A restricted r-k class estimator in the mixed regression model with autocorrelated disturbances

Shalini Chandra[a]*, Nityananda Sarkar[b]
[a]Centre for Mathematical Sciences, Banasthali University, Rajasthan-304022, INDIA
[b]Indian Statistical Institute, 203 B.T. Road, Kolkata-700108, INDIA

Abstract
In this paper, a new estimator called the restricted r-k class estimator, is introduced when the linear restrictions binding the regression coefficients are stochastic in nature, by combining the ordinary ridge regression estimator and principal component regression estimator for a regression model suffering from the problem of multicollinearity. The performance of the proposed r-k class estimator in the mixed regression model is compared with that of the mixed regression estimator and the stochastic ridge regression estimator in terms of the mean square error matrix criterion. Tests for verifying the conditions of dominance of the proposed estimator over the two others are also proposed. Furthermore, a Monte Carlo study and a numerical evaluation are carried out to study the performance of the tests involving conditions of superiority of the proposed estimator over the other two.

Keywords r-k class estimator, mixed regression estimator, stochastic linear constraints, autocorrelated errors.

Mathematics Subject Classification Primary 62J05; Secondary 62J07.

1 Introduction

In a linear regression model, the ill conditioning of the design matrix, arises, *inter alia*, due to multicollinearity, and consequently the ordinary least squares estimator (OLSE) becomes unreliable. So far, a number of alternative estimators have been proposed to overcome this problem. One popular technique is to consider suitable biased estimators, such as the ordinary ridge regression estimator (ORRE) proposed by Hoerl and Kennard(1970) and the principal component regression estimator (PCRE). Considering the fact that combination of two different estimators might inherit the advantages of both the estimators, Baye and

*This work was carried out when the first author was a visiting scientist at Indian Statistical Institute, North-East Center, Tezpur, Assam, India during July 2011-June 2012.

Parker (1984) proposed the r-k class estimator by combining the ORRE and PCRE. Likewise, Sarkar (1992) combined the ORRE and the restricted least squares estimator to introduce the restricted ridge regression estimator, and Kaçiranlar and Sakallioğlu (2001) introduced the r-d class estimator combining the estimator proposed by Liu and PCRE. The performance of such biased estimators have been evaluated by different criteria like the mean squared error (MSE), MSE matrix , Mahalanobis's D^2 and Pitman's closeness. Sarkar(1996), for instance, studied the properties of the r-k class estimator under the mean squared error matrix (MSEM) criterion and Özkale and Kaçiranlar (2008) compared the r-k class estimator with the OLSE under the Pitman's closeness criterion.

An alternative technique to deal with the multicollinearity problem is to consider parameter estimation with some restrictions on the unknown parameters, which may be exact or stochastic in nature. By grafting the ORRE into the restricted least squares estimator (RLSE) procedure, Sarkar (1992) introduced the restricted ridge regression estimator (RRRE). Özkale(2009) presented stochastic restricted ridge regression estimator with autocorrelated errors, and recently Xu and Yang (2011)have studied the properties of the r-k and r-d class estimators in the presence of exact linear restrictions under the MSE criterion.

In this paper, we propose to study the r-k class estimator in the presence of stochastic prior information. We shall compare the proposed estimator with the MRE and other competing estimators using the MSEM criterion. The rest of the paper is organized as follows: The new estimator is introduced in Section 2. The superiority of the proposed estimator by the MSEM criterion is discussed in Section 3. Section 4 describes the tests for verifying the conditions of dominance. A Monte Carlo simulation study is given to illustrate the performance of the tests for verifying the conditions of dominance in Section 5. A numerical evaluation is presented in Section 6. Finally, some concluding remarks are given in Section 7.

2 The Model and the Estimators

Let us consider the classical linear regression model

$$y = X\beta + u \tag{2.1}$$

where y is an $n \times 1$ vector of observations on the variable to be explained, X is an ill-conditioned $n \times p$ matrix of n observations on p explanatory variables, β is a $p \times 1$ vector of regression coefficients associated with them and u is an $n \times 1$ vector of disturbances, the elements of which are assumed to have mean zero and variance covariance matrix $\sigma^2\Omega$ where Ω is an $n \times n$ known positive definite (p.d.) and symmetric matrix. The generalized least squares estimator

(GLSE) of β is given by

$$\widehat{\beta}_g = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y \tag{2.2}$$

which is unbiased and has minimum variance. If the explanatory variables used in the regression model are collinear, the matrix $(X'\Omega^{-1}X)$ in (2.2) may become rank deficient. For that situation, Trenkler(1984) proposed the ridge estimator of β as

$$\widehat{\beta}_g(k) = (X'\Omega^{-1}X + kI)^{-1}X'\Omega^{-1}y, k > 0 \tag{2.3}$$

which was obtained in a way similar to that of Hoerl and Kennard who derived the ORRE for the classical linear regression model.

Alternatively, the use of prior information also solves the problem of multicollinearity. Xu *et al.*(2011) used the exact prior information available in the form of linear restrictions and proposed the restricted *r-k* and *r-d* class of estimators. If uncertainty exists about the prior information specifications, one alternative is to make use of stochastic linear restrictions given by

$$r = R\beta + \epsilon \tag{2.4}$$

where r is a known $m \times 1$ random vector, R is a known $m \times p$ matrix of prior information of rank $m \leq p$ and ϵ is a random vector independent of u, with $E(\epsilon) = 0$ and $Var(\epsilon) = \sigma^2 W$, where W is a known p.d. matrix.

Theil and Goldberger (1961) introduced the mixed regression estimator (MRE) by unifying the sample information and the prior information in (2.4) in a single model as

$$\begin{pmatrix} y \\ r \end{pmatrix} = \begin{pmatrix} X \\ R \end{pmatrix} \beta + \begin{pmatrix} u \\ \epsilon \end{pmatrix}. \tag{2.5}$$

A necessary assumption is to suppose that the two random errors are uncorrelated to each other i.e., $E(u\epsilon') = 0$. The augmented matrices and vectors in (2.5) can be written, in a more compact form, as

$$y_M = X_M\beta + u_M \tag{2.6}$$

with u_M following $(0, \sigma^2 W_M)$ where

$$W_M = \begin{pmatrix} \Omega & 0 \\ 0 & W \end{pmatrix} \tag{2.7}$$

is a p.d. matrix. Since rank of the augmented matrix X_M is assumed to be full, the model in (2.6) is a generalized linear model. Following Theil *et al.* (1961), the MRE is given by

$$\widehat{\beta}_M = (X'\Omega^{-1}X + R'W^{-1}R)^{-1}(X'\Omega^{-1}y + R'W^{-1}r). \tag{2.8}$$

It is desirable, from our consideration, to consider the case where the variables forming a augmented matrix X_M are correlated. For that situation, Özkale(2009) combined the approaches of obtaining the mixed estimator and the ORRE in a linear regression model and called that estimator a stochastic restricted ridge regression estimator (SRRRE), given by

$$\widehat{\beta}_m(k) = (X'\Omega^{-1}X + R'W^{-1}R + kI_p)^{-1}(X'\Omega^{-1}y + R'W^{-1}r), k > 0. \quad (2.9)$$

In the next section, following Sarkar(1992) and Özkale(2009), we propose the *r-k* class estimator in the presence of stochastic linear constraints so that another estimator with improved performance in situation of multicollinearity can be obtained.

2.1 The Proposed Estimator

In order to propose such an estimator, we first obtain the PCRE for the model given in (2.6). Since W_M is a p.d. matrix, there exists a nonsingular matrix E such that $W_M = E'E$. Premultiplying both sides of (2.6) by E^{-1}, we get,

$$E^{-1}y_M = E^{-1}X_M\beta + E^{-1}u_M$$

which is redesignated as

$$y^* = X^*\beta + u^* \quad (2.10)$$

with $*$ indicating the respective transformations with the property that u^* follows $(0, \sigma^2 I)$. Thus, the model in (2.10) reduced to a classical linear regression model. To obtain the PCRE for the model in (2.10), suppose T is an $p \times p$ orthogonal matrix i.e., $TT' = I$ and $X^{*\prime}X^* = S^*$. Suppose $T = (t_1, t_2, \ldots, t_p)$ such that it diagonalizes S^* i.e., $T'S^*T = T'(X'\Omega^{-1}X + R'W^{-1}R)T = G$, say, where $G = diag(g_1, g_2, \ldots, g_p)$, with $g_1, g_2, \ldots, g_p$ being the eigenvalues of $(X'\Omega^{-1}X + R'W^{-1}R)$. Further, let $T_r = (t_1, t_2, \ldots, t_r)$ where $r \leq p$, and hence $T_r'(X'\Omega^{-1}X + R'W^{-1}R)T_r = G_r = diag(g_1, g_2, \ldots, g_r)$, and $T_{p-r}'(X'\Omega^{-1}X + R'W^{-1}R)T_{p-r} = G_{p-r}$ where $G_{p-r} = diag(g_{r+1}, g_{r+2}, \ldots, g_p)$ and $T_{p-r} = (t_{r+1}, t_{r+2}, \ldots, t_p)$. Therefore, the model in (2.10) can be written as

$$y^* = X^*TT'\beta + u^*,$$

which, in turn, can be expressed as

$$y^* = Z_r\alpha_r + Z_{p-r}\alpha_{p-r} + u^* \quad (2.11)$$

where $Z_r = X^*T_r$, $Z_{p-r} = X^*T_{p-r}$, $\alpha_r = T_r'\beta$ and $\alpha_{p-r} = T_{p-r}'\beta$. The OLSE of α_r with Z_{p-r} omitted is $\widehat{\alpha}_r = (Z_r'Z_r)^{-1}Z_r'y^*$. The PCRE of β based on the model (2.11) can be obtained by transforming it into the original variables, and this is given by

$$\widehat{\beta}_M(r) = T_r(T_r'X^{*\prime}X^*T_r)T_r'X^{*\prime}y^*. \quad (2.12)$$

Substituting the values of X^* and y^* in (2.12), we get

$$\widehat{\beta}_M(r) = T_r[T_r'(X'\Omega^{-1}X + R'W^{-1}R)T_r]^{-1}T_r'(X'\Omega^{-1}y + R'W^{-1}r). \quad (2.13)$$

Following Baye and Parker (1984), a combination of (2.9) and (2.13) yields a new r-k class estimator under stochastic restrictions. This estimator is given below in (2.14) as

$$\widehat{\beta}_M(r,k) = T_r[T_r'(X'\Omega^{-1}X + R'W^{-1}R)T_r + kI_r]^{-1}T_r'(X'\Omega^{-1}y + R'W^{-1}r). \tag{2.14}$$

This is a general estimator which has special cases as:

(i) for $r = p$, $\widehat{\beta}_M(r,k) = \widehat{\beta}_M(k)$, which is the SRRRE,

(ii) for $r = p$, $k \to 0$, $\widehat{\beta}_M(r,k) = \widehat{\beta}_M$, which is the MRE,

(iii) for $\Omega = I$, $\lim_{W\to 0} \widehat{\beta}_M(r,k) = \widehat{\beta}_R(r,k)$, (see Rao et al.(1999)), which is the r-k class estimator under exact restrictions.

In the next section, the MSEM comparisons of the proposed estimator with those of mixed estimator and other biased estimators are presented.

3 Comparisons by mean square error matrix

It is well-known that the MSE matrix (MSEM) of an estimator $\widehat{\beta}$ of β is defined as

$$M(\widehat{\beta}) = E[(\widehat{\beta} - \beta)(\widehat{\beta} - \beta)'] = D(\widehat{\beta}) + [bias(\widehat{\beta})][[bias(\widehat{\beta})]' \tag{3.1}$$

where $D(\widehat{\beta})$ stands for the variance-covariance matrix of $\widehat{\beta}$. From (2.14), we find that

$$Bias(\widehat{\beta}_M(r,k)) = (T_rG_{rk}T_r'S^* - I_p)\beta \tag{3.2}$$

where $G_{rk} = [G_r + kI_r]^{-1}$, and

$$Var(\widehat{\beta}_M(r,k)) = \sigma^2 T_rG_{rk}G_rG_{rk}T_r'. \tag{3.3}$$

Therefore, the MSE of $\widehat{\beta}_M(r,k)$ is given by

$$M(\widehat{\beta}_M(r,k)) = \sigma^2 T_rG_{rk}G_rG_{rk}T_r' + (T_rG_{rk}G_rT_r' - I_p)\beta\beta'(T_rG_{rk}G_rT_r' - I_p)'. \tag{3.4}$$

The MSE of MRE ($\widehat{\beta}_M$), SRRRE ($\widehat{\beta}_M(k)$), the restricted r-k class estimator ($\widehat{\beta}_R(r,k)$) can be obtained from (3.4) by putting $r = p, k = 0; r = p$; and $\Omega = I, W = 0$, respectively.

3.1 MSE matrix comparison between $\widehat{\beta}_M(r,k)$ and $\widehat{\beta}_M$

We can obtain MSEM of $\widehat{\beta}_M$ from (3.4) by putting $r = p$ and $k = 0$ and this is obtained as

$$M(\widehat{\beta}_M) = \sigma^2 TG^{-1}T'. \tag{3.5}$$

In order to find out under what condition $\widehat{\beta}_M(r,k)$ is better than $\widehat{\beta}_M$, we take the difference, $\triangle_1$, say, between $M(\widehat{\beta}_M)$ and $M(\widehat{\beta}_M(r,k))$. Thus,

$$\begin{aligned}\triangle_1 = M(\widehat{\beta}_M) - M(\widehat{\beta}_M(r,k)) = \sigma^2(TG^{-1}T' - T_rG_{rk}G_rG_{rk}T_r') \\ - (T_rG_{rk}G_rT_r' - I_p)\beta\beta'(T_rG_{rk}G_rT_r' - I_p)'\end{aligned} \tag{3.6}$$

Following Xu *et al.*(2011) and Sarkar(1996), the equation in (3.6) can be simplified to

$$\begin{aligned}\triangle_1 = {} & k^2T_rG_{rk}[\sigma^2(\frac{2}{k}I_r + G_r) - T_r'\beta\beta'T_r)G_{rk}T_r' \\ & + T_{p-r}[\sigma^2G_{p-r} - T_{p-r}'\beta\beta'T_{p-r}]T_{p-r}' \\ & - kT_rG_{rk}T_r'\beta\beta'T_{p-r}T_{p-r}' \\ & - kT_{p-r}T_{p-r}'\beta\beta'T_rG_{rk}T_r',\end{aligned}$$

which can be compactly expressed as

$$\triangle_1 = TS_k^{-1}[\sigma^2(G^*)^{-1} - T'\beta\beta'T]G_k^{-1}T' \tag{3.7}$$

where

$$S_k = \begin{pmatrix} \frac{1}{k}(G_{rk})^{-1} & 0 \\ 0' & I_{p-r} \end{pmatrix}$$

and

$$(G^*)^{-1} = \begin{pmatrix} \frac{2}{k}I_r + G_r^{-1} & 0 \\ 0' & G_{p-r}^{-1} \end{pmatrix}.$$

From (3.7), it is clear that $\triangle_1 \geq 0$ if and only if $\sigma^2(G^*)^{-1} - T'\beta\beta'T$ is a non negative definite (n.n.d.) matrix. Now, $\sigma^2(G^*)^{-1} - T'\beta\beta'T$ is a n.n.d. matrix if and only if $\beta'TG^*T'\beta \leq \sigma^2$. Further, by substituting the values for G^*, the condition of dominance can be further simplified, which is stated in the form of the following theorem.

Theorem 3.1 *A necessary and sufficient condition for the dominance of $\widehat{\beta}_M(r,k)$ over $\widehat{\beta}_M$ by the MSEM criterion is given by*

$$[\beta'T_r(\frac{2}{k}I_r + G_r^{-1})^{-1}T_r'\beta + \beta'T_{p-r}G_{p-r}T_{p-r}'\beta] \leq \sigma^2 \tag{3.8}$$

for $k > 0$.

In the absence of the restrictions and for $\Omega = I$, the condition of dominance stated in (3.8) is same as that of obtained by Sarkar(1996) for the superiority of the *r-k* class estimator over the OLSE under MSEM criterion.

3.2 Comparison between $\widehat{\beta}_M(r,k)$ and $\widehat{\beta}_M(k)$

The MSEM of $\widehat{\beta}_M(k)$ can be obtained from (3.4) by putting r=p . It is given by

$$M(\widehat{\beta}_M(k)) = \sigma^2 G_k G G_k + k^2 G_k \beta\beta' G_k \tag{3.9}$$

where $G_k = G + kI_p$. In order to compare the MSE of both the estimators, we consider the difference, $\triangle_2$, as given below. It is easy to check that $\triangle_2$ can be expressed as

$$\begin{aligned}\triangle_2 = M(\widehat{\beta}_M(k)) - M(\widehat{\beta}_M(r,k)) &= \sigma^2[TG_kGG_kT' - T_rG_{rk}GG_{rk}T_r'] \\ &+ (TG_kGT' - I_p)\beta\beta'(TG_kGT' - I_p)' \\ &- (T_rG_{rk}G_rT_r' - I_p)\beta\beta'(T_rG_{rk}G_rT_r' - I_p)'.\end{aligned} \tag{3.10}$$

For the superiority of $\widehat{\beta}_M(r,k)$ over $\widehat{\beta}_M(k)$, $\triangle_2$ has to be a n.n.d matrix for all $k > 0$. Note that the expression (3.10) can be written as

$$\triangle_2 = A + a_1a_1' + a_2a_2' \tag{3.11}$$

where

$$A = \sigma^2[TG_kGG_kT' - T_rG_{rk}GG_{rk}T_r']$$

$$a_1 = (TG_kGT' - I_p)\beta$$
$$and, a_2 = (T_rG_{rk}G_rT_r' - I_p)\beta$$

After some algebraic simplifications, A can be written as

$$A = \sigma^2[T_{p-r}G_{(p-r)k}G_{p-r}G_{(p-r)k}T_{p-r}']$$

In order to find the condition for superiority of $\widehat{\beta}_M(r,k)$ over $\widehat{\beta}_M(k)$, we would make use of a theorem by Baksalary and Trenkler (1991, p.172) which is given below.

Theorem 3.2 *Let $\zeta_{n\times p}$ be the set of $n \times p$ complex matrices and let H_n be the set of $\zeta_{n\times n}$consisting of Hermitian matrices. Further, given $L \in \zeta_{n\times p}$, the symbols L^*, $\Re(L)$ and $\varsigma(L)$ stand for the conjugate transpose, the range and the set of all generalized inverses, respectively of L. Now, let $A \in H_n$, a_1 and $a_2 \in \zeta_{n\times 1}$ be linearly independent, $f_{ij} = a_i^*A^-a_j$, $i, j = 1, 2$ for $A^- \in \varsigma(A)$, and $s = \frac{[a_1^*(I_n-AA^-)^*(I_n-AA^-)a_2]}{[a_1^*(I_n-AA^-)*(I_n-AA^-)a_1]}$, provided that $a_1 \notin \Re(A)$. Then $A + a_1a_1^* - a_2a_2^*$ is n.n.d. if and only if any one of the following sets of conditions hold.*

1. *A is n.n.d., $a_1 \in \Re(A)$, $a_2 \in \Re(A)$ and $(f_{11} + 1)(f_{22} - 1) \le |f_{12}|^2$;*
2. *A is n.n.d., $a_1 \notin \Re(A)$, $a_2 \in \Re(A : a_1)$ and $(a_2 - sa_1)^*A^-(a_2 - sa_1) \le 1 - |s|^2$;*

3. $A = UU^* - \delta vv^*$, $a_1 \in \Re(A)$, $a_2 \in \Re(A)$, $v^* a_1 \neq 0$ *and* $(f_{11}+1) \leq 0$, $(f_{22} - 1 \leq 0)$, $(f_{11}+1)(f_{22}-1) \leq |f_{12}|^2$*;*

where (U:v) (with U possibly absent) is a sub unitary matrix, $\triangle$ *is a p.d. diagonal matrix (occurring when U is present) and* δ *is a positive scalar. Further, the conditions* $(a) - (c)$ *are all independent of the choice of* $A^- \in \zeta(A)$*.*

Now, the Moore Penrose inverse of A , say A^+, is given by

$$A^+ = \sigma^{-2}[T_{p-r} G^{-1}_{(p-r)k} G^{-1}_{p-r} G^{-1}_{(p-r)k} T'_{p-r}]$$

and we have $AA^+ = T_{p-r}T'_{p-r} = I_p - T_r T'_r$. From the above equations, it could be easily seen that $a_1 \in R(A)$ iff $a_1 = 0$ where R(A) is the range of A. Furthermore, $a_2 = a_1 + A\eta$, where

$$\eta = -\sigma^{-2}[T_{p-r} G^{-1}_{(p-r)k} T'_{p-r}]\beta$$

which indicates that $a_2 \in R(A : a_1)$. Since in converse $a_1 \notin A$ and $a_2 \in R(A : a_1)$, we note that the second conditions stated in (b) of the theorem is satisfied and we can conclude that $\triangle_2$ is a n.n.d. matrix iff

$$(a_2 - sa_1)' A^+ (a_2 - sa_1) \leq 1 - (\det s)^2 \tag{3.12}$$

where

$$s = \frac{a'_1 (I - AA^+)'(I - AA^+) a_2}{a'_1 (I - AA^+)'(I - AA^+) a_1}.$$

In our case, s=1. Therefore, a necessary and sufficient condition for $\triangle_2$ to be a n.n.d. matrix as given in (3.12), reduces to $(a_2 - a_1)' A^+ (a_2 - a_1) \leq 0$ i.e., $\eta' AA^+ A\eta \leq 0$. However, it is noteworthy that since $A^+$is a n.n.d. matrix, the condition of dominance turns out to be $\eta' AA^+ A\eta = 0$. Substituting the value of η, the condition simplifies to

$$\beta' T_{p-r} G_{p-r} T'_{p-r} \beta = 0, \tag{3.13}$$

which holds iff $T'_{p-r}\beta = 0$.

Theorem 3.3 *A necessary and sufficient condition for the dominance of* $\widehat{\beta}_M(r,k)$ *over* $\widehat{\beta}_M(k)$ *by the MSEM criterion is* $T'_{p-r}\beta = 0$*.*

When $R = 0$ and $\Omega = I$, the criterion in (3.13) reduces to the condition obtained by Sarkar(1996) for the superiority of *r-k* class estimator over ORRE under the MSEM criterion.

4 Tests for Verifying the Conditions

In this section tests are proposed for verifying the conditions stated in Theorems 3.1 and 3.3 in a given situation since the dominance conditions stated in these two theorems depend on the unknown parameters β and σ^2.

Now, looking carefully for the necessary and sufficient conditions for the dominance of $\widehat{\beta}_M(r,k)$ over $\widehat{\beta}_M$ as well as $\widehat{\beta}_M(k)$, it can be clearly seen that it is not possible to give test for verifying the condition to hold stated in Theorem 3.1. However, it is possible to give test for a condition which is sufficient one for the condition of Theorem 3.1 to hold. Since, $G - G^*$ is a n.n.d. matrix, the condition in Theorem 3.1 can be restated in terms of G, *viz,* $\beta' TG^*T'\beta \leq \sigma^2$ holds if $\beta' TGT'\beta \leq \sigma^2$. The test statistic for testing the two conditions stated in Theorems 3.1 and 3.3, are given below.

4.1 Testing between $\widehat{\beta}_M(r,k)$ and $\widehat{\beta}_M$

To test the null hypothesis $H_{01} : \frac{\beta' TGT'\beta}{\sigma^2} \leq 1$, it may be noted that since $\widehat{\beta}_M \sim N(\beta, \sigma^2 TG^{-1}T')$ under the assumption of normality of errors in (2.1) and (2.4), $G^{1/2}T'\widehat{\beta}_M \sim N(G^{1/2}T'\beta, \sigma^2 I_p)$. Thus, we have the following test statistic for testing the null hypothesis H_{01} against $H_{11} : \frac{\beta' TGT'\beta}{\sigma^2} > 1$;

$$F_1 = \frac{\widehat{\beta}'_M TGT'\widehat{\beta}_M/m}{e'e/n-p} \tag{4.1}$$

where $e'e/(n-p)$ is the GLSE of σ^2 with e being the vector of residuals. Here, F_1 follows a non-central F distribution with d.f. m and $n-p$ with non-centrality parameter $\lambda = \frac{\beta' TGT'\beta}{\sigma^2}$ where $\lambda \leq 1$ under null (H_{01}) and $\lambda > 1$ for alternative hypothesis (H_{11}).

4.2 Test involving $\widehat{\beta}_M(r,k)$ and $\widehat{\beta}_M(k)$

To test the condition stated in Theorem 3.3, the null hypothesis is $H_{02} : T'_{p-r}\beta = 0$. Under the assumption of normality of the disturbances in (2.1) and (2.4), it follows from (3.2) and (3.3) that

$$\widehat{\beta}_M(r,k) \sim N(T_r G_{rk} T'_r S^* \beta, \sigma^2 T_r G_{rk} G_r G_{rk} T'_r),$$

and

$$T'_{p-r}\widehat{\beta}_M(r,k) \sim N(G_{(p-r)k} T'_{p-r} S^* \beta, \sigma^2 G_{(p-r)k} G_p - r G_{(p-r)k}). \tag{4.2}$$

Since, $G_{(p-r)k} T_p - r' S^* \beta = G_{(p-r)k} T_p - r'(T_r G_r T'_r + T_{p-r} G_{p-r} T'_{p-r})\beta = G_{(p-r)k} G_{p-r} T'_{p-r}\beta$ is an unbiased estimator of $T'_{p-r}\beta$ under H_{02}(see Sarkar(1996)). Therefore, the test statistic for testing H_{02} is given by

$$F_2 = \frac{\widehat{\beta}_M(r,k)' T_{p-r} (G_{(p-r)k} G_{p-r} G_{(p-r)k})^{-1} T'_r \widehat{\beta}_M(r,k)/p-r}{e'e/n-p}. \tag{4.3}$$

Clearly, F_2 follows a F distribution with d.f. *p-r* and $n-p$ under H_{02}.

5 A Simulation Study

In this section, we present the results of a simulation study undertaken for verifying the conditions of dominance of $\widehat{\beta}_M(r,k)$ over each of $\widehat{\beta}_M(k)$ and $\widehat{\beta}_M$ under the MSEM criterion. In this study, the X matrix is generated following the method given by Newhouse (1971).

$$X_{ij} = U + \alpha V_{ij}, \quad i = 1, 2 \ldots n, \quad and \quad j = 1, 2 \ldots p \tag{5.1}$$

where U and V_{ij} are independent sequences of standard random normal variates. Here α is so chosen as to result in a desired value of the sample correlation (ρ) among explanatory variables, and the relationship involving α and ρ is given by $\alpha = \sqrt{\frac{1-\rho}{\rho}}$. For this study, we consider the values for ρ to be $0.7, 0.8$ and 0.9. The normalized eigenvector corresponding to the largest eigenvalue of a matrix G is chosen as a coefficient vector. Observations on the dependent variable y is then generated by

$$y_i = \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots \beta_p X_{ip} + u_i, \quad i = 1, 2, \ldots n.$$

The R matrix is generated with its columns as independent standard normal variates. The r vector is generated by

$$r_i = \beta_1 R_{i1} + \beta_2 R_{i2} + \ldots \beta_p R_{ip} + v_i, \quad i = 1, 2, \ldots m.$$

Following Firinguetti(1970) and Judge *et al.*(1985), both u_i and v_i are generated by two different autocorrelation process *viz.*, AR(1) and MA(1) processes and these are given as

$$u_i = \phi u_{i-1} + e_{1i}, \quad and \quad v_j = \phi v_{j-1} + e_{2j}, |\phi| < 1, \quad i = 1, 2, \ldots n; \quad and \quad j = 1, 2 \ldots m, \tag{5.2}$$

and

$$u_i = \tau e_{1i-1} + e_{1i}, \quad and \quad v_j = \tau e_{2j-1} + e_{2j}, |\tau| < 1, \quad i = 1, 2, \ldots n; \quad and \quad j = 1, 2 \ldots m, \tag{5.3}$$

where e_1 and e_2 are independent normal random variables with mean 0 and same variance σ_e^2. The Ω matrix and the error variances under the two processes are given by

$$\Omega = ((\omega_{ij})), \quad \omega_{ij} = \phi^{|i-j|}, \quad and \quad \sigma^2 = \frac{\sigma_e^2}{1-\phi^2}; \quad i, j = 1, 2 \ldots n; \tag{5.4}$$

and

$$\Omega = \frac{1}{1+\tau^2} \begin{pmatrix} 1+\tau^2 & \tau & 0 & 0 & \ldots & 0 & 0 \\ \tau & 1+\tau^2 & \tau & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & \tau & 1+\tau^2 \end{pmatrix} \tag{5.5}$$

and $\sigma^2 = \sigma_e^2(1+\tau^2)$ for the equations (5.2) and (5.3) respectively. The number of explanatory variables, p, in the study is taken to be 3, the size of sample as 10. The number of extraneous information affects the performance of mixed estimators, therefore variation in m will also be taken into consideration. The value of σ is taken as 10, and that of r to be 2. W matrix can be obtained from Ω matrix after its construction. The values of the test statistics for testing the conditions stated in Theorems 3.1 and 3.3 are calculated for $k = 0.001, 0.01, 0.1, 0.5, 0.7, 0.9, 1$. This process has been repeated for 1000 times and the proportion of the cases for which the relevant null hypothesis is not rejected, is reported below.

$\widehat{\beta}_M(r,k)$ and $\widehat{\beta}_M$

For testing the condition for the dominance of $\widehat{\beta}_M(r,k)$ over the mixed regression estimator,$\widehat{\beta}_M$, the test statistic is F_1 and it follows a non-central F distribution with degrees of freedom m and $n-p$ and a non-centrality parameter λ. Following Johnson and Kotz(2004), a non-central F distribution i.e., $F_{(m,n-p)}(\lambda)$ can be approximated by $(1+(\lambda/m))\ F_{(m^*,n-p)}$ where $m^* = \frac{(m+\lambda)^2}{m+2\lambda} = m + \frac{\lambda^2}{m+2\lambda}$ is greater than m and approximated to the nearest integer value. Thus, for a test of size α,

$$P_{\lambda=1}(F_{(m,n-p)}(\lambda) > C_r) = \alpha,$$

then $P_\lambda(F_{(m,n-p)}(\lambda) > C_r) \leq \alpha$ for all $\lambda \leq 1$ where C_r is the critical value at $\alpha\%$ level of significance. Hence, we carry out test with the value of the non-centrality parameter $\lambda = 1$ (see Sarkar and Chandra(2012), for details).

The value of the first degree of freedom of F distribution, m^*, and the critical values of the approximate central F distribution at 5% level of significance are given in Table 1.

Table 1: Approximation of non-central F distribution with central F distributions

m	m^*	Approx. central F distribution	Critical value at 5% level of significance
3	3.20	F(3,7)	4.35
5	5.30	F(5,7)	3.97

For testing the null hypothesis H_{01} against the alternative H_{11}, the value of the test statistic in (4.1) is computed. The proportion of number of times when the null hypothesis is not rejected at 5% level of significance, say P_{01}, is reported in Table 2.

From Table 2, it can seen that the proportions of cases when H_{01} is not rejected at 5% level of significance are quite high for both $m = 3$ and $m = 5$, and for all values of ρ considered here. Therefore, it can be concluded that it

Table 2: Proportion of times (P_{01}) when the null hypothesis is not rejected at 5% level of significance

	P_{01}		
m/ ρ	**0.70**	**0.80**	**0.90**
3	0.988	0.984	0.838
5	0.987	0.993	0.928

is very likely that the condition stated in Theorem 3.1 holds true i.e., $\widehat{\beta}_M(r,k)$ estimator dominates MRE under MSEM criterion at 5% level of significance. In fact, the same conclusion holds at 1% level of significance as well.

$\widehat{\beta}_M(r,k)$ and $\widehat{\beta}_M(k)$

For testing the condition of dominance of $\widehat{\beta}_M(r,k)$ estimator over stochastic restricted ridge regression estimator (SRRRE), the values of the test statistic F_2 given in (4.3) is computed. When compared with the critical value of $F(1,7)$ at 5% level of significance, the proportions of number of times the null hypothesis H_{02} is not rejected, is found to be 1 for all values of k and ρ considered here. It thus suggests that $\widehat{\beta}_M(r,k)$ estimator performs better than SRRRE under MSEM criterion at 5% level of significance. The same conclusion holds for 1% level of significance also.

The simulation study has been carried out for both the error processes and for various possible values of ϕ and τ, and it is found that the results do not show much variations. Therefore, simulation results has been reported in this section only for AR(1) process for both u_i and v_i with $\phi = 0.5$.

6 Numerical Illustration

The quarterly US data on GDP growth (y), personal disposable income (X_1), personal consumption expenditure (X_2), corporate tax after profits (X_3) for the years 1971-1978 and 1982-1991 have been taken from Gujarati (2002). It is found that the independent variables show collinear structure. The data for the year 1971-1978 is taken as the old data set providing stochastic prior information and the second data set is used as a new data to verify the tests for the conditions of dominance obtained in Section 4. The old data is used to estimate the variance covariance matrix of the errors and also a subset of this data (2 observations) is taken as extraneous information in the mixed model.

The data set has been standardized. To detect the presence of autocorrelation, the Durbin-Watson (DW) test statistic is applied. The value of the DW statistic for the new data set comes out to be 1.2845, which indicates the presence of autocorrelation at a significance level 0.01 with the critical values $d_L = 1.20$ and $d_U = 1.40$ for $n = 40$. For the selection of the form of the error structure, minimum Akaike's Information Criterion (AIC) is used which confirms that the error structure of the new data follows AR(1) process. The same procedure has been applied for the old data set and it is found that the old data set also has the error structure following an AR(1) process.

The Ω matrix will be estimated by the equation (5.4) from the old data with ϕ estimated as 0.7471, so that it can be proclaimed that it is a known matrix (see Bayhan and Bayhan(1998)). To check for the multicollinearity in the explanatory variables, the condition index (CI) of $X'\Omega^{-1}X$ matrix is computed and it is found out to be 124.65, which is a very high value. Therefore, it can be concluded that the data are autocorrelated and multicollinear as well.

Now, insofar as stochastic linear restrictions are concerned, we note that these arise from prior statistical information, usually in the form of previous estimates of parameters, and take the form of an additional linear model, therefore the old data set can be used to provide prior information. Here, the 3^{rd} and 4^{th} observations from the standardized old data are taken to form the R matrix and r vector, and these are given by

$$R = \begin{pmatrix} 5.7958965 & -1.438879 & -1.299954 \\ -0.3410506 & -1.458109 & -1.379626 \\ & & \end{pmatrix}$$

$$r = \begin{pmatrix} -1.415274 \\ -1.5056155 \end{pmatrix}.$$

Following Bayhan and Bayhan (1998) and Özkale (2009), the W matrix is taken as first two rows and columns of Ω matrix. The value of σ^2 is estimated by using GLS estimate of β in (2.2) which comes out to be 0.48951. Here, r is taken to be 2.

$\widehat{\beta}_M(r,k)$ versus $\widehat{\beta}_M$

The value of the test statistic in (4.1) for testing H_{01} against H_{11} comes out to be around 1. When compared with the critical value 3.26 of $F_{2,37}$ at 5% level of significance, the value of the statistic suggests that the null hypothesis is not rejected at 5% level of significance. Hence, it can be concluded that $\widehat{\beta}_M(r,k)$ estimator outperforms MRE under MSEM criterion for this data set.

$\widehat{\beta}_M(r,k)$ and $\widehat{\beta}_M(k)$

For testing the superiority of $\widehat{\beta}_M(r,k)$ estimator over stochastic restricted ridge regression estimator, $\widehat{\beta}_M(k)$, the value of the statistic F_2 is calculated to be very low for all values of k considered here. When compared with the critical values of $F(1,37)$ at both 5% and 1% levels of significance, the null hypothesis H_{02} is not rejected. Hence, it implies that for the given of data set, $\widehat{\beta}_M(r,k)$ dominates the stochastic restricted ridge regression estimator for all chosen values of k under the MSEM criterion.

7 Concluding Remarks

In this paper, a new *r-k* class estimator, $\widehat{\beta}_M(r,k)$, is proposed when the prior information is in the form of stochastic linear constraints binding the regression coefficients is available. Its performances *vis-a-vis* MRE and SRRRE have been compared under the MSEM criterion. Tests for verifying the conditions have also been suggested. A simulation study as well as numerical illustration have been done to study the performance of the tests for verifying the conditions of dominance of $\widehat{\beta}_M(r,k)$ estimator over the other two under the MSEM criterion. The empirical findings suggest that $\widehat{\beta}_M(r,k)$ is a better estimator than each of $\widehat{\beta}_M$ and $\widehat{\beta}_M(k)$ by the criterion of MSEM.

References

[1] Baye, M.R., Parker D.F. (1984). *Combining ridge and principal components regression:a money demand illustration*, Communications in Statistics-Theory and Methods 13(2), 197-205.

[2] Bayhan, G.M., Bayhan, M. (1998). *Forecasting using autocorrelated errors and multicollinear predictor variables*, Computers and Industrial Engineering 34, 413-421.

[3] Firinguetti, L. (1989). *A simulation study of ridge regression estimators with autocorrelated errors*, Communications in Statistics- Simulation and Computation 18(2), 673-702.

[4] Gujarati, D.N.(2002) *Basic Econometrics*, Tata McGraw-Hill, Fourth edition 794.

[5] Guler, H. and Kaciranlar, S. (2009). *A Comparison of Mixed and Ridge Estimators of Linear Models*, Communications in Statistics- Simulation and Computation 38, 368-401.

[6] Hoerl, A.E., Kennard, R.W. (1970). *Ridge regression:biased estimation for nonorthogonal problems*, Technometrics 12,55-67.

[7] Judge, G.G., Griffiths, W.E., Hill, R.C., Ltkepohl, H., Lee,T.C. (1985) *The Theory and Practice of Econometrics*, John Wiley and Sons, New York.

[8] Kaçiranlar, S. and Sakallioğlu, S. (2001). *Combining the Liu estimator and the principal component regression estimator*, Communications in Statistics- Theory and Methods 30(12), 2699-2705.

[9] Masayuki, J. and Nobuo, I. (1996). *r-k class estimation in regression model with concomitant variables*,Annals of Institute of Statistical Mathematics 48(1), 89-95.

[10] Newhouse, J.P., Oman, S.D. (1971). *An evaluation of ridge estimators*, Rand corporation, 1-29.

[11] Özkale, M. R.,Kaçiranlar, S. (2008). *Comparisons of the r-k class estimator to the ordinary least squares estimator under the Pitman's closeness criterion*, Statistical Papers 49, 503-512.

[12] Özkale, M.R. (2009). *A stochastic restricted ridge regression estimator*, Journal of Multivariate Analysis 100, 1700-1706.

[13] Rao, C. and Toutenbourgh, H. (1999). *Linear Models:Least squares and Alternatives*, Second Edition, Springer-Verlag, New York.

[14] Sarkar, N. (1989). *Comparisons among some estimators in misspecified linear models with Multicollinearity*, Annals of Institute of Statistical Mathematics 41(4), 717-724.

[15] 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, 1987-2000.

[16] Sarkar, N. (1996). *Mean square error matrix comparison of some estimators in linear regressions with multicollinearity*, Statistics and Probability Letters 30,133-138.

[17] Sarkar, N., Chandra, S. (2012). *Comparison of the r-(k,d) class estimator with some estimators under the Mahalanobis loss function* manuscript.

[18] Theil, H. and Goldberger, A.S. (1961). *On pure and mixed statistical estimation in economics*, International Economic Review 2, 65-78.

[19] Trenkler, G. (1984). *On the performance of biased estimators in the linear regression model with correlated or heteroscedastic errors*, Journal of Econometrics 25, 179-190.

[20] Xu, J. and Yang, H. (2011) *On the restricted r-k class estimator and the restricted r-d class estimator in linear regression*, Journal of Statistical Computation and Simulation 81(6),679-691.