

A NEW STOCHASTIC RESTRICTED BIASED ESTIMATOR UNDER HETEROSCEDASTIC OR CORRELATED ERROR

MUSTAFA ISMAEEL ALHEETY¹

Abstract. In this paper, under the linear regression model with heteroscedastic and/or correlated errors when the stochastic linear restrictions on the parameter vector are assumed to be held, a generalization of the ordinary mixed estimator (GOME), ordinary ridge regression estimator (GORR) and Generalized least squares estimator (GLSE) is proposed. The performance of this new estimator against GOME, GORR, GLS and the stochastic restricted Liu estimator (SRLE) [Yang and Xu, *Statist. Papers* **50** (2007) 639–647] are examined in terms of matrix mean square error criterion. A numerical example is considered to illustrate the theoretical results.

Mathematics Subject Classification. 62J05, 62J07.

Received September 17, 2009. Revised January 9, 2009.

1. INTRODUCTION

Consider the multiple linear regression model

$$Y = X\beta + \epsilon, \quad (1.1)$$

where Y is an $n \times 1$ vector of observations, X is an $n \times p$ matrix of rank p , β is a $p \times 1$ vector of unknown parameters, and ϵ is an $n \times 1$ vector of nonobservable errors with $E(\epsilon) = 0$ and variance-covariance matrix $\text{Cov}(\epsilon) = \sigma^2 I_n$.

The method of least squares is generally used to estimate the coefficients in the model (1.1)

$$\hat{\beta} = (X'X)^{-1}X'Y. \quad (1.2)$$

$\hat{\beta}$ has minimum variance in the class of unbiased estimators. While this property is theoretically satisfying, in practice the results of $\hat{\beta}$ are often unacceptable, particularly when there are near-linear dependencies among the regressors variables. This is a problem of multicollinearity. To overcome this problem, different remedial actions have been proposed. We believe that biased estimation methods are useful techniques, that they should be considered, when the analyst deals with such a problem.

Keywords and phrases. Heteroscedasticity, generalized least squares estimator, stochastic restricted Liu estimator.

¹ Department of Mathematics, Al-Anbar University, Ramadi, Iraq; alheety@yahoo.com

Hoerl and Kennard [4,5] suggested the ordinary ridge regression estimator (ORR)

$$\hat{\beta}(k) = (X'X + kI)^{-1}X'Y, \quad (1.3)$$

where $k > 0$ is a scalar constant, often called the biasing parameter.

Liu [7] combined the Stein estimator [10] with ORR estimator which is called (LE). The motivation of LE is came by the fact: the combining of two estimators may inherit advantages of both estimators. LE is defined as follows

$$\hat{\beta}(d) = (X'X + I)^{-1}(X'X + dI)\hat{\beta}, \quad (1.4)$$

where $0 < d < 1$ is a biasing parameter.

An alternative technique to combat multicollinearity is to consider the parameter estimation of the model with some restrictions on the unknown parameters.

Suppose there are some prior information with the model (1.1) about β as a set of q independent stochastic linear restrictions as follows

$$r = R\beta + \phi, \quad \phi \sim (0, \sigma^2V) \quad (1.5)$$

where r is a $q \times 1$ known vector, R is a $q \times p$ known matrix of rank q and ϕ is a $q \times 1$ vector of disturbances with mean 0 and variance-covariance matrix σ^2V , V may be assumed to be known and positive definite (p.d.). Theil and Goldberger [12] and Theil [11] introduced the mixed estimator (OME) technique by bringing together the sample and the prior information (1.5)

$$\begin{pmatrix} Y \\ r \end{pmatrix} = \begin{pmatrix} X \\ R \end{pmatrix} \beta + \begin{pmatrix} \epsilon \\ \phi \end{pmatrix}, \quad (1.6)$$

where both random errors are uncorrelated

$$E(\epsilon\phi') = 0.$$

Since $E(\epsilon\epsilon') = \sigma^2I_n$, therefore, with (1.5), the variance-covariance matrix becomes

$$E \begin{pmatrix} \epsilon \\ \phi \end{pmatrix} (\epsilon \quad \phi)' = \sigma^2 \begin{pmatrix} I_n & 0 \\ 0 & V \end{pmatrix}. \quad (1.7)$$

Calling the augmented matrices and vectors in the mixed model (1.6) Y_* , X_* and ϵ_* , that is

$$Y_* = \begin{pmatrix} Y \\ r \end{pmatrix}, \quad X_* = \begin{pmatrix} X \\ R \end{pmatrix}, \quad \epsilon_* = \begin{pmatrix} \epsilon \\ \phi \end{pmatrix}.$$

Therefore, we may write

$$Y_* = X_*\beta + \epsilon_*, \quad \epsilon_* \sim (0, \sigma^2W_*), \quad (1.8)$$

where

$$W_* = \begin{pmatrix} I_n & 0 \\ 0 & V \end{pmatrix}$$

is positive definite (p.d.). The best linear unbiased estimator of β in the mixed model (1.8) which is called ordinary mixed estimator (OME) is given as follows:

$$\begin{aligned} \hat{\beta}_{\text{OME}} &= (X'X + R'V^{-1}R)^{-1}(X'Y + R'V^{-1}r) \\ &= \hat{\beta} + (X'X)^{-1}R'(V + R(X'X)^{-1}R')^{-1}(r - R\hat{\beta}). \end{aligned} \quad (1.9)$$

In order to improve the OME, Hubert and Wijekoon [6] inserted the Liu estimator into the OME procedure and obtained a new type of stochastic restricted Liu estimators as

$$\hat{\beta}_{srd}(d) = J_d\hat{\beta}_{\text{OME}},$$

where $J_d = (X'X + I)^{-1}(X'X + dI)$. In practical applications, the homoscedasticity ($\text{Var}(\epsilon) = \sigma^2 I_n$) and/or uncorrelatedness ($\text{Var}(\epsilon_i, \epsilon_j) = 0$), for $i \neq j$ can not be satisfied. Then with model (1.1) we consider the error assumption which is called heteroscedasticity and/or correlated errors as follows

$$E(\epsilon) = 0, \text{Cov}(\epsilon) = \sigma^2 W, \quad (1.10)$$

where W is a known $n \times n$ symmetric positive definite (p.d.) matrix.

Since W is symmetric and p.d., there exists a nonsingular $n \times n$ matrix T such that $W^{-1} = T'T$. Transforming the model (1.1) by premultiplication with T :

$$TY = TX\beta + T\epsilon.$$

Let $Y^* = TY$, $X^* = TX$, $\epsilon^* = T\epsilon$, it is seen that $E(\epsilon^*) = 0$ and $\text{Cov}(\epsilon^*) = \sigma^2 I$. Therefore, the transformed model

$$Y^* = X^*\beta + \epsilon^* \quad (1.11)$$

satisfies the assumptions of the homoscedasticity and uncorrelated errors. The OLS estimator of β in (1.11) is

$$\hat{\beta}_G = (X^{*'}X^*)^{-1}X^{*'}Y^* = (X'W^{-1}X)^{-1}X'W^{-1}Y \quad (1.12)$$

which is well known as the generalized least squares (GLS) estimator.

Hence, the best linear unbiased estimator of β in the model (1.10) with the assumptions (1.9) and the stochastic restrictions (1.4) is defined as:

$$\hat{\beta}_{\text{GOME}} = (S + R'V^{-1}R)^{-1}(X'W^{-1}Y + R'V^{-1}r), \quad (1.13)$$

where $S = X'W^{-1}X$. The estimator in (1.13) also called mixed estimator but under heteroscedasticity and/or correlated errors and we denoted as (GOME).

Trenkler [13] specified the ridge estimator of β in the model (1.10) as

$$\hat{\beta}_G(k) = (S + kI)^{-1}X'W^{-1}Y, k > 0 \quad (1.14)$$

to reduce the effect of the multicollinearity problem on the variance of the estimated model (1.1) with assumption (1.10), we denoted as (GORR).

Consider the LE of β in the model (1.11) as

$$\hat{\beta}_G(d) = (S + I)^{-1}(S + dI)\hat{\beta}_G \quad (1.15)$$

and we denoted as (GOLE).

Yang and Xu [14] introduced the stochastic restricted Liu estimator (SRLE) under the linear model (1.1) with assumption of homoscedasticity and uncorrelated errors as follows

$$\hat{\beta}_{\text{SRLE}}(d) = (X'X + R'V^{-1}R)^{-1}(J_d X'Y + R'V^{-1}r). \quad (1.16)$$

Since we are interesting in the model (1.11) with assumption (1.10) and stochastic restrictions (1.4), $\hat{\beta}_{\text{SRLE}}$ will be considered as follows

$$\hat{\beta}_{\text{GSRLE}}(d) = (S + R'V^{-1}R)^{-1}(F_d X'W^{-1}Y + R'V^{-1}r), \quad (1.17)$$

where $F_d = (S + I)^{-1}(S + dI)$ and we denoted as (GSRLE) to be not confused with (SRLE).

In this paper, we introduce a new mixed estimator under stochastic restrictions with heteroscedastic and/or correlated errors. In Section 2 we give some Lemmas with the concept of the new estimator. The performance of the new estimator compared with the GOME, GORR GLS and GSRLE is shown in Section 3. Section 4 gives the numerical example to illustrate the theoretical results.

2. SOME LEMMAS AND THE NEW ESTIMATOR

In this section we need to give some lemmas which are needed to do and prove our theoretical results. Also we consider the new estimator with its properties.

Lemma 2.1 (Rao *et al.* [9]). *Suppose that the square matrices $A : p \times p$, $C : n \times n$ are not singular, and $B : p \times n$, $D : n \times p$ be any two matrices, then*

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

Lemma 2.2 (Farebrother [2]). *Suppose that A is $p.d.$ and a is a column vector. Then $A - aa'$ is non negative definite (n.n.d.) if and only if*

$$a'A^{-1}a \leq 1.$$

Lemma 2.3 (Rao [8]). *Let B a $p.d.$ matrix and A a n.n.d. matrix, and $\Lambda = \text{diag}(\lambda_i^B(A))$ the diagonal matrix of the eigen values of A in the B . Then there exists a non singular matrix H such that*

$$B = H'H, \quad A = H'\Lambda H.$$

Lemma 2.4 (Rao *et al.* [9]). *Suppose that $D = \text{Cov}(\hat{\beta}_1) - \text{Cov}(\hat{\beta}_2)$ is $p.d.$ Then $\Delta = \text{MSEM}(\hat{\beta}_1) - \text{MSEM}(\hat{\beta}_2)$ is n.n.d. if and only if*

$$(1 + d_{11})(d_{22} - 1) \leq d_{12}^2,$$

where $d_{ij} = B_i'D^{-1}B_j$, $B_i = \text{Bias}(\hat{\beta}_i)$ $i, j = 1, 2$.

By using Lemma 2.1, we may rewrite $\hat{\beta}_{\text{GOME}}$ in (1.12) as follows:

$$\hat{\beta}_{\text{GOME}} = \hat{\beta}_G + S^{-1}R'(V + RS^{-1}R')^{-1}(r - R\hat{\beta}_G). \quad (2.1)$$

It is clear that from (2.1) and when the stochastic restrictions are correct,

$$\text{Cov}(\hat{\beta}_G) - \text{Cov}(\hat{\beta}_{\text{GOME}}) = \sigma^2 S^{-1}R'(V + RS^{-1}R')^{-1}RS^{-1}$$

is n.n.d. Therefore, the GOME is unbiased and has smaller variance compared with GLS.

For the sake, we are ready to introduce a new mixed estimator with stochastic restrictions when the heteroscedasticity and/or correlated errors are exists. This estimator simply coming by combining the GOME and GORR as follows

$$\begin{aligned} \hat{\beta}_{\text{GSRRE}}(k) &= \hat{\beta}_G(k) + S^{-1}R'(V + RS^{-1}R')^{-1}(r - R\hat{\beta}_G(k)) \\ &= (S + R'V^{-1}R)^{-1}(F_k X'W^{-1}Y + R'V^{-1}r), \end{aligned} \quad (2.2)$$

where $F_k = (S + kI)^{-1}S$ and $S^{-1}R'(V + RS^{-1}R')^{-1}r = (V + RS^{-1}R')^{-1}R'V^{-1}r$. In fact, $\hat{\beta}_{\text{GSRRE}}(k)$ is a general estimator which includes the GOME, GORR and GLS estimators as special cases, that is if $k = 0$, then

$$\hat{\beta}_{\text{GSRRE}}(0) = \hat{\beta}_{\text{GOME}}$$

if $k = 0$ and $R = 0_{q \times p}$, then

$$\hat{\beta}_{\text{GSRRE}}(0) = \hat{\beta}_G$$

if $R = 0_{q \times p}$, then

$$\hat{\beta}_{\text{GSRRE}}(k) = \hat{\beta}_G(k).$$

The expectation and the variance-covariance matrix of $\hat{\beta}_{\text{GSRRE}}(k)$ are

$$\begin{aligned} E\left(\hat{\beta}_{\text{GSRRE}}(k)\right) &= A^* (F_k S + R' V^{-1} R) \beta \\ &= \beta + A^* (F_k - I) S \beta, \end{aligned} \quad (2.3)$$

$$\text{Cov}\left(\hat{\beta}_{\text{GSRRE}}(k)\right) = \sigma^2 A^* (F_k S F_k' + R' V^{-1} R) A^* \quad (2.4)$$

where $A^* = (S + R' V^{-1} R)^{-1}$. From (2.3) we can obtain the bias of $\hat{\beta}_{\text{GSRRE}}(k)$

$$\text{Bias}\left(\hat{\beta}_{\text{GSRRE}}(k)\right) = A^* (F_k - I) S \beta. \quad (2.5)$$

The $\hat{\beta}_{\text{GSRRE}}(k)$ is always biased unless $k = 0$.

3. SUPERIORITY OF THE NEW ESTIMATOR

The mean square error matrix (MSEM) of any estimator β^* of β is defined as follows

$$\begin{aligned} \text{MSEM}(\beta^*) &= E(\beta^* - \beta)(\beta^* - \beta)' \\ &= \text{Cov}(\beta^*) + \text{Bias}(\beta^*) \text{Bias}(\beta^*)'. \end{aligned} \quad (3.1)$$

It is well known that when we have biased estimators, the MSEM is a good measure for checking the performance of any two estimators. Therefore, for any two estimators $\hat{\beta}_j$, $j = 1, 2$ of β , we said that $\hat{\beta}_1$ is superior to $\hat{\beta}_2$ with respect to MSEM if and only if

$$\Delta = \text{MSEM}\left(\hat{\beta}_1\right) - \text{MSEM}\left(\hat{\beta}_2\right)$$

is n.n.d.

For the sake, from (1.12), (1.13), (1.17) and (2.2) we obtain the MSEM of $\hat{\beta}_{\text{GOME}}$, $\hat{\beta}_G(k)$, $\hat{\beta}_{\text{GSLRE}}(d)$ and $\hat{\beta}_{\text{GSRRE}}(k)$ respectively as follows

$$\text{MSEM}\left(\hat{\beta}_{\text{GOME}}\right) = \sigma^2 A^* \quad (3.2)$$

$$\text{MSEM}\left(\hat{\beta}_G(k)\right) = \sigma^2 S^{-1}(k) S S^{-1}(k) + k^2 S^{-1}(k) \beta \beta' S^{-1}(k) \quad (3.3)$$

$$\text{MSEM}\left(\hat{\beta}_{\text{GSLRE}}(d)\right) = \sigma^2 A^* (F_d S F_d' + R' V^{-1} R) A^* + A^* (F_d - I) S \beta \beta' S (F_d - I) A^* \quad (3.4)$$

$$\text{MSEM}\left(\hat{\beta}_{\text{GSRRE}}(k)\right) = \sigma^2 A^* (F_k S F_k' + R' V^{-1} R) A^* + A^* (F_k - I) S \beta \beta' S (F_k - I) A^*. \quad (3.5)$$

The differences between the mean square error matrix of $\hat{\beta}_{\text{GOME}}$, $\hat{\beta}_G(k)$, $\hat{\beta}_{\text{GSRLE}}(d)$ and $\hat{\beta}_{\text{GSRRE}}(k)$ are obtained as follows

$$\begin{aligned}\Delta_1 &= \text{MSEM}\left(\hat{\beta}_{\text{GOME}}\right) - \text{MSEM}\left(\hat{\beta}_{\text{GSRRE}}(k)\right) = \sigma^2 D_1 - B_2 B_2' \\ \Delta_2 &= \text{MSEM}\left(\hat{\beta}_G(k)\right) - \text{MSEM}\left(\hat{\beta}_{\text{GSRRE}}(k)\right) = \sigma^2 D_2 + B_1 B_1' - B_2 B_2', \\ \Delta_3 &= \text{MSEM}\left(\hat{\beta}_{\text{GSRRE}}(k)\right) - \text{MSEM}\left(\hat{\beta}_{\text{GSRLE}}(d)\right) = \sigma^2 D_3 + B_2 B_2' - B_3 B_3', \\ \Delta_4 &= \text{MSEM}\left(\hat{\beta}_{\text{GSRLE}}(d)\right) - \text{MSEM}\left(\hat{\beta}_{\text{GSRRE}}(k)\right) = \sigma^2 D_4 + B_3 B_3' - B_2 B_2',\end{aligned}$$

where $D_1 = A^*(S - F_k S F_k') A^*$, $D_2 = S^{-1}(k) S S^{-1}(k) - A^*(F_k S F_k' - R' V^{-1} R) A^*$, $D_3 = A^*(F_k S F_k' - F_d S F_d') A^*$, $D_4 = A^*(F_d S F_d' - F_k S F_k') A^*$ and $S^{-1}(k) = (S + kI)^{-1}$. Since we want to check the performance of the new estimator, we should prove that Δ_1 , Δ_2 , Δ_3 and Δ_4 are n.n.d. Therefore, we can find the necessary and sufficient condition for that purpose.

Since $S = X'W^{-1}X$ is p.d., there exist an orthogonal matrix P such that $S = P\Lambda P'$ where $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_p\}$ is the diagonal matrix of the eigenvalues of S .

So, $S - F_k S F_k' = P\Gamma P'$, where $\Gamma = \text{diag}\{\gamma_1, \dots, \gamma_p\}$ is the diagonal matrix of $S - F_k S F_k'$. Now,

$$\begin{aligned}\gamma_i &= \lambda_i - \frac{\lambda_i^3}{(\lambda_i + k)^2} \\ &= \lambda_i \left(1 - \frac{\lambda_i^2}{(\lambda_i + k)^2}\right).\end{aligned}$$

Since $\lambda_i > 0$ and $k > 0$ for all $i = 1, 2, \dots, p$, it is clear that $S - F_k S F_k'$ is p.d., therefore, D_1 is p.d. We can state the following theorem after applying Lemma 2.2.

Theorem 3.1. *Under the linear regression model (1.11) with the stochastic linear restrictions (1.5) and for $k > 0$, Δ_1 is n.n.d. if and only if*

$$B_2' D_1^{-1} B_2 \leq \sigma^2.$$

Corollary. *Under the linear regression model (1.11) with the stochastic linear restrictions (1.5), when Theorem (1.1) holds, the $\hat{\beta}_{\text{GSRRE}}(k)$ is superior than of $\hat{\beta}_G$ in the mean square error matrix sense.*

By Lemma 2.3, we may rewrite D_2 as

$$D_2 = H'(I - \Omega)H,$$

where $\Omega = \text{diag}\{\omega_i^{S^{-1}(k)SS^{-1}(k)}(A^*(F_k S F_k' - R' V^{-1} R)A^*)\}$ the diagonal matrix of the eigenvalues of $A^*(F_k S F_k' - R' V^{-1} R)A^*$ in the $S^{-1}(k)SS^{-1}(k)$ and we denote it as $\omega_i(D_2)$.

So,

$$\begin{aligned}x' D_2 x &= x' H' (I - \Omega) H x \\ &= y' (I - \Omega) y \\ &= \sum_{i=1}^p (1 - \omega_i(D_2)) y_i^2,\end{aligned}$$

where $y = Hx$. Therefore, if $\omega_i(D_2) < 1$, then $x' D_2 x$ is p.d. Therefore, after applying Lemma 2.4 we can give the following theorem.

Theorem 3.2. *Under the linear regression model (1.11) with the stochastic linear restrictions (1.5). If $\omega_i(D_2) < 1$ then Δ_2 is n.n.d. if and only if*

$$(1 + d_{11})(d_{22} - 1) \leq d_{12}^2.$$

To prove that D_3 is p.d., we request to show that $(F_k S F_k' - F_d S F_d')$ is p.d. Consider d to be fixed. Therefore, $(F_k S F_k' - F_d S F_d')$ is p.d. if and only if

$$\begin{aligned} & \text{diag} \left\{ \frac{\lambda_i^3}{(\lambda_i + k)^2} - \frac{(\lambda_i + d)^2 \lambda_i}{(\lambda_i + 1)^2} \right\}_{i=1}^p \geq 0 \\ \Leftrightarrow & \frac{\lambda_i^3 (\lambda_i + 1)^2 - (\lambda_i + k)^2 (\lambda_i + d)^2 \lambda_i}{(\lambda_i + 1)^2 (\lambda_i + k)^2} \geq 0 \quad \forall i \\ \Leftrightarrow & \lambda_i (\lambda_i + 1) - (\lambda_i + k) (\lambda_i + d) \geq 0 \\ \Leftrightarrow & (\lambda_i + k) \leq \frac{\lambda_i (\lambda_i + 1)}{(\lambda_i + d)} \\ \Leftrightarrow & k \leq k^* = \frac{\lambda_i (1 - d)}{(\lambda_i + d)}. \end{aligned}$$

For the sake, for $k \leq k^* = \frac{\lambda_i(1-d)}{(\lambda_i+d)}$, D_3 will be p.d. After applying Lemma 2.4, we can state the following theorem.

Theorem 3.3. *Under the linear regression model (1.11) with the stochastic linear restrictions (1.5)*

1) *For $k \leq k^*$, Δ_3 will be n.n.d. if and only if*

$$(1 + d_{22})(d_{33} - 1) \leq d_{23}^2.$$

2) *For $k \geq k^*$, Δ_4 will be n.n.d. if and only if*

$$(1 + d_{33})(d_{22} - 1) \leq d_{32}^2$$

where $k^* = \frac{\lambda_i(1-d)}{(\lambda_i+d)}$.

Now, let us consider k to be fixed, by repeating the same proof of Theorem 3.3, we obtain the following theorem

Theorem 3.4. *Under the linear regression model (1.11) with the stochastic linear restrictions (1.5)*

1) *For $d \leq d^*$, Δ_4 will be n.n.d. if and only if*

$$(1 + d_{33})(d_{22} - 1) \leq d_{32}^2.$$

2) *For $d \geq d^*$, Δ_3 will be n.n.d. if and only if*

$$(1 + d_{22})(d_{33} - 1) \leq d_{23}^2$$

where $d^* = \frac{\lambda_i(1-k)}{(\lambda_i+k)}$.

We can note from our theorems that the comparison results depend on the unknown parameters β and σ^2 . In consequence of that, we can not exclude that our results obtained in the theorems will be held and the results

may be changeable. So, we replace them (β and σ^2) by their unbiased estimators. Since W is rarely known, the estimate of W can be used. Trenkler [13] gave some estimates of W as

$$W = (v_{ij}), v_{ij} = \rho^{|i-j|}, i, j = 1, 2, \dots, n \quad (3.6)$$

and

$$W = \frac{1}{1 + \rho^2} = \begin{pmatrix} 1 + \rho^2 & \rho & 0 & \cdots & 0 \\ \rho & 1 + \rho^2 & \rho & \cdots & 0 \\ 0 & \rho & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 + \rho^2 & \rho \\ 0 & 0 & \cdots & \rho & 1 + \rho^2 \end{pmatrix}$$

there are other estimates of W matrix are given by Firinguetti [3] and Bayhan and Bayhan [1].

4. NUMERICAL EXAMPLE

To illustrate our theoretical results, we consider a dataset which discussed by Bayhan and Bayhan [1]. Tables 1 and 2 give 75 weekly observations of sales. Sixty observations in Table 1 are taken as historical data and fifteen observations in Table 2 are taken from the last 15 weeks as a fresh data. These two Tables y_i and y_j denote weekly quantities of shampoos sold, while x_{i1}, x_{i2} and x_{j1}, x_{j2} denote the weekly list prices (averages from selected supermarkets) of the firm's shampoos and of a certain brand of soap, substituted from shampoos, respectively.

Using the data in Table 2, the matrix $X'X$ has eigenvalues $\lambda_1 = 41392.6, \lambda_2 = 0.5$. In the literature several diagnostic procedures are suggested for detecting multicollinearity. One of these diagnostic is the condition number $K = \lambda_{\max}/\lambda_{\min}$. Therefore the condition number in this case is $K = 91873.2$. The condition number here is very large, that means this data has strong multicollinearity. The Durbin-Watson statistics $d = \frac{\sum_{i=2}^n (\hat{e}_i - \hat{e}_{i-1})^2}{\sum_{i=1}^n \hat{e}_i^2}$ can be used to detect the presence of autocorrelation where \hat{e}_i denote the residuals of a linear regression model. If we use the data in Table 2 for computing d , we find that $d = 0.38$. For a significance level of 0.05 and for $n = 15$, the critical values of the Durbin-Watson statistic are $d_L = 0.95$ and $d_U = 1.54$. since $d < d_L$, it is concluded that autocorrelation is present for this data. The estimated value of W is given as follows: 0.72284, 0.42003, 0.28663, 0.15967, 0.10987, 0.16687, 0.20766, 0.20862, 0.17573, 0.17168, 0.12975, -0.02658, -0.13842 and -0.12330.

We may improve our estimators by adding stochastic linear restrictions to the model. consider the following stochastic linear restrictions:

$$R = \begin{pmatrix} 92.8 & 23.0 \\ 76.0 & 25.4 \end{pmatrix}, r = \begin{pmatrix} 33.144 \\ 33.294 \end{pmatrix}.$$

Also the matrix V is determined as:

$$V = \begin{pmatrix} 1 & 0.7228 \\ 0.7228 & 1 \end{pmatrix}$$

the ordinary mixed estimator (GOME) of the regression coefficients are:

$$\hat{\beta}_{\text{GOME}} = \begin{pmatrix} 0.091096 \\ 0.962606 \end{pmatrix}.$$

Therefore the estimator of σ^2 is obtained as

$$\hat{\sigma}^2 = \frac{(Y - X\hat{\beta}_{\text{GOME}})'(Y - X\hat{\beta}_{\text{GOME}})}{n - p} = 1.5297.$$

TABLE 1. Historical data for weekly sales of shampoos and prices.

Obs.	y_{i1}	x_{i1}	x_{i2}	Obs.	y_{i1}	x_{i1}	x_{i2}
1	28.445	49	12.5	31	31.446	84.4	21.2
2	28.547	49	12.5	32	31.549	85	21.2
3	28.644	51.2	13	33	31.641	85	21.2
4	28.746	51.2	13	34	31.743	78	20.1
5	28.849	40.3	13	35	31.848	78	20.1
6	28.94	52	13	36	31.94	81.3	20.1
7	29.045	52.3	13.8	37	32.043	83.1	21
8	29.142	58	14.4	38	32.146	83	21
9	29.248	58	14.4	39	32.25	88.5	22.3
10	29.25	58	14.4	40	32.344	88.5	22.3
11	29.443	62	16	41	32.441	88.5	22.3
12	29.545	62	16	42	32.545	68.7	22.9
13	29.644	62	16	43	32.643	68.7	22.9
14	29.747	52	17.1	44	32.748	91.3	22.9
15	29.841	67.2	17.1	45	32.842	91.3	22.9
16	29.045	67.2	17.1	46	32.95	91.3	23
17	30.046	67.2	18	47	33.039	92.8	23
18	30.142	67.2	18	48	33.144	92.8	23
19	30.245	72.4	18	49	33.249	76	25.4
20	30.348	72.4	18	50	33.347	76	26
21	30.441	72.4	18	51	33.442	93.4	24.1
22	30.549	72.4	21	52	33.543	93.4	24.1
23	30.641	80	21	53	33.647	93.4	24.1
24	30.739	72	18.3	54	33.746	96.3	24.1
25	30.849	72	18.3	55	33.849	96.3	24.3
26	30.949	55	19	56	33.94	97.2	24.3
27	31.051	48	19	57	34.041	97.2	24.3
28	31.148	80.1	19.4	58	34.143	75.2	25.1
29	31.245	80.1	21.2	59	34.248	100	25.1
30	31.342	84.4	21.2	60	34.345	101.5	25.4

To evaluate the value of the shrinkage parameter k in (2.2), Hoerl *et al.* (1975) proposed that an appropriate choice for k is:

$$\hat{k} = \frac{p\hat{\sigma}^{2*}}{\hat{\beta}'\hat{\beta}},$$

where $\hat{\sigma}^{2*}$ and $\hat{\beta}$ are found from the method of least squares. Using the method of least squares for the standard observations of the data in Table 2, we obtain $\hat{\sigma}^{2*} = 0.01379$ and $\hat{\beta} = (4.987 - 4.099)'$. Hence, $\hat{k} = 0.00065$.

Interpretation of the results

1. From Figure 1, when $k > 0$, the value of the trace MSEM (trMSEM) of the GSRRE is smaller than the value of the trace MSEM of GOME. Also, by calculating γ_i , we find that $\gamma_i = 0.0013000; 0.0012972$, $i = 1, 2$ is positive. Therefore Theorem 3.1 is held.
2. We compute the following values of $\omega_i(D_2)$: 0.4760; -0.0053. Since $\omega_i(D_2) < 1$, D_2 is p.d. Therefore using Theorem 3.2, the trMSEM value of GORR is 3.2859 which is greater than the trMSEM value of GSRRE 0.01811.

TABLE 2. Fresh data for weekly sales of shampoos and prices.

Obs.	y_j1	x_j1	x_j2
1	34.481	101.3	25.3
2	34.369	102	25.5
3	34.268	102.7	25.7
4	34.16	102.35	25.9
5	34.215	104.2	26.1
6	34.308	104.9	26.2
7	34.402	105.6	26.4
8	34.479	106.9	26.6
9	34.58	107	26.8
10	34.682	107.7	27
11	34.78	108.5	27.1
12	34.875	109.1	27.3
13	34.963	109.9	27.5
14	35.054	110.6	27.7
15	35.173	111.3	27.8

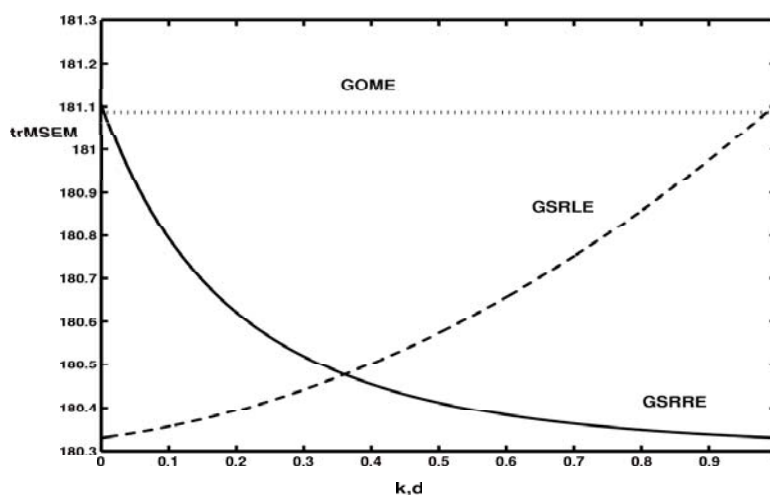


FIGURE 1. Estimated trMSEM values of the estimators GOME, GSRLE and GSRRE (y -axis $\times 10^4$).

- Let $d = 0.9$ be fixed. Using k_i^* in Theorem 3.3, we find the following values of k_i^* : 0.099; 0.0334. Comparing the MSEM value of GSRLE $\text{trMSEM}(\hat{\beta}_{\text{GSRLE}}(d = .9)) = 0.0182$ with $\text{trMSEM}(\hat{\beta}_{\text{GSRRE}}(k = 0.4)) = 0.0181$ for $k = 0.4 < \min(k_i^*) = 0.0334$. Since the condition in Theorem 3 (1.1) does not hold, $\text{trMSEM}(\hat{\beta}_{\text{GSRRE}}(k = 0.4)) = 0.0181 < \text{trMSEM}(\hat{\beta}_{\text{GSRLE}}(d = .9)) = 0.0182$. That means Δ_3 is not positive. We have $\text{trMSEM}(\hat{\beta}_{\text{GSRLE}}(d = .9)) = 0.0182$ with $\text{trMSEM}(\hat{\beta}_{\text{GSRRE}}(k = 0.1)) = 0.018$ for $k = 0.1 > \max(k_i^*) = 0.0999$. We see that $\text{trMSEM}(\hat{\beta}_{\text{GSRRE}}(k = 0.1)) = 0.018$ is smaller than $\text{trMSEM}(\hat{\beta}_{\text{GSRLE}}(d = .9)) = 0.0182$ as stated in Theorem 3 (1.2). That means Δ_4 is positive (see also Fig. 2).
- Let $k = 0.5$ be fixed. Using d_i^* in Theorem 3.4, we find the following values of d_i^* : 0.78; 0.554. For $d = 0.4 < \min(d_i^*) = 0.554$, we have the following inequality from Theorem 3.4 (1.1): $\text{trMSEM}(\hat{\beta}_{\text{GSRLE}}(d = .4)) = 0.01811 > \text{trMSEM}(\hat{\beta}_{\text{GSRRE}}(k = 0.5)) = 0.01806$. Since the condition in Theorem 3.4 (1.2) does

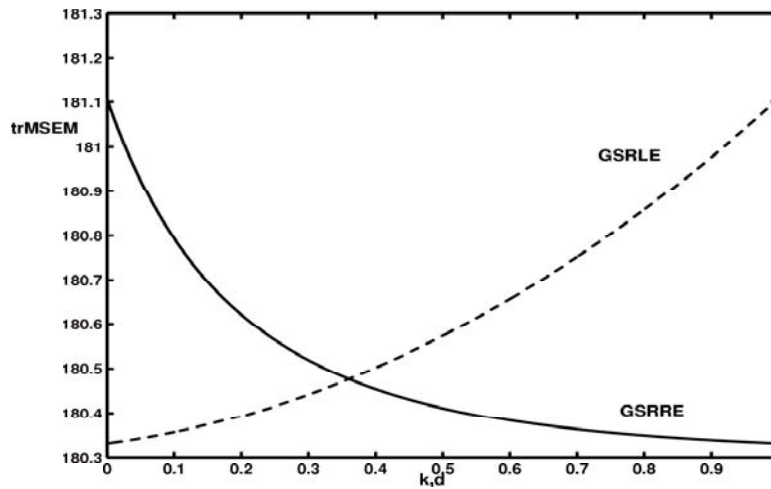


FIGURE 2. Estimated trMSEM values of the estimators GSRLE and GSRRE (y -axis $\times 10^4$).

not hold, $\text{trMSEM}(\hat{\beta}_{\text{GSRLE}}(d = .8)) = 0.018105$ is greater than $\text{trMSEM}(\hat{\beta}_{\text{GSRRE}}(k = 0.5)) = 0.01806$ which means that Δ_3 is not positive.

5. CONCLUSIONS

In this paper, a new estimator under linear regression model with heteroscedastic and/or correlated error when stochastic linear restrictions on the unknown parameter vector β available is introduced. The performance of the new estimator GSRRE against GLS, GORR and GSRLE is examined with respect to mean square error matrix sense. Under some conditions, the new estimator has mean square error matrix less than other estimators.

REFERENCES

- [1] G.M. Bayhan and M. Bayhan, Forecasting using autocorrelated errors and multicollinear predictor variables. *Comput. Ind. Eng.* **34** (1998) 413–421.
- [2] R.W. Farebrother, Further results on the mean square error of ridge regression. *J. R. Stat. Soc. B* **38** (1976) 284–250.
- [3] L. Firinguetti, A simulation study of ridge regression estimators with autocorrelated errors. *Commun. Stat. Simul.* **18** (1989) 673–702.
- [4] A.E. Hoerl and R.W. Kennard, Ridge Regression: Biased estimation for non-orthogonal problem. *Technometrics* **12** (1970) 55–67.
- [5] A.E. Hoerl and R.W. Kennard, Ridge Regression: Application for non-orthogonal problem. *Technometrics* **12** (1970) 69–82.
- [6] M.H. Hubert and P. Wijekoon, Improvement of the Liu estimator in linear regression model. *Statist. Papers* **47** (2006) 471–479.
- [7] K. Liu, A new class of biased estimate in linear regression. *Commun. Stat. – Theory Meth.* **22** (1993) 393–402.
- [8] C.R. Rao, *Linear Statistics Inference and its applications*. Second edn. John Wiley and Sons (1973).
- [9] C.R. Rao, H. Toubtenburg and S.C. Heumann, *Linear Models and Generalizations: Least squares and alternatives*. Springer Ser. Statist. Springer-Verlag, New York (2008).
- [10] C. Stein, Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, in *Proc. Third Berkeley Symp. on Mathematics, Statistics and Probability*. University of California, Berkeley, 1956, pp. 197–206.
- [11] H. Theil, On the use of incomplete prior information in regression analysis. *J. Am. Stat. Assoc.* **58** (1963) 401–414.
- [12] H. Theil and A.S. Goldberger, On pure and mixed estimation in econometrics. *Int. Econ. Rev.* **2** (1961) 65–78.
- [13] G. Trenkler, On the performance of biased estimators in the linear regression model with correlated or heteroscedastic errors. *J. Econometrics* **25** (1984) 179–190.
- [14] H. Yang and J. Xu, An alternative stochastic restricted Liu estimator in linear regression. *Statist. Papers* **50** (2007) 639–647.