

## A New Ridge Estimator in Linear Measurement Error Model with Stochastic Linear Restrictions

F. Ghapani <sup>1</sup>, B. Babadi <sup>2</sup>

<sup>1</sup>Department of Mathematics and Statistics, Shoushtar Branch, Islamic Azad University, Shoushtar, Iran

<sup>2</sup>Department of Statistics, Shahid Chamran University of Ahvaz, Ahvaz, Iran

**Abstract.** In this paper, a new ridge-type estimator is proposed and termed as the new mixed ridge estimator (NMRE) which is obtained by unifying the sample and prior information in linear measurement error model with additional stochastic linear restrictions. The new estimator is a generalization of the mixed estimator (ME) and ridge estimator (RE). The performances of this new estimator and mixed ridge estimator (MRE) with respect to the ME are examined under the criterion of mean squared error matrix. Finally, a numerical example and a Monte Carlo simulation are also presented to analyze.

**Keywords.** Mean squared error matrix; Mixed estimator; Mixed ridge estimator.

**MSC:** Primary 62J05; Secondary 62J07.

### 1 Introduction

One of the basic assumptions in regression analysis is that all the observations are correctly observed. The presence of measurement errors in the observations violates the essential properties of estimators. An im-

---

Corresponding Author: F. Ghapani(f-ghapani@phdstu.scu.ac.ir)  
B. Babadi (b.babadi@scu.ac.ir)

portant issue in the area of measurement errors is to find the consistent estimators of the parameters. Several approaches have been developed for the measurement error problems (see for more details (Cheng , 1999; Fuller , 1987, see for example). In order to correct the effects of measurement error on parameters estimation, Nakamura (1990) considered an approach based on the correction of score function. This approach makes it possible to draw inference as well as estimation of parameters without additional assumptions.

Also, when the columns of the regression matrix are nearly dependent, the problem of collinearity enters into the data. The existence of collinearity in the linear regression model affects the method of least squares to generally produce poor estimates of parameters. In order to resolve this problem, several approaches have been suggested and, among them, the ridge estimator (RE) was proposed by Hoerl (1970) combat the problem of collinearity for the estimation of regression parameters, also see Belsley *et al.* (1980), Mason and Gunst (1985) and Belsley (1991) for more details. Saleh *et al.* (2014) considered ridge regression and five quasi-empirical Bayes estimators of the regression parameters of a measurement error model.

An alternative technique to overcome the collinearity problem is to consider parameter estimation in addition to the sample information, such as some exact or stochastic restrictions on unknown parameter vector (Rao *et al.* , 2008). When additional stochastic linear restrictions on unknown parameter vector are assumed to be held, Durbin (1953), Theil and Goldberger (1961), and Theill (1963) proposed the ME. Sarkar (1992) proposed a new restricted estimator by combining the restricted least squares estimate with RE. Kaciranlar *et al.* (1998) compared the estimator introduced by Sarkar (1992) and the modified ridge regression estimator based on prior information proposed by Swindel (1976). Grob (2003) studied another restricted ridge estimator by combing the estimator given by Swindel (1976) and the restricted least squares method. A new restricted ridge estimation method is proposed by minimizing the sum of squared residuals by Zhong and Yang (2007). Li and Yang (210) derived the MRE in linear regression models. He and Wu (2014) proposed a new estimator to combat the collinearity in the linear model when there were stochastic linear restrictions on the regression coefficients. Arumairajan *et al.* (2014) improvement of ridge estimator in the linear regression model with stochastic restrictions. Alkhamisi *et al.* (2015) derived the necessary and sufficient conditions for superiority of the restricted ridge estimator over the restricted

least squares estimator by trace of the mean square error. Shalabh *et al.* (2007) obtained consistent estimator of regression coefficients in a measurement error model when the prior information is available in the form of exact linear restrictions. In the context of multiple linear regression models, the additional information in the form of known covariance matrix of measurement errors associated with explanatory variables and known matrix of reliability ratios of explanatory variables have been extensively utilized in Shalabh *et al.* (2009). Also, Shalabh *et al.* (2010) obtained consistent estimation of regression coefficients in ultrastructural measurement error model using stochastic prior information. When the prior information and the sample information are not equally important, Ghapani *et al.* (2016) introduce weighted mixed ridge estimator for regression coefficients in measurement error models.

In this paper, by combining the ME and the RE in a special way, we introduce a new mixed ridge estimator (NMRE) for unknown parameter vector in linear measurement error models when additional stochastic linear restrictions is supposed to be held. Furthermore, the performances of this new estimator and MRE with respect to the ME are examined based on the MSEM criterion. The paper is structured as follows. In Section 2, the model specification and the estimators are given. The properties of the new estimator are discussed in Section 3. The performances of the NMRE and MRE over ME are examined based on the MSEM criterion in Section 4. In Section 5, the selection of ridge parameter is discussed. A numerical example and a simulation study are provided to illustrate some of the theoretical results in Section 6. Finally, some conclusions are given in Section 7.

## 2 Model Specification and the Estimators

We consider the linear measurement error model

$$\begin{aligned} y &= Z\beta + \varepsilon, \varepsilon \sim N(0, \sigma^2 I_n), \\ X &= Z + \Phi, \Phi \sim N(0, I_n \otimes \Sigma), \end{aligned} \quad (2.1)$$

where  $y = (y_1, y_2, \dots, y_n)'$  is an  $n \times 1$  vector of response variables,  $\beta$  is a  $p \times 1$  vector of unknown parameters,  $\varepsilon$  is an  $n \times 1$  vector of unobservable random errors and  $Z$  is an  $n \times p$  matrix of unobservable values of explanatory variables which can be observed through the matrix  $X$  with the measurement error  $\Phi$ . Furthermore, we assume that  $\varepsilon$  and  $\Phi$  are

independent and  $\Sigma$  is a  $p \times p$  matrix of known values with nonnegative diagonal elements (Fuller, 1987). The log-likelihood function for model (2.1) is

$$l(\beta, \sigma^2, Z, y) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} [(y - Z\beta)'(y - Z\beta)].$$

If we replace  $Z$  by  $X$  without considering the measurement errors, then the estimates obtained from the score functions are not consistent in general. To correct the effects of measurement errors on parameters estimation, we use the corrected score method proposed by Nakamura (1990). The corrected log-likelihood function for model (1) is suggested by Nakamura (1990) as

$$l^*(\beta, \sigma^2, X, y) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} [(y - X\beta)'(y - X\beta) - n\beta'\Sigma\beta].$$

The following equation is useful to find the corrected log-likelihood function

$$E^*(X'X) = Z'Z + n\Sigma,$$

where  $E^*$  denotes the conditional mean with respect to  $X$  given  $y$ . The corrected log-likelihood function  $l^*(\beta, \sigma^2, X, y)$  satisfies  $E^* \left[ \frac{\partial}{\partial \beta} l^*(\beta, \sigma^2, X, y) \right] = \frac{\partial}{\partial \beta} l(\beta, \sigma^2, Z, y)$  and  $E^* \left[ \frac{\partial}{\partial \sigma^2} l^*(\beta, \sigma^2, X, y) \right] = \frac{\partial}{\partial \sigma^2} l(\beta, \sigma^2, Z, y)$ . By solving the equations  $\frac{\partial}{\partial \beta} l^*(\beta, \sigma^2, X, y) = 0$  and  $\frac{\partial}{\partial \sigma^2} l^*(\beta, \sigma^2, X, y) = 0$ , the corrected score estimates of  $\beta$  and  $\sigma^2$  respectively, are given by  $\hat{\beta} = (X'X - n\Sigma)^{-1}X'y$  and  $\hat{\sigma}^2 = \frac{1}{n} [(y - X\hat{\beta})'(y - X\hat{\beta}) - n\hat{\beta}'\Sigma\hat{\beta}]$ .

Now, we assume that the vector of parameters is subject to the stochastic linear restrictions

$$r = R\beta + e, e \sim N(0, \sigma^2 W), \quad (2.2)$$

where  $r$  is a  $q \times 1$  observable random vector,  $R$  is a  $q \times p$  known matrix with  $\text{rank}(R) = q$  for  $q \leq p$  and  $e$  is a  $q \times 1$  error vector with  $E(e) = 0$  and  $\text{Var}(e) = \sigma^2 W$ , in which  $W$  is supposed to be known and positive definite matrix. Further it is also assumed that random vector  $e$  is stochastically independent of  $\epsilon$  and  $\Phi$ . The log-likelihood function and the corrected log-likelihood function for model (2.1) with stochastic linear restrictions (2.2) are given by

$$l(\beta, \sigma^2, Z, y, r) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2} \log |W| - \frac{1}{2\sigma^2} [(y - Z\beta)'(y - Z\beta)] - \frac{1}{2\sigma^2} [(r - R\beta)'W^{-1}(r - R\beta)],$$

$$l^*(\beta, \sigma^2, X, y, r) = -\frac{N}{2} \log(2\pi\sigma^2) - \frac{1}{2} \log |W| - \frac{1}{2\sigma^2} [(y - X\beta)'(y - X\beta) - n\beta'\Sigma\beta] - \frac{1}{2\sigma^2} [(r - R\beta)'W^{-1}(r - R\beta)],$$

respectively, where  $N = n + q$ . The  $l^*(\beta, \sigma^2, X, y, r)$  have the following properties:

$$E^* \left[ \frac{\partial}{\partial \beta} l^*(\beta, \sigma^2, X, y, r) \right] = \frac{\partial}{\partial \beta} l(\beta, \sigma^2, Z, y, r),$$

$$E^* \left[ \frac{\partial}{\partial \sigma^2} l^*(\beta, \sigma^2, X, y, r) \right] = \frac{\partial}{\partial \sigma^2} l(\beta, \sigma^2, Z, y, r).$$

The corrected score estimates of  $\beta$  and  $\sigma^2$  for model (2.1) with stochastic linear restrictions (2.2) will be obtained with differentiating from  $l^*(\beta, \sigma^2, X, y, r)$  with respect to the  $\beta$  and  $\sigma^2$ . We call these estimators, denoted by  $\hat{\beta}_{ME}$  and  $\hat{\sigma}_{ME}^2$ , respectively, as the mixed estimators. Then we have

$$\hat{\beta}_{ME} = (X'X + R'W^{-1}R - n\Sigma)^{-1}(X'y + R'W^{-1}r),$$

$$\hat{\sigma}_{ME}^2 = \frac{1}{N}(y'y - \hat{\beta}'_{ME}X'y - \hat{\beta}'_{ME}R'W^{-1}r + r'W^{-1}r). \quad (2.3)$$

Using Lemma 1 (presented in Appendix), we can write

$$(X'X + R'W^{-1}R - n\Sigma)^{-1} = S^{-1} - S^{-1}R'(W + RS^{-1}R')^{-1}RS^{-1},$$

where,  $S = X'X - n\Sigma$  and

$$\left[ S^{-1} - S^{-1}R'(W + RS^{-1}R')^{-1}RS^{-1} \right] R'W^{-1}r = S^{-1}R'(W + RS^{-1}R')^{-1}r.$$

Then, the  $\hat{\beta}_{ME}$  can be rewritten as

$$\hat{\beta}_{ME} = \hat{\beta} + S^{-1}R'(W + RS^{-1}R')^{-1}(r - R\hat{\beta}). \quad (2.4)$$

## 2.1 Ridge Estimator under the Stochastic Linear Restrictions

To deal with collinearity, we propose the ridge estimator of  $\beta$  under the stochastic linear restrictions. We consider the augmented model

$$u = U\beta + \eta, \quad (2.5)$$

in which  $u_{m \times 1} = [y, r, 0]'$ ,  $U_{m \times p} = [X, R, \sqrt{k}I_p]'$ ,  $[\varepsilon - \Phi\beta, e, \varphi]'$  for  $m = n + q + p$  and  $\eta$  is a random vector with  $E(\eta) = 0$  and  $Var(\eta) = Blockdiag[(\sigma^2 + \beta'\Sigma\beta)I_n, \sigma^2W, \sigma^2I_p]$ . The parameter  $0 < k < 1$  denotes the

ridge parameter,  $\varphi$  is an error vector with  $E(\varphi) = 0$  and  $Var(\varphi) = \sigma^2 I_p$ . The corrected log-likelihood function for model (2.5) is given by

$$l_k^*(\beta, \sigma^2, X, y, r) = l^*(\beta, \sigma^2, X, y, r) - \frac{k\beta'\beta}{2\sigma^2}.$$

The corrected score estimates of  $\beta$  and  $\sigma^2$  will be obtained with differentiating from the  $l_k^*(\beta, \sigma^2, X, y, r)$  with respect to the  $\beta$  and  $\sigma^2$ . We call these estimators as mixed ridge estimators and denote by  $\hat{\beta}_{MRE}(k)$  and  $\hat{\sigma}_{MRE}^2(k)$ , respectively. Then we obtain

$$\begin{aligned} \hat{\beta}_{MRE}(k) &= (X'X + R'W^{-1}R - n\Sigma + kI_p)^{-1}(X'y + R'W^{-1}r), 0 < k < 1, \\ \hat{\sigma}_{MRE}^2(k) &= \frac{1}{m} \left[ y'y - \hat{\beta}'_{MRE}(k)X'y - \hat{\beta}'_{MRE}(k)R'W^{-1}r + r'W^{-1}r \right]. \end{aligned} \quad (2.6)$$

Using Lemma 1 (presented in Appendix), we can write

$$\hat{\beta}_{MRE}(k) = \hat{\beta}_{RE}(k) + S_k^{-1}R'(W + RS_k^{-1}R')^{-1} \left[ r - R\hat{\beta}_{RE}(k) \right],$$

where  $\hat{\beta}_{RE}(k) = (X'X - n\Sigma + kI_p)^{-1}X'y = S_k^{-1}X'y$  and  $S_k = X'X - n\Sigma + kI_p$ . Additionally, based on Li and Yang (2011) a new mixed ridge estimator is obtained by combining the ME and RE. Substituting  $\hat{\beta}$  with  $\hat{\beta}_{RE}(k)$  in the (2.4), we can get the NMRE as

$$\hat{\beta}_{NMRE}(k) = \hat{\beta}_{RE}(k) + S^{-1}R'(W + RS^{-1}R')^{-1} \left[ r - R\hat{\beta}_{RE}(k) \right].$$

Let  $T_k = (I + kS^{-1})^{-1} = S_k^{-1}S = SS_k^{-1}$ . So we can write the NMRE as

$$\begin{aligned} \hat{\beta}_{NMRE}(k) &= S^{-1}T_kX'y + S^{-1}R'(W + RS^{-1}R')^{-1}(r - RS^{-1}T_kX'y) \\ &= \left[ S^{-1} - S^{-1}R'(W + RS^{-1}R')^{-1}RS^{-1} \right] \\ &\quad \times T_kX'y + S^{-1}R'(W + RS^{-1}R')^{-1}r \quad (2.7) \\ &= \left[ S^{-1} - S^{-1}R'(W + RS^{-1}R')^{-1}RS^{-1} \right] (T_kX'y + R'W^{-1}r) \\ &= (S + R'W^{-1}R)^{-1}(T_kX'y + R'W^{-1}r). \end{aligned}$$

### 3 Asymptotic Properties of Estimators

The exact distribution and finite sample properties of the ME, MRE and NMRE estimators are difficult to derive. Therefore, we propose to employ the large sample asymptotic approximation theory to study the

asymptotic distribution of the estimators. We assume the parameter  $\beta$  is identifiable and, as  $n$  tends to infinity, the limits of  $n^{-1}Z'Z$ ,  $n^{-1}(Z'Z + kI_p)$ ,  $n^{-1}(Z'Z + R'W^{-1}R)$ ,  $n^{-1}(Z'Z + R'W^{-1}R + kI_p)$  and  $n^{-1}(G_kZ'Z + R'W^{-1}R)$ , where  $G_k = [I + k(Z'Z)^{-1}]^{-1}$ , exist and  $E$  denotes the global expectation taken at the true value  $\beta$ .

**Theorem 3.1.** *The asymptotic distribution of  $\sqrt{n} [\hat{\beta}_{NMRE}(k) - M^{-1}M_k\beta]$  is normal with mean vector zero and*

$$AVar [\hat{\beta}_{NMRE}(k)] = M^{-1} [G_kBG_k + \sigma^2(G_kZ'ZG_k + R'W^{-1}R)] M^{-1},$$

where  $M_k = G_kZ'Z + R'W^{-1}R$ ,  $M = M_{k=0}$  and  $B = (n\sigma^2 + \beta'Z'Z\beta)\Sigma$ .

*Proof.* Since  $E(X'X) = Z'Z + n\Sigma$  (Fung *et al.* , 2003) we have

$$X'X = Z'Z + n\Sigma + O_p(n^{\frac{1}{2}}).$$

Then

$$\begin{aligned} n^{-1}X'X &= n^{-1}Z'Z + \Sigma + O_p(n^{-\frac{1}{2}}), \\ n^{-1}(X'X + kI_p) &= n^{-1}(Z'Z + kI_p) + \Sigma + O_p(n^{-\frac{1}{2}}). \end{aligned}$$

Moreover, we can obtain  $T_k = G_k + O_p(n^{-\frac{1}{2}})$ . Furthermore, we have

$$n^{-1}(X'X + RW^{-1}R) = n^{-1}(Z'Z + R'W^{-1}R) + \Sigma + O_p(n^{-\frac{1}{2}}). \quad (3.1)$$

So, it follows from (2.7) and (3.1) that

$$\begin{aligned} \sqrt{n}\hat{\beta}_{NMRE}(k) &= [I_p + O_p(n^{-\frac{1}{2}})]^{-1} C^{-1} n^{-\frac{1}{2}} \{ [G_k + O_p(n^{-\frac{1}{2}})] X'y + R'W^{-1}r \} \\ &= [I_p + O_p(n^{-\frac{1}{2}})] C^{-1} n^{-\frac{1}{2}} [G_kX'y + R'W^{-1}r + O_p(n^{\frac{1}{2}})]. \end{aligned} \quad (3.2)$$

Moreover, since the limit of  $C = n^{-1}(Z'Z + R'W^{-1}R)$  exists, then (3.2) can be written as

$$\sqrt{n}\hat{\beta}_{NMRE}(k) = C^{-1}\xi + O_p(n^{-\frac{1}{2}}), \quad (3.3)$$

where  $\xi = n^{-\frac{1}{2}}(G_kX'y + R'W^{-1}r)$  is asymptotically normal (Fung *et al.* , 2003). So it follows from  $E(G_kX'y + R'W^{-1}r) = (G_kZ'Z + R'W^{-1}R)\beta$  that we have  $E(\xi) = n^{-\frac{1}{2}}M_k\beta$ . Consequently, we asymptotically have

$$\sqrt{n} [\hat{\beta}_{NMRE}(k) - M^{-1}M_k\beta] = C^{-1} [\xi - E(\xi)] + O_p(n^{-\frac{1}{2}}),$$

where  $\sqrt{n}[\hat{\beta}_{NMRE}(k) - M^{-1}M_k\beta]$  is asymptotically normal with mean vector zero. So, we can asymptotically write

$$\begin{aligned} E[\hat{\beta}_{NMRE}(k)] &= M^{-1}G_kZ'Z\beta + M^{-1}R'W^{-1}R\beta \\ &= \beta + M^{-1}(G_k - I)Z'Z\beta \end{aligned}$$

Furthermore, from (3.3) we have  $AVar[\sqrt{n}\hat{\beta}_{MRE}(k)] = C^{-1}Var(\xi)C^{-1}$ . The variance of  $\xi$  becomes

$$\begin{aligned} Var(\xi) &= E_{y^*} [Var(\xi | y^*)] + Var_{y^*} [E(\xi | y^*)] \\ &= n^{-1}E_{y^*}(G_k y' y \Sigma G_k) + n^{-1}Var_{y^*}(G_k Z' y + R' W^{-1} r), \end{aligned}$$

where  $E_{y^*}$  and  $Var_{y^*}$  denote the expectation and variance with respect to the random vector  $y^{*'} = (y', r')$ . Since  $E(y'y) = n\sigma^2 + \beta'Z'Z\beta$  and  $Var_{y^*}(G_k Z' y + R' W^{-1} r) = \sigma^2(G_k Z' Z G_k + R' W^{-1} R)$ , therefore  $Var(\xi) = n^{-1}[G_k B G_k + \sigma^2(G_k Z' Z G_k + R' W^{-1} R)]$  and  $AVar[\hat{\beta}_{NMRE}(k)] = M^{-1}[G_k B G_k + \sigma^2(G_k Z' Z G_k + R' W^{-1} R)]M^{-1}$ .  $\square$

*Corollary 3.1.*  $\sqrt{n}[\hat{\beta}_{MRE}(k) - A_k^{-1}M\beta]$  has asymptotic normal distribution with mean vector zero and  $AVar[\hat{\beta}_{MRE}(k)] = A_k^{-1}(B + \sigma^2 M)A_k^{-1}$ , where  $A_k = Z'Z + R'W^{-1}R + kI_p$ .

*Corollary 3.2.*  $\sqrt{n}(\hat{\beta}_{ME} - \beta)$  has asymptotic normal distribution with mean vector zero and  $AVar(\hat{\beta}_{ME}) = M^{-1}(B + \sigma^2 M)M^{-1}$ .

## 4 Mean Squared Error Matrix

In order to compare the performance of any estimator with others a criterion for measuring the goodness of an estimator is required. For this purpose, the mean square error (MSEM) is used to measure the goodness of an estimator. We note that, for any estimator  $\hat{\beta}$  of  $\beta$ , its MSEM is defined as

$$MSEM(\hat{\beta}) = E(\hat{\beta} - \beta)'(\hat{\beta} - \beta) = Var(\hat{\beta}) + Bias(\hat{\beta})Bias(\hat{\beta})',$$

where  $Bias(\hat{\beta})$  is the bias vector. For any two given estimators  $\hat{\beta}_1$  and  $\hat{\beta}_2$ , we say that  $\hat{\beta}_2$  is superior to  $\hat{\beta}_1$  with respect to MSEM criterion, if and only if

$$D(\hat{\beta}_1, \hat{\beta}_2) = MSEM(\hat{\beta}_1) - MSEM(\hat{\beta}_2) \geq 0.$$

Now, we can obtain the asymptotic MSEM of the estimators  $\hat{\beta}_{ME}$ ,  $\hat{\beta}_{MRE}(k)$  and  $\hat{\beta}_{NMRE}(k)$  as

$$AMSEM(\hat{\beta}_{ME}) = M^{-1}(B + \sigma^2 M)M^{-1},$$

$$AMSEM[\hat{\beta}_{MRE}(k)] = A_k^{-1}(B + \sigma^2 M)A_k^{-1} + b_1 b_1',$$

$$AMSEM[\hat{\beta}_{NMRE}(k)] = M^{-1} [G_k B G_k + \sigma^2 (G_k Z' Z G_k + R' W^{-1} R)] M^{-1} + b_2 b_2',$$

where  $b_1 = -k A_k^{-1} \beta$  and  $b_2 = M^{-1} (G_k - I) Z' Z \beta$ .

In order to compare  $\hat{\beta}_{MRE}(k)$  and  $\hat{\beta}_{NMRE}(k)$  in the MSEM criterion, we consider the asymptotic MSEM differences

$$\begin{aligned} \Delta_1 &= AMSEM(\hat{\beta}_{ME}) - AMSEM(\hat{\beta}_{MRE}) = D_1 - b_1 b_1' \\ &= A_k^{-1} [2\sigma^2 k I_p + k^2 \sigma^2 M^{-1} + k B M^{-1} + k M^{-1} B + k^2 M^{-1} B M^{-1} - k^2 \beta \beta'] A_k^{-1}, \end{aligned}$$

$$\begin{aligned} \Delta_2 &= AMSEM(\hat{\beta}_{ME}) - AMSEM(\hat{\beta}_{NMRE}) = D_2 - b_2 b_2', \\ &= M^{-1} [B - G_k B G_k + \sigma^2 (Z' Z - G_k Z' Z G_k) - (G_k - I) Z' Z \beta \beta' Z' Z (G_k - I)] \\ &\quad \times M^{-1}, \end{aligned}$$

where

$$D_1 = A_k^{-1} [2\sigma^2 k I_p + k^2 \sigma^2 M^{-1} + k B M^{-1} + k M^{-1} B + k^2 M^{-1} B M^{-1}] A_k^{-1},$$

$$D_2 = M^{-1} [B - G_k B G_k + \sigma^2 (Z' Z - G_k Z' Z G_k)] M^{-1}.$$

**Theorem 4.1.** *The  $\hat{\beta}_{NMRE}(k)$  is superior to the estimator  $\hat{\beta}_{ME}$  in the MSEM criterion, if and only if  $b_2' D_2^{-1} b_2 \leq 1$ .*

*Proof.* To apply Lemma 2 (presented in Appendix), we need to prove that  $D_2$  is a positive definite (p.d.) matrix. For  $A = Z' Z > 0$ , there exists some orthogonal matrix  $Q$ , such that  $A = Q \Lambda Q'$ , where  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$ ,  $\lambda_i > 0$ . Therefore, we can compute that  $Z' Z - G_k Z' Z G_k = Q \Gamma Q' = Q \text{diag}(\gamma_1, \dots, \gamma_p) Q'$ , where  $\Gamma = \Lambda - (I + k \Lambda^{-1})^{-1} \Lambda (I + k \Lambda^{-1})^{-1}$ , for  $k > 0$ ,  $\lambda_i > 0$ , so  $\gamma_i > 0$  which means  $Z' Z - G_k Z' Z G_k > 0$ . Furthermore,  $B - G_k B G_k > 0$  and  $M^{-1} = (Z' Z + R' W^{-1} R)^{-1} > 0$ . This implies that  $D_2 > 0$ . Hence according to Lemma 2 (presented in Appendix), we can get  $\Delta_2 > 0$  if and only if  $b_2' D_2^{-1} b_2 \leq 1$ .  $\square$

**Theorem 4.2.** *The  $\hat{\beta}_{MRE}(k)$  is superior to the estimator  $\hat{\beta}_{ME}$  in the MSEM criterion, if and only if  $b_1' D_1^{-1} b_1 \leq 1$ .*

*Proof.* We note that  $D_1 > 0$ . Hence according to Lemma 2 (presented in Appendix), we can get  $\Delta_1 \geq 0$  if and only if  $b_1' D_1^{-1} b_1 \leq 1$ .  $\square$

## 5 Selection of Ridge Parameter $k$

An important issue in the study of ridge regression is how to find an appropriate parameter  $k$ . In this section, we discuss the selection of parameter  $k$  based on the MSEM criterion. In this method we derive ridge parameter  $k$  such that  $\Delta_1$  and  $\Delta_2$  are p.d. matrices.

First, we consider the matrix difference  $\Delta_1$  and note that  $k^2\sigma^2M^{-1} + kBM^{-1} + kM^{-1}B + k^2M^{-1}BM^{-1}$  is a p.d. matrix. Therefore, using Lemma 2 (presented in Appendix), we have that  $\Delta_1$  is p.d. if  $2\sigma^2kI_p - k^2\beta\beta'$  is positive semi-definite (p.s.d.). Thus a sufficient condition for  $\hat{\beta}_{MRE}$  to be superior over  $\hat{\beta}_{ME}$  is  $k \leq \frac{2\sigma^2}{\beta'\beta}$ . Since  $k$  depends on the unknown parameters  $\beta$  and  $\sigma^2$ , we replace them by appropriate estimators and the estimated value is  $\hat{k}_{MRE} = \frac{2\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}}$ .

Next, consider matrix difference  $\Delta_2$  and note that  $B - G_kBG_k$  is a p.d. matrix. Thus by using Lemma 2 (presented in Appendix), we have that  $\Delta_2$  is a p.d. matrix if  $\sigma^2(Z'Z - G_kZ'ZG_k) - (G_k - I)Z'Z\beta\beta'Z'Z(G_k - I)$  is a p.s.d. matrix. Denoting  $\alpha = Q'\beta$ , we have  $\beta = Q\alpha$ . In this case, we can write  $\sigma^2(Z'Z - G_kZ'ZG_k) - (G_k - I)Z'Z\beta\beta'Z'Z(G_k - I) = QTQ' = Qdiag(\tau_1, \dots, \tau_p)Q'$ , where  $\tau_i = (\lambda_i + k)^{-2} [k\sigma^2\lambda_i(2\lambda_i + k) - k^2\lambda_i^2\alpha_i^2]$ . For  $k > 0$  and  $\lambda_i > 0$ , we have  $\tau_i > 0$  if and only if  $k < \frac{2\sigma^2}{\alpha_i^2 - \lambda_i^{-1}\sigma^2}$ . We suggest the replacement of  $\sigma^2$ ,  $\alpha_i$  and  $\lambda_i$  by their corresponding appropriate estimators. Therefore, we have  $\hat{k}_{NMRE} = \frac{2\hat{\sigma}^2}{\hat{\alpha}_{\max}^2 - \lambda_{\max}^{-1}\hat{\sigma}^2}$ , where  $\hat{\alpha}_{\max}^2$  and  $\lambda_{\max}$  are denote the maximum element of  $Q\hat{\beta}$  and  $\Lambda$ .

We are now interested in knowing that under which conditions  $\hat{\beta}_{NMRE}$  and  $\hat{\beta}_{MRE}$  are superior over  $\hat{\beta}_{ME}$ . Note that  $\hat{k}_{MRE} < \hat{k}_{NMRE}$  and, therefore, in this paper, we suggest to choose  $\hat{k} = \hat{k}_{MRE}$ .

## 6 Numerical Example and Monte Carlo Simulation

To illustrate our theoretical results, we consider the data known as the Egyptian pottery data. This data set arises from an extensive archaeological survey of pottery reduction and distribution in the ancient Egyptian city of Al-Amarna. This data set consists of measurements of chemical contents (mineral elements) made on many samples of pottery using two different techniques, NAA and ICP. For descriptions of techniques, see Smith *et al.* (1998). The set of pottery was collected from different locations around the city. The group structure among the objects arises from two main sources, fabric code and location of

pottery. Both of these subdivisions are important to the archaeologists. Consequently, according to this group structure, the selected vessels have been divided into 25 groups and three groups of imported vessels are selected as stochastic linear restrictions Ghapani *et al.* (2015, see). In each group, there are different numbers of vessels from the same fabric code and provenance, which can essentially be regarded as replicated observations. Among all mineral elements, our interest is in the relation between Na measured with NAA as response variable versus mineral elements Na, Al, K, V, Cr and Nm measured with ICP as predictor variables. The results are computed by R software. We fit a linear measurement error model with stochastic linear restrictions. We firstly obtain the estimates  $\hat{\beta}$ ,  $\hat{\sigma}^2$ ,  $\hat{k}$  and  $\hat{Z}$ . An estimate of  $Z$  can be derived as  $\hat{Z} = X + \hat{\sigma}_v^{-2} \hat{\sigma} \hat{\beta}' \Sigma$ , where,  $\hat{v} = y - X\hat{\beta}$  and  $\hat{\sigma}_v^2 = \hat{\sigma}^2 + \hat{\beta}' \Sigma \hat{\beta}$  (Rasekh, 2006). The estimated MSE values of ME, MRE and NMRE and values of  $b_1' D_1^{-1} b_1$  and  $b_2' D_2^{-1} b_2$  are presented in Table 1.

Table 1: Estimated MSE values of the ME, MRE and NMRE

ME	MRE	NMRE	$b_1' D_1^{-1} b_1$	$b_2' D_2^{-1} b_2$
705.917	684.359	525.851	0.348	7.72e-5

From Table 1, we can find that estimated MSE value of the NMRE and MRE are indeed smaller than of the ME, and the NMRE has smaller estimated MSE values than the MRE. Moreover,  $b_1' D_1^{-1} b_1 \leq 1$  and  $b_2' D_2^{-1} b_2 \leq 1$ , which agree with our theoretical findings in Theorem 4.1 and Theorem 4.2. This can imply that the proposed  $\beta_{MRE}$  and  $\beta_{NMRE}$  estimators can perform better than the  $\hat{\beta}_{ME}$ . For more convenience, plots of the estimated MSE values of the ME, MRE and NMRE versus  $k$  in the interval  $[0, 0.1]$  with increment 0.0001 are given in Figure 1. The MSE values of MRE and NMRE decreases when  $k$  increases. It is obvious that NMRE and MRE are superior to the ME.

In order to further illustrate the behavior of the new estimator, we performed a simulation study. The explanatory variables are computed following Kibria (2003) as

$$z_{ij} = (1 - \rho^2)^{\frac{1}{2}} u_{ij} + \rho u_{i,p+1}, i = 1, \dots, n, j = 1, \dots, p,$$

where  $u_{ij}$  are independent standard normal random numbers and  $\rho^2$  represents the correlation between explanatory variables. Three different sets of correlation are considered according to the values of

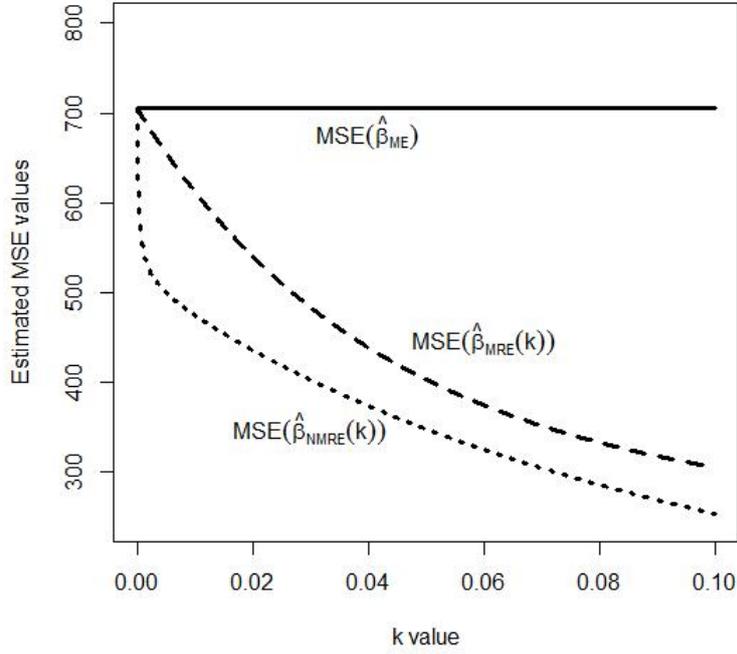


Figure 1: Estimated MSE values of ME, MRE and NMRE

$\rho = 0.8, 0.9$  and  $0.95$ . Moreover, the explanatory variables are standardized to obtain  $Z'Z$  in correlation form. In this study,  $p = 3$  is the number of the explanatory variables,  $n = 25$  and  $n = 70$  and  $q = 2$ . We generated the  $l$ -th set of simulated data as

$$\begin{aligned} y_l &= Z\beta + \varepsilon_l, \\ X_l &= Z + \Delta_l, l = 1, \dots, 2000, \\ r_l &= R\beta + e_l, \end{aligned}$$

where  $y_l = (y_{1l}, \dots, y_{nl})'$ ,  $Z = (z^{(1)}, z^{(2)}, z^{(3)})$ ,  $z^{(j)} = (z_{1j}, \dots, z_{nj})'$ ,  $j = 1, 2, 3$ , and  $\varepsilon_l \sim N(0, \sigma^2 I_n)$  is rewritten in accordance with  $y_l$ . Furthermore,  $r_l = (r_{1l}, \dots, r_{ql})'$ ,  $R = (R^{(1)}, R^{(2)}, R^{(3)})$ ,  $R^{(j)} = (R_{1j}, \dots, R_{qj})'$ ,  $j = 1, 2, 3$ , and  $e_l \sim N(0, \sigma^2 I_q)$  is rewritten in accordance with  $r_l$ . Moreover, we assume that  $R_{sj} \sim N(0, 1)$ ,  $s = 1, \dots, q$ ,  $j = 1, \dots, p$ ,  $\sigma^2 = 0.16$ ,  $\sigma^2 = 0.36$ ,  $\Sigma = (0.1, 0.1, 0.1)$  and  $\Sigma = (0.15, 0.15, 0.15)$ . For each set of explanatory variables, we considered the coefficient vector that corresponded to the

largest eigenvalues of  $Z'Z$ . We first calculated estimators  $\beta, \sigma^2, k$  and  $Z$ . Then the experiment was repeated 2000 times by generating new error terms. Let us consider the ME, MRE and NMRE and compute their respective estimated MSE values which is defined by

$$\widehat{MSE}(\tilde{\beta}) = \frac{1}{2000} \sum_{l=1}^{2000} \sum_{j=1}^3 (\tilde{\beta}_{jl} - \beta_j)^2,$$

where  $\tilde{\beta}$  is any estimator that used in this study for making a comparison. The simulated MSE for all estimators are presented in Table 2.

Table 2: Estimated MSE values of the ME, MRE and NMRE

$\sigma^2$	$\Sigma = (0.1, 0.1, 0.1)$		$\Sigma = (0.15, 0.15, 0.15)$		
	$n = 25$	$n = 70$	$n = 25$	$n = 70$	
$\rho = 0.8$					
	ME	0.1612	0.0229	2.0130	0.0268
	MRE	0.1481	0.0225	1.6531	0.0264
0.16	NMRE	0.1332	0.0225	0.8475	0.0264
	$b'_1 D_1^{-1} b_1$	0.6488	0.7011	0.5985	0.6151
	$b'_2 D_2^{-1} b_2$	0.6451	0.6988	0.5950	0.6131
	ME	0.2483	0.0411	2.2439	0.0452
	MRE	0.2287	0.0400	1.8822	0.0441
0.36	NMRE	0.2216	0.0400	1.1204	0.0440
	$b'_1 D_1^{-1} b_1$	0.6628	0.7008	0.5957	0.6181
	$b'_2 D_2^{-1} b_2$	0.6565	0.6988	0.5906	0.6160
$\rho = 0.9$					
	ME	1.1328	0.0432	4.9075	0.0550
	MRE	0.9865	0.0421	4.0242	0.0538
0.16	NMRE	0.6427	0.0421	2.4779	0.0538
	$b'_1 D_1^{-1} b_1$	0.6164	0.6457	0.5493	0.5714
	$b'_2 D_2^{-1} b_2$	0.6116	0.6424	0.5441	0.5686
	ME	2.3966	0.0774	8.2243	0.0930
	MRE	2.1943	0.0743	7.1429	0.0900
0.36	NMRE	1.3474	0.0742	3.7939	0.0899
	$b'_1 D_1^{-1} b_1$	0.6287	0.6402	0.5794	0.5808
	$b'_2 D_2^{-1} b_2$	0.6229	0.6365	0.5742	0.5771
$\rho = 0.95$					
	ME	3.1774	0.1118	6.3263	1.2711
	MRE	2.8041	0.1086	5.4690	1.2040
0.16	NMRE	1.9041	0.1081	3.1526	1.0796
	$b'_1 D_1^{-1} b_1$	0.5820	0.5906	0.5047	0.5395
	$b'_2 D_2^{-1} b_2$	0.5763	0.5849	0.4929	0.5354
	ME	8.0245	0.2050	12.7951	3.4758
	MRE	7.4260	0.1964	11.4682	3.3761
0.36	NMRE	4.0465	0.1957	6.1155	2.9116
	$b'_1 D_1^{-1} b_1$	0.5689	0.5940	0.5264	0.5246
	$b'_2 D_2^{-1} b_2$	0.5580	0.5881	0.5175	0.5207

From the simulation results shown in Tables 2, we can see that, by the increase of the levels of collinearity,  $\Sigma$  and  $\sigma^2$  the estimated values

of MSE of the different estimators ME, MRE and NMRE will increase in general. Moreover, we can see that, for all cases, the MRE and NMRE have smaller estimated MSE values than the ME. In addition, for all cases,  $b_1' D_1^{-1} b_1 \leq 1$  and  $b_2' D_2^{-1} b_2 \leq 1$ . Therefore, we can conclude that the ridge estimators MRE and NMRE can perform better than the ME estimator. Furthermore, we can see that the new estimator is superior to the ME and MRE in the MSEM sense under certain conditions.

## 7 Conclusion

In this paper, we propose the NMRE for the vector of parameters in a linear measurement error model by unifying the sample and the prior information. We investigated the performance of the NMRE and MRE over the ME of the parameters using the asymptotic MSEM criterion. Furthermore, we show that the new estimator is superior to the ME and MRE in the MSEM sense under certain conditions. Finally, we illustrate our findings with a numerical example and a simulation study.

## Acknowledgements

The authors are most grateful to the Associate Editor and anonymous referees for valuable comments and suggestions which helped us to improve the quality of the paper. This research was supported by the Islamic Azad University, Shoushtar Branch. The authors would like to thank for this support.

## References

- Alkhamisi, M. A. and MacNeill, I. B. (2015). Recent results in ridge regression methods. *METRON* 73(359). DOI: 10.1007/s40300-015-0065-4.
- Arumairajan, S. and Wijekoon, P. (2014). Improvement of ridge estimator when stochastic restrictions are available in the linear regression model. *Journal of Statistical and Econometric Methods*. 3, 35-48.
- Belsley, D. A., Kuh, E. and Welsch, R. E. (1980). *Regression diagnostics: identifying data and sources of collinearity*. Wiley, New York.
- Belsley, D. A. (1991). *Conditioning Diagnostics: Collinearity and Weak Data in Regression*. 1 ed., John Wiley and Sons, Inc, New York.

- Cheng, C. L. and Van Ness, J. W. (1999). *Statistical regression with measurement error*, Arnold, London.
- Durbin, J (1953). A note on regression when there is extraneous information about one of the coefficients. *Journal of the American Statistical Association*. **48**, 799-808.
- Farebrother, R.W. (1976). Further results on the mean square error of ridge regression. *Journal of the Royal Statistical Society*. **38**, 248-250.
- Fuller, W. A. (1987). *Measurement error models*. Wiley, New York.
- Fung, W. K., Zhong, X. P. and Wei, B. C. (2003). On estimation and influence diagnostics in linear mixed measurement error models. *American Journal of Mathematical and Management Sciences*. **23**, 37-59.
- Ghapani, F., Rasekh, A. R. and Babadi, B. (2015). Mean shift and influence measures in linear measurement error models with stochastic linear restrictions. *Communications in Statistics - Simulation and Computation*: doi/full/10.1080/03610918.2015.1122047.
- Ghapani, F., Rasekh, A. R. and Babadi, B. (2016). The weighted ridge estimator in stochastic restricted linear measurement error models. *Statistical Papers*. doi 10.1007/s00362-016-0786-3.
- Grob, J. (2003). Restricted ridge estimation. *Statistics and probability letters*, **65**, 57-64.
- He, D. J. and Wu, Y. (2014). A Stochastic Restricted principal Components Regression Estimator in the Linear Model. *The Scientific Worlds Journal*, Article ID 231506.
- Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: biased estimation for non-orthogonal problems. *Technometrics*. **12**, 69-82.
- Kaciranlar, S., Sakallioglus, S. and Akdeniz, F. (1998). Mean squared error comparisons of the modified ridge regression estimator and the restricted ridge regression estimator. *Communications in Statistics - Theory and Methods*. **27**, 131-138.
- Kibria, B. M. G. (2003). Performance of some new ridge regression estimators. *Communications in Statistics - Simulation and Computation*. **32**, 419-435.

- Mason, R. L. and Gunst, R. F. (1985). Outlier-induced collinearities. *Technometrics*. **27**, 401-407.
- Nakamura, T. (1990). Corrected score function for errors-in-variables models: Methodology and application to generalized linear models. *Biometrika*. **77**, 127-137.
- Rao, C. R., Toutenburg, H., Shalabh and Heumann, C. (2008). *Linear Models and Generalizations*. Springer, Berlin.
- Rasekh, A. R. (2006). Local influence in measurement error models with ridge estimate. *Computational Statistics and Data Analysis*. **50**, 2822-2834.
- Saleh, A. K. Md. E and Shalabh (2014). A Ridge regression estimation approach to the measurement error model. *Journal of Multivariate Analysis*. **123**, 68-84.
- 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.
- Shalabh, Garg, G. and Misra, N. (2007). Restricted regression estimation in measurement error models. *Computational Statistics and Data Analysis*. **52**, 1149-1166.
- Shalabh, Garg, G. and Misra, N. (2009). Use of prior information in the consistent estimation of regression coefficients in measurement error models. *Journal of Multivariate Analysis*. **100**, 1498-1520.
- Shalabh, Garg, G. and Misra, N. (2010). Consistent estimation of regression coefficients in ultrastructural measurement error model using stochastic prior information. *Statistical Papers*. **51**, 717-748.
- Smith, N. D. M., Hart, F. A., Symond, R. D. and Walsh, J. N. (1988). Analysis of Roman pottery from Colchester by Inductively Coupled Plasma Spectrometry. In: Slatet, E. A., Tate, J.O.(Eds.), *Science and Archaeology Glasgow*, vol. **196**(I). B.A.R., 41-55.
- Swindel, B. F. (1976). Good estimators based on prior information. *Communications in Statistics-Theory and Methods*. **5**, 1065-1075.
- Theil, H. and Goldberger, A. S. (1961). On pure and mixed estimation in econometrics. *International Economic Review*. **2**, 65-78.

Theil, H. (1963). The use of incomplete prior information in regression analysis. *Journal of the American Statistical Association*. **58**, 401-414.

Li, Y. L. and Yang, Hu. (2010). A new stochastic mixed ridge estimator in linear regression model. *Statistical Papers*. **51**, 315-323.

Li, Y. L. and Yang, Hu. (2011). A new ridge-type estimator in stochastic restricted linear regression. *Statistics*. **45**, 123-130.

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

## Appendix

**Lemma 1.** Assume that square matrices  $A$  and  $C$  are not singular and  $B$  and  $D$  are matrices with proper orders. Then  $(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$ . Rao et al. (2008, Theorem A. 18)

**Lemma 2.** Let  $M$  be a positive definite matrix, namely  $M > 0$  and  $\alpha$  be some vector. Then  $M - \alpha\alpha' \geq 0$  if and only if  $\alpha'M^{-1}\alpha \leq 1$  (Farebrother, 1976).