CUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

PHD THESIS

Issam DAWOUD

PREDICTIONS IN LINEAR MODELS

DEPARTMENT OF STATISTICS

ADANA, 2016

CUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

PREDICTIONS IN LINEAR MODELS

Issam DAWOUD

PHD THESIS

DEPARTMENT OF STATISTICS

We certify that the thesis titled above was reviewed and approved for the award of degree of the Doctorate of Philosophy by the board of jury on 28/06/2016.

…….....	…….....	……..…….....
Prof. Dr. Selahattin KAÇIRANLAR	Prof. Dr. M. Revan ÖZKALE	Asst. Prof. Dr. Hüseyin GÜLER
SUPERVISOR	MEMBER	MEMBER

…….....	……
Prof. Dr. Zeki YILDIZ	Prof. Dr. M. Akif BAKIR
MEMBER	MEMBER

This PHD Thesis is written in the Institute of Natural and Applied Sciences at Cukurova University.
Registration Number:

Prof. Dr. Mustafa GÖK
Director
Institute of Natural and Applied Sciences

Note: The usage of the presented specific declarations, tables, figures, and photographs either in this thesis or in any other reference without citation is subject to "The law of Arts and Intellectual Products" number of 5846 of Turkish Republic.

ABSTRACT

PHD THESIS

PREDICTIONS IN LINEAR MODELS

Issam DAWOUD

CUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF STATISTICS

Supervisor : Prof. Dr. Selahattin KAÇIRANLAR
Year: 2016, Pages: 99
Jury : Prof. Dr. Selahattin KAÇIRANLAR
: Prof. Dr. M. Revan ÖZKALE
: Asst. Prof. Dr. Hüseyin GÜLER
: Prof. Dr. Zeki YILDIZ
: Prof. Dr. M. Akif BAKIR

Regression models are widely used in prediction to predict new future values which provides information about the potential future events and their consequences. Although the ordinary least squares estimator is the best linear unbiased estimator of regression parameters in a linear regression model, we can improve upon the variability of an estimator or a predictor of a regression coefficient when the unbiasedness criterion can be skipped. Also, the predictive performance of a regression model can be adversely affected by both multicollinearity and correlated errors. In spite of biased estimation procedures have been proposed as an alternative to least squares, there has been little analysis of the predictive performance of the resulting equations. This study finds the optimal estimators and predictors of the extended balanced loss function comparing them according to some criteria, and discusses the predictive performance of various biased estimators under only multicollinearity and under multicollinearity and correlated errors simultaneously in terms of the prediction mean squared error. A simulation and numerical examples are conducted to compare the resulting equations.

Keywords: Regression Model, Optimal Estimator, Multicollinearity, Correlated Errors, Prediction Mean Square Error

ÖZ

DOKTORA TEZİ

LİNEER MODELLERDE KESTİRİMLER

Issam DAWOUD

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İSTATİSTİK ANABİLİM DALI

Danışman : Prof. Dr. Selahattin KAÇIRANLAR
Yıl: 2016, Sayfa: 99
Jüri : Prof. Dr. Selahattin KAÇIRANLAR
: Prof. Dr. M. Revan ÖZKALE
: Yrd. Doç. Dr. Hüseyin GÜLER
: Prof. Dr. Zeki YILDIZ
: Prof. Dr. M. Akif BAKIR

Regresyon modelleri, gelecekteki muhtemel olaylar ve bu olayların sonuçları ile ilgili bilgiler içeren yeni değerlerin kestirimini yapmak için yaygın olarak kullanılmaktadır. En küçük kareler tahmin edici lineer regresyon modelindeki regresyon parametrelerinin tahmini için en iyi lineer yansız tahmin edicidir. Öte yandan, yansızlık kriteri göz ardı edilirse, regresyon parametrelerinin tahmini ya da kestirimi için çeşitli tahmin yöntemleri geliştirilebilir. Ayrıca, lineer regresyon modelindeki açıklayıcı değişkenler arasında çoklu iç ilişki varlığında ve hataların ilişkili olması durumunda kestirim performansı bu durumlardan kötü bir şekilde etkilenebilmektedir. Yanlı tahmin yöntemleri en küçük kareler tahmin ediciye alternatif olarak kullanılmasına rağmen kestirim performansları ile ilgili çok az çalışma bulunmaktadır. Bu çalışmada, genişletilmiş dengeli kayıp fonksiyonuna göre en iyi tahmin ediciler ile kestirimciler elde edilmiş ve farklı kriterlere göre karşılaştırılmıştır. Ayrıca, çeşitli yanlı tahmin edicilerin kestirim performansları çoklu iç ilişki varlığında ve hem çoklu iç ilişki hem de hataların ilişkili olması durumunda kestirimci hata kareler ortalaması kriterine göre incelenmiştir. Son olarak, elde edilen sonuçları incelemek için sayısal örnek ve Monte Carlo simülasyon çalışması yapılmıştır.

Anahtar Kelimeler: Regresyon Modeli, En İyi Tahmin Edici, Çoklu İç İlişki, İlişkili Hatalar, Kestirimci Hata Kareler Ortalaması

ACKNOWLEDGEMENTS

First, I would like to thank my supervisor, Prof. Dr. Selahattin KAÇIRANLAR. I am grateful to him for his guidance and the opportunities he has afforded me. He is incredibly organized and a great problem solver, both of these qualities were immensely helpful in moving my thesis forward.

Second, I most gratitude to my thesis committee members for their contributions to this work.

Third, I am grateful to Çukurova University in Turkey for offering me the opportunity to get a Ph.D. Degree in Statistics, and my thanks to all Professors who teach me in the Statistics department.

Fourth, I would also like to thank my colleagues in the Statistics program for their encouragement and friendship.

Last but not least, special and great thank goes to my parents for their endless love, support, and encouragement.

PAGE

CONTENTS

ABSTRACT.....I
ÖZ.....II
ACKNOWLEDGEMENTS.....III
CONTENTS.....IV
LIST OF TABLES.....VI
LIST OF FIGURES.....VIII
LIST OF ABBREVIATIONS.....X
1. INTRODUCTION.....1
2. THE PERFORMANCE OF THE OPTIMAL EXTENDED BALANCED LOSS FUNCTION ESTIMATORS AND PREDICTORS.....5
2.1. Introduction.....5
2.2. Linear Regression Model with Some Concepts and the Loss Functions.....6
2.3. The Optimal Extended Balanced Loss Function Estimators and Predictors and Their Performances.....10
2.4. The Monte Carlo Simulation Study.....21
2.5. Conclusion.....23
3. PREDICTION OF THE LINEAR CONVEX COMBINATION OF BIASED REGRESSION ESTIMATORS.....27
3.1. Introduction.....27
3.2. Evaluations of the Prediction Mean Squared Errors.....32
3.3. Comparisons of Prediction Mean Squared Errors in Two Dimensional Space.....36
3.4. Numerical Example.....51
3.5. Conclusion.....58
4. THE PREDICTIVE PERFORMANCE EVALUATION OF BIASED REGRESSION PREDICTORS WITH CORRELATED ERRORS.....61

4.1. Introduction.....61
4.2. The Prediction Problem.....62
4.3. The Properties of the Predictors.....65
4.4. The Comparisons of the Prediction Mean Squared Errors in Two Dimensional Space.....71
4.5. Numerical Example.....82
4.6. Conclusion.....89
5. CONCLUSIONS.....91
REFERENCES.....93
CURRICULUM VITAE.....99

LIST OF TABLES **PAGE**

Table 2.1. The mse estimates when $\sigma = 0.1$.....24

Table 2.2. The mse estimates when $\sigma = 1$.....24

Table 2.3. The mse estimates when $\sigma = 10$.....25

Table 3.1. c and α_2^2 values for the LOR vs. the OLS.....52

Table 3.2. c and α_2^2 values for the LOR vs. the ORR.....54

Table 3.3. c and α_2^2 values for the LOL vs. the OLS.....55

Table 3.4. c and α_2^2 values for the LOL vs. the Liu.....56

Table 3.5. c and α_2^2 values for the LOL vs. the LOR.....58

Table 4.1. ρ and α_2^2 values for the LRP vs. the BLUP.....84

Table 4.2. ρ and α_2^2 values for the LLP vs. the BLUP.....86

Table 4.3. ρ and α_2^2 values for the LLP vs. the LRP.....88

LIST OF FIGURES **PAGE**

Figure 3.1. The pmse comparison of the LOR and the OLS.....52
Figure 3.2. The pmse comparison of the LOR and the ORR.....53
Figure 3.3. The pmse comparison of the LOL and the OLS.....55
Figure 3.4. The pmse comparison of the LOL and the Liu.....56
Figure 3.5. The pmse comparison of the LOL and the LOR 57
Figure 4.1. The pmse comparisons between the LRP and the BLUP with different RHOs.....83
Figure 4.2. The pmse comparisons between the LLP and the BLUP with different RHOs.....85
Figure 4.3. The pmse comparisons between the LLP and the LRP with different RHOs.....87

LIST OF ABBREVIATONS

OLS	: Ordinary Least Squares
MSE	: Mean Square Error
mse	: Scalar MSE
ORR	: Ordinary Ridge Regression
GLS	: Generalized Least Squares
BLUE	: Best Linear Unbiased Estimator
PMSE	: Prediction Mean Square Error
pmse	: Scalar PMSE
BLUP	: Best Linear Unbiased Predictor
ZBLF	: Zellner's Balanced Loss Function
PLF	: Predictive Loss Function
EBLF	: Extended Balanced Loss Function
LOR	: Combination of OLS and ORR
LOL	: Combination of OLS and Liu
pd	: Positive Definite
LRP	: Linear Ridge Predictor
LLP	: Linear Liu Predictor
HRR	: Heteroscedastic Ridge Regression
GLE	: Generalized Liu Estimator
AR(1)	: First-Order Autoregressive
LB	: Linear Bayes
n.n.d.	: Non Negative Definite

1. INTRODUCTION

The prediction problem of linear models belongs to the most important problems of the statistical inference of linear models. Predicting future events based on the historical and present data, captured people's attention and specialists dramatically. Several techniques have been developed to address this problem and predict the future behavior of a particular series of events. Prediction is one of the main statistical tools used by researchers in order to find the best prediction model under specific conditions. Clearly, no method of prediction can be perfect, but in a fixed situation one method may be better than some alternatives.

The linear regression model plays an important role in prediction. Although the Ordinary Least Squares (OLS) estimator is the Best Linear Unbiased Estimator (BLUE) of regression parameters in a linear regression model, we can improve upon the variability of an estimator or a predictor of a regression coefficient when the unbiasedness criterion can be skipped. So, the optimal heterogeneous, homogenous and homogenous-unbiased estimators of the Mean Square Error (MSE) are introduced and discussed by many authors. Since the goodness of fitted model criterion is quite often ignored which is used to investigate the performance of estimators, Shalabh et al. (2009) proposed the Extended Balanced Loss Function (EBLF) in which the MSE, the Zellner's Balanced Loss Function (ZBLF) and the Predictive Loss Function (PLF) are just special cases of it. So, we derive the optimal heterogeneous, homogenous and homogenous unbiased estimators of the EBLF and discuss the performances of these estimators and their predictors in Chapter 2.

It is well-known that the predictive performance of a regression equation can be adversely affected by multicollinearity and correlated errors. Multicollinearity, linear or near linear dependency among the explanatory variables in the regression model is an important problem faced in application. The presence of multicollinearity has serious effects on the OLS estimates of the regression coefficients. If multicollinearity exists, in the method of least squares, the estimates of parameters will be poor generally, the

variance of the regression coefficients may be inflated, the absolute values of the estimates will be too large and they will be unstable. To overcome the problem of multicollinearity, alternative estimators are designed. The most popular method in the presence of multicollinearity is the Ordinary Ridge Regression (ORR) estimator which was proposed by Hoerl and Kennard (1970).

Therefore, when multicollinearity occurred under the linear regression model, Friedman and Montgomery (1985) focused on the prediction of a new response using different biased regression estimators by finding their Scalar Predicition Mean Square Error (pmse) and making comparisons among them as well as evaluating the predictive ability of the estimators at a particular observation. Then, Özbey and Kaçıranlar (2015) used a similar method to compare the Liu estimator (Liu, 1993) with other biased estimators according to the pmse criterion. So, it is reasonable to extend this idea by introducing two linear convex combination estimators of the operational optimal heterogeneous estimator of the EBLF to examine their predictive performances compared to OLS, ORR and Liu estimators in Chapter 3.

Autocorrelation in the errors in the regression model is an important problem faced in application. The presence of autocorrelation has serious effects on the OLS regression procedure. When autocorrelation is present in the errors, the OLS estimator will not be efficient and the usual estimator of the variance–covariance matrix will be biased. If variance–covariance matrix estimator is biased, then it may be misleading to use these variances for the confidence intervals and hypothesis tests (see Griffiths et al., 1993). To combat these effects, alternative methods were proposed. Aitken (1935) derived the Generalized Least Squares (GLS) estimator. Moreover, the problem of multicollinearity may also happens such that the GLS estimator gives unreliable estimates, because of the large total variance. As a result, Trenkler (1984) extended the idea of ridge regression to reduce the effects of autocorrelation and multicollinearity problems.

In the general linear model, the GLS estimator is the BLUE. However, in view of the interdependence of the disturbances, the pattern of sample residuals contains information which may be used to modify the regression function in order to reduce the Prediction Mean Square Error (PMSE). So, Goldberger (1962) derived the Best Linear Unbiased Predictor (BLUP) in the general linear regression model and found its pmse. However, when multicollinearity occurred among regressors, it is useful to use the biased regression estimators to find the biased predictors in order to reduce the pmse. So, it is reasonable to extend this idea to derive the biased predictors. Also, the predictive performance of various biased predictors with correlated errors is discussed and all pair-wise comparisons are made among these predictors in Chapter 4.

2. THE PERFORMANCE OF THE OPTIMAL EXTENDED BALANCED LOSS FUNCTION ESTIMATORS AND PREDICTORS

2.1. Introduction

Although the OLS estimator is the BLUE of regression parameters in a linear regression model, we can improve upon the variability of an estimator of a regression coefficient when the unbiasedness criterion can be skipped. With this manner, the minimum homogenous and the minimum heterogeneous mean squared error estimators are introduced and discussed by Theil (1958), Toutenburg (1968), Rao (1971), Theil (1971), Rao (1973) and Schaffrin (1985, 1986, 1987) among others. Then, the minimum homogenous and the minimum heterogeneous prediction mean squared error predictors related to the minimum homogenous and the minimum heterogeneous mean squared error estimators are introduced and examined by Bibby and Toutenburg (1977), Toutenburg (1968, 1970a, 1970b, 1970c), Goldberger (1962) and Toutenburg and Trenkler (1990) among others. Since the goodness of fitted model criterion is quite often ignored in which it is used to investigate the performance of estimators, Zellner (1994) introduced the ZBLF which concentrates of estimates around true parameter and goodness of fit of model. Further, Shalabh (1995) presented a PLF that not only incorporates the ZBLF as its particular case but also measures the correlation between the goodness of fit of model and concentration of estimates around the true parameter. The ZBLF has received considerable attention in the literature under different setups, see for example, Rodrigues and Zellner (1994), Giles et al. (1996), Ohtani et al. (1997), Ohtani (1998, 1999) and Gruber (2004); see also Toutenburg and Shalabh (2005) for the application of the ZBLF in some other areas of linear models. Appreciating the popularity of the ZBLF, Shalabh et al. (2009) extended it further and presented a general loss function called an EBLF to judge the performance of the OLS and the Stein-rule estimation.

As a consequence, it appears reasonable to extend the idea of finding optimal estimators according to the general loss function, EBLF. So, we present some fundamental concepts and loss functions. Then, we introduce and derive the optimal EBLF estimators and predictors and discuss their performances. Finally, we give a simulation study.

2.2. Linear Regression Model with Some Concepts and the Loss Functions

Let us consider the following linear model:

$$y = X\beta + \varepsilon, \tag{2.1.}$$

where y is an $n \times 1$ vector of responses, X is an $n \times p$ full column rank matrix of non-stochastic predetermined regressors, β is a $p \times 1$ vector of unknown parameters, and ε is an $n \times 1$ vector of $iid(0, \sigma^2)$ random errors.

The OLS estimator of β in model (2.1.) is

$$\hat{\beta} = (X'X)^{-1}X'y. \tag{2.2.}$$

This estimator is the most widely used technique for estimating the parameters of the linear regression models. The main reason for focusing on the OLS estimator is because it is unbiased and it has the minimum variance among all linear unbiased estimators under some set of assumptions.

The optimal heterogeneous, homogenous and homogenous unbiased estimators are found under the following criteria (Rao and Toutenburg, 1995),

$$R_1(\tilde{\beta}, \beta, G) = E(\tilde{\beta} - \beta)'G(\tilde{\beta} - \beta), \tag{2.3.}$$

where G is a pd $p \times p$-matrix.

$$R_2(\tilde{\beta}, \beta, g) = E\left[(\tilde{\beta} - \beta)' g\right]^2, \tag{2.4.}$$

where g is a fixed $p \times 1$-vector and $g \neq 0$.

$$R_3(\tilde{\beta}, \beta) = E(y - X\tilde{\beta})'(y - X\tilde{\beta}). \tag{2.5.}$$

ZBLF for any estimator $\tilde{\beta}$ of β is introduced by Zellner (1994) as follows:

$$ZBLF(\tilde{\beta}, \beta, t) = t\,(X\tilde{\beta} - y)'(X\tilde{\beta} - y) + (1-t)(\tilde{\beta} - \beta)' X'X (\tilde{\beta} - \beta), \tag{2.6.}$$

where t is a scalar lying between 0 and 1 which provides the weight assigned to the goodness of fit of model. The component $(X\tilde{\beta} - y)'(X\tilde{\beta} - y)$ reflects the goodness of fit of $\tilde{\beta}$ and the component $(\tilde{\beta} - \beta)' X'X (\tilde{\beta} - \beta)$ reflects the precision of estimation of $\tilde{\beta}$.

Shalabh (1995) defined a target function for the purpose of simultaneous prediction of actual and average values of y as

$$T(t) = t\,y + (1-t)\,E(y). \tag{2.7.}$$

The following PLF arises when we use the predictor $X\tilde{\beta}$ for simultaneous prediction of actual and average values of y through the target function (2.7.):

$$\begin{aligned} PLF(\tilde{\beta}, \beta, t) = {} & t^{\,2}(X\tilde{\beta} - y)'(X\tilde{\beta} - y) + (1-t)^{\,2}(\tilde{\beta} - \beta)' X'X (\tilde{\beta} - \beta) \\ & + 2t(1-t)(X\tilde{\beta} - y)' X (\tilde{\beta} - \beta). \end{aligned} \tag{2.8.}$$

Looking at the functional forms of the ZBLF and the PLF, Shalabh et al. (2009) proposed the EBLF as follows:

$$EBLF(\tilde{\beta},\beta,t_1,t_2) = t_1(X\tilde{\beta}-y)'(X\tilde{\beta}-y) + t_2(\tilde{\beta}-\beta)'X'X(\tilde{\beta}-\beta) + (1-t_1-t_2)(X\tilde{\beta}-y)'X(\tilde{\beta}-\beta), \quad (2.9.)$$

where t_1 and t_2 are the scalars lying between 0 and 1 characterizing the loss functions and $t_1+t_2 \le 1$. Obviously, the function (2.9.) encompasses the loss functions for goodness of fit, precision of estimation and the equations (2.3.), (2.4.), (2.5.), (2.6.) and (2.8.) as its particular cases. Then, from equations (2.1.) and (2.9.), Shalabh et al. (2009) observed that

$$EBLF(\tilde{\beta},\beta,t_1,t_2) = t_1\varepsilon'\varepsilon - (1+t_1-t_2)\varepsilon'X(\tilde{\beta}-\beta) + (\tilde{\beta}-\beta)'X'X(\tilde{\beta}-\beta), \quad (2.10.)$$

and the risk of any estimator $\tilde{\beta}$ is the expected value of the EBLF which is calculated by

$$R(\tilde{\beta},\beta,t_1,t_2) = E(EBLF(\tilde{\beta},\beta,t_1,t_2)). \quad (2.11.)$$

The main advantage of the EBLF is that it is more flexible in comparison to (2.6.) and (2.8.). The ZBLF in (2.6.) takes care only either the precision of estimation and the goodness of fit whereas (2.8.) extends it by considering the interaction or covariation between the precision of estimation and goodness of fit. Ignoring this covariation in the formulation of loss function may lead to wrong statistical inferences. The weights assigned in (2.8.) for precision of estimation, goodness of fit and their covariation depends only on one factor t assigned to one of the criterion only. The EBLF in (2.9.) provides a more flexibile option of choosing different weights for

precision of estimation, goodness of fit and covariation terms and, in the presence of covatiation, it will obviously lead to improved statistical inferences (Chaturvedi and Shalabh, 2014).

We use the following Lemma which was given by Rao (1973) and some definitions were given by Rao and Toutenburg (1995, pp. 165-166) in the Theorems and Corollaries.

Lemma 2.1: If $A: p \times p$, $B: p \times n$, $C: n \times n$ and $D: n \times p$ and all the inverses exist, we have

$$(A+BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}. \quad (2.12.)$$

Definition 2.1. The MSE:

The MSE of the estimator $\tilde{\beta}$ is defined as

$$MSE(\tilde{\beta}, \beta) = E(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)' = Cov(\tilde{\beta}) + Bias(\tilde{\beta})Bias(\tilde{\beta})', \quad (2.13.)$$

where $Cov(\tilde{\beta}) = E(\tilde{\beta} - E(\tilde{\beta}))(\tilde{\beta} - E(\tilde{\beta}))'$ and $Bias(\tilde{\beta}) = E(\tilde{\beta}) - \beta$,

and the Scalar MSE (mse) of $\tilde{\beta}$ is defined as

$$mse(\tilde{\beta}, \beta) = E(\tilde{\beta} - \beta)'(\tilde{\beta} - \beta) = tr\, Cov(\tilde{\beta}) + Bias(\tilde{\beta})'\, Bias(\tilde{\beta}). \quad (2.14.)$$

Definition 2.2. The PMSE:

The PMSE of the predictor p is defined as

$$PMSE(p, y_*) = E(p - y_*)(p - y_*)' = Cov(p) + Bias(p)Bias(p)', \quad (2.15.)$$

where $y_* = X_*\beta + \varepsilon_*, \varepsilon_* \sim (0, \sigma^2 \mathrm{I})$ is the linear regression model of future observations, $Cov(p) = E(p - E(p))(p - E(p))'$ and $Bias(p) = E(p) - X_*\beta.$

Definition 2.3. ($X_*\beta$- superiority). Consider two estimators $\tilde{\beta}_1$ and $\tilde{\beta}_2$. Then the predictor $X_*\tilde{\beta}_2$ of y_* is said to be $X_*\beta$-superior to $X_*\tilde{\beta}_1$ if

$$PMSE(X_*\tilde{\beta}_1, X_*\beta) - PMSE(X_*\tilde{\beta}_2, X_*\beta) \geq 0. \quad (2.16.)$$

Definition 2.4. (y_*- superiority). Consider two predictors p_1 and p_2 of y_*. The predictor p_2 is said to be y_*-superior to p_1 if

$$PMSE(p_1, y_*) - PMSE(p_2, y_*) \geq 0. \quad (2.17.)$$

***Note*:** when $PMSE(p_1, y_*) - PMSE(p_2, y_*) = PMSE(p_1, X_*\beta) - PMSE(p_2, X_*\beta)$, the $X_*\beta$- superiority implies y_*- superiority, and vice versa.

2.3. The Optimal Extended Balanced Loss Function Estimators and Predictors and Their Performances

Following Rao and Toutenburg (1995), we use the same forms of the heterogeneous, homogenous and homogenous unbiased estimators to derive these estimators for β under the EBLF criterion.

Theorem 2.1: Assume the linear regression model (2.1.) and the EBLF risk in (2.11.), then the optimal linear estimators for β are

(a) heterogeneous: $\tilde{\beta}_1 = c\hat{\beta} + (1-c)\beta$,

(b) homogenous: $\tilde{\beta}_2 = c\hat{\beta} + (1-c)r\beta$,

(c) homogenous unbiased: $\tilde{\beta}_3 = \hat{\beta}$,

where $c = \frac{(1+t_1-t_2)}{2}$, $0 \le c \le 1$ and $r = \frac{\beta' X' y}{\sigma^2 + \beta' X' X \beta}$.

Proof: (a) Firstly, we find the optimal, R-heterogeneous estimator $\tilde{\beta}_1$ as follows:

let

$$\tilde{\beta}_1 = Ay + a,$$

where the matrix $A(p \times n)$ and the vector $a(p \times 1)$ are non-stochastic and determined through optimization of the EBLF risk in (2.11.).

Then the EBLF in (2.10.) becomes

$$\begin{aligned} EBLF(\tilde{\beta}_1) = t_1 \varepsilon'\varepsilon - (1+t_1-t_2)\varepsilon' X([Ay+a]-\beta) \\ + ([Ay+a]-\beta)' X' X([Ay+a]-\beta). \end{aligned} \qquad (2.18.)$$

So, the risk of $\tilde{\beta}_1$ is

$$\begin{aligned} R(\tilde{\beta}_1) &= t_1 E(\varepsilon'\varepsilon) - (1+t_1-t_2)\, E(\varepsilon' X([Ay+a]-\beta)) \\ &\quad + E(([Ay+a]-\beta)' X' X([Ay+a]-\beta)) \\ &= t_1 n\sigma^2 - (1+t_1-t_2)\sigma^2 tr(XA) + \sigma^2 tr(A' X' XA) + \beta' X' A' X' XAX\beta \\ &\quad + a' X' XAX\beta + \beta' X' A' X' Xa + a' X' Xa - \beta' X' A' X' X\beta - a' X' X\beta \\ &\quad - \beta' X' XAX\beta - \beta' X' Xa + \beta' X' X\beta. \end{aligned} \tag{2.19.}$$

After differentiating $R(\tilde{\beta}_1)$ with respect to the matrix A, we obtain

$$\begin{aligned} \frac{\partial R(\tilde{\beta}_1)}{\partial A} &= -(1+t_1-t_2)\sigma^2 X' + 2\sigma^2 X' XA + 2X' XAX\beta\beta' X' \\ &\quad + 2X' Xa\beta' X' - 2X' X\beta\beta' X'. \end{aligned} \tag{2.20.}$$

Then, after differentiating $R(\tilde{\beta}_1)$ with respect to the vector a, we obtain

$$\frac{\partial R(\tilde{\beta}_1)}{\partial a} = 2X' XAX\beta + 2X' Xa - 2X' X\beta. \tag{2.21.}$$

After equating equations (2.20.) and (2.21.) to null matrix and zero vector, we get

$$\hat{A} = c(X' X)^{-1} X' \text{ and } \hat{a} = (1-c)\beta, \tag{2.22.}$$

where $c = \dfrac{(1+t_1-t_2)}{2}$ and $0 \le c \le 1$.

Therefore, the optimal R - heterogeneous estimator is as follows:

$$\tilde{\beta}_1 = c\hat{\beta} + (1-c)\beta. \tag{2.23.}$$

(b) Secondly, we find the optimal, R-homogenous estimator $\tilde{\beta}_2$ as follows:

let $\tilde{\beta}_2 = Ay$, then

$$EBLF(\tilde{\beta}_2) = t_1\varepsilon'\varepsilon - (1+t_1-t_2)\,\varepsilon' X(Ay-\beta) + (Ay-\beta)' X' X(Ay-\beta). \qquad (2.24.)$$

So, the risk of $\tilde{\beta}_2$ is

$$\begin{aligned} R(\tilde{\beta}_2) &= t_1 E(\varepsilon'\varepsilon) - (1+t_1-t_2)\,E(\varepsilon' X(Ay-\beta)) + E((Ay-\beta)' X' X(Ay-\beta)) \\ &= t_1 n\sigma^2 - (1+t_1-t_2)\sigma^2 tr(XA) + \sigma^2 tr(A' X' XA) + \beta' X' A' X' XAX\beta \\ &\quad - \beta' X' A' X' X\beta - \beta' X' XAX\beta + \beta' X' X\beta. \end{aligned} \qquad (2.25.)$$

After differentiating $R(\tilde{\beta}_2)$ with respect to the matrix A, we obtain

$$\frac{\partial R(\tilde{\beta}_2)}{\partial A} = -(1+t_1-t_2)\,\sigma^2 X' + 2\sigma^2 X' XA + 2X' XAX\beta\beta' X' - 2X' X\beta\beta' X'. \qquad (2.26.)$$

After equating equation (2.26.) to null matrix, we get

$$\hat{A} = c\sigma^2 (X' X)^{-1} X' (\sigma^2 \mathrm{I} + X\beta\beta' X')^{-1} + \beta\beta' X' (\sigma^2 \mathrm{I} + X\beta\beta' X')^{-1}. \qquad (2.27.)$$

Then,

$$\tilde{\beta}_2 = \hat{A}y = \left[c\sigma^2 (X' X)^{-1} X' (\sigma^2 \mathrm{I} + X\beta\beta' X')^{-1} + \beta\beta' X' (\sigma^2 \mathrm{I} + X\beta\beta' X')^{-1}\right] y. \qquad (2.28.)$$

Applying Lemma 2.1, we may simplify the expression for $\hat{A}$ by noting that

$$(\sigma^2 \mathrm{I} + X\beta\beta' X')^{-1} = \sigma^{-2}\mathrm{I} - \frac{\sigma^{-4} X\beta\beta' X'}{1+\sigma^{-2}\beta' X' X\beta}. \qquad (2.29.)$$

So, equation (2.28.) is

$$\tilde{\beta}_2 = \hat{A}y = c\sigma^2\left(\sigma^{-2}(X'X)^{-1}X'y - \frac{\sigma^{-4}\beta\beta' X'y}{1+\sigma^{-2}\beta' X' X\beta}\right) + \beta\left(\frac{\beta' X'y}{\sigma^2 + \beta' X' X\beta}\right). \qquad (2.30.)$$

Therefore, the optimal R - homogenous estimator is as follows:

$$\tilde{\beta}_2 = c\hat{\beta} + (1-c)r\beta, \qquad (2.31.)$$

where $r = \dfrac{\beta' X'y}{\sigma^2 + \beta' X' X\beta}$.

(c) Finally, we find the optimal unbiased, R-homogenous estimator $\tilde{\beta}_3$ as follows:

a homogenous linear estimator is unbiased if

$$AX - \mathrm{I} = 0 \qquad (2.32.)$$

or, equivalently, if

$$a_i' X - e_i' = 0 \quad (i = 1, 2, ..., p),$$

where e_i' and a_i' are the ith row vectors of I and A, respectively. Using (2.32.) in (2.25.), we find that

$$R(\tilde{\beta}_3) = t_1 n\sigma^2 - 2c\sigma^2 tr(XA) + \sigma^2 tr(A'X'XA) + \beta'X'X\beta\,. \tag{2.33.}$$

Therefore, the optimal A in this case is the solution obtained from

$$\min_A R(\tilde{\beta}_3) = t_1 n\sigma^2 - 2c\sigma^2 tr(XA) + \sigma^2 tr(A'X'XA) + \beta'X'X\beta - 2\sum_{i=1}^{p} \lambda_i'(a_i'X - e_i')', \tag{2.34.}$$

where $\lambda_1, \lambda_2, ..., \lambda_p$ are p-vectors of Lagrangian multipliers. Writing $\Lambda'_{p\times p} = (\lambda_1, \lambda_2, ..., \lambda_p)$, differentiating with respect to A and Λ, and equating to null matrices, we get

$$\frac{\partial R(\tilde{\beta}_3)}{\partial A} = -2c\sigma^2 X' + 2\sigma^2 X'XA - 2\Lambda X' = 0\,, \tag{2.35.}$$

$$\frac{\partial R(\tilde{\beta}_3)}{\partial \Lambda} = 2(AX - \mathrm{I}) = 0\,, \tag{2.36.}$$

which yields the optimal A as

$$\hat{A} = (X'X)^{-1}X'\,. \tag{2.37.}$$

The matrix $\hat{A}$ is consistent with the condition (2.32.):

$$\hat{A}X = (X'X)^{-1}X'X = \mathrm{I}\,. \tag{2.38.}$$

Therefore, the optimal unbiased, R-homogenous estimator is specified by

$$\tilde{\beta}_3 = \hat{\beta} = (X'X)^{-1}X'y. \tag{2.39.}$$

Corollary 2.1: $R(\tilde{\beta}_1) < R(\tilde{\beta}_2)$ iff $\beta'X'X\beta < \sigma^2$.

Proof: After substituting $\tilde{\beta}_1$ in (2.11.), we get

$$\begin{aligned} R(\tilde{\beta}_1) &= t_1E(\varepsilon'\varepsilon) - 2c^2E(\varepsilon'X(\hat{\beta}-\beta)) + c^2E[((\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta))] \\ &= t_1\sigma^2 n - 2c^2\sigma^2 tr(X(X'X)^{-1}X') + c^2\sigma^2 tr(X(X'X)^{-1}X') \\ &= t_1\sigma^2 n - c^2\sigma^2 p, \end{aligned} \tag{2.40.}$$

where $p = tr(X(X'X)^{-1}X')$.

After substituting $\tilde{\beta}_2$ in (2.11.), we get

$$R(\tilde{\beta}_2) = t_1\sigma^2 n - c^2\sigma^2 p + k(1-c)\left[(1-c)\beta'\beta - 2c\sigma^2 + \frac{\sigma^2}{1-k}(1+c-2k)\right], \tag{2.41.}$$

where $k = \dfrac{\beta'X'X\beta}{\sigma^2 + \beta'X'X\beta} < 1$.

Then,

$$R(\tilde{\beta}_2) - R(\tilde{\beta}_1) = k(1-c)\left[(1-c)\beta'\beta - 2c\sigma^2 + \frac{\sigma^2}{1-k}(1+c-2k)\right]. \tag{2.42.}$$

Therefore, $R(\tilde{\beta}_2) - R(\tilde{\beta}_1)$ is positive iff $\beta'X'X\beta < \sigma^2$.

Corollary 2.2: $\tilde{\beta}_3 = \hat{\beta}$ is a special case of the operational estimator of the estimator $\tilde{\beta}_1$ which does not depend on any unknown parameter.

Proof: Since the operational estimator of the estimator $\tilde{\beta}_1$ is $\hat{\tilde{\beta}}_1 = c\hat{\beta} + (1-c)\tilde{\beta}$, where $\tilde{\beta}$ is any estimator of β. Therefore, if $\tilde{\beta}$ is the OLS estimator, then $\hat{\tilde{\beta}}_1 = \hat{\beta} = \tilde{\beta}_3$.

Corollary 2.3: If $c = 1$, then $\tilde{\beta}_1$, $\tilde{\beta}_2$ and $\tilde{\beta}_3$ have the same risk. (i.e. $R = t_1\sigma^2 n - \sigma^2 p$).

Proof: If $c = 1$, then $\tilde{\beta}_1 = \tilde{\beta}_2 = \tilde{\beta}_3 = \hat{\beta}$. Therefore, $\tilde{\beta}_1$, $\tilde{\beta}_2$ and $\tilde{\beta}_3$ have the same risk. Also, substituting $c = 1$ in the equation (2.40.) or (2.41.), we get $R = t_1\sigma^2 n - \sigma^2 p$.

Corollary 2.4: If $c = 0$, then

- $\tilde{\beta}_2 = r\beta = \beta^*$, where β^* is the optimal mean squared error homogenous estimator of β.

- The operational estimator of $\tilde{\beta}_2$ is $\hat{\tilde{\beta}}_2 = \hat{r}\hat{\beta}$ (Dwivedi and Srivastava, 1978), where $\hat{r} = \dfrac{\hat{\beta}'X'y}{\hat{\sigma}^2 + \hat{\beta}'X'X\hat{\beta}}$.

Proof: The proof is evident by substituting $c = 0$ in Theorem 2.1.(b).

Theorem 2.2: The optimal heterogeneous predictor $p_1 = X_*\tilde{\beta}_1$ is always superior to the optimal homogenous predictor $p_2 = X_*\tilde{\beta}_2$ in terms of the PMSE criterion.

Proof: Since

$$PMSE(p_1, y_*) = X_* MSE(\tilde{\beta}_1, \beta) X_*' + \sigma^2 \mathrm{I} \tag{2.43.}$$

and

$$PMSE(p_2, y_*) = X_* MSE(\tilde{\beta}_2, \beta) X_*' + \sigma^2 \mathrm{I}, \tag{2.44.}$$

then

$$PMSE(p_2, y_*) - PMSE(p_1, y_*) = X_* \left[MSE(\tilde{\beta}_2, \beta) - MSE(\tilde{\beta}_1, \beta) \right] X_*' \tag{2.45.}$$

to be pd, the sufficient condition is $MSE(\tilde{\beta}_2, \beta) - MSE(\tilde{\beta}_1, \beta)$ to be pd.

So,

$$MSE(\tilde{\beta}_1) = Cov(\tilde{\beta}_1) = \sigma^2 c^2 (X'X)^{-1}, \tag{2.46.}$$

where $Bias(\tilde{\beta}_1) = 0$.

Also, after substituting $\tilde{\beta}_2$ in equation (2.13.), we get

$$MSE(\tilde{\beta}_2) = \sigma^2 c^2 (X'X)^{-1} + (1-k)(1-c^2)\beta\beta', \tag{2.47.}$$

where $Cov(\tilde{\beta}_2) = \sigma^2 c^2 (X'X)^{-1} + (1-k)(1-c)(2c+(1-c)k)\beta\beta'$ and

$Bias(\tilde{\beta}_2) = -(1-c)(1-k)\beta$.

So,

$$MSE(\tilde{\beta}_2, \beta) - MSE(\tilde{\beta}_1, \beta) = (1-k)(1-c^2)\beta\beta', \tag{2.48.}$$

where $(1-k)(1-c^2)\beta\beta'$ is always pd.

That means, $MSE(\tilde{\beta}_1, \beta) < MSE(\tilde{\beta}_2, \beta)$. Therefore, $PMSE(p_1, y_*) < PMSE(p_2, y_*)$.

Theorem 2.3: The optimal heterogeneous predictor $p_1 = X_*\tilde{\beta}_1$ is always superior to the optimal homogenous-unbiased predictor $p_3 = X_*\tilde{\beta}_3$ in terms of the PMSE criterion.

Proof: From Theorem 2.2, we have $PMSE(p_1, y_*) = X_* MSE(\tilde{\beta}_1, \beta)X_*' + \sigma^2 \mathrm{I}$.

Also, we obtain

$$PMSE(p_3, y_*) = X_* MSE(\tilde{\beta}_3, \beta)X_*' + \sigma^2 \mathrm{I}, \tag{2.49.}$$

then

$$PMSE(p_3, y_*) - PMSE(p_1, y_*) = X_* \left\lfloor MSE(\hat{\beta}_3, \beta) - MSE(\hat{\beta}_1, \beta) \right\rfloor X_*' \tag{2.50.}$$

to be pd, the sufficient condition is $MSE(\tilde{\beta}_3, \beta) - MSE(\tilde{\beta}_1, \beta)$ to be positive.

From Theorem 2.2, we have $MSE(\tilde{\beta}_1) = Cov(\tilde{\beta}_1) = \sigma^2 c^2 (X'X)^{-1}$.

Also, we obtain

$$MSE(\tilde{\beta}_3) = Cov(\tilde{\beta}_3) = \sigma^2 (X'X)^{-1}, \tag{2.51.}$$

where $Bias(\tilde{\beta}_3) = 0$.

So,

$$MSE(\tilde{\beta}_3, \beta) - MSE(\tilde{\beta}_1, \beta) = \sigma^2(1-c^2)(X'X)^{-1}, \qquad (2.52.)$$

where $\sigma^2(1-c^2)(X'X)^{-1}$ is always pd.

That means, $MSE(\tilde{\beta}_1, \beta) < MSE(\tilde{\beta}_3, \beta)$. Therefore, $PMSE(p_1, y_*) < PMSE(p_3, y_*)$.

Theorem 2.4: The optimal homogenous predictor $p_2 = X_*\tilde{\beta}_2$ is superior to the optimal homogenous-unbiased predictor $p_3 = X_*\tilde{\beta}_3$ when $\beta\beta' < \sigma^2(1-k)^{-1}(X'X)^{-1}$.

Proof: From Theorem 2.2, we have $PMSE(p_2, y_*) = X_* MSE(\tilde{\beta}_2, \beta)X_*' + \sigma^2 \mathrm{I}$.

Also, from Theorem 2.3, we have $PMSE(p_3, y_*) = X_* MSE(\tilde{\beta}_3, \beta)X_*' + \sigma^2 \mathrm{I}$.

Then, we obtain

$$PMSE(p_3, y_*) - PMSE(p_2, y_*) = X_* \left[MSE(\tilde{\beta}_3, \beta) - MSE(\tilde{\beta}_2, \beta)\right]X_*' \qquad (2.53.)$$

to be pd, the sufficient condition is $MSE(\tilde{\beta}_3, \beta) - MSE(\tilde{\beta}_2, \beta)$ to be pd.

From Theorem 2.2, we have $MSE(\tilde{\beta}_2) = \sigma^2 c^2 (X'X)^{-1} + (1-k)(1-c^2)\beta\beta'$.

Also, from Theorem 2.3, we have $MSE(\tilde{\beta}_3) = \sigma^2 (X'X)^{-1}$.

So,

$$MSE(\tilde{\beta}_3,\beta)-MSE(\tilde{\beta}_2,\beta)=(1-c^2)\left[\sigma^2(X'X)^{-1}-(1-k)\beta\beta'\right], \qquad (2.54.)$$

to be pd, the sufficient condition is $\sigma^2(X'X)^{-1}-(1-k)\beta\beta'>0$ where $(1-c^2)$ is always pd.

Therefore, the sufficient condition to be $\tilde{\beta}_2$ superior to $\tilde{\beta}_3$ and p_2 superior to p_3 is $\beta\beta'<\sigma^2(1-k)^{-1}(X'X)^{-1}$.

That means, $MSE(\tilde{\beta}_2,\beta)<MSE(\tilde{\beta}_3,\beta)$ and $PMSE(p_2,y_*)<PMSE(p_3,y_*)$ when $\beta\beta'<\sigma^2(1-k)^{-1}(X'X)^{-1}$.

To examine the performances of the heterogeneous, homogenous and homogenous unbiased predictors, it is sufficient to compare the performances of their estimators using the estimated *mse* criterion.

2.4. The Monte Carlo Simulation Study

In this section, we present a simulation study to compare the performances of the heterogeneous, homogenous and homogenous unbiased estimators. MATLAB is used for the simulation experiment. Following McDonald and Galarneau (1975) and Kibria (2003), the explanatory variables are generated by

$$x_{ij}=(1-\gamma^2)^{1/2}z_{ij}+\gamma z_{ip+1},\quad i=1,2,...,n\quad j=1,2,...,p,$$

where z_{ij} are independent standard normal pseudo-random numbers, γ is specified so that the correlation between any two explanatory variables is given by γ^2. Four different sets of correlation are considered, corresponding to $\gamma = 0.0, 0.1, 0.5, 0.9$. The explanatory variables are then standardized so that $X'X$ is in correlation form.

Following Kibria (2003), we choose β as the normalized eigenvector corresponding to the largest eigenvalue of the $X'X$ matrix. Observations on the dependent variable are determined by

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i, \quad i = 1, 2, ..., n$$

where ε_i are independent normal pseudo-random numbers with 0 mean and σ^2 variance and β_0 is taken to be identically zero. Three values of σ are considered which are 0.1, 1 and 10. Then the dependent variable is standardized so that $X'y$ is the vector of correlations of the dependent variable with each explanatory variable.

In this study, we choose $n = 20, 60, 100$, $c = 0.0, 0.1, 0.5, 0.9$ and $p = 4$. Then the experiment is replicated 5,000 times by generating new error terms.

We use the *mse* criterion to investigate the performance of the heterogeneous, homogenous and homogenous unbiased estimators. The estimated *mse* for any estimator $\hat{\beta}^*$ is calculated as follows (Güler and Kaçıranlar, 2009):

$$\hat{mse} = \frac{1}{MCN} \sum_{r=1}^{MCN} \left(\hat{\beta}_r^* - \beta\right)' \left(\hat{\beta}_r^* - \beta\right), \tag{2.55.}$$

where $\hat{\beta}_r^*$ is the computed value of β for the *r*-th replication of the experiment and *MCN* is the number of replications. The results of the simulation study are summarized in the Tables 2.1. – 2.3.. We have the following comments.

Firstly, we discuss the results for σ. As σ increases, the estimated mse of the mentioned estimators also increases.

Now, we investigate the effect of γ. For each value of σ, as γ increases, the estimated mse of all mentioned estimators is fluctuated except $\tilde{\beta}_1$, $\tilde{\beta}_2$ (when $c: 0.5 \to 0.9$) and $\tilde{\beta}_3$ for $n = 20$, the estimated mse increases.

Also, for each value of σ, γ and n; as c increases, the estimated mse of the mentioned estimators also increases.

In addition to that, we interpret the results for the sample size (n). In each given value of σ, γ and c; as the sample size increases, the estimated mse of the mentioned estimators gets smaller.

Finally, the inequality $\hat{mse}(\tilde{\beta}_1) < \hat{mse}(\tilde{\beta}_2) < \hat{mse}(\tilde{\beta}_3)$ always holds for each value of σ, γ, n and c.

2.5. Conclusion

We derive the optimal heterogeneous, homogenous and homogenous unbiased estimators of the EBLF criterion and we make comparisons between the optimal heterogeneous and the optimal homogenous estimators as well as other comparisons are introduced by giving some corollaries under some conditions in terms of the EBLF risk. In addition, we make all pair-wise comparisons among the optimal heterogeneous, homogenous and homogenous-unbiased predictors in terms of the PMSE criterion. A monte carlo simulation study is conducted to compare the resulting equations.

Table 2.1.: The mse estimates when $\sigma = 0.1$

		$\tilde{\beta}_1$				$\tilde{\beta}_2$				$\tilde{\beta}_3$
n	γ / c	0.0	0.1	0.5	0.9	0.0	0.1	0.5	0.9	--------
20	0.0	0.0000000	0.0000000	0.0004000	0.0014000	0.0003000	0.0003000	0.0006000	0.0014000	0.0017000
	0.1	0.0000000	0.0000000	0.0005000	0.0015000	0.0003000	0.0003000	0.0007000	0.0015000	0.0018000
	0.5	0.0000000	0.0000000	0.0006000	0.0018000	0.0003000	0.0004000	0.0008000	0.0019000	0.0022000
	0.9	0.0000000	0.0001000	0.0016000	0.0051000	0.0002000	0.0003000	0.0017000	0.0051000	0.0063000
60	0.0	0.0000000	0.0000073	0.0001828	0.0005922	0.0001265	0.0001326	0.0002776	0.0006162	0.0007311
	0.1	0.0000000	0.0000071	0.0001780	0.0005769	0.0001260	0.0001319	0.0002726	0.0006008	0.0007122
	0.5	0.0000000	0.0000078	0.0001940	0.0006285	0.0000872	0.0000940	0.0002594	0.0006451	0.0007760
	0.9	0.0000000	0.0000000	0.0007000	0.0022000	0.0000000	0.0001000	0.0007000	0.0023000	0.0028000
100	0.0	0.0000000	0.0000042	0.0001057	0.0003426	0.0000806	0.0000840	0.0001662	0.0003579	0.0004229
	0.1	0.0000000	0.0000042	0.0001044	0.0003384	0.0000785	0.0000819	0.0001633	0.0003533	0.0004177
	0.5	0.0000000	0.0000046	0.0001156	0.0003745	0.0000462	0.0000504	0.0001502	0.0003833	0.0004623
	0.9	0.0000000	0.0000000	0.0004000	0.0013000	0.0000000	0.0000000	0.0004000	0.0013000	0.0017000

Table 2.2.: The mse estimates when $\sigma = 1$

		$\tilde{\beta}_1$				$\tilde{\beta}_2$				$\tilde{\beta}_3$
n	γ / c	0.0	0.1	0.5	0.9	0.0	0.1	0.5	0.9	--------
20	0.0	0.0000000	0.0017000	0.0423000	0.1372000	0.0265000	0.0279000	0.0622000	0.1422000	0.1694000
	0.1	0.0000000	0.0018000	0.0453000	0.1469000	0.0267000	0.0282000	0.0653000	0.1520000	0.1813000
	0.5	0.0000000	0.0022000	0.0551000	0.1785000	0.0334000	0.0353000	0.0802000	0.1848000	0.2203000
	0.9	0.0000000	0.0063000	0.1574000	0.5099000	0.0225000	0.0285000	0.1742000	0.5142000	0.6295000
60	0.0	0.0000000	0.0007000	0.0183000	0.0592000	0.0125000	0.0131000	0.0277000	0.0616000	0.0731000
	0.1	0.0000000	0.0007000	0.0178000	0.0577000	0.0124000	0.0130000	0.0271000	0.0600000	0.0712000
	0.5	0.0000000	0.0008000	0.0194000	0.0629000	0.0087000	0.0093000	0.0259000	0.0645000	0.0776000
	0.9	0.0000000	0.0028000	0.0693000	0.2245000	0.0044000	0.0071000	0.0726000	0.2253000	0.2771000
100	0.0	0.0000000	0.0004000	0.0106000	0.0343000	0.0080000	0.0084000	0.0166000	0.0358000	0.0423000
	0.1	0.0000000	0.0004000	0.0104000	0.0338000	0.0078000	0.0082000	0.0163000	0.0353000	0.0418000
	0.5	0.0000000	0.0005000	0.0116000	0.0374000	0.0046000	0.0050000	0.0150000	0.0383000	0.0462000
	0.9	0.0000000	0.0017000	0.0414000	0.1341000	0.0022000	0.0039000	0.0431000	0.1345000	0.1655000

Table 2.3.: The mse estimates when $\sigma = 10$

		$\tilde{\beta}_1$				$\tilde{\beta}_2$				$\tilde{\beta}_3$
n	γ / c	0.0	0.1	0.5	0.9	0.0	0.1	0.5	0.9	--------
20	0.0	0.0000000	0.1694000	4.2338000	13.717500	0.7396000	0.8998000	4.7836000	13.856300	16.935200
	0.1	0.0000000	0.1813000	4.5337000	14.689300	0.7411000	0.9133000	5.0847000	14.828400	18.135000
	0.5	0.0000000	0.2203000	5.5080000	17.846100	0.7680000	0.9750000	6.0683000	17.986300	22.032200
	0.9	0.0000000	0.6295000	15.737300	50.988900	0.6935000	1.3102000	16.241000	51.114700	62.949200
60	0.0	0.0000000	0.0731000	1.8276000	5.9216000	0.5642000	0.6310000	2.2488000	6.0281000	7.3106000
	0.1	0.0000000	0.0712000	1.7804000	5.7686000	0.5522000	0.6169000	2.1918000	5.8725000	7.1217000
	0.5	0.0000000	0.0776000	1.9400000	6.2855000	0.4687000	0.5406000	2.2888000	6.3736000	7.7599000
	0.9	0.0000000	0.2771000	6.9278000	22.446000	0.3073000	0.5812000	7.1578000	22.504200	27.711100
100	0.0	0.0000000	0.0423000	1.0573000	3.4256000	0.4547000	0.4926000	1.3989000	3.5122000	4.2291000
	0.1	0.0000000	0.0418000	1.0443000	3.3835000	0.4506000	0.4878000	1.3822000	3.4691000	4.1772000
	0.5	0.0000000	0.0462000	1.1558000	3.7449000	0.3256000	0.3679000	1.3982000	3.8061000	4.6233000
	0.9	0.0000000	0.1655000	4.1378000	13.406500	0.1862000	0.3496000	4.2768000	13.441600	16.551200

3. PREDICTION OF THE LINEAR CONVEX COMBINATION OF BIASED REGRESSION ESTIMATORS

3.1. Introduction

One of the most popular estimator dealing with multicollinearity is the ORR estimator which is proposed by Hoerl and Kennard (1970) and is defined as

$$\hat{\beta}_k = (X'X + k\mathrm{I})^{-1} X'y, \qquad k \geq 0 \tag{3.1.}$$

where k is the biasing parameter. The purpose of adding the constant k to the diagonal elements of the matrix $X'X$ is the reduction of the sensitivity of the OLS estimates when $X'X$ is not close to the identity matrix.

Another alternative biased estimator dealing with multicollinearity is defined by Liu (1993) as

$$\begin{aligned} \hat{\beta}_d &= (X'X + \mathrm{I})^{-1}(X'y + d\hat{\beta}) \\ &= (X'X + \mathrm{I})^{-1}(X'X + d\mathrm{I})\hat{\beta}, \qquad 0 < d < 1 \end{aligned} \tag{3.2.}$$

where d is the biasing parameter. $\hat{\beta}_d = \hat{\beta}$ when $d = 1$. It is shown by Liu (1993) that this estimator is superior to the OLS estimator in terms of the MSE criterion and has an advantage over the ORR estimator because it is a linear function of d. So, the selection of biasing parameter is simpler. Later, this estimator is called the Liu estimator by Akdeniz and Kaçıranlar (1995). Sakallıoğlu et al. (2001) compared the Liu estimator to the ORR and the iterative estimators in terms of the MSE criterion.

Since we introduced the non-operational optimal heterogeneous estimator of the EBLF in equation (2.23.). So, we introduce the operational optimal heterogeneous estimator of the EBLF as follows:

$$\hat{\tilde{\beta}}_1 = c\hat{\beta} + (1-c)\tilde{\beta} \tag{3.3.}$$

where $0 \le c \le 1$ and $\tilde{\beta}$ is any regression estimator of β.

The Linear Bayes (LB) estimator is given by

$$\hat{\beta}_c = \theta + HX'(XHX' + \sigma^2 \mathrm{I})^{-1}(y - X\theta), \tag{3.4.}$$

where β is a random variable with $E(\beta) = \theta$, $Cov(\beta) = H$ and H is a Non Negative Definite (n.n.d.) matrix (see in details, Gruber (2010, 2012) and Rao (1973)).

The alternative form of $\hat{\beta}_c$ is given by Gruber (2012) as

$$\hat{\beta}_c = (X'X + \sigma^2 (UU'HUU')^{+})^{+}(X'y + (UU'HUU')^{+}\theta\sigma^2), \tag{3.5.}$$

where $UU' = \mathrm{I}$ and $(.)^{+}$ is the Moore Penrose inverse.

The form in (3.5.) is a generalization of the ridge estimator and called ridge-type estimator.

Similarly, the Liu-type estimator is given as (Gruber, 2012),

$$\hat{\beta}_{LT} = d\, b^{*} + (1-d)\,\hat{\beta}_{c}. \tag{3.6.}$$

where $b^{*} = (X'X)^{+}X'y$.

The Liu-type estimator is a convex combination of the OLS and the ridge-type with respect to the prior assumptions $E(\beta) = \theta$, $Cov(\beta) = F$ and F is a n.n.d. matrix. It is also the optimal estimator for the ZBLF. Using the alternative forms of the LB estimator were derived in Gruber (2010), the Liu-type estimator may be rewritten in the form

$$\begin{aligned}\hat{\beta}_{LT} = {} & (X'X + \sigma^{2}(UU'FUU')^{+})^{+}(X'y + \sigma^{2}(UU'FUU')^{+}\theta \\ & + d\sigma^{2}(UU'FUU')^{+}(\hat{\beta} - \theta)).\end{aligned} \tag{3.7.}$$

Therefore, for the full rank case if the LB estimator in equation (3.6.) is replaced with $\tilde{\beta}$ in equation (3.3.), then the ridge-type and the Liu-type estimators when $c \to 0$ and $c \to d$ respectively are just special cases of the operational optimal heterogeneous estimator of the EBLF.

A convex combination of two estimators can be useful when both estimators appear to be appropriate in a specific situation. So, as a special case, we replace the ORR estimator in equation (3.1.) and the Liu estimator in equation (3.2.) with $\tilde{\beta}$ in equation (3.3) respectively and obtain two linear convex combination estimators: the Combination of OLS and ORR (LOR) estimator and the Combination of OLS and Liu (LOL) estimator as follows:

$$\hat{\beta}_{LOR} = \hat{\beta}(k,c) = c\hat{\beta} + (1-c)\hat{\beta}_k, \tag{3.8.}$$

$$\hat{\beta}_{LOL} = \hat{\beta}(d,c) = c\hat{\beta} + (1-c)\hat{\beta}_d. \tag{3.9.}$$

We find that $\hat{\beta}_{LOR} = \hat{\beta}(k,c)$ in equation (3.8.) includes $\hat{\beta}$, $\hat{\beta}_k$, $\hat{\beta}_d$ and $\hat{\beta}(k,d)$ as special cases:

1. $\lim_{c\to 1} \hat{\beta}(k,c) = \lim_{\substack{k\to 0 \\ c\to d}} \hat{\beta}(k,c) = \hat{\beta}$
2. $\lim_{c\to 0} \hat{\beta}(k,c) = \hat{\beta}_k$
3. $\lim_{\substack{k\to 1 \\ c\to d}} \hat{\beta}(k,c) = \hat{\beta}_d$
4. $\lim_{c\to d} \hat{\beta}(k,c) = \hat{\beta}(k,d)$,

where $\hat{\beta}(k,d)$ is a two-parameter variation of the Liu estimator which is proposed by Özkale and Kaçıranlar (2007) and takes the form

$$\hat{\beta}(k,d) = (X'X + k\mathrm{I})^{-1}(X'y + kd\hat{\beta}) \quad k \geq 0,\ 0 < d < 1. \tag{3.10.}$$

For the full rank case in (3.7.), the two-parameter Liu estimator is obtained when $\theta = 0$ and $F = \frac{\sigma^2}{k} I$. Furthermore, Özkale (2013) has also noted that $\hat{\beta}(k,d)$ can also be demonstrated as

$$\hat{\beta}(k,d) = d\hat{\beta} + (1-d)\hat{\beta}_k. \tag{3.11.}$$

The two-parameter regression estimator is a convex combination estimator of the OLS and the ORR estimators. For the full rank case, $\hat{\beta}_c$ in (3.5.) simplifies to $\hat{\beta}_k$ when $\theta = 0$ and $H = \frac{\sigma^2}{k} I$. In this case, $\hat{\beta}_{LT}$ simplifies to $\hat{\beta}(k,d)$.

Although $\hat{\beta}(k,d)$ is called as the two-parameter estimator by Özkale and Kaçıranlar (2007), on account of the written form given in equation (3.11.), it is called the "affine combination type" estimator by Özkale (2013). Özkale and Kaçıranlar (2007) have also noted that $\hat{\beta}(k,d)$ can also be obtained by using the prior information on β in the form of linear stochastic restrictions. Then, the estimator $\hat{\beta}(k,d)$ is also derived from the mixed estimation method suggested by Theil (1963) and Theil and Goldberger (1961). $\hat{\beta}(k,d)$ is more stable than the OLS estimator and is a special case of the linear Bayes, mixed and minimax estimators (see Gruber (2010, 2012)).

Also, $\hat{\beta}_{LOL} = \hat{\beta}(d,c)$ in equation (3.9.) includes $\hat{\beta}$ and $\hat{\beta}_d$ as special cases:

1. $\lim_{c \to 1} \hat{\beta}(d,c) = \hat{\beta}$
2. $\lim_{c \to 0} \hat{\beta}(d,c) = \hat{\beta}_d$

$\hat{\beta}_{LOL}$ can also be obtained by using the following prior information. For the full rank case, $\hat{\beta}_c$ in (3.5.) simplifies to $\hat{\beta}_d$ when $\theta = d\hat{\beta}$ and $H = \sigma^2 I$. In this case, $\hat{\beta}_{LT}$ simplifies to $\hat{\beta}_{LOL}$.

As a consequence, it appears reasonable to evaluate the predictive performance of the proposed two linear convex combination estimators. In this chapter, the predictive performance of the LOR and the LOL estimators is discussed according to the pmse criterion. A numerical example is given to demonstrate the theoretical results.

3.2. Evaluations of the Prediction Mean Squared Errors

Following Friedman and Montgomery (1985) and Özbey and Kaçıranlar (2015) we obtain the pmse of the LOR and the LOL estimators.

The pmse is a measure of the closeness of a predictor to the response being predicted

$$pmse = E(y_0 - \hat{y}_0)^2 . \tag{3.12.}$$

Let J represents the pmse. J is the sum of the variance (V) and the squared bias (B) of the prediction:

$$J = V + B \tag{3.13.}$$

If y_0 is the value to be predicted, and $\hat{y}_0$ is the prediction of that value, then the variance and the bias of the prediction error are

$$V(y_0 - \hat{y}_0) = V(y_0) + V(\hat{y}_0) \tag{3.14.}$$

and

$$Bias = E(y_0 - \hat{y}_0) . \tag{3.15.}$$

For convenience, the orthogonal form of model (2.1.)

$$y = Z\alpha + \varepsilon \tag{3.16.}$$

will be used. Here $Z = XU$ and $\alpha = U'\beta$. Then the OLS estimator of α in model (3.16.) is

$$\hat{\alpha} = (Z'Z)^{-1}Z'y = \Lambda^{-1}Z'y, \tag{3.17.}$$

where $\Lambda = diag(\lambda_1, \lambda_2, ..., \lambda_p)$ is the matrix of the eigenvalues of $Z'Z$ and the lambdas are in descending order.

If z_0 is the orthonormalized point at which the prediction $\hat{y}_0$ is made, the pmse of the OLS estimator is

$$J_{OLS} = \sigma^2\left(1 + \sum_{i=1}^{p}\frac{z_{0i}^2}{\lambda_i}\right). \tag{3.18.}$$

Note that, since the OLS estimator is unbiased, its pmse is equal to its prediction variance

$$J_{OLS} = V_{OLS}. \tag{3.19.}$$

The ORR estimator of α in model (3.16.) is

$$\hat{\alpha}_k = (Z'Z + k\mathrm{I})^{-1}Z'y = (\Lambda + k\mathrm{I})^{-1}Z'y, \qquad k \geq 0 \tag{3.20.}$$

and the pmse of the ORR estimator is

$$J_k = \sigma^2\left(1 + \sum_{i=1}^{p}\frac{z_{0i}^2\,\lambda_i}{a_i^2}\right) + k^2\left(\sum_{i=1}^{p}\frac{z_{oi}\alpha_i}{a_i}\right)^2, \tag{3.21.}$$

where $a_i = \lambda_i + k$ (see also, Friedman and Montgomery (1985)).

The Liu estimator of α in model (3.16.) is

$$\begin{aligned}\hat{\alpha}_d &= (Z'Z + \mathrm{I})^{-1}(Z'y + d\,\hat{\alpha}) \\ &= (\Lambda + \mathrm{I})^{-1}(\Lambda + d\,\mathrm{I})\hat{\alpha}, \quad 0 < d < 1\end{aligned} \tag{3.22.}$$

and the pmse of the Liu estimator is

$$J_d = \sigma^2\left(1 + \sum_{i=1}^{p}\frac{z_{0i}^2\, c_i^2}{\lambda_i b_i^2}\right) + (1-d)^2\left(\sum_{i=1}^{p}\frac{z_{oi}\alpha_i}{b_i}\right)^2, \tag{3.23.}$$

where $b_i = \lambda_i + 1$ and $c_i = \lambda_i + d$ (see also, Özbey and Kaçıranlar (2015)).

The LOR estimator of α in model (3.16.) is

$$\begin{aligned}\hat{\alpha}_{LOR} &= \left[c(Z'Z)^{-1} + (1-c)(Z'Z + k\mathrm{I})^{-1}\right]Z'y \\ &= \left[c\Lambda^{-1} + (1-c)(\Lambda + k\mathrm{I})^{-1}\right]Z'y, \quad k \geq 0.\end{aligned} \tag{3.24.}$$

The variance of the prediction error of the LOR estimator is

$$\begin{aligned}V_{\mathrm{LOR}}(y_0 - \hat{y}_0) &= V(y_0) + V_{\mathrm{LOR}}(\hat{y}_0) \\ &= \sigma^2 + V(z_0'\,\hat{\alpha}_{\mathrm{LOR}}) \\ &= \sigma^2\left(1 + \sum_{i=1}^{p}\frac{\left[(1-c)\lambda_i + ca_i\right]^2 z_{0i}^2}{\lambda_i a_i^2}\right).\end{aligned} \tag{3.25.}$$

The bias of the prediction error of the LOR estimator is

$$\begin{aligned} Bias_{\mathrm{LOR}} &= E(y_0 - \hat{y}_0) = z_0'\alpha - z_0' E(\hat{\alpha}_{\mathrm{LOR}}) \\ &= k(1-c)\sum_{i=1}^{p} \frac{z_{oi}\alpha_i}{a_i}, \end{aligned} \tag{3.26.}$$

so, the squared bias is

$$B_{\mathrm{LOR}} = Bias^2_{\mathrm{LOR}} = k^2(1-c)^2\left(\sum_{i=1}^{p} \frac{z_{oi}\alpha_i}{a_i}\right)^2. \tag{3.27.}$$

By summing up the variance and the squared bias of the LOR estimator we obtain

$$\begin{aligned} J_{\mathrm{LOR}} &= V_{\mathrm{LOR}} + B_{\mathrm{LOR}} \\ &= \sigma^2\left(1 + \sum_{i=1}^{p} \frac{\left[(1-c)\lambda_i + ca_i\right]^2 z_{0i}^2}{\lambda_i a_i^2}\right) + k^2(1-c)^2\left(\sum_{i=1}^{p} \frac{z_{oi}\alpha_i}{a_i}\right)^2. \end{aligned} \tag{3.28.}$$

The LOL estimator of α in model (3.16.) is

$$\begin{aligned} \hat{\alpha}_{\mathrm{LOL}} &= \left[c\mathrm{I} + (1-c)(Z'Z + \mathrm{I})^{-1}(Z'Z + d\mathrm{I})\right]\hat{\alpha} \\ &= \left[c\mathrm{I} + (1-c)(\Lambda + \mathrm{I})^{-1}(\Lambda + d\mathrm{I})\right]\Lambda^{-1}Z'y, \quad 0 < d < 1. \end{aligned} \tag{3.29.}$$

The variance of the prediction error of the LOL estimator is

$$\begin{aligned} V_{\mathrm{LOL}}(y_0 - \hat{y}_0) &= V(y_0) + V_{\mathrm{LOL}}(\hat{y}_0) \\ &= \sigma^2 + V(z_0'\hat{\alpha}_{\mathrm{LOL}}) \\ &= \sigma^2\left(1 + \sum_{i=1}^{p} \frac{\left[cb_i + (1-c)c_i\right]^2 z_{0i}^2}{\lambda_i b_i^2}\right). \end{aligned} \tag{3.30.}$$

The bias of the prediction error of the LOL estimator is

$$\begin{aligned} Bias_{\mathrm{LOL}} &= E(y_0 - \hat{y}_0) = z_0'\alpha - z_0' E(\hat{\alpha}_{\mathrm{LOL}}) \\ &= (1-c)(1-d)\sum_{i=1}^{p} \frac{z_{oi}\alpha_i}{b_i}, \end{aligned} \tag{3.31.}$$

so, the squared bias is

$$B_{\mathrm{LOL}} = Bias^2_{\mathrm{LOL}} = (1-c)^2(1-d)^2\left(\sum_{i=1}^{p}\frac{z_{oi}\alpha_i}{b_i}\right)^2. \tag{3.32.}$$

By summing up the variance and the squared bias of the LOL estimator we obtain

$$\begin{aligned} J_{\mathrm{LOL}} &= V_{\mathrm{LOL}} + B_{\mathrm{LOL}} \\ &= \sigma^2\left(1+\sum_{i=1}^{p}\frac{\left[cb_i+(1-c)c_i\right]^2 z_{0i}^2}{\lambda_i b_i^2}\right) + (1-c)^2(1-d)^2\left(\sum_{i=1}^{p}\frac{z_{oi}\alpha_i}{b_i}\right)^2. \end{aligned} \tag{3.33.}$$

3.3. Comparisons of Prediction Mean Squared Errors in Two Dimensional Space

Friedman and Montgomery (1985) used the pmse approach for evaluating the predictive ability of different biased estimators in the linear regression model. By using the two-dimensional space, they got the opportunity to obtain the ratio z_{02}^2 / z_{01}^2, and use it as the reference point in their comparisons. Özbey and Kaçıranlar (2015) also used the same method for the Liu estimator. We use the same method here for the LOR and the LOL estimators. So, our inferences is based on the subspace of the observation to be predicted (i.e., the ratio z_{02}^2 / z_{01}^2). As in Friedman and Montgomery (1985) and Özbey and Kaçıranlar (2015) α_1^2 is set to zero because, non-zero value of α_1^2 increase only the intercept values for the J_{LOR} and the J_{LOL} but leave the curve for the J_{OLS} unchanged.

So, comparisons of J_{LOR} with J_{OLS} and J_k and J_{LOL} with J_{OLS}, J_d and J_{LOR} are made and stated in the following five theorems.

Theorem 3.1:

1. If $\alpha_2^2 > \dfrac{\sigma^2\left(a_2^2 - ((1-c)\lambda_2 + ca_2)^2\right)}{\lambda_2 k^2 (1-c)^2}$, then

- $J_{LOR} < J_{OLS}$ for $a_1^2 < ((1-c)\lambda_1 + ca_1)^2$,

- $J_{LOR} < J_{OLS} \Leftrightarrow \dfrac{z_{02}^2}{z_{01}^2} < f_1(\alpha_2^2)$ for $a_1^2 > ((1-c)\lambda_1 + ca_1)^2$.

2. If $\alpha_2^2 < \dfrac{\sigma^2\left(a_2^2 - ((1-c)\lambda_2 + ca_2)^2\right)}{\lambda_2 k^2 (1-c)^2}$, then

- $J_{LOR} < J_{OLS}$ for $a_1^2 > ((1-c)\lambda_1 + ca_1)^2$,

- $J_{LOR} < J_{OLS} \Leftrightarrow \dfrac{z_{02}^2}{z_{01}^2} < f_1(\alpha_2^2)$ for $a_1^2 < ((1-c)\lambda_1 + ca_1)^2$.

Where

$$f_1(\alpha_2^2) = \frac{\sigma^2\left(\dfrac{1}{\lambda_1} - \dfrac{((1-c)\lambda_1 + ca_1)^2}{\lambda_1 a_1^2}\right)}{\left(\dfrac{\sigma^2((1-c)\lambda_2 + ca_2)^2}{\lambda_2 a_2^2} + \dfrac{k^2(1-c)^2\alpha_2^2}{a_2^2} - \dfrac{\sigma^2}{\lambda_2}\right)}. \qquad (3.34.)$$

Proof: If the LOR estimator is superior to the OLS estimator in terms of the pmse criterion, we have $J_{LOR} < J_{OLS}$. That is,

$$\sigma^2+\sigma^2\left[\frac{((1-c)\lambda_1+ca_1)^2 z_{01}^2}{\lambda_1 a_1^2}+\frac{((1-c)\lambda_2+ca_2)^2 z_{02}^2}{\lambda_2 a_2^2}\right] + \frac{k^2(1-c)^2\alpha_2^2 z_{02}^2}{a_2^2}<\sigma^2+\sigma^2\left(\frac{z_{01}^2}{\lambda_1}+\frac{z_{02}^2}{\lambda_2}\right).$$

Rearranging this inequality, we obtain

$$z_{02}^2\left(\frac{\sigma^2((1-c)\lambda_2+ca_2)^2}{\lambda_2 a_2^2}+\frac{k^2(1-c)^2\alpha_2^2}{a_2^2}-\frac{\sigma^2}{\lambda_2}\right)< z_{01}^2\sigma^2\left(\frac{1}{\lambda_1}-\frac{((1-c)\lambda_1+ca_1)^2}{\lambda_1 a_1^2}\right).$$

If both

$$\frac{\sigma^2((1-c)\lambda_2+ca_2)^2}{\lambda_2 a_2^2}+\frac{k^2(1-c)^2\alpha_2^2}{a_2^2}-\frac{\sigma^2}{\lambda_2} \tag{3.35.}$$

and

$$\sigma^2\left(\frac{1}{\lambda_1}-\frac{((1-c)\lambda_1+ca_1)^2}{\lambda_1 a_1^2}\right) \tag{3.36.}$$

have the same signs, the condition for the superiority of the LOR estimator over the OLS estimator is

$$\frac{z_{02}^2}{z_{01}^2}<f_1(\alpha_2^2). \tag{3.37.}$$

If (3.35.) and (3.36.) have opposite signs, the condition for the superiority of the LOR estimator over the OLS estimator is

$$\frac{z_{02}^2}{z_{01}^2} > f_1(\alpha_2^2). \tag{3.38.}$$

It is obvious that if (3.35.) and (3.36.) have opposite signs, the right hand side of (3.38.) is negative, thus (3.38.) always holds. Consequently, at that region the LOR estimator is uniformly superior to the OLS estimator. The condition for the positiveness of (3.35.) can be written as

$$\alpha_2^2 > \frac{\sigma^2\left(a_2^2 - ((1-c)\lambda_2 + ca_2)^2\right)}{\lambda_2 k^2 (1-c)^2} \tag{3.39.}$$

and the condition for the positiveness of (3.36.) can be written as

$$a_1^2 > ((1-c)\lambda_1 + ca_1)^2. \tag{3.40.}$$

Of course, the opposite conditions are needed for the negativeness of (3.35.) and (3.36.). The vertical asymptote of the hyperbola $f_1(\alpha_2^2)$ is at the point

$$\alpha_2^2 = \frac{\sigma^2\left(a_2^2 - ((1-c)\lambda_2 + ca_2)^2\right)}{\lambda_2 k^2 (1-c)^2}. \tag{3.41.}$$

Corollary 3.1: If $c = 0$ in Theorem 3.1, we get the Friedman and Montgomery's (1985) results.

Corollary 3.2: If $c = d$ and $k = 1$ in Theorem 3.1, we get the Özbey and Kaçıranlar's (2015) results.

Theorem 3.2:

1. If $\alpha_2^2 > \dfrac{\sigma^2\left(\lambda_2^2 - ((1-c)\lambda_2 + ca_2)^2\right)}{\lambda_2 k^2[(1-c)^2 - 1]}$, then

- $J_{LOR} < J_k$ for $\lambda_1^2 < ((1-c)\lambda_1 + ca_1)^2$,

- $J_{LOR} < J_k \Leftrightarrow \dfrac{z_{02}^2}{z_{01}^2} < f_2(\alpha_2^2)$ for $\lambda_1^2 > ((1-c)\lambda_1 + ca_1)^2$.

2. If $\alpha_2^2 < \dfrac{\sigma^2\left(\lambda_2^2 - ((1-c)\lambda_2 + ca_2)^2\right)}{\lambda_2 k^2[(1-c)^2 - 1]}$, then

- $J_{LOR} < J_k$ for $\lambda_1^2 > ((1-c)\lambda_1 + ca_1)^2$,

- $J_{LOR} < J_k \Leftrightarrow \dfrac{z_{02}^2}{z_{01}^2} < f_2(\alpha_2^2)$ for $\lambda_1^2 < ((1-c)\lambda_1 + ca_1)^2$.

Where

$$f_2(\alpha_2^2) = \frac{\sigma^2\left(\dfrac{\lambda_1}{a_1^2} - \dfrac{((1-c)\lambda_1 + ca_1)^2}{\lambda_1 a_1^2}\right)}{\left(\dfrac{\sigma^2((1-c)\lambda_2 + ca_2)^2}{\lambda_2 a_2^2} + \dfrac{k^2(1-c)^2\alpha_2^2}{a_2^2} - \dfrac{\sigma^2\lambda_2}{a_2^2} - \dfrac{k^2\alpha_2^2}{a_2^2}\right)}. \qquad (3.42.)$$

Proof: If the LOR estimator is superior to the ORR estimator in terms of the pmse criterion, we have $J_{LOR} < J_k$. That is,

$$\sigma^2 + \sigma^2\left[\frac{((1-c)\lambda_1 + ca_1)^2 z_{01}^2}{\lambda_1 a_1^2} + \frac{((1-c)\lambda_2 + ca_2)^2 z_{02}^2}{\lambda_2 a_2^2}\right] + \frac{k^2(1-c)^2\alpha_2^2 z_{02}^2}{a_2^2} <$$
$$\sigma^2 + \sigma^2\left(\frac{\lambda_1 z_{01}^2}{a_1^2} + \frac{\lambda_2 z_{02}^2}{a_2^2}\right) + \frac{k^2\alpha_2^2 z_{02}^2}{a_2^2}.$$

Rearranging this inequality, we obtain

$$z_{02}^2\left(\frac{\sigma^2((1-c)\lambda_2+ca_2)^2}{\lambda_2 a_2^2}+\frac{k^2(1-c)^2\alpha_2^2}{a_2^2}-\frac{\sigma^2\lambda_2}{a_2^2}-\frac{k^2\alpha_2^2}{a_2^2}\right)< z_{01}^2\sigma^2\left(\frac{\lambda_1}{a_1^2}-\frac{((1-c)\lambda_1+ca_1)^2}{\lambda_1 a_1^2}\right).$$

If both

$$\left(\frac{\sigma^2((1-c)\lambda_2+ca_2)^2}{\lambda_2 a_2^2}+\frac{k^2(1-c)^2\alpha_2^2}{a_2^2}-\frac{\sigma^2\lambda_2}{a_2^2}-\frac{k^2\alpha_2^2}{a_2^2}\right) \tag{3.43.}$$

and

$$\sigma^2\left(\frac{\lambda_1}{a_1^2}-\frac{((1-c)\lambda_1+ca_1)^2}{\lambda_1 a_1^2}\right) \tag{3.44.}$$

have the same signs, the condition for the superiority of the LOR estimator over the ORR estimator is

$$\frac{z_{02}^2}{z_{01}^2}< f_2(\alpha_2^2). \tag{3.45.}$$

If (3.43.) and (3.44.) have opposite signs, the condition for the superiority of the LOR estimator over the ORR estimator is

$$\frac{z_{02}^2}{z_{01}^2}> f_2(\alpha_2^2). \tag{3.46.}$$

It is obvious that if (3.43.) and (3.44.) have opposite signs, the right hand side of (3.46.) is negative, thus (3.46.) always holds. Consequently, at that region the LOR estimator is uniformly superior to the ORR estimator. The condition for the positiveness of (3.43.) can be written as

$$\alpha_2^2 > \frac{\sigma^2\left(\lambda_2^2 - ((1-c)\lambda_2 + ca_2)^2\right)}{\lambda_2 k^2[(1-c)^2 - 1]} \tag{3.47.}$$

and the condition for the positiveness of (3.44.) can be written as

$$\lambda_1^2 > ((1-c)\lambda_1 + ca_1)^2. \tag{3.48.}$$

Of course, the opposite conditions are needed for the negativeness of (3.43.) and (3.44.). The vertical asymptote of the hyperbola $f_2(\alpha_2^2)$ is at the point

$$\alpha_2^2 = \frac{\sigma^2\left(\lambda_2^2 - ((1-c)\lambda_2 + ca_2)^2\right)}{\lambda_2 k^2[(1-c)^2 - 1]}. \tag{3.49.}$$

Theorem 3.3:

1. If $\alpha_2^2 > \frac{\sigma^2\left(b_2^2 - (cb_2 + (1-c)c_2)^2\right)}{\lambda_2(1-c)^2(1-d)^2}$, then

- $J_{LOL} < J_{OLS}$ for $b_1^2 < (cb_1 + (1-c)c_1)^2$,

- $J_{LOL} < J_{OLS} \Leftrightarrow \frac{z_{02}^2}{z_{01}^2} < f_3(\alpha_2^2)$ for $b_1^2 > (cb_1 + (1-c)c_1)^2$.

2. If $\alpha_2^2 < \dfrac{\sigma^2\left(b_2^2-(cb_2+(1-c)c_2)^2\right)}{\lambda_2(1-c)^2(1-d)^2}$, then

- $J_{LOL} < J_{OLS}$ for $b_1^2 > (cb_1+(1-c)c_1)^2$,

- $J_{LOL} < J_{OLS} \Leftrightarrow \dfrac{z_{02}^2}{z_{01}^2} < f_3(\alpha_2^2)$ for $b_1^2 < (cb_1+(1-c)c_1)^2$.

Where

$$f_3(\alpha_2^2) = \frac{\sigma^2\left(\dfrac{1}{\lambda_1}-\dfrac{(cb_1+(1-c)c_1)^2}{\lambda_1 b_1^2}\right)}{\left(\dfrac{\sigma^2(cb_2+(1-c)c_2)^2}{\lambda_2 b_2^2}+\dfrac{(1-c)^2(1-d)^2\alpha_2^2}{b_2^2}-\dfrac{\sigma^2}{\lambda_2}\right)}. \qquad (3.50.)$$

Proof: If the LOL estimator is superior to the OLS estimator in terms of the pmse criterion, we have $J_{LOL} < J_{OLS}$. That is,

$$\sigma^2+\sigma^2\left[\frac{(cb_1+(1-c)c_1)^2 z_{01}^2}{\lambda_1 b_1^2}+\frac{(cb_2+(1-c)c_2)^2 z_{02}^2}{\lambda_2 b_2^2}\right]+\frac{(1-c)^2(1-d)^2\alpha_2^2 z_{02}^2}{b_2^2} < \sigma^2+\sigma^2\left(\frac{z_{01}^2}{\lambda_1}+\frac{z_{02}^2}{\lambda_2}\right).$$

Rearranging this inequality, we obtain

$$z_{02}^2\left(\frac{\sigma^2(cb_2+(1-c)c_2)^2}{\lambda_2 b_2^2}+\frac{(1-c)^2(1-d)^2\alpha_2^2}{b_2^2}-\frac{\sigma^2}{\lambda_2}\right) < z_{01}^2\sigma^2\left(\frac{1}{\lambda_1}-\frac{(cb_1+(1-c)c_1)^2}{\lambda_1 b_1^2}\right).$$

If both

$$\frac{\sigma^2(cb_2+(1-c)c_2)^2}{\lambda_2 b_2^2}+\frac{(1-c)^2(1-d)^2\alpha_2^2}{b_2^2}-\frac{\sigma^2}{\lambda_2} \tag{3.51.}$$

and

$$\sigma^2\left(\frac{1}{\lambda_1}-\frac{(cb_1+(1-c)c_1)^2}{\lambda_1 b_1^2}\right) \tag{3.52.}$$

have the same signs, the condition for the superiority of the LOL estimator over the OLS estimator is

$$\frac{z_{02}^2}{z_{01}^2}<f_3(\alpha_2^2)\,. \tag{3.53.}$$

If (3.51.) and (3.52.) have opposite signs, the condition for the superiority of the LOL estimator over the OLS estimator is

$$\frac{z_{02}^2}{z_{01}^2}>f_3(\alpha_2^2)\,. \tag{3.54.}$$

It is obvious that if (3.51.) and (3.52.) have opposite signs, the right hand side of (3.54.) is negative, thus (3.54.) always holds. Consequently, at that region the LOL estimator is uniformly superior to the OLS estimator. The condition for the positiveness of (3.51.) can be written as

$$\alpha_2^2>\frac{\sigma^2\left(b_2^2-(cb_2+(1-c)c_2)^2\right)}{\lambda_2(1-c)^2(1-d)^2} \tag{3.55.}$$

and the condition for the positiveness of (3.52.) can be written as

$$b_1^2 > (cb_1 + (1-c)c_1)^2 . \tag{3.56.}$$

Of course, the opposite conditions are needed for the negativeness of (3.51.) and (3.52.). The vertical asymptote of the hyperbola $f_3(\alpha_2^2)$ is at the point

$$\alpha_2^2 = \frac{\sigma^2\left(b_2^2 - (cb_2 + (1-c)c_2)^2\right)}{\lambda_2(1-c)^2(1-d)^2}. \tag{3.57.}$$

Corollary 3.3: If $c = 0$ in Theorem 3.3, we get the Özbey and Kaçıranlar's (2015) results.

Theorem 3.4:

1. If $\alpha_2^2 > \frac{\sigma^2\left(c_2^2 - (cb_2 + (1-c)c_2)^2\right)}{\lambda_2(1-d)^2[(1-c)^2 - 1]}$, then

- $J_{LOL} < J_d$ for $c_1^2 < (cb_1 + (1-c)c_1)^2$,

- $J_{LOL} < J_d \Leftrightarrow \frac{z_{02}^2}{z_{01}^2} < f_4(\alpha_2^2)$ for $c_1^2 > (cb_1 + (1-c)c_1)^2$.

2. If $\alpha_2^2 < \frac{\sigma^2\left(c_2^2 - (cb_2 + (1-c)c_2)^2\right)}{\lambda_2(1-d)^2[(1-c)^2 - 1]}$, then

- $J_{LOL} < J_d$ for $c_1^2 > (cb_1 + (1-c)c_1)^2$,

- $J_{LOL} < J_d \Leftrightarrow \frac{z_{02}^2}{z_{01}^2} < f_4(\alpha_2^2)$ for $c_1^2 < (cb_1 + (1-c)c_1)^2$.

Where

$$f_4(\alpha_2^2)=\frac{\sigma^2\left(\frac{c_1^2}{\lambda_1 b_1^2}-\frac{(cb_1+(1-c)c_1)^2}{\lambda_1 b_1^2}\right)}{\left(\frac{\sigma^2(cb_2+(1-c)c_2)^2}{\lambda_2 b_2^2}+\frac{(1-c)^2(1-d)^2\alpha_2^2}{b_2^2}-\frac{\sigma^2 c_2^2}{\lambda_2 b_2^2}-\frac{(1-d)^2\alpha_2^2}{b_2^2}\right)}. \quad (3.58.)$$

Proof: If the LOL estimator is superior to the Liu estimator in terms of the pmse criterion, we have $J_{LOL} < J_d$. That is,

$$\sigma^2+\sigma^2\left[\frac{(cb_1+(1-c)c_1)^2 z_{01}^2}{\lambda_1 b_1^2}+\frac{(cb_2+(1-c)c_2)^2 z_{02}^2}{\lambda_2 b_2^2}\right]+\frac{(1-c)^2(1-d)^2\alpha_2^2 z_{02}^2}{b_2^2}<$$
$$\sigma^2+\sigma^2\left(\frac{c_1^2 z_{01}^2}{\lambda_1 b_1^2}+\frac{c_2^2 z_{02}^2}{\lambda_2 b_2^2}\right)+\frac{(1-d)^2\alpha_2^2 z_{02}^2}{b_2^2}.$$

Rearranging this inequality, we obtain

$$z_{02}^2\left(\frac{\sigma^2(cb_2+(1-c)c_2)^2}{\lambda_2 b_2^2}+\frac{(1-c)^2(1-d)^2\alpha_2^2}{b_2^2}-\frac{\sigma^2 c_2^2}{\lambda_2 b_2^2}-\frac{(1-d)^2\alpha_2^2}{b_2^2}\right)<$$
$$z_{01}^2\sigma^2\left(\frac{c_1^2}{\lambda_1 b_1^2}-\frac{(cb_1+(1-c)c_1)^2}{\lambda_1 b_1^2}\right).$$

If both

$$\left(\frac{\sigma^2(cb_2+(1-c)c_2)^2}{\lambda_2 b_2^2}+\frac{(1-c)^2(1-d)^2\alpha_2^2}{b_2^2}-\frac{\sigma^2 c_2^2}{\lambda_2 b_2^2}-\frac{(1-d)^2\alpha_2^2}{b_2^2}\right) \quad (3.59.)$$

and

$$\sigma^2\left(\frac{c_1^2}{\lambda_1 b_1^2}-\frac{(cb_1+(1-c)c_1)^2}{\lambda_1 b_1^2}\right) \tag{3.60.}$$

have the same signs, the condition for the superiority of the LOL estimator over the Liu estimator is

$$\frac{z_{02}^2}{z_{01}^2} < f_4(\alpha_2^2). \tag{3.61.}$$

If (3.59.) and (3.60.) have opposite signs, the condition for the superiority of the LOL estimator over the Liu estimator is

$$\frac{z_{02}^2}{z_{01}^2} > f_4(\alpha_2^2). \tag{3.62.}$$

It is obvious that if (3.59.) and (3.60.) have opposite signs, the right hand side of (3.62.) is negative, thus (3.62.) always holds. Consequently, at that region the LOL estimator is uniformly superior to the Liu estimator. The condition for the positiveness of (3.59.) can be written as

$$\alpha_2^2 > \frac{\sigma^2\left(c_2^2-(cb_2+(1-c)c_2)^2\right)}{\lambda_2(1-d)^2[(1-c)^2-1]} \tag{3.63.}$$

and the condition for the positiveness of (3.60.) can be written as

$$c_1^2 > (cb_1+(1-c)c_1)^2. \tag{3.64.}$$

Of course, the opposite conditions are needed for the negativeness of (3.59.) and (3.60.). The vertical asymptote of the hyperbola $f_4(\alpha_2^2)$ is at the point

$$\alpha_2^2 = \frac{\sigma^2\left(c_2^2 - (cb_2 + (1-c)c_2)^2\right)}{\lambda_2(1-d)^2[(1-c)^2 - 1]}. \tag{3.65}$$

Theorem 3.5:

1. If $\alpha_2^2 > \frac{\sigma^2\left(((1-c)\lambda_2 + ca_2)^2 b_2^2 - (cb_2 + (1-c)c_2)^2 a_2^2\right)}{\lambda_2(1-c)^2[(1-d)^2 a_2^2 - k^2 b_2^2]}$, then

- $J_{LOL} < J_{LOR}$ for $((1-c)\lambda_1 + ca_1)^2 b_1^2 < (cb_1 + (1-c)c_1)^2 a_1^2$,

- $J_{LOL} < J_{LOR} \Leftrightarrow \frac{z_{02}^2}{z_{01}^2} < f_5(\alpha_2^2)$ for $((1-c)\lambda_1 + ca_1)^2 b_1^2 > (cb_1 + (1-c)c_1)^2 a_1^2$.

2. If $\alpha_2^2 < \frac{\sigma^2\left(((1-c)\lambda_2 + ca_2)^2 b_2^2 - (cb_2 + (1-c)c_2)^2 a_2^2\right)}{\lambda_2(1-c)^2[(1-d)^2 a_2^2 - k^2 b_2^2]}$, then

- $J_{LOL} < J_{LOR}$ for $((1-c)\lambda_1 + ca_1)^2 b_1^2 > (cb_1 + (1-c)c_1)^2 a_1^2$,

- $J_{LOL} < J_{LOR} \Leftrightarrow \frac{z_{02}^2}{z_{01}^2} < f_5(\alpha_2^2)$ for $((1-c)\lambda_1 + ca_1)^2 b_1^2 < (cb_1 + (1-c)c_1)^2 a_1^2$.

Where

$$f_5(\alpha_2^2) = \frac{\sigma^2\left(\frac{((1-c)\lambda_1 + ca_1)^2}{\lambda_1 a_1^2} - \frac{(cb_1 + (1-c)c_1)^2}{\lambda_1 b_1^2}\right)}{\left(\frac{\sigma^2(cb_2 + (1-c)c_2)^2}{\lambda_2 b_2^2} + \frac{(1-c)^2(1-d)^2\alpha_2^2}{b_2^2} - \frac{\sigma^2((1-c)\lambda_2 + ca_2)^2}{\lambda_2 a_2^2}\right) + \left(-\frac{k^2(1-c)^2\alpha_2^2}{a_2^2}\right)}. \tag{3.66.}$$

Proof: If the LOL estimator is superior to the LOR estimator in terms of the pmse criterion, we have $J_{LOL} < J_{LOR}$. That is,

$$\sigma^2 + \sigma^2\left[\frac{(cb_1+(1-c)c_1)^2 z_{01}^2}{\lambda_1 b_1^2} + \frac{(cb_2+(1-c)c_2)^2 z_{02}^2}{\lambda_2 b_2^2}\right] + \frac{(1-c)^2(1-d)^2\alpha_2^2 z_{02}^2}{b_2^2} <$$
$$\sigma^2 + \sigma^2\left(\frac{((1-c)\lambda_1 + ca_1)^2 z_{01}^2}{\lambda_1 a_1^2} + \frac{((1-c)\lambda_2 + ca_2)^2 z_{02}^2}{\lambda_2 a_2^2}\right) + \frac{k^2(1-c)^2\alpha_2^2 z_{02}^2}{a_2^2}.$$

Rearranging this inequality, we obtain

$$z_{02}^2\left(\frac{\sigma^2(cb_2+(1-c)c_2)^2}{\lambda_2 b_2^2} + \frac{(1-c)^2(1-d)^2\alpha_2^2}{b_2^2} - \frac{\sigma^2((1-c)\lambda_2 + ca_2)^2}{\lambda_2 a_2^2}\right) +$$
$$z_{02}^2\left(-\frac{k^2(1-c)^2\alpha_2^2}{a_2^2}\right) < z_{01}^2\sigma^2\left(\frac{((1-c)\lambda_1 + ca_1)^2}{\lambda_1 a_1^2} - \frac{(cb_1+(1-c)c_1)^2}{\lambda_1 b_1^2}\right).$$

If both

$$\left(\frac{\sigma^2(cb_2+(1-c)c_2)^2}{\lambda_2 b_2^2} + \frac{(1-c)^2(1-d)^2\alpha_2^2}{b_2^2}\right) +$$
$$\left(-\frac{\sigma^2((1-c)\lambda_2 + ca_2)^2}{\lambda_2 a_2^2} - \frac{k^2(1-c)^2\alpha_2^2}{a_2^2}\right) \qquad (3.67.)$$

and

$$\sigma^2\left(\frac{((1-c)\lambda_1 + ca_1)^2}{\lambda_1 a_1^2} - \frac{(cb_1+(1-c)c_1)^2}{\lambda_1 b_1^2}\right) \qquad (3.68.)$$

have the same signs, the condition for the superiority of the LOL estimator over the LOR estimator is

$$\frac{z_{02}^2}{z_{01}^2} < f_5(\alpha_2^2). \tag{3.69.}$$

If (3.67.) and (3.68.) have opposite signs, the condition for the superiority of the LOL estimator over the LOR estimator is

$$\frac{z_{02}^2}{z_{01}^2} > f_5(\alpha_2^2). \tag{3.70.}$$

It is obvious that if (3.67.) and (3.68.) have opposite signs, the right hand side of (3.70.) is negative, thus (3.70.) always holds. Consequently, at that region the LOL estimator is uniformly superior to the LOR estimator. The condition for the positiveness of (3.67.) can be written as

$$\alpha_2^2 > \frac{\sigma^2\left(((1-c)\lambda_2 + ca_2)^2 b_2^2 - (cb_2 + (1-c)c_2)^2 a_2^2\right)}{\lambda_2(1-c)^2[(1-d)^2 a_2^2 - k^2 b_2^2]} \tag{3.71.}$$

and the condition for the positiveness of (3.68.) can be written as

$$((1-c)\lambda_1 + ca_1)^2 b_1^2 > (cb_1 + (1-c)c_1)^2 a_1^2. \tag{3.72.}$$

Of course, the opposite conditions are needed for the negativeness of (3.67.) and (3.68.). The vertical asymptote of the hyperbola $f_5(\alpha_2^2)$ is at the point

$$\alpha_2^2 = \frac{\sigma^2\left(((1-c)\lambda_2 + ca_2)^2 b_2^2 - (cb_2 + (1-c)c_2)^2 a_2^2\right)}{\lambda_2(1-c)^2[(1-d)^2 a_2^2 - k^2 b_2^2]}. \quad (3.73.)$$

Corollary 3.4: If $c = 0$ in Theorem 3.5, we get the Özbey and Kaçıranlar's (2015) results.

The estimation of the biasing parameters k and d are the important issue for a-multicollinear regression model. We have not made any attempt to estimate them. However, we refer our readers to Hoerl and Kennard (1970), Kibria (2003), Khalaf and Shukur (2005), Muniz and Kibria (2009) and Liu (1993) among others.

3.4. Numerical Example

In this section, we illustrate our theoretical results using the example given by Friedman and Montgomery (1985) (i.e., $\sigma^2 = 1$, $k = 0.1$, $\lambda_1 = 1.95$ and $\lambda_2 = 0.05$) and Özbey and Kaçıranlar (2015) (i.e., $d = 0.9$) as well as we let $c = 0.5$.

Firstly, let us consider the predictive performances of the LOR and the OLS estimators. From (3.34.), we get

$$f_1(\alpha_2^2) = \frac{0.024711}{0.1111\alpha_2^2 - 11.1111}, \quad (3.74.)$$

which is a hyperbola with a vertical asymptote at

$$\alpha_2^2 = 100. \quad (3.75.)$$

Because of both z_{02}^2 / z_{01}^2 and α_2^2 are positive, we are interested only in the points which lie in the first quadrant. Figure 3.1. illustrates this situation. For values of α_2^2 smaller than 100, the LOR estimator is uniformly superior to the OLS estimator. For larger values of α_2^2 there is a trade-off between these two estimators. If the value of the ratio z_{02}^2 / z_{01}^2 is smaller than the value of $f_1(\alpha_2^2)$, then the LOR estimator is superior to the OLS estimator, otherwise the OLS estimator is superior to the LOR estimator.

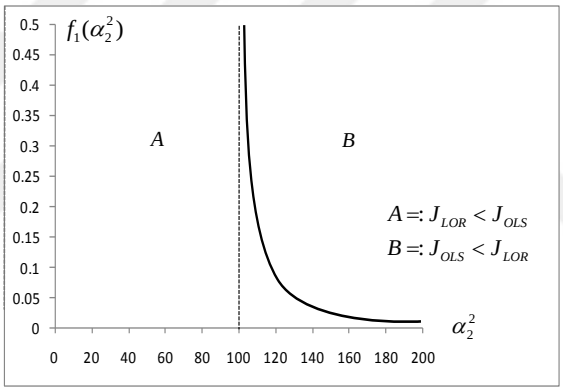


Figure 3.1. The pmse comparison of the LOR and the OLS

We take different values of c as $0.1, 0.2, ..., 0.9$ to determine the effect of c on the predictive performance of the LOR estimator and the OLS estimator.

Table 3.1. c and α_2^2 values for the LOR vs. the OLS.

c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
α_2^2	46.7	55.0	65.7	80.0	100.0	130.0	180.0	280.0	580.0

Table 3.1. shows that if c increases, the value of α_2^2 increases. That means, when α_2^2 is increasing, the region where the LOR estimator is uniformly superior to the OLS estimator is expanding.

-In this part, we get the same results of the example given by Friedman and Montgomery (1985) if $c=0$. Also, we get the same results of the example given by Özbey and Kaçıranlar (2015) if $c=d$ and $k=1$.

Secondly, let us consider the predictive performances of the LOR and the ORR estimators. From (3.42.), we get

$$f_2(\alpha_2^2)=\frac{0.0241}{0.3333\alpha_2^2-6.6667}, \tag{3.76.}$$

which is a hyperbola with a vertical asymptote at

$$\alpha_2^2=20. \tag{3.77.}$$

Figure 3.2. illustrates this situation. For values of α_2^2 smaller than 20, the LOR estimator is uniformly superior to the ORR estimator. For larger values of α_2^2 there is a trade-off between these two estimators. If the value of the ratio z_{02}^2/z_{01}^2 is smaller than the value of $f_2(\alpha_2^2)$, then the LOR estimator is superior to the ORR estimator, otherwise the ORR estimator is superior to the LOR estimator.

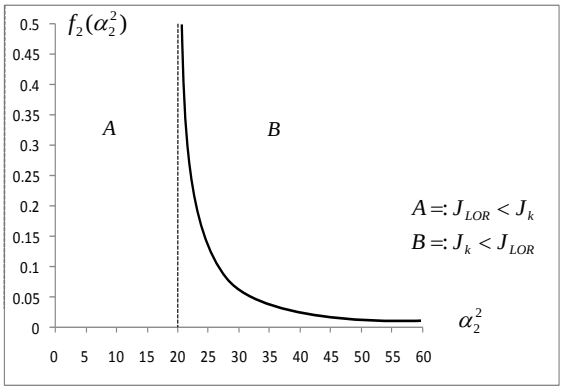


Figure 3.2. The pmse comparison of the LOR and the ORR

-The effect of c on the predictive performance of the LOR estimator and the ORR estimator is described in Table 3.2.

Table 3.2. c and α_2^2 values for the LOR vs. the ORR.

c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
α_2^2	11.6	13.3	15.3	17.5	20.0	22.9	26.2	30.0	34.5

Table 3.2. shows that if c increases, the value of α_2^2 increases. That means, when α_2^2 is increasing, the region where the LOR estimator is uniformly superior to the ORR estimator is expanding.

Thirdly, let us consider the predictive performances of the LOL and the OLS estimators. From (3.50.), we get

$$f_3(\alpha_2^2) = \frac{0.017236}{0.00226\,\alpha_2^2 - 1.8595}, \tag{3.78.}$$

which is a hyperbola with a vertical asymptote at

$$\alpha_2^2 = 820. \tag{3.79.}$$

Figure 3.3. illustrates this situation. For values of α_2^2 smaller than 820, the LOL estimator is uniformly superior to the OLS estimator. For larger values of α_2^2 there is a trade-off between these two estimators. If the value of the ratio z_{02}^2 / z_{01}^2 is smaller than the value of $f_3(\alpha_2^2)$, then the LOL estimator is superior to the OLS estimator, otherwise the OLS estimator is superior to the LOL estimator.

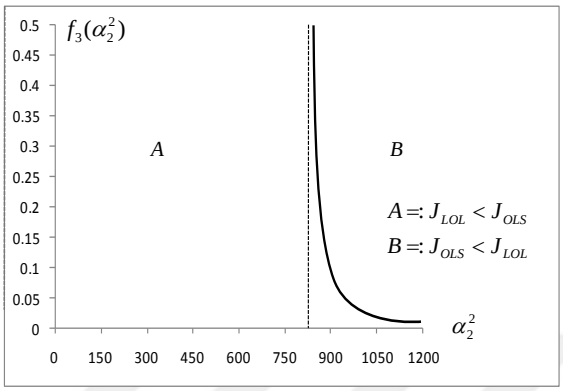


Figure 3.3. The pmse comparison of the LOL and the OLS

-The effect of c on the predictive performance of the LOL estimator and the OLS estimator is described in Table 3.3.

Table 3.3. c and α_2^2 values for the LOL vs. the OLS

c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
α_2^2	446.7	505.0	580.0	680.0	820.0	1030.0	1380.0	2080.0	4180.0

Table 3.3. shows that if c increases, the value of α_2^2 increases. That means, when α_2^2 is increasing, the region where the LOL estimator is uniformly superior to the OLS estimator is expanding.

-In this part, we get the same results of the example given by Özbey and Kaçıranlar (2015) if $c = 0$.

Fourthly, let us consider the predictive performances of the LOL and the Liu estimators. From (3.58.), we get

$$f_4(\alpha_2^2) = \frac{0.01694}{0.00681\alpha_2^2 - 1.7687}, \tag{3.80.}$$

which is a hyperbola with a vertical asymptote at

$$\alpha_2^2 = 260. \qquad (3.81.)$$

Figure 3.4. illustrates this situation. For values of α_2^2 smaller than 260, the LOL estimator is uniformly superior to the Liu estimator. For larger values of α_2^2 there is a trade-off between these two estimators. If the value of the ratio z_{02}^2 / z_{01}^2 is smaller than the value of $f_4(\alpha_2^2)$, then the LOL estimator is superior to the Liu estimator, otherwise the Liu estimator is superior to the LOL estimator.

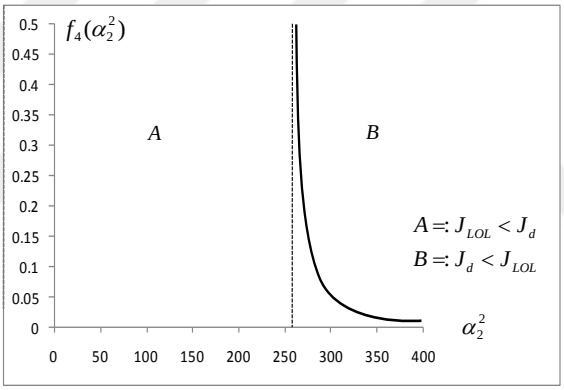


Figure 3.4. The pmse comparison of the LOL and the Liu

-The effect of c on the predictive performance of the LOL estimator and the Liu estimator is described in Table 3.4.

Table 3.4. c and α_2^2 values for the LOL vs. the Liu

c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
α_2^2	201.1	213.3	227.1	242.5	260.0	280.0	303.1	330.0	361.8

Table 3.4. shows that if c increases, the value of α_2^2 increases. That means, when α_2^2 is increasing, the region where the LOL estimator is uniformly superior to the Liu estimator is expanding.

Finally, let us consider the predictive performances of the LOL and the LOR estimators. From (3.66.), we get

$$f_5(\alpha_2^2) = \frac{0.00747}{0.10884\,\alpha_2^2 - 9.2516}, \tag{3.82.}$$

which is a hyperbola with a vertical asymptote at

$$\alpha_2^2 = 85. \tag{3.83.}$$

Figure 3.5. illustrates this situation. For values of α_2^2 smaller than 85, the LOL estimator is uniformly superior to the LOR estimator. For larger values of α_2^2 there is a trade-off between these two estimators. If the value of the ratio z_{02}^2 / z_{01}^2 is smaller than the value of $f_5(\alpha_2^2)$, then the LOL estimator is superior to the LOR estimator, otherwise the LOR estimator is superior to the LOL estimator.

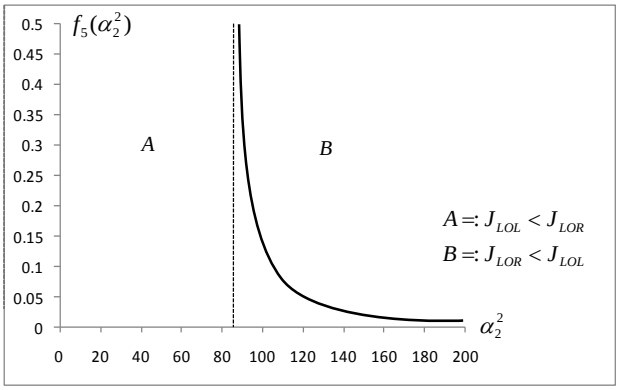


Figure 3.5. The pmse comparison of the LOL and the LOR

-The effect of c on the predictive performance of the LOL estimator and the LOR estimator is described in Table 3.5.

Table 3.5. c and α_2^2 values for the LOL vs. the LOR

c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
α_2^2	38.3	45.6	55.0	67.5	85.0	111.3	155.0	242.5	505.0

Table 3.5. shows that if c increases, the value of α_2^2 increases. That means, when α_2^2 is increasing, the region where the LOL estimator is uniformly superior to the LOR estimator is expanding. In this part, we get the same results of the example given by Özbey and Kaçıranlar (2015) if $c=0$.

3.5. Conclusion

We investigate the predictive performance of the proposed LOR estimator compared to the OLS and the ORR estimators and the proposed LOL estimator compared to the OLS, the Liu and the LOR estimators. The comparisons of these estimators are in terms of the pmse criterion at a specific point in the two-dimensional regressor variable spaces. In this context, the pmse of the LOR and the LOL estimators are developed and five theorems are given. In addition, three corollaries are given here illustrating that the theorems given by Friedman and Montgomery (1985) and Özbey and Kaçıranlar (2015) are just special cases of the Theorems 3.1, 3.3 and 3.5. The theoretical consequences are illustrated by a numerical example and the regions are assigned for the superiority of the proposed estimators. For some values of α_2^2, there are trade-offs between the relative effectiveness of the estimators. The OLS estimator is effective only when the value of α_2^2 is large compared to the LOR and the LOL estimators. The effectiveness of these techniques is also affected by the location of the prediction point and by the scalar c such that when c is increasing, the region where the proposed estimator is uniformly superior to the other mentioned estimators is increasing and if $c=0$, we get the same results of the example given by Friedman and Montgomery (1985) and Özbey and Kaçıranlar (2015). In the numerical example, a region is

established where the LOR and the LOL estimators are uniformly superior to the mentioned estimators above. This implies that it is theoretically possible to determine such a region.

3. PREDICTION OF THE LINEAR CONVEX COMBINATION OF BIASED REGRESSION ESTIMATORS

Issam DAWOUD

4. THE PREDICTIVE PERFORMANCE EVALUATION OF BIASED REGRESSION PREDICTORS WITH CORRELATED ERRORS

4.1. Introduction

The general linear regression model is given by

$$Y = X\beta + u, \tag{4.1.}$$

where Y is an $n \times 1$ vector of responses, X is an $n \times p$ full column rank matrix of non-stochastic predetermined regressors, β is a $p \times 1$ vector of unknown parameters, and u is an $n \times 1$ vector of $IID(0,V)$ random errors and V is a known $n \times n$ Positive Definite (pd) matrix (Trenkler, 1984).

Consider the problem of predicting a single drawing of the regressand given the vector of regressors. The actual drawing will be

$$y_0 = x'_0 \beta + \varepsilon_0, \tag{4.2.}$$

where y_0 is the 1×1 scalar value of the regressand, x_0 is the $p \times 1$ vector of prediction regressors; and ε_0 is the 1×1 scalar value of the prediction disturbances.

A more appealing set of assumptions is one which allows the prediction disturbance to be correlated with the sample disturbances. Therefore, we shall assume

$$\varepsilon_0 \sim (0, \sigma_0^2) \text{ and } Cov(u, \varepsilon_0) = w \tag{4.3.}$$

where w is the $n \times 1$ vector of covariances of the prediction disturbance with the vector of sample disturbances.

Goldberger (1962) used Lagrangian multipliers to find the BLUP which is

$$\hat{p}_0 = x'_0 b + w'V^{-1}(Y - Xb) = x'_0 b + w'V^{-1}e, \tag{4.4.}$$

where $b = (X'V^{-1}X)^{-1}X'V^{-1}Y$ (Aitken, 1935) and e is simply the $n \times 1$ vector of sample residuals from the generalized least squares regression. Also, he gave the pmse of the linear predictor $p = c'Y$ as follows:

$$\begin{aligned} J &= V(p) + (bias(p))^2 \\ &= c'Vc + \sigma_0^2 - 2c'w + (x'_0\beta - E(p))^2, \end{aligned} \tag{4.5.}$$

where c is a $n \times 1$ vector of constants.

Then, he derived the pmse of the BLUP as follows:

$$\begin{aligned} J_{BLUP} = \sigma_0^2 &+ x'_0(X'V^{-1}X)^{-1}x_0 - 2x'_0(X'V^{-1}X)^{-1}X'V^{-1}w \\ &- w'[V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}]w. \end{aligned} \tag{4.6.}$$

In this chapter, we derive the Linear Ridge Predictor (LRP) and the Linear Liu Predictor (LLP) when the disturbances are correlated and then discuss the predictive performance of these biased predictors with all pair-wise comparisons among these predictors. A numerical example is given to demonstrate the theoretical results.

4.2. The Prediction Problem

In practical applications of the general linear regression model, there is often severe but not exact multicollinearity. When multicollinearity is present, the method of GLS generally produces poor estimates of parameters. The variance of the GLS

estimates of the regression coefficients may be inflated, and the absolute value of the GLS estimates will be too large and they will be unstable; that is, their magnitudes and signs may change considerably on a different sample.

To overcome these problems, alternative biased estimators (i.e. ORR and Liu estimator) combined with the GLS estimator are proposed (see Trenkler, 1984 and Kaçıranlar, 2003). Now, we introduce the LRP and the LLP, respectively.

From Goldberger (1962), after applying Lagrangian multipliers, we have

$$2Vc - 2w - 2X\lambda = 0, \tag{4.7.}$$

$$2X'c - 2x_0 = 0, \tag{4.8.}$$

where λ is a $p \times 1$ vector of Lagrangian multipliers.
After solving the equations (4.7.) and (4.8.), we get

$$c = V^{-1}w + V^{-1}X\lambda, \tag{4.9.}$$

and

$$X'V^{-1}w + X'V^{-1}X\,\lambda = x_0. \tag{4.10.}$$

A popular numerical technique to deal with multicollinearity is the ridge regression. The Heteroscedastic Ridge Regression (HRR) estimator which is proposed by Trenkler (1984) is perhaps the first biased estimator combined the ORR estimator with the GLS estimator. The use of this estimator is motivated by the fact that the matrix $X'V^{-1}X$ in equation (4.10.) is stabilized when a positive number k is added to each of its main diagonal elements. Small relative changes in the matrix $X'V^{-1}X + k\mathrm{I}$ should

involve small relative changes in the inverse. In other words, augmenting the diagonal of $X'V^{-1}X$ with small positive quantities portrays the sensitivity of the solution to $X'V^{-1}X$.

Therefore, we get

$$\lambda = (X'V^{-1}X + k\mathrm{I})^{-1}(x_0 - X'V^{-1}w)$$

and

$$\hat{c}_1 = V^{-1}w + V^{-1}X(X'V^{-1}X + k\mathrm{I})^{-1}x_0 - V^{-1}X(X'V^{-1}X + k\mathrm{I})^{-1}X'V^{-1}w. \qquad (4.11.)$$

So, we define the LRP as

$$\begin{aligned}\hat{p}_1 = \hat{c}'_1 Y = x'_0(X'V^{-1}X + k\mathrm{I})^{-1}X'V^{-1}Y + w'V^{-1}Y - \\ w'V^{-1}X(X'V^{-1}X + k\mathrm{I})^{-1}X'V^{-1}Y \\ = x'_0 b_k + w'V^{-1}(Y - Xb_k) = x'_0 b_k + w'V^{-1}e_k,\end{aligned} \qquad (4.12.)$$

where e_k is simply the $n \times 1$ vector of sample residuals from $b_k = (X'V^{-1}X + k\mathrm{I})^{-1}X'V^{-1}Y$ which is the HRR estimator.

Following the method above, we define the LLP as

$$\hat{p}_2 = \hat{c}'_2 Y = x'_0 b_d + w'V^{-1}(Y - Xb_d) = x'_0 b_d + w'V^{-1}e_d, \qquad (4.13.)$$

where

$$\hat{c}_2 = V^{-1}w + V^{-1}X(X'V^{-1}X)^{-1}(X'V^{-1}X + d\mathrm{I})(X'V^{-1}X + \mathrm{I})^{-1}x_0$$
$$-V^{-1}X(X'V^{-1}X)^{-1}(X'V^{-1}X + d\mathrm{I})(X'V^{-1}X + \mathrm{I})^{-1}X'V^{-1}w$$

and e_d is simply the $n \times 1$ vector of sample residuals from $b_d = (X'V^{-1}X + \mathrm{I})^{-1}(X'V^{-1}X + d\,\mathrm{I})b$ which is the Generalized Liu Estimator (GLE) (Kaçıranlar, 2003).

4.3. The Properties of the Predictors

We discuss the properties of the predictors under $w = 0$, $w \neq 0$ and the stationary First-Order Autoregressive (AR(1)) process.

Assume the prediction disturbances are independent of the sample disturbances so that $w = 0$. This independence assumption would seem to be the appropriate one for the classical case where $V = \mathrm{I}$. So, the LRP reduces to $\hat{p}_1 = x'_0\hat{\beta}_k$, Friedman and Montgomery (1985) and the LLP reduces to $\hat{p}_2 = x'_0\hat{\beta}_d$ (Özbey and Kaçıranlar, 2015).

Since the reduction of the pmse associated with the use of our predictors (4.12.) and (4.13.) we obtain the pmse of each predictor under $w \neq 0$, respectively.

-For the LRP, we first obtain its variance by substitution of the $\hat{c}_1$ of (4.11.) into the general variance expression $V_1 = \hat{c}'_1 V\hat{c}_1 + \sigma_0^2 - 2\hat{c}'_1 w$, which reduces to

$$V_1 = \sigma_0^2 + x'_0 Ax_0 - 2x'_0 AX'V^{-1}w - w'[V^{-1} - V^{-1}X\,AX'V^{-1}]w,$$

where $A = (X'V^{-1}X + k\mathrm{I})^{-1}(X'V^{-1}X)(X'V^{-1}X + k\mathrm{I})^{-1}$,

and its bias is calculated by finding

$$\begin{aligned} bias_1 &= x'_0\beta - E(\hat{p}_1) \\ &= x'_0\beta - E(x'_0 b_k + w'V^{-1}(Y - Xb_k)) \\ &= k[x'_0 - w'V^{-1}X](X'V^{-1}X + k\mathrm{I})^{-1}\beta. \end{aligned}$$

Therefore, the pmse of the LRP is given by

$$J_{LRP} = V_1 + (bias_1)^2 .$$

$$\begin{aligned} J_{LRP} &= \sigma_0^2 + x'_0 A x_0 - 2x'_0 A X'V^{-1}w - w'[V^{-1} - V^{-1}X\,AX'V^{-1}]w \\ &\quad + (k[x'_0 - w'V^{-1}X](X'V^{-1}X + k\mathrm{I})^{-1}\beta)^2. \end{aligned} \qquad (4.14.)$$

-For the LLP, we first obtain its variance by substitution of the $\hat{c}_2$ of (4.13.) into the general variance expression $V_2 = \hat{c}'_2 V\hat{c}_2 + \sigma_0^2 - 2\hat{c}'_2 w$, which reduces to

$$V_2 = \sigma_0^2 + x'_0 B x_0 - 2x'_0 B X'V^{-1}w - w'[V^{-1} - V^{-1}X\,BX'V^{-1}]w,$$

where $B = (X'V^{-1}X + \mathrm{I})^{-1}(X'V^{-1}X + d\mathrm{I})(X'V^{-1}X)^{-1}(X'V^{-1}X + d\mathrm{I})(X'V^{-1}X + \mathrm{I})^{-1}$,

and its bias is calculated by finding

$$\begin{aligned} bias_2 &= x'_0\beta - E(\hat{p}_2) \\ &= x'_0\beta - E(x'_0 b_d + w'V^{-1}(Y - Xb_d)) \\ &= (1-d)[x'_0 - w'V^{-1}X](X'V^{-1}X + \mathrm{I})^{-1}\beta. \end{aligned}$$

Therefore, the pmse of the LLP is given by

$$J_{LLP} = V_2 + (bias_2)^2 .$$

$$\begin{aligned} J_{LLP} = \sigma_0^2 + x'_0 B x_0 - 2x'_0 B X'V^{-1}w - w'[V^{-1} - V^{-1}X B X'V^{-1}]w \\ + ((1-d)[x'_0 - w'V^{-1}X](X'V^{-1}X + \mathrm{I})^{-1}\beta)^2 . \end{aligned} \tag{4.15.}$$

Now, we discuss the above mentioned predictors under the stationary AR(1) process in which the disturbances are generated by

$$u_t = \rho u_{t-1} + \upsilon_t , \tag{4.16.}$$

where υ_t is a circular random disturbance with $E\upsilon_t = 0$, $E\upsilon_t^2 = (1-\rho^2)\sigma^2$ and $|\rho| < 1$. Applied to the sample period disturbances, that is, for $t = 1,...,n$, (4.16.) means that the sample disturbances variance-covariance matrix is

$$V = \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ \rho & 1 & \cdots & \rho^{n-2} \\ \vdots & \vdots & & \vdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & 1 \end{pmatrix}. \tag{4.17.}$$

Suppose that the autoregressive specification (4.16.) is maintained for the post-sample observations $t = n+1, n+2, ..., n+s$ and consider prediction for the period $n+s$.

Application of (4.16.) gives

$E\varepsilon_0 = 0;$ $E\varepsilon_0^2 = \sigma_0^2 = \sigma^2;$ and

$$w = E\varepsilon_0 u = \sigma^2 \begin{pmatrix} \rho^{n-1+s} \\ \vdots \\ \rho^s \end{pmatrix} = \rho^s V_n , \tag{4.18.}$$

where V_n is the final column of V (Goldberger, 1962).

The BLUP reduces to

$$\begin{aligned} \hat{p}_0 &= x'_0 b + \rho^s V'_n V^{-1} e \\ &= x'_0 b + \rho^s (0\ 0 \cdots 0\ 1) e \\ &= x'_0 b + \rho^s e_n , \end{aligned} \tag{4.19.}$$

where e_n is the final element of e, Goldberger (1962).

Then, the LRP reduces to

$$\begin{aligned} \hat{p}_1 &= x'_0 b_k + \rho^s V'_n V^{-1} e_k \\ &= x'_0 b_k + \rho^s (0\ 0 \cdots 0\ 1) e_k \\ &= x'_0 b_k + \rho^s e_{kn} , \end{aligned} \tag{4.20.}$$

where e_{kn} is the final element of e_k.

Also, the LLP reduces to

$$\begin{aligned} \hat{p}_2 &= x'_0 b_d + \rho^s V'_n V^{-1} e_d \\ &= x'_0 b_d + \rho^s (0\ 0 \cdots 0\ 1) e_d \\ &= x'_0 b_d + \rho^s e_{dn} , \end{aligned} \tag{4.21.}$$

where e_{dn} is the final element of e_d.

In particular, to predict the first post-sample drawing, we let $s=1$. So, (4.19.), (4.20.) and (4.21.) reduce respectively to

$$\hat{p}_0 = x'_0\, b + \rho e_n, \tag{4.22.}$$

$$\hat{p}_1 = x'_0\, b_k + \rho e_{kn}, \tag{4.23.}$$

$$\hat{p}_2 = x'_0\, b_d + \rho e_{dn}. \tag{4.24.}$$

Now, we examine the pmses of our predictors under AR(1) process which were given in (4.20.) and (4.21.). In this case, we have

$$X'V^{-1}w = X'V^{-1}\rho^s V_n = \rho^s X' \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \rho^s X_n, \tag{4.25.}$$

where X_n is the final row of X, arranged in a column.

$$w'V^{-1}w = \rho^s V'_n V^{-1} V_n \rho^s = \rho^{2s}\sigma^2, \tag{4.26.}$$

$$V^{-1} = \frac{1}{\sigma^2(1-\rho^2)} \begin{bmatrix} 1 & -\rho & \cdots 0 \\ -\rho & 1+\rho^2 & \cdots 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \end{bmatrix}, \tag{4.27.}$$

and also $\sigma_0^2 = \sigma^2$.

For convenience, the orthogonal form of (4.1.)

$$Y = Z\alpha + u \tag{4.28.}$$

will be used. Where $Z = XU$ and $\alpha = U'\beta$. Then the GLS estimator of α is

$$\hat{\alpha} = (Z'V^{-1}Z)^{-1}Z'V^{-1}y = \Gamma^{-1}Z'V^{-1}Y \tag{4.29.}$$

where $\Gamma = diag(\gamma_1, \gamma_2, ..., \gamma_p)$ is the diagonal matrix of the eigenvalues of $Z'V^{-1}Z$ and γ's are in descending order.

The vector of the original parameters is given by

$$b = U\hat{\alpha}. \tag{4.30.}$$

-Specialized to the case $s = 1$:

-For the BLUP, inserting these into (4.6.), we get

$$J_{BLUP} = (1-\rho^2)\sigma^2 + \sum_{i=1}^{p}\frac{z_{oi}^2}{\gamma_i} - 2\rho\sum_{i=1}^{p}\frac{N_i}{\gamma_i} + \rho^2\sum_{i=1}^{p}\frac{Z_{ni}^2}{\gamma_i} \tag{4.31.}$$

where Z_n is the final row of Z, arranged in a column such that $Z_{ni}'Z_{ni} = Z_i^2$ and $z_{0i}Z_{ni} = N_i$, $i = 1, 2, ..., p$.

-For the LRP, inserting these into (4.14.), we get

$$J_{LRP} = (1-\rho^2)\sigma^2 + \sum_{i=1}^{p} \frac{\gamma_i z_{oi}^2}{(\gamma_i + k)^2} - 2\rho \sum_{i=1}^{p} \frac{\gamma_i N_i}{(\gamma_i + k)^2} + \rho^2 \sum_{i=1}^{p} \frac{\gamma_i Z_{ni}^2}{(\gamma_i + k)^2} + k^2 \left[\sum_{i=1}^{p} \frac{(z_{oi} - \rho Z_{ni})\alpha_i}{(\gamma_i + k)} \right]^2 . \quad (4.32.)$$

-For the LLP, inserting these into (4.15.), we get

$$J_{LLP} = (1-\rho^2)\sigma^2 + \sum_{i=1}^{p} \frac{(\gamma_i + d)^2 z_{oi}^2}{\gamma_i(\gamma_i + 1)^2} - 2\rho \sum_{i=1}^{p} \frac{(\gamma_i + d)^2 N_i}{\gamma_i(\gamma_i + 1)^2} + \rho^2 \sum_{i=1}^{p} \frac{(\gamma_i + d)^2 Z_{ni}^2}{\gamma_i(\gamma_i + 1)^2} + (1-d)^2 \left[\sum_{i=1}^{p} \frac{(z_{oi} - \rho Z_{ni})\alpha_i}{(\gamma_i + 1)} \right]^2 . \quad (4.33.)$$

4.4. The Comparisons of the Prediction Mean Squared Errors in Two Dimensional Space

Following Friedman and Montgomery (1985) and Özbey and Kaçıranlar (2015), we discuss the predictive performance of the BLUP, the LRP and the LLP. So, comparisons of J_{LRP} with J_{BLUP} and J_{LLP} with J_{BLUP} and J_{LRP} are made and stated in the following three theorems.

Theorem 4.1:

1. If $\alpha_2^2 > \frac{(2\gamma_2 + k)}{k\gamma_2}$, then $J_{LRP} < J_{BLUP}$

for $\alpha_2^2 < \frac{(\gamma_2 + k)^2[2\rho(h_1N_1 + h_2N_2) - h_1(z_{01}^2 + \rho^2 Z_1^2) - \rho^2 h_2 Z_2^2]}{\rho^2 k^2 (4N_2 - \rho^2 Z_2^2)}$

or for $\alpha_2^2 > \frac{(\gamma_2 + k)^2[2\rho(h_1N_1 + h_2N_2) - h_1(z_{01}^2 + \rho^2 Z_1^2) - \rho^2 h_2 Z_2^2]}{\rho^2 k^2 (4N_2 - \rho^2 Z_2^2)}$ iff $\frac{z_{02}^2}{z_{01}^2} < f_6(\alpha_2^2)$.

2. If $\alpha_2^2 < \frac{(2\gamma_2 + k)}{k\gamma_2}$, then $J_{LRP} < J_{BLUP}$

for $\alpha_2^2 > \frac{(\gamma_2 + k)^2[2\rho(h_1N_1 + h_2N_2) - h_1(z_{01}^2 + \rho^2 Z_1^2) - \rho^2 h_2 Z_2^2]}{\rho^2 k^2 (4N_2 - \rho^2 Z_2^2)}$

or for $\alpha_2^2 < \frac{(\gamma_2 + k)^2[2\rho(h_1N_1 + h_2N_2) - h_1(z_{01}^2 + \rho^2 Z_1^2) - \rho^2 h_2 Z_2^2]}{\rho^2 k^2 (4N_2 - \rho^2 Z_2^2)}$ iff $\frac{z_{02}^2}{z_{01}^2} < f_6(\alpha_2^2)$.

3. If $\rho = 0$, we get the Friedman and Montgomery's (1985) results.

Where

$$f_6(\alpha_2^2) = \frac{h_1 + \frac{M}{z_{01}^2}}{k^2 \frac{\alpha_2^2}{(\gamma_2 + k)^2} - h_2},\ h_1 = \left[\frac{1}{\gamma_1} - \frac{\gamma_1}{(\gamma_1 + k)^2}\right], h_2 = \left[\frac{1}{\gamma_2} - \frac{\gamma_2}{(\gamma_2 + k)^2}\right]$$

and

$$M = \rho^2 h_1 Z_1^2 + \rho^2\left[h_2 - \frac{k^2\rho^2\alpha_2^2}{(\gamma_2 + k)^2}\right]Z_2^2 - 2\rho h_1 N_1 - 2\rho\left[h_2 - \frac{2\rho k^2 \alpha_2^2}{(\gamma_2 + k)^2}\right]N_2.$$

Proof: If the LRP is superior to the BLUP, then $J_{LRP} < J_{BLUP}$. That is

$$z_{02}^2\left(\frac{k^2\alpha_2^2}{(\gamma_2 + k)^2} - h_2\right) < z_{01}^2 h_1 + M$$

where

$$M = \rho^2 h_1 Z_1^2 + \rho^2\left[h_2 - \frac{k^2\rho^2\alpha_2^2}{(\gamma_2 + k)^2}\right]Z_2^2 - 2\rho h_1 N_1 - 2\rho\left[h_2 - \frac{2\rho k^2 \alpha_2^2}{(\gamma_2 + k)^2}\right]N_2,$$

$$h_1 = \left[\frac{1}{\gamma_1} - \frac{\gamma_1}{(\gamma_1 + k)^2}\right],\ h_2 = \left[\frac{1}{\gamma_2} - \frac{\gamma_2}{(\gamma_2 + k)^2}\right]$$

and

$$f_6(\alpha_2^2) = \frac{h_1 + \dfrac{M}{z_{01}^2}}{k^2 \dfrac{\alpha_2^2}{(\gamma_2 + k)^2} - h_2}. \tag{4.34}$$

If $\rho = 0$, then $M = 0$, $V = \mathrm{I}$ and γ becomes λ which is the eigenvalue of $(Z'Z)$.

So,

$$f_6^*(\alpha_2^2) = \frac{\left[\dfrac{1}{\lambda_1} - \dfrac{\lambda_1}{(\lambda_1 + k)^2}\right]}{\dfrac{\lambda_2}{(\lambda_2 + k)^2} + k^2 \dfrac{\alpha_2^2}{(\lambda_2 + k)^2} - \dfrac{1}{\lambda_2}} \tag{4.35.}$$

(Friedman and Montgomery's (1985) results).

If both

$$h_1 + \frac{M}{z_{01}^2} \tag{4.36.}$$

and

$$\left(\frac{k^2\alpha_2^2}{(\gamma_2 + k)^2} - h_2\right) \tag{4.37.}$$

have the same signs, the condition for the superiority of the LRP over the BLUP is

$$z_{02}^2 / z_{01}^2 < f_6(\alpha_2^2). \tag{4.38.}$$

If (4.36.) and (4.37.) have opposite signs, the condition for the superiority of the LRP over the BLUP is

$$z_{02}^2 / z_{01}^2 > f_6(\alpha_2^2). \tag{4.39.}$$

It is obvious that if (4.36.) and (4.37.) have opposite signs, the right hand side of (4.39.) is negative, thus (4.39.) always holds. Consequently, at the region the LRP is uniformly superior to the BLUP. The condition for the positiveness of (4.36.) can be written as

$$h_1 + \frac{M}{z_{01}^2} > 0. \tag{4.40.}$$

After substituting the value of M, we get

$$\alpha_2^2 > \frac{(\gamma_2 + k)^2[2\rho(h_1 N_1 + h_2 N_2) - h_1(z_{01}^2 + \rho^2 Z_1^2) - \rho^2 h_2 Z_2^2]}{\rho^2 k^2 (4N_2 - \rho^2 Z_2^2)}. \tag{4.41.}$$

And the condition for the positiveness of (4.37.) can be written as

$$\alpha_2^2 > \frac{(2\gamma_2 + k)}{k\gamma_2}. \tag{4.42.}$$

Of course, the opposite conditions are needed for the negativeness of (4.36.) and (4.37.). The vertical asymptotes of the hyperbola $f_6(\alpha_2^2)$ are at the points

$$\alpha_2^2 = \frac{(\gamma_2 + k)^2[2\rho(h_1 N_1 + h_2 N_2) - h_1(z_{01}^2 + \rho^2 Z_1^2) - \rho^2 h_2 Z_2^2]}{\rho^2 k^2 (4N_2 - \rho^2 Z_2^2)} \tag{4.43.}$$

and

$$\alpha_2^2 = \frac{(2\gamma_2 + k)}{k\gamma_2}. \tag{4.44.}$$

Theorem 4.2:

1. If $\alpha_2^2 > \frac{2\gamma_2 + (1+d)}{(1-d)\gamma_2}$, then $J_{LLP} < J_{BLUP}$

for $\alpha_2^2 < \frac{(\gamma_2+1)^2[2\rho(s_1N_1 + s_2N_2) - s_1(z_{01}^2 + \rho^2 Z_1^2) - \rho^2 s_2 Z_2^2]}{\rho^2(1-d)^2(4N_2 - \rho^2 Z_2^2)}$

or for $\alpha_2^2 > \frac{(\gamma_2+1)^2[2\rho(s_1N_1 + s_2N_2) - s_1(z_{01}^2 + \rho^2 Z_1^2) - \rho^2 s_2 Z_2^2]}{\rho^2(1-d)^2(4N_2 - \rho^2 Z_2^2)}$ iff $\frac{z_{02}^2}{z_{01}^2} < f_7(\alpha_2^2)$.

2. If $\alpha_2^2 < \frac{2\gamma_2 + (1+d)}{(1-d)\gamma_2}$, then $J_{LLP} < J_{BLUP}$

for $\alpha_2^2 > \frac{(\gamma_2+1)^2[2\rho(s_1N_1 + s_2N_2) - s_1(z_{01}^2 + \rho^2 Z_1^2) - \rho^2 s_2 Z_2^2]}{\rho^2(1-d)^2(4N_2 - \rho^2 Z_2^2)}$

or for $\alpha_2^2 < \frac{(\gamma_2+1)^2[2\rho(s_1N_1 + s_2N_2) - s_1(z_{01}^2 + \rho^2 Z_1^2) - \rho^2 s_2 Z_2^2]}{\rho^2(1-d)^2(4N_2 - \rho^2 Z_2^2)}$ iff $\frac{z_{02}^2}{z_{01}^2} < f_7(\alpha_2^2)$.

3. If $\rho = 0$, we get the Özbey and Kaçıranlar's (2015) results.

Where

$$f_7(\alpha_2^2) = \frac{s_1 + \frac{M}{z_{01}^2}}{\frac{(1-d)^2\alpha_2^2}{(\gamma_2+1)^2} - s_2},\ s_1 = \left[\frac{1}{\gamma_1} - \frac{(\gamma_1+d)^2}{\gamma_1(\gamma_1+1)^2}\right],\ s_2 = \left[\frac{1}{\gamma_2} - \frac{(\gamma_2+d)^2}{\gamma_2(\gamma_2+1)^2}\right] \text{ and}$$

$$M = \rho^2 s_1 Z_1^2 + \rho^2\left[s_2 - \frac{(1-d)^2\rho^2\alpha_2^2}{(\gamma_2+1)^2}\right]Z_2^2 - 2\rho s_1 N_1 - 2\rho\left[s_2 - \frac{2\rho(1-d)^2\alpha_2^2}{(\gamma_2+1)^2}\right]N_2.$$

Proof: If the LLP is superior to the BLUP, then $J_{LLP} < J_{BLUP}$. That is

$$z_{02}^2\left[\frac{(1-d)^2\alpha_2^2}{(\gamma_2+1)^2}-s_2\right]<z_{01}^2 s_1+M$$

where

$$M=\rho^2 s_1 Z_1^2+\rho^2\left[s_2-\frac{(1-d)^2\rho^2\alpha_2^2}{(\gamma_2+1)^2}\right]Z_2^2-2\rho s_1 N_1-2\rho\left[s_2-\frac{2\rho(1-d)^2\alpha_2^2}{(\gamma_2+1)^2}\right]N_2,$$

$$s_1=\left[\frac{1}{\gamma_1}-\frac{(\gamma_1+d)^2}{\gamma_1(\gamma_1+1)^2}\right],\ s_2=\left[\frac{1}{\gamma_2}-\frac{(\gamma_2+d)^2}{\gamma_2(\gamma_2+1)^2}\right]$$

and

$$f_7(\alpha_2^2)=\frac{s_1+\frac{M}{z_{01}^2}}{\frac{(1-d)^2\alpha_2^2}{(\gamma_2+1)^2}-s_2}. \qquad (4.45.)$$

If $\rho=0$, then $M=0$, $V=\mathrm{I}$ and γ becomes λ.

So,

$$f_7^*(\alpha_2^2)=\frac{\left[\frac{1}{\lambda_1}-\frac{(\lambda_1+d)^2}{\lambda_1(\lambda_1+1)^2}\right]}{\frac{(\lambda_2+d)^2}{\lambda_2(\lambda_2+1)^2}+\frac{(1-d)^2\alpha_2^2}{(\lambda_2+1)^2}-\frac{1}{\lambda_2}} \qquad (4.46.)$$

(Özbey and Kaçıranlar's (2015) results).

If both

$$s_1 + \frac{M}{z_{01}^2} \tag{4.47.}$$

and

$$\frac{(1-d)^2\alpha_2^2}{(\gamma_2+1)^2} - s_2 \tag{4.48.}$$

have the same signs, the condition for the superiority of the LLP over the BLUP is

$$z_{02}^2 / z_{01}^2 < f_7(\alpha_2^2). \tag{4.49.}$$

If (4.47.) and (4.48.) have opposite signs, the condition for the superiority of the LLP over the BLUP is

$$z_{02}^2 / z_{01}^2 > f_7(\alpha_2^2). \tag{4.50.}$$

It is obvious that if (4.47.) and (4.48.) have opposite signs, the right hand side of (4.50.) is negative, thus (4.50.) always holds. Consequently, at the region the LLP is uniformly superior to the BLUP. The condition for the positiveness of (4.47.) can be written as

$$s_1 + \frac{M}{z_{01}^2} > 0.$$

After substituting the value of M , we get

$$\alpha_2^2 > \frac{(\gamma_2+1)^2[2\rho(s_1N_1+s_2N_2)-s_1(z_{01}^2+\rho^2Z_1^2)-\rho^2s_2Z_2^2]}{\rho^2(1-d)^2(4N_2-\rho^2Z_2^2)}. \quad (4.51.)$$

And the condition for the positiveness of (4.48.) can be written as

$$\alpha_2^2 > \frac{2\gamma_2+(1+d)}{(1-d)\gamma_2}. \quad (4.52.)$$

Of course, the opposite conditions are needed for the negativeness of (4.47.) and (4.48.).

The vertical asymptotes of the hyperbola $f_7(\alpha_2^2)$ are at the points

$$\alpha_2^2 = \frac{(\gamma_2+1)^2[2\rho(s_1N_1+s_2N_2)-s_1(z_{01}^2+\rho^2Z_1^2)-\rho^2s_2Z_2^2]}{\rho^2(1-d)^2(4N_2-\rho^2Z_2^2)} \quad (4.53.)$$

and

$$\alpha_2^2 = \frac{2\gamma_2+(1+d)}{(1-d)\gamma_2}. \quad (4.54.)$$

Theorem 4.3:

1. If $\alpha_2^2 > \frac{\gamma_2^2(\gamma_2+1)^2-(\gamma_2+k)^2(\gamma_2+d)^2}{\gamma_2[(1-d)^2(\gamma_2+k)^2-k^2(\gamma_2+1)^2]}$, then $J_{LLP} < J_{LRP}$

for $\alpha_2^2 < \frac{(\gamma_2+1)^2[2\rho(r_1N_1+r_2N_2)-r_1(z_{01}^2+\rho^2Z_1^2)-\rho^2r_2Z_2^2]}{\rho^2(1-d)^2(4N_2-\rho^2Z_2^2)}$

or for $\alpha_2^2 > \frac{(\gamma_2+1)^2[2\rho(r_1N_1+r_2N_2)-r_1(z_{01}^2+\rho^2Z_1^2)-\rho^2r_2Z_2^2]}{\rho^2(1-d)^2(4N_2-\rho^2Z_2^2)}$ iff $\frac{z_{02}^2}{z_{01}^2} < f_8(\alpha_2^2)$.

2. If $\alpha_2^2 < \frac{\gamma_2^2(\gamma_2+1)^2-(\gamma_2+k)^2(\gamma_2+d)^2}{\gamma_2[(1-d)^2(\gamma_2+k)^2-k^2(\gamma_2+1)^2]}$, then $J_{LLP} < J_{LRP}$

for $\alpha_2^2 > \dfrac{(\gamma_2+1)^2[2\rho(r_1N_1+r_2N_2)-r_1(z_{01}^2+\rho^2Z_1^2)-\rho^2r_2Z_2^2]}{\rho^2(1-d)^2(4N_2-\rho^2Z_2^2)}$

or for $\alpha_2^2 < \dfrac{(\gamma_2+1)^2[2\rho(r_1N_1+r_2N_2)-r_1(z_{01}^2+\rho^2Z_1^2)-\rho^2r_2Z_2^2]}{\rho^2(1-d)^2(4N_2-\rho^2Z_2^2)}$ iff $\dfrac{z_{02}^2}{z_{01}^2} < f_8(\alpha_2^2)$.

3. If $\rho = 0$, we get the Özbey and Kaçıranlar's (2015) results.

Where $f_8(\alpha_2^2) = \dfrac{r_1 + \dfrac{M}{z_{01}^2}}{\alpha_2^2\left(\dfrac{(1-d)^2}{(\gamma_2+1)^2} - \dfrac{k^2}{(\gamma_2+k)^2}\right) - r_2}$,

$$r_1 = \left[\frac{\gamma_1}{(\gamma_1+k)^2} - \frac{(\gamma_1+d)^2}{\gamma_1(\gamma_1+1)^2}\right],\ r_2 = \left[\frac{\gamma_2}{(\gamma_2+k)^2} - \frac{(\gamma_2+d)^2}{\gamma_2(\gamma_2+1)^2}\right] \text{ and}$$

$$M = \rho^2 r_1 Z_1^2 + \rho^2\left[r_2 - \frac{(1-d)^2\rho^2\alpha_2^2}{(\gamma_2+1)^2}\right]Z_2^2 - 2\rho r_1 N_1 - 2\rho\left[r_2 - \frac{2\rho(1-d)^2\alpha_2^2}{(\gamma_2+1)^2}\right]N_2.$$

Proof: If the LLP is superior to the LRP, then $J_{LLP} < J_{LRP}$. That is

$$z_{02}^2\left[\frac{(1-d)^2\alpha_2^2}{(\gamma_2+1)^2} - \frac{k^2\alpha_2^2}{(\gamma_2+k)^2} - r_2\right] < z_{01}^2 r_1 + M$$

where

$$M = \rho^2 r_1 Z_1^2 + \rho^2\left[r_2 - \frac{(1-d)^2\rho^2\alpha_2^2}{(\gamma_2+1)^2}\right]Z_2^2 - 2\rho r_1 N_1 - 2\rho\left[r_2 - \frac{2\rho(1-d)^2\alpha_2^2}{(\gamma_2+1)^2}\right]N_2,$$

$$r_1 = \left[\frac{\gamma_1}{(\gamma_1+k)^2} - \frac{(\gamma_1+d)^2}{\gamma_1(\gamma_1+1)^2}\right],\ r_2 = \left[\frac{\gamma_2}{(\gamma_2+k)^2} - \frac{(\gamma_2+d)^2}{\gamma_2(\gamma_2+1)^2}\right]$$

and

$$f_8(\alpha_2^2)=\frac{r_1+\dfrac{M}{z_{01}^2}}{\alpha_2^2\left(\dfrac{(1-d)^2}{(\gamma_2+1)^2}-\dfrac{k^2}{(\gamma_2+k)^2}\right)-r_2}. \tag{4.55.}$$

If $\rho=0$, then $M=0$, $V=\mathrm{I}$ and γ becomes λ.

So,

$$f_8^*(\alpha_2^2)=\frac{\left[\dfrac{\lambda_1}{(\lambda_1+k)^2}-\dfrac{(\lambda_1+d)^2}{\lambda_1(\lambda_1+1)^2}\right]}{\dfrac{(\lambda_2+d)^2}{\lambda_2(\lambda_2+1)^2}+\dfrac{(1-d)^2\alpha_2^2}{(\lambda_2+1)^2}-\dfrac{\lambda_1}{(\lambda_1+k)^2}-\dfrac{k^2\alpha_2^2}{(\lambda_1+k)^2}} \tag{4.56.}$$

(Özbey and Kaçıranlar's (2015) results).

If both

$$r_1+\frac{M}{z_{01}^2} \tag{4.57.}$$

and

$$\alpha_2^2\left(\frac{(1-d)^2}{(\gamma_2+1)^2}-\frac{k^2}{(\gamma_2+k)^2}\right)-r_2 \tag{4.58.}$$

have the same signs, the condition for the superiority of the LLP over the LRP is

$$z_{02}^2/z_{01}^2<f_8(\alpha_2^2) \tag{4.59.}$$

If (4.57.) and (4.58.) have opposite signs, the condition for the superiority of the LLP over the LRP is

$$z_{02}^2 / z_{01}^2 > f_8(\alpha_2^2). \tag{4.60.}$$

It is obvious that if (4.57.) and (4.58.) have opposite signs, the right hand side of (4.60.) is negative, thus (4.60.) always holds. Consequently, at the region the LLP is uniformly superior to the LRP. The condition for the positiveness of (4.57.) can be written as

$$r_1 + \frac{M}{z_{01}^2} > 0.$$

After substituting the value of M, we get

$$\alpha_2^2 > \frac{(\gamma_2+1)^2[2\rho(r_1N_1+r_2N_2) - r_1(z_{01}^2+\rho^2Z_1^2) - \rho^2 r_2 Z_2^2]}{\rho^2(1-d)^2(4N_2-\rho^2Z_2^2)}. \tag{4.61.}$$

And the condition for the positiveness of (4.58.) can be written as

$$\alpha_2^2 > \frac{\gamma_2^2(\gamma_2+1)^2 - (\gamma_2+k)^2(\gamma_2+d)^2}{\gamma_2[(1-d)^2(\gamma_2+k)^2 - k^2(\gamma_2+1)^2]}. \tag{4.62.}$$

Of course, the opposite conditions are needed for the negativeness of (4.57.) and (4.58.). The vertical asymptotes of the hyperbola $f_8(\alpha_2^2)$ are at the points

$$\alpha_2^2 = \frac{(\gamma_2+1)^2[2\rho(r_1N_1+r_2N_2) - r_1(z_{01}^2+\rho^2Z_1^2) - \rho^2 r_2 Z_2^2]}{\rho^2(1-d)^2(4N_2-\rho^2Z_2^2)} \tag{4.63.}$$

and

$$\alpha_2^2 = \frac{\gamma_2^2(\gamma_2+1)^2 - (\gamma_2+k)^2(\gamma_2+d)^2}{\gamma_2[(1-d)^2(\gamma_2+k)^2 - k^2(\gamma_2+1)^2]}. \tag{4.64.}$$

4.5. Numerical Example

In this section, we illustrate our theoretical results using the example given by Friedman and Montgomery (1985) (i.e., $k = 0.1$ and $r_{12} = 0.95$) and Özbey and Kaçıranlar (2015) (i.e., $d = 0.9$).

We choose the correlation error as $\rho = 0.5$, $(\gamma_1 = 1.3, \gamma_2 = 0.1)$, $z_{01} = -1, z_{02} = 2, Z_1 = 1$ and $Z_2 = 1.5$.

Firstly, let us consider the predictive performances of the LRP and the BLUP. From (4.34.), we get

$$f_6(\alpha_2^2) = \frac{0.714844\alpha_2^2 - 18.0428}{0.25\alpha_2^2 - 7.5} \tag{4.65.}$$

which is a hyperbola with a horizontal asymptote at $\alpha_2^2 = 25$ and a vertical asymptote at

$$\alpha_2^2 = 30. \tag{4.66.}$$

Because both z_{02}^2 / z_{01}^2 and α_2^2 are positive, we are interested only in the points which lie in the first quadrant. Figure 4.1. ($\rho = 0.5$) illustrates the regions for the superiority of the predictors. As it seems, in this example, the trade-off occurs from the value $\alpha_2^2 = 0$ till $\alpha_2^2 = 25$ and over the value $\alpha_2^2 = 30$. For smaller values of α_2^2, the LRP is

uniformly superior to the BLUP. Again, for small values of the ratio z_{02}^2 / z_{01}^2, the trade-off is on the LRP's side.

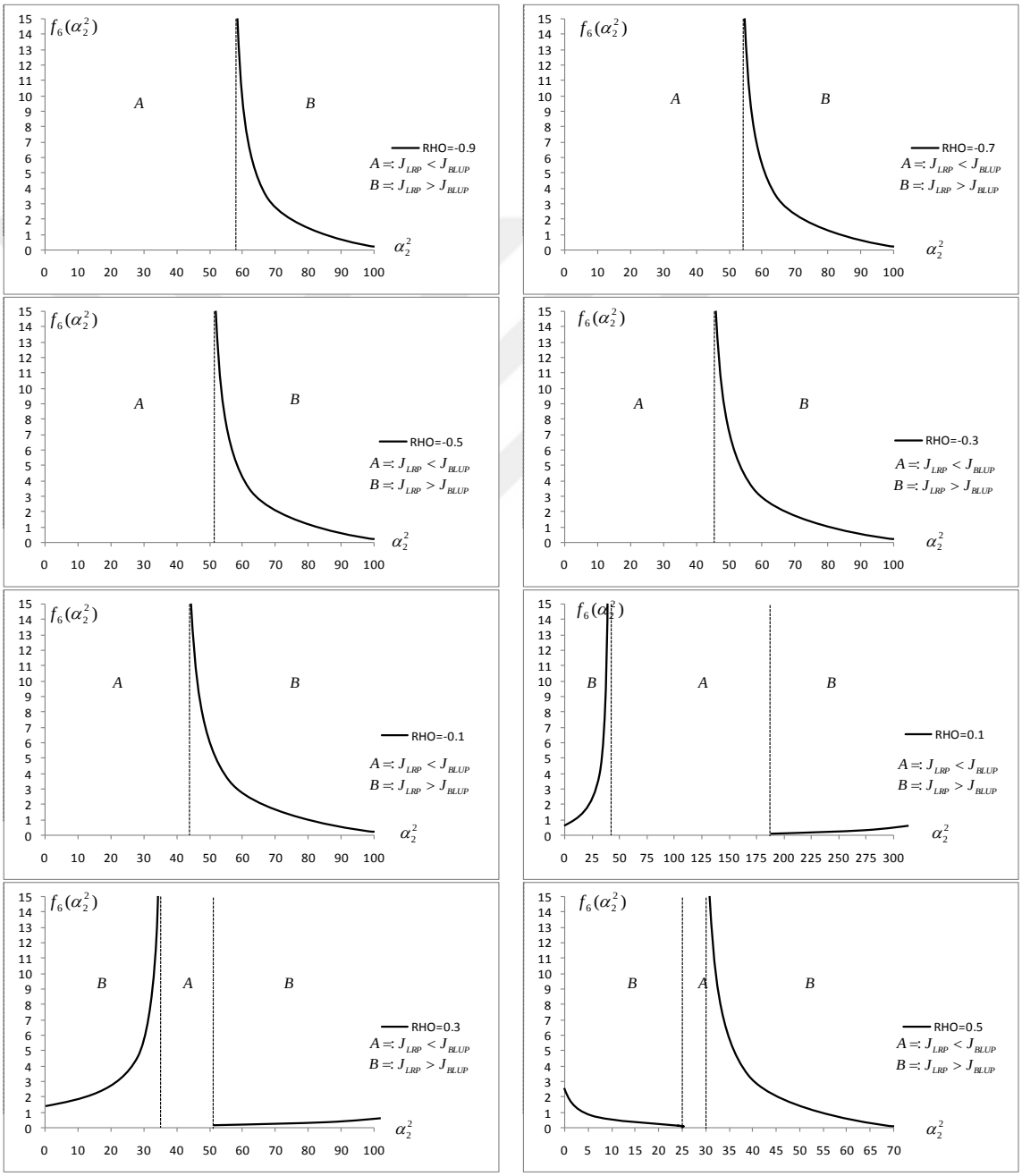


Figure 4.1. The pmse comparisons between the LRP and the BLUP with different RHOs

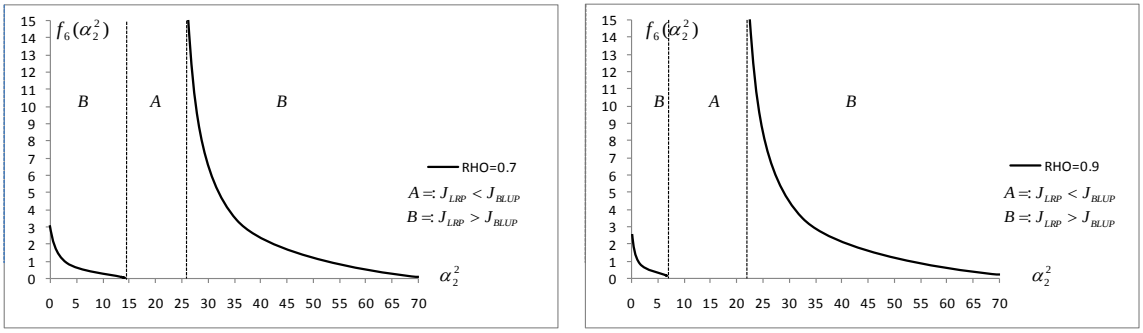


Figure 4.1. The pmse comparisons between the LRP and the BLUP with different RHOs (Continued)

-The effect of ρ on the predictive performance of the LRP estimator and the BLUP estimator is described in Table 4.1.

Table 4.1. ρ and α_2^2 values for the LRP vs. the BLUP

ρ	-0.9	-0.7	-0.5	-0.3	-0.1
α_2^2	(0, 58.5)	(0, 54.0)	(0, 50.3)	(0, 46.0)	(0, 42.2)
ρ	0.1	0.3	0.5	0.7	0.9
α_2^2	(37.9, 181.1)	(34.0, 50.8)	(25, 30.0)	(14, 26.0)	(6.9, 22)

Table 4.1. shows that if $\rho = -0.9, ..., -0.1$ gets larger as a value, the interval of α_2^2 gets smaller. That means, the subspace of observation to be predicted when the LRP is uniformly superior to the BLUP gets smaller. And the interval of α_2^2 when $\rho = 0.1, ..., 0.9$ is fluctuating from large to small and from small to large because it depends on the value of ρ.

Secondly, let us consider predictive performances of the LLP and the BLUP. From (4.45.), we have the function

$$f_7(\alpha_2^2) = \frac{0.023631\alpha_2^2 - 4.08314}{0.008264\alpha_2^2 - 1.73554} \tag{4.67.}$$

with a horizontal asymptote at $\alpha_2^2 = 172.786$ and a vertical asymptote at

$$\alpha_2^2 = 210 \tag{4.68.}$$

Figures 4.2. ($\rho = 0.5$) illustrates the regions for the superiority of the predictors. As it seems, in this example, the trade-off occurs from the value $\alpha_2^2 = 0$ till $\alpha_2^2 = 172.786$ and over the value $\alpha_2^2 = 210$. For smaller values of α_2^2, the LLP is uniformly superior to the BLUP. Again, for small values of the ratio z_{02}^2 / z_{01}^2, the trade-off is on the LLP's side.

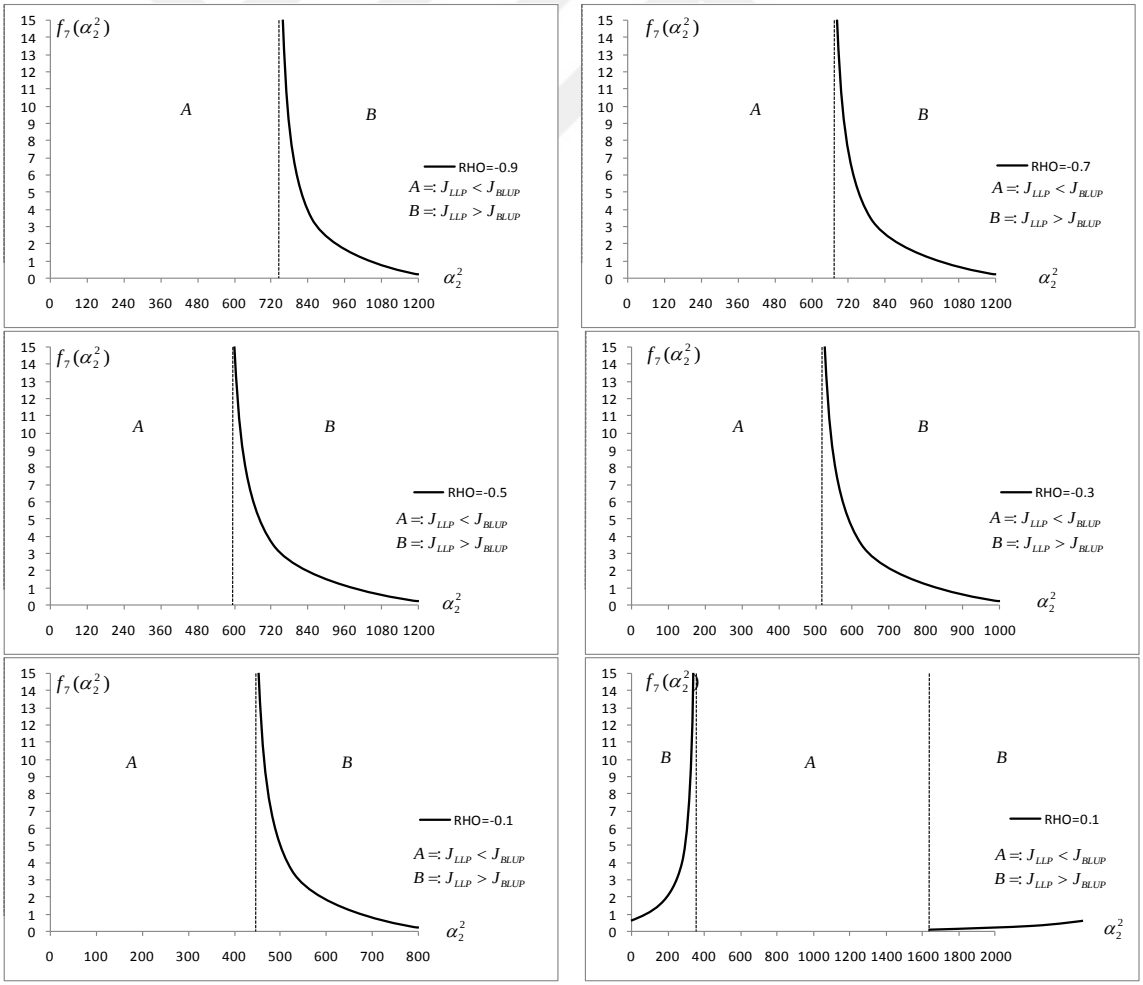


Figure 4.2. The pmse comparisons between the LLP and the BLUP with different RHOs

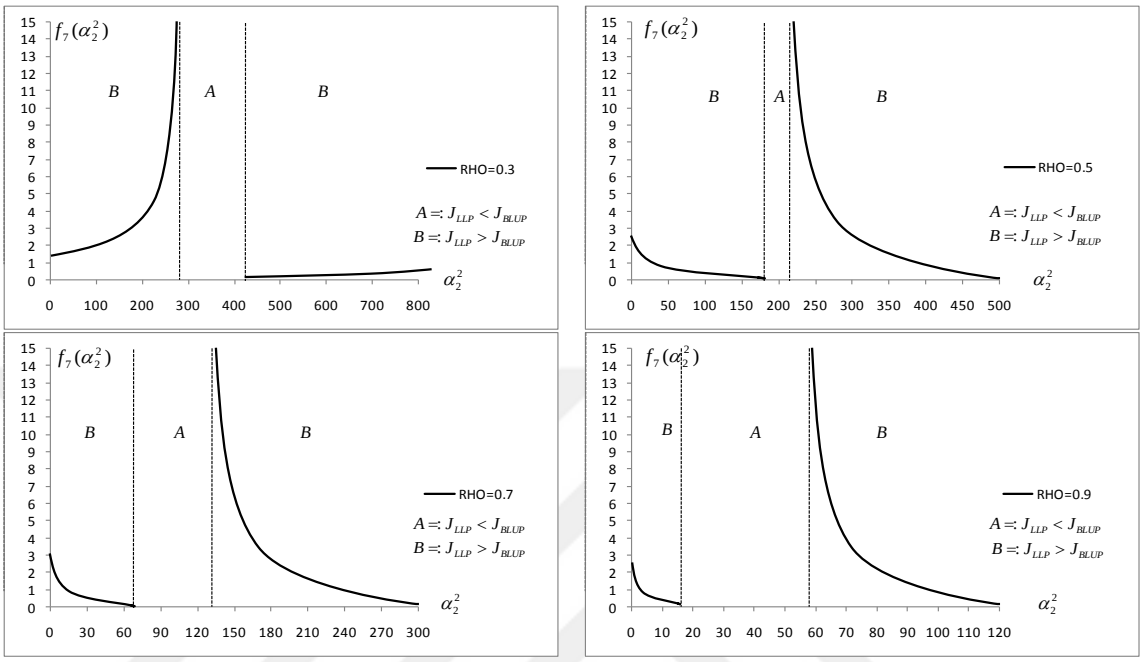


Figure 4.2. The pmse comparisons between the LLP and the BLUP with different RHOs (Continued)

-The effect of ρ on the predictive performance of the LRP estimator and the BLUP estimator is described in Table 4.2.

Table 4.2. ρ and α_2^2 values for the LLP vs. the BLUP

ρ	-0.9	-0.7	-0.5	-0.3	-0.1
α_2^2	(0, 750.8)	(0, 666.3)	(0, 595.8)	(0, 513.5)	(0, 442.2)
ρ	0.1	0.3	0.5	0.7	0.9
α_2^2	(359.3, 1687.2)	(286.1, 420.9)	(172.7, 210)	(71.8, 133.8)	(15.9, 58)

Table 4.2. shows that if $\rho = -0.9,...,-0.1$ gets larger as a value, the interval of α_2^2 gets smaller. That means, the subspace of observation to be predicted when the LLP is uniformly superior to the BLUP gets smaller. And the interval of α_2^2 when $\rho = 0.1,...,0.9$ is fluctuating from large to small and from small to large because it depends on the value of ρ.

Finally, let us consider the predictive performances of the LLP and the LRP. From (4.55.), we get

$$f_8(\alpha_2^2) = \frac{5.764463 - 0.24174\alpha_2^2}{0.023631\alpha_2^2 + 13.95969} \tag{4.69.}$$

which is a hyperbola with a vertical asymptote at

$$\alpha_2^2 = 23.84. \tag{4.70.}$$

Figures 4.3. ($\rho = 0.5$) illustrates the regions for the superiority of the predictors. As it seems, in this example, the trade-off occurs below the value $\alpha_2^2 = 23.84$ till $\alpha_2^2 = 0$. For smaller values of α_2^2, the LLP is uniformly superior to the LRP. Again, for small values of the ratio z_{02}^2 / z_{01}^2, the trade-off is on the LLP's side.

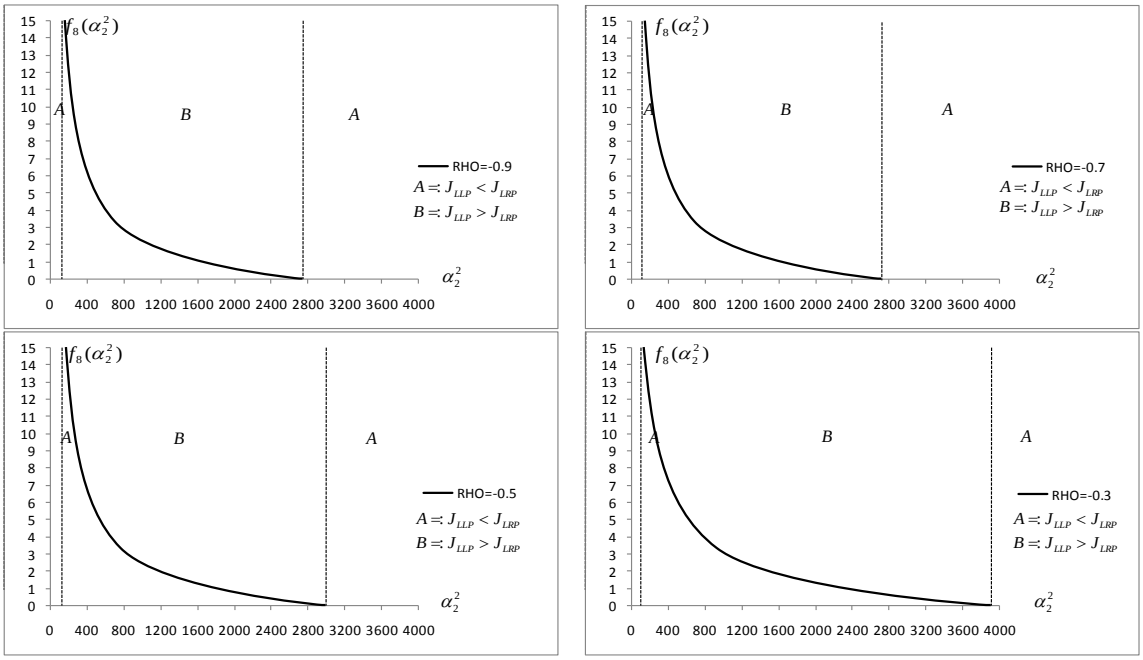


Figure 4.3. The pmse comparisons between the LLP and the LRP with different RHOs

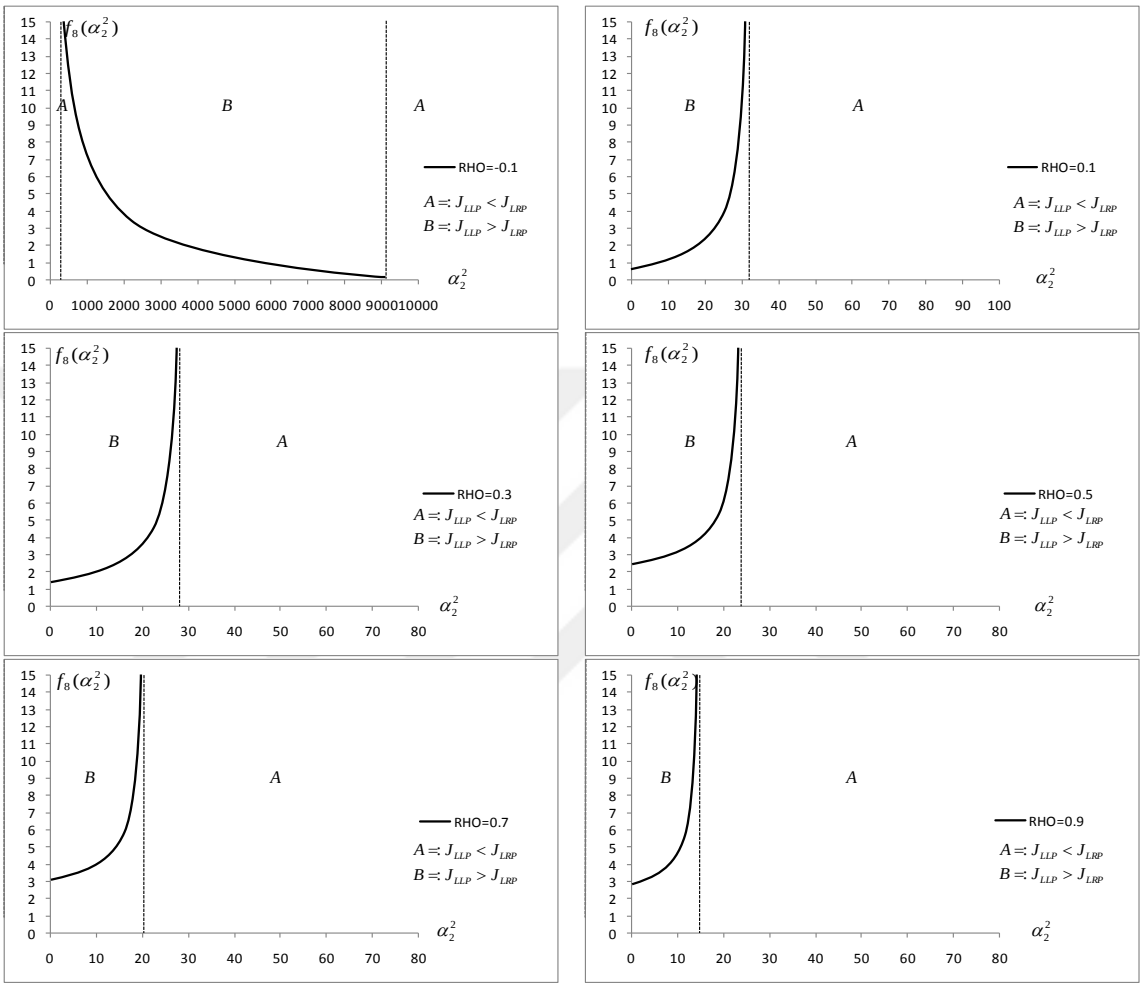


Figure 4.3. The pmse comparisons between the LLP and the LRP with different RHOs (Continued)

-The effect of ρ on the predictive performance of the LRP estimator and the BLUP estimator is described in Table 4.3.

Table 4.3. ρ and α_2^2 values for the LLP vs. the LRP

ρ	-0.9	-0.7	-0.5	-0.3	-0.1
α_2^2	(47.9, 2738.4)	(44.2, 2732.8)	(41.1, 3038.5)	(37.5, 3906)	(34.4, 9107.4)
ρ	0.1	0.3	0.5	0.7	0.9
α_2^2	(0, 30.7)	(0, 27.4)	(0, 23.8)	(0, 20.0)	(0, 15.1)

Table 4.3. shows that if $\rho = -0.9, ..., -0.1$ gets larger as a value, the interval of α_2^2 gets larger. That means, the subspace of observation to be predicted when the LLP is uniformly superior to the LRP gets larger. And if $\rho = 0.1, ..., 0.9$ gets larger as a value, the interval of α_2^2 gets smaller. That means, the subspace of observation to be predicted when the LLP is uniformly superior to the LRP gets smaller.

4.6. Conclusion

This chaper derived the LRP and the LLP predictors and investigates the predictive performance of the BLUP, the LRP and the LLP predictors. The pair-wise comparisons of these predictors are made in terms of the pmse criterion at a specific point in the two-dimensional regressor variable space.

The predictive performances of the LRP compared to the BLUP as well as the LLP compared to the BLUP and the LRP. In this context, the PMSE of the above mentioned predictors are developed and three theorems are given.

The theoretical consequences are illustrated by a numerical example, and the regions are assigned for the superiority of the proposed predictors. For some values of α_2^2, there are trade-offs between the relative effectiveness of the predictors. The effectiveness of these techniques is also affected by the location of the prediction point. Hence, the choice of the predictor may depend on the location of the point to be predicted and on the degree of multicollinearity, as well as on the degree of correlated errors (ρ). This implies that it is theoretically possible to determine such a region.

4. THE PREDICTIVE PERFORMANCE EVALUATION OF BIASED REGRESSION PREDICTORS WITH CORRELATED ERRORS Issam DAWOUD

5. CONCLUSIONS

The prediction problem of linear models belongs to the most important problems of the statistical inference of linear models. Since the linear regression model plays an important role in prediction, so we discuss some issues about it.

Firstly, although the OLS estimator is the BLUE of regression parameters in a linear regression model, we can improve upon the variability of an estimator or a predictor of a regression coefficient when the unbiasedness criterion can be skipped. So, the optimal heterogeneous, homogenous and homogenous-unbiased estimators of the MSE are introduced and discussed by many authors. Since the goodness of fitted model criterion is quite often ignored which is used to investigate the performance of estimators, EBLF is proposed in which the MSE, the ZBLF and the PLF are just special cases of it. Therfore, we derive the optimal heterogeneous, homogenous and homogenous unbiased estimators of the EBLF criterion and we make comparisons between the optimal heterogeneous and the optimal homogenous estimators as well as other comparisons are introduced by giving some corollaries under some conditions in terms of the EBLF risk. In addition, we make all pair-wise comparisons among the optimal heterogeneous, homogenous and homogenous-unbiased predictors in terms of the PMSE criterion.

Secondly, it is well-known that the predictive performance of a regression equation can be adversely affected by multicollinearity and correlated errors such that the presence of multicollinearity has serious effects on the OLS estimates of the regression coefficients. The popular methods in the presence of multicollinearity are the ORR and the Liu estimators. Under multicollinearity we extend the idea of Friedman and Montgomery (1985) and Özbey and Kaçıranlar (2015) after introducing two linear convex combination estimators of the operational optimal heterogeneous estimator; the LOR estimator compared to the OLS and the ORR estimators and the LOL estimator compared to the OLS, the Liu and the LOR estimators. The comparisons of these

estimators are in terms of the pmse criterion at a specific point in the two-dimensional regressor variable spaces. In this context, the pmse of the LOR and the LOL estimators are developed and five theorems are given. In addition, three corollaries are given here illustrating that the theorems given by Friedman and Montgomery (1985) and Özbey and Kaçıranlar (2015) are just special cases of the Theorems 3.1, 3.3 and 3.5. The theoretical consequences are illustrated by a numerical example and the regions are assigned for the superiority of the proposed estimators.

Finallly, in the general linear model, the GLS estimator is the BLUE. However, in view of the interdependence of the disturbances, the pattern of sample residuals contains information which may be used to modify the regression function in order to reduce the pmse. So, Goldberger (1962) derived the BLUP in the general linear regression model and found its pmse but when multicollinearity occurred among regressors, it is useful to use the biased regression estimators to find the biased predictors in order to reduce the pmse. This information exploited in the biased predictors derived. So, we extend this idea by introducing the LRP and the LLP predictors.. Then, the pmse of the above mentioned predictors are developed and the pair-wise comparisons of these predictors are made in terms of the pmse criterion at a specific point in the two-dimensional regressor variable space by investigating the predictive performances of the LRP compared to the BLUP and the LLP compared to the BLUP and the LRP in three theorems. Also, the theoretical consequences are illustrated by a numerical example, and the regions are assigned for the superiority of the proposed predictors.

REFERENCES

AITKEN, A. C., 1935. On least squares and linear combinations of observations. Proc R Soc. Edinb., 55: 42–48.

AKDENIZ, F. and KAÇIRANLAR, S., 1995. On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE. Commun. Statist. Theory Methods, 24: 1789-1797.

BIBBY, J. M. and TOUTENBURG, H., 1977. Prediction and Improved Estimation in Linear Models. Wiley, New York .

CHATURVEDI, A. and SHALABH, 2014. Bayesian Estimation of Regression Coefficients Under Extended Balanced Loss Function. Comm. Stat. Theory Methods, 43: 4253-4264.

DWIVEDI, T. D. and SRIVASTAVA, V. K., 1978. On the minimum mean squared error estimators in a regression model. Comm. Stat. Theory Methods, 7: 487-494.

FRIEDMAN, D. J. and MONTGOMERY, D. C., 1985. Evaluation of the predictive performance of biased regression estimators. J. Forecast., 4: 153-163.

GILES, J. A., GILES D. E. A. and OHTANI K., 1996. The exact risks of some pre-test and Stein-type regression estimates under balanced loss. Comm. Stat. Theory Methods, 25: 2901–2924.

GOLDBERGER, A. S., 1962. Best linear unbiased prediction in the generalized regression model. J. American Stat. Assoc., 57: 369–375.

GRIFFITHS, W. E., HILL, R. C. and JUDGE, G. C. 1993. Learning and practicing econometrics. Wiley, New York.

GRUBER, M. H. J., 2004. The efficiency of shrinkage estimators with respect to Zellner's balanced loss function. Comm. Stat. Theory Methods 33: 235–249.

_____, 2010. Regression Estimators: A Comparative Study, 2nd ed., John Hopkins University Press, Baltimore, M. D.

_____, 2012. Liu and Ridge Estimators- A comparison. Commun. Statist Theory Methods, 41: 3739-3749.

GÜLER, H. and KAÇIRANLAR, S., 2009. A Comparison of Mixed and Ridge Estimators of Linear Models. Communications in Statistics-Simulation and Computation, 38: 368-401.

HOERL, A. E. and KENNARD, R. W., 1970. Ridge regression: biased estimation for nonorthogonal problems. Technometrics, 12: 55–76.

KAÇIRANLAR, S., 2003. Liu estimator in the general linear regression model. J. Appl. Statist. Science, 13: 229-234.

KHALAF, G. and SHUKUR, G., 2005. Choosing ridge parameters for regression problems. Commun. Statist Theory Methods, 34: 1177-1182.

KIBRIA, B. M. G., 2003. Performance of some new ridge regression estimators. Commun. Statist. Simul. Comput., 32: 419-435.

LIU, K., 1993. A new class of biased estimate in linear regression. Commun. Statist. Theory Methods, 22: 393-402.

MCDONALD, G. C. and GALARNEAU, D. I., 1975. A monte carlo evaluation of some ridge-type estimators. J. American Stat. Assoc., 20: 407–416.

MUNIZ, G. and KIBRIA, B. M. G., 2009. On some ridge regression estimators: An empirical comparison. Commun. Statist. Simul. Comput., 38: 621-630

OHTANI, K., GILES, D. E. A. and GILES, J. A., 1997. The exact risk performance of a pre-test estimator in a heteroskedastic linear regression model under the balanced loss function, Econometric Rev., 16: 119–130.

_____, 1998. The exact risk of a weighted average estimator of the OLS and Stein-rule estimators in regression under balanced loss. Stat. Decis., 16: 35–45.

_____, 1999. Inadmissibility of the Stein-rule estimator under the balanced loss function. J. Econometrics, 88: 193–201.

ÖZBEY, F. and KAÇIRANLAR, S., 2015. Evaluation of the Predictive Performance of the Liu Estimator. Commun. Statist. Theory Methods, 44: 1981-1993.

ÖZKALE, M. R., 2013. Influence measures in affine combination type regression. Journal of Applied Statistics, 40: 2219-2243.

ÖZKALE, M. R. and KAÇIRANLAR, S., 2007. The Restricted and Unrestricted Two Parameter Estimators. Commun. Statist. Theory Methods, 36: 2707–2725.

RAO, C. R., 1971. Unified Theory of linear estimation, Sankhya A 33: 371-394.

_____, 1973. Linear Statistical Inference and Its Applications. 2nd edition, Wiley, New York.

RAO, C. R. and TOUTENBURG, H., 1995. Linear Models: Least Squares and Alternatives. Springer-Verlag, New York.

RODRIGUES, J. and ZELLNER, A., 1994. Weighted balanced loss function and estimation of the mean time to failure. Commun. Statist. Theory Methods, 23: 3609-3616.

SAKALLIOĞLU, S., KAÇIRANLAR, S. and AKDENIZ, F., 2001. Mean squared error comparisons of some biased regression estimator. Commun. Statist. Theory Methods, 30: 347-361.

SCHAFFRIN, B., 1985. A note on linear prediction within a Gauss–Markov model linearized with respect to a random approximation. in T. Pukkila and S. Puntanen (eds.), Proceedings of the First International Tampere Seminar on Linear Statistical Models and Their Applications 285–300, Tampere, Finland.

_____, 1986. New estimation/prediction techniques for the determination of crustal deformations in the presence of geophysical prior information. Technometrics, 130: 361–367.

_____, 1987. Less sensitive tests by introducing stochastic linear hypotheses. in T. Pukkila and S. Puntanen (eds.), Proceedings of the Second International Tampere Conference in Statistics, 647–664, Tampere, Finland.

SHALABH, 1995. Performance of Stein-rule procedure for simultaneous prediction of actual and average values of study variable in linear regression model. Jan Tinbergen

Award Paper, Bulletin of the 50th Session of the International Statistical Institute, 1375–1390.

SHALABH, TOUTENBURG, H. and HEUMANN, C., 2009. Stein-rule estimation under an extended balanced loss function. J. Stat. Comput. Simul., 79: 1259-1273.

THEIL, H., 1958. Economic Forecasts and Policy. North-Holland, Amsterdam.

_____, 1963. On The Use of Incomplete Prior Information in Regression Analysis. J. Am. Stat. Assoc. 58: 401–414.

_____, 1971. Principles of Econometrics. Wiley, New York.

THEIL, H. and GOLDBERGER, A. S., 1961. On Pure and Mixed Statistical Estimation in Economics. Int. Econ. Rev. 2: 65–78.

TOUTENBURG, H., 1968. Vorhersage im allgemeinen linearen Regressionsmodell mit Zusatzinformation ¨uber die Koeffizienten. Operationsforschung Mathematische Statistik, 1: 107–120.

_____, 1970a. Probleme linearer Vorhersagen im allgemeinen linearen Regressionsmodell. Biometrische Zeitschrift, 12: 242–252.

_____, 1970b. Uber die Wahl zwischen erwartungstreuen und nichterwartungstreuen Vorhersagen. Operationsforschung Mathematische Statistik, 2: 107–118.

_____, 1970c. Vorhersage im allgemeinen Regressionsmodell mit stochastischen Regressoren. Mathematische Operationsforschung und Statistik, 1: 105–116.

TOUTENBURG, H. and SHALABH, 2005. Estimation of regression coefficients subject to exact linear restrictions when some observations are missing and quadratic error balanced loss function is used. Test, 14: 385–396.

TOUTENBURG, H. and TRENKLER, G., 1990. Mean square error matrix comparisons of optimal and classical predictors and estimators in linear regression. Comput. Stat. Data Analysis, 10: 297–305.

TRENKLER, G., 1984. On the performance of biased estimators in the linear regression model with correlated or heteroscedastic errors. J. Econometrics, 25: 179-190.

ZELLNER, A., 1994. Bayesian and non-Bayesian estimation using balanced loss functions. Statistical Decision Theory and Related Topics, V: 377–390.

CURRICULUM VITAE

I was born in Palestine in 1987. I graduated from Abu-Obieda High School in 2005. Then, I graduated from the Department of Mathematics/Statistics (Bachelor Degree) at Islamic University in 2009. After that, I graduated from the Department of Statistics (Master degree) at Al-Azhar University in 2012. Now, I have been working as a lecturer in the Department of Statistics at Al-Aqsa University since 2010.