m}{\chi^2_q(\Delta)}\right)\right] \\
&= -\frac{k\sigma}{n}\,C^{-1}H'(HC^{-1}H')^{-\frac{1}{2}}\,\mathrm{E}\left[\boldsymbol{Z}\left\{1+\frac{\chi^2_m}{\chi^2_q(\Delta)}\right\}\right] \\
&= -\frac{k\sigma}{n}\,C^{-1}H'(HC^{-1}H')^{-\frac{1}{2}}\big\{\mathrm{E}[\boldsymbol{Z}] + m\,\mathrm{E}\big[\boldsymbol{Z}\,(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\big]\big\}. \qquad (5.3.8)
\end{aligned}
$$

Applying Theorem 1, Appendix B2, Judge and Bock (1978) to (5.3.8), we get

$$
\begin{aligned}
\mathrm{B}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LM}};\, \boldsymbol{\beta}\Big] &= -\frac{k}{n}\,C^{-1}H'(HC^{-1}H')^{-1}(H\boldsymbol{\beta} - \boldsymbol{h}) \\
&\quad - \frac{km}{n}\,C^{-1}H'(HC^{-1}H')^{-1}(H\boldsymbol{\beta} - \boldsymbol{h})\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] \\
&= \frac{k\boldsymbol{\eta}}{n}\big\{1 + m\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big]\big\} \qquad (5.3.9)
\end{aligned}
$$

which is the bias function of the SE of $\boldsymbol{\beta}$ under the LM test statistic. This completes the proof of the theorem.

5.3.1 Analysis of the Bias Functions

As the bias function of any estimator of a parameter vector is also a vector, the direct comparison among the biases is not meaningful. Therefore, using the definition in (4.2.9) the quadratic bias (QB) functions of the three SEs of $\boldsymbol{\beta}$ are derived and presented in the following theorem.

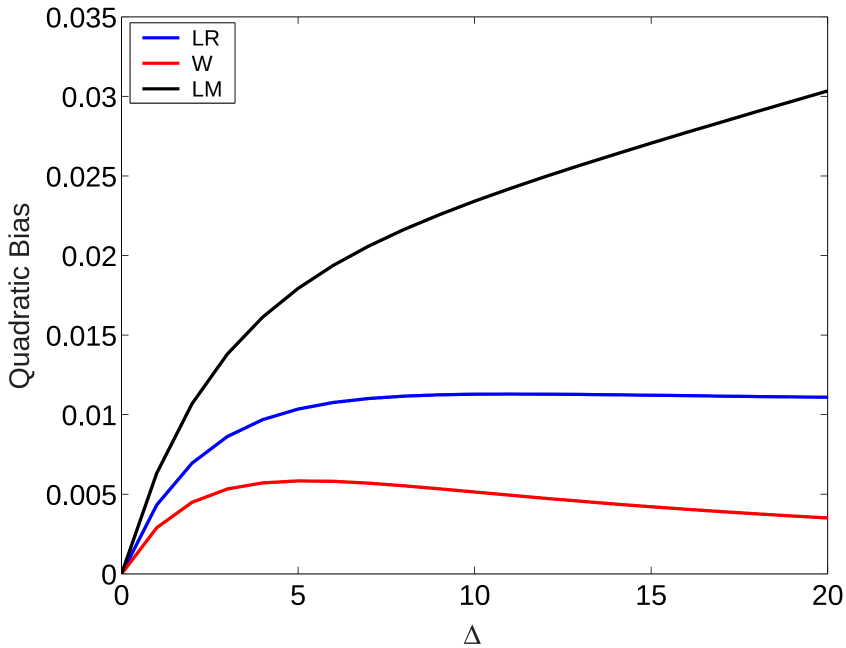

Figure 5.1: The QBs of the SEs of $\boldsymbol{\beta}$ for $n = 20$, $p = 8$ and $q = 5$

Theorem 5.24. *The quadratic bias functions of the shrinkage estimators of $\boldsymbol{\beta}$ under the W, LR, and LM test statistics are given by*

$$\mathrm{QB}\left[\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SE}};\, \boldsymbol{\beta}\right] = \frac{m^2k^2}{n^2}\,\boldsymbol{\eta}\boldsymbol{\eta}'\big(\mathrm{E}\big[\chi_{q+2}^{-2}(\Delta)\big]\big)^2, \tag{5.3.10}$$

$$\mathrm{QB}\left[\hat{\boldsymbol{\beta}}_{\mathrm{LR}}^{\mathrm{SE}};\, \boldsymbol{\beta}\right] = \frac{k^2}{n^2}\,\boldsymbol{\eta}\boldsymbol{\eta}'\left(\mathrm{E}\left[\ln\left(1 + \frac{q+2}{m}F_{q+2,\,m}(\Delta)\right)\right]^{-1}\right)^2 \tag{5.3.11}$$

and

$$\mathrm{QB}\left[\hat{\boldsymbol{\beta}}_{\mathrm{LM}}^{\mathrm{SE}};\, \boldsymbol{\beta}\right] = \frac{k^2}{n^2}\,\boldsymbol{\eta}\boldsymbol{\eta}'\big(\big\{1 + m\mathrm{E}\big[\chi_{q+2}^{-2}(\Delta)\big]\big\}\big)^2 \tag{5.3.12}$$

respectively, where $\chi_{q+2}^2(\Delta)$ is the noncentral chi-square random variable with $(q+2)$ d.f. and non-centrality parameter Δ, and $F_{q+2,\,m}(\Delta)$ is the noncentral F variable with $(q+2,\, m)$ d.f. and non-centrality parameter Δ.

Under the null hypothesis, both $\boldsymbol{\eta}\boldsymbol{\eta}'$ and Δ are zero. Therefore, the quadratic biases of the SEs under all three test statistics, are zero. As Δ deviates from zero, the QBs grow larger. However, as $\Delta \to \infty$, both $\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big]$ and $\mathrm{E}\big[\ln\big(1 + \frac{q+2}{m} F_{q+2,\,m}(\Delta)\big)\big]^{-1}$ tend to 0. Therefore, the QBs of the SEs under the W and LR test statistics approach zero, and that of the SE under the LM test statistic grows unboundedly large.

Figure 5.1 displays the behaviour of the QBs of the three SEs under the W, LR and LM test statistics, for selected values of n, p and q. The computation of the QB of the SE under the LR test statistic has been done by using recursive adaptive Simpson quadrature in MATLAB, Release 12. From the expressions of the QBs in (5.3.10) — (5.3.12) as well as Figure 5.1 the inequality relation $\mathrm{QB}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{W}};\, \boldsymbol{\beta}\Big] \leq \mathrm{QB}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LR}};\, \boldsymbol{\beta}\Big] \leq \mathrm{QB}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LM}};\, \boldsymbol{\beta}\Big]$ is observed.

5.4 The MSE Matrices and QR Functions

In this section, the mse matrices and quadratic risk functions of the SEs of $\boldsymbol{\beta}$ under the W, LR and LM test statistics are derived and analysed.

Theorem 5.25. *The mean square error matrices of the shrinkage estimators of $\boldsymbol{\beta}$ for the multiple linear regression model with iid normal error under the W, LR, and LM test statistics are given by*

$$\begin{aligned}\mathrm{M}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{W}};\, \boldsymbol{\beta}\Big] &= \sigma^2 C^{-1} + \frac{k^2(m^2+2m)}{n^2}\Big\{\Lambda\sigma^2\mathrm{E}\big[\chi^{-4}_{q+2}(\Delta)\big] + \boldsymbol{\eta}\boldsymbol{\eta}' E\left[\chi^{-4}_{q+4}(\Delta)\right]\Big\} \\ &\quad - \frac{2mk}{n}\Big\{\Lambda\sigma^2\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] - 2\boldsymbol{\eta}\boldsymbol{\eta}'\mathrm{E}\big[\chi^{-4}_{q+4}(\Delta)\big]\Big\}, \qquad (5.4.13)\end{aligned}$$

$$\mathrm{M}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LR}};\,\boldsymbol{\beta}\Big] = \sigma^2 C^{-1} + \frac{k^2}{n^2}\Bigg\{\sigma^2 \Lambda\,\mathrm{E}\Big[\ln\Big(1+\frac{q+2}{m}F_{q+2,\,m}(\Delta)\Big)\Big]^{-2}$$
$$+\,\boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\Big[\ln\Big(1+\frac{q+2}{m}F_{q+2,\,m}(\Delta)\Big)\Big]^{-2}\Bigg\}$$
$$-\,\frac{2k}{n}\Bigg\{\sigma^2\Lambda\mathrm{E}\Big[\ln\Big(1+\frac{q+2}{m}F_{q+2,\,m}(\Delta)\Big)\Big]^{-1}$$
$$+\,\boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\Big[\ln\Big(1+\frac{q+4}{m}F_{q+4,\,m}(\Delta)\Big)\Big]^{-1}$$
$$-\,\boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\Big[\ln\Big(1+\frac{q+2}{m}F_{q+2,\,m}(\Delta)\Big)\Big]^{-1}\Bigg\} \qquad (5.4.14)$$

and

$$\mathrm{M}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LM}};\,\boldsymbol{\beta}\Big] = \sigma^2 C^{-1} + \frac{k^2}{n^2}\Big[\Lambda\sigma^2 + 2m\left\{\Lambda\sigma^2\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] + \boldsymbol{\eta}\boldsymbol{\eta}'\mathrm{E}\big[\chi^{-2}_{q+4}(\Delta)\big]\right\}$$
$$+\,m(m+2)\left\{\Lambda\sigma^2\mathrm{E}\big[\chi^{-4}_{q+2}(\Delta)\big] + \boldsymbol{\eta}\boldsymbol{\eta}'\mathrm{E}\big[\chi^{-4}_{q+4}(\Delta)\big]\right\}\Big]$$
$$-\,\frac{2k}{n}\Big[\Lambda\sigma^2\left\{1+m\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big]\right\} - \boldsymbol{\eta}\boldsymbol{\eta}'\left\{1+2m\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big]\right\}\Big] \qquad (5.4.15)$$

respectively.

Proof. By definition, the mse matrix of the SE of $\boldsymbol{\beta}$ under the W test statistic is

$$\mathrm{M}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{W}};\,\boldsymbol{\beta}\Big] = \mathrm{E}\Big[(\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{W}}-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{W}}-\boldsymbol{\beta})'\Big]$$
$$= \mathrm{E}\Big[\big\{(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta}) - (\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})k\,\xi_{\mathrm{W}}^{-1}\big\}\big\{(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta}) - (\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})k\,\xi_{\mathrm{W}}^{-1}\big\}'\Big]$$
$$= \sigma^2 C^{-1} + k^2\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\xi_{\mathrm{W}}^{-2}\Big] - 2k\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\xi_{\mathrm{W}}^{-1}\Big]. \qquad (5.4.16)$$

The second term of (5.4.16) is

$$k^2\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\xi_{\mathrm{W}}^{-2}\Big]$$
$$= k^2\,C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'\left\{\frac{nq}{m}F\right\}^{-2}\Big]\,(HC^{-1}H')^{-1}HC^{-1}$$

$$= \left(\frac{km}{nq}\right)^2 C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'F^{-2}\Big]\,(HC^{-1}H')^{-1}HC^{-1}$$

$$= \left(\frac{km\sigma}{nq}\right)^2 C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\left[\boldsymbol{Z}\boldsymbol{Z}'\left(\frac{m\boldsymbol{Z}'\boldsymbol{Z}}{q\chi^2_m}\right)^{-2}\right] HC^{-1}. \qquad (5.4.17)$$

Applying Theorem 3, Appendix B2, Judge and Bock (1978) to (5.4.17), we get

$$k^2\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\xi_{\mathrm{w}}^{-2}\Big]$$

$$= \left(\frac{km\sigma}{n}\right)^2 \frac{(m^2+2m)}{m^2} C^{-1}H'(HC^{-1}H')^{-1}\Big\{\mathrm{E}\big[\chi^{-4}_{q+2}(\Delta)\big] + \sigma^{-2}(HC^{-1}H')^{-\frac{1}{2}}$$
$$(H\boldsymbol{\beta}-\boldsymbol{h})(H\boldsymbol{\beta}-\boldsymbol{h})'(HC^{-1}H')^{-\frac{1}{2}}\mathrm{E}\big[\chi^{-4}_{q+4}(\Delta)\big]\Big\}\,HC^{-1}$$

$$= \frac{k^2(m^2+2m)}{n^2}\Big\{\sigma^2 C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\mathrm{E}\big[\chi^{-4}_{q+2}(\Delta)\big] + C^{-1}H'(HC^{-1}H')^{-1}$$
$$(H\boldsymbol{\beta}-\boldsymbol{h})(H\boldsymbol{\beta}-\boldsymbol{h})'(HC^{-1}H')^{-1}H'C^{-1}\mathrm{E}\big[\chi^{-4}_{q+4}(\Delta)\big]\Big\}$$

$$= \frac{k^2(m^2+2m)}{n^2}\Big\{\sigma^2\Lambda\,\mathrm{E}\big[\chi^{-4}_{q+2}(\Delta)\big] + \boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\big[\chi^{-4}_{q+4}(\Delta)\big]\Big\}. \qquad (5.4.18)$$

The third term of (5.4.16) is

$$-2k\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\xi_{\mathrm{w}}^{-1}\Big]$$

$$= -\frac{2km}{nq}\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'F^{-1}\Big]$$

$$= -\frac{2km}{nq}\mathrm{E}\Big[\mathrm{E}\Big\{(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})/(H\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})\Big\}\,(H\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})'F^{-1}\Big](HC^{-1}H')^{-1}HC^{-1}$$

$$= -\frac{2km}{nq}\mathrm{E}\Big[\sigma^2C^{-1}H'(\sigma^2HC^{-1}H')^{-1}\Big\{(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})-(H\boldsymbol{\beta}-\boldsymbol{h})\Big\}\,(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'F^{-1}\Big]$$
$$\times\,(HC^{-1}H')^{-1}HC^{-1}$$

$$= -\frac{2km}{nq}\Big\{C^{-1}H'(HC^{-1}H')^{-1}E\,\Big[(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'F^{-1}\Big]$$
$$-\,C^{-1}H'(HC^{-1}H')^{-1}\mathrm{E}\Big[(H\boldsymbol{\beta}-\boldsymbol{h})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'F^{-1}\Big]\Big\}(HC^{-1}H')^{-1}HC^{-1}$$

$$= -\frac{2km}{nq}\Big\{C^{-1}H'(HC^{-1}H')^{-1}\sigma^2(HC^{-1}H')^{\frac{1}{2}}\mathrm{E}\left[\boldsymbol{Z}\boldsymbol{Z}'\left(\frac{m\boldsymbol{Z}'\boldsymbol{Z}}{q\chi^2_m}\right)^{-1}\right](HC^{-1}H')^{\frac{1}{2}}$$

$$- C^{-1}H'(HC^{-1}H')^{-1}(H\boldsymbol{\beta}-\boldsymbol{h})\sigma\,\mathrm{E}\left[\boldsymbol{Z}'\left(\frac{m\boldsymbol{Z}'\boldsymbol{Z}}{q\chi^2_m}\right)^{-1}\right](HC^{-1}H')^{\frac{1}{2}}\Big\}$$

$$(HC^{-1}H')^{-1}HC^{-1}. \tag{5.4.19}$$

Applying Theorems 1 and 3, Appendix B2, Judge and Bock (1978), we get

$$- 2k\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\xi_{\mathrm{W}}^{-1}\Big]$$

$$= -\frac{2km}{nq}\Big\{C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}\,q\sigma^2\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] + C^{-1}H'(HC^{-1}H')^{-1}$$

$$\times (H\boldsymbol{\beta}-\boldsymbol{h})(H\boldsymbol{\beta}-\boldsymbol{h})'(HC^{-1}H')^{-1}C^{-1}H'q\,\mathrm{E}\big[\chi^{-2}_{q+4}(\Delta)\big] - C^{-1}H'(HC^{-1}H')^{-1}$$

$$\times (H\boldsymbol{\beta}-\boldsymbol{h})(H\boldsymbol{\beta}-\boldsymbol{h})'(HC^{-1}H')^{-\frac{1}{2}}q\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big](HC^{-1}H')^{-\frac{1}{2}}HC^{-1}\Big\}$$

$$= -\frac{2km}{n}\Big\{\Lambda\sigma^2\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] + \boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\big[\chi^{-2}_{q+4}(\Delta)\big] - \boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big]\Big\}$$

$$= -\frac{2km}{n}\Big\{\Lambda\sigma^2\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] - 2\boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\big[\chi^{-4}_{q+4}(\Delta)\big]\Big\}, \tag{5.4.20}$$

as $\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] - \mathrm{E}\big[\chi^{-2}_{q+4}(\Delta)\big] = 2\,\mathrm{E}\big[\chi^{-4}_{q+2}(\Delta)\big]$.

Collecting the expressions of the second and third terms of (5.4.16) from (5.4.18) and (5.4.20) respectively, the mean square error matrix of the SE of $\boldsymbol{\beta}$ under the W test statistic is obtained as

$$\mathrm{M}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{W}};\,\boldsymbol{\beta}\Big] = \sigma^2C^{-1} + \frac{k^2(m^2+2m)}{n^2}\Big\{\Lambda\sigma^2\mathrm{E}\big[\chi^{-4}_{q+2}(\Delta)\big] + \boldsymbol{\eta}\boldsymbol{\eta}'E\left[\chi^{-4}_{q+4}(\Delta)\right]\Big\}$$
$$- \frac{2mk}{n}\Big\{\Lambda\sigma^2\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] - 2\boldsymbol{\eta}\boldsymbol{\eta}'\mathrm{E}\big[\chi^{-4}_{q+4}(\Delta)\big]\Big\}. \tag{5.4.21}$$

Now we derive the mse matrix of the SE of $\boldsymbol{\beta}$ under the LR test statistic.

By definition, the mse matrix of the SE of $\boldsymbol{\beta}$ under the LR test statistic is

$$\begin{aligned}
\mathrm{M}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LR}};\boldsymbol{\beta}\Big] &= \mathrm{E}\Big[(\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LR}}-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LR}}-\boldsymbol{\beta})'\Big] \\
&= \mathrm{E}\Big[\big\{(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})-(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})k\,\xi^{-1}_{\mathrm{LR}}\big\}\big\{(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})-(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})k\,\xi^{-1}_{\mathrm{LR}}\big\}'\Big] \\
&= \mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})'\Big]+k^2\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\xi^{-2}_{\mathrm{LR}}\Big] \\
&\quad -2k\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\xi^{-1}_{\mathrm{LR}}\Big] \\
&= \sigma^2C^{-1}+k^2\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\left\{n\ln\left(1+\frac{q}{m}F\right)\right\}^{-2}\right] \\
&\quad -2k\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\left\{n\ln\left(1+\frac{q}{m}F\right)\right\}^{-1}\right].
\end{aligned} \tag{5.4.22}$$

The second term of the right hand side of (5.4.22) is

$$\begin{aligned}
&k^2\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\left\{n\ln\left(1+\frac{q}{m}F\right)\right\}^{-2}\right] \\
&= \frac{k^2}{n^2}\,C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\left[(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'\left\{\ln\left(1+\frac{q}{m}F\right)\right\}^{-2}\right] \\
&\quad\times(HC^{-1}H')^{-1}HC^{-1} \\
&= \frac{k^2}{n^2}\sigma^2C^{-1}H'(HC^{-1}H')^{-1}\,\mathrm{E}\left[\boldsymbol{Z}\boldsymbol{Z}'\left\{\ln\left(1+\frac{\boldsymbol{Z}'\boldsymbol{Z}}{\chi^2_m}\right)\right\}^{-2}\right]HC^{-1}.
\end{aligned} \tag{5.4.23}$$

Applying Theorem 3, Appendix B2, Judge and Bock (1978), to (5.4.30) we get

$$\begin{aligned}
&k^2\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\left\{n\ln\left(1+\frac{q}{m}F\right)\right\}^{-2}\right] \\
&= \frac{k^2}{n^2}C^{-1}H'(HC^{-1}H')^{-1}\left(\mathrm{E}\left[\ln\left(1+\frac{q+2}{m}F_{q+2,\,m}(\Delta)\right)\right]^{-2}+\sigma^{-2}(HC^{-1}H')^{-1/2}\right. \\
&\quad\times(H\boldsymbol{\beta}-\boldsymbol{h})(H\boldsymbol{\beta}-\boldsymbol{h})'(HC^{-1}H')^{-1/2}\mathrm{E}\left[\ln\left(1+\frac{q+4}{m}F_{q+4,\,m}(\Delta)\right)\right]^{-2}HC^{-1} \\
&= \frac{k^2}{n^2}\,(\sigma^2\Lambda)\,\mathrm{E}\left[\ln\left(1+\frac{q+2}{m}F_{q+2,\,m}(\Delta)\right)\right]^{-2}+\frac{k^2}{n^2}\,\boldsymbol{\eta}\boldsymbol{\eta}' \\
&\quad\times\mathrm{E}\left[\ln\left(1+\frac{q+4}{m}F_{q+4,\,m}(\Delta)\right)\right]^{-2}.
\end{aligned} \tag{5.4.24}$$

The last term of the right hand side of (5.4.22) is

$$
\begin{aligned}
&- 2k\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})' \left\{n \ln\left(1 + \frac{q}{m}F\right)\right\}^{-1}\right] \\
&= -2k\,\mathrm{E}\Big[\mathrm{E}\left\{(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})/(H\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})\right\}(H\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})' \left\{n \ln\left(1 + \frac{q}{m}F\right)\right\}^{-1}\Big] \\
&\quad \times (HC^{-1}H')^{-1}HC^{-1}. \qquad (5.4.25)
\end{aligned}
$$

Following the computation of the third term of (5.4.16), and applying Theorems 1 and 3, Appendix B2, Judge and Bock (1978) to (5.4.25) we get

$$
\begin{aligned}
&- 2k\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})' \left\{n \ln\left(1 + \frac{q}{m}F\right)\right\}^{-1}\right] \\
&= -\frac{2k}{n}\Bigg\{\sigma^2\Lambda\,\mathrm{E}\left[\ln\left(1 + \frac{q+2}{m}F_{q+2,\,m}(\Delta)\right)\right]^{-1} + \boldsymbol{\eta}\boldsymbol{\eta}' \\
&\quad \times \mathrm{E}\left[\ln\left(1 + \frac{q+4}{m}F_{q+4,\,m}(\Delta)\right)\right]^{-1} - \boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\left[\ln\left(1 + \frac{q+2}{m}F_{q+2,\,m}(\Delta)\right)\right]^{-1}\Bigg\}. \\
&\qquad (5.4.26)
\end{aligned}
$$

Collecting the results from (5.4.24) and (5.4.26), and substituting into (5.4.22), the mse matrix of the SE of $\boldsymbol{\beta}$ under the LR test statistic is obtained as

$$
\begin{aligned}
\mathrm{M}\left[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LR}};\, \boldsymbol{\beta}\right] &= \sigma^2 C^{-1} + \frac{k^2}{n^2}\Bigg\{\sigma^2\Lambda\,\mathrm{E}\left[\ln\left(1 + \frac{q+2}{m}F_{q+2,\,m}(\Delta)\right)\right]^{-2} \\
&\quad + \boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\left[\ln\left(1 + \frac{q+2}{m}F_{q+2,\,m}(\Delta)\right)\right]^{-2}\Bigg\} - \frac{2k}{n}\Bigg\{\sigma^2\Lambda \\
&\quad \mathrm{E}\left[\ln\left(1 + \frac{q+2}{m}F_{q+2,\,m}(\Delta)\right)\right]^{-1} + \boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\left[\ln\left(1 + \frac{q+4}{m}F_{q+4,\,m}(\Delta)\right)\right]^{-1} \\
&\quad - \boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\left[\ln\left(1 + \frac{q+2}{m}F_{q+2,\,m}(\Delta)\right)\right]^{-1}\Bigg\}. \qquad (5.4.27)
\end{aligned}
$$

Finally, we derive the mse matrix of the SE of $\boldsymbol{\beta}$ under the LM test statistic.

By definition, the mse matrix of the SE of $\boldsymbol{\beta}$ under the LM test statistic is

$$\mathrm{M}\Big[\hat{\boldsymbol{\beta}}^{\text{SE}}_{\text{LM}};\,\boldsymbol{\beta}\Big] = \mathrm{E}\Big[(\hat{\boldsymbol{\beta}}^{\text{SE}}_{\text{LM}}-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}^{\text{SE}}_{\text{LM}}-\boldsymbol{\beta})'\Big]$$

$$\begin{aligned}
&= \mathrm{E}\Big[\big\{(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})-(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})k\,\xi^{-1}_{\text{LM}}\big\}\big\{(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})-(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})k\,\xi^{-1}_{\text{LM}}\big\}'\Big]\\
&= \sigma^2 C^{-1} + k^2\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\xi^{-2}_{\text{LM}}\Big] - 2k\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\xi^{-1}_{\text{LM}}\Big].
\end{aligned} \tag{5.4.28}$$

The second term of the right hand side of (5.4.28) is

$$\begin{aligned}
&k^2\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\xi^{-2}_{\text{LM}}\Big]\\
&= \frac{k^2}{n^2}\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\left(1+\frac{m}{q}F^{-1}\right)^2\right]\\
&= \frac{k^2}{n^2}\,\mathrm{E}\left[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\left(1+\frac{2m}{q}F^{-1}+\frac{m^2}{q^2}F^{-2}\right)\right]\\
&= \frac{k^2}{n^2}\Big\{\sigma^2\Lambda + \frac{2m}{q}\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'F^{-1}\Big] + \frac{m^2}{q^2}\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'F^{-2}\Big]\Big\}.
\end{aligned}$$

Now,

$$\begin{aligned}
&\frac{2m}{q}\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'F^{-1}\Big]\\
&= \frac{2m}{q}\Big\{C^{-1}H'(HC^{-1}H')\,\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'F^{-1}\Big]\,(HC^{-1}H')^{-1}HC^{-1}\Big\}\\
&= \frac{2m}{q}\left\{C^{-1}H'(HC^{-1}H')^{-1}\sigma^2\,\mathrm{E}\left[\boldsymbol{Z}\boldsymbol{Z}'\left(\frac{m\boldsymbol{Z}'\boldsymbol{Z}}{q\chi^2_m}\right)^{-1}\right]HC^{-1}\right\}\\
&= \frac{2m}{q}\Big\{C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}q\sigma^2\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] + C^{-1}H'(HC^{-1}H')^{-1}\\
&\quad\times(H\boldsymbol{\beta}-\boldsymbol{h})(H\boldsymbol{\beta}-\boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\,\mathrm{E}\big[\chi^{-2}_{q+4}(\Delta)\big]\Big\}\\
&= 2m\Big\{\sigma^2\Lambda\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] + \boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\big[\chi^{-2}_{q+4}(\Delta)\big]\Big\}.
\end{aligned} \tag{5.4.29}$$

Also, using the result from (5.4.18), we get

$$\frac{m^2}{q^2}\,E\left[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'F^{-2}\right] = m(m+2)\left\{\sigma^2\Lambda\,\mathrm{E}\big[\chi^{-4}_{q+2}(\Delta)\big] + \boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\big[\chi^{-4}_{q+4}(\Delta)\big]\right\}.$$

Therefore, the second term of (5.4.28) is obtained as

$$k^2\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\xi_{\mathrm{LM}}^{-2}\Big] = \frac{k^2}{n^2}\Big[\sigma^2\Lambda + 2m\left\{\sigma^2\Lambda\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] + \boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\big[\chi^{-2}_{q+4}(\Delta)\big]\right\} + m(m+2)\left\{\sigma^2\Lambda\,\mathrm{E}\big[\chi^{-4}_{q+2}(\Delta)\big] + \boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\big[\chi^{-4}_{q+4}(\Delta)\big]\right\}\Big]. \tag{5.4.30}$$

The third term of (5.4.28) is

$$-2k\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\xi_{\mathrm{LM}}^{-1}\Big] = -2k\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\,\frac{1}{n}\Big(1+\frac{m}{q}F^{-1}\Big)\Big]$$
$$= -\frac{2k}{n}\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\Big] - \frac{2mk}{nq}\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'F^{-1}\Big]. \tag{5.4.31}$$

Now,

$$\begin{aligned}
&-\frac{2k}{n}\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\Big]\\
&= -\frac{2k}{n}\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'\Big]\,(HC^{-1}H')^{-1}HC^{-1}\\
&= -\frac{2k}{n}\mathrm{E}\Big[\mathrm{E}\Big\{(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})/(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})\Big\}\,(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\Big]\\
&= -\frac{2k}{n}\mathrm{E}\Big[\sigma^2C^{-1}H'(\sigma^2HC^{-1}H')^{-1}\Big\{(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})-(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})\Big\}\,(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'\\
&\qquad (HC^{-1}H')^{-1}HC^{-1}\Big]\\
&= -\frac{2k}{n}\Big\{C^{-1}H'(HC^{-1}H')^{-1}\mathrm{E}\Big[(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'\Big]\,(HC^{-1}H')^{-1}HC^{-1}\\
&\qquad - C^{-1}H'(HC^{-1}H')^{-1}\mathrm{E}\Big[(H\boldsymbol{\beta}-\boldsymbol{h})(H\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'\Big]\,(HC^{-1}H')^{-1}HC^{-1}\Big\}\\
&= -\frac{2k}{n}\Big\{C^{-1}H'(HC^{-1}H')^{-1}(\sigma^2HC^{-1}H')(HC^{-1}H')^{-1}HC^{-1}\\
&\qquad - C^{-1}H'(HC^{-1}H')^{-1}(H\boldsymbol{\beta}-\boldsymbol{h})(H\boldsymbol{\beta}-\boldsymbol{h})'(HC^{-1}H')^{-1}HC^{-1}\Big\}\\
&= -\frac{2k}{n}\left\{\sigma^2\Lambda - \boldsymbol{\eta}\boldsymbol{\eta}'\right\}.
\end{aligned} \tag{5.4.32}$$

Also,

$$-\frac{2mk}{nq}\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'F^{-1}\Big] = -\frac{2mk}{n}\Big\{\sigma^2\Lambda\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] + \boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\big[\chi^{-2}_{q+4}(\Delta)\big] - \boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big]\Big\}. \tag{5.4.33}$$

Therefore, the third term of (5.4.28) is

$$
\begin{aligned}
&-2k\,\mathrm{E}\Big[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\xi_{\mathrm{LM}}^{-1}\Big]\\
&=-\frac{2k}{n}\left\{\sigma^2\Lambda-\boldsymbol{\eta}\boldsymbol{\eta}'+m\sigma^2\Lambda\,\mathrm{E}\big[\chi_{q+2}^{-2}(\Delta)\big]-\boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\big[\chi_{q+4}^{-2}(\Delta)\big]+\boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\big[\chi_{q+2}^{-2}(\Delta)\big]\right\}\\
&=-\frac{2k}{n}\Big[\sigma^2\Lambda\left\{1+m\,\mathrm{E}\big[\chi_{q+2}^{-2}(\Delta)\big]\right\}-\boldsymbol{\eta}\boldsymbol{\eta}'\left\{1+m\,\mathrm{E}\big[\chi_{q+4}^{-2}(\Delta)\big]-m\,\mathrm{E}\big[\chi_{q+2}^{-2}(\Delta)\big]\right\}\Big]
\end{aligned}
$$

$$
=-\frac{2k}{n}\Big[\sigma^2\Lambda\left\{1+m\,\mathrm{E}\big[\chi_{q+2}^{-2}(\Delta)\big]\right\}-\boldsymbol{\eta}\boldsymbol{\eta}'\left\{1+2m\,\mathrm{E}\big[\chi_{q+4}^{-2}(\Delta)\big]\right\}\Big]. \qquad (5.4.34)
$$

Collecting the expressions of the second and third terms of (5.4.28) from (5.4.30) and (5.4.34), the mse matrix of the SE of $\boldsymbol{\beta}$ under the LM test statistic is obtained as

$$
\begin{aligned}
\mathrm{M}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{LM}}^{\mathrm{SE}};\boldsymbol{\beta}\Big]=\sigma^2C^{-1}&+\frac{k^2}{n^2}\Big[\sigma^2\Lambda+2m\left\{\sigma^2\Lambda\,\mathrm{E}\big[\chi_{q+2}^{-2}(\Delta)\big]+\boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\big[\chi_{q+4}^{-2}(\Delta)\big]\right\}\\
&+m(m+2)\left\{\sigma^2\Lambda\,\mathrm{E}\big[\chi_{q+2}^{-4}(\Delta)\big]+\boldsymbol{\eta}\boldsymbol{\eta}'\,\mathrm{E}\big[\chi_{q+4}^{-4}(\Delta)\big]\right\}\Big]\\
&-\frac{2k}{n}\Big[\sigma^2\Lambda\left\{1+m\,\mathrm{E}\big[\chi_{q+2}^{-2}(\Delta)\big]\right\}-\boldsymbol{\eta}\boldsymbol{\eta}'\left\{1+2m\,\mathrm{E}\big[\chi_{q+4}^{-2}(\Delta)\big]\right\}\Big].
\end{aligned}
\qquad (5.4.35)
$$

This completes the proof of the theorem.

By using the definition of the QR function in (4.2.11) and assuming $\mathcal{W}=\sigma^{-2}C$, the risk functions of the SEs are stated in the following theorem.

Theorem 5.26. *The quadratic risk functions of the SEs of $\boldsymbol{\beta}$ for the multiple linear regression model with iid normal error under the W, LR, and LM test statistics are*

$$
\begin{aligned}
\mathrm{R}\Big[\hat{\boldsymbol{\beta}}_{\mathrm{W}}^{\mathrm{SE}};\boldsymbol{\beta}\Big]=p&+\frac{k^2(m^2+2m)}{n^2}\Big\{q\,\mathrm{E}\big[\chi_{q+2}^{-4}(\Delta)\big]+\Delta\,\mathrm{E}\big[\chi_{q+4}^{-4}(\Delta)\big]\Big\}\\
&-\frac{2mk}{n}\Big\{q\,\mathrm{E}\big[\chi_{q+2}^{-2}(\Delta)\big]-2\Delta\,\mathrm{E}\big[\chi_{q+4}^{-4}(\Delta)\big]\Big\},
\end{aligned}
\qquad (5.4.36)
$$

$$\begin{aligned} \mathrm{R}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LR}};\,\boldsymbol{\beta}\Big] &= p + \frac{k^2}{n^2}\Bigg\{q\,\mathrm{E}\Bigg[\ln\Bigg(1+\frac{q+2}{m}F_{q+2,\,m}(\Delta)\Bigg)\Bigg]^{-2} \\ &\quad + \Delta\,\mathrm{E}\Bigg[\ln\Bigg(1+\frac{q+4}{m}F_{q+4,\,m}(\Delta)\Bigg)\Bigg]^{-2}\Bigg\} \\ &\quad - \frac{2k}{n}\Bigg\{(q-\Delta)\mathrm{E}\Bigg[\ln\Bigg(1+\frac{q+2}{m}F_{q+2,\,m}(\Delta)\Bigg)\Bigg]^{-1} \\ &\quad + \Delta\,\mathrm{E}\Bigg[\ln\Bigg(1+\frac{q+4}{m}F_{q+4,\,m}(\Delta)\Bigg)\Bigg]^{-1}\Bigg\}. \end{aligned} \tag{5.4.37}$$

and

$$\begin{aligned} \mathrm{R}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LM}};\,\boldsymbol{\beta}\Big] &= p + \frac{k^2}{n^2}\Big[q + 2m\Big\{q\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] + \Delta\,\mathrm{E}\big[\chi^{-2}_{q+4}(\Delta)\big]\Big\} \\ &\quad + m(m+2)\left\{q\,\mathrm{E}\big[\chi^{-4}_{q+2}(\Delta)\big] + \Delta\,\mathrm{E}\big[\chi^{-4}_{q+4}(\Delta)\big]\right\}\Big] \\ &\quad - \frac{2k}{n}\Big[q\left\{1 + m\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big]\right\} - \Delta\left\{1 + 2m\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big]\right\}\Big] \end{aligned} \tag{5.4.38}$$

respectively.

5.4.1 Analysis of the QR Functions

We express the quadratic risk functions of the SEs of $\boldsymbol{\beta}$ as the efficiency functions of the SEs relative to the UE. Therefore, the efficiency functions of the SEs of $\boldsymbol{\beta}$ under the W, LR, and LM test statistics, relative to the UE are

$$\begin{aligned} \mathrm{Eff}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{W}};\,\tilde{\boldsymbol{\beta}}\Big] &= p\Big(p + \frac{k^2(m^2+2m)}{n^2}\Big\{q\,\mathrm{E}\big[\chi^{-4}_{q+2}(\Delta)\big] + \Delta\,\mathrm{E}\big[\chi^{-4}_{q+4}(\Delta)\big]\Big\} \\ &\quad - \frac{2mk}{n}\Big\{q\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] - 2\Delta\,\mathrm{E}\big[\chi^{-4}_{q+4}(\Delta)\big]\Big\}\Big)^{-1}, \end{aligned} \tag{5.4.39}$$

$$\begin{aligned}\mathrm{Eff}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LR}};\, \tilde{\boldsymbol{\beta}}\Big] = p\Bigg(p + \frac{k^2}{n^2}\Bigg\{q\,\mathrm{E}\Bigg[\ln\Bigg(1+\frac{q+2}{m}F_{q+2,\,m}(\Delta)\Bigg)\Bigg]^{-2}\\ + \Delta\,\mathrm{E}\Bigg[\ln\Bigg(1+\frac{q+4}{m}F_{q+4,\,m}(\Delta)\Bigg)\Bigg]^{-2}\Bigg\}\\ - \frac{2k}{n}\Bigg\{(q-\Delta)\mathrm{E}\Bigg[\ln\Bigg(1+\frac{q+2}{m}F_{q+2,\,m}(\Delta)\Bigg)\Bigg]^{-1}\\ + \Delta\,\mathrm{E}\Bigg[\ln\Bigg(1+\frac{q+4}{m}F_{q+4,\,m}(\Delta)\Bigg)\Bigg]^{-1}\Bigg\}\Bigg)^{-1}\end{aligned} \tag{5.4.40}$$

and

$$\begin{aligned}\mathrm{Eff}\Big[\hat{\boldsymbol{\beta}}^{\mathrm{SE}}_{\mathrm{LM}};\, \tilde{\boldsymbol{\beta}}\Big] = p\Big(p + \frac{k^2}{n^2}\Big[q + 2m\Big\{q\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big] + \Delta\,\mathrm{E}\big[\chi^{-2}_{q+4}(\Delta)\big]\Big\}\\ + m(m+2)\left\{q\,\mathrm{E}\big[\chi^{-4}_{q+2}(\Delta)\big] + \Delta\,\mathrm{E}\big[\chi^{-4}_{q+4}(\Delta)\big]\right\}\Big]\\ - \frac{2k}{n}\Big[q\left\{1 + m\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big]\right\} - \Delta\left\{1 + 2m\,\mathrm{E}\big[\chi^{-2}_{q+2}(\Delta)\big]\right\}\Big]\Big)^{-1}\end{aligned} \tag{5.4.41}$$

respectively.

Due to very complex mathematical expressions of the relative efficiency functions, their analytical analysis appears to be difficult. However, the graphical and numerical analyses are pursued. For the computation of the relative efficiency of the LR test statistic based SE, the adaptive Simpson quadrature formula has been applied by using MATLAB, Release 12. From Figure 5.2 and Table 5.4.1 it is evident that none of the three SEs uniformly dominates the other two for all values of Δ, n, p and q.

The SEs based on all three test statistics achieve their highest relative efficiencies (> 1, the efficiency of the UE) at $\Delta = 0$. As Δ increases, the relative efficiencies decrease. For smaller n and from 0 to some small value of Δ (say Δ_1), the performance of the SE based on the LM test statistic is the best followed by that based on the LR and W test statistics, respectively. For

Table 5.1: The efficiencies of the SEs under the W, LR and LM test statistics relative to the UE for selected values of n, $p = 8$ and $q = 3$

q	Δ	$n = 10$			$n = 20$			$n = 30$		
		W	LR	LM	W	LR	LM	W	LR	LM
5	0	1.0261	1.0125	1.0517	1.0928	1.0368	1.0716	1.1139	1.0355	1.0600
	1	1.0236	1.0095	1.0369	1.0785	1.0185	1.0273	1.0948	1.0106	1.0069
	2	1.0215	1.0073	1.0237	1.0676	1.0048	0.9929	1.0804	0.9922	0.9669
	3	1.0198	1.0056	1.0118	1.0592	0.9945	0.9654	1.0694	0.9784	0.9358
	4	1.0183	1.0042	1.0008	1.0526	0.9864	0.9428	1.0608	0.9676	0.9109
	5	1.0172	1.0032	0.9906	1.0474	0.9800	0.9238	1.0541	0.9592	0.8905
	6	1.0162	1.0023	0.9811	1.0432	0.9749	0.9076	1.0487	0.9524	0.8735
	7	1.0153	1.0016	0.9720	1.0397	0.9708	0.8935	1.0443	0.9469	0.8591
	8	1.0146	1.0010	0.9634	1.0369	0.9673	0.8810	1.0407	0.9424	0.8466
	9	1.0140	1.0005	0.9552	1.0345	0.9645	0.8698	1.0377	0.9386	0.8357
	10	1.0135	1.0001	0.9472	1.0325	0.9620	0.8597	1.0352	0.9355	0.8260
	11	1.0131	0.9997	0.9395	1.0308	0.9600	0.8505	1.0331	0.9327	0.8173
	12	1.0127	0.9994	0.9321	1.0293	0.9582	0.8419	1.0312	0.9304	0.8094
	13	1.0124	0.9992	0.9249	1.0281	0.9566	0.8340	1.0296	0.9283	0.8022
	14	1.0121	0.9989	0.9178	1.0270	0.9552	0.8266	1.0283	0.9265	0.7955
	15	1.0118	0.9987	0.9110	1.0261	0.9540	0.8195	1.0271	0.9250	0.7894
	17	1.0114	0.9984	0.8977	1.0245	0.9520	0.8066	1.0251	0.9223	0.7782
	20	1.0110	0.9980	0.8788	1.0230	0.9497	0.7892	1.0229	0.9192	0.7637
3	0	1.0110	1.0081	1.0208	1.0432	1.0401	1.0515	1.0539	1.0517	1.0600
	1	1.0093	1.0057	1.0119	1.0332	1.0257	1.0217	1.0406	1.0320	1.0224
	2	1.0080	1.0040	1.0042	1.0262	1.0161	0.9993	1.0313	1.0189	0.9950
	3	1.0070	1.0029	0.9974	1.0211	1.0095	0.9819	1.0248	1.0100	0.9744
	4	1.0062	1.0021	0.9912	1.0175	1.0048	0.9680	1.0201	1.0037	0.9583
	5	1.0057	1.0015	0.9856	1.0149	1.0015	0.9565	1.0167	0.9993	0.9455
	6	1.0052	1.0011	0.9803	1.0129	0.9991	0.9469	1.0142	0.9960	0.9350
	7	1.0049	1.0008	0.9752	1.0115	0.9973	0.9386	1.0123	09936	0.9262
	8	1.0046	1.0006	0.9705	1.0104	0.9960	0.9314	1.0109	0.9918	0.9187
	9	1.0044	1.0004	0.9659	1.0095	0.9949	0.9249	1.0099	0.9903	0.9122
	10	1.0043	1.0003	0.9614	1.0089	0.9941	0.9190	1.0091	0.9892	0.9064
	11	1.0042	1.0001	0.9571	1.0084	0.9934	0.9137	1.0084	0.9883	0.9013
	12	1.0041	1.0001	0.9529	1.0081	0.9929	0.9087	1.0079	0.9875	0.8966
	13	1.0041	1.0000	0.9488	1.0078	0.9924	0.9040	1.0076	0.9869	0.8924
	14	1.0040	0.9999	0.9447	1.0077	0.9920	0.8996	1.0073	0.9863	0.8884
	15	1.0040	0.9999	0.9407	1.0076	0.9917	0.8954	1.0071	0.9859	0.8847
	17	1.0041	0.9998	0.9330	1.0075	0.9911	0.8875	1.0069	0.9851	0.8779
	20	1.0042	0.9997	0.9217	1.0080	0.9905	0.8767	1.0072	0.9843	0.8689

$\Delta > \Delta_1$, the performance of the LR test statistic based SE is the best followed by that based on the W and LM test statistics, respectively.

For larger n and q, the LR test statistic based SE uniformly dominates the W and LM test statistics based SEs (see upper right graph of Figure 5.2). From 0 to some small value of Δ (say Δ_2), the LM test statistic based SE dominates the W test statistic based SE. For $\Delta > \Delta_2$, the situation is reversed.

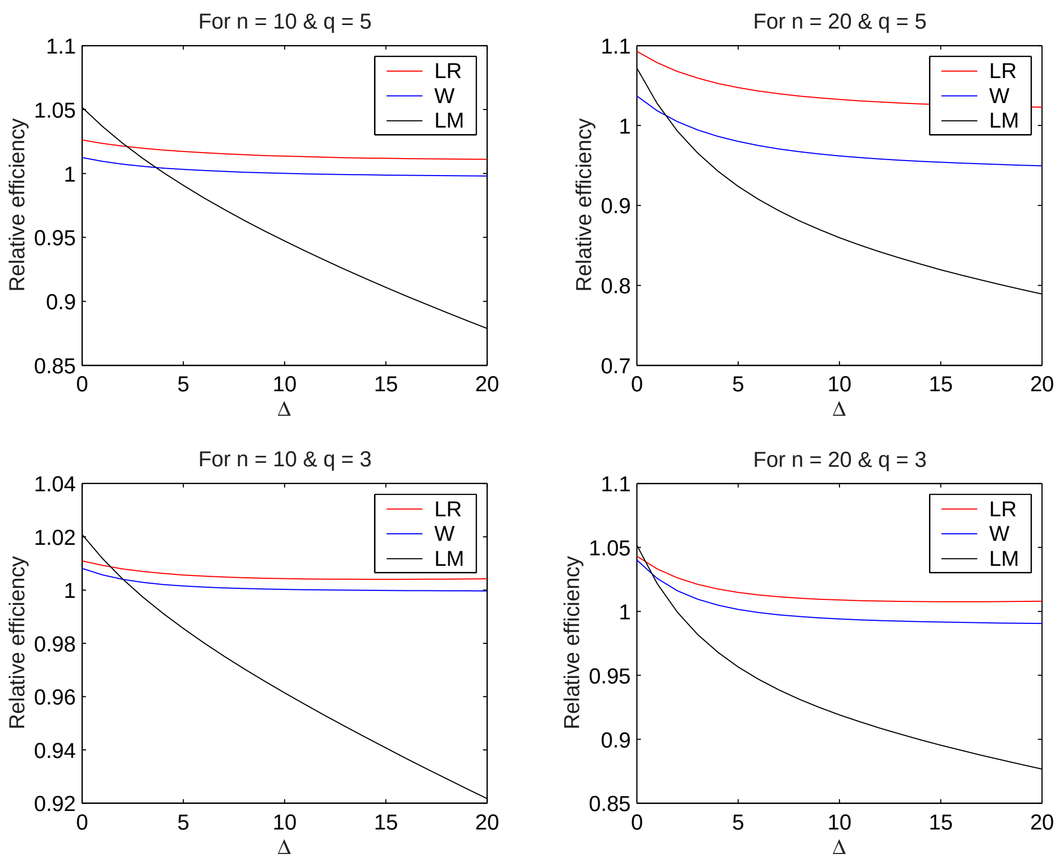

Figure 5.2: The relative efficiencies of the SEs relative to the UE for $n = 20$, $p = 8$ and $q = 5$

For larger n and smaller q, from 0 to some small value of Δ (say Δ_3), the performance of the LM test statistic based SE is the best followed by that based on the LR and W test statistics, respectively (see lower right graph of Figure 5.2). For $\Delta > \Delta_3$, the performance of the LR test statistic based SE is the best followed by those based on the W and LM test statistics.

However, as $\Delta \to \infty$, the relative efficiencies of the SEs based on the W and LR test statistics approach some constant value, say Eff_1 where $0 < \text{Eff}_1 < 1$, and that of the SE based on the LM test statistic approaches zero.

5.5 Concluding remarks

From the foregoing analyses of the quadratic bias and relative efficiency functions it is clear that with respect to QB, W test statistic based SE is the best choice followed by those based on the LR and LM test statistics. With respect to the relative efficiencies of the SEs relative to the UE, there is no uniform domination of one estimator over the others, for all values of n, p, q and Δ. As the prior information is usually obtained from some reliable sources, the value of Δ is likely to be small. Therefore, except for large sample sizes and large values of q, the LM test statistic based SE is the best choice. But this can be the worst if the value of Δ is moderate or large. For large sample sizes and large values of q, the LR test statistic based SE is the best choice. Although there is no uniform superiority of the SE based on one test statistic over those based on the other test statistics, for Δ near zero, the LM test statistic based SE is better if both n and q are not too large. For very large values of Δ, the LR test statistic based SE is the best while that based on the LM test statistic is the worst.

5.A Appendix

The following MATLAB codes are used for producing Figure 5.1.

- File 1

```
function z=f(x);
global n m p q delta
z=1./log(1+((q+2)./m).*x).*ncfpdf(x,q+2,m,delta);
```

- File 2

```
global n m p q delta n=20; p=8; q=5; m=n-p;
d=((q-2).*m)./(q.*(m+2)); Qstore=[];domain=0:20;
for delta=domain
h=0.01; x=6:h:1/h; Q=quad(@f,h,1/h); Qstore=[Qstore,Q];
end
QBLR=d.^2.*domain.*Qstore.^2./n.^2;
plot(domain, QBLR, 'b')
hold on
D=0:1:20; r=0:1:100;
for delta=0:1:20;
y1(delta+1)=sum(poisspdf(r,delta./2)./(q+2.*r))
end
QBW=D.*(m.*d.*y1./n).^2;plot(D, QBW, 'r-')
QBLM=D.*(d.*(1+m.*y1)./n).^2; plot(D, QBLM, 'k-')
legend('LR', 'W','LM', 1) xlabel('\Delta');
ylabel('Quadratic bias'); title('For q = 5')
```

Part III

Estimation Under Linex Loss Function

Chapter 6

Preliminary Test Estimator of the Slope Parameter of Simple Linear Regression Model

6.1 Introduction

In Chapters 2 and 3 we studied the performances of the SRE, SPTE and SE of the slope and intercept parameters of the simple linear regression model under a squared error loss function that is symmetric in nature. As there are many cases where underestimation of the parameter is more serious than its overestimation and vice-versa, there is growing criticism against the appropriateness of the symmetric squared error loss function (cf. Pandey and Rai, 1996) because it gives equal importance to both underestimation and overestimation. The criticism against the appropriateness of the squared error loss function has grown since the introduction of the linex loss function by Varian (1975).

The linex loss function is appropriate to assign unequal weights to underestimation and overestimation (see Zellner, 1986). Therefore, it is of interest to study the performance of the improved estimator(s) of the parameters of simple linear regression model under the linex loss function. A detailed discussion on the linex loss function is provided in Section 6.2. In this chapter we study the performance of the slope parameter of the simple linear regression model under the linex loss function.

From (2.2.4) in Chapter 2, the exclusively sample information based unrestricted estimator (UE) of the slope β_1 is given by

$$\tilde{\beta}_1 = S_{xx}^{-1} S_{xy} \tag{6.1.1}$$

where $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2$ and $S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$. The uncertain non-sample prior information based restricted estimator (RE) of β_1 is given by

$$\hat{\beta}_1^{\text{RE}} = \beta_{10}. \tag{6.1.2}$$

Following Ahsanullah and Saleh (1972), a simple form of the preliminary test estimator of β_1 is

$$\hat{\beta}_1^{\text{PTE}} = \tilde{\beta}_1 - (\tilde{\beta}_1 - \hat{\beta}_1^{\text{RE}})\, I(F_{1,\nu} < F_{1,\nu}(\alpha)) \tag{6.1.3}$$

where $I(A)$ is an indicator function of the set A and $F_{1,\nu}(\alpha)$ is the critical value chosen for the α-level test of the null hypothesis $H_0 : \beta_1 = \beta_{10}$ based on the F distribution with $(1,\, \nu)$ degrees of freedom. Under the alternative hypothesis, $H_a : \beta_1 \neq \beta_{10}$, the distribution of F is the non-central F with $(1,\, \nu)$ d.f. and non-centrality parameter Δ^2 given by (2.2.8).

Under the squared error loss function, the performance of the PTE is better than those of the UE and RE in the neighborhood of $\Delta^2 = 0$ (see Ahsanullah and Saleh, 1972). As Δ^2 deviates further from 0, the performance of PTE becomes worse than those of the UE and RE. However, as Δ^2 approaches a very large value, the performance of the PTE becomes the same as that of the UE. On the other hand, as Δ^2 increases, the performance of the RE worsen. Therefore, in the literature, with respect to the squared error loss function, the PTE is regarded as an improved estimator if the value of Δ^2 is not too far from zero, see for instance, Billah and Saleh (1998, 2002a). However, the performance of the PTE of β_1 under the linex loss function is not investigated yet.

The layout of this chapter is as follows. Section 6.2 illustrates the linex loss function and its important features. Some important lemmas are stated and proved in Section 6.3. The risk functions of different estimators of the slope parameter are derived and analysed in Section 6.4. Some concluding remarks are presented in Section 6.5. Derivation of some results, tables of efficiencies of the PTE relative to the UE and Selected MATLAB codes, used for producing graphs, are presented in Appendix 6.A.

6.2 The Linex Loss Function

The linex loss function, proposed by Varian (1975) for estimating any parameter θ by θ^*, is given by

$$\mathrm{L}(\delta) = b\left[e^{a\delta} - a\delta - 1\right], \qquad \text{for} \quad a \neq 0,\ b > 0 \tag{6.2.1}$$

where $\delta = (\theta^* - \theta)$ is the estimation error. The two parameters a and b in $\mathrm{L}(\delta)$ serve to determine the shape and scale, respectively, of $\mathrm{L}(\delta)$. A positive value of a indicates that overestimation is more serious than underestimation and a negative value of a represents the reverse situation. The magnitude of a reflects the degree of asymmetry about $\delta = 0$. This asymmetric loss function

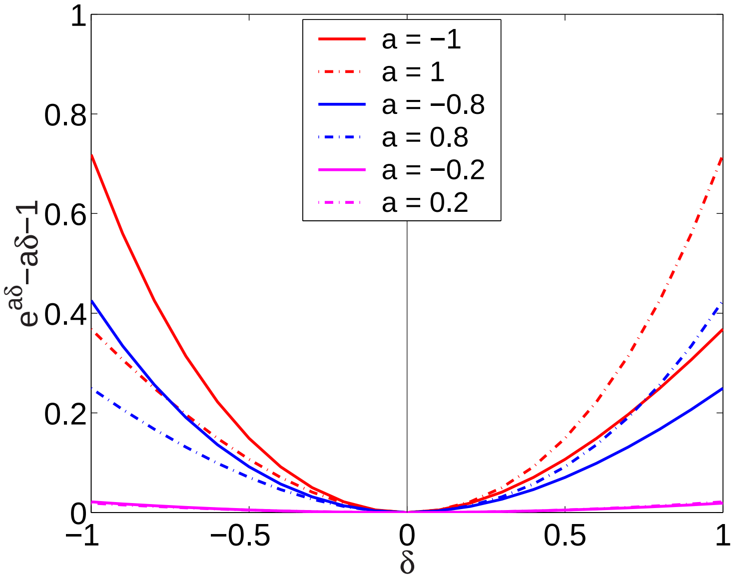

Figure 6.1: The linex loss function for selected values of a.

grows approximately linearly on one side of $\delta = 0$ and grows approximately exponentially on the other side. If $a \to 0$, then the linex loss function reduces to the squared error loss function.

Figure 6.1 displays the form of linex loss function for selected values of a against a range of values of δ. It is clear that if $a = 1$ the growth of the loss is approximately linear for negative values of δ, while for positive values of δ it is approximately exponential. For $a = -1$, the situation is reversed. As a approaches 0, the growth pattern of linex loss becomes similar for both positive and negative errors of estimation and approaches the quadratic loss. Hence, the linex loss function is more general than the quadratic loss function. More details about this loss function can be found in Zellner (1986).

6.3 Important Lemmas

In this section we derive some important results those are used in the following section to derive the risk functions of different estimators of β_1 under the linex loss function.

Lemma 6.27. *If $Z \sim N(0,1)$, and Z and S are independent then for any Borel measurable function $\phi : \Re \times (0,\infty) \to \Re$ and for any $c \in \Re$,*

$$\mathrm{E}\big[e^{cZ}\phi(Z,S)\big] = e^{c^2/2}\,\mathrm{E}\left[\phi(Z+c,\, S)\right] \tag{6.3.1}$$

provided $e^{cZ}\phi(Z,S)$ is integrable.

Proof. By definition

$$\begin{aligned}\mathrm{E}\big[e^{cZ}\phi(Z,S)\big] &= \mathrm{E}\big[\mathrm{E}\left[e^{cZ}\phi(Z,S)|S\right]\big] \\ &= \mathrm{E}\left[\frac{1}{\sqrt{2\pi}}\int_{\Re}\phi(z,S)\,e^{cz-z^2/2}dz\right] \\ &= e^{c^2/2}\,\mathrm{E}\left[\frac{1}{\sqrt{2\pi}}\int_{\Re}\phi(z,S)\,e^{-\frac{1}{2}(z-c)^2}dz\right].\end{aligned} \tag{6.3.2}$$

Consider $U = Z - c$. The Jacobian of the transformation is $|J| = 1$. Therefore,

$$\begin{aligned}\mathrm{E}\big[e^{cZ}\phi(Z,S)\big] &= e^{c^2/2}\,\mathrm{E}\left[\frac{1}{\sqrt{2\pi}}\int_{\Re}\phi(u+c,\, S)\,e^{-u^2/2}du\right] \\ &= e^{c^2/2}\,\mathrm{E}\left[\phi(Z+c,\, S)\right].\end{aligned} \tag{6.3.3}$$

This completes the proof of the lemma.

Lemma 6.28. *For any positive integers m, n*

$$\frac{\partial f_{F(m,\,n,\,D)}(x)}{\partial D} = -\frac{1}{2}f_{F(m,\,n,\,D)}(x) + \frac{m}{2(m+2)}f_{F(m+2,\,n,\,D)}\left(\frac{mx}{m+2}\right), \quad X,\, D \in \Re \tag{6.3.4}$$

where $f_{F(k,l,D)}$ denotes the density function of the non-central F with (k, l) d.f. and non-centrality parameter D.

Proof. The density function of the non-central F with (m, n) d.f. and non-centrality parameter D is given by (see Evans *et al.*, 2000, p.95)

$$f_{F(m,n,D)}(x) = \frac{e^{-D/2}m^{m/2}n^{n/2}}{\Gamma(n/2)} \frac{x^{\frac{m}{2}-1}}{(n+mx)^{\frac{m+n}{2}}} \sum_{j=0}^{\infty} \left[\frac{mxD}{2(n+mx)}\right]^j \frac{\Gamma\left(\frac{m+n}{2}+j\right)}{\Gamma\left(\frac{m}{2}+j\right) j!}. \tag{6.3.5}$$

Differentiating both sides of (6.3.5) with respect to D, we get

$$\begin{aligned}
\frac{\partial f_{F(m,n,D)}(x)}{\partial D} &= -\frac{1}{2} f_{F(m,n,D)}(x) + \frac{e^{-D/2}m^{m/2}n^{n/2}}{\Gamma(n/2)} \frac{x^{\frac{m}{2}-1}}{(n+mx)^{\frac{m+n}{2}}} \\
&\quad \times \sum_{j=1}^{\infty} \left[\frac{mx}{2(n+mx)}\right]^j \frac{D^{j-1}}{(j-1)!} \frac{\Gamma\left(\frac{m+n}{2}+j\right)}{\Gamma\left(\frac{m}{2}+j\right)} \\
&= -\frac{1}{2} f_{F(m,n,D)}(x) + \frac{e^{-D/2}m^{m/2}n^{n/2}}{\Gamma(n/2)} \frac{x^{\frac{m}{2}-1}}{(n+mx)^{\frac{m+n}{2}}} \\
&\quad \times \sum_{i=0}^{\infty} \left[\frac{mx}{2(n+mx)}\right]^{i+1} \frac{D^i}{i!} \frac{\Gamma\left(\frac{m+n+2}{2}+i\right)}{\Gamma\left(\frac{m+2}{2}+i\right)} \\
&= -\frac{1}{2} f_{F(m,n,D)}(x) + \frac{e^{-D/2}m^{m/2+1}n^{n/2}}{2\,\Gamma(n/2)} \frac{x^{\frac{m+2}{2}-1}}{(n+mx)^{\frac{m+n+2}{2}}} \\
&\quad \times \sum_{i=0}^{\infty} \left[\frac{mxD}{2(n+mx)}\right]^i \frac{\Gamma\left(\frac{m+n+2}{2}+i\right)}{i!\,\Gamma\left(\frac{m+2}{2}+i\right)} \\
&= -\frac{1}{2} f_{F(m,n,D)}(x) + \frac{e^{-D/2}m^{\frac{m+2}{2}}n^{n/2}}{2\Gamma(n/2)} \frac{\left(\frac{mx}{m+2}\right)^{\frac{m+2}{2}-1}\left(\frac{m+2}{m}\right)^{m/2}}{\left\{n+(m+2)\frac{mx}{m+2}\right\}^{\frac{m+n+2}{2}}} \\
&\quad \times \sum_{i=0}^{\infty} \left[\frac{(m+2)\frac{mxD}{m+2}}{2\left\{n+(m+2)\frac{mx}{m+2}\right\}}\right]^i \frac{\Gamma\left(\frac{m+n+2}{2}+i\right)}{i!\,\Gamma\left(\frac{m+2}{2}+i\right)} \\
&= -\frac{1}{2} f_{F(m,n,D)}(x) + \frac{me^{-D/2}(m+2)^{\frac{m+2}{2}}\,n^{n/2}\left(\frac{mx}{m+2}\right)^{\frac{m+2}{2}-1}}{2(m+2)\Gamma(n/2)\left\{n+(m+2)\frac{mx}{m+2}\right\}^{\frac{m+n+2}{2}}} \\
&\quad \times \sum_{i=0}^{\infty} \left[\frac{(m+2)\frac{mx}{m+2}D}{2\left\{n+(m+2)\frac{mx}{m+2}\right\}}\right]^i \frac{\Gamma\left(\frac{m+n+2}{2}+i\right)}{i!\,\Gamma\left(\frac{m+2}{2}+i\right)}.
\end{aligned} \tag{6.3.6}$$

Therefore,

$$\frac{\partial f_{F(m,\,n,\,D)}(x)}{\partial D} = -\frac{1}{2} f_{F(m,\,n,\,D)}(x) + \frac{m}{2(m+2)} f_{F(m+2,\,n,\,D)}\left(\frac{mx}{m+2}\right). \quad (6.3.7)$$

This completes the proof of the lemma.

6.4 Risk of the Estimators

In this section the risk function of the unrestricted and preliminary test estimators of the slope parameter β_1 are derived and presented in the following theorems.

Theorem 6.29. *The risk function of the unrestricted estimator $\tilde{\beta}_1$ of β_1 under the linex loss function is given by*

$$\mathrm{R}\Big[\tilde{\beta}_1\Big] = e^{a_1^2/2} - 1 \quad (6.4.1)$$

where $a_1 = a\sigma/\sqrt{S_{xx}}$.

Proof. By definition, the risk function of the UE of β_1 under the linex loss function in (6.2.1) is

$$\mathrm{R}\Big[\tilde{\beta}_1\Big] = \mathrm{E}\Big[e^{a(\tilde{\beta}_1 - \beta_1)}\Big] - a\,\mathrm{E}\Big[\tilde{\beta}_1 - \beta_1\Big] - 1\,. \quad (6.4.2)$$

The first term of the right hand side of (6.4.2) is

$$\begin{aligned} \mathrm{E}\Big[e^{a(\tilde{\beta}_1 - \beta_1)}\Big] &= \mathrm{E}\Bigg[e^{\frac{a(\tilde{\beta}_1 - \beta_1)\sigma/\sqrt{S_{xx}}}{\sigma/\sqrt{S_{xx}}}}\Bigg] \\ &= \mathrm{E}\Big[e^{a_1 Z}\Big] \end{aligned} \quad (6.4.3)$$

where $a_1 = a\sigma/\sqrt{S_{xx}}$ and $Z \sim N(0,\,1)$.

Applying Lemma 6.27. with ϕ as identity to (6.4.3), we get

$$\mathrm{E}\left[e^{a(\tilde{\beta}_1-\beta_1)}\right] = e^{a_1^2/2}. \tag{6.4.4}$$

As $\tilde{\beta}_1$ is an unbiased estimator of β_1, the second term of the right hand side of (6.4.2) is 0.

Collecting the results from (6.4.4) and substituting in (6.4.2), the risk function of the unrestricted estimator of β_1 is obtained as

$$\mathrm{R}\left[\tilde{\beta}_1\right] = e^{a_1^2/2} - 1. \tag{6.4.5}$$

This completes the proof of the theorem.

Theorem 6.30. *The risk function of the preliminary test estimator of the slope parameter β_1 under the linex loss function is given by*

$$\begin{aligned}\mathrm{R}\left[\hat{\beta}_1^{\mathrm{PTE}}\right] &= e^{-a_1\Delta} G_{1,\nu}\left(c;\Delta^2\right) + e^{a_1^2/2}\left[1 - G_{1,\nu}\left(c;(\Delta+a_1)^2\right)\right] \\ &\quad + a_1\Delta\, G_{3,\nu}(c/3;\,\Delta^2) - 1 \end{aligned} \tag{6.4.6}$$

where $c = F_{1,\nu}(\alpha)$ and $G_{a,b}(c;\theta)$ is the cumulative distribution function of the non-central F distribution with $(a,\, b)$ d.f., non-centrality parameter θ, and evaluated at c.

Proof. By definition, the risk function of the PTE of β_1 under the linex loss function in (6.2.1) is

$$\mathrm{R}\left[\hat{\beta}_1^{\mathrm{PTE}}\right] = \mathrm{E}\left[e^{a\Phi}\right] - a\,\mathrm{E}[\Phi] - 1 \tag{6.4.7}$$

where $\Phi = \hat{\beta}_1^{\mathrm{PTE}} - \beta_1$.

The first term of the right hand side of (6.4.7) is

$$\begin{aligned}
\mathrm{E}[e^{a\Phi}] &= \mathrm{E}\Big[e^{a(\hat{\beta}_1^{\mathrm{PTE}}-\beta_1)}\Big] \\
&= \mathrm{E}\Big[e^{a\{(\tilde{\beta}_1-\beta_1)-(\tilde{\beta}_1-\beta_{10})\,I(F_{1,\nu}<F_{1,\nu}(\alpha))\}}\Big] \\
&= \mathrm{E}\big[e^{a\{(\tilde{\beta}_1-\beta_1)-(\tilde{\beta}_1-\beta_{10})\,I(F_{1,\nu}<F_{1,\nu}(\alpha))\}}\big] \\
&\quad\times\big[I(F_{1,\nu}<F_{1,\nu}(\alpha))+I(F_{1,\nu}\geq F_{1,\nu}(\alpha))\big] \\
&= e^{a(\beta_{10}-\beta_1)}P(F_{1,\nu}<F_{1,\nu}(\alpha))+\mathrm{E}\big[e^{a(\tilde{\beta}_1-\beta_1)}\,I(F_{1,\nu}\geq F_{1,\nu}(\alpha))\big]. \quad (6.4.8)
\end{aligned}$$

The first term of the right hand side of (6.4.8) is

$$e^{a(\beta_{10}-\beta_1)}P(F_{1,\nu}<F_{1,\nu}(\alpha)) = e^{-a_1\Delta}G_{1,\nu}\left(F_{1,\nu}(\alpha);\Delta^2\right). \quad (6.4.9)$$

The second term of the right hand side of (6.4.8) is

$$\begin{aligned}
&\mathrm{E}\Big[e^{a(\tilde{\beta}_1-\beta_1)}\,I(F_{1,\nu}\geq F_{1,\nu}(\alpha))\Big] \\
&= \mathrm{E}\left[e^{a(\tilde{\beta}_1-\beta_1)}\,I\left(\frac{(\tilde{\beta}_1-\beta_{10})^2S_{xx}}{\nu S_n^2}\geq F_{1,\nu}(\alpha)\right)\right] \\
&= \mathrm{E}\left[e^{\frac{a\sigma(\tilde{\beta}_1-\beta_1)\sqrt{S_{xx}}}{\sqrt{S_{xx}}\sigma}}I\left(\frac{\left\{\frac{(\tilde{\beta}_1-\beta_1)}{\sigma/\sqrt{S_{xx}}}-\frac{(\beta_{10}-\beta_1)}{\sigma/\sqrt{S_{xx}}}\right\}^2}{\nu S_n^2/\sigma^2}\geq F_{1,\nu}(\alpha)\right)\right] \\
&= \mathrm{E}\left[e^{a_1Z}\,I\left(\frac{(Z+\Delta)^2}{\nu S_n^2/\sigma^2}\geq F_{1,\nu}(\alpha)\right)\right] \quad (6.4.10)
\end{aligned}$$

where $a_1 = a\sigma/\sqrt{S_{xx}}$, $Z = \frac{\tilde{\beta}_1-\beta_1}{\sigma/\sqrt{S_{xx}}} \sim N(0,1)$ and $(\nu S_n^2/\sigma^2) \sim \chi^2_\nu$. Z and S are independent.

Applying Lemma 6.27. with $\phi(X,Y) = I\left(\frac{(X+\Delta)^2}{Y}\geq F_{1,\nu}(\alpha)\right)$ to (6.4.10), we get

$$\begin{aligned}
\mathrm{E}\Big[e^{a(\tilde{\beta}_1-\beta_1)}\,I(F_{1,\nu}\geq F_{1,\nu}(\alpha))\Big] &= e^{a_1^2/2}\,\mathrm{E}\left[I\left(\frac{(Z-(-\Delta-a_1))^2}{\nu S_n^2/\sigma^2}\geq F_{1,\nu}(\alpha)\right)\right] \\
&= e^{a_1^2/2}\,\mathrm{E}\left[I\left(\frac{(Z-(\Delta+a_1))^2}{\nu S_n^2/\sigma^2}\geq F_{1,\nu}(\alpha)\right)\right]
\end{aligned}$$

$$
\begin{aligned}
&= e^{a_1^2/2}\,\mathrm{E}\Big[I\big(F_{1,\nu}(\Delta + a_1)^2 \geq F_{1,\nu}(\alpha)\big)\Big] \\
&= e^{a_1^2/2}\,\mathrm{E}\Big[1 - I\big(F_{1,\nu}(\Delta + a_1)^2 < F_{1,\nu}(\alpha)\big)\Big] \\
&= e^{a_1^2/2}\left[1 - G_{1,\nu}\left(F_{1,\nu}(\alpha); (\Delta + a_1)^2\right)\right]. \qquad (6.4.11)
\end{aligned}
$$

Combining (6.4.9) and (6.4.11), the first term of the right hand side of (6.4.7) yields

$$
\mathrm{E}\big[e^{a\Phi}\big] = e^{-a_1\Delta} G_{1,\nu}\left(F_{1,\nu}(\alpha); \Delta^2\right) + e^{a_1^2/2}\left[1 - G_{1,\nu}\left(F_{1,\nu}(\alpha); (\Delta + a_1)^2\right)\right]. \qquad (6.4.12)
$$

The second term of the right hand side of (6.4.7) is

$$
a\,\mathrm{E}[\Phi] = -a_1\Delta\, G_{3,\nu}\left(\frac{1}{3}F_{1,\nu}(\alpha);\ \Delta^2\right). \qquad (6.4.13)
$$

Note, the derivation of the expression in (6.4.13) along with the second order expectation is provided in the Appendix B (at the end of this chapter).

Collecting the expressions in (6.4.12) and (6.4.13) and plugging into (6.4.7), the risk function of the PTE of β_1, under the linex loss function is obtained as

$$
\begin{aligned}
\mathrm{R}\Big[\hat{\beta}_1^{\mathrm{PTE}}\Big] &= e^{-a_1\Delta} G_{1,\nu}\left(c; \Delta^2\right) + e^{a_1^2/2}\left[1 - G_{1,\nu}\left(c; (\Delta + a_1)^2\right)\right] \\
&\quad + a_1\Delta\, G_{3,\nu}(c/3;\ \Delta^2) - 1. \qquad (6.4.14)
\end{aligned}
$$

This completes the proof of the theorem.

6.4.1 Analysis of the Risks

In this subsection, we discuss some important features of the risk functions of the UE and PTE of the slope β_1 relative to the change of $\Delta = S_{xx}^{-1}(\beta_1 - \beta_{10})\sigma$ and a, the shape parameter of the linex loss function.

6.4.1.1 The Risk of the UE

Clearly, the risk function of the PTE of β_1 is independent of δ, and hence of Δ. However, it depends on the magnitude of a, but not on its sign. From the functional form of the risk of the UE it is evident that as $|a|$ grows larger, the risk of the UE also grows larger.

6.4.1.2 The Risk of the PTE

From Theorems 6.29. and 6.30., for any non-zero value of Δ, the risk function of the PTE of β_1 can be written as

$$\mathrm{R}\left[\hat{\beta}_1^{\mathrm{PTE}}\right] = \mathrm{R}\left[\tilde{\beta}_1\right] + g(\Delta) \tag{6.4.15}$$

where $h(\Delta) = e^{-a_1\Delta} G_{1,\nu}\,(c; \Delta^2) + a_1\Delta\, G_{3,\nu}(c/3;\, \Delta^2) - e^{a_1^2/2} G_{1,\nu}\,(c; (\Delta + a_1)^2)$.

- Under the null hypothesis, $\Delta = 0$ and hence the risk of the PTE of β_1 is

  $$\mathrm{R}\left[\hat{\beta}_1^{\mathrm{PTE}}\right] = \mathrm{R}\left[\tilde{\beta}_1\right] + G_{1,\nu}\,(c; 0) - e^{a_1^2/2} G_{1,\nu}\left(c; a_1^2\right). \tag{6.4.16}$$

  For any $a \neq 0$, $G_{1,\nu}\,(c; 0) - e^{a_1^2/2} G_{1,\nu}\,(c; a_1^2) < 0$. Therefore, at $\Delta = 0$, the risk of the PTE is less than that of the UE.

- For any positive (negative) value of a, if the value of $\Delta = \frac{\sqrt{S_{xx}}(\beta_1 - \beta_{10})}{\sigma}$ is positive, the value of $a_1\Delta\, G_{3,\nu}(c/3;\, \Delta^2)$ is also positive (negative). Therefore, for positive (negative) values of a, as Δ grows larger from zero, the risk of the PTE grows larger (smaller, reaching its minimum at some Δ depending on the magnitude of a, and then starts growing larger) and crosses the risk of the UE at

  $$\Delta = \frac{e^{a_1^2/2} G_{1,\nu}\,(c; (\Delta + a_1)^2) - e^{-a_1\Delta} G_{1,\nu}\,(c; \Delta^2)}{a_1\, G_{3,\nu}(c/3;\, \Delta^2)} \tag{6.4.17}$$

regardless of the value of a.

- As $\Delta \to \infty$, the values of $G_{1,\nu}(c; \Delta^2)$, $G_{1,\nu}(c; (\Delta + a_1)^2)$ and $G_{3,\nu}(c/3; \Delta^2)$ tend to zero, and hence, $\mathrm{R}\left[\hat{\beta}_1^{\mathrm{PTE}}\right] \to \left(e^{a_1^2/2} - 1\right)$, the risk of the UE of β_1. Therefore, starting from a certain large value of Δ, the risk of the PTE is no different from that of the UE.

- For very small values of a, the growth pattern of the risk for both positive and negative values of Δ are similar, because for very small values of a, the linex loss function reduces to the squared error loss function.

From the foregoing analyses and Figure 6.2 it is clear that the risk function of the PTE, and hence the efficiency of the PTE relative to the UE, depend on the three factors, namely, the level of significance α, non-centrality parameter Δ, and shape parameter a of the linex loss function. The value of a is deter-

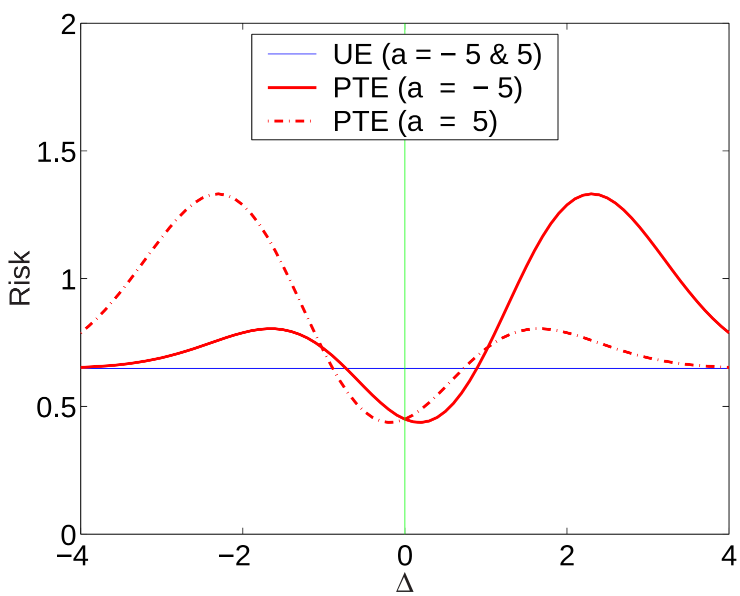

Figure 6.2: The risk of the PTE of the slope under the linex loss function for $\alpha = 0.20$, $n = 25$ $\sigma = 1$, and selected values of a.

mined by the experimenter according to the potential impact of the positive and negative errors of estimation, and the value of Δ is usually unknown to the experimenter. Regardless of the values of a and Δ, the risk of the PTE is a function of α. The question is which value of α should be used for the preliminary test? To answer this question, the efficiency of the PTE relative to the UE is used in the following subsection.

6.4.2 Determination of the Optimum Value of α

As a function of Δ and α, the efficiency function of the PTE of β_1 relative to the UE is given by

$$\mathrm{Eff}\left[\hat{\beta}_1^{\mathrm{PTE}}; \alpha, \Delta\right] = \left[e^{a_1^2/2} - 1\right]\left[e^{a_1^2/2} - 1 + h(\Delta)\right]^{-1} \tag{6.4.18}$$

where $h(\Delta) = e^{-a_1\Delta} G_{1,\nu}\,(c; \Delta^2) + a_1\Delta\, G_{3,\nu}(c/3;\, \Delta^2) - e^{a_1^2/2} G_{1,\nu}\,(c; (\Delta + a_1)^2)$. Figure 6.3 displays the efficiency of the PTE of the slope parameter relative to the UE, for a range of values of Δ.

From the analyses of the risk functions of the UE and PTE it is evident that the PTE does not have uniform domination over the UE for all values of Δ, and the value of Δ is usually unknown to the experimenter. Thus, we pre-assign a value of the relative efficiency, say Eff_o, that we are willing to accept. Consider the set

$$A_\alpha = \left(\alpha \mid \mathrm{Eff}\left[\hat{\beta}_1^{\mathrm{PTE}}; \alpha, \Delta\right] \geq \mathrm{Eff}_o\right) \tag{6.4.19}$$

for all Δ. An estimator $\hat{\beta}_1^{\mathrm{PTE}}$ is chosen which maximizes $\mathrm{Eff}\left[\hat{\beta}_1^{\mathrm{PTE}}; \alpha, \Delta\right]$ over all $\alpha \in A_\alpha$ and for all values of Δ. Thus we solve

$$\max_\alpha \min_\Delta \mathrm{Eff}\left[\hat{\beta}_1^{\mathrm{PTE}}; \alpha, \Delta\right] = \mathrm{Eff}_o \tag{6.4.20}$$

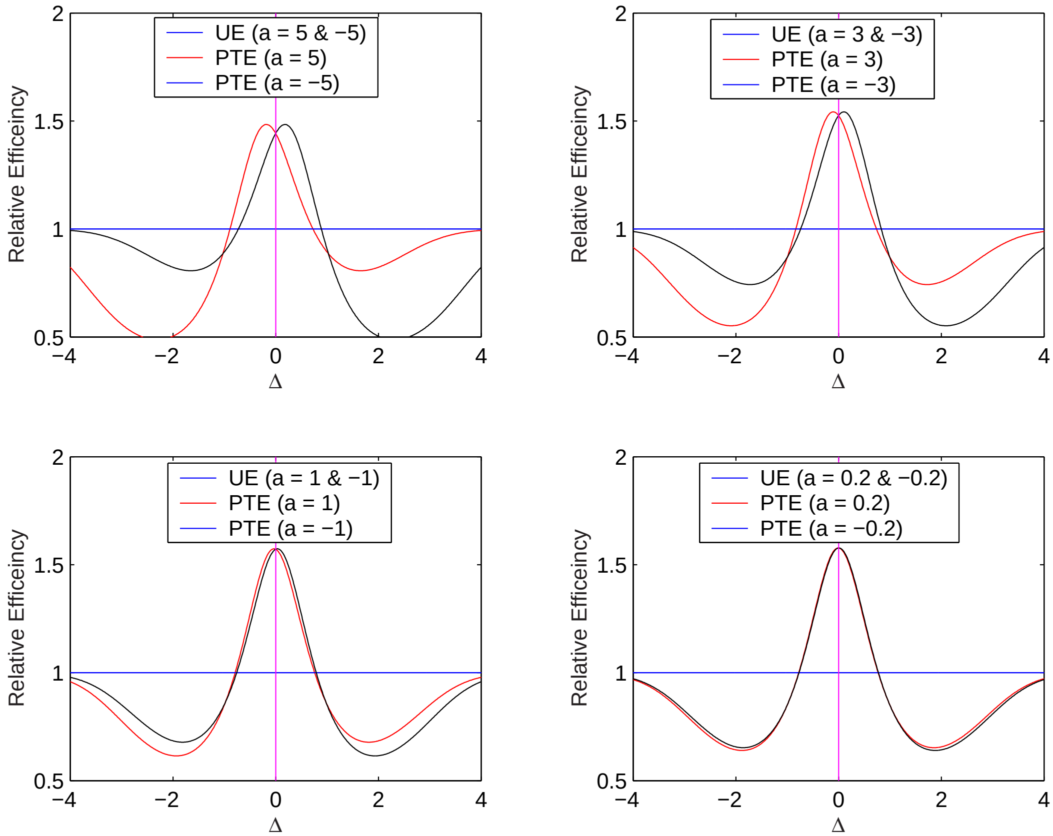

Figure 6.3: The efficiency of the PTE relative to the UE of β_1 under the linex loss function for $\alpha = 0.20$, $n = 25$ $\sigma = 1$, and selected values of a.

for α. The solution of this equation provides the maximum and minimum guaranteed efficiencies of the PTE of β_1 relative to the UE, for any selected values of n and Δ. Tables 6.1 and 6.2 present the maximum guaranteed efficiency (Eff*) and minimum guaranteed efficiency (Eff$_0$) of the PTE of β_1 relative to the UE, and the value of Δ (Δ_o) at which the minimum guaranteed efficiency occurs, for selected values of a, n and α. For example, if $a = 1$ and $n = 20$, and the experimenter wishes to achieve the minimum guaranteed efficiency 0.6055 of the PTE of β_1, the recommended value of α is 0.20.

6.5 Concluding Remarks

The risk functions of the UE and PTE of the slope parameter under the linex loss function are derived. The analytical, graphical, as well as the numerical analyses of the risk functions are also pursued. From the foregoing analysis it is revealed that under the linex loss function the PTE of β_1 outperforms the UE in the neighbourhood of Δ regardless of the value of a. Like the linex loss function, the growth of the risk of the PTE is also approximately linear on one side of $\delta = 0$ (left for $a > 0$), and approximately exponential on the other side (right for $a < 0$). However, if the magnitude of the shape parameter of the linex loss function is very close to zero, the growth pattern of the risk function of the PTE is nearly symmetrical.

6.A Appendix

6.A.1 Proof of Some Results

Here detailed proof of the result in (6.4.13) as well as some generalization of second order expectation is provided.

From (6.4.12), the moment generating function of the PTE of β_1 is given by

$$m(a) = e^{-a_1\Delta} \int_0^c f_{F(1,\nu,\Delta^2)}(y)\, dy + e^{a_1^2/2} \int_c^\infty f_{F(1,\nu,(\Delta+a_1)^2)}(y)\, dy \qquad (6.A.1)$$

where $c = F_{1,\nu}(\alpha)$.

Differentiating (6.A.1) with respect to a and using Lemma 6.28. gives

$$
\begin{aligned}
m'(a) = \frac{\sigma}{\sqrt{S_{xx}}}\Big[&-\Delta\, e^{-a_1\Delta}\int_0^c f_{F(1,\nu,\Delta^2)}(y)\,dy + a_1 e^{a_1^2/2}\int_c^\infty f_{F(1,\nu,(\Delta+a_1)^2)}(y)\,dy \\
&+ 2(\Delta+a_1)e^{a_1^2/2}\int_c^\infty\left(-\frac{1}{2}f_{F(1,\nu,(\Delta+a_1)^2)}(y) + \frac{1}{6}f_{F(3,\nu,(\Delta+a_1)^2)}(y/3)\right)dy \\
= \frac{\sigma}{\sqrt{S_{xx}}}\Big[&-\Delta\, e^{-a_1\Delta}\int_0^c f_{F(1,\nu,\Delta^2)}(y)\,dy + a_1 e^{a_1^2/2}\int_c^\infty f_{F(1,\nu,(\Delta+a_1)^2)}(y)\,dy \\
&- (\Delta+a_1)e^{a_1^2/2}\int_c^\infty f_{F(1,\nu,(\Delta+a_1)^2)}(y)\,dy \\
&+ e^{a_1^2/2}(\Delta+a_1)\int_{c/3}^\infty f_{F(3,\nu,(\Delta+a_1)^2)}(y)\,dy\Big] \\
= \frac{\sigma}{\sqrt{S_{xx}}}\Big[&-\Delta\, e^{-a_1\Delta}\int_0^c f_{F(1,\nu,\Delta^2)}(y)\,dy - \Delta e^{a_1^2/2}\int_c^\infty f_{F(1,\nu,(\Delta+a_1)^2)}(y)\,dy \\
&+ e^{a_1^2/2}(\Delta+a_1)\int_{c/3}^\infty f_{F(3,\nu,(\Delta+a_1)^2)}(y)\,dy\Big]. \qquad (6.A.2)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
m'(0) &= \frac{\sigma}{\sqrt{S_{xx}}}\left[-\Delta\int_0^c f_{F(1,\nu,\Delta^2)}(y)\,dy - \Delta\int_c^\infty f_{F(1,\nu,\Delta^2)}(y)\,dy \right.\\
&\qquad \left. + \Delta\int_{c/3}^0 f_{F(3,\nu,\Delta^2)}(y)\,dy\right] \\
&= \frac{\sigma}{\sqrt{S_{xx}}}\left[-\Delta\int_0^c f_{F(1,\nu,\Delta^2)}(y)\,dy - \Delta\left(1-\int_0^c f_{F(1,\nu,\Delta^2)}(y)\,dy\right)\right.\\
&\qquad \left. + \Delta\left(1-\int_0^{c/3} f_{F(3,\nu,\Delta^2)}(y)\,dy\right)\right]. \qquad (6.A.3)
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\mathrm{E}[\Phi] &= -(\beta_1-\beta_{10})\int_0^{c/3} f_{F(3,\nu,\Delta^2)}(y)\,dy \\
&= -\frac{\sigma}{\sqrt{S_{xx}}}\Delta\, G_{3,\nu}(F_{1,\nu}(\alpha);\,\Delta^2) \qquad (6.A.4)
\end{aligned}
$$

is the bias function of the PTE of β_1.

Now, we find the second moment, or equivalently, the mean square error function of the PTE of β_1.

Differentiating both sides of (6.A.2) with respect to a, we get

$$m''(a) = \frac{\sigma^2}{S_{xx}} \Bigg[\Delta^2 e^{-a_1\Delta} \int_0^c f_{F(1,\nu,\Delta^2)}(y)\, dy - \Delta a_1 e^{a_1^2/2} \int_c^\infty f_{F(1,\nu,(\Delta+a_1)^2)}(y)\, dy$$
$$- 2\Delta(\Delta+a_1) e^{a_1^2/2} \int_c^\infty \Big(-\frac{1}{2} f_{F(1,\nu,(\Delta+a_1)^2)}(y) + \frac{1}{6} f_{F(3,\nu,(\Delta+a_1)^2)}(y/3) \Big)\, dy$$
$$+ a_1 e^{a_1^2/2}(\Delta+a_1) \int_{c/3}^\infty f_{F(3,\nu,(\Delta+a_1)^2)}(y)\, dy + e^{a_1^2/2} \int_{c/3}^\infty f_{F(3,\nu,(\Delta+a_1)^2)}(y)\, dy$$
$$+ 2e^{a_1^2/2}(\Delta+a_1)^2 \int_{c/3}^\infty \Big(-\frac{1}{2} f_{F(3,\nu,(\Delta+a_1)^2)}(y) + \frac{3}{10} f_{F(5,\nu,(\Delta+a_1)^2)}(3y/5) \Big) dy \Bigg]. \tag{6.A.5}$$

Putting $a = 0$ in (6.A.5), we get

$$m''(0) = \frac{\sigma^2}{S_{xx}} \Bigg[\Delta^2 \int_0^c f_{F(1,\nu,\Delta^2)}(y)\, dy + \Delta^2 \int_c^\infty \Big\{ f_{F(1,\nu,\Delta^2)}(y)$$
$$- \frac{1}{3} f_{F(3,\nu,\Delta^2)}(y/3) \Big\}\, dy + 1 - \int_0^{c/3} f_{F(3,\nu,\Delta^2)}(y)\, dy$$
$$- \Delta^2 \int_{c/3}^\infty \Big\{ f_{F(3,\nu,\Delta^2)}(y) - \frac{3}{5} f_{F(5,\nu,\Delta^2)}(3y/5) \Big\}\, dy \Bigg]$$
$$= \frac{\sigma^2}{S_{xx}} \Bigg[\Delta^2 - \Delta^2 \int_{c/3}^\infty f_{F(3,\nu,\Delta^2)}(y)\, dy + 1 - \int_0^{c/3} f_{F(3,\nu,\Delta^2)}(y)\, dy$$
$$- \Delta^2 \int_{c/3}^\infty f_{F(3,\nu,\Delta^2)}(y)\, dy + \Delta^2 \int_{c/5}^\infty f_{F(5,\nu,\Delta^2)}(y)\, dy \Bigg]$$
$$= \frac{\sigma^2}{S_{xx}} \Bigg[\Delta^2 - 2\Delta^2 + 2\Delta^2 \int_0^{c/3} f_{F(3,\nu,\Delta^2)}(y)\, dy + 1$$
$$- \int_0^{c/3} f_{F(3,\nu,\Delta^2)}(u)\, du + \Delta^2 - \Delta^2 \int_0^{c/5} f_{F(5,\nu,\Delta^2)}(y)\, dy \Bigg]$$
$$= \frac{\sigma^2}{S_{xx}} \Bigg[1 - G_{3,\nu}\left(\frac{1}{3} F_{1,\nu}(\alpha);\, \Delta^2 \right) + \Delta^2 \Bigg\{ 2\, G_{3,\nu}\left(\frac{1}{3} F_{1,\nu}(\alpha);\, \Delta^2 \right)$$
$$- G_{5,\nu}\left(\frac{1}{5} F_{1,\nu}(\alpha);\, \Delta^2 \right) \Bigg\} \Bigg] \tag{6.A.6}$$

which is the second moment, or the mean square error function of the PTE of β_1. Similarly, equating the r^{th}-order derivative of (6.A.1) to zero, the r^{th} moment of the PTE of β_1 can be obtained.

6.A.2 Tables of Efficiencies of PTE Relative to UE

Table 6.1: Maximum and minimum efficiencies of the PTE of the slope relative to the UE for $a = 3$.

		Sample size, n						
α		10	15	20	25	30	35	40
0.05	Eff*	3.7012	3.6151	3.5913	3.5824	3.5785	3.5768	3.5761
	Eff_o	0.1583	0.2221	0.2497	0.2699	0.2841	0.2947	0.3033
	Δ_o	−3.1400	−3.0010	−2.6240	−2.5500	−2.4970	−2.4710	−2.3750
0.10	Eff*	2.3302	2.2955	2.2851	2.2809	2.2788	2.2776	2.2769
	Eff_o	0.2639	0.3269	0.3592	0.3793	0.3889	0.4040	0.4122
	Δ_o	−2.7400	−2.4710	−2.3480	−2.2990	−2.2680	−2.2370	−2.1955
0.15	Eff*	1.8259	1.8064	1.8003	1.7976	1.7963	1.7954	1.7949
	Eff_o	0.3589	0.4205	0.4515	0.4707	0.4841	0.4940	0.5017
	Δ_o	−2.5250	−2.3000	−2.2200	−2.1800	−2.1500	−2.1200	−2.0900
0.20	Eff*	1.5604	1.5480	1.5440	1.5422	1.5412	1.5406	1.5402
	Eff_o	0.4475	0.5054	0.5341	0.5519	0.5641	0.5732	0.5803
	Δ_o	−2.3855	−2.2190	−2.1430	−2.0790	−2.0489	−2.0400	−2.0190
0.25	Eff*	1.3971	1.3886	1.3886	1.3845	1.3838	1.3834	1.3831
	Eff_o	0.5308	0.5834	0.6092	0.6251	0.6361	0.6442	0.6504
	Δ_o	−2.2900	−2.1480	−2.0760	−2.0190	−1.9990	−1.9620	−1.9610
0.30	Eff*	1.2875	1.2816	1.2796	1.2786	1.2781	1.2778	1.2776
	Eff_o	0.6083	0.6547	0.6773	0.6911	0.7006	0.7077	0.7131
	Δ_o	−2.2200	−2.0900	−2.0200	−1.9790	−1.9790	−1.9490	−1.9190
0.35	Eff*	1.2101	1.2058	1.2044	1.2037	1.2033	1.2031	1.2029
	Eff_o	0.6795	0.7192	0.7384	0.7501	0.7581	0.7640	0.7686
	Δ_o	−2.1600	−2.0300	−1.9700	−1.9500	−1.9100	−1.8900	−1.8800
0.40	Eff*	1.1536	1.1505	1.1495	1.1490	1.1487	1.1485	1.1484
	Eff_o	0.7437	0.7765	0.7923	0.8019	0.8085	0.8134	0.8171
	Δ_o	−2.1100	−2.0000	−1.9500	−1.9050	−1.8790	−1.8830	−1.8560
0.45	Eff*	1.1115	1.1093	1.1085	1.1082	1.1079	1.1078	1.1077
	Eff_o	0.8002	0.8265	0.8391	0.8468	0.8520	0.8559	0.8588
	Δ_o	−2.0900	−1.9600	−1.9100	−1.8700	−1.8600	−1.8400	−1.8350
0.50	Eff*	1.0798	1.0783	1.0777	1.1445	1.1614	1.1776	1.1934
	Eff_o	0.8488	0.8691	0.8788	0.8847	0.8887	0.8917	0.8940
	Δ_o	−2.0500	−1.9500	−1.9000	−1.8500	−1.8400	−1.8300	−1.7900

Table 6.2: Maximum and minimum efficiencies of the PTE of the slope relative to the UE for $a = 1$.

α		Sample size, n						
		10	15	20	25	30	35	40
0.05	Eff*	4.2674	3.9892	3.8713	3.80640	3.7653	3.7369	3.7162
	Eff$_o$	0.2661	0.3076	0.3279	0.3279	0.3401	0.3401	0.3545
	Δ_o	−2.6300	−2.4700	−2.3850	−2.3950	−2.3550	−0.3250	−2.3004
0.10	Eff*	2.5615	2.4468	2.3978	2.3706	2.3534	2.3415	2.3328
	Eff$_o$	0.3807	0.4202	0.4393	0.4507	0.4585	0.4641	0.4684
	Δ_o	−2.3550	−2.2255	−2.1515	−2.1448	−2.1050	−2.1105	−2.0850
0.15	Eff*	1.9556	1.8907	1.8629	1.8474	1.8376	1.8308	1.8258
	Eff$_o$	0.4748	0.5112	0.5285	0.5390	0.5462	0.5511	0.5550
	Δ_o	−2.1750	−2.0850	−2.0390	−2.0100	−2.0055	−2.0025	−1.9750
0.20	Eff*	1.6429	1.6014	1.5835	1.5736	1.5673	1.5629	1.5597
	Eff$_o$	0.5573	0.5900	0.6055	0.6148	0.6211	0.6256	0.6291
	Δ_o	−2.0610	−1.9800	−1.9500	−1.9302	−1.9100	−1.9000	−1.8950
0.25	Eff*	1.4530	1.4247	1.4125	1.4057	1.4014	1.3984	1.3962
	Eff$_o$	0.6310	0.6597	0.6733	0.6814	0.6868	0.6908	0.6938
	Δ_o	−1.9850	−1.9250	−1.8950	−1.8755	−1.8609	−1.8550	−1.8500
0.30	Eff*	1.3268	1.3068	1.2982	1.2934	1.2904	1.2883	1.2867
	Eff$_o$	0.6969	0.7215	0.7331	0.7400	0.7446	0.7480	0.7506
	Δ_o	−1.925	−1.8650	−1.8459	−1.8255	−1.8100	−1.8080	−1.7968
0.35	Eff*	1.2382	1.2239	1.2177	1.2143	1.2121	1.2105	1.2094
	Eff$_o$	0.7553	0.7759	0.7856	0.7913	0.7952	0.7980	0.8001
	Δ_o	−1.8759	−1.8260	−1.7992	−1.7900	−1.7890	−1.7645	−1.7700
0.40	Eff*	1.1739	1.1635	1.1590	1.1565	1.1549	1.1538	1.1530
	Eff$_o$	0.8064	0.8232	0.8311	0.8357	0.8389	0.8412	0.8429
	Δ_o	−1.8352	−1.7989	−1.7777	−1.7657	−1.7559	−1.7325	−1.7300
0.45	Eff*	1.1261	1.1186	1.1154	1.1136	1.1124	1.1116	1.1111
	Eff$_o$	0.8429	0.8637	0.8629	0.8716	0.8700	0.8619	0.8600
	Δ_o	−1.8020	−1.7659	−1.7592	−1.7434	−1.7413	−1.7400	−1.7375
0.50	Eff*	1.0902	1.0849	1.0826	1.0813	1.0805	1.0799	1.0795
	Eff$_o$	0.8876	0.8978	0.9025	0.9053	0.9072	0.9086	0.9096
	Δ_o	−1.7750	−1.7450	−1.7236	−1.7100	−1.70001	−1.6959	−1.7100

6.A.3 MATLAB Codes

- The following MATLAB codes are used for producing Figure 6.2.

```
n=20; v=n-2; d=-4:.1:4; s=1; D=d.^2;
a=-5; a1=a.*s./sqrt(n);
x=a1.^2./2; r1=exp(x)-1; Q=r1.*ones(1, length(d));
plot(d, Q, 'b')
hold on
a2=0.8; c1=finv(a2,1,v);
G1=ncfcdf(c1, 1, v, D); G2=ncfcdf(c1, 1,
v, (d+a1).^2); G3=ncfcdf(c1./3, 3, v, D);
r3=exp(-a1.*d).*G1+exp(x).*(1-G2)+a1.*d.*G3-1;
plot(d, r3, 'r')
a=5; a1=a.*s./sqrt(n); x=a1.^2./2; a2=0.8;c1=finv(a2,1,v);
G1=ncfcdf(c1, 1, v, D); G2=ncfcdf(c1, 1, v, (d+a1).^2);
G3=ncfcdf(c1./3, 3, v, D);
r3=exp(-a1.*d).*G1+exp(x).*(1-G2)+a1.*d.*G3-1;
plot(d, r3, 'r-.')
xlabel('\Delta'); ylabel('Risk');
legend('UE (a = -5 & 5','PTE (a= -5)', 'PTE(a= 5)', 0);
ly = linspace(0, max(r3), 10000); lx =
linspace(0, 0, 10000);
hold on;
plot(lx,ly, 'g')
```

Chapter 7

Conclusions

In this thesis we have considered the UE, SRE, SPTE and SE of each of the slope and intercept parameters of the simple linear regression model with normal errors. With respect to the bias and mean square error criteria the performances of the SRE, SPTE and SE were compared with those of the exclusive sample information based UE. It is revealed that although there is no uniform domination of one estimator over the others, under certain conditions the SE outperforms the other estimators considered in this thesis. This study would guide the practitioners of the estimators in choosing an appropriate estimator of the slope or intercept parameter of simple linear regression model with minimum bias and mean square error.

We defined the SPTE of the coefficient vector of multiple linear regression model with normal errors under the original, modified and Edgeworth size-corrected W, LR and LM tests. With respect to the quadratic bias and quadratic risk criteria the performances of the SPTEs were investigated. It is revealed that there is a great deal of conflict among the properties of the original W, LR and LM test based SPTEs. The use of the modified test reduces the conflict among the properties of the SPTEs to some extent but remains considerable. However, the use of size-corrected tests almost eliminates the conflict among the SPTEs. This study recommends the practitioners of the SPTE that they should choose any of the three size-corrected tests without risking of conflicting performance of the estimator.

Using the preliminary test approach we defined the SE of the coefficient vector of the multiple linear regression model under the original W, LR and LM test statistics. The performances of the estimators were investigated under the quadratic bias and quadratic risk criteria. It is revealed that with respect to the bias of the estimators, the W test statistic-based SE outperforms the LR and LM test statistics based SEs. However, with respect to the quadratic risk criterion, the LM test statistic based SE outperforms the LR and LM test statistics-based SEs, under certain conditions. This study would guide the practitioners of the SE of the multiple linear regression model in choosing an optimal test statistic in the sense of having minimum quadratic risk of the estimator.

Finally, the performance of the PTE of the slope parameter of the simple linear regression model was investigated under the linex loss function. It is revealed that if the non-sample prior information about the value of the parameter is not too far from the true value, the PTE outperforms the UE. Like the linex loss function, the risk of the PTE under the linex loss function is asymmetric in nature. However, for small values of the shape parameter of the loss function, the risk of the PTE is symmetric. This study enables the practitioners of the PTE of the slope parameter in determining the risk of the estimator under the asymmetric loss that takes into account the problem of underestimation and overestimation.

Further work in this area may include investigations on the following issues.

1. Under certain conditions the SE of the slope parameter of the simple linear regression model outperforms the UE with respect to the squared error loss function. Does the SE outperforms the UE with respect to the linex loss function? Does the PTE and SE of the intercept parameter outperform the UE with respect to the linex loss function?

2. The SPTE and SE of the slope and intercept parameters of the simple linear regression model have been defined under the LR test. If these estimators are defined under the W and LM tests, would there be any conflict among the statistical properties of the estimators? If so, how might that that could be minimized?

3. How do the SPTE and SE of the coefficient vector of multiple linear regression model perform under the linex loss function?

4. The moment generating function of the PTE of the slope parameter of the simple linear regression model has been derived, from which the first four moments of the PTE are easily obtainable. What are the shapes of the sampling distributions of the PTE of the slope and intercept parameter?

5. It is well known that the normal distribution is not robust and is not suitable to model all problems. On the other hand, the Student's t distribution is robust and more typical member of the elliptical class of distributions. The studies undertaken in this thesis can be extended to the models with non-normal errors, such as Student's t errors.

Bibliography

Ahmed, S.E. and Saleh, A.K.Md.E. (1989). Pooling multivariate data, *Journal of Statistical Computation and Simulation*, **31**, 149-167.

Ahsanullah, M. and Saleh, A.K.Md.E. (1972). Estimation of intercept in a linear regression model with one dependent variable after a preliminary test on the regression coefficient. *International Statistical Review*, **40**, 139-145.

Aitchison, J. and Dunsmore, I.R. (1975). Statistical Prediction Analysis, *Cambridge University Press*, London.

Aitchison, J. and Silvey, S.D. (1958). Maximum likelihood estimation of parameters subject to restraints, *Annals of Mathematical Statistics*, **29**, 813-828.

Ashish, S. and Srivastava, M. (1997). Regression Analysis, Theory, Methods and Applications, *Springer*, New York.

Bancroft, T.A. (1944). On biases in estimation due to the use of the preliminary test of significance. *Annals of Mathematical Statistics*, **15**, 190-204.

Berndt, E. and Savin, N.E. (1977). Conflict among criteria for testing hypotheses in the multivariate linear regression model, *Econometrica*, **45**, 1263-1278.

Berger, J.O. (1980). Statistical Decision Theory: Foundations, Concepts, and Methods, *Springer-Verlag*, New York.

Bhattacharaya, D., Samaniego, F.J. and Vestrup, E.M. (2002). On the comperative performance of bayesian and classical point estimators under asymmetric loss, *Sankhya*, Series B, **64**(3), 239-266

Bhoj, D.S. and Ahsanullah, M. (1994). Estimation of a conditional mean in a linear regression model after a preliminary test on regression coefficient. *Biometrical Jourrnal*, **36**, 153-163.

Bhoj, D.S. and Ahsanullah, M. (1993). Estimation of a conditional mean in a regression model. *Biometrical Journal*, **35**, 791-799.

Billah, M.B. (1997). Small sample studies of the W, LR, and LM based SPTEs of regression parameters, *Journal of Statistical Research*, **30**(2), 93-108.

Billah, M.B. and Saleh, A.K.Md.E. (1998). Conflict between pretest estimators induced by three large sample tests under a regression model with Students t-error, *the Statistician*, **47**(4), 593-606.

Billah, M.B. and Saleh, A.K.Md.E. (2000a). Performance of the PTE based on the conflicting W, LR and LM tests in regression model, *Advances on Methodological and Applied Aspects of Probability and Statistics*, N. Balakrishnan (ed.), Gordon and Breach Publishing Company, 263-276.

Billah, M.B. and Saleh, A.K.Md.E. (2000b). Performance of three large sample tests in the actual formation of the pretest estimators for regression model with t-error, *Journal of Applied Statistical Science*, **9**(3), 237-251.

Bolfarine, H. and Zacks, S. (1992). Prediction Theory for Finite Populations, *Springer-Verlag*, New York.

Breusch, T.S. (1979). Conflict among criteria among testing hypothesis: Extensions and comments,*Econometrica*, **47**, 203-207.

Chiou, P. and Saleh, A.K.Md.E. (2002). Preliminary test confidence sets for the mean of a multivariate normal distribution, *Journal of Propagation in Probability and Statistics*, **2**, 177-189.

Engle, R.F. (1984). Wald, likelihood ratio, and Lagrange multiplier tests in econometrics, in *Handbook of Econometrics*, **II**, Edited by Z. Griliches and M.D. Intriligator, 776-826, North-Holland.

Evans, G.B.A. and Savin, N.E. (1980). Conflict among testing procedures in a linear regression model with lagged dependent variables, invited paper presented at *Word Congress of Econometric Society*, Aix-en-Provence.

Evans, G.B.A. and Savin, N.E. (1982). Conflict among the criteria revisited: the W, LR and LM tests, *Econometrica*, **50**(3), 737-748.

Evans, M. Hastings, N. and Peacock, B. (2000). Statistical Distributions, *John Wiley & Sons*.

Ferguson, T.S. (1967). Mathematical Statistics: A Decision Theoretic Approach, *Academic Press*, New York.

Gallant, A.R. (1975). Seemingly unrelated nonlinear regression, *Journal of Econometrics*, **3**, 35-50.

Han, C.P. and Bancroft, T.A. (1968). On pooling means when variance is unknown. *Journal of American Statistical Association*, **63**, 133-1342.

James, W. and Stein, C. (1961). Estimation with quadratic loss, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Ulcal Press, Berkeley, California, 361-379.

Judge, G.G. and Bock, M.E. (1978). The Statistical Implication of Pre-test and Stein-rule Estimators in Econometrics, *North-Holland*, New York.

Khan, S. (1998). On the estimation of the mean vector of Student-t population with uncertain prior information, *Pakistan Journal Statistics*, **14**, 161-175.

Khan, S. and Saleh, A.K.Md.E. (1995). Preliminary test estimators of the mean based on p-samples from multivariate Student-t populations, *Bulletin of the International Statistical Institute*, 50th Session of the ISI, Beijing, 599-600.

Khan, S. and Saleh, A.K.Md.E. (1997). Shrinkage pre-test estimator of the intercept parameter for a regression model with multivariate Student-t errors, *Biom Jour*, **39**, 1-17.

Khan, S. and Saleh, A.K.Md.E. (1998). Comparison of the estimators of the mean based on p-samples from multivariate student-t population, *Communications in Statistics, Theory and Methods*, **27**(1), 193-210.

Khan, S. and Saleh, A.K.Md.E. (2001). On the comparison of the pre-test and shrinkage estimators for the univariate normal mean, *Statistical Papers*, **42**(4), 451-473.

Khan, S. and Hoque, Z. (2002). James-Stein estimators for the mean vector of a multivariate normal population based on independent samples from two normal populations with common covariance structure, *Pakistan Journal of Statistics*, **18**(3), p. 359-381.

Khan, S. and Hoque, Z. (2003). Preliminary test estimators for the multivariate normal mean based on three large sample tests, *Journal of Statistical Research*, **37**(1), p. 43-55.

Khan, S., Hoque, Z. and Saleh, A.K.Md.E. (2002). Estimation of the slope parameter for linear regression model with uncertain prior information, *Journal of Statistical Research*, **36**(1), 55-73.

Kibria, B.M.G. (2002). Simultaneous shrinkage estimation of the mean vector of a multivariate normal distribution, *Journal of Statistical Research*, **36**(1), 29-44.

Maatta, J.M. and Casella, G. (1990). Developments in decision-theoretic variance estimation, *Statistical Science*, **5**, 90-101.

Pandey, B.N. and Rai, O. (1996). Testimation of mean of a normal under linex loss function. *Aligarh Journal of Statistics*, **15** and **16**, 35-40.

Pandey, B.N. and Rai, O. (1992). Bayesian estimation of mean and square of mean of normal distribution using linex loss function, *Communications in Statistics, Theory and Methods*, **21**(12), 3369-3391

Parsian, A. and Farispour, N.S. (1993). On the admissibility and inadmissibility of estimators of scale parameters using an asymmetric loss function, *Communications in Statistics, Theory and Methods*, **22**(10), 2877-2901.

Parsian, A., Farispour, N.S. and Nematollahi, N. (1993). On the minimaxity of pitman type estimator under a linex loss function, *Communicatipns in Statistics, Theory and Methods*, **22**(1), 97-113.

Rao, C.R. (1947). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation, *Proceedings of the Cambridge Philosophical Society*, **44**, 50-57.

Rojo, J. (1987). On the admissibility of $c\bar{X} + d$ with respect to the linex loss function, *Communicatipns in Statistics, Theory and Methods*, **16**(12), 3745-3748.

Rothenberg, T.J. (1977). Edgeworth expansions for multivariate test statistics, *Working Paper in Economic Theory and Econometrics IP-255*, Center for Research in Management Science, Institute of Business and Economic Research, University of California, Berkeley.

Saleh, A.K.Md.E. and Sen, P.K. (1978). Nonparametric estimation of location parameter after a preliminary test on regression, *Annals of Statistics*, **6**, 154-168.

Saleh, A.K.Md.E. and Sen, P.K. (1979). Nonparametric estimation of location parameter after a preliminary test on regression in the multivariate case, *Journal of Multivaraite Analysis*, **9**, 154-168.

Saleh, A.K.Md.E. and Sen, P.K. (1984). Nonparametric preliminary test inference, *Handbook of Statistics*, **4**, Krishnaiah, P.R. and Sen, P.K. (eds.), North-Holland,

Saleh, A.K.Md.E. and Sen, P.K. (1985). Shrinkage least squares estimation in a general multivariate linear model, *Proceedings of the Fifth Pannonian Symposium on Mathematical Statistics*, 307-325. Amsterdam, 275-297.

Saleh, A.K.Md.E. and Sen, P.K. (1986). On shrinkage R-estimation in a multiple regression model, *Communication in Statistics, Theory and Methods*, **15**(7), 2229-2244.

Savin, N.E. (1976). Conflict among testing procedures in a linear regression model with autoregressive disturbances, *Econometrica*, **44**, 1303-1315.

Sclove,S.L., Morris, C. and Rao, C.R. (1972). Non-optimality of preliminary-test estimators for the mean of a multivariate normal distribution, *Annals Mathematical Statistics*, **43**, 1481-1490.

Silvey, S.D. (1959). The lagrange multiplier test, *Annals of Mathematical Statistics*, **30**, 389-407.

Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Loss Angeles, Uni. of California Press, **1**, 197-206.

Stein, C. (1981). Estimation of the mean of a multivariate normal distribution, *Annals of Statistics*, **9**, 1135-1151.

Ullah, A. and Zinde-Walsh, V. (1984). On the robustness of LM, LR and W tests in regression models, *Econometrica*, **52**, 1055-1066.

Varian, H.R. (1975). A Bayesian approach to real estate assessment, in *Studies in Bayesian Econometrics and Statistics in honor of L.J. Savage, eds., Feinberg and A. Zellner*, North Holland, Amsterdam, 195-208.

Wald, A. (1943). Tests of hypotheses concerning several parameters when the number of observations is large, *Transaction of the American Mathematical Society*, **54**, 426-482.

Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions, *Journal of the American Statistical Association, Theory and Methods*, **81**(394), 446-451.

Zellner, A. and Geisel, M.S. (1968). Sensitivity of control to uncertainty and form of the criterion foundation, *in the Future of Statistics, ed. Donald G. Watts*, New York: Academic Press, 269-289.