Yayın Geliş Tarihi (Submitted): **15/01/2024**

Yayın Kabul Tarihi (Accepted): **22/12/2024**

Makele Türü (PaperType): **Araştırma Makalesi – ResearchPaper**

PleaseCite As/Atıf için: Shaltoot, I. M., İnce, N., and Şentürk, S. (2024), Comparison of restricted and unrestricted estimators with a Monte-Carlo simulation study, *Nicel Bilimler Dergisi*, 6(2), 228-273. *doi:10.51541/nicel.1375616*

COMPARISON OF RESTRICTED AND UNRESTRICTED ESTIMATORS WITH A MONTE-CARLO SIMULATION STUDY

Israa Mahmoud SHALTOOT[1], Nihal İNCE[2]and Sevil ŞENTÜRK[3]

ABSTRACT

This study deals with the problem of multicollinearity in the linear regression model. Restricted and unrestricted parameter estimates are chosen among biased estimators to be studied and compared as two corresponding groups, with the aim of identifying which group gives better parameter estimates in the case of multicollinearity. Estimators' performance is compared according to matrix mean square error and scalar mean square error. Proceeding from this, it has been shown that, in the sense of Scalar Mean Square Error (SMSE), the Restricted Ridge regression (RRR) estimator outperforms all constrained and unconstrained estimators, while the Ridge regression is superior to the unconstrained set of estimators. A real-life application and Monte-Carlo simulation study are conducted to compare the performance of restricted and unrestricted estimators. As a result, it was decided that the most effective estimators are the restricted biased estimators when it comes to the state of multicollinearity.

Keywords: Multicollinearity, Mean square error, Biased estimators, Monte-Carlo simulation, linear restrictions.

[1] Department of Statistics, Faculty of Science, Eskişehir Technical University, Eskişehir, Türkiye. ORCID ID: https://orcid.org/0000-0002-5428-397X
[2]Corresponding Author, Dr., Department of Statistics, Faculty of Science, Eskişehir Technical University, Eskişehir, Türkiye. ORCID ID: https://orcid.org/0000-0001-6684-5848
[3]Prof.Dr., Department of Statistics, Faculty of Science, Eskişehir Technical University, Eskişehir, Türkiye. ORCID ID: https://orcid.org/0000-0002-9503-7388

MONTE-CARLO SİMÜLASYON ÇALIŞMASI İLE SINIRLI VE SINIRSIZ TAHMİN EDİCİLERİN KARŞILAŞTIRILMASI

ÖZ

Bu çalışma, doğrusal regresyon modelinde çoklu doğrusallık sorununu ele almaktadır. Çoklu doğrusal bağlantı durumunda hangi grubun daha iyi parametre tahminleri verdiğini belirlemek amacıyla, incelenecek ve karşılaştırılacak yanlı tahmin ediciler arasından kısıtlı ve kısıtlı olmayan parametre tahminleri iki grup olarak seçilmiştir. Tahmin edicilerin performansı matris ortalama kare hatası ve skaler ortalama kare hatasına göre karşılaştırılmıştır. Buradan hareketle, Skaler Ortalama Kare Hata (SMSE) anlamında, Kısıtlı Ridge regresyonu (RRR) tahmin edicisinin tüm kısıtlı ve kısıtlı olmayan tahmin edicilerden daha iyi performans gösterdiği, Ridge regresyonunun ise kısıtlı olmayan tahmin ediciler kümesinden daha üstün olduğu gösterilmiştir. Kısıtlı ve kısıtlı olmayan tahmin edicilerin performansını karşılaştırmak için gerçek hayattan bir uygulama ve Monte-Carlo simülasyon çalışması yapılmıştır. Sonuç olarak, çoklu doğrusal bağlantı durumu söz konusu olduğunda en etkin tahmin edicilerin kısıtlı yanlı tahmin ediciler olduğuna karar verilmiştir.

Anahtar Kelimeler: Çoklu doğrusallık, Ortalama karesel hata, Yanlı tahminciler, Monte-Carlo simülasyonu, doğrusal kısıtlamalar

1. INTRODUCTION

The linear regression model consisting of more than independent variable is referred to as multiple. The standard multiple linear regression model is expressed as

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{e}, \qquad (1)$$

where $\boldsymbol{y}$ is an $n \times 1$ observation vector of dependent variable, $\boldsymbol{X}$ is an $n \times p$ full column rank observation matrix of p independent variables, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown regression coefficients, and $\boldsymbol{e}$ is an $n \times 1$ vector of independent and identically distributed $(0, \sigma^2)$ random errors. Whereas n is the observations' number, p is the independent variables' number (Montgomery et al., 2021).

The ordinary least squares (OLS) estimate of model (1) is obtained as

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}, \tag{2}$$

where $\boldsymbol{X}'\boldsymbol{X}$ is a $p \times p$ correlation matrix between the independent variables and $\boldsymbol{X}'\boldsymbol{y}$ is a $p \times 1$ correlation vector between the independent variables and the dependent variable.

The OLS estimator $\widehat{\boldsymbol{\beta}}$ is an unbiased estimator that has the smallest variance among all other linear unbiased estimators. And according to the Gauss-Markov theorem the ordinary least squares (OLS) estimator is considered the best linear unbiased estimator (BLUE) for $\boldsymbol{\beta}$ parameter (Albert, 1973). Therefore, due to its suitable statistical properties it has been used for a long time. But, if the model's assumptions are not met, the parameters estimated by the OLS method may not be reliable. Although theoretically some results can be obtained under the assumption that the columns of the $\boldsymbol{X}$ matrix are linear independent, applications are generally encountered in cases where the columns of the design matrix are linearly dependent. In this case, the problem of multicollinearity arises, and the matrix $\boldsymbol{X}'\boldsymbol{X}$ becomes ill-conditioned. In datasets with multicollinearity, the OLS estimates of the parameters are known to be great in absolute values(Montgomery et al., 2012). To overcome these problems, many authors have introduced different kinds of one and two-parameter estimators: to mention a few, Stein (1956), Massy (1965), Hoerl and Kennard (1970), Swindel (1976), Liu (1993), Akdeniz and Kaçiranlar (1995), Ozkale and Kaçiranlar (2007), Sakallıoglu and Kaçıranlar (2008), Yang and Chang (2010), Dorugade (2014), Roozbeh (2018), Akdeniz and Roozbeh (2019), Lukman et al. (2019), Lukman et al. (2020), and, very recently, Kibria and Lukman (2020), among others.

The purpose of this study is to examine several different classes of biased estimators when the input matrix of the design is ill-conditioned. These estimators are the ridge regression estimator (1970), the contraction estimator (1973), the Liu estimator (1993), the two-parameter estimator (2007), and their restricted cases subject to the exact constraint Rβ=r. Then with the aim of conducting a distinct comparison from comparisons found profusely in the literature, the concerned estimators have been compared after separating them into two key groups one under the classic linear regression model the other under the restricted linear regression model. The comparison was applied once theoretically by using the criterion of the matrix mean square error (MMSE), and once practically by conducting a simulation study and a real-life application based on the discussed theories. While numerous studies and articles in the literature have examined biased estimators, no research has specifically compared the Linear regression model with the restricted regression model. This study is the first to provide a comparative analysis of certain unrestricted biased estimators and their corresponding restricted biased estimators in the presence of multicollinearity. Deserves to be mentioned that Månsson and Kibria (2021) compared between the restricted and unrestricted Liu estimators in the presence of multicollinearity but for the Poisson regression model.

This article is organized as follows. In Sec. 2, some restricted and unrestricted biased estimators were presented and discussed. Some theories that summarize the comparisons between the restricted and the unrestricted estimators under the MMSE criterion were shown in Sec. 3. A numerical example based on theoretical comparisons provided in the previous chapter were undertaken to evaluate the performance of estimators in respect of both the scalar mean square error (SMSE) and MMSE criteria is given in Sec. 4. To compare the performance of the estimators, a Monte-Carlo simulation study has been conducted in Sec. 5. All obtained results have been summarized in the conclusion in Sec. 6. The acronyms used in this study are listed in Table 1.

Table 1. A list of acronyms and definitions used throughout the paper

CN	: ConditionNumber
LE	: LiuEstimator
MRL	: Modified RestrictedLiu
MRR	: Modified RidgeRegression
MMSE	: MatrixMeanSquareError
SMSE	: ScalarMean SquareError

Table 1 (Continued). A list of acronyms and definitions used throughout the paper

nnd	: NonnegativeDefined
OLS	: OrdinaryLeastSquares
RL	: RestrictedLiu
RLM	: RestrictedLinear Model
RLS	: RestrictedLeastSquares
RR	: RidgeRegression
RRR	: RestrictedRidgeRegression
RTP	: Restricted Two-Parameter
SRTP	: StochasticRestrictedTwo-Parameter
TP	: Two-Parameter
VIF	: VarianceInflationFactor
RC	: RestrictedContraction
$\boldsymbol{S}$	: TheCorrelationMatrix ($\boldsymbol{X'X}$)
EMSE	: EstimatedMean Square Error
$E(\widetilde{\boldsymbol{\beta}})$	: TheExpected Value of $\widetilde{\boldsymbol{\beta}}$ Estimator
$Var(\widetilde{\boldsymbol{\beta}})$	: TheVariance-CovarianceMatrix of $\widetilde{\boldsymbol{\beta}}$ Estimator
$Bias(\widetilde{\boldsymbol{\beta}})$	: TheBias of $\widetilde{\boldsymbol{\beta}}$ Estimator
$\boldsymbol{\beta}$	: TheRegressionCoefficientsVector
$\widetilde{\boldsymbol{\beta}}$	: AnyEstimator of $\boldsymbol{\beta}$ Parameter
$\widehat{\boldsymbol{\beta}}$	: The OLS Estimator of $\boldsymbol{\beta}$ Parameter

Table 1 (Continued). A list of acronyms and definitions used throughout the paper

$\boldsymbol{\alpha}$	: TheRegressionCoefficients in theCanonical Form
$\tilde{\boldsymbol{\alpha}}$	: AnyEstimator of $\boldsymbol{\alpha}$ Parameter
$\hat{\boldsymbol{\alpha}}$	: The OLS Estimator in theCanonical Form
$\boldsymbol{Q}$	: An OrthogonalMatrix
$\boldsymbol{k}$	: RidgeBiasParameter
$\boldsymbol{d}$	: LiuBiasParameter
$\boldsymbol{\lambda_i}$	: TheEigenvalues of $\boldsymbol{X'X}$ Matrix
$\boldsymbol{\Lambda}$	: A DiagonalMatrixConsisting of theEigenvalues of $\boldsymbol{X'X}$ Matrix
$\hat{\boldsymbol{\beta}}(k)$	: RidgeRegressionEstimator
$\hat{\boldsymbol{\beta}}(k,b)$	: ModifiedRidgeRegressionEstimator
$\hat{\boldsymbol{\beta}}(d)$	: ContractionEstimator
$\hat{\boldsymbol{\beta}}_d$	: LiuEstimator
$\hat{\boldsymbol{\beta}}(k,d)$	: TwoParameterEstimator
$\hat{\boldsymbol{\beta}}_{\boldsymbol{r}}$	: RestrictedLeastSquareEstimator
$\hat{\boldsymbol{\beta}}_{\boldsymbol{r}}(k)$	: RestrictedRidgeRegressionEstimator
$\hat{\boldsymbol{\beta}}_{\boldsymbol{R}}(d)$	: RestrictedContractionEstimator
$\hat{\boldsymbol{\beta}}_{\boldsymbol{r}}(d)$	: RestrictedLiuEstimator
$\hat{\boldsymbol{\beta}}_{\boldsymbol{r}}(k,d)$	: RestrictedTwo-ParameterEstimator

Following a brief introduction and some preliminary discussions, we will now give the description of the linear regression model, linear estimators, and some standard criteria for evaluating the goodness of estimators.

2. SOME BIASED LINEAR ESTIMATORS

In statistics, an estimator's bias is the discrepancy between the estimated parameter's real value and its expected value. If an estimator is either underestimate or overestimate a population parameter, then it is considered to be biased. So, in this chapter, some biased estimators that are widely used in statistics and econometrics will be included and divided into two main groups. In addition, the statistical properties of these estimators will be presented.

One of the best-known approaches to circumvent plenty of the obstacles associated with the OLS estimates is the ridge regression (RR) estimator that was originally introduced by Hoerl and Kennard (1970) through the acceptance of some bias to reduce variance as follows:

$$\widehat{\boldsymbol{\beta}}_k = (\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}\boldsymbol{X}'\boldsymbol{y}\,, k \geq 0. \tag{3}$$

The expected value of $\widehat{\boldsymbol{\beta}}_k$ is

$$E(\widehat{\boldsymbol{\beta}}_k) = E[(\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}\boldsymbol{X}'\boldsymbol{y}] = \boldsymbol{\beta} - k\,\boldsymbol{S}_k{}^{-1}\boldsymbol{\beta}, \tag{4}$$

the part $(-\,k\,\boldsymbol{S}_k{}^{-1}\boldsymbol{\beta})$ of this equality expresses bias. So,

$$Bias(\widehat{\boldsymbol{\beta}}_k) = E(\widehat{\boldsymbol{\beta}}_k) - \boldsymbol{\beta} = -\,k\,\boldsymbol{S}_k{}^{-1}\boldsymbol{\beta}, \tag{5}$$

while the variance-covariance matrix is as follows:

$$Var(\widehat{\boldsymbol{\beta}}_{\boldsymbol{k}}) = Var[(\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}\boldsymbol{X}'\boldsymbol{y}] = \sigma^2\boldsymbol{S}_k{}^{-\boldsymbol{1}}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{S}_k{}^{-\boldsymbol{1}}. \tag{6}$$

By using (5) and (6) the MMSE matrix of ridge regression estimator is

$$\begin{aligned} MMSE(\widehat{\boldsymbol{\beta}}_k) &= Var(\widehat{\boldsymbol{\beta}}_k) + Bias(\widehat{\boldsymbol{\beta}}_{\boldsymbol{k}})Bias(\widehat{\boldsymbol{\beta}}_k)' \\ &= \boldsymbol{S}_k{}^{-1}(\sigma^2\boldsymbol{X}'\boldsymbol{X} + k^2\boldsymbol{\beta}\boldsymbol{\beta}')\boldsymbol{S}_k{}^{-1}. \end{aligned} \tag{7}$$

Now, let λ_1, λ_2,…, λ_p to be the eigenvalues of the $\boldsymbol{X}'\boldsymbol{X}$ matrix, then the SMSE equation of the RR estimator is

$$SMSE(\widehat{\boldsymbol{\beta}}_k) = tr[MMSE(\widehat{\boldsymbol{\beta}}_k)] = \sum_{j=1}^{p} \frac{\sigma^2\lambda_j + k^2\beta_j^2}{(\lambda_j + k)^2}. \tag{8}$$

In spite of ridge estimator efficiency in practice, k is a complex function. The bias parameter k has numerous selection methods and its preference depends on the analyzer. Since there is no consensus on the manner of choosing k, new estimation methods have been investigated in the case of linear model morbidity. One of those methods is the Liu estimator, by combining the ridge and the contraction estimators' advantages Kejian (1993) introduced a new estimator. This estimator was later named Liu estimator by Akdeniz and Kaçiranlar (1995) and M. H. J. Gruber (1998). The estimator of Liu was obtained by using the OLS estimate on the model caused by augmenting $d\hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + \boldsymbol{\varepsilon}'$ to (1). The advantage of $\hat{\boldsymbol{\beta}}_d$ over $\hat{\boldsymbol{\beta}}_k$ is because d is a linear function, so the selection of d is easier than the selection of k. $\hat{\boldsymbol{\beta}}_d$also can be written as a linear transformation of the OLS estimator in the following form:

$$\hat{\boldsymbol{\beta}}_d = (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{X} + d\boldsymbol{I})\,\hat{\boldsymbol{\beta}}. \tag{9}$$

From this, we can see that when $d = 1$, $\hat{\boldsymbol{\beta}}_d = \hat{\boldsymbol{\beta}}$ and $\|\hat{\boldsymbol{\beta}}_d\| < \|\hat{\boldsymbol{\beta}}\|$. Also, the expected value of $\hat{\boldsymbol{\beta}}_d$

$$E(\hat{\boldsymbol{\beta}}_d) = (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{X} + d\boldsymbol{I})\boldsymbol{\beta}\,, \qquad ((10)$$

demonstrates that $\hat{\boldsymbol{\beta}}_d$ is a biased estimator. Furthermore, with supposing that $\boldsymbol{A}_d = (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{X} + d\boldsymbol{I})$, the bias and the variance-covariance matrix of the Liu estimator are respectively.

$$Bias(\hat{\boldsymbol{\beta}}_d) = (\boldsymbol{A}_d - 1)\boldsymbol{\beta}, \tag{11}$$

$$Var(\hat{\boldsymbol{\beta}}_d) = \sigma^2 \boldsymbol{A}_d(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{A}_d'. \tag{12}$$

By using the equations (11) and (12), Liu estimator's MMSE matrix is as follows:

$$MMSE(\hat{\boldsymbol{\beta}}_d) = \sigma^2 \boldsymbol{A}_d(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{A}_d' + (\boldsymbol{A}_d - 1)\boldsymbol{\beta}\boldsymbol{\beta}'(\boldsymbol{A}_d - 1). \tag{13}$$

As well as the SMSE value is

$$SMSE(\hat{\boldsymbol{\beta}}_d) = \sigma^2 \sum_{i=1}^{p} \frac{(\lambda_i + d)^2}{\lambda_i(\lambda_i + 1)} + (d-1)^2 \sum_{i=1}^{p} \frac{\alpha_i^2}{(\lambda_i + 1)^2}\,, \tag{14}$$

where α_i is the i-th element of $\boldsymbol{\alpha} = \boldsymbol{T}'\boldsymbol{\beta}$ vector used to write the model (1) in the canonical form, given the orthogonal matrix $\boldsymbol{T}$ of the orthonormal eigenvectors of the $\boldsymbol{X}'\boldsymbol{X}$ matrix. The equation of the bias parameter that makes the $SMSE(\hat{\boldsymbol{\beta}}_d)$ minimum is given by Liu (1993) as follows:

$$d_{opt} = \sum_{i=1}^{p} \frac{\alpha_i^2 - \sigma^2}{(\lambda_i + 1)^2} \Big/ \sum_{i=1}^{p} \frac{\sigma^2 + \lambda_i \alpha_i^2}{\lambda_i (\lambda_i + 1)^2} = 1 - \frac{\sigma^2 \left(\sum_{i=1}^{p} \frac{1}{\lambda_i(\lambda_i+1)} \right)}{\sum_{i=1}^{p} \frac{\sigma^2 + \lambda_i \alpha_i^2}{\lambda_i (\lambda_i+1)^2}}. \tag{15}$$

However, d_{opt} is not very handy in practice since it depends on the unknown parameters σ^2 and α_i^2. If the unbiased estimators $\hat{\alpha}_i^2 - \hat{\sigma}^2/\lambda_i$ and $\hat{\sigma}^2$ are used instead, then the following $\hat{d}_{opt}$ will be obtained as follows:

$$\hat{d}_{opt} = 1 - \hat{\sigma}^2 \left(\sum_{i=1}^{p} \frac{1}{\lambda_i(\lambda_i + 1)} \Big/ \sum_{i=1}^{p} \frac{\hat{\alpha}_i^2}{(\lambda_i + 1)^2} \right). \tag{16}$$

Liu (1993) named (16) the minimum MMSE estimate and gave a generalized form of it. In addition, different selection methods of d were conducted.

With the aim of reaching new estimators that are more balanced and less biased, statisticians have started a new trend towards finding new estimation methods with more than one parameter to find new estimators encountering multicollinearity whose length is closer to $\boldsymbol{\beta}$ than $\widehat{\boldsymbol{\beta}}$. For instance, Özkale and Kaciranlar (2007) introduced the following two-parameter estimator by grafting the contraction estimator into the modified ridge regression (MRR) estimator:

$$\widehat{\boldsymbol{\beta}}(k,d) = (\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{y} + kd\widehat{\boldsymbol{\beta}}), \quad k > 0, 0 < d < 1. \tag{17}$$

Another two parameter estimator was obtained by Yang and Chang (2010)

$$\widehat{\boldsymbol{\beta}}^*(k,d) = (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{X} + d\boldsymbol{I})(\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}\boldsymbol{X}'\boldsymbol{y}, \quad k > 0, 0 < d < 1. \tag{18}$$

Many other two type parameter estimators were suggested separately in the literature, see (Sakallıoğlu and Kaçıranlar, 2008;Dorugade, 2014 and Idowu et al, 2023). Both two-parameter (TP) estimators in (17) and (18) can be obtained as a solution of the minimization problems, or similar to that of the Liu estimator it can be obtained by augmenting the linear stochastic constraints $kd\widehat{\boldsymbol{\beta}} = k\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ and $(d-k)\widehat{\boldsymbol{\beta}}(k) = \boldsymbol{\beta} + \boldsymbol{\varepsilon}'$to (1) model one for (17) and the other for (18) respectively, and then using the least squares estimator. Here $\boldsymbol{\varepsilon}: p \times 1$is a random vector with a mean of zero, variance-covariance matrix of $\sigma^2\boldsymbol{I}$ and both random errors $\boldsymbol{e}$ and $\boldsymbol{\varepsilon}$ are uncorrelated i.e., $E(\boldsymbol{e}\boldsymbol{\varepsilon}') = 0$.

From this section on, the name two parameter estimator will refer to the two-parameter estimator $\widehat{\boldsymbol{\beta}}(k,d)$ of Özkale and Kaciranlar in (17) that will be discussed throughout the study. The TP estimator unifies the advantages of ridge and contraction estimators, also considers a general estimator that contains the ordinary least squares, the ridge regression, the Liu, and the contraction estimators as special cases. This feature can be shown by choosing different values for k and d in $\widehat{\boldsymbol{\beta}}(k,d)$, as in the following equations:

- $\lim_{d\to 1} \widehat{\boldsymbol{\beta}}(k,d) = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$which is the OLS estimator.
- $\lim_{k\to 0} \widehat{\boldsymbol{\beta}}(k,d) = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$ which is the OLS estimator.
- $\lim_{d\to 0} \widehat{\boldsymbol{\beta}}(k,d) = (\boldsymbol{X}'\boldsymbol{X}+k\boldsymbol{I})^{-1}\boldsymbol{X}'\boldsymbol{y}$ is the RR estimator.
- $\widehat{\boldsymbol{\beta}}(1,d) = (\boldsymbol{X}'\boldsymbol{X}+\boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{y}+d\widehat{\boldsymbol{\beta}})$ is the Liu estimator.
- $\widehat{\boldsymbol{\beta}}(k,d) = (\boldsymbol{X}'\boldsymbol{X}+k\boldsymbol{I})^{-1}\left(\boldsymbol{X}'\boldsymbol{y}+kd\widehat{\boldsymbol{\beta}}\right)$

$$= [\boldsymbol{I}+k(\boldsymbol{X}'\boldsymbol{X})^{-1}]^{-1}(\boldsymbol{X}'\boldsymbol{X})^{-1}\left(\boldsymbol{X}'\boldsymbol{y}+kd\widehat{\boldsymbol{\beta}}\right)$$
$$= [\boldsymbol{I}+k(\boldsymbol{X}'\boldsymbol{X})^{-1}]^{-1}\left(\widehat{\boldsymbol{\beta}}+kd(\boldsymbol{X}'\boldsymbol{X})^{-1}\widehat{\boldsymbol{\beta}}\right)$$
$$= [\boldsymbol{I}+k(\boldsymbol{X}'\boldsymbol{X})^{-1}]^{-1}\left(\widehat{\boldsymbol{\beta}}-d\widehat{\boldsymbol{\beta}}\right)+d\widehat{\boldsymbol{\beta}}.$$

After putting the $\widehat{\boldsymbol{\beta}}(k,d)$ on the previous form it can be said that $\lim_{k\to\infty} \widehat{\boldsymbol{\beta}}(k,d) = d\widehat{\boldsymbol{\beta}}$ is the contraction estimator

$$\widehat{\boldsymbol{\beta}}(k,d) = (\boldsymbol{X}'\boldsymbol{X}+k\boldsymbol{I})^{-1}\left(\boldsymbol{X}'\boldsymbol{y}+kd\widehat{\boldsymbol{\beta}}\right)$$
$$= (\boldsymbol{X}'\boldsymbol{X}+k\boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{y}+kd\boldsymbol{I})\widehat{\boldsymbol{\beta}}.$$

We see that the parameters of shrinkage related to each $\widehat{\boldsymbol{\beta}}$ differ. Therefore, the contraction estimator's drawback was overcome.

For the representation, $\widehat{\boldsymbol{\beta}}(k,d) = \boldsymbol{S}_k^{-1}\boldsymbol{S}_{kd}^{-1}\widehat{\boldsymbol{\beta}}$, where $\boldsymbol{S}_k = \boldsymbol{S}+k\boldsymbol{I}$, $\boldsymbol{S}_{kd} = \boldsymbol{S}+kd\boldsymbol{I}$ and $\boldsymbol{S} = \boldsymbol{X}'\boldsymbol{X}$ the biased and dispersion matrix respectively are as follows:

$$Bais\left(\widehat{\boldsymbol{\beta}}(k,d)\right) = (\boldsymbol{S}_k^{-1}\boldsymbol{S}_{kd}-\boldsymbol{I})\boldsymbol{\beta} = k(d-1)\boldsymbol{S}_k^{-1}\boldsymbol{\beta}. \tag{19}$$

$$Var\left(\widehat{\boldsymbol{\beta}}(k,d)\right) = \sigma^2\boldsymbol{S}_k^{-1}\boldsymbol{S}_{kd}\boldsymbol{S}^{-1}\boldsymbol{S}_{kd}\boldsymbol{S}_k^{-1}. \tag{20}$$

By using (19) and (20) equations the MMSE for the TP estimator is obtained

$$MMSE\left(\widehat{\boldsymbol{\beta}}(k,d)\right) = \sigma^2 \boldsymbol{S}_k^{-1}\boldsymbol{S}_{kd}\boldsymbol{S}^{-1}\boldsymbol{S}_{kd}\boldsymbol{S}_k^{-1} + k^2(d-1)^2\boldsymbol{S}_k^{-1}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{S}_k^{-1}. \quad (21)$$

Özkale and Kaciranlar (2007) presented the necessary and sufficient conditions for the two-parameter estimator to be superior to the OLS estimator according to the MMSE criterion. They also introduced a selection method for k and d bias parameters. Besides M. H. J. Gruber (2010) noticed that the two-parameter estimator $\hat{\beta}(k,d)$is a special case of the Bayes, the minimax, generalized ridge regression estimators, and the mixed estimators.

In some cases, the classical linear regression model given in (1) is estimated subject to certain prior limitations on the model's unknown parameter vector $\boldsymbol{\beta}$ in the hope to provide better estimators than the OLS estimator. These prior limitations may be stated in the form of the linear equality restrictions

$$\boldsymbol{R\beta} = \boldsymbol{r}, \quad (22)$$

where $\boldsymbol{R}$ is an$m \times p$ matrix of known prior information, its $rank(\boldsymbol{R}) = m < p$, $\boldsymbol{r}$ is an $m \times 1$ vector, and the linear constraints "m" involved in (1) model are independent of each other. In this research, the restricted linear model (RLM) denotes incorporation between (1) and (22).

The restricted least square (RLS) estimator is obtained

$$\widehat{\boldsymbol{\beta}}_r = \widehat{\boldsymbol{\beta}} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{r} - \boldsymbol{R}\widehat{\boldsymbol{\beta}}). \quad (23)$$

This approach generates an unbiased estimator when the constraints are correct, and contributes to a smaller variance in sampling than the OLS estimator. If the constraints are incorrect, then the sampling variance remains reduced, but the estimator is a biased one. The drawback of using exact restrictions is that the estimator will have more risk than the OLS estimator because of this potential bias (Güler and Kaçıranlar, 2009).

The expected value of the RLS

$$E(\widehat{\boldsymbol{\beta}}_r) = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{r} - \boldsymbol{R}\boldsymbol{\beta}), \quad (24)$$

this demonstrates that if the constraint is true then $\boldsymbol{\beta}r$is unbiased, otherwise it is biased. Also, the variance-covariance matrix is

$$Var\left(\widehat{\boldsymbol{\beta}}_r\right) = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1} - \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}. \quad (25)$$

The MMSE of the restricted least squares estimator is

$$MMSE(\hat{\boldsymbol{\beta}}_r) = \sigma^2 \boldsymbol{M}_0 \boldsymbol{X}'\boldsymbol{X}\boldsymbol{M}_0 = \sigma^2 \boldsymbol{S}^{-1} - \sigma^2 \boldsymbol{S}^{-1}\boldsymbol{R}'[\boldsymbol{R}\boldsymbol{S}^{-1}\boldsymbol{R}']^{-1}\boldsymbol{R}\boldsymbol{S}^{-1}, \quad (26)$$

where $\boldsymbol{M}_0 = \boldsymbol{S}^{-1} - \boldsymbol{S}^{-1}\boldsymbol{R}'[\boldsymbol{R}\boldsymbol{S}^{-1}\boldsymbol{R}']^{-1}\boldsymbol{R}\boldsymbol{S}^{-1}$, $\boldsymbol{S} = \boldsymbol{X}'\boldsymbol{X}$.

When it comes to the restricted biased estimators, Zhong and Yang (2007) showed that the restricted ridge regression estimator can be obtained by minimizing the sum of squared residuals with a spherical restriction and a linear restriction (22). The restricted ridge regression (RRR) estimator may be introduced as follows:

$$\hat{\boldsymbol{\beta}}_r(k) = \hat{\boldsymbol{\beta}}(k) + \boldsymbol{S}_k^{-1}\boldsymbol{R}'[\boldsymbol{R}\boldsymbol{S}_k^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{r} - \boldsymbol{R}\hat{\boldsymbol{\beta}}(k)), k \geq 0, \quad (27)$$

where $\hat{\boldsymbol{\beta}}(k)$ is the RR estimator of $\boldsymbol{\beta}$. We refer to (27) as the RRR estimator. If linear restrictions $\boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{r}$ are assumed to hold, then it is more appropriate to shrink towards $\boldsymbol{\beta}_0$ the shortest vector which satisfies the restrictions, rather than towards the zero vector. Groß (2003) proved that $\hat{\boldsymbol{\beta}}_r(0) = \hat{\boldsymbol{\beta}}_r$ and $\lim_{k\to\infty} \hat{\boldsymbol{\beta}}_r(k) = \boldsymbol{\beta}_0$. The expected value of $\hat{\boldsymbol{\beta}}_r(k)$ is

$$E(\hat{\boldsymbol{\beta}}_r(k)) = \boldsymbol{\beta} - k\boldsymbol{M}_k\boldsymbol{\beta}. \quad (28)$$

Also, the bias of $\hat{\boldsymbol{\beta}}_r(k)$

$$Bias(\hat{\boldsymbol{\beta}}_r(k)) = E(\hat{\boldsymbol{\beta}}_r(k)) - \boldsymbol{\beta} = -k\boldsymbol{M}_k\boldsymbol{\beta}, \quad (29)$$

where $\boldsymbol{M}_k$ is a non-zero symmetric matrix, equals to $\boldsymbol{S}_k^{-1} - \boldsymbol{S}_k^{-1}\boldsymbol{R}'[\boldsymbol{R}\boldsymbol{S}_k^{-1}\boldsymbol{R}']^{-1}\boldsymbol{R}\boldsymbol{S}_k^{-1}$, and the variance-covariance matrix is

$$Var(\hat{\boldsymbol{\beta}}_r(k)) = E[(\hat{\boldsymbol{\beta}}_r(k) - E(\hat{\boldsymbol{\beta}}_r(k)))(\hat{\boldsymbol{\beta}}_r(k) - E(\hat{\boldsymbol{\beta}}_r(k)))'] = \sigma^2 \boldsymbol{M}_k \boldsymbol{X}'\boldsymbol{X}\boldsymbol{M} \quad (30)$$

For the proofs of (28), (29) and (30) see (Zhong and Yang, 2007). By making use of the two previous equations the MMSE of $\hat{\boldsymbol{\beta}}_r(k)$ is obtained as

$$MMSE\left(\hat{\boldsymbol{\beta}}_r(k)\right) = \sigma^2 \boldsymbol{M}_k \boldsymbol{X}'\boldsymbol{X}\boldsymbol{M}_k + k^2 \boldsymbol{M}_k \boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{M}_k, \quad (31)$$

where $\boldsymbol{M}_k = \boldsymbol{S}_k^{-1} - \boldsymbol{S}_k^{-1}\boldsymbol{R}'[\boldsymbol{R}\boldsymbol{S}_k^{-1}\boldsymbol{R}']^{-1}\boldsymbol{R}\boldsymbol{S}_k^{-1}$. The mean square error of $\hat{\boldsymbol{\beta}}_r(k)$ is

$$SMSE\left(\widehat{\boldsymbol{\beta}}_r(k)\right) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i(\lambda_i + k - r_{ii}^*)^2}{(\lambda_i + k)^4} + k^2 \left[\sum_{i=1}^{p} \frac{\alpha_i(\lambda_i + k - r_{ii}^*)}{(\lambda_i + k)^2}\right]^2. \tag{32}$$

Although this estimator has fewer MMSE compared to the restricted least squares (RLS) estimator, it is very complex in practical implementation and a compacter estimator is supposed to be obtained. Therefore, there had remained a need to find a better estimator to overcome the multicollinearity of the restricted linear model (RLM). When the RRR estimator combines the two estimators MR proposed by Swindel (1976) and the RLS estimator; the second restricted liu(RL) estimator combines the modified Liu estimator with the RLS estimator

$$\widehat{\boldsymbol{\beta}}_r(d) = \widehat{\boldsymbol{\beta}}_d + (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I})^{-1}\boldsymbol{R}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{r} - \boldsymbol{R}\widehat{\boldsymbol{\beta}}_d), \tag{33}$$

where $\widehat{\boldsymbol{\beta}}_d$ is the Liu estimator. In order to avoid overlapping of names of the two restricted Liu estimators, Özkale and Kaciranlar suggested calling $\widehat{\boldsymbol{\beta}}_{\boldsymbol{r}}(d)$ as modified restricted Liu (MRL) estimator, see (Özkale and Kaciranlar, 2007). If the bias of $\widehat{\boldsymbol{\beta}}_r(d)$ is given as follows:

$$Bias(\widehat{\boldsymbol{\beta}}_r(d)) = (d-1)\boldsymbol{M}_1\boldsymbol{\beta}, \tag{34}$$

where $\boldsymbol{M}_1 = \boldsymbol{S}_1^{-1} - \boldsymbol{S}_1^{-1}\boldsymbol{R}'[\boldsymbol{R}\boldsymbol{S}_1^{-1}\boldsymbol{R}']^{-1}\boldsymbol{R}\boldsymbol{S}_1^{-1}$ and $\boldsymbol{S}_1 = \boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I}$. The variance-covariance matrix is

$$Var\left(\widehat{\boldsymbol{\beta}}_r(d)\right) = \sigma^2\boldsymbol{M}_1\boldsymbol{S}_d\boldsymbol{S}^{-1}\boldsymbol{S}_d\boldsymbol{M}_1. \tag{35}$$

Then the matrix mean square error of the modified restricted Liu is obtained as

$$MMSE\left(\widehat{\boldsymbol{\beta}}_r(d)\right) = \sigma^2\boldsymbol{M}_1\boldsymbol{S}_d\boldsymbol{S}^{-1}\boldsymbol{S}_d\boldsymbol{M}_1 + (d-1)^2\boldsymbol{M}_1\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{M}_1. \tag{36}$$

Under the same logic they had used obtaining the two-parameter (TP) estimator before Özkale and Kaçiranlar (2007) derived the restricted two-parameter (RTP) estimator in order to find an improved estimator that minimizes the squared distance of $\widehat{\boldsymbol{\beta}}$ toward the regression parameter. The new estimator was derived as a solution to minimization problems with the addition of additional restrictions on the parameter, thus the resulting estimator is as follows:

$$\widehat{\boldsymbol{\beta}}_r(k,d) = \widehat{\boldsymbol{\beta}}(k,d) + \boldsymbol{S}_k^{-1}\boldsymbol{R}'[\boldsymbol{R}\boldsymbol{S}_k^{-1}\boldsymbol{R}']^{-1}\left(\boldsymbol{r} - \boldsymbol{R}\widehat{\boldsymbol{\beta}}(k\,,d)\right), k > 0, 0 < d < 1. \tag{37}$$

As the TP estimator considers a general case of the OLS, the RR, the Liu, and the contraction estimator, the restricted version of it, that is the RTP estimator considers a general case of the RLS, the RRR introduced by GroB (2003), the MRL mentioned by Özkale and Kaciranlar (2007), and the restricted contraction (RC) estimator. The following equations illustrate some relationships expressed as

- $\lim_{k\to 0} \widehat{\boldsymbol{\beta}}_r(k,d) = \widehat{\boldsymbol{\beta}} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{R}']^{-1}(\boldsymbol{r} - \boldsymbol{R}\widehat{\boldsymbol{\beta}})$ is the RLS estimator.
- $\lim_{d\to 0} \widehat{\boldsymbol{\beta}}_r(k,d) = \widehat{\boldsymbol{\beta}}(k) + \boldsymbol{S}_k^{-1}\boldsymbol{R}'[\boldsymbol{R}\boldsymbol{S}_k^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{r} - \boldsymbol{R}\widehat{\boldsymbol{\beta}}(k))$ is the RRR estimator.
- $\lim_{k\to \infty} \widehat{\boldsymbol{\beta}}_r(k,d) = \widehat{\boldsymbol{\beta}}_R(d) = d\widehat{\boldsymbol{\beta}} + \boldsymbol{R}'[\boldsymbol{R}\boldsymbol{R}']^{-1}(\boldsymbol{r} - \boldsymbol{R}d\widehat{\boldsymbol{\beta}})$ is the RC estimator.
- For $k = 1$,
- $\widehat{\boldsymbol{\beta}}_r(1,d) = \widehat{\boldsymbol{\beta}}_r(d) = \widehat{\boldsymbol{\beta}}_d + (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I})^{-1}\boldsymbol{R}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{r} - \boldsymbol{R}\widehat{\boldsymbol{\beta}}_{\boldsymbol{d}})$ is the MRL estimator.

By putting the RTP estimator given in (37) on another form

$$\widehat{\boldsymbol{\beta}}_r(k,d) = \boldsymbol{M}(k)\boldsymbol{S}_{kd}\widehat{\boldsymbol{\beta}} + \boldsymbol{S}_k^{-1}\boldsymbol{R}'[\boldsymbol{R}\boldsymbol{S}_k^{-1}\boldsymbol{R}']^{-1}\boldsymbol{r}, \tag{38}$$

where $\boldsymbol{M}_k = \boldsymbol{S}_k^{-1} - \boldsymbol{S}_k^{-1}\boldsymbol{R}'[\boldsymbol{R}\boldsymbol{S}_k^{-1}\boldsymbol{R}']^{-1}\boldsymbol{R}\boldsymbol{S}_k^{-1}$, then the variance of $\hat{\beta}_r(k,d)$ is

$$Var\left(\widehat{\boldsymbol{\beta}}_r(k,d)\right) = \sigma^2\boldsymbol{M}_k\boldsymbol{S}_{kd}\boldsymbol{S}^{-1}\boldsymbol{S}_{kd}\boldsymbol{M}_k. \tag{39}$$

And the bias is

$$Bias\left(\widehat{\boldsymbol{\beta}}_r(k,d)\right) = \boldsymbol{M}_k[\boldsymbol{S}_{kd} - \boldsymbol{S}_k]\boldsymbol{\beta} = k(d-1)\boldsymbol{M}_k\boldsymbol{\beta}. \tag{40}$$

Then the MMSE of $\widehat{\boldsymbol{\beta}}_r(k,d)$, when the condition $\boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{r}$ is presumed to hold true, is

$$MMSE\left(\widehat{\boldsymbol{\beta}}_r(k,d)\right) = \sigma^2\boldsymbol{M}_k\boldsymbol{S}_{kd}\boldsymbol{S}^{-1}\boldsymbol{S}_{kd}\boldsymbol{M}_k + k^2(d-1)^2\boldsymbol{M}_k\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{M}_k. \tag{41}$$

Following our discussion of the preliminary steps, we now present a deep explanation of the problem of multicollinearity includes its definition, reasons of occurrence, effects on the linear model, and some ways to deal with and overcome it were presented.

3. COMPARISONS ACCORDING TO MMSE CRITERION

There are many criteria to compare two estimators of an unknown parameter. Some of these criteria have been mentioned previously in chapter two. In this research, the author believes that the most appropriate criterion for use is the MMSE, which is a commonly accepted criterion for gauging an estimator's performance of $\boldsymbol{\beta}$ because it contains all relevant estimator quality information; also it includes the comparison in terms of bias and in terms of dispersion matrix. The MMSE of an estimator $\widetilde{\boldsymbol{\beta}}$ is defined as

$$\begin{aligned} MMSE(\widetilde{\boldsymbol{\beta}}) &= E[(\widetilde{\boldsymbol{\beta}}-\boldsymbol{\beta})(\widetilde{\boldsymbol{\beta}}-\boldsymbol{\beta})'] \\ &= Var(\widetilde{\boldsymbol{\beta}}) + [Bias(\widetilde{\boldsymbol{\beta}})][Bias(\widetilde{\boldsymbol{\beta}})]'. \end{aligned} \tag{42}$$

Let $\widetilde{\boldsymbol{\beta}}_1$and $\widetilde{\boldsymbol{\beta}}_2$be two estimators of $\boldsymbol{\beta}$ parameter, in the sense of MMSE criterion the estimator $\widetilde{\boldsymbol{\beta}}_2$ is said to be superior to $\widetilde{\boldsymbol{\beta}}_1$ if and only if $MMSE(\widetilde{\boldsymbol{\beta}}_1) - MMSE(\widetilde{\boldsymbol{\beta}}_2) \geq 0$. Since the SMSE is the trace of MMSE

$$SMSE(\widetilde{\boldsymbol{\beta}}) = tr[MMSE(\widetilde{\boldsymbol{\beta}})] = tr\left(Var(\widetilde{\boldsymbol{\beta}})\right) + [Bias(\widetilde{\boldsymbol{\beta}})]'[Bias(\widetilde{\boldsymbol{\beta}})]. \tag{43}$$

Then if the estimator $\widetilde{\boldsymbol{\beta}}_2$ dominates the estimator $\widetilde{\boldsymbol{\beta}}_1$ in the sense of the MMSE criterion, $\widetilde{\boldsymbol{\beta}}_2$ dominates $\widetilde{\boldsymbol{\beta}}_1$in the sense of the SMSE, i.e. $SMSE(\widetilde{\boldsymbol{\beta}}_1) \geq SMSE(\widetilde{\boldsymbol{\beta}}_2)$. The reverse conclusion does not necessarily hold true. Hence, the MMSE is regarded as a stronger criterion than the SMSE, see (Rao and Toutenburg, 1995).Using the SMSE as a search basic criterion implies that we disregard MMSE's off-diagonal components. Therefore, it may be more rational to compare two estimators in the sense of the MMSE criterion.

3.1. Comparison of $\widehat{\boldsymbol{\beta}}_r(\boldsymbol{k},\boldsymbol{d})$ to $\widehat{\boldsymbol{\beta}}(\boldsymbol{k},\boldsymbol{d})$ by the MMSE Criterion

The expected loss or the MMSE of the two-parameter estimator and the restricted two parameter estimator are given respectively by (21) and (41). The difference between the two estimator's MMSEs is

$$\Delta = MMSE\left(\widehat{\boldsymbol{\beta}}(k,d)\right) - MMSE\left(\widehat{\boldsymbol{\beta}}_r(k,d)\right) = \boldsymbol{B} - \boldsymbol{M}_k\boldsymbol{S}_k\boldsymbol{B}\boldsymbol{S}_k\boldsymbol{M}_k\ , \tag{44}$$

where $\boldsymbol{B} = MMSE(\widehat{\boldsymbol{\beta}}(k,d))$.

Özkale and Kaciranlar (2007) derived the necessary and sufficient condition for $\boldsymbol{\Delta}$ to be nnd matrix by applying a theorem from (Graybill, 1976; Graybill, 1983) and gave the following theorem:

Theorem 1. *The $\widehat{\boldsymbol{\beta}}_r(k,d)$ estimator of $\boldsymbol{\beta}$ is superior to the $\widehat{\boldsymbol{\beta}}(k,d)$ estimator by the criterion of MMSE if and only if $\lambda_{max}(\boldsymbol{B}^{-1}\boldsymbol{M}_k\boldsymbol{S}_k\boldsymbol{B}\boldsymbol{S}_k\boldsymbol{M}_k) \leq 1$, where λ_{max} is the maximum eigenvalue of $\boldsymbol{B}^{-1}\boldsymbol{M}_k\boldsymbol{S}_k\boldsymbol{B}\boldsymbol{S}_k\boldsymbol{M}_k$.*

The MMSE criterion comparison between $\widehat{\boldsymbol{\beta}}(k,d)$ and $\widehat{\boldsymbol{\beta}}_{\boldsymbol{r}}(k,d)$ includes the MMSE comparisons between the OLS and the RLS as k approaches zero, the ridge regression and the restricted ridge regression as d approaches zero, the Liu and the MRL when $k = 1$. Hence, we have the following theorems on the comparability of these estimators using the MMSE criterion:

Theorem 2. *A necessary and sufficient condition for the RLS estimator of $\boldsymbol{\beta}$ to be superior to the OLS estimator by the MMSE criterion is $\lambda_{max}(1 - \boldsymbol{R}'[\boldsymbol{R}\boldsymbol{S}^{-1}\boldsymbol{R}']^{-1}\boldsymbol{R}\boldsymbol{S}^{-1}) \leq 1$.*

Theorem 3. *The restricted ridge estimator of $\boldsymbol{\beta}$ dominates the ridge regression estimator by the criterion of MMSE if and only if $\lambda_{max}(\boldsymbol{A}^{-1}\boldsymbol{M}_k\boldsymbol{S}_k\boldsymbol{A}\boldsymbol{S}_k\boldsymbol{M}_k) \leq 1$, where $A = \sigma^2\boldsymbol{S}_k^{-1}\boldsymbol{S}\boldsymbol{S}_k^{-1} + k^2\boldsymbol{S}_k^{-1}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{S}_k^{-1}$.*

Theorem 4. *The necessary and sufficient condition for the MRL estimator of $\boldsymbol{\beta}$ to dominate the Liu estimator by the MMSE criterion is given by $\lambda_{max}(\boldsymbol{C}^{-1}\boldsymbol{M}_1\boldsymbol{S}_1\boldsymbol{C}\boldsymbol{S}_1\boldsymbol{M}_1) \leq 1$, where $\boldsymbol{C} = \sigma^2\boldsymbol{S}_1^{-1}\boldsymbol{S}_d\boldsymbol{S}^{-1}\boldsymbol{S}_d\boldsymbol{S}_1^{-1} + (d-1)^2\boldsymbol{S}_1^{-1}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{S}_1^{-1}$.*

Theorem 2 was introduced by Toro-Vizcarrondo and Wallace (1968) when it is assumed that the restrictions in (22) are not true, while Theorems 3 and 4 were introduced for the first time by Özkale and Kaciranlar (2007). Notice that the results of the comparison depend on the unknown parameters $\boldsymbol{\beta}$ and σ^2, besides the d and k choices. We cannot exclude, because of such unknown parameters that the conditions obtained in the theorems will hold. Therefore, for certain values of $\boldsymbol{\beta}$, σ^2, d, and k the restricted estimators will perform better than the other estimators. Consequently, using a suitable estimate of d and k or utilizing prior information about these parameters, and replacing $\boldsymbol{\beta}$ and σ^2 with their unbiased estimators' leads to feasible results.

3.2. Selection of the Parameters k and d

Trying to find appropriate bias parameters is an integral part of any study of biased estimators in the linear model. One of the most common methods used to estimate bias parameters is to propose the estimators of the biasing parameters that yield minimum MMSE. Nevertheless, due to the inability to diagonalizable$\boldsymbol{M}_k$ and $\boldsymbol{S}_k$ by the same orthogonal matrix, no particular rule for selecting k and d may be suggested in the sense of $\widehat{\boldsymbol{\beta}}_{\boldsymbol{r}}(k,d)$ that is guaranteed to generate minimum MMSE, see (Özkale, 2014).

It is well known that orthogonal transformation can transform a linear regression model into a canonical form. Let $\boldsymbol{Z}=\boldsymbol{XQ}$,$\boldsymbol{\alpha}=\boldsymbol{Q}'\boldsymbol{\beta}$ and $\boldsymbol{Q}$ is $p\times p$ orthogonal matrix such that $\boldsymbol{Z}'\boldsymbol{Z}=\boldsymbol{Q}'\boldsymbol{X}'\boldsymbol{XQ}=\boldsymbol{Q}'\boldsymbol{SQ}=\boldsymbol{\Lambda}=diag(\lambda_1,\dots,\lambda_p)$, where $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_p\geq 0$ are the order eigenvalues of $\boldsymbol{X}'\boldsymbol{X}$, then the canonical form of model (1) can be obtained as

$$\boldsymbol{y}=\boldsymbol{Z\alpha}+\boldsymbol{e}, \tag{45}$$

since $\widehat{\boldsymbol{\alpha}}(k,d)=\boldsymbol{Q}'\widehat{\boldsymbol{\beta}}(k,d)$ and $MMSE(\widehat{\boldsymbol{\alpha}}(k,d))=\boldsymbol{Q}'MMSE\left(\widehat{\boldsymbol{\beta}}(k,d)\right)\boldsymbol{Q}$, based on (21) the MMSE of $\widehat{\boldsymbol{\alpha}}(k,d)$ can be written as

$$\begin{aligned}MMSE\big(\widehat{\boldsymbol{\alpha}}(k,d)\big)&\\ &=\sigma^2(\boldsymbol{\Lambda}+k\boldsymbol{I})^{-1}(\boldsymbol{\Lambda}+kd\boldsymbol{I})\boldsymbol{\Lambda}^{-1}(\boldsymbol{\Lambda}+kd\boldsymbol{I})(\boldsymbol{\Lambda}+k\boldsymbol{I})^{-1}\\ &+\ k^2(d-1)^2(\boldsymbol{\Lambda}+k\boldsymbol{I})^{-1}\boldsymbol{\alpha}\boldsymbol{\alpha}'(\boldsymbol{\Lambda}+k\boldsymbol{I})^{-1}.\end{aligned} \tag{46}$$

Optimal values for$d$ and $k$can be derived by minimizing the quadratic function of d, which can be written as

$$f(k,d)=tr\big[MMSE\big(\widehat{\boldsymbol{\alpha}}(k,d)\big)\big]=\sum_{i=1}^{p}\frac{\sigma^2(\lambda_i+kd)^2+k^2(\mathrm{d}-1)^2\alpha_i^2\lambda_i}{\lambda_i(\lambda_i+k)^2}. \tag{47}$$

The d-value that minimizes the function $f(k,d)$ can be found by differentiating $f(k,d)$ with respect to d when k is fixed.

$$\frac{\partial f(k,d)}{\partial d}=\sum_{i=1}^{p}\frac{2\sigma^2k(\lambda_i+kd)+2k^2(d-1)\alpha_i^2\lambda_i}{\lambda_i(\lambda_i+k)^2}. \tag{48}$$

By equating it to zero, and after replacing the unknown parameters σ^2 and α_i^2 with their unbiased estimators, we get the optimal estimator of d for the fixed value k

$$\hat{d}_{opt} = \frac{\sum_i^p \frac{(k\hat{\alpha}_i^2 - \hat{\sigma}^2)}{(\lambda_i + k)^2}}{\sum_i^p \frac{k(\hat{\sigma}^2 + \hat{\alpha}_i^2 \lambda_i)}{\lambda_i(\lambda_i + k)^2}}. \qquad (49)$$

As was mentioned before, the TP estimator leads to the Liu estimator when $k = 1$. Therefore, when $k = 1$the value of $\hat{d}_{opt}$ in (49) decreases to be equal to

$$\hat{d}_{opt} = \frac{\sum_i^p \frac{(\hat{\alpha}_i^2 - \hat{\sigma}^2)}{(\lambda_i + 1)^2}}{\sum_i^p \frac{(\hat{\sigma}^2 + \hat{\alpha}_i^2 \lambda_i)}{\lambda_i(\lambda_i + 1)^2}}. \qquad (50)$$

This estimate of d was given by Liu (1993). Similarly, the k value that minimizes the function $f(k, d)$ for fixed d-value can be derived by differentiating $f(k, d)$ with respect to d

$$\frac{\partial f(k,d)}{\partial k} = \sum_{i=1}^{p} \frac{2\sigma^2(\lambda_i + kd)(\mathrm{d} - 1) + 2k(d-1)^2\alpha_i^2\lambda_i}{(\lambda_i + k)^3}, \qquad (51)$$

and then equating the numerator to zero. The orientation toward equating only the numerator to zero not all the fraction came from what Hoerl and Kennard (1970) have done during the procedures of finding the optimal k value as $= \frac{\sigma^2}{\alpha_i^2}$. This value of k was obtained by minimizing or partial deriving $tr[MMSE(\hat{\boldsymbol{\alpha}}(k))]$ in respect to k, where $\hat{\boldsymbol{\alpha}}(k)$ is the ridge regression estimator in canonical form, then equating the numerator to zero. Similarly,Özkale and Kaciranlar (2007) derived the k-value by equating the numerator of $\frac{\partial f(k,d)}{\partial k}$ to zero which can be concluded as

$$k = \frac{\sigma^2}{\alpha_i^2 - d(\frac{\sigma^2}{\lambda_i} + \alpha_i^2)}. \qquad (52)$$

Due to the full dependency of the optimal value of k on the unknown parameters σ^2 and α_i^2, when d is fixed. They must be estimated from the observed data and be replaced by their unbiased estimators as suggested by Hoerl and Kennard (1970) and Kibria (2003), so the estimated optimal k value is

$$\hat{k}=\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2-d(\frac{\hat{\sigma}^2}{\lambda_i}+\hat{\alpha}_i^2)}. \tag{53}$$

Another estimator of k was proposed by Hoerl et al. (1975) by taking the harmonic mean of k values that had been found by Hoerl and Kennard (1970b),whereas Kibria (2003) proposed estimators of k by using the arithmetic and geometric means of k values found by Hoerl and Kennard (1970) too. With the same idea, Özkale and Kaciranlar (2007) proposed estimators of k which minimize$f(k,d)$. And they are as follows:

$$\hat{k}_{HM}=\frac{p\hat{\sigma}^2}{\sum_{i=1}^{p}\left[\hat{\alpha}_i^2-d(\frac{\hat{\sigma}^2}{\lambda_i}+\hat{\alpha}_i^2)\right]}, \tag{54}$$

$$\hat{k}_{AM}=\frac{1}{p}\sum_{i=1}^{p}\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2-d(\frac{\hat{\sigma}^2}{\lambda_i}+\hat{\alpha}_i^2)}, \tag{55}$$

$$\hat{k}_{GM}=\frac{\hat{\sigma}^2}{\left(\prod_{i=1}^{p}\left[\hat{\alpha}_i^2-d(\frac{\hat{\sigma}^2}{\lambda_i}+\hat{\alpha}_i^2)\right]\right)^{1/p}}, \tag{56}$$

which respectively are the harmonic mean, the arithmetic mean and the geometric mean of $\hat{k}$ values in (53). Lukman et al. (2020) developed a new estimator for the bias estimator k, which we will refer to as the Lukman bias estimator from now on

$$\hat{k}_{LM}=\frac{p\hat{\sigma}^2}{(1+d)[(\hat{\beta}-J)(\hat{\beta}-J)'-\hat{\sigma}^2tr(\Lambda^{-1})]}, \tag{57}$$

where $J=\frac{\sum_{i=1}^{p}\hat{\beta}_i}{p}$. When k is negative, they suggested using:

$$\hat{k}_{rep}=\frac{p\hat{\sigma}^2}{\sum_{\boldsymbol{i=1}}^{\boldsymbol{p}}\alpha_i^2}. \tag{58}$$

As was shown before when d approaches zero, $\hat{\beta}(k,d)$ approaches the RR estimator. Thus, when d approaches zero, the k estimators in (54), (55), and (56) decrease to k estimators given by Hoerl et al. (1975) and Kibria (2003). Other values of k may be found in Muniz and Kibria, (2009); Aslam (2014) and Kibria and Banik (2016).

Because k must always be positive, the following theorem shows the situation of the positivity of the estimators in Eqs. (54) - (56). Where if k values in (52) are restricted to be positive, the positivity of the estimators may be achieved.

Theorem 5. If

$$\hat{d} < min\left\{\frac{\hat{\alpha}_i^2}{\frac{\hat{\sigma}^2}{\lambda_i} + \hat{\alpha}_i^2}\right\}, \qquad (55)$$

for all i, then the $\hat{k}_{HM}$, $\hat{k}_{AM}$ and $\hat{k}_{GM}$ are always positive. Check the proof in (Özkale and Kaciranlar, 2007).

It is apparent that $\hat{d}_{opt}$ in (49) is dependent on the value of k and the estimators of k are dependent on the value of d. One of the used procedures to avoid looping in choosing and estimating k and d parameters is the iterative method.

(1) Calculate $\hat{d}$ from (59).
(2) By using calculated $\hat{d}$, estimate $\hat{k}_{HM}$, $\hat{k}_{AM}$, $\hat{k}_{GM}$or $\hat{k}_{LM}$.
(3) Find $\hat{d}_{opt}$ in (49) by using obtained $\hat{k}_{HM}$, $\hat{k}_{AM}$, $\hat{k}_{GM}$or $\hat{k}_{LM}$.
(4) If $\hat{d}_{opt}$ is negative replace $\hat{d}_{opt} = \hat{d}$. It should be noted that $\hat{d}_{opt}$ is always less than one, although it may be less than zero. In addition, $\hat{d}$ is always less than one and greater than zero.
(5) If any value of $\hat{k}_{HM}$, $\hat{k}_{AM}$, $\hat{k}_{GM}$or $\hat{k}_{LM}$is negative replace it with $\hat{k}_{rep}$ from (58).

Following a discussion of predictors' theoretical characteristics, *we now*undertake to evaluatea numerical example ofthe performance of estimators in respect of both the scalar mean square error (SMSE) and the matrix mean square error (MMSE) criteria.

4. A REAL-LIFE EXAMPLE

To illustrate the theories shown in the previous chapter, we consider the data set on the Total National Research and Development Expenditures'(TNRDE) data as a Percent of Gross National Product (GNP) for some countries in the time period between 1972 and 1986, were originally given by (Gruber, 1998) and also cited by (Akdeniz and Erol, 2003), (Zhong and Yang, 2007), (Yang and Cui, 2011), (Najarian et al., 2013) and (Şiray, 2014) to compare some biased estimators. The regression on this data set represents the relationship between the dependent variable $\boldsymbol{y}$ which stands for the percentage spent by the United States and another four independent variables $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, $\boldsymbol{X}_3$ and $\boldsymbol{X}_4$. The variable $\boldsymbol{X}_1$ represents the percent spent by France, $\boldsymbol{X}_2$ represents the percent spent by West Germany, $\boldsymbol{X}_3$ represents the percent spent by Japan, and $\boldsymbol{X}_4$ represents the percent spent by the former Soviet Union. The data set consists of 10 observations shown in Table 2.

Table 2. Total national research and development expenditures as a present of GNP by country: 1972-1986

Year	$\boldsymbol{y}$	$\boldsymbol{X_1}$	$\boldsymbol{X_2}$	$\boldsymbol{X_3}$	$\boldsymbol{X_4}$
1972	2.3	1.9	2.2	1.9	3.7
1975	2.2	1.8	2.2	2.0	3.8
1979	2.2	1.8	2.4	2.1	3.6
1980	2.3	1.8	2.4	2.2	3.8
1981	2.4	2.0	2.5	2.3	3.8
1982	2.5	2.1	2.6	2.4	3.7
1983	2.6	2.1	2.6	2.6	3.8
1984	2.6	2.2	2.6	2.6	4.0
1985	2.7	2.3	2.8	2.8	3.7
1986	2.7	2.3	2.7	2.8	3.8

If we fit a linear model without intercept (homogeneous model) to the data, then the $CN = 8776.382$ which means that the $\boldsymbol{X}$ matrix is ill-conditioned and the regressors suffer a serious level of multicollinearity. Under this model the eigenvalues of $\boldsymbol{X}'\boldsymbol{X}$ are

$$\lambda_1 = 302.9626, \lambda_2 = 0.7283\,, \lambda_3 = 0.0446\,, \lambda_4 = \; 0.0345.$$

The variance inflation factors computed from the correlation matrix of the independent variables are

$$VIF_1 = 6.910, VIF_2 = 21.581, VIF_3 = 29.756\,, VIF_4 = \; 1.795,$$

and these factors indicate that the estimates of $\boldsymbol{\beta}_2$ and $\boldsymbol{\beta}_3$ would be affected by the very near-singularity in $\boldsymbol{X}$ matrix. In this case, the near-singularity is known to be due the near-redundancy between $\boldsymbol{X}_2$ and $\boldsymbol{X}_3$. When we fit a linear model with intercept (inhomogeneous model) by adding an $n \times 1$ vector all of its elements equal one to the design matrix, the size of design matrix becomes 10×5. Here we still have $n = 10$ observations, but there are now $p = 5$ unknown regression coefficients. Under this case the eigenvalues of $\boldsymbol{X}'\boldsymbol{X}$ are

$$\lambda_1 = 312.932, \lambda_2 = 0.7536, \lambda_3 = 0.0453, \lambda_4 = 0.0372, \lambda_5 = 0.0019.$$

And the condition number is 168129.285 this indicates the existence of a severe degree of multicollinearity among the regressors. In this section the analysis and results are given only for the homogeneous model. By setting the initial value of d equals to 0.3, the selected k and d values are as follows in Table 3.

Table 3. The selected *k* and *d* values for total national research and development expenditures data

	k **values**	***d*** **optimum**
Harmonic mean	0.0202	**0.1940**
Arithmetic mean	0.0988	**0.6378**
Geometric mean	0.0406	**0.4742**

For the linear restriction $\boldsymbol{R\beta} = \boldsymbol{r}$, we use the $\boldsymbol{R}$ vector suggested by (Yang and Cui, 2011) and extract the value of $\boldsymbol{r}$

$$\boldsymbol{R} = [1 \quad -2 \quad -2 \quad -2\ ], \qquad \boldsymbol{r} = [-2.775].$$

The values of all our biased estimators are obtained and their respective SMSE values are computed then summarized in Table 4.

Table 4. The coefficient estimates and the SMSE values of OLS, RR, Liu, TP, RLS, RRR, MRL, and RTP estimators for TNRDE data

	$(\boldsymbol{k}\,,\boldsymbol{d})$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$	$\boldsymbol{\beta_3}$	$\boldsymbol{\beta_4}$	**SMSE**
OLS	(0, 0)	0.6455	0.0896	0.1436	0.1526	**0.0808**
RR	(0.0202, 0)	0.5053	0.1293	0.1965	0.1685	**0.0360**
	(0.0988, 0)	0.3370	0.1832	0.2464	0.1918	**0.0079**
	(0.0406, 0)	0.4319	0.1525	0.2208	0.1773	**0.0207**
Liu	(0,0.1940)	0.2877	0.1871	0.2161	0.2335	**9.0235**
	(0,0.6378)	0.4846	0.1334	0.1762	0.1890	**9.0424**
	(0,0.4742)	0.4120	0.1532	0.1909	0.2054	**9.0321**
TP	(0.0202, 0.1940)	0.5325	0.1216	0.1862	0.1654	**0.0432**
	(0.0988, 0.6378)	0.5337	0.1235	0.1808	0.1668	**0.0450**
	(0.0406, 0.4742)	0.5332	0.1227	0.1842	0.1656	**0.0440**
RLS	(0,0)	-0.3848	0.3699	0.5741	0.2511	**0.0451**
RRR	(0.0202, 0)	-0.3836	0.3973	0.5525	0.2459	**0.0195**
	(0.0988, 0)	-0.3851	0.4235	0.5240	0.2475	**0.0047**
	(0.0406, 0)	-0.3836	0.4098	0.5411	0.2449	**0.0113**
MRL	(0,0.1940)	-0.3953	0.4321	0.4806	0.2772	**0.0028**
	(0,0.6378)	-0.3790	0.4433	0.5106	0.2442	**0.0196**
	(0,0.4742)	-0.3850	0.4392	0.4995	0.2564	**0.0115**
RTP	(0.0202, 0.1940)	-0.3833	0.3977	0.5530	0.2452	**0.0236**
	(0.0988, 0.6378)	-0.3805	0.4277	0.5323	0.2373	**0.0252**
	(0.0406, 0.4742)	-0.3821	0.4115	0.5436	0.2414	**0.0243**

Corresponding to the results obtained in the preceding example, from Table 4 we can observe that the restricted estimators are performing better than the unrestricted estimators in the sense of smaller SMSE. Puzzlingly, the Liu estimator provides estimates with SMSE values that are remarkably far away from the other unrestricted estimators. The RR estimator is superior to the unrestricted estimator set, and the RRR estimator is superior to all the restricted and unrestricted estimators in the sense of SMSE criterion. Also, it was noted that the RR and RRR estimators of $\hat{k}_{AM}$ perform better comparing to the RR and RRR estimators of $\hat{k}_{HM}$ and $\hat{k}_{GM}$ as well as outperform all other biased estimators.

Table 5. The superiority according to MMSE

Superiority of by MMSE criterion	Condition of superiority	Programme's results
RLS to OLS by MMSE criterion	$\lambda_{max}(1-\boldsymbol{R}'[\boldsymbol{R}\boldsymbol{S}^{-1}\boldsymbol{R}']^{-1}\boldsymbol{R}\boldsymbol{S}^{-1}) \leq 1.$	**1**
RRR to RR by MMSE criterion	$\lambda_{max}(\boldsymbol{A}^{-1}\boldsymbol{M}_k\boldsymbol{S}_k\boldsymbol{A}\boldsymbol{S}_k\boldsymbol{M}_k) \leq 1$	**1.0009**
		1.0047
		1.0020
MRL to LE by MMSE criterion	$\lambda_{max}(\boldsymbol{C}^{-1}\boldsymbol{M}_1\boldsymbol{S}_1\boldsymbol{C}\boldsymbol{S}_1\boldsymbol{M}_1) \leq 1$	**1.0036**
		1.0314
		1.0169
RTP to TP by MMSE criterion	$\lambda_{max}(\boldsymbol{B}^{-1}\boldsymbol{M}_k\boldsymbol{S}_k\boldsymbol{B}\boldsymbol{S}_k\boldsymbol{M}_k) \leq 1$	**1.0002**
		1.0021
		1.0002

As seen in Table 5,when the estimators are compared according to MMSE criterion, the RLS estimator is superior to the OLS estimator, the RRR estimators is superior to the RR estimators, the MRL estimators is superior to the Liu estimators, and the RTP estimators is superior to the TP estimators, since the data fulfills the necessary and sufficient conditions mentioned in the Theorems (1), (2), (3) and (4).

After discussing the theoretical aspects of the estimators using a real-world example, we proceed to a simulation study to empirically compare their performance under various multicollinearity situations.

5. MONTE-CARLO SIMULATION STUDY

For the sake of drawing an extensive and generalizable conclusion of relative characteristics of our pre-discussed estimators further than the results of the investigation that had been on the real data, a simulation study has been conducted.

5.1. Simulation Essence

In this section, we perform a Monte-Carlo simulation study by using MATLAB program. By following McDonald and Galarneau (1975) as many other researchers like Kejian (1993) the generation of independent variables was based on subsequent equation

$$X_{ij} = (1-\gamma^2)^{1/2} Z_{ij} + \gamma Z_{ip}, i = 1,2, \dots, n, j = 1,2, \dots, p, \quad (60)$$

where Z_{ij} are independent standard normal pseudo-random numbers and γ is defined in such a way that the correlation between any two explanatory variables is provided by γ^2. The independent variables were standardized resulting in $\boldsymbol{X'X}$ being in correlation form. The observations on the dependent variables were generated by

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + e_i, i = 1,2, \dots, n, \quad (61)$$

where e_i are independent normal $(0, \sigma^2)$ pseudo-random numbers. Also, $\boldsymbol{y}$ is standardized such that $\boldsymbol{X'y}$ represents the vector of the dependent variable's correlations with each explanatory variable. In this study a comprehensive simulation is intended; therefore, three different sample sizes $n = 25, 50, 100$ for two different independent variables numbers $p = 4, 7$ are adopted, and following different authors such as DG Gibbons (1981) Four distinctvalues of correlations (γ) are considered, corresponding to γ = 0.8, 0.9, 0.95, 0.99to demonstrate a weak, strong and severe level of multicollinearity among the regressors. Also, for the error standard deviation two different values $\sigma = 1$, 5 are investigated, and the intercept β_0 is presumed to be identically zero. n, p and σ was selected at varying values to examine its impact on the estimation. These values were chosen to be either equal to or closely aligned with those utilized in simulations reported in the literature.

According to Newhouse and Oman (1971) if the mean square error is a function of $\boldsymbol{\beta}$, σ^2 and k and if the independent variables are fixed, then the MMSE is minimized when $\boldsymbol{\beta}$ is the normalized eigenvector corresponding to the largest eigenvalue of $\boldsymbol{X'X}$ matrix subject to constraint that $\boldsymbol{\beta'\beta} = 1$. Consequently, we selected the eigenvector that makes $\boldsymbol{\beta'\beta} = 1$ as a parameter vector. Now coming to the exact linear restrictions, when $p = 4$ the $\boldsymbol{R\beta} = \boldsymbol{r}$ was specified supposing $\boldsymbol{r} = [0]$ then $\boldsymbol{R} = [1 \quad -1 \quad 1 \quad 0]$, and when $p = 7$ the linear restriction was specified supposing that $\boldsymbol{R} = [1 \quad 0 \quad 1 \quad -1 \quad 1 \quad 1 \quad 0]$ as a result $\boldsymbol{r} = [1.0640]$. Notice that to ensure $rank(\boldsymbol{R}) = m < p$ we designed $\boldsymbol{R}$ matrix to be $1 \times p$.

In simulation study, for each σ, γ, p and n the experiment is performed 3000 times. The estimated mean square error (EMSE) is computed as

$$EMSE(\widehat{\boldsymbol{\alpha}}) = \frac{1}{RN} \sum_{i=1}^{RN} (\widehat{\boldsymbol{\alpha}}_i - \boldsymbol{\alpha})'(\widehat{\boldsymbol{\alpha}}_{\boldsymbol{i}} - \boldsymbol{\alpha}), \quad (62)$$

where $\boldsymbol{\alpha}$ is the parameter vector in the canonical form, $\widehat{\boldsymbol{\alpha}}_{\boldsymbol{i}}$ is one of α estimators in the ith replication, and RN is the replications number of the experiment which in our study end up to 3000.

5.2. SimulationResults

The values of the parameter coefficients and the eigen values of $\boldsymbol{X'X}$ for the different values of p and n resulting from the simulation are obtained.

Tables 6-11 describe the simulation findings of 3000 replications. On the tables, the condition number (CN) of each n and p altered experiment is also given. Deserves to be mentioned that the k's used in the simulation are the harmonic ($\hat{k}_{HM}$), the arithmetic ($\hat{k}_{AM}$) means of $\hat{k}$ and the Lukman bias estimator ($\hat{k}_{LM}$), respectively.

Table 6. Estimated $\boldsymbol{k}$ and $\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n = 25,\ p = 4}$

					$\boldsymbol{\gamma = 0.80}$				$\boldsymbol{CN = 22.869}$	
			EMSE of the unrestricted biased estimators				**EMSE of the restricted biased estimators**			
$\boldsymbol{\sigma}$	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**2.793**	**0.019**	0.598	0.293	0.375	0.254	0.576	0.330	0.402	0.298
	511.083	**0.377**	0.598	0.291	0.490	0.301	0.576	0.338	0.487	0.333
	34.831	**0.191**	0.598	0.177	0.456	0.257	0.576	0.237	0.461	0.294
5	**0.705**	**0.285**	373.965	338.165	293.513	348.819	270.899	244.860	213.885	251.926
	249.299	**0.801**	373.965	128.566	360.486	348.161	270.899	94.396	261.938	262.714
	4.063	**0.330**	373.965	335.978	296.369	348.759	270.899	243.165	215.954	251.774
					$\boldsymbol{\gamma = 0.90}$				$\boldsymbol{CN}$ **=54.017**	
$\boldsymbol{\sigma}$	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**2.235**	**0.048**	1.171	0.462	0.597	0.485	0.996	0.440	0.554	0.458
	534.776	**0.320**	1.171	0.279	0.819	0.512	0.996	0.308	0.721	0.494
	15.147	**0.204**	1.171	0.239	0.763	0.470	0.996	0.267	0.678	0.447
5	**0.388**	**0.299**	731.801	658.704	484.220	682.113	538.550	483.644	359.029	500.031
	338.892	**0.830**	731.801	185.110	698.035	680.775	538.550	134.155	517.653	524.578
	3.780	**0.344**	731.801	654.337	492.289	681.983	538.550	480.206	364.974	499.712
					$\boldsymbol{\gamma = 0.95}$				$\boldsymbol{CN}$ **=118.234**	
$\boldsymbol{\sigma}$	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**

Table 6 (Continued). Estimated $\boldsymbol{k}$ and $\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n = 25,\ p = 4}$

1	**1.663**	**0.078**	**2.321**	**0.780**	**0.883**	**0.949**	**1.860**	**0.662**	**0.757**	**0.797**
	3.717e+3	0.296	2.321	0.324	1.324	0.936	1.860	0.323	1.099	0.830
	7.715	**0.250**	2.321	0.379	1.229	0.908	1.860	0.361	1.021	0.777
5	**0.206**	**0.308**	1.451e+3	1.302e+3	759.952	1.352e+3	1.081e+3	967.253	575.340	1.003e+3
	1.021e+3	**0.856**	1.451e+3	251.251	1.370e+3	1.349e+3	1.081e+3	179.788	1.039e+3	1.058e+3
	1.582	**0.352**	1.451e+3	1.294e+3	780.193	1.351e+3	1.081e+3	960.315	590.385	1.002e+3
				$\boldsymbol{\gamma = 0.99}$				$\boldsymbol{CN = 642.202}$		
$\boldsymbol{\sigma}$	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**0.640**	**0.129**	11.531	3.261	1.959	4.661	8.876	2.445	1.633	3.579
	34.767	**0.319**	11.531	0.705	4.385	4.396	8.876	0.533	3.604	3.629
	3.622	**0.331**	11.531	1.563	3.673	4.436	8.876	1.238	2.961	3.519
5	**0.045**	**0.320**	7.207e+3	6.454e+3	2.2822e+3	6.711e+3	5.469e+3	4.872e+3	1.855e+3	5.068e+3
	600.704	**0.890**	7.207e+3	349.139	6.712e+3	6.696e+3	5.469e+3	228.653	5.359e+3	5.405e+3
	1.470	**0.362**	7.207e+3	6.410e+3	2.369e+3	6.710e+3	5.469e+3	4.837e+3	1.926e+3	5.065e+3

Table 7. Estimated $\boldsymbol{k}$ and$\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n = 50,\ p = 4}$

					$\boldsymbol{\gamma = 0.80}$				$\boldsymbol{CN = 18.222}$		
$\boldsymbol{\sigma}$	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**3.391**	**0.005**	0.234	0.154	0.175	0.108	0.260	0.191	0.211	0.156
	1.420e+3	**0.468**	0.234	0.277	0.216	0.141	0.260	0.313	0.245	0.182
	248.282	**0.193**	0.234	0.116	0.201	0.113	0.260	0.163	0.232	0.160
5	**1.688**	**0.263**	146.379	132.884	131.807	136.289	113.348	102.422	102.079	105.894
	757.506	**0.776**	146.379	57.127	143.658	136.141	113.348	41.939	111.272	106.084
	4.992	**0.301**	146.379	132.004	132.415	136.272	113.348	101.756	102.520	105.880
					$\boldsymbol{\gamma = 0.90}$				$\boldsymbol{CN = 44.081}$		
$\boldsymbol{\sigma}$	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**2.971**	**0.023**	0.457	0.240	0.308	0.202	0.443	0.260	0.320	0.238
	1.44e+3	**0.359**	0.457	0.229	0.388	0.227	0.443	0.267	0.385	0.258
	29.133	**0.176**	0.457	0.144	0.360	0.198	0.443	0.186	0.362	0.235
5	**0.947**	**0.295**	285.580	257.636	235.923	265.795	223.284	200.214	184.526	208.516
	1.07e+3	**0.809**	285.580	83.568	278.122	265.481	223.284	59.806	217.596	209.152
	2.492	**0.332**	285.580	255.912	237.800	265.753	223.284	198.904	185.890	208.488
					$\boldsymbol{\gamma = 0.95}$				$\boldsymbol{CN = 97.476}$		
$\boldsymbol{\sigma}$	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**

Table 7 (Continued). Estimated $\boldsymbol{k}$ and$\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n = 50},\ \boldsymbol{p = 4}$

1	2.440	0.050	**0.911**	**0.385**	**0.517**	**0.392**	**0.804**	**0.369**	**0.481**	**0.393**
	660.995	**0.302**	0.911	0.215	0.674	0.398	0.804	0.245	0.610	0.401
	160.254	**0.186**	0.911	0.206	0.630	0.374	0.804	0.236	0.572	0.383
5	**0.506**	**0.309**	569.090	511.692	406.871	529.574	446.939	399.147	321.003	417.243
	852.577	**0.838**	569.090	120.004	548.970	528.806	446.939	83.164	431.763	418.920
	4.337	**0.347**	569.090	508.260	412.491	529.482	446.939	396.537	325.054	417.186
					$\boldsymbol{\gamma = 0.99}$				$\boldsymbol{CN = 532.792}$	
$\boldsymbol{\sigma}$	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**1.152**	**0.108**	4.592	1.411	1.285	1.905	3.707	1.132	1.093	1.600
	193.966	**0.292**	4.592	0.429	2.158	1.797	3.707	0.367	1.800	1.546
	6.382	**0.292**	4.592	0.685	1.964	1.805	3.707	0.622	1.605	1.555
5	**0.110**	**0.321**	2.870e+3	2.574e+3	1.206e+3	2.670e+3	2.260e+3	2.011e+3	1.004e+3	2.109e+3
	3.921e+3	**0.885**	2.870e+3	196.285	2.708e+3	2.666e+3	2.260e+3	125.942	2.146e+3	2.123e+3
	1.621	**0.359**	2.870e+3	2.556e+3	1.246e+3	2.669e+3	2.260e+3	1.998e+3	1.033e+3	2.108e+3

Table 8. Estimated $\boldsymbol{k}$ and $\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n = 100},\ \boldsymbol{p = 4}$

					$\boldsymbol{\gamma = 0.80}$				$\boldsymbol{CN = 41.203}$	
$\boldsymbol{\sigma}$	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**

Table 8 (Continued). Estimated $\boldsymbol{k}$ and $\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n = 100,\ p = 4}$

1	3.423	0.006	**0.206**	**0.134**	**0.157**	**0.097**	**0.268**	**0.202**	**0.226**	**0.172**
	1.43e+3	**0.430**	0.206	0.199	0.190	0.114	0.268	0.253	0.252	0.173
	70.316	**0.200**	0.206	0.088	0.178	0.096	0.268	0.161	0.243	0.166
5	**2.117**	**0.327**	128.571	115.592	117.419	119.910	94.347	84.570	86.266	87.569
	1.04e+3	**0.794**	128.571	42.266	126.721	119.701	94.377	30.901	93.118	91.482
	7.546	**0.364**	128.571	114.769	117.878	119.881	94.347	83.922	86.602	87.504
				$\boldsymbol{\gamma = 0.90}$				$\boldsymbol{CN = 95.929}$		
$\boldsymbol{\sigma}$	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**3.001**	**0.029**	0.423	0.218	0.291	0.192	0.430	0.257	0.322	0.238
	491.882	**0.341**	0.423	0.163	0.359	0.198	0.430	0.215	0.375	0.239
	31.587	**0.184**	0.423	0.118	0.337	0.182	0.430	0.176	0.357	0.229
5	**1.107**	**0.343**	264.429	236.678	222.487	246.565	196.469	175.072	165.650	182.368
	3.70e+3	**0.825**	264.429	64.720	258.783	246.040	196.469	46.594	192.962	191.132
	3.375	**0.380**	264.429	235.037	224.042	246.503	196.469	173.777	166.800	182.246
				$\boldsymbol{\gamma = 0.95}$				$\boldsymbol{CN = 205.239}$		
$\boldsymbol{\sigma}$	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**2.476**	**0.062**	0.855	0.353	0.498	0.376	0.756	0.349	0.472	0.373
	560.278	**0.304**	0.855	0.171	0.636	0.363	0.756	0.212	0.580	0.373

Table 8 (Continued). Estimated $\boldsymbol{k}$ and $\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n=100,\ p=4}$

	46.987	0.197	**0.855**	**0.180**	**0.598**	**0.353**	**0.756**	**0.216**	**0.550**	**0.359**
5	**0.577**	**0.351**	534.328	477.241	393.739	498.150	400.932	356.203	297.104	372.319
	5.48e+3	**0.851**	534.328	90.640	518.368	496.873	400.932	63.567	392.257	390.818
	2.311	**0.388**	534.328	473.986	398.530	498.023	400.932	353.635	300.638	372.092
				$\boldsymbol{\gamma=0.99}$				$\boldsymbol{CN=1072.73}$		
σ	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**1.205**	**0.124**	4.276	1.274	1.297	1.798	3.366	1.127	1.078	1.455
	157.228	**0.296**	4.276	0.401	2.050	1.672	3.366	0.372	1.684	1.431
	4.943	**0.293**	4.276	0.593	1.893	1.682	3.366	0.522	1.536	1.404
5	**0.128**	**0.357**	2.672e+3	2.383e+3	1.229e+3	2.491e+3	2.032e+3	1.801e+3	984.607	1.889e+3
	1.254e+4	**0.889**	2.672e+3	126.367	2.536e+3	2.484e+3	2.032e+3	81.385	1.989e+3	1.992e+3
	1.840	**0.394**	2.672e+3	2.367e+3	1.263e+3	2.490e+3	2.032e+3	1.788e+3	1.010e+3	1.888e+3

Table 9. Estimated $\boldsymbol{k}$ and $\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n=25,\ p=7}$

				$\boldsymbol{\gamma=0.80}$				$\boldsymbol{CN=\ 48.402}$		
σ	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**

Table 9 (Continued). Estimated $\boldsymbol{k}$ and $\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n = 25,\ p = 7}$

1	3.897	0.073	**1.040**	**0.392**	**0.705**	**0.440**	**0.737**	**0.312**	**0.527**	**0.334**
	635.578	**0.435**	1.040	0.319	0.859	0.467	0.737	0.279	0.626	0.377
	35.491	**0.222**	1.040	0.209	0.796	0.412	0.737	0.178	0.585	0.314
5	**0.474**	**0.376**	650.111	576.111	525.541	607.231	460.729	413.780	379.841	430.028
	357.531	**0.870**	650.111	150.168	632.075	603.625	460.729	121.305	449.895	453.801
	0.732	**0.403**	650.111	572.566	529.967	607.124	460.729	411.209	382.942	429.779
					$\boldsymbol{\gamma = 0.90}$			$\boldsymbol{CN = 127.210}$		
σ	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**2.920**	**0.119**	1.963	0.605	1.116	1.028	1.388	0.478	0.845	0.617
	1.023e+3	**0.401**	1.963	0.254	1.417	0.972	1.388	0.222	1.044	0.623
	17.216	**0.267**	1.963	0.279	1.303	0.750	1.388	0.228	0.969	0.563
5	**0.266**	**0.402**	1.227e+3	1.077e+3	884.755	1.146e+3	867.562	770.563	646.895	808.315
	174.219	**0.881**	1.227e+3	212.230	1.180e+3	1.139e+3	867.562	171.363	845.920	866.450
	0.504	**0.427**	1.227e+3	1.070e+3	896.137	1.146e+3	867.562	765.404	654.890	807.696
					$\boldsymbol{\gamma = 0.95}$			$\boldsymbol{CN}$ **=294.289**		
σ	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**2.002**	**0.153**	3.816	1.026	1.727	1.595	2.699	0.806	1.332	1.185
	408.994	**0.386**	3.816	0.227	2.337	1.401	2.699	0.200	1.754	1.133
	10.928	**0.314**	3.816	0.427	2.128	1.431	2.699	0.338	1.610	1.068

Table 9 (Continued). Estimated $\boldsymbol{k}$ and $\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n = 25},\ \boldsymbol{p = 7}$

5	0.141	0.416	**2.385e+3**	**2.081e+3**	**1.493e+3**	**2.229e+3**	**1.686e+3**	**1.488e+3**	**1.117e+3**	**1.569e+3**
	536.684	**0.892**	2.385e+3	297.815	2.268e+3	2.212e+3	1.686e+3	241.884	1.657e+3	1.709e+3
	0.471	**0.441**	2.385e+3	2.067e+3	1.520e+3	2.228e+3	1.686e+3	1.477e+3	1.137e+3	1.568e+3
				$\boldsymbol{\gamma = 0.99}$				**CN =1662.300**		
σ	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**0.613**	**0.190**	18.647	4.358	4.615	7.722	13.198	3.401	3.742	5.718
	52.212	**0.385**	18.647	0.427	7.810	6.580	13.198	0.373	6.174	5.364
	6.046	**0.376**	18.647	1.671	6.865	6.893	13.198	1.277	5.453	5.138
5	**0.030**	**0.429**	1.165e+4	1.011e+4	5.450e+3	1.089e+4	8.249e+3	7.233e+3	4.420e+3	7.666e+3
	391.269	**0.913**	1.165e+4	531.585	1.087e+4	1.080e+4	8.249e+3	0.445e+3	8.459e+3	8.614e+3
	0.429	**0.453**	1.165e+4	1.005e+4	5.615e+3	1.089e+4	8.249e+3	7.181e+3	4.511e+3	7.658e+3

Table 10. Estimated $\boldsymbol{k}$ and $\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n = 50},\ \boldsymbol{p = 7}$

				$\boldsymbol{\gamma = 0.80}$				**CN =117.413**		
σ	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**4.097**	**0.081**	0.888	0.358	0.635	0.394	0.783	0.310	0.559	0.346
	790.860	**0.408**	0.888	0.221	0.747	0.377	0.783	0.189	0.658	0.342
	41.622	**0.211**	0.888	0.185	0.702	0.355	0.783	0.161	0.618	0.316

Table 10 (Continued). Estimated $\boldsymbol{k}$ and $\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n = 50,\ p = 7}$

5	0.537	0.381	**554.793**	**492.055**	**463.605**	**518.334**	**488.500**	**432.061**	**408.541**	**455.440**
	878.785	**0.872**	554.793	119.744	542.017	514.952	488.500	102.961	477.652	461.550
	0.651	**0.401**	554.793	489.118	466.951	518.225	488.500	429.385	411.476	455.265
					$\boldsymbol{\gamma = 0.90}$				$\boldsymbol{CN = 294.220}$	
σ	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**2.937**	**0.120**	1.893	0.615	1.122	0.819	1.668	0.530	0.984	0.720
	487.588	**0.383**	1.893	0.190	1.385	0.718	1.668	0.162	1.220	0.653
	15.251	**0.255**	1.893	0.274	1.290	0.727	1.668	0.237	1.134	0.649
5	**0.260**	**0.398**	1.183e+3	1.043e+3	869.456	1.106e+3	1.042e+3	915.676	768.721	971.182
	594.012	**0.888**	1.183e+3	184.612	1.142e+3	1.098e+3	1.042e+3	156.894	1.010e+3	986.973
	0.315	**0.417**	1.183e+3	1.037e+3	880.062	1.106e+3	1.042e+3	909.915	778.004	970.745
					$\boldsymbol{\gamma = 0.95}$				$\boldsymbol{CN}$ **= 652.400**	
σ	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**1.929**	**0.144**	3.907	1.109	1.809	1.659	3.441	0.952	1.589	1.460
	236.112	**0.374**	3.907	0.206	2.407	1.411	3.441	0.171	2.124	1.286
	10.185	**0.312**	3.907	0.444	2.225	1.468	3.441	0.385	1.957	1.314
5	**0.128**	**0.405**	2.442e+3	2.148e+3	1.529e+3	2.284e+3	2.150e+3	1.884e+3	1.363e+3	2.004e+3
	1.351e+3	**0.901**	2.442e+3	262.643	2.330e+3	2.264e+3	2.150e+3	0.222e+3	2.072e+3	2.042e+3
	0.160	**0.425**	2.442e+3	2.135e+3	1.556e+3	2.283e+3	2.150e+3	1.872e+3	1.387e+3	2.003e+3

Table 10 (Continued). Estimated $\boldsymbol{k}$ and $\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n = 50},\ \boldsymbol{p = 7}$

					$\gamma = 0.99$				CN =3523.237		
σ	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**0.547**	**0.172**	19.995	4.981	4.701	8.329	17.597	4.264	4.252	7.335
	85.680	**0.382**	19.995	0.567	8.167	7.038	17.597	0.457	7.367	6.433
	11.122	**0.382**	19.995	1.893	7.302	7.417	17.597	1.655	6.536	6.661
5	**0.025**	**0.411**	1.250e+4	1.097e+4	5.970e+3	1.169e+4	1.100e+4	9.620e+3	5.438e+3	1.025e+4
	2.625e+4	**0.923**	1.250e+4	388.799	1.169e+4	1.159e+4	1.100e+4	319.652	1.054e+4	1.050e+4
	0.067	**0.431**	1.250e+4	1.091e+4	6.042e+3	1.169e+4	1.100e+4	9.559e+3	5.507e+3	1.025e+4

Table 11. Estimated $\boldsymbol{k}$ and $\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n = 100},\ \boldsymbol{p = 7}$

					$\gamma = 0.80$				**CN =79.782**	
σ	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**5.490**	**0.039**	0.317	0.174	0.263	0.155	0.245	0.147	0.210	0.127
	4614.253	**0.483**	0.317	0.241	0.296	0.167	0.245	0.203	0.232	0.144
	102.477	**0.200**	0.317	0.111	0.282	0.144	0.245	0.098	0.222	0.120
5	**1.553**	**0.377**	198.163	175.597	184.073	185.082	150.989	135.542	141.164	140.212
	2167.309	**0.853**	198.163	51.744	195.962	183.987	150.989	43.690	149.572	149.713
	1.966	**0.401**	198.163	174.457	184.648	185.034	150.989	134.613	141.598	140.072

Table 11 (Continued). Estimated $\boldsymbol{k}$ and $\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n = 100,\ p = 7}$

					$\boldsymbol{\gamma = 0.90}$				$\boldsymbol{CN = 205.589}$		
σ	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**4.577**	**0.081**	0.653	0.278	0.494	0.309	0.498	0.233	0.394	0.246
	1702.663	**0.404**	0.653	0.171	0.565	0.282	0.498	0.147	0.442	0.244
	27.895	**0.197**	0.653	0.150	0.534	0.276	0.498	0.130	0.421	0.225
5	**0.787**	**0.405**	408.011	358.634	356.263	381.420	309.070	275.237	273.318	286.680
	702.633	**0.869**	408.011	82.828	400.538	378.630	309.070	70.223	305.145	311.437
	1.038	**0.426**	408.011	356.316	358.202	381.298	309.070	273.319	274.803	286.347
					$\boldsymbol{\gamma = 0.95}$				$\boldsymbol{CN = 463.703}$		
σ	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**3.520**	**0.117**	1.327	0.455	0.868	0.602	1.010	0.380	0.699	0.476
	976.972	**0.373**	1.327	0.139	1.028	0.512	1.010	0.120	0.815	0.443
	18.627	**0.227**	1.327	0.221	0.964	0.533	1.010	0.187	0.769	0.429
5	**0.396**	**0.416**	829.409	726.137	657.913	775.736	628.939	557.535	510.236	583.092
	791.103	**0.884**	829.409	126.481	806.416	769.208	628.939	107.405	621.393	640.870
	0.567	**0.437**	829.409	721.469	663.919	775.468	628.939	553.625	514.961	582.375
					$\boldsymbol{\gamma = 0.99}$				$\boldsymbol{CN = 2543.556}$		
σ	$\boldsymbol{k}$	$\boldsymbol{d}$	**OLS**	**RR**	**LE**	**TP**	**RLS**	**RRR**	**MRL**	**RTP**
1	**1.360**	**0.163**	6.725	1.746	2.571	2.886	5.140	1.444	2.134	2.278

Table 11 (Continued). Estimated $\boldsymbol{k}$ and $\boldsymbol{d}$ and EMSE for the restricted and unrestricted biased estimators when $\boldsymbol{n = 100,\ p = 7}$

	185.821	0.370	**6.725**	**0.238**	**3.594**	**2.408**	**5.140**	**0.200**	**2.973**	**2.102**
	9.594	**0.342**	6.725	0.668	3.314	2.540	5.140	0.553	2.742	2.043
5	**0.079**	**0.424**	4203.171	3668.386	2407.509	3932.684	3209.213	2832.084	1974.738	2974.577
	751.011	**0.911**	4203.171	246.996	3987.508	3894.249	3209.213	207.005	3267.161	3335.638
	0.272	**0.444**	4203.171	3644.953	2454.065	3931.223	3209.213	2812.175	2013.169	2970.834

The results in Tables 6-11 summarize the average of the mean square errors values resulting from the simulation for each estimator, and show the averaged values for simulated harmonic, arithmetic means, and Lukman's bias estimator of k and their corresponding average of simulated d-optimal for each estimate method. We have the following results of Monte-Carlo simulation study.

1. Performance as a Function of the Restrictions: This efficiency in performance in detail applies for all the study's estimators regardless of n, p, k, d and for fixed σ and γ values. i.e, the $\widehat{\boldsymbol{\beta}}_r$ estimator is superior to the $\widehat{\boldsymbol{\beta}}$, the $\widehat{\boldsymbol{\beta}}_r(k)$ estimator is superior to the $\widehat{\boldsymbol{\beta}}_k$ estimator, the $\widehat{\boldsymbol{\beta}}_r(d)$ estimator is superior to the $\widehat{\boldsymbol{\beta}}_d$ estimator, and the $\widehat{\boldsymbol{\beta}}_r(k,d)$ estimator is superior to the $\widehat{\boldsymbol{\beta}}(k,d)$ estimator.Nonetheless, some exceptions were noticed in Tables 6-9 when $\gamma = 0.80, 0.90$ and $\sigma = 1$, in these specific parts of Tables the unrestricted biased estimators show better performance than the restricted estimators by having fewer EMSE values. This might be explained by the fact that these cases create null to low degrees of multicollinearity, resulting in modest condition numbers. It is worth noting that this does not occur when σ = 5, so we conclude thatthe restricted estimators do perform better when the error has a considerable standard deviation even in the absence of the multicollinearity. In other words, as σ gets larger not only the EMSE values grows, but also does the degree of superiority of the restricted estimators over the unrestricted estimators.

2. Performance as a Function of $\boldsymbol{\gamma}$*:* For fixed values of σ, n and p, every increase in the correlation value represented in γ resulted in an increase in the amount of averaged mean square error for both the restricted and unrestricted biased estimators. Moreover, as γ increases the performance gap between the OLS estimator and the other unrestricted estimators becomes wider, as well as the performance gap between the RLS estimator and the other restricted estimators. Since by γ growing the EMSE values of the OLS and the RLS estimators increase substantially unlike the other estimators. This may explain the genuine influence of the multicollinearity on the OLS and RLS estimation by making them unstable, have large variance, and have longer vector than $\boldsymbol{\beta}$'s vector $\|\boldsymbol{\beta}\| < \|\widehat{\boldsymbol{\beta}}\|$ and $\|\boldsymbol{\beta}\| < \|\widehat{\boldsymbol{\beta}}_r\|$. Generally *w*hen correlation confections increase the restricted estimators perform better, because they impose constraints that can help stabilize the estimates. If restrictions align with the true underlying relationships, they can lead to more precise estimations in the presence of multicollinearity. However, if the restrictions are incorrect or overly strict the restricted estimators will be biased and inefficient.

3. Performance as a Function of $\boldsymbol{k}$ *and* $\boldsymbol{d}$*:* The EMSE for each level of correlations decreases down as k grows. Which justifies the RR of $\hat{k}_{AM}$with the smallest EMSE value among the RR estimators and all other unrestricted biased estimators. Similarly, having the RRR of $\hat{k}_{AM}$the smallest EMSE value among the other RRR estimators and outperforming all restricted biased estimators become more understandable. Regarding $d$what happened is the exact opposite, the largest estimate of $d$corresponds to the largest EMSE values of the Liu and MRL estimators. Although, a negative link between k and the EMSE is concluded, on the contrary to Lukman et al. (2019)a positive link between d and EMSE is concluded.

4. Performance as a Function of $\boldsymbol{\sigma}$:A positive direct relationship between the σ and the EMSE have been determined based on the results of the Tables. Regardless of γ value, dramatic changes have been noticed occurring in the EMSE values after a small change in σ. In the unrestricted biased estimator group, when $\sigma = 1$, the RR estimators outperform the other unrestricted estimators in virtually all instances, Specifically, the RR of $\hat{k}_{AM}$. The second acceptable performance is shown by the TP estimator. Exceptions are unavoidable; for example, in Table 6, when $\gamma = 0.99$, some of Liu's estimations take center stage. However, when $\sigma =$ 5the best performance has occurred by $\hat{k}_{AM}$ of RR, then by the first and third estimated values of d at Liu, followed by the other ridge and the other Liu estimations, indicating an overlapped level of efficiency between Liu and ridge except for the second estimated d due of being larger than the other values.

5. Performance as a Function of $\boldsymbol{n}$*:* The rise in the number of observations leads to a decrease in the amount of averaged mean square error for both the restricted and unrestricted biased estimators for fixed values of σ, γ, and p. Regarding the differences between restricted and unrestricted estimators, no correlation can be established between the number of observations or sample size and the degree of the unrestricted estimators' superiority over the restricted estimators. It means that the unrestricted estimators outperform the restricted biased estimators regardless of n. Generally as the sample size increases: Unrestricted estimators tend to perform better because they can utilize the full information in the data; the variance of the estimators tends to decrease, and the estimators become more efficient. As the sample size grows, the unrestricted estimators tend to converge to the true parameter values and exhibit less bias. Restricted estimators might perform worse in larger samples if the restrictions imposed are not true (e.g., incorrectly assuming that a coefficient is zero). The bias of the restricted estimator may not shrink with larger sample sizes, leading to inefficiency. However, if the restrictions are correct, restricted estimators can still be efficient, but this requires a strong assumption of correctness about the restrictions. Based on the Simulation results, the restricted estimators perform efficiently and even outperform the unrestricted estimators, which mean that the constraints we have chosen are correct.

6. Performance as a Function of $\boldsymbol{p}$*:* The mean square error of both groups, the restricted and unrestricted biased estimators' values grow as the number of regressors p increases for any given $n,\ \gamma,$ and σ. Furthermore, the increase in the number of regressors$p$ leads to a decrease in the regression coefficients values (Shaltoot, 2021).

5. CONCLUSION

In order to identify and compare the performance of linear estimators when the regression model suffers from the multicollinearity problem, this applied study has been conducted. And since many biased alternatives to the ordinary least squares have been suggested in literature with the aim of obtaining a substantial reduction in variance, many comparison studies emerged as a result. With the aim of making a different comparison we have decided to to separate some biased estimators into two main groups, the act that transfers the comparison to a radically different level of comparing between two linear regression models instead of only comparing the performance of a random group of estimators. These two models are the restricted linear regression model and the classical linear regression model. So, what differentiates this study is the comparison core itself, as well as the simulation study that was applied to it.

The OLS, the RR, the Liu, the TP, the RRR, the MRL and the RTP estimators were broadly presented, and their properties as well were discussed. Also, the issue of selecting the biasing parameters k and d have been considered. Based on the MMSE criterion, some theories of comparing the restricted and the unrestricted estimators have been represented and discussed. Then the theoretical findings have been illustrated in terms of both the SMSE and the MMSE criteria, by using a real-life data set followed by a comprehensive simulation study controlled by several dimensions: σ, γ, p, n, k and d. The commentary was considered separately under the previous core areas. From the real applications we may highlight the following results:

- Rather than the unrestricted estimators, the restricted estimators have been demonstrated as a noble alternative to the OLS in estimating when the problem of multicollinearity is existent in the linear regression model.

- In the sense of SMSE, the RRR estimator has been shown outperform all the restricted and unrestricted estimators, whereas the RR is superior to the unrestricted estimators set.
- The RR and RRR estimators of $\hat{k}_{AM}$ perform better comparing to the RR and RRR estimators of $\hat{k}_{HM}$ and $\hat{k}_{GM}$.

From the simulation study we may highlight the following results:

- The restricted biased estimators outperform the unrestricted biased estimators when the $\boldsymbol{X'X}$ is ill-conditioned in the linear regression model. On the other hand, In the absence of multicollinearity, the unrestricted estimators outperform the restricted estimators.
- Almost everywhere, the ridge regression estimator performs best in the unrestricted estimators' group and as the restricted ridge regression does in the restricted estimators' group.
- An increase in the value of the correlation coefficient results to an increase in the EMSE for all the estimators.
- As the standard deviation of the error grows so does the EMSE for each estimator.
- As k increases, the EMSE for each level of correlations decreases.
- As d increases, the EMSE for each level of correlations increases.
- An increase in the sample size n leads to a decrease in the EMSE for all the estimators.
- For every given n, $\gamma$and σ, the values of the restricted and unrestricted biased estimators grow as the number of regressors p increases.

Given that the results in both practical examples and the Monte-Carlo simulation showed the superiority of the restricted estimators to the unrestricted estimators, the researcher believes that restricted estimation as a contemporary science deserves more attention from statisticians. This research can be further developed by including more estimators in the comparison or by including the stochastic restricted estimators beside or instead of the exact restricted estimators.

ETHICAL DECLARATION

In the writing process of the study titled "Comparison of Restricted and Unrestricted estimators with a Monte-Carlo Simulation Study", there were followed the scientific, ethical and the citation rules; was not made any falsification on the collected data and this study was not sent to any other academic media for evaluation.

REFERENCES

Akdeniz, F., and Erol, H. (2003), Mean squared error matrix comparisons of some biased estimators in linear regression. *Communications in statistics-theory and methods*, 32(12), 2389-2413.

Akdeniz, F. and Kaçiranlar, S. (1995), On the almost unbiased generalized liu estimator and unbiased estimation of the bias and mse, *Communications in Statistics - Theory and Methods*, 24(7), 1789-1797.

Akdeniz, F., and Roozbeh, M. (2019), Generalized difference-based weighted mixed almost unbiased ridge estimator in partially linear models, *Statistical Papers*, 60(5), 1717-1739.

Albert, A. (1973), The Gauss–Markov theorem for regression models with possibly singular covariances, *SIAM Journal on Applied Mathematics*, 24(2), 182-187.

Aslam, M. (2014). Performance of Kibria's Method for the Heteroscedastic Ridge Regression Model: Some Monte Carlo Evidence, *Communications in Statistics - Simulation and Computation*, 43(4), 673-686.

Dorugade, A. V. (2014), A modified two-parameter estimator in linear regression, *Statistics in Transition. New Series*, 15(1), 23-36.

Gibbons, D. G. (1981), A simulation study of some ridge estimators, *Journal of the American Statistical Association*, 76(373), 131-139.

Graybill, F. A. (1976), Theory and application of the linear model, 183, Duxbury press North Scituate, MA.

Graybill, F. A. (1983),*Matrices with Applications in Statistics (Belmont, CA: Wadsworth)*, Graybill2Matrices With Applications in Statistics1983.

Groß, J. (2003), Restricted ridge estimation, *Statistics &Probability letters*, 65(1), 57-64.

Gruber, M. H. J. (1998),*Improving efficiency by shrinkage, Statistics: Textbooks and Monographs*,156. In: Marcel Dekker Inc., New York. Gruber, M. H. J., Regression estimators: A comparative study: JHU Press.

Hoerl, A. E., Kannard, R. W.and Baldwin, K. F. (1975), Ridge regression: Some simulations, *Communications in Statistics*, 4(2), 105-123.

Hoerl, A. E.,and Kennard, R. W. (1970a), Ridge Regression: Applications to Nonorthogonal Problems. *Technometrics*, 12(1), 69-82.

Hoerl, A. E.,and Kennard, R. W. (1970b), Ridge Regression: Biased Estimation for Nonorthogonal Problems, *Technometrics*, 12(1), 55-67.

Idowu, J. I., Oladapo, O. J., Owolabi, A. T., Ayinde, K. and Akinmoju, O. (2023), Combating multicollinearity: A new two-parameter approach, *Nicel Bilimler Dergisi*, 5(1), 90-116.

Kejian, L. (1993), A new class of blased estimate in linear regression, *Communications in Statistics - Theory and Methods*, 22(2), 393-402.

Kibria, B. M. G. (2003), Performance of Some New Ridge Regression Estimators. *Communications in Statistics-Simulation and Computation,* 32(2), 419-435.

Kibria, B. M. G., and Banik, S. (2016), Some ridge regression estimators and their performances. *Journal of Modern Applied Statistical Methods*, 15, 206-238.

Kibria, B. M., and Lukman, A. F. (2020), A new ridge-type estimator for the linear regression model: Simulations and applications, *Scientifica.*

Liu, K. (2003), Using Liu-Type Estimator to Combat Collinearity. *Communications in Statistics-Theory and Methods,* 32(5), 1009-1020.

Lukman, A. F., Ayinde, K., Aladeitan, B. and Bamidele, R. (2020), An unbiased estimator with prior information, *Arab Journal of Basic and Applied Sciences*, 27(1), 45-55.

Lukman, A. F., Ayinde, K., Binuomote, S. and Clement, O. A. (2019), Modified ridge -type estimator to combat multicollinearity: Application to chemical data. *Journal of Chemometrics*, 33(5), e3125.

Månsson, K. and Kibria, B. G. (2021), Estimating the unrestricted and restricted Liu estimators for the Poisson regression model: Method and application, *Computational Economics*, 58, 311-326.

Massy, W. F. (1965), Principal components regression in exploratory statistical research, *Journal of the American Statistical Association*, 60(309), 234-256.

McDonald, G. C., and Galarneau, D. I. (1975), A Monte Carlo Evaluation of Some Ridge type Estimators, *Journal of the American Statistical Association*, 70(350), 407- 416.

Montgomery, D. C., Peck, E. A.and Vining, G. G. (2012),*Introduction to linear regression analysis*,821, John Wiley & Sons.

Montgomery, D. C., Peck, E. A.and Vining, G. G. (2021), *Introduction to linear regression analysis,* John Wiley & Sons.

Muniz, G.and Kibria, B. M. G. (2009), On Some Ridge Regression Estimators: An Empirical Comparisons, *Communications in Statistics - Simulation and Computation*, 38(3), 621-630.

Najarian, S., Arashi, M.and Kibria, B. M. G. (2013), A Simulation Study on Some Restricted Ridge Regression Estimators, *Communications in Statistics - Simulation and Computation,* 42(4), 871-890.

Newhouse, J. P.and Oman, S. D. (1971),*An Evaluation Of Ridge Estimators.*

Özkale, M. R. (2014), The relative efficiency of the restricted estimators in linear regression models. *Journal of Applied Statistics,* 41(5), 998-1027.

Özkale, M. R., and Kaciranlar, S. (2007). The restricted and unrestricted two-parameter estimators, *Communications in Statistics—Theory and Methods*, 36(15), 2707- 2725.

Rao, C. R.and Toutenburg, H. (1995),*Linear models.* In Linear models (pp. 3-18): Springer.

Roozbeh, M. (2018), Optimal QR-based estimation in partially linear regression models with correlated errors using GCV criterion, *Computational Statistics & Data Analysis*, 117, 45-61.

Sakallıoğlu, S.and Kaçıranlar, S. (2008), A new biased estimator based on ridge estimation, *Statistical Papers,* 49(4), 669-689.

Shaltoot, I. (2021), *Comparison of restricted and unrestricted estimators in the multiple linear regression analysis,* Unpublished Master's thesis, Eskisehir Technical University.

Şiray, G. Ü. (2014). Modified and Restricted r-k Class Estimators. *Communications in Statistics-Theory and Methods*, 43(24), 5130-5155.

Stein, C. (1956, January). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proceedings of the Third Berkeley symposium on mathematical statistics and probability*, 1(1), 197-206.

Swindel, B. F. (1976), Good ridge estimators based on prior information, *Communications in Statistics - Theory and Methods, 5*(11), 1065-1075.

Toutenburg, H., and Trenkler, G. (1990). Mean square error matrix comparisons of optimal and classical predictors and estimators in linear regression, *Computational Statistics and Data Analysis*, 10(3), 297-305.

Trenkler, G. (1983), A note on superiority comparisons of homogeneous linear estimators. *Communications in statistics-theory and methods, 12*(7), 799-808.

Vinod, H. D., and Ullah, A. (1981),*Recent advances in regression methods,* 41, Marcel Dekker Incorporated.

Vizcarrondo, C., and Wallace, T. D. (1968), A test of the mean square error criterion for restrictions in linear regression, *Journal of the American Statistical Association,* 63(322), 558-572.

Wencheko, E. (2000), Estimation of the signal-to-noise in the linear regression model. *Statistical Papers, 41*(3), 327.

Woods, H., Steinour, H. H., and Starke, H. R. (1932), Effect of composition of Portland cement on heat evolved during hardening,*Industrial & Engineering Chemistry*, *24*(11), 1207-1214.

Yang, H., and Chang, X. (2010), A new two-parameter estimator in linear regression. *Communications in Statistics—Theory and Methods, 39*(6), 923-934.

Yang, H., and Cui, J. (2011), A stochastic restricted two-parameter estimator in linear regression model. *Communications in statistics-theory and methods*, *40*(13), 2318-2325.

Zhong, Z., andYang, H. (2007). Ridge estimation to the restricted linear model. *Communications in Statistics—Theory and Methods, 36*(11), 2099-2115.