

More on the Preliminary Test Stochastic Restricted Liu Estimator in Linear Regression Model

Sivarajah Arumairajan^{1,2}, Pushpakanthie Wijekoon³

¹Postgraduate Institute of Science, University of Peradeniya, Peradeniya, Sri Lanka

²Department of Mathematics and Statistics, Faculty of Science, University of Jaffna, Jaffna, Sri Lanka

³Department of Statistics & Computer Science, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka

Email: arumais@gmail.com, pushpaw@pdn.ac.lk

Received 18 May 2015; accepted 23 June 2015; published 29 June 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In this paper we compare recently developed preliminary test estimator called Preliminary Test Stochastic Restricted Liu Estimator (PTSRL) with Ordinary Least Square Estimator (OLSE) and Mixed Estimator (ME) in the Mean Square Error Matrix (MSEM) sense for the two cases in which the stochastic restrictions are correct and not correct. Finally a numerical example and a Monte Carlo simulation study are done to illustrate the theoretical findings.

Keywords

Multicollinearity, Stochastic Restrictions, Ordinary Least Square Estimator, Mixed Estimator, Preliminary Test Estimator, Mean Square Error Matrix

1. Introduction

To overcome the multicollinearity problem arises in the Ordinary Least Squares Estimation (OLSE) procedure, different methods have been proposed in the literature. One of the most important estimation methods is to consider biased estimators, such as the Ridge Estimator (RE) by Hoerl and Kennard [1], the Liu Estimator (LE) by Liu [2], and the Almost Unbiased Liu Estimator (AULE) by Akdeniz and Kaçiranlar [3]. Alternative method to solve the multicollinearity problem is to consider parameter estimation with some restrictions on the unknown parameters, which may be exact or stochastic. When the stochastic restrictions are available in addition to sample model, Theil and Goldberger [4] introduced the Mixed Estimator (ME). Replacing OLSE by ME in the Liu Estimator, the Stochastic Restricted Liu Estimator (SRLE) has been proposed by Hubert and Wijekoon [5].

When different estimators are available, the preliminary test estimation procedure is adopted to select a suitable estimator, and it can also be used as another estimator with combining properties of both estimators. The preliminary test approach was first proposed by Bancroft [6], and then has been studied by many researchers, such as Judge and Bock [7], Wijekoon [8] and Arumairajan and Wijekoon [9]. By combining OLSE and ME, the Ordinary Stochastic Preliminary Test Estimator (OSPE) was proposed by Wijekoon [8]. Recently, Arumairajan and Wijekoon [9] introduced the Preliminary Test Stochastic Restricted Liu Estimator (PTSRLE) by combining the Stochastic Restricted Liu Estimator and Liu Estimator. In their study, they compared PTSRLE with SRLE by using the Mean Square Error Matrix (MSEM) and Scalar Mean Square Error (SMSE) criterions.

In this research we further compare the mean square error matrix of PTSRLE with OLSE and ME. The rest of the paper is organized as follows. The model specification and estimation are given in section 2. In section 3, the mean square error matrix comparisons between PTSRLE with OLSE and ME are performed. A numerical example and a Monte Carlo simulation are used to illustrate the theoretical findings in section 4, and in section 5 we state the conclusions.

2. Model Specification and Estimation

First we consider the multiple linear regression model

$$y = X\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I), \quad (1)$$

where y is an $n \times 1$ observable random vector, X is an $n \times p$ known design matrix of rank p , β is a $p \times 1$ vector of unknown parameters and ε is an $n \times 1$ vector of disturbances.

In addition to sample model (1), let us be given some prior information about β in the form of a set of m independent stochastic linear restrictions as follows;

$$r = R\beta + \delta + v, \quad v \sim (0, \sigma^2 \Omega) \quad (2)$$

where r is an $m \times 1$ stochastic known vector R is a $m \times p$ of full row rank $m \leq p$ with known elements, δ is non zero $m \times 1$ unknown vector and v is an $m \times 1$ random vector of disturbances and Ω is assumed to be known and positive definite. Further it is assumed that v is stochastically independent of ε . i.e. $E(\varepsilon v') = 0$.

The Ordinary Least Squares Estimator (OLSE) for model (1) and the Mixed Estimator (ME) (Theil and Goldberger [4]) due to a stochastic prior restriction (2) are given by

$$\hat{\beta}_{OLSE} = S^{-1} X'y \quad (3)$$

and

$$\hat{\beta}_{ME} = \hat{\beta}_{OLSE} + S^{-1} R' (\Omega + RS^{-1} R')^{-1} (r - R\hat{\beta}_{OLSE}) \quad (4)$$

respectively, where $S = X'X$.

The expectation vector, and the mean square error matrix of $\hat{\beta}$ are given as

$$E(\hat{\beta}_{OLSE}) = \beta \quad (5)$$

and

$$MSE(\hat{\beta}_{OLSE}) = \sigma^2 S^{-1} \quad (6)$$

respectively.

The expectation vector, dispersion matrix, and the mean square error matrix of $\hat{\beta}_{ME}$ are given as

$$E(\hat{\beta}_{ME}) = \beta + H\delta, \quad (7)$$

$$D(\hat{\beta}_{ME}) = \sigma^2 (S^{-1} - G) \quad (8)$$

and

$$MSE(\hat{\beta}_{ME}) = \sigma^2 (S^{-1} - G) + H\delta\delta'H' \quad (9)$$

respectively, where, $G = S^{-1}R'(\Omega + RS^{-1}R')^{-1}RS^{-1}$, $H = S^{-1}R'(\Omega + RS^{-1}R')^{-1}$ and $\delta = E(r) - R\beta$. Note that when the stochastic restrictions are correct then $\delta = E(r) - R\beta = 0$, and consequently the Mixed Estimator becomes an unbiased estimator.

The Liu Estimator (Liu, [2]) is given as

$$\hat{\beta}_{LE}(d) = F_d \hat{\beta}_{OLSE} \text{ where } F_d = (S + I)^{-1}(S + dI) \text{ for } 0 < d < 1. \quad (10)$$

Replacing OLSE by ME in the Liu Estimator, Hubert and Wijekoon [5] introduced the Stochastic Restricted Liu Estimator (SRLE), and is given by

$$\hat{\beta}_{SRLE}(d) = F_d \hat{\beta}_{ME}. \quad (11)$$

when different estimators are available for the same parameter vector β in the linear regression model one must solve the problem of their comparison. Usually as a simultaneous measure of covariance and bias, the mean square error matrix of $\hat{\beta}$ is used, and is defined by

$$MSE(\hat{\beta}, \beta) = E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'] = D(\hat{\beta}) + B(\hat{\beta})B(\hat{\beta})', \quad (12)$$

where $D(\hat{\beta})$ is the dispersion matrix, and $B(\hat{\beta}) = E(\hat{\beta}) - \beta$ denotes the bias vector. We recall that the Scalar Mean Square Error of $\hat{\beta}$ $SMSE(\hat{\beta}, \beta) = \text{trace}(MSE(\hat{\beta}, \beta))$.

For two given estimators $\hat{\beta}_1$ and $\hat{\beta}_2$, the estimator $\hat{\beta}_2$ is said to be superior to $\hat{\beta}_1$ under the MSEM criterion if and only if

$$M(\hat{\beta}_1, \hat{\beta}_2) = MSE(\hat{\beta}_1, \beta) - MSE(\hat{\beta}_2, \beta) \geq 0. \quad (13)$$

Let us now turn to the question of the statistical evaluation of the compatibility of sample and stochastic information. The classical procedures is to test the hypothesis

$$H_0 : \delta = 0 \text{ against } H_1 : \delta \neq 0 \quad (14)$$

under linear model (1) and stochastic prior information (2).

The Ordinary Stochastic Preliminary Test Estimator (OSPE) of β (Wijekoon [8]) is defined as

$$\hat{\beta}_{OSPE} = \begin{cases} \hat{\beta}_{ME} & \text{if } H_0 : \delta = 0 \\ \hat{\beta}_{OLSE} & \text{if } H_1 : \delta \neq 0. \end{cases} \quad (15)$$

Further, we can write (15) as

$$\hat{\beta}_{OSPE} = \hat{\beta}_{ME} I_{[0, F_{m,n-p}(\alpha))}(F) + \hat{\beta}_{OLSE} I_{[F_{m,n-p}(\alpha), \infty)}(F), \quad (16)$$

where,

$$F = \frac{(r - R\hat{\beta}_{OLSE})'(\Omega + RS^{-1}R')^{-1}(r - R\hat{\beta}_{OLSE})}{m\hat{\sigma}^2} \quad (17)$$

which has a non-central $F_{m,n-p,\lambda}$ distribution under $H_1 : \delta \neq 0$, with non-centrality parameter

$$\lambda = \frac{\delta'(\Omega + RS^{-1}R')^{-1}\delta}{2\sigma^2} \text{ with } \hat{\sigma}^2 = \frac{(Y - X\hat{\beta}_{OLSE})'(Y - X\hat{\beta}_{OLSE})}{n-p}, \quad (18)$$

and $I_{[0, F_{m,n-p}(\alpha))}(F)$ and $I_{[F_{m,n-p}(\alpha), \infty)}(F)$ are indicator functions which take the value one if F falls in the sub-

scripted interval, and zero otherwise. $F_{m,n-p}(\alpha)$ is the upper α -level critical value from the central F distribution $F_{m,n-p,0}$.

The expectation vector, dispersion matrix, and the mean square error matrix of $\hat{\beta}_{OSPE}$ are derived by Wijekoon [8], and given by

$$E(\hat{\beta}_{OSPE}) = \beta + h_{\lambda}(2)H\delta, \quad (19)$$

$$D(\hat{\beta}_{OSPE}) = \sigma^2 S^{-1} - \sigma^2 h_{\lambda}(2)G + [2h_{\lambda}(2) - h_{\lambda}(4) - h_{\lambda}^2(2)]H\delta\delta'H' \quad (20)$$

an

$$MSE(\hat{\beta}_{OSPE}) = \sigma^2 S^{-1} - \sigma^2 h_{\lambda}(2)G + [2h_{\lambda}(2) - h_{\lambda}(4)]H\delta\delta'H' \quad (21)$$

respectively, where, $h_{\lambda}(\ell) = \Pr\left(\frac{\chi_{m+\ell,\lambda}^2}{\chi_{n-p}^2} \leq \frac{mF_{m,n-p}(\alpha)}{n-p}\right)$ for $\ell \in \mathbb{N}$.

Recently, Arumairajan and Wijekoon [9] proposed the Preliminary Test Stochastic Restricted Liu Estimator (PTSRLE) by combining the Liu Estimator and Stochastic Restricted Liu Estimator, and is given by

$$\tilde{\beta}_{PTSRLE} = \begin{cases} F_d \hat{\beta}_{ME} & \text{if } H_0 : \delta = 0 \\ F_d \hat{\beta}_{OLSE} & \text{if } H_1 : \delta \neq 0. \end{cases} \quad (22)$$

Note that the PTSRLE can be rewritten as follows:

$$\tilde{\beta}_{PTSRLE}(d) = F_d \hat{\beta}_{ME} I_{[0, F_{m,n-p}(\alpha)]}(F) + F_d \hat{\beta}_{OLSE} I_{[F_{m,n-p}(\alpha), \infty)}(F) = F_d \hat{\beta}_{OSPE}. \quad (23)$$

By using Equations (19), (20) and (21), Arumairajan and Wijekoon [9] derived the expectation vector, bias vector, dispersion matrix and mean square error matrix of PTSRLE as follows:

$$E[\tilde{\beta}_{PTSRLE}(d)] = F_d E[\hat{\beta}_{OSPE}] = F_d \beta + h_{\lambda}(2)F_d H\delta, \quad (24)$$

$$B[\tilde{\beta}_{PTSRLE}(d)] = (d-1)(S+I)^{-1} \beta + h_{\lambda}(2)F_d H\delta, \quad (25)$$

$$\begin{aligned} D[\tilde{\beta}_{PTSRLE}(d)] &= F_d D[\hat{\beta}_{OSPE}] F_d' \\ &= \sigma^2 F_d S^{-1} F_d' - \sigma^2 h_{\lambda}(2)F_d G F_d' + \xi F_d H\delta\delta'H' F_d' \end{aligned} \quad (26)$$

and

$$\begin{aligned} MSE[\tilde{\beta}_{PTSRLE}(d)] &= \sigma^2 F_d S^{-1} F_d' - \sigma^2 h_{\lambda}(2)F_d G F_d' + \xi F_d H\delta\delta'H' F_d' \\ &\quad + [(d-1)(S+I)^{-1} \beta + h_{\lambda}(2)F_d H\delta][[(d-1)(S+I)^{-1} \beta + h_{\lambda}(2)F_d H\delta]]' \end{aligned} \quad (27)$$

respectively, where $\xi = 2h_{\lambda}(2) - h_{\lambda}(4) - h_{\lambda}^2(2) \geq 0$.

3. Mean Square Error Matrix Comparisons

In this section we compare the PTSRLE with OLSE and ME in the mean square error matrix sense for the two cases in which the stochastic restrictions are correct and not correct.

3.1. Comparison between the PTSRLE and OLSE

The mean square error matrix difference between OLSE and PTSRLE can be written as

$$MSE[\hat{\beta}_{OLSE}] - MSE[\tilde{\beta}_{PTSRLE}(d)] = D - d_1 d_1' \quad (28)$$

where, $D = D[\hat{\beta}_{OLSE}] - D[\tilde{\beta}_{PTSRLE}(d)] = D_1 - D_2$

with $D_1 = \sigma^2 S^{-1} - \sigma^2 F_d S^{-1} F_d' + \sigma^2 h_{\lambda}(2) F_d G F_d'$, $D_2 = \xi F_d H \delta \delta' H' F_d'$ and $d_1 = (d-1)(S+I)^{-1} \beta + h_{\lambda}(2) F_d H \delta$.

Now the following theorem can be stated

Theorem 1:

1) When the stochastic restrictions are true (*i.e.* $\delta = 0$), the PTSRLE is superior to OLSE in the mean square error matrix sense if and only if $d_1^{*'} D_1^{*-1} d_1^* \leq 1$, where $d_1^* = (d-1)(S+I)^{-1} \beta$, $D_1^* = \sigma^2 S^{-1} - \sigma^2 F_d S^{-1} F_d' + \sigma^2 h_{\lambda^*}(2) F_d G F_d'$, and $h_{\lambda^*}(2)$ is the value of $h_{\lambda}(2)$ when $\delta = 0$.

2) When the stochastic restrictions are not true (*i.e.* $\delta \neq 0$), and if the maximum eigenvalue of $D_2 D_1^{-1}$ is less than one, the PTSRLE is superior to OLSE if and only if $d_1' D^{-1} d_1 \leq 1$.

Proof:

1) If stochastic restrictions are correct then $\delta = E(r) - R\beta = 0$, and the equation (28) reduced to

$$\sigma^2 S^{-1} - \sigma^2 F_d S^{-1} F_d' + \sigma^2 h_{\lambda^*}(2) F_d G F_d' - d_1^{*'} d_1^*.$$

To apply lemma 2 (Appendix), we have to show that

$$D_1^* = \sigma^2 S^{-1} - \sigma^2 F_d S^{-1} F_d' + \sigma^2 h_{\lambda^*}(2) F_d G F_d' > 0.$$

Since $F_d = (S+I)^{-1}(S+dI)$ for $0 < d < 1$ we have

$$\begin{aligned} & \sigma^2 S^{-1} - \sigma^2 F_d S^{-1} F_d' \\ &= \sigma^2 S^{-1} - \sigma^2 (S+I)^{-1} (S+dI) S^{-1} (S+dI) (S+I)^{-1} \\ &= \sigma^2 (S+I)^{-1} [(S+I) S^{-1} (S+I) + (S+dI) S^{-1} (S+dI)] (S+I)^{-1} \\ &= \sigma^2 (1-d) (S+I)^{-1} [2I + (1+d) S^{-1}] (S+I)^{-1} \end{aligned}$$

Hence $(\sigma^2 S^{-1} - \sigma^2 F_d S^{-1} F_d')$ is positive definite matrix for all $0 < d < 1$. According to lemma 1 (Appendix), D_1^* is positive definite matrix since $(\sigma^2 S^{-1} - \sigma^2 F_d S^{-1} F_d')$ is positive definite matrix and $\sigma^2 h_{\lambda^*}(2) F_d G F_d'$ is nonnegative definite matrix. By applying lemma 2 (Appendix), $D_1^* - d_1^{*'} d_1^* \geq 0$ if and only if $d_1^{*'} D_1^{*-1} d_1^* \leq 1$. Therefore the PTSRLE is superior to OLSE in the mean square error matrix sense when stochastic restrictions are correct if and only if $d_1^{*'} D_1^{*-1} d_1^* \leq 1$.

2) If stochastic restrictions are not correct then $E(r) - R\beta \neq 0$, and consequently with respect to the MSE matrix criterion $\tilde{\beta}_{PTSRLE}(d)$ is superior to $\hat{\beta}_{OLSE}$ if and only if $D - d_1 d_1' \geq 0$. To apply lemma 2 (Appendix) in order to prove that $D - d_1 d_1' \geq 0$, one required condition is that $D > 0$.

We have already proved that $D_1 = \sigma^2 S^{-1} - \sigma^2 F_d S^{-1} F_d' + \sigma^2 h_{\lambda}(2) F_d G F_d' > 0$. Also $D_2 = \xi F_d H \delta \delta' H' F_d'$ is positive definite since $H \delta \delta' H' > 0$, $F_d > 0$ and $\xi > 0$. Now according to lemma 3 (Appendix),

$D[\hat{\beta}_{OLSE}] - D[\tilde{\beta}_{PTSRLE}(d)] = D_1 - D_2 > 0$ if and only if $\lambda_{\max}(D_2 D_1^{-1}) < 1$, where $\lambda_{\max}(D_2 D_1^{-1})$ is the maximum eigenvalue of $D_2 D_1^{-1}$. Now according to lemma 2, $MSE[\hat{\beta}_{OLSE}] - MSE[\tilde{\beta}_{PTSRLE}(d)] \geq 0$ if and only if $d_1' D^{-1} d_1 \leq 1$. This completes the proof.

3.2. Comparison between the PTSRLE and ME

The mean square error matrix difference between ME and PTSRLE is

$$MSE[\hat{\beta}_{ME}] - MSE[\tilde{\beta}_{PTSRLE}(d)] = C + c_1 c_1' - d_1 d_1' \quad (29)$$

where $C = D_1 - D_3$ with $D_1 = \sigma^2 S^{-1} - \sigma^2 F_d S^{-1} F_d' + \sigma^2 h_{\lambda}(2) F_d G F_d'$, $D_3 = \sigma^2 G + \xi F_d H \delta \delta' H' F_d'$, $c_1 = H \delta$ and $d_1 = (d-1)(S+I)^{-1} \beta + h_{\lambda}(2) F_d H \delta$.

Now we can give the following theorem.

Theorem 2:

1) When the stochastic restrictions are true (*i.e.* $\delta = 0$), and if the maximum eigen value of $D_3^* D_1^{*-1}$ is less than one, the PTSRLE is superior to ME in the mean square error matrix sense if and only if $d_1^{*'} C^{*-1} d_1^* \leq 1$ where $C^* = D_1^* - D_3^*$ with $D_1^* = \sigma^2 S^{-1} - \sigma^2 F_d S^{-1} F_d' + \sigma^2 h_{\lambda^*}(2) F_d G F_d'$ and $D_3^* = \sigma^2 G$.

2) When the stochastic restrictions are not true (*i.e.* $\delta \neq 0$), and if the maximum eigen value of $D_3 D_1^{-1}$ is less than one, the PTSRLE is superior to ME in the mean square error matrix sense if and only if $d_1'(C + c_1 c_1')^{-1} d_1 \leq 1$.

Proof:

1) When stochastic restrictions are true then $\delta = E(r) - R\beta = 0$, and hence $D_3^* = \sigma^2 G$, which is clearly a nonnegative matrix. We have already proved that D_1^* is positive definite matrix. Now according to lemma 3, $C^* = D_1^* - D_3^* > 0$ if and only if the maximum eigen values of $D_3^* D_1^{*-1}$ is less than one. Now according to lemma 2, the PTSRLE is superior to ME if and only if $d_1^{*'} C^{*-1} d_1^* \leq 1$, and if $D_3^* D_1^{*-1}$.

2) When stochastic restrictions are not true (*i.e.* $\delta \neq 0$), $MSE[\hat{\beta}_{ME}] - MSE[\tilde{\beta}_{PTSRLE}(d)] \geq 0$ if and only if $C + c_1 c_1' - d_1 d_1' \geq 0$. To prove that $C + c_1 c_1' - d_1 d_1' \geq 0$ lemma 4 (Appendix) can be used. To use this lemma first we prove that C is positive definite. According to lemma 3, C is a positive definite matrix if and only if the maximum eigenvalue of $D_3 D_1^{-1}$ is less than one. By applying lemma 4, we can prove that $C + c_1 c_1' - d_1 d_1' \geq 0$ if and only if $d_1'(C + c_1 c_1')^{-1} d_1 \leq 1$. This completes the proof.

According to theorem 1 and 2 it is clear that PTSRLE is superior to OLSE and ME under certain conditions.

4. Numerical Illustrations

In this section the comparison of PTSRLE with OLSE and ME are demonstrated using a numerical example, and a simulation study.

4.1. Numerical Example

To illustrate our theoretical result, we consider the data set on Total National Research and Development Expenditures as a percent of Gross National product due to Gruber [10]. This data set is used by Akdeniz and Erol [11], Li and Yang [12] and Wu and Yang [13] to verify the theoretical results. Data shows Total National Research and Development Expenditures as a Percent of Gross National Product by Country: 1972-1986. It represents the relationship between the dependent variable y the percentage spent by the United States and the four other independent variables x_1, x_2, x_3 and x_4 . The variable x_1 represents the percent spent by France, x_2 that spent by West Germany, x_3 that spent by Japan, and x_4 that spent by the former Soviet Union. The data set is given below:

$$X = \begin{pmatrix} 1.9 & 2.2 & 1.9 & 3.7 \\ 1.8 & 2.2 & 2.0 & 3.8 \\ 1.8 & 2.4 & 2.1 & 3.6 \\ 1.8 & 2.4 & 2.2 & 3.8 \\ 2.0 & 2.5 & 2.3 & 3.8 \\ 2.1 & 2.6 & 2.4 & 3.7 \\ 2.1 & 2.6 & 2.6 & 3.8 \\ 2.2 & 2.6 & 2.6 & 4.0 \\ 2.3 & 2.8 & 2.8 & 3.7 \\ 2.3 & 2.7 & 2.8 & 3.8 \end{pmatrix}, \quad y = \begin{pmatrix} 2.3 \\ 2.2 \\ 2.2 \\ 2.3 \\ 2.4 \\ 2.5 \\ 2.6 \\ 2.6 \\ 2.7 \\ 2.7 \end{pmatrix}.$$

The four column of the 10×4 matrix X comprise the data on x_1, x_2, x_3 and x_4 respectively, and y is the response variable. From this data, we obtain the following results:

- 1) The eigen values of $X'X$: 302.9626, 0.7283, 0.0447, 0.0345.
- 2) The OLS estimator of β : $\hat{\beta}_{OLSE} = (0.6455, 0.0896, 0.1436, 0.1526)'$.
- 3) The OLS estimator of σ^2 : $\hat{\sigma}^2 = 0.0015$.
- 4) The condition number of $X'X$: 8781.53.

The condition number implies the existence of multicollinearity in the data set. We consider the following stochastic restrictions (Li and Yang, [12])

$$R = (1, -2, -2, -2)', \quad r = 0, \quad v \sim (0, \hat{\sigma}^2 = 0.0015) \text{ and } \hat{\delta} = r - R\hat{\beta}.$$

Further the significance level is taken as $\alpha = 0.05$. **Figure 1** is drawn by using the SMSE obtained by using Equations (6), (9) and (27).

Based on **Figure 1**, we can say that the SMSE of PTSRLLE is larger than the SMSE of both ME and OLSE when d is small. But the PTSRLLE has the smallest SMSE than both ME and OLSE for $0.6029 < d < 0.9386$. Moreover the ME is always superior to OLSE.

4.2. Monte Carlo Simulation

To illustrate the behavior of our proposed estimators, we perform the Monte Carlo Simulation study by considering different levels of multicollinearity. Following McDonald and Galarneau [14] we generate explanatory variables as follows:

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

where z_{ij} is an independent standard normal pseudo random number, and γ is specified so that the theoretical correlation between any two explanatory variables is given by γ^2 . A dependent variable is generated by using the equation.

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \varepsilon_i, \quad i = 1, 2, \dots, n,$$

where ε_i is a normal pseudo random number with mean zero and variance σ_i^2 . Newhouse and Oman [15] have noted that if the MSE is a function of σ^2 and β , and if the explanatory variables are fixed, then subject to the constraint $\beta'\beta = 1$, the MSE is minimized when β is the normalized eigenvector corresponding to the largest eigenvalue of the $X'X$ matrix. In this study we choose the normalized eigenvector corresponding to the largest eigenvalue of $X'X$ as the coefficient vector β , $n = 100$, $p = 4$ and $\sigma_i^2 = 1$. Four different sets of correlations are considered by selecting the values as $\gamma = 0.9, 0.99, 0.999$ and 0.9999 , and the significance level is taken as $\alpha = 0.05$. The condition numbers for $\gamma = 0.9, 0.99, 0.999$ and 0.9999 are 19.49, 223.67, 2297.89 and 23137.88 respectively. **Figures 2-5** are drawn by using the SMSE obtained by using Equation (6), (9) and (27).

According to **Figure 2**, the PTSRLLE has the smallest SMSE than ME except d is small. However the OLSE has the smallest SMSE than ME and PTSRLLE. From **Figure 3**, the SMSE of PTSRLLE is larger than the SMSE of both OLSE and ME when d is small. From **Figure 4** and **Figure 5**, we can notice that most of the cases, the PTSRLLE has the smallest SMSE than ME and OLSE.

5. Conclusion

In this paper we have shown that the Preliminary Test Stochastic Restricted Liu Estimator is superior to Mixed Estimator and Ordinary Least Square Estimator in the mean square error matrix sense under certain conditions.

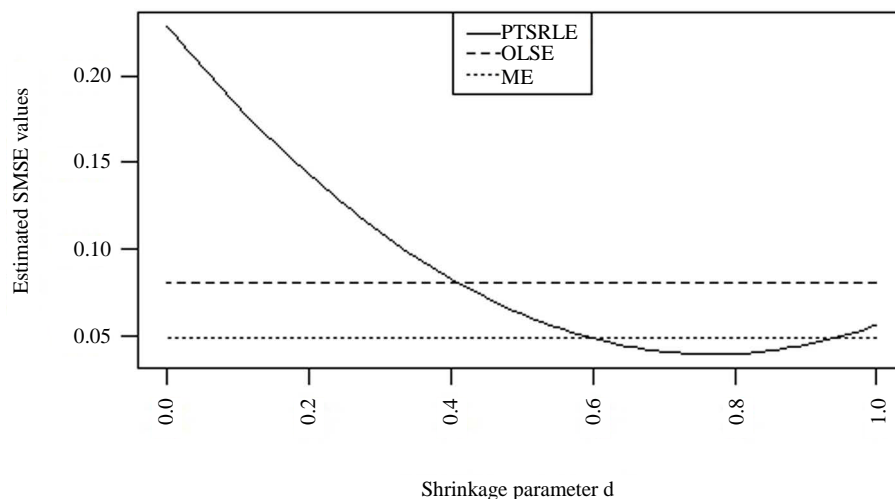


Figure 1. Estimated SMSE for PTSRLLE, ME and OLSE.

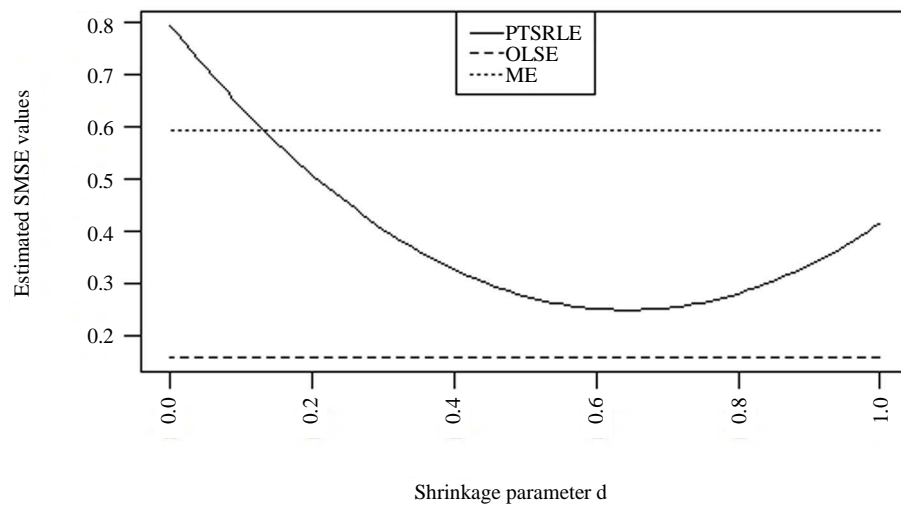


Figure 2. Estimated SMSE for PTSRLE, ME and OLSE when $\gamma = 0.9$.

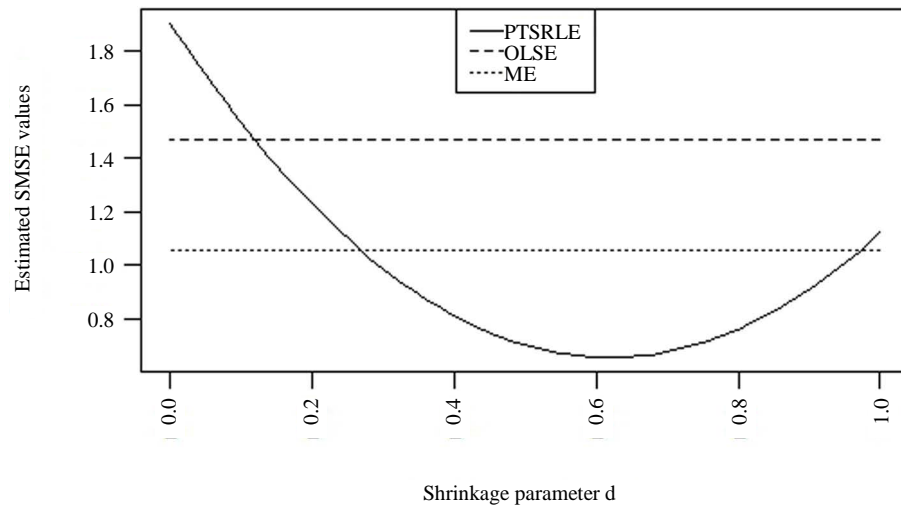


Figure 3. Estimated SMSE for PTSRLE, ME and OLSE when $\gamma = 0.99$.

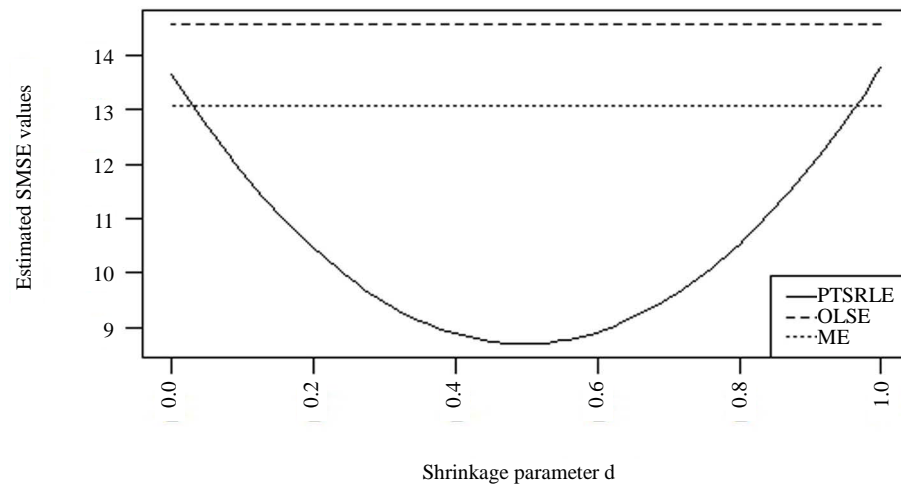


Figure 4. Estimated SMSE for PTSRLE, ME and OLSE when $\gamma = 0.999$.

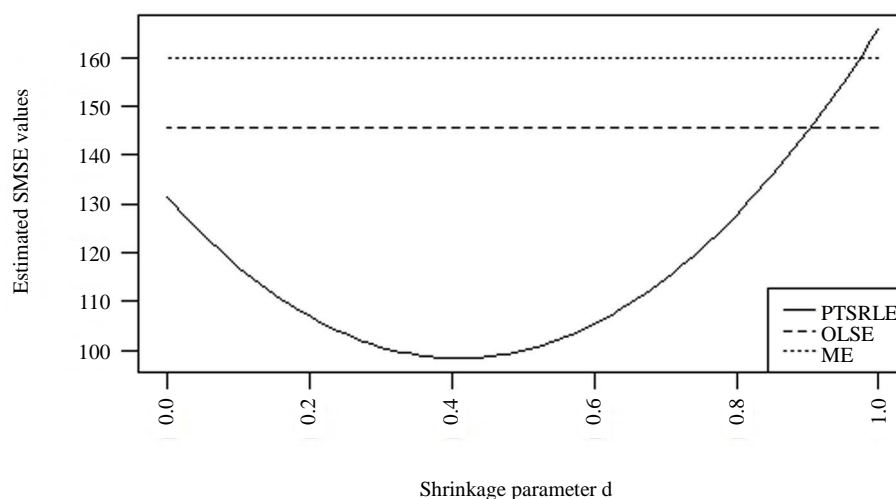


Figure 5. Estimated SMSE for PTSRLE, ME and OLSE when $\gamma = 0.9999$.

From the simulation study and the numerical illustration we notice that the PTSRLE has the smallest SMSE than ME and OLSE when multicollinearity among the predictor variables is large.

Acknowledgements

We thank the Postgraduate Institute of Science, University of Peradeniya, Sri Lanka for providing all facilities to do this research.

References

- [1] Hoerl, E. and Kennard, W. (1970) Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, **12**, 55-67. <http://dx.doi.org/10.1080/00401706.1970.10488634>
- [2] Liu, K. (1993) A New Class of Biased Estimate in Linear Regression. *Communications in Statistics—Theory and Methods*, **22**, 393-402. <http://dx.doi.org/10.1080/03610929308831027>
- [3] 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*, **34**, 1789-1797. <http://dx.doi.org/10.1080/03610929508831585>
- [4] Theil, H. and Goldberger, A.S. (1961) On Pure and Mixed Estimation in Economics. *International Economic Review*, **2**, 65-77. <http://dx.doi.org/10.2307/2525589>
- [5] Hubert, M.H. and Wijekoon, P. (2006) Improvement of the Liu Estimator in Linear Regression Model. *Statistical Papers*, **47**, 471-479. <http://dx.doi.org/10.1007/s00362-006-0300-4>
- [6] Bancroft, A. (1944) On Biases in Estimation Due to Use of Preliminary Tests of Significance. *Annals of Mathematical Statistics*, **15**, 190-204. <http://dx.doi.org/10.1214/aoms/1177731284>
- [7] Judge, G. and Bock, E. (1978) *The Statistical Implications of Pre-Test and Stein-Rule Estimators in Econometrics*. North Holland, New York.
- [8] Wijekoon, P. (1990) *Mixed Estimation and Preliminary Test Estimation in the Linear Regression Model*. Ph.D. Thesis, University of Dortmund, Dortmund.
- [9] Arumairajan, S. and Wijekoon, P. (2013) Improvement of the Preliminary Test Estimator When Stochastic Restrictions are Available in Linear Regression Model. *Open Journal of Statistics*, **3**, 283-292. <http://dx.doi.org/10.4236/ojs.2013.34033>
- [10] Gruber, M.H.J. (1998) *Improving Efficiency by Shrinkage: The James-Stein and Ridge Regression Estimators*. Dekker, Inc., New York.
- [11] 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**, 2389-2413. <http://dx.doi.org/10.1081/STA-120025385>
- [12] Li, Y. and Yang, H. (2010) A New Stochastic Mixed Ridge Estimator in Linear Regression. *Statistical Papers*, **51**,

- 315-323. <http://dx.doi.org/10.1007/s00362-008-0169-5>
- [13] Wu, J. and Yang, H. (2013) Two Stochastic Restricted Principal Components Regression Estimator in Linear Regression. *Communications in Statistics—Theory and Methods*, **42**, 3793-3804. <http://dx.doi.org/10.1080/03610926.2011.639004>
- [14] McDonald, C. and Galarneau, A. (1975) A Monte Carlo Evaluation of some Ridge-Type Estimators. *Journal of American Statistical Association*, **70**, 407-416. <http://dx.doi.org/10.1080/01621459.1975.10479882>
- [15] Newhouse, J.P. and Oman, S.D. (1971) An Evaluation of Ridge Estimators. Rand Report, No. R-716-Pr, 1-28.
- [16] Rao, C.R. and Toutenburg, H. (1995) Linear Models, Least Squares and Alternatives. Springer Verlag, Berlin. <http://dx.doi.org/10.1007/978-1-4899-0024-1>
- [17] Farebrother, R.W. (1976) Further Results on the Mean Square Error of Ridge Regression. *Journal of the Royal Statistical Society*, **38**, 248-250.
- [18] Wang, S.G., Wu, M.X. and Jia, Z.Z. (2006) Matrix Inequalities. 2nd Edition, Chinese Science Press, Beijing.
- [19] Trenkler, G. and Toutenburg, H. (1990) Mean Square Error Matrix Comparisons between Biased Estimators—An Overview of Recent Results. *Statistical Papers*, **31**, 165-179. <http://dx.doi.org/10.1007/BF02924687>

Appendix

Lemma 1: (Rao and Toutenburg, [16])

Let A and B be $(n \times n)$ matrices such that $A > 0$ and $B \geq 0$. Then $A + B > 0$.

Lemma 2: (Farebrother, [17])

Let $A > 0$ be an $(n \times n)$ matrix, b an $(n \times 1)$ vector. Then $A - b_1 b_1' \geq 0$ if and only if $b_1' A^{-1} b_1 \leq 1$.

Lemma 3: (Wang *et al.*, [18])

Let $n \times n$ matrices $M > 0$, $N > 0$ (or $N \geq 0$), then $M > N$ if and only if $\lambda_1(NM^{-1}) < 1$, where $\lambda_1(NM^{-1})$ is the largest eigenvalue of the matrix NM^{-1} .

Lemma 4: (Trenkler and Toutenburg, [19])

Let $\hat{\beta}_1$ and $\hat{\beta}_2$ be two linear estimator of β . Suppose that $D = D(\hat{\beta}_1) - D(\hat{\beta}_2)$ is positive definite then $\Delta = MSE(\hat{\beta}_1) - MSE(\hat{\beta}_2)$ is nonnegative definite if and only if $b_2'(D + b_1 b_1')^{-1} b_2 \leq 1$, where b_j denotes the bias vector of $\hat{\beta}_j$, $j = 1, 2$.