AN OVERVIEW OF BIASED ESTIMATORS

Ng Set Foong[1], Low Heng Chin[2] and Quah Soon Hoe[2]

[1]Department of Information Technology and Quantitative Sciences,
Universiti Teknologi MARA, Jalan Permatang Pauh,
13500 Permatang Pauh, Pulau Pinang, Malaysia

[2]School of Mathematical Sciences, Universiti Sains Malaysia,
11800 USM, Pulau Pinang, Malaysia

*Corresponding author: hclow@cs.usm.my/shquah@gmail.com

Abstrak: *Penganggar pincang telah dicadangkan sebagai satu cara untuk meningkatkan kejituan anggaran parameter dalam model regresi apabila kekolinearan wujud dalam model tersebut. Sebab-sebab untuk menggunakan penganggar pincang telah dibincangkan dalam kertas kerja ini. Satu senarai penganggar-penganggar pincang juga dirumuskan dalam kertas kerja ini.*

Abstract: *Some biased estimators have been suggested as a means of improving the accuracy of parameter estimates in a regression model when multicollinearity exists. The rationale for using biased estimators instead of unbiased estimators when multicollinearity exists is given in this paper. A summary for a list of biased estimators is also given in this paper.*

Keywords: multicollinearity, regression, unbiased estimor

1. INTRODUCTION

When serious multicollinearity is detected in the data, some corrective actions should be taken in order to reduce its impact. The remedies for the problem of multicollinearity depend on the objective of the regression analysis. Multicollinearity causes no serious problem if the objective is to predict. However, multicollinearity is a problem when our primary interest is in the estimation of parameters.[1] The variances of parameter estimate, when multicollinearity exists, can become very large. Hence, the accuracy of the parameter estimate is reduced.

One obvious solution is to eliminate the regressors that are causing the multicollinearity. However, selecting regressors to delete for the purpose of removing or reducing multicollinearity is not a safe strategy. Even with extensive examination of different subsets of the available regressors, one might still select a subset of regressors that is far from optimal. This is because a small amount of sampling variability in the regressors or the dependent variable in a multicollinear data can result in a different subset being selected.[2]

An alternative to regressor deletion is to retain all of the regressors, and to use a biased estimator instead of a least squares estimator in the regression analysis. The least squares estimator is an unbiased estimator that is frequently used in the regression analysis. When the primary interest of the regression analysis is in the parameter estimation, some biased estimators have been suggested as a means to improve the accuracy of the parameter estimate in the model when multicollinearity exists.

The rationale for using biased estimators instead of unbiased estimators in a regression model when multicollinearity exists is presented in Section 2 while an overview of biased estimators is presented in Section 3. Some hybrids of the biased estimators are presented in Section 4. A comparison of the biased estimators is presented in Section 5.

2. THE RATIONALE FOR USING BIASED ESTIMATORS

Suppose there are n observations. A linear regression model with p standardized independent variables, $z_1, z_2, ..., z_p$, and a standardized dependent variable, y, can be written in the matrix form

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{Y}$ is an $n \times 1$ vector of standardized dependent variables, $\mathbf{Z}$ is an $n \times p$ matrix of standardized independent variables, $\boldsymbol{\gamma}$ is a $p \times 1$ vector of parameters, $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of errors such that $\boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0}, \sigma^2\mathbf{I}_n)$ and $\mathbf{I}_n$ is an identity matrix of dimension $n \times n$.

Let $\hat{\boldsymbol{\gamma}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}$ be the least squares estimator of the parameter $\boldsymbol{\gamma}$. The least squares estimator, $\hat{\boldsymbol{\gamma}}$, is an unbiased estimator of $\boldsymbol{\gamma}$ because the expected value of $\hat{\boldsymbol{\gamma}}$ is equal to $\boldsymbol{\gamma}$. Furthermore, it is the best linear unbiased estimator of the parameter, $\boldsymbol{\gamma}$.

Instead of using the least squares estimator, biased estimators are considered in the regression analysis in the presence of multicollinearity. When the expected value of the estimator is equal to the parameter which is supposed to be estimated, then the estimator is said to be unbiased; otherwise, it is said to be biased.

The mean squared error of an estimator is a measure of the goodness of the estimator. The least squares estimator (which is an unbiased estimator) has no bias. Thus, its mean squared error is equal to its variance. However, the variance of the least squares estimator may be very large in the presence of multicollinearity. Thus, its mean squared error may be unacceptably large, too. This would reduce the accuracy of parameter estimate in the regression model. Although the biased estimators have a certain amount of bias, it is possible for the variance of a biased estimator to be sufficiently smaller than the variance of the unbiased estimator to compensate for the bias introduced. Therefore, it is possible to find a biased estimator where its mean squared error is smaller than the mean squared error of the least squares estimator.[1] Hence, by allowing for some bias in the biased estimator, its smaller variance would lead to a smaller spread of the probability distribution of the estimator. Thus, the biased estimator is closer on average to the parameter being estimated.[1]

3. THE BIASED ESTIMATORS

There are several biased estimators that have been proposed as alternatives to the least squares estimator in the presence of multicollinearity. By combining these biased estimators, some hybrids of these biased estimators are formed. Before presenting the details of biased estimators, a linear regression model which is in canonical form is introduced.

Let $\boldsymbol{\lambda}$ be a $p \times p$ diagonal matrix whose diagonal elements are eigenvalues of $\mathbf{Z}'\mathbf{Z}$. The eigenvalues of $\mathbf{Z}'\mathbf{Z}$ are denoted by $\lambda_1, \lambda_2, ..., \lambda_p$. Let the matrix $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_p]$ be a $p \times p$ orthonormal matrix consisting of the p eigenvectors of $\mathbf{Z}'\mathbf{Z}$, where $\mathbf{t}_j$, $j = 1, 2, ..., p$, is the j-th eigenvector of $\mathbf{Z}'\mathbf{Z}$. Note that matrix $\mathbf{T}$ and matrix $\boldsymbol{\lambda}$ satisfy $\mathbf{T}'\mathbf{Z}'\mathbf{Z}\mathbf{T} = \boldsymbol{\lambda}$ and $\mathbf{T}'\mathbf{T} = \mathbf{T}\mathbf{T}' = \mathbf{I}$, where $\mathbf{I}$ is a $p \times p$ identity matrix. By using matrix $\boldsymbol{\lambda}$ and matrix $\mathbf{T}$, the linear regression model, $\mathbf{Y} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$, as given by equation (1), can be transformed into a canonical form

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2}$$

where $\mathbf{X} = \mathbf{Z}\mathbf{T}$ is an $n \times p$ matrix, $\boldsymbol{\beta} = \mathbf{T}'\boldsymbol{\gamma}$ is a $p \times 1$ vector of parameters and $\mathbf{X}'\mathbf{X} = \boldsymbol{\lambda}$.

The least squares estimator of the parameter $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \tag{3}$$

The least squares estimator, $\hat{\boldsymbol{\beta}}$, is an unbiased estimator of $\boldsymbol{\beta}$ and is often called the Ordinary Least Squares Estimator (OLSE) of parameter $\boldsymbol{\beta}$.

In the presence of multicollinearity, biased estimators are proposed as alternatives to the OLSE (which is an unbiased estimator) in order to increase the accuracy of the parameter estimate. The details of these biased estimators are given below. The Principal Component Regression Estimator (PCRE) is one of the proposed biased estimators. The PCRE is also known as the Generalized Inverse Estimator.[3–6] Principal component regression approaches the problem of multicollinearity by dropping the dimension defined by a linear combination of the independent variables but not by a single independent variable. The idea behind principal component regression is to eliminate those dimensions that cause multicollinearity. These dimensions usually correspond to eigenvalues that are very small. The PCRE of parameter $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}}_r = \mathbf{T}_r'\hat{\boldsymbol{\gamma}}_r \tag{4}$$

where $\hat{\boldsymbol{\gamma}}_r = \mathbf{T}_r(\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r)^{-1}\mathbf{T}_r'\mathbf{Z}'\mathbf{Y}$, is the PCRE of parameter $\boldsymbol{\gamma}$, $\mathbf{T}_r = (\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_r)$ is the matrix of the remaining eigenvectors of $\mathbf{Z}'\mathbf{Z}$ after we have deleted $p-r$ of the columns of $\mathbf{T}$ and it satisfies $\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r = \boldsymbol{\lambda}_r = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_p)$.

The Shrunken Estimator, or the Stein Estimator, is another biased estimator. It was proposed by Stein.[7,8] It is further discussed by Sclove (1968)[9] and Mayer and Willke.[10] The Shrunken Estimator is given by

$$\hat{\boldsymbol{\beta}}_s = s\hat{\boldsymbol{\beta}} \tag{5}$$

where $0 < s < 1$.

Trenkler proposed the Iteration Estimator.[11] The Iteration Estimator is given by

$$\hat{\boldsymbol{\beta}}_{m,\delta} = \mathbf{X}_{m,\delta}\mathbf{Y} \tag{6}$$

where the series $\mathbf{X}_{m,\delta} = \delta\sum_{i=0}^{m}(\mathbf{I} - \delta\mathbf{X}'\mathbf{X})^{i}\mathbf{X}'$, $m = 0, 1, 2, \ldots$, $0 < \delta < \frac{1}{\lambda_{\max}}$ and $\lambda_{\max}$ refers to the largest eigenvalue.

Trenkler stated that $\mathbf{X}_{m,\delta}$ converges to the Moore-Penrose inverse $\mathbf{X}^{+} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ of $\mathbf{X}$.[11]

Due to the fact that the least squares estimator based on minimum residual sum of squares has a high probability of being unsatisfactory when multicollinearity exists in the data, Hoerl and Kennard proposed the Ordinary Ridge Regression Estimator (ORRE) and the Generalized Ridge Regression Estimator (GRRE).[12,13] The proposed estimation procedure is based on adding small positive quantities to the diagonal of $\mathbf{X}'\mathbf{X}$. The GRRE is given by

$$\hat{\boldsymbol{\beta}}_{\mathbf{K}} = (\mathbf{X}'\mathbf{X} + \mathbf{K})^{-1}\mathbf{X}'\mathbf{Y} \tag{7}$$

where $\mathbf{K} = \mathrm{diag}(k_i)$ is a diagonal matrix of biasing factors $k_i > 0,\ i = 1, 2, \ldots, p$. When all diagonal elements of the matrix, $\mathbf{K}$, in the GRRE have values that are equal to k, the GRRE can be written as the ORRE. The ORRE Estimator is given by

$$\hat{\boldsymbol{\beta}}_{k} = (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}'\mathbf{Y} \tag{8}$$

where $k > 0$.

Authors proved that the ORRE has a smaller mean squared error compared to the OLSE.[12] The following existence theorem is stated in their paper, "There always exists a $k > 0$ such that the mean squared error of $\hat{\boldsymbol{\beta}}_{k}$ is less than the mean squared error of $\hat{\boldsymbol{\beta}}$". There is also an equivalent existence theorem for the GRRE.[12]

The ORRE and the GRRE turn out to be popular biased estimators. Many studies based on the ORRE and the GRRE have been done since the work of Hoerl and Kennard.[12,13] Some methods have been proposed for choosing the value of k.[14,15] In 1986, Singh et al.[16] proposed the Almost Unbiased Generalized Ridge Regression Estimator (AUGRRE) by using the jack-knife procedure. This estimator reduces the bias uniformly for all components of the parameter vector. The AUGRRE is given as

$$\tilde{\boldsymbol{\beta}}^{*}_{\mathbf{K}} = (\mathbf{I} - (\mathbf{X}'\mathbf{X} + \mathbf{K})^{-2}\mathbf{K}^{2})\hat{\boldsymbol{\beta}} \tag{9}$$

where $\mathbf{K} = \mathrm{diag}(k_i)$, $k_i > 0,\ i = 1,\ 2,\ ...,\ p$.

In the case where all diagonal elements of the matrix, $\mathbf{K}$, in the AUGRRE have values that are equal to k, then we may write the Almost Unbiased Ridge Regression Estimator (AURRE) as[17]

$$\tilde{\boldsymbol{\beta}}_k^* = (\mathbf{I} - (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-2} k^2)\hat{\boldsymbol{\beta}} \tag{10}$$

where $k > 0$.

On the other hand, Akdeniz *et al.*[18](2004) derived general expressions for the moments of the Lawless and Wang operational AURRE for individual regression coefficients.[18,19]

There are some other biased estimators developed based on the ORRE, such as the Modified Ridge Regression Estimator (MRRE) introduced by Swindel.[20,21] and the Restricted Ridge Regression Estimator (RRRE) proposed by Sarkar[22,23] The MRRE and the RRRE are given in equations (11) and (12), respectively.

$$\mathbf{b}(k, \mathbf{b}^*) = (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}(\mathbf{X}'\mathbf{Y} + k\mathbf{b}^*) \tag{11}$$

where $\mathbf{b}^*$ is a prior mean and it is assumed that $\mathbf{b}^* \neq \hat{\boldsymbol{\beta}}$, $k > 0$.

$$\boldsymbol{\beta}^*(k) = [\mathbf{I} + k(\mathbf{X}'\mathbf{X})^{-1}]^{-1}\boldsymbol{\beta}^* \tag{12}$$

where $k > 0$, $\boldsymbol{\beta}^* = \hat{\boldsymbol{\beta}} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})$ is the restricted least squares estimator and the set of linear restrictions on the parameters are represented by $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$.

4. HYBRIDS OF THE BIASED ESTIMATORS

Biased estimators have been proposed as alternatives to the OLSE when multicollinearity exists in the data. Major types of the proposed biased estimators are the PCRE, the Shrunken Estimator, the Iteration Estimator, the ORRE and the GRRE. Some studies have been done on combining the biased estimators. Thus, some hybrids of these biased estimators have been proposed.

Baye and Parker proposed the $r-k$ Class Estimator which combined the techniques of the ORRE and the PCRE.[24] They proved that there exists a $k > 0$

where the mean squared error of the $r-k$ Class Estimator is smaller than the mean squared error of the PCRE. The $r-k$ Class Estimator of parameter $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}}_r(k) = \mathbf{T}_r'[\hat{\boldsymbol{\gamma}}_r(k)] \tag{13}$$

where $r \le p,\ k > 0,\ \hat{\boldsymbol{\gamma}}_r(k) = \mathbf{T}_r(\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r + k\mathbf{I}_r)^{-1}\mathbf{T}_r'\mathbf{Z}'\mathbf{Y}$ is the $r-k$ Class Estimator of parameter $\boldsymbol{\gamma}$, $\mathbf{T}_r$ is the remaining eigenvectors of $\mathbf{Z}'\mathbf{Z}$ after having deleted $p-r$ of the columns of $\mathbf{T}$ and satisfying $\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r = \boldsymbol{\lambda}_r = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_p)$.

Liu introduced a biased estimator by combining the advantages of the ORRE and the Shrunken Estimator.[25] This new biased estimator is known as the Liu Estimator. The Liu Estimator can also be generalized to the Generalized Liu Estimator (GLE). The Liu Estimator and the GLE are given in equations (14) and (15), respectively.

$$\hat{\boldsymbol{\beta}}_d = (\mathbf{X}'\mathbf{X} + \mathbf{I})^{-1}(\mathbf{X}'\mathbf{Y} + d\hat{\boldsymbol{\beta}}) \tag{14}$$

where $0 < d < 1$.

$$\hat{\boldsymbol{\beta}}_{\mathbf{D}} = (\mathbf{X}'\mathbf{X} + \mathbf{I})^{-1}(\mathbf{X}'\mathbf{Y} + \mathbf{D}\hat{\boldsymbol{\beta}}) \tag{15}$$

where $\mathbf{D} = \mathrm{diag}(d_i)$ is a diagonal matrix of the biasing factors, d_i, and $0 < d_i < 1$, $i = 1, 2, ..., p$.

When all the diagonal elements of matrix $\mathbf{D}$ in the GLE have values that are equal to d, the GLE can be written as the Liu Estimator. Liu showed that the Liu Estimator is preferable to the OLSE in terms of the mean squared error criterion.[25] The advantage of the Liu Estimator over the ORRE is that the Liu Estimator is a linear function of d. Hence, it is easy to choose d. Recently, Akdeniz and Ozturk derived the density function of the stochastic shrinkage parameters of the operational Liu Estimator by assuming normality.[26]

Some studies based on the Liu Estimator and the GLE have been done. Akdeniz and Kaciranlar introduced the Almost Unbiased Generalized Liu Estimator (AUGLE).[21] This estimator is a bias corrected GLE. When all the diagonal elements of the matrix $\mathbf{D}$ in the AUGLE have values that are equal to *d*, then the Almost Unbiased Generalized Liu Estimator can be written as the Almost Unbiased Liu Estimator (AULE).[17] The AUGLE and the AULE are given by equations (16) and (17), respectively.

$$\tilde{\boldsymbol{\beta}}_{\mathbf{D}}^{*} = [\mathbf{I} - (\mathbf{X}'\mathbf{X} + \mathbf{I})^{-2}(\mathbf{I} - \mathbf{D})^2]\hat{\boldsymbol{\beta}} \tag{16}$$

where $\mathbf{D} = \text{diag}(d_i)$ and $0 < d_i < 1$, $i = 1, 2, ..., p$.

$$\tilde{\boldsymbol{\beta}}_d^* = [\mathbf{I} - (\mathbf{X}'\mathbf{X} + \mathbf{I})^{-2}(1-d)^2]\hat{\boldsymbol{\beta}} \tag{17}$$

where $0 < d < 1$.

Kaciranlar et al. introduced a new estimator by replacing the OLSE, $\hat{\boldsymbol{\beta}}$, in the Liu Estimator, by the restricted least squares estimator, $\boldsymbol{\beta}^*$.[27] They called it the Restricted Liu Estimator (RLE) and it is given as

$$\hat{\boldsymbol{\beta}}_{rd} = (\mathbf{X}'\mathbf{X} + \mathbf{I})^{-1}(\mathbf{X}'\mathbf{X} + d\mathbf{I})\boldsymbol{\beta}^* \tag{18}$$

where $\boldsymbol{\beta}^* = \hat{\boldsymbol{\beta}} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})$ is the restricted least squares estimator and the set of linear restrictions on the parameters are represented by $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$.

In 2001, Kaciranlar and Sakallioglu[28] proposed the $r-d$ Class Estimator by combining the Liu Estimator and the PCRE. The $r-d$ Class Estimator is a general estimator which includes the OLSE, the PCRE and the Liu Estimator as a special case. Kaciranlar and Sakallioglu have shown that the $r-d$ Class Estimator is superior to the PCRE in terms of mean squared error.[28] The $r-d$ Class Estimator of parameter $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}}_r(d) = \mathbf{T}_r'[\hat{\boldsymbol{\gamma}}_r(d)] \tag{19}$$

where $r \le p,\ 0 < d < 1,\ \hat{\boldsymbol{\gamma}}_r(d) = \mathbf{T}_r(\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r + \mathbf{I}_r)^{-1}(\mathbf{T}_r'\mathbf{Z}'\mathbf{Y} + d\mathbf{T}_r'\hat{\boldsymbol{\gamma}}_r)$ is the $r-d$ Class Estimator of parameter $\boldsymbol{\gamma}$, $\hat{\boldsymbol{\gamma}}_\mathbf{r} = \mathbf{T}_\mathbf{r}(\mathbf{T}_\mathbf{r}'\mathbf{Z}'\mathbf{Z}\mathbf{T}_\mathbf{r})^{-1}\mathbf{T}_\mathbf{r}'\mathbf{Z}'\mathbf{Y}$ is the PCRE of parameter $\boldsymbol{\gamma}$, $\mathbf{T}_r$ is the remaining eigenvectors of $\mathbf{Z}'\mathbf{Z}$ after having deleted $p-r$ of the columns of $\mathbf{T}$ and satisfying $\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r = \boldsymbol{\lambda}_r = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_p)$.

Table 1 displays a matrix showing the biased estimators and the hybrids that have been proposed. The hybrids that have been proposed are the $r-k$ Class Estimator, the Liu Estimator and the $r-d$ Class Estimator. The Liu Estimator combines the advantages of the ORRE and the Shrunken Estimator. The $r-k$ Class Estimator combined the techniques of the ORRE and the PCRE while the $r-d$ Class Estimator combined the techniques of the Liu Estimator and the PCRE. There are some biased estimators developed based on the ORRE, the GRRE, the Liu Estimator and the GLE. The MRRE, the RRE, the AUGRRE and the AURRE are the biased estimators developed based on the ORRE and the

GRRE while the AUGLE, the AULE and the RLE were developed based on the Liu Estimator and the GLE. The equations for the biased estimators presented in Sections 3 and 4 are summarized in Table 2.

Table 1: Matrix of the biased estimators and the hybrids.

	PCRE	GRRE, ORRE	Shrunken Estimator	Iteration Estimator	GLE, Liu Estimator	*r-k* Class Estimator	*r-d* Class Estimator
PCRE		*r-k* Class Estimator			*r-d* Class Estimator		
GRRE, ORRE		MRRE, RRRE, AUGRRE, AURRE	Liu Estimator				
Shrunken Estimator							
Iteration Estimator							
GLE, Liu Estimator					AUGLE, AULE, RLE		
r-k Class Estimator							
r-d Class Estimator							

5. REVIEW ON THE COMPARISONS BETWEEN THE BIASED ESTIMATORS

The comparisons among the biased estimators as well as the OLSE are found in several papers. Most of the comparisons were done in terms of the mean squared error. An estimator is superior to the another if its mean squared error is less than the other.

Table 2: Summary of a list of estimators.

No.	Estimators*	Equation	Relevant References
1	OLSE	$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$	Belsley 1991[34]
2	PCRE	$\hat{\boldsymbol{\beta}}_r = \mathbf{T}_r'\hat{\boldsymbol{\gamma}}_r$ where $\hat{\boldsymbol{\gamma}}_r = \mathbf{T}_r(\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r)^{-1}\mathbf{T}_r'\mathbf{Z}'\mathbf{Y}$, is the PCRE of parameter $\boldsymbol{\gamma}$, $\mathbf{T}_r = [\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_r]$ is the remaining eigenvectors of $\mathbf{Z}'\mathbf{Z}$ after having deleted $p-r$ of the columns of $\mathbf{T}$ and satisfying $\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r = \boldsymbol{\lambda}_r = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_p)$	Massy 1965; Marquardt 1970; Hawkins 1973; Greenberg 1975
3	Shrunken Estimator	$\hat{\boldsymbol{\beta}}_s = s\hat{\boldsymbol{\beta}}$ where $0 < s < 1$	Stein 1960; cited by Hocking et al. 1976; Sclove 1968; Mayer & Willke 1973
4	Iteration Estimator	$\hat{\boldsymbol{\beta}}_{m,\delta} = \mathbf{X}_{m,\delta}\mathbf{Y}$ where the series $\mathbf{X}_{m,\delta} = \delta\sum_{i=0}^{m}(\mathbf{I}-\delta\mathbf{X}'\mathbf{X})^i\mathbf{X}'$, $m = 0, 1, 2, \ldots$, $0 < \delta < \frac{1}{\lambda_{\max}}$ and $\lambda_{\max}$ refers to the largest eigenvalue	Trenkler 1978
5	GRRE	$\hat{\boldsymbol{\beta}}_{\mathbf{K}} = (\mathbf{X}'\mathbf{X}+\mathbf{K})^{-1}\mathbf{X}'\mathbf{Y}$ where $\mathbf{K} = \text{diag}(k_i)$ is a diagonal matrix with biasing factors $k_i > 0,\ i = 1, 2, ..., p$	Hoerl & Kennard 1970a,b
6	ORRE	$\hat{\boldsymbol{\beta}}_k = (\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}\mathbf{X}'\mathbf{Y}$ where $k > 0$	Hoerl & Kennard 1970a,b
7	AUGRRE	$\tilde{\boldsymbol{\beta}}_{\mathbf{K}}^* = (\mathbf{I}-(\mathbf{X}'\mathbf{X}+\mathbf{K})^{-2}\mathbf{K}^2)\hat{\boldsymbol{\beta}}$ where $\mathbf{K} = \text{diag}(k_i)$, $k_i > 0,\ i = 1, 2, ..., p$	Singh et al. 1986
8	AURRE	$\tilde{\boldsymbol{\beta}}_k^* = (\mathbf{I}-(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-2}k^2)\hat{\boldsymbol{\beta}}$ where $k > 0$	Akdeniz & Erol 2003

(continue on next page)

Table 2: *(continued)*

No.	Estimators*	Equation	Relevant References
9	MRRE	$\mathbf{b}(k,\mathbf{b}^*) = (\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}(\mathbf{X}'\mathbf{Y}+k\mathbf{b}^*)$ where $\mathbf{b}^*$ is a prior mean and it is assumed that $\mathbf{b}^* \neq \hat{\boldsymbol{\beta}}$, $k > 0$	Swindel 1976; cited by Akdeniz & Kaciranlar 1995
10	RRRE	$\boldsymbol{\beta}^*(k) = [\mathbf{I}+k(\mathbf{X}'\mathbf{X})^{-1}]^{-1}\boldsymbol{\beta}^*$ where $k > 0$, $\boldsymbol{\beta}^* = \hat{\boldsymbol{\beta}} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{r}-\mathbf{R}\hat{\boldsymbol{\beta}})$ is the restricted least squares estimator and the set of linear restrictions on the parameters are represented by $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$	Sarkar, 1992; cited by Kaciranlar et al. 1998
11	$r-k$ Class Estimator	$\hat{\boldsymbol{\beta}}_r(k) = \mathbf{T}_r'[\hat{\boldsymbol{\gamma}}_r(k)]$ where $r \leq p$, $k > 0$, $\hat{\gamma}_r(k) = \mathbf{T}_r(\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r + k\mathbf{I}_r)^{-1}\mathbf{T}_r'\mathbf{Z}'\mathbf{Y}$ is the $r-k$ Class Estimator of parameter $\boldsymbol{\gamma}$, $\mathbf{T}_r$ is the remaining eigenvectors of $\mathbf{Z}'\mathbf{Z}$ after having deleted $p-r$ of the columns of $\mathbf{T}$ and satisfying $\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r = \boldsymbol{\lambda}_r = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_p)$	Baye & Parker 1984
12	GLE	$\hat{\boldsymbol{\beta}}_{\mathbf{D}} = (\mathbf{X}'\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}'\mathbf{Y}+\mathbf{D}\hat{\boldsymbol{\beta}})$ where $\mathbf{D} = \text{diag}(d_i)$ is a diagonal matrix of biasing factors d_i and $0 < d_i < 1$, $i = 1, 2, ..., p$	Liu 1993
13	Liu Estimator	$\hat{\boldsymbol{\beta}}_d = (\mathbf{X}'\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}'\mathbf{Y}+d\hat{\boldsymbol{\beta}})$ where $0 < d < 1$	Liu 1993
14	AUGLE	$\tilde{\boldsymbol{\beta}}^*_{\mathbf{D}} = [\mathbf{I}-(\mathbf{X}'\mathbf{X}+\mathbf{I})^{-2}(\mathbf{I}-\mathbf{D})^2]\hat{\boldsymbol{\beta}}$, where $\mathbf{D} = \text{diag}(d_i)$ and $0 < d_i < 1$, $i = 1, 2, ..., p$	Akdeniz & Kaciranlar 1995
15	AULE	$\tilde{\boldsymbol{\beta}}^*_d = [\mathbf{I}-(\mathbf{X}'\mathbf{X}+\mathbf{I})^{-2}(1-d)^2]\hat{\boldsymbol{\beta}}$ where $0 < d < 1$	Akdeniz & Erol 2003

(continue on next page)

Tabel 2: *(continued)*

No.	Estimators*	Equation	Relevant References
16	RLE	$\hat{\boldsymbol{\beta}}_{rd} = (\mathbf{X}'\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}'\mathbf{X}+d\mathbf{I})\boldsymbol{\beta}^*$ where $\boldsymbol{\beta}^* = \hat{\boldsymbol{\beta}} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{r}-\mathbf{R}\hat{\boldsymbol{\beta}})$ is the restricted least squares estimator and the set of linear restrictions on the parameters are represented by $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$	Kaciranlar et al. 1999
17	$r-d$ Class Estimator	$\hat{\boldsymbol{\beta}}_r(d) = \mathbf{T}_r'[\hat{\boldsymbol{\gamma}}_r(d)]$ where $r \le p,\ 0 < d < 1$, $\hat{\boldsymbol{\gamma}}_r(d) = \mathbf{T}_r(\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r + \mathbf{I}_r)^{-1}(\mathbf{T}_r'\mathbf{Z}'\mathbf{Y} + d\mathbf{T}_r'\hat{\boldsymbol{\gamma}}_r)$ is the $r-d$ Class Estimator of parameter $\boldsymbol{\gamma}$, $\hat{\boldsymbol{\gamma}}_\mathbf{r} = \mathbf{T_r}(\mathbf{T_r'Z'ZT_r})^{-1}\mathbf{T_r'Z'Y}$ is the PCRE of parameter $\boldsymbol{\gamma}$, $\mathbf{T}_r$ is the remaining eigenvectors of $\mathbf{Z}'\mathbf{Z}$ after having deleted $p-r$ of the columns of $\mathbf{T}$ and satisfying $\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r = \boldsymbol{\lambda}_r = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_p)$	Kaciranlar & Sakallioglu 2001

* No. 1 is an unbiased estimator while No.2 – No. 17 are biased estimators

However, Singh et al.[16] compared the GRRE and the AUGRRE in terms of bias. It is found that there is a reduction in the bias of the AUGRRE when compared with the bias of the GRRE in terms of absolute value.

Table 3 gives a summary of the comparisons among the biased estimators and the OLSE (which is an unbiased estimator) while Table 4 gives the relevant references of the comparisons.

Hoerl and Kennard compared the OLSE, $\hat{\boldsymbol{\beta}}$, with the ORRE, $\hat{\boldsymbol{\beta}}_k$, and the GRRE, $\hat{\boldsymbol{\beta}}_\mathbf{K}$.[12] It is found that there exists a $k > 0$ such that the mean squared error of $\hat{\boldsymbol{\beta}}_k$ is less than the mean squared error of $\hat{\boldsymbol{\beta}}$. There is also an equivalent existence theorem for the GRRE.[12]

Trenkler compared the Iteration Estimator $\hat{\boldsymbol{\beta}}_{m,\delta}$ and the OLSE, $\hat{\boldsymbol{\beta}}$.[11,29] It was found that the mean squared error of $\hat{\boldsymbol{\beta}}_{m,\delta}$ is less than the mean squared error of $\hat{\boldsymbol{\beta}}$.

Table 3: Summary of the comparisons among the estimators.

	OLSE	PCRE	Shrunken	Iteration Estimator	GRRE, ORRE	UGRRE, AUREE	MRRE	RRRE	*r-k* Class Etimator	GLE, Liu Estimator	AUGLE, AULE	RLE	*r-d* Class Etimator
OLSE				ii, iii	i	vii			x	viii	ix		xiv
PCRE				iii					iv, x				xiv
Shrunken Estimator				iii									
Iteration Estimator					iii					xiii			
GRRE, ORRE						v, vii	vi		x	xiii, xv			
AUGRRE, AURRE											xv		
MRRE								xi					
RRRE													
r-k Class Estimator													
GLE, Liu Estimator											ix	xii	xiv
AUGLE, AULE													
RLE													
r-d Class Estimator													

In 1984, Baye and Parker[24] compared the $r-k$ Class Estimator, $\hat{\gamma}_r(k)$, with the PCRE, $\hat{\gamma}_r$. They showed that there exists a $k>0$ such that the mean squared error of $\hat{\gamma}_r(k)$ is less than the mean squared error of $\hat{\gamma}_r$ for $0<r\le p$.

Table 4: References for the comparisons among the estimators.

References	Comparison between the Estimators
(i) Hoerl & Kennard 1970a	(a) OLSE and ORRE (b) OLSE and GRRE
(ii) Trenkler 1978	(a) Iteration Estimator and OLSE
(iii) Trenkler 1980	(a) Iteration Estimator and OLSE (b) Iteration Estimator and ORRE (c) Iteration Estimator and Shrunken Estimator (d) Iteration Estimator and PCRE
(iv) Baye & Parker 1984	(a) PCRE and $r-k$ Class Estimator
(v) Singh et al. 1986	(a) GRRE and AUGRRE
(vi) Pliskin 1987	(a) MRRE and ORRE
(vii) Nomura 1988	(a) GRRE and AUGRRE (b) OLSE and AUGRRE
(viii) Liu 1993	(a) OLSE and Liu Estimator
(ix) Akdeniz & Kaciranlar 1995	(a) GLE and AUGLE (b) OLSE and AUGLE
(x) Sarkar 1996	(a) $r-k$ Class Estimator and PCRE (b) $r-k$ Class Estimator and OLSE (c) $r-k$ Class Estimator and ORRE
(xi) Kaciranlar et al. 1998	(a) MRRE and RRRE
(xii) Kaciranlar et al. 1999	(a) RLE and Liu Estimator
(xiii) Sakallioglu et al. 2001	(a) ORRE and Liu Estimator (b) Iteration Estimator and Liu Estimator
(xiv) Kaciranlar & Sakallioglu 2001	(a) $r-d$ Class Estimator and PCRE (b) $r-d$ Class Estimator and Liu Estimator (c) $r-d$ Class Estimator and OLSE
(xv) Akdeniz & Erol 2003	(a) GRRE and GLE (b) AUGRRE and AUGLE

A comparison between the MRRE, $\mathbf{b}(k,\mathbf{b}^*)$, and the ORRE, $\hat{\boldsymbol{\beta}}_k$, was done by Pliskin.[30] A necessary and sufficient condition for the mean squared error matrix of $\hat{\boldsymbol{\beta}}_k$ minus the mean squared error matrix of $\mathbf{b}(k,\mathbf{b}^*)$ to be positive semidefinite when both estimators are computed using the same value of k was developed. The author suggested that researchers who are inclined to use the conventional ORRE should consider the MRRE if prior information is available.

Liu made a comparison between the OLSE, $\hat{\boldsymbol{\beta}}$, and the Liu Estimator, $\hat{\boldsymbol{\beta}}_d$.[25] He showed that there exists $0 < d < 1$ such that the mean squared error of $\hat{\boldsymbol{\beta}}_d$ is less than mean squared error of $\hat{\boldsymbol{\beta}}$.

A comparison between the $r-k$ Class Estimator and the OLSE, the PCRE, the ORRE was done by Sarkar (1996).[31] Necessary and sufficient conditions for the superiority of the $r-k$ Class Estimator over each of the other three estimators using the mean squared error matrix criterion were obtained.

Kaciranlar et al.[23] compared the RRRE and the MRRE. They proved that the RRRE is superior to the MRRE using the mean squared error matrix criterion whether the linear restrictions are true or not. Kaciranlar et al.[27] introduced the RLE and showed that the RLE is superior in the scalar mean squared error sense, to both the restricted least squares estimator and to the Liu Estimator when the restrictions are indeed correct. They also derived conditions for the superiority of the RLE over both the restricted least squares estimator and the Liu Estimator when the restrictions are not correct.

Kaciranlar and Sakallioglu made a comparison between the $r-d$ Class Estimator, $\hat{\boldsymbol{\gamma}}_r(d)$, with the PCRE, $\hat{\boldsymbol{\gamma}}_r$, the Liu Estimator and the OLSE respectively.[28] They showed that there exist $0 < d < 1$ such that the mean squared error of $\hat{\boldsymbol{\gamma}}_r(d)$ is less than the mean squared error of $\hat{\boldsymbol{\gamma}}_r$. The comparisons between the $r-d$ Class Estimator and the Liu Estimator as well as the $r-d$ Class Estimator with the OLSE show that which estimator is better depends on the unknown parameters, the variance of the error term in the linear regression model and the choice of biased factor, d, in the biased estimators.

In addition, there are also several comparisons in terms of mean squared error which produced similar conclusions. Trenkler compared the Iteration Estimator with the ORRE, the Shrunken Estimator and the PCRE respectively.[29] Nomura compared the AUGRRE with the GRRE and the OLSE respectively.[32] Akdeniz and Kaciranlar made a comparison between the GLE and the AUGLE.[21] They also compared the OLSE and the AUGLE. Recently, Sakallioglu et al.[33] compared the Liu Estimator with the ORRE and the Iteration Estimator respectively. Akdeniz and Erol made a comparison between the GRRE and the GLE.[17] They also compared the AUGRRE and the AUGLE. These comparisons showed that the better estimator depends on the unknown parameters, the variance of the error term in the linear regression model and the choice of the biasing factors in biased estimators.

6. CONCLUSION

Multicollinearity is one of the problems that arise in regression analysis. Thus, multicollinearity diagnostics should be carried out to detect the problem of multicollinearity in the data. The remedies for the problem of multicollinearity depend on the objective of the regression analysis. Multicollinearity causes no serious problems if the objective is prediction. However, multicollinearity is a problem when the primary interest is in the estimation of the parameters in a regression model.

In the presence of multicollinearity, the minimum variance of the least squares estimator may be unacceptably large and hence reduces the accuracy of the parameter estimates. Some biased estimators have been suggested as a means to improve the accuracy of the parameter estimate in the model when multicollinearity exists. There are several biased estimators that have been proposed, such as the PCRE, the Shrunken Estimator, the Iteration Estimator, the ORRE and the GRRE. In addition, the MRRE, the RRRE, the AUGRRE and the AUGRRE are biased estimators developed based on the ORRE and the GRRE.

By combining these biased estimators, some hybrids of these biased estimators, such as the $r-k$ Class Estimator, the Liu Estimator, the GLE and the $r-d$ Class Estimator are obtained. Furthermore, the AUGLE, the AUGLE and the RLE were developed based on the Liu Estimator and the GLE.

From most of the comparisons between the biased estimators,[17,21,29,32,33] we find that the better estimator depends on the unknown parameters and the variance of error term in the linear regression model as well as the choice of the biased factors in biased estimators. Therefore, there is still room for improvement where new classes of biased estimators could be developed in order to provide a better solution.

7. REFERENCES

1. Rawlings, J.O., Pantula, S.G. & Dickey, D.A. (1998). *Applied regression analysis–A research tool*. New York: Springer-Verlag.
2. Ryan, T.P. (1997). *Modern regression methods*. New York: John Wiley & Sons, Inc.
3. Massy, W.F. (1965). Principal components regression in exploratory statistical research. *Journal of the American Statistical Association*, 60, 234–246.
4. Marquardt, D.W. (1970). Generalized inverses, ridge regression, biased linear estimation and nonlinear estimation. *Technometrics*, 12, 591–612.

5. Hawkins, D.M. (1973). On the investigation of alternative regressions by principal component analysis. *Applied Statistics*, 22, 275–286.
6. Greenberg, E. (1975). Minimum variance properties of principal component regression. *Journal of the American Statistical Association*, 70, 194–197.
7. Stein, C.M. (1960). Multiple regression. In I. Ikin (Ed.). *Contributions to probability and statistics: Essays in Honor of Harold Hotelling.* CA: Stanford University Press, 424–443
8. Hocking, R.R., Speed, F.M. & Lynn, M.J. (1976). A class of biased estimators in linear regression. *Technometrics*, 18, 425–437.
9. Sclove, S.L. (1968). Improved estimators for coefficients in linear regression. *Journal of the American Statistical Association*, 63, 597–606.
10. Mayer, L.S. & Willke, T.A. (1973). On biased estimation in linear models. *Technometrics*, 15, 497–508.
11. Trenkler, G. (1978). An iteration estimator for the linear model. *Compstat*, 125–131.
12. Hoerl, A.E. & Kennard, R.W. (1970a). Ridge regression: Biased estimation for non-orthogonal problems. *Technometrics*, 12, 55–67.
13. Hoerl, A.E. & Kennard, R.W. (1970b). Ridge regression: Applications to non-orthogonal problems. *Technometrics*, 12, 69–82.
14. McDonald, G.C. & Galarneau, D.I. (1975). A Monte Carlo evaluation of some ridge type estimators. *Journal of the American Statistical Association*, 70, 407–416.
15. Hemmerle, W.J. & Brantle, T.F. (1978). An explicit and constrained generalized ridge estimation. *Technometrics*, 20, 109–120.
16. Singh, B., Chaubey, Y.P. & Dwivedi, T.D. (1986). An almost unbiased ridge estimator. *Sankhya: The Indian Journal of Statistics*, 48(B), 342–346.
17. Akdeniz, F. & 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.
18. Akdeniz, F., Yüksel, G. & Wan, A.T.K. (2004). The moments of the operational almost unbiased ridge regression estimator. *Applied Mathematics and Computation*, 153(3), 673–684.
19. Lawless, J.F. & Wang, P. (1976). A simulation study of ridge and other regression estimators. *Communications in Statistics-Theory and Methods*, A5, 307–323.
20. Swindel, B.F. (1976). Good ridge estimators based on prior information. *Communications in Statistics-Theory and Methods*, A5, 1065–1075.
21. Akdeniz, F. & Kaciranlar, 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.

22. Sarkar, N. (1992). A new estimator combining the ridge regression and the restricted least squares methods of estimation. *Communications in Statistics-Theory and Methods*, 21(7), 1987–2000.
23. Kaciranlar, S., Sakallioglu, S. & Akdeniz, F. (1998). Mean squared error comparisons of the modified ridge regression estimator and the restricted ridge regression estimator. *Communications in Statistics*, 27(1), 131–138.
24. Baye, M.R. & Parker, D.F. (1984). Combining ridge and principal component regression: A money demand illustration. *Communications in Statistics-Theory and Methods*, 13(2), 197–205.
25. Liu, K. (1993). A new class of biased estimate in linear regression. *Communications in Statistics-Theory and Methods*, 22(2), 393–402.
26. Akdeniz, F. & Ozturk, F. (2005). The distribution of stochastic shrinkage biasing parameters of the Liu type estimator. *Applied Mathematics and Computation*, 163, 29–38.
27. Kaciranlar, S., Sakallioglu, S., Akdeniz, F., Styan, G.P.H. & Werner, H.J. (1999). A new biased estimator in linear regression and a detailed analysis of the widely-analysed dataset on Portland cement. *Sankhya: The Indian Journal of Statistics*, 61(B3), 443–459.
28. Kaciranlar, S. & Sakallioglu, S. (2001). Combining the Liu estimator and the principal component regression estimator. *Communications in Statistics-Theory and Methods*, 30(12), 2699–2705.
29. Trenkler, G. (1980). Generalized mean squared error comparisons of biased regression estimators. *Communications in Statistics-Theory and Methods*, A9(12), 1247–1259.
30. Pliskin, J.L. (1987). A ridge-type estimator and good prior means. *Communications in Statistics-Theory and Methods*, 16(12), 3429–3437.
31. Sarkar, N. (1996). Mean square error matrix comparison of some estimators in linear regressions with multicollinearity. *Statistics & Probability Letters*, 30(2), 133–138
32. Nomura, M. (1988). On the almost unbiased ridge regression estimator. *Communications in Statistics-Simulation and Computation*, 17, 729–743.
33. Sakallioglu, S., Kaciranlar, S. & Akdeniz, F. (2001). Mean squared error comparisons of some biased regression estimators. *Communications in Statistics-Theory and Methods*, 30(2), 347–361.
34. Belsley, D.A. (1991). *Conditioning diagnostics: Collinearity and weak data in regression*. New York: John Wiley & Sons.