



LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

INSTITUT FÜR STATISTIK



Christian Heumann & Shalabh

Weighted Mixed Regression Estimation Under Biased Stochastic Restrictions

Technical Report Number 010, 2007
Department of Statistics
University of Munich

<http://www.stat.uni-muenchen.de>



Weighted Mixed Regression Estimation Under Biased Stochastic Restrictions

Christian Heumann
Department of Statistics
University of Munich, Akademiestrasse 1,
80799 Munich, Germany
Email-chris@stat.uni-muenchen.de

Shalabh
Department of Mathematics & Statistics
Indian Institute of Technology
Kanpur - 208016 (India)
Email- shalab@iitk.ac.in; shalabh1@yahoo.com

1 Introduction

The use of prior information in linear regression analysis is well known to provide more efficient estimators of regression coefficients. Such prior information can be available in different forms from various sources like as past experience of the experimenter, similar kind of experiments conducted in the past, etc. The available prior information sometimes can be expressed in the form of exact, stochastic or inequality restrictions. The methods of restricted regression estimation, mixed estimation (Theil and Goldberger (1961)) and minimax estimation are preferred when prior information is available in the form of exact, stochastic and inequality restrictions, respectively. More details about these estimation procedures can be found in Rao, Toutenburg, Shalabh and Heumann (2008).

When the prior information is available in the form of stochastic restrictions, then in many applications a systematic bias is also present. Such systematic bias can arise from different sources and due to various reasons like personal judgements of the persons involved in the experiment, in testing of general linear hypothesis in linear models when null hypothesis is rejected, in imputation of missing values through regression approach etc. Teräsvirta (1980) and Hill and Ziemer (1983) have given some interesting examples for this type of information. How to in-

corporate such systematic bias in the estimation procedure is an issue which is addressed in this article. The method of weighted mixed regression estimation is utilized for the purpose. How to choose the weights in this estimation procedure so as to have gain in efficiency under the criterion of mean dispersion error matrix is also addressed.

The plan of the paper is as follows. The model description and the estimation of parameters are discussed in Section 2. The properties of the estimators are derived and analyzed in Section 3. Some conclusions are placed in Section 4.

2 Model Specification and Estimation of Parameters

Consider the linear regression model

$$y = X\beta + \epsilon \quad (2.1)$$

where y is a $(T \times 1)$ vector of study variable, X is a $(T \times K)$ full column rank matrix of T observations on each of the K explanatory variables, β is a $(K \times 1)$ vector of regression parameters and ϵ is a $(T \times 1)$ vector of random disturbances with $E(\epsilon) = 0$ and $V(\epsilon) = \sigma^2 I_T$ where σ^2 is unknown.

The application of Gauss–Markov theory on (2.1) yields the ordinary least squares estimator (OLSE) of β as

$$\hat{\beta} = S^{-1}X'y \quad (2.2)$$

where $S = X'X$. The OLSE is the best linear unbiased estimator of β with covariance matrix

$$V(\hat{\beta}) = \sigma^2 S^{-1} .$$

Further we assume that some prior information about the regression coefficients is available which is stochastic in nature and contains systematic bias. We use the framework of linear stochastic restrictions to present the available prior information and systematic bias as

$$r = R\beta + \delta + \phi \quad (2.3)$$

where r is a $(J \times 1)$ vector and R is a $(J \times K)$ matrix of known elements; δ is a $(J \times 1)$ vector that expresses the unknown systematic but nonstochastic bias in the restrictions; and ϕ is a $(J \times 1)$ vector representing the stochastic nature of prior information. We assume that $E(\phi) = 0$ and $V(\phi) = \sigma^2 I_J$. For R , we assume full row rank (if $J < K$) or full column rank (if $J \geq K$).

Note that in many statistical applications, the assumption that the prior information is unbiased, i.e., $E(r) = R\beta$ is violated. Under those cases, the set up of (2.3) fits well. Also, the mixed regression estimator (Theil and Goldberger (1961)) which is an unbiased estimator of β when $\delta = 0$ becomes biased when $\delta \neq 0$. Wijekoon and Trenkler (1995) have used the framework of (2.3) in pre-test estimation of parameters.

The criterion of mean dispersion error matrix (MDEM) comparison allows a more general view on the properties of estimators in the linear regression model when additional and possibly biased stochastic restrictions are available. A motivation can be given as follows. Using a quadratic loss function

$$L(\hat{\beta}, \beta, A) = (\hat{\beta} - \beta)'A(\hat{\beta} - \beta)$$

where $A \geq 0$ is a symmetric and nonnegative definite matrix, the (quadratic) risk function $R(\hat{\beta}, \beta, A)$ of an estimator $\hat{\beta}$ of β is the expected loss as

$$R(\hat{\beta}, \beta, A) = EL(\hat{\beta}, \beta, A) = E(\hat{\beta} - \beta)'A(\hat{\beta} - \beta) .$$

A theorem by Theobald (1974) and Trenkler (1985) gives a necessary and sufficient condition that if an estimator is superior over other estimators under the criterion of MDE matrix (MDEM) (often called as MSE-I superiority), then the same estimator remains uniformly superior over other estimators under the criterion of risk function also for all nonnegative definite matrix A . The MDEM superiority means that an estimator $\hat{\beta}_2$ of β is better than an estimator $\hat{\beta}_1$ of β when

$$\Delta(\hat{\beta}_1, \hat{\beta}_2) = M(\hat{\beta}_1, \beta) - M(\hat{\beta}_2, \beta) \geq 0 , \quad (2.4)$$

i.e., $\Delta(\hat{\beta}_1, \hat{\beta}_2)$ is nonnegative definite where MDEM of $\hat{\beta}$ is

$$\begin{aligned} M(\hat{\beta}, \beta) &= E(\hat{\beta} - \beta)(\hat{\beta} - \beta)' \\ &= V(\hat{\beta}) + \text{Bias}(\hat{\beta}, \beta)\text{Bias}(\hat{\beta}, \beta)' , \end{aligned}$$

covariance matrix of $\hat{\beta}$ is

$$V(\hat{\beta}) = E[(\hat{\beta} - E(\hat{\beta}))(\hat{\beta} - E(\hat{\beta}))'] ,$$

and bias of $\hat{\beta}$ is

$$\text{Bias}(\hat{\beta}, \beta) = E(\hat{\beta}) - \beta .$$

The techniques of MDEM comparisons have been studied and illustrated, e.g. by Trenkler (1981), Teräsvirta (1982), Trenkler and Toutenburg (1990) and Toutenburg and Trenkler (1990). An overview can be found in Rao, Toutenburg, Shalabh and Heumann (2008). In order to incorporate the restrictions (2.3) in the estimation of parameters, we minimize

$$(y - X\beta)'(y - X\beta) + w (r - R\beta)'(r - R\beta)$$

with respect to β where w is the weight lying between 0 and 1 such that $w \neq 0$ ($w = 0$ would lead to OLSE). The resulting estimator of β is given by

$$\hat{\beta}_w = (S + wR'R)^{-1}(X'y + wR'r) = Z_w^{-1}(X'y + wR'r) \quad (2.5)$$

where $Z_w = S + wR'R$ and $\hat{\beta}_w$ is termed as weighted mixed regression estimator (WMRE).

3 Properties and Efficiency of WMRE Over OLSE

Now we study the efficiency properties of weighted mixed regression estimator and the dominance conditions for the MDEM superiority of WMRE over OLSE.

The bias of $\hat{\beta}_w$ is

$$\begin{aligned} \text{Bias}(\hat{\beta}_w, \beta) &= E(\hat{\beta}_w) - \beta \\ &= wZ_w^{-1}R'\delta \end{aligned} \quad (3.1)$$

and MDEM of $\hat{\beta}_w$ is

$$M(\hat{\beta}_w, \beta) = \sigma^2 Z_w^{-1}(S + w^2 R'R)Z_w^{-1} + w^2 Z_w^{-1}R'\delta\delta'RZ_w^{-1}. \quad (3.2)$$

The covariance matrix of $\hat{\beta}_w$ is

$$V(\hat{\beta}_w) = \sigma^2 Z_w^{-1}(S + w^2 R'R)Z_w^{-1}. \quad (3.3)$$

The difference in the covariance matrices of OLSE and WMRE is

$$\begin{aligned} D(\hat{\beta}, \hat{\beta}_w) &= V(\hat{\beta}) - V(\hat{\beta}_w) \\ &= \sigma^2 S^{-1} - \sigma^2 Z_w^{-1}(S + w^2 R'R)Z_w^{-1} \\ &= \sigma^2 Z_w^{-1} [Z_w S^{-1} Z_w - S - w^2 R'R] Z_w^{-1} \\ &= w^2 \sigma^2 Z_w^{-1} R' \left[\left(\frac{2}{w} - 1 \right) I + RS^{-1}R' \right] RZ_w^{-1}. \end{aligned} \quad (3.4)$$

The difference in (3.4) is positive definite when Z_w is positive definite and

$$\left(\frac{2}{w} - 1 \right) I + RS^{-1}R' > 0, \quad (3.5)$$

which is possible as long as $w < 2$.

Now there are two possible cases:

1. When $J < K$, R has full row rank, therefore R' has full column rank and it follows that in this case we can only conclude that

$$D(\hat{\beta}, \hat{\beta}_w) \geq 0 .$$

2. When $J \geq K$, R has full column rank and it is concluded that

$$D(\hat{\beta}, \hat{\beta}_w) > 0 .$$

3.1 Case 1: When $J < K$

Now we study the necessary and sufficient condition for the MDEM superiority of WMRE over OLSE in case when $J < K$. The next theorem presents a necessary and sufficient condition for such superiority.

Theorem 1 *The WMRE $\hat{\beta}_w$ is MDEM superior to OLSE $\hat{\beta}$ in case of $J < K$ if and only if*

$$\rho(w) = \sigma^{-2} \delta' [(2w^{-1} - 1)I + Rs^{-1}R']^{-1} \delta \leq 1 . \quad (3.6)$$

Thereby we assume a priori, that $0 < w \leq 1$.

See Toutenburg (1989) for the derivation of (3.6). Note that for $w = 0$, we get the OLSE and condition (3.6) is trivial, since $\rho(0) = 0$. Now we show, that $\rho(w)$ is monotone in w .

Theorem 2 *$\rho(w)$ is monotonic increasing in w .*

Proof:

$$\begin{aligned} \frac{\partial \rho(w)}{\partial w} &= \\ &= \frac{\partial}{\partial w} \sigma^{-2} \delta' [(2w^{-1} - 1)I + RS^{-1}R']^{-1} \delta \\ &= \sigma^{-2} \delta' \left\{ \frac{\partial}{\partial w} [(2w^{-1} - 1)I + Rs^{-1}R']^{-1} \right\} \delta \\ &= -\sigma^{-2} \delta' [(2w^{-1} - 1)I + RS^{-1}R']^{-1} \left\{ \frac{\partial}{\partial w} [(2w^{-1} - 1)I + Rs^{-1}R'] \right\} \\ &\quad \times [(2w^{-1} - 1)I + RS^{-1}R']^{-1} \delta \\ &= 2w^{-2} \sigma^{-2} \delta' [(2w^{-1} - 1)I + RS^{-1}R']^{-2} \delta > 0 . \end{aligned} \quad (3.7)$$

This completes the proof.

Now we derive the sufficient condition for the MDEM superiority of WMRE over OLSE in case when $J < K$. To derive a sufficient dominance condition, we use the following theorems, see e.g. Rao, Toutenburg, Shalabh and Heumann (2008, Theorems A.39 and A.44):

1. If A is a $(n \times n)$ symmetric matrix and λ_1 is its maximum eigenvalue (the eigenvalues of a symmetric matrix are all real), then for the quadratic form $h'Ah$,

$$\sup_h \frac{h'Ah}{h'h} = \lambda_1 .$$

2. If $A > 0$, then all the eigenvalues of A are positive.

Now, if $\mu_1 \geq \dots \geq \mu_J > 0$ are the real eigenvalues of positive definite matrix $RS^{-1}R'$, then the eigenvalues of the matrix

$$Q = (2w^{-1} - 1)I + RS^{-1}R' \quad (3.8)$$

are also all positive as $0 < w \leq 1$.

Applying the spectral decomposition on $RS^{-1}R'$ using Rao, Toutenburg, Shalabh and Heumann (2008, Theorem A.30), the matrix Q in (3.8) becomes

$$\begin{aligned} Q &= (2w^{-1} - 1)I + P \operatorname{diag}(\mu_j) P' \\ &= (2w^{-1} - 1)PP' + P \operatorname{diag}(\mu_j) P' \\ &= P [\operatorname{diag}(2w^{-1} - 1) + \operatorname{diag}(\mu_j)] P' \\ &= P \operatorname{diag}(2w^{-1} - 1 + \mu_j) P' \end{aligned} \quad (3.9)$$

where P is an orthogonal matrix.

So we obtain the condition (3.6) as

$$\begin{aligned} \rho_w &= \sigma^{-2} \delta' Q^{-1} \delta \\ &= \sigma^{-2} \delta' (P')^{-1} \operatorname{diag}(2w^{-1} - 1 + \mu_j)^{-1} (P')^{-1} \delta \\ &= \sigma^{-2} \tilde{\delta}' \operatorname{diag} \left(\frac{1}{2w^{-1} - 1 + \mu_j} \right) \tilde{\delta} \leq 1 , \end{aligned} \quad (3.10)$$

where

$$\tilde{\delta} = P' \delta.$$

An equivalent transformation of (3.10) is

$$\frac{\tilde{\delta}' \operatorname{diag} \left(\frac{1}{2w^{-1} - 1 + \mu_j} \right) \tilde{\delta}}{\tilde{\delta}' \tilde{\delta}} \leq \sigma^2 \frac{1}{\tilde{\delta}' \tilde{\delta}} .$$

Using

$$\tilde{\delta}'\tilde{\delta} = \delta'PP'\delta = \delta'\delta$$

and a result that the eigenvalues of a diagonal matrix are the diagonal elements themselves, we can derive the following condition for the superiority of $\hat{\beta}_w$ over $\hat{\beta}$ with respect to w . The WMRE $\hat{\beta}_w$ is MDEM superior to OLSE $\hat{\beta}$ if

$$\sup_{\tilde{\delta}} \frac{\tilde{\delta}'\text{diag}(\frac{1}{2w^{-1}-1+\mu_j})\tilde{\delta}}{\tilde{\delta}'\tilde{\delta}} \leq \sigma^2 \frac{1}{\tilde{\delta}'\tilde{\delta}}$$

or if

$$\sup_{\tilde{\delta}} \frac{\tilde{\delta}'\text{diag}(\frac{1}{2w^{-1}-1+\mu_j})\tilde{\delta}}{\tilde{\delta}'\tilde{\delta}} \leq \frac{\sigma^2}{\tilde{\delta}'\tilde{\delta}}$$

or if

$$\frac{1}{2w^{-1} - 1 + \mu_J} \leq \frac{\sigma^2}{\tilde{\delta}'\tilde{\delta}}$$

with $2w^{-1} - 1 + \mu_J > 0$ ($0 < w \leq 1$) and μ_J being the smallest eigenvalue of $RS^{-1}R'$.

This can further be transformed into

$$\frac{2}{w} \geq \sigma^{-2} \delta'\delta + 1 - \mu_J . \quad (3.11)$$

Now we have two cases:

1. When $\sigma^{-2}\delta'\delta + 1 - \mu_J \leq 0$, i.e.,

$$\mu_J \geq \sigma^{-2}\delta'\delta + 1 . \quad (3.12)$$

Then we obtain that every $w \in (0, 1]$ can be chosen to obtain the MDEM superiority of $\hat{\beta}_w$ over $\hat{\beta}$.

2. When $\sigma^{-2}\delta'\delta + 1 - \mu_J > 0$, then (3.11) can be transformed into

$$w \leq \frac{2}{\sigma^{-2}\delta'\delta + 1 - \mu_J} . \quad (3.13)$$

We again have two subcases:

- When $2/(\sigma^{-2}\delta'\delta + 1 - \mu_J) \geq 1$, then

$$\mu_J \geq \sigma^{-2}\delta'\delta - 1 . \quad (3.14)$$

Then we obtain that every $w \in (0, 1]$ is selectable to obtain the MDEM superiority of $\hat{\beta}_w$ over $\hat{\beta}$.

- When $2/(\sigma^{-2}\delta'\delta + 1 - \mu_J) < 1$, then

$$w \in \left(0, \frac{2}{\sigma^{-2}\delta'\delta + 1 - \mu_J}\right] \quad (3.15)$$

can be selected to obtain the MDEM superiority of $\hat{\beta}_w$ over $\hat{\beta}$.

Now we can formulate the following theorem.

Theorem 3 *A sufficient condition for the MDEM superiority of WMRE $\hat{\beta}_w$ over OLSE $\hat{\beta}$ when $J < K$ is given by the following choice of w :*

$$\left. \begin{array}{l} w \in (0, 1] \\ w \in \left(0, \frac{2}{\sigma^{-2}\delta'\delta + 1 - \mu_J}\right] \end{array} \right\} \begin{array}{l} \text{if } \sigma^{-2}\delta'\delta \leq 1 + \mu_J \\ \text{otherwise .} \end{array} \quad (3.16)$$

Now we look for the existence of weight w^* which guarantees the MDEM superiority of WMRE over OLSE in case of $J < K$. This is stated in the following theorem.

Theorem 4

$$w^* = \frac{1}{1 + \sigma^{-2}\delta'\delta} \quad (3.17)$$

always fulfills condition (3.6).

Proof: It suffices to show that w^* fulfills the condition (3.16). In case $\sigma^{-2}\delta'\delta \leq 1 + \mu_J$, then this is fulfilled because of $0 < w^* < 1$ and in the other case, we only have to show that

$$w^* = \frac{1}{1 + \sigma^{-2}\delta'\delta} \leq \frac{2}{\sigma^{-2}\delta'\delta + 1 - \mu_J} .$$

This is also fulfilled because $\mu_J > 0$. Therefore, independent of J ($J < K$), there exists a superiority guaranteed by w^* .

This completes the proof.

Next we obtain the distribution of estimated w^* under the assumption of normal distribution when $J < K$.

We assume that the error vectors ϵ and ϕ are independently and normally distributed as follows.

$$\epsilon \sim N(0, \sigma^2 I_T)$$

and

$$\phi \sim N(0, \sigma^2 I_J) .$$

In order to find the distribution of estimated w^* , we replace the unknown parameters δ and σ^2 in (3.17) by their unbiased estimators

$$\widehat{\delta} = r - R\widehat{\beta}$$

and

$$\widehat{\sigma}^2 = \frac{1}{T-K} \widehat{\epsilon}' \widehat{\epsilon} ,$$

respectively.

We note that for $\widehat{\delta}$,

$$\begin{aligned} E(\widehat{\delta}) &= E(r - R\widehat{\beta}) \\ &= \delta \end{aligned} \tag{3.18}$$

and

$$\begin{aligned} V(\widehat{\delta}) &= V(r - R\widehat{\beta}) \\ &= V(\phi) + V(R(X'X)^{-1}X'\epsilon) \\ &= \sigma^2 [I_J + RS^{-1}R'] . \end{aligned} \tag{3.19}$$

So we obtain

$$\widehat{\delta} \sim N(\delta, \sigma^2 [I_J + RS^{-1}R'])$$

and it follows that

$$\sigma^{-2}(\widehat{\delta} - \delta)' [I_J + RS^{-1}R']^{-1} (\widehat{\delta} - \delta) \sim \chi_J^2.$$

For the estimation of σ^2 , we use its unbiased estimator as $\widehat{\sigma}^2$. We obtain under the normal distribution assumption that

$$\frac{(T-K)\widehat{\sigma}^2}{\sigma^2} \sim \chi_{T-K}^2 .$$

Further we note that $\widehat{\delta}$ and $\widehat{\sigma}^2$ are also independent because $\widehat{\delta}$ depends on $\widehat{\beta}$ and; $\widehat{\beta}$ and $\widehat{\sigma}^2$ are independent.

Then

$$\begin{aligned} \widehat{w}^* &= \frac{1}{1 + \frac{\widehat{\delta}'\widehat{\delta}}{\widehat{\sigma}^2}} \\ &= \frac{\widehat{\sigma}^2}{\widehat{\sigma}^2 + \widehat{\delta}'\widehat{\delta}} \end{aligned}$$

and it follows that

$$\widehat{w}^* \sim \frac{\frac{\sigma^2}{T-K}Z}{\frac{\sigma^2}{T-K}Z + Y'Y} \quad (3.20)$$

or

$$\widehat{w}^* \sim \frac{\sigma^2 Z}{\sigma^2 Z + (T - K)Y'Y} \quad (3.21)$$

where $Z \sim \chi_{T-K}^2$ and Y is distributed as $\widehat{\delta} \sim N(\delta, \sigma^2 [I_J + RS^{-1}R'])$.

For a better representation of this distribution, we choose another illustration for the quadratic form $Y'Y$ from Mathai and Provost (1992, p. 29, Representation 3.1a.1), with the special case $A = I$.

We use again the spectral decomposition for this as

$$\sigma^2(I_J + RS^{-1}R') = P \text{diag}[\sigma^2(1 + \mu_j)]P',$$

where μ_j , ($j = 1, \dots, J$) are the eigenvalues of $RS^{-1}R$.

Then we obtain

$$[\sigma^2 (I_J + RS^{-1}R')]^{-1/2} = P \text{diag} \left(\frac{1}{\sigma \sqrt{1 + \mu_j}} \right) P'.$$

Further, let

$$\begin{aligned} b' &= P' [\sigma^2 (I_J + RS^{-1}R')]^{-1/2} \delta \\ &= P' P \text{diag} \left(\frac{1}{\sigma \sqrt{1 + \mu_j}} \right) P' \delta \\ &= \text{diag} \left(\frac{1}{\sigma \sqrt{1 + \mu_j}} \right) P' \delta \\ &= \text{diag} \left(\frac{1}{\sigma \sqrt{1 + \mu_j}} \right) \tilde{\delta} \end{aligned}$$

so that

$$\begin{aligned} b &= \left(\frac{\tilde{\delta}_1}{\sigma \sqrt{1 + \mu_1}}, \dots, \frac{\tilde{\delta}_J}{\sigma \sqrt{1 + \mu_J}} \right)' \\ &= (b_1, \dots, b_J)' . \end{aligned}$$

with $\tilde{\delta} = (\tilde{\delta}_1, \dots, \tilde{\delta}_J)' = P' \delta$.

Then we obtain

$$\begin{aligned} Y'Y &= \sum_{j=1}^J \sigma^2(1 + \mu_j)(U_j + b_j)^2 \\ &= \sigma^2 \sum_{j=1}^J (1 + \mu_j) \left(U_j + \frac{\tilde{\delta}_j}{\sigma\sqrt{1 + \mu_j}} \right)^2 \end{aligned}$$

where

$$U = (U_1, \dots, U_J)',$$

and we have

$$E(U) = 0 \quad \text{and} \quad V(U) = I_J .$$

Because Y is assumed to be normally distributed, so U_j 's are also independent and standard normally distributed random variables. Therefore $Y'Y$ is a linear combination of independent non-central χ^2 -variables.

The distribution of \hat{w}^* from (3.21) is then

$$\hat{w}^* \sim \frac{\sigma^2 \chi_{T-K}^2}{\sigma^2 \chi_{T-K}^2 + (T - K) \sigma^2 \sum_{j=1}^J (1 + \mu_j) \chi_1^2(b_j^2)} \quad (3.22)$$

or

$$\hat{w}^* \sim \frac{\chi_{T-K}^2}{\chi_{T-K}^2 + (T - K) \sum_{j=1}^J (1 + \mu_j) \chi_1^2 \left(\frac{\tilde{\delta}_j^2}{\sigma^2(1 + \mu_j)} \right)}$$

where $\chi_1^2(b_j^2)$ indicates the non-central χ^2 -distribution with non-centrality parameter

$$b_j^2 = \frac{\tilde{\delta}_j^2}{\sigma^2(1 + \mu_j)} \quad , \quad (j = 1, \dots, J).$$

So the distribution of (3.22) depends on T , K , J , σ^2 , δ and the eigenvalues of $RS^{-1}R'$.

3.2 Case 2: When $J \geq K$

Now we discuss the superiority of WMRE and OLSE over each other when $J \geq K$.

The necessary and sufficient condition for the MDEM superiority of $\hat{\beta}_w$ over $\hat{\beta}$ in case when $J \geq K$ is mentioned in the next theorem.

Theorem 5 *The WMRE $\hat{\beta}_w$ is MDEM superior over OLSE $\hat{\beta}$ in case when $J \geq K$ if and only if*

$$\begin{aligned}
\rho_w &= \text{Bias}(\hat{\beta}_w)' \text{D}(\hat{\beta}, \hat{\beta}_w)^{-1} \text{Bias}(\hat{\beta}_w) \\
&= w^2 \delta' R Z_w^{-1} \left\{ w^2 \sigma^2 Z_w^{-1} R' \left[\left(\frac{2}{w} - 1 \right) I + R S^{-1} R' \right] R Z_w^{-1} \right\}^{-1} \\
&\quad \times Z_w^{-1} R' \delta \\
&= \sigma^{-2} \delta' R \left\{ \left(\frac{2}{w} - 1 \right) R' R + R' R S^{-1} R' R \right\}^{-1} R' \delta \\
&\leq 1 .
\end{aligned} \tag{3.23}$$

Because R is assumed to have full column rank, so $R' R$ is positive definite and invertible. A transformation of (3.23) provides

$$\begin{aligned}
\rho_w &= w \sigma^{-2} \delta' R \left[(2 - w) R' R + w R' R S^{-1} R' R \right]^{-1} R' \delta \\
&= w \sigma^{-2} \delta' R \left[2 R' R - w R' R + w R' R S^{-1} R' R \right]^{-1} R' \delta
\end{aligned} \tag{3.24}$$

and we obtain that

$$\rho_w = 0 , \quad \text{if } w = 0 .$$

Now we show the monotonicity of ρ_w with respect to w in case of $J \geq K$. For this, we differentiate ρ_w in (3.24) with respect to w and use again Rao, Toutenburg, Shalabh and Heumann (2008, Theorems A.94 and A.96).

Let

$$Q_* = \left(\frac{2}{w} - 1 \right) R' R + R' R S^{-1} R' R \tag{3.25}$$

and so

$$\frac{\partial}{\partial w} Q_* = -\frac{2}{w^2} R' R .$$

Then we obtain for $0 < w \leq 1$:

$$\begin{aligned}
\frac{\partial \rho_w}{\partial w} &= \frac{\partial}{\partial w} \sigma^{-2} \delta' R Q_*^{-1} R' \delta \\
&= \sigma^{-2} \delta' R \left\{ \frac{\partial}{\partial w} Q_*^{-1} \right\} R' \delta \\
&= -\sigma^{-2} \delta' R Q_*^{-1} \left\{ \frac{\partial}{\partial w} Q_* \right\} Q_*^{-1} R' \delta \\
&= \frac{2}{w^2} \sigma^{-2} \delta' R Q_*^{-1} R' R Q_*^{-1} R' \delta \geq 0 .
\end{aligned} \tag{3.26}$$

We note from (3.26) that ρ_w is monotonic increasing in w .

Also $\rho_w = 0$ for $w = 0$. So there exists a w^* which guarantees the MDEM superiority of $\hat{\beta}_w$ over $\hat{\beta}$.

Now we derive the sufficient condition for the MDEM superiority of $\hat{\beta}_w$ over $\hat{\beta}$ in case when $J \geq K$.

As in the case of $J < K$, we tried to obtain a sufficient condition with the help of eigenvalue system. For this case, now we use again the spectral decomposition for positive semi-definite matrix $RS^{-1}R'$. To accomplish this, first we use two times the inversion formula for matrices from Rao, Toutenburg, Shalabh and Heumann (2008, Theorem A.18 (iii)) on ρ_w .

Let

$$q = \frac{2}{w} - 1 (\geq 1)$$

and

$$S_R = R'R.$$

Then we obtain:

$$\begin{aligned} \rho_w &= \sigma^{-2} \delta' R \{qS_R + S_R S^{-1} S_R\}^{-1} R' \delta \\ &= \sigma^{-2} \delta' R \left\{ \frac{1}{q} S_R^{-1} - \frac{1}{q} S_R^{-1} S_R \left[S + S_R \frac{1}{q} S_R^{-1} S_R \right]^{-1} \right. \\ &\quad \left. \times S_R \frac{1}{q} S_R^{-1} \right\} R' \delta \\ &= \frac{1}{q} \sigma^{-2} \delta' R S_R^{-1} R' \delta - \frac{1}{q^2} \sigma^{-2} \delta' R \left[S + \frac{1}{q} S_R \right]^{-1} R' \delta \\ &= \frac{1}{q} \sigma^{-2} \delta' R (R'R)^{-1} R' \delta \\ &\quad - \frac{1}{q^2} \sigma^{-2} \delta' R \left[S^{-1} - S^{-1} R' (qI + RS^{-1}R')^{-1} RS^{-1} \right] R' \delta. \end{aligned} \tag{3.27}$$

Now we use the spectral decomposition of $RS^{-1}R' = P \text{diag}(\mu_j) P'$. Thereby $\mu_1 \geq \dots \geq \mu_J \geq 0$ are again the eigenvalues of $RS^{-1}R'$. In contrast to the case $J < K$, some eigenvalues are zero now, and so in particular, we can assume $\mu_J = 0$.

With this, we obtain

$$\begin{aligned}
\rho_w &= \frac{1}{q} \sigma^{-2} \delta' R (R' R)^{-1} R' \delta \\
&\quad - \frac{1}{q^2} \sigma^{-2} \delta' P \text{diag} \left(\mu_j - \frac{\mu_j^2}{q + \mu_j} \right) P' \delta \\
&= \frac{1}{q} \sigma^{-2} \delta' R (R' R)^{-1} R' \delta \\
&\quad - \sigma^{-2} \delta' P \text{diag} \left(\frac{\mu_j}{q(q + \mu_j)} \right) P' \delta.
\end{aligned} \tag{3.28}$$

With

$$\tilde{\delta} = P' \delta \quad \text{and} \quad \tilde{\delta}' \tilde{\delta} = \delta' \delta,$$

we can derive the following condition. The WMRE $\hat{\beta}_w$ is MDEM superior to the OLSE $\hat{\beta}$, if

$$\frac{\rho_w}{\delta' \delta} = \frac{\frac{1}{q} \delta' R (R' R)^{-1} R' \delta}{\delta' \delta} - \frac{\delta' P \text{diag} \left(\frac{\mu_j}{q(q + \mu_j)} \right) P' \delta}{\delta' \delta} \leq \frac{\sigma^2}{\delta' \delta}$$

or

$$\frac{\rho_w}{\delta' \delta} = \frac{\frac{1}{q} \delta' R (R' R)^{-1} R' \delta}{\delta' \delta} - \frac{\tilde{\delta}' \text{diag} \left(\frac{\mu_j}{q(q + \mu_j)} \right) \tilde{\delta}}{\tilde{\delta}' \tilde{\delta}} \leq \frac{\sigma^2}{\delta' \delta}. \tag{3.29}$$

Because the second term in (3.29) after the minus sign is positive, we consider the worst case, viz., the first and second terms in (3.29) are maximum and minimum, respectively.

The first term in (3.29) contains the idempotent matrix $\delta' R (R' R)^{-1} R' \delta$, whose eigenvalues are only zero and one (Rao, Toutenburg, Shalabh and Heumann (2008, Theorem A.61 (i))).

The minimum eigenvalue of the diagonal matrix in the second term of (3.29) is

$$\frac{\mu_J}{q(q + \mu_J)},$$

since $\frac{\partial}{\partial x} \frac{x}{(c^2 + cx)} > 0$.

So a sufficient condition is

$$\frac{1}{q} - \frac{\mu_J}{q(q + \mu_J)} \leq \frac{\sigma^2}{\delta' \delta}$$

or

$$\frac{1}{2w^{-1} - 1 + \mu_J} \leq \frac{\sigma^2}{\delta' \delta}$$

is sufficient for $\rho_w \leq 1$.

Because $\mu_J = 0$ surely, this condition simplifies more in contrast to the case when $J < K$. So we obtain the following theorem.

Theorem 6 *A sufficient condition for the MDEM superiority of $\hat{\beta}_w$ against $\hat{\beta}$ in case when $J \geq K$ is*

$$w \leq \min \left\{ 1 ; \frac{2}{1 + \sigma^{-2} \delta' \delta} \right\} . \quad (3.30)$$

4 Conclusions

We have considered the method of weighted mixed regression estimation to incorporate the systematic bias and randomness in the prior information to estimate the regression coefficients in a linear regression model. The weighted mixed regression estimator is derived and its dominance over the ordinary least squares estimator is studied under the criterion of mean dispersion error matrix. The choice of weight is found with which the MDEM dominance of WMRE over OLSE is obtained. We find that the MDEM dominance depends on the range of weight which itself depends on the model settings. The distribution of the estimated weight is obtained. The choice of weight which guarantees the MDEM dominance of WMRE over OLSE is found as well as its distribution is derived which is a function of central and noncentral χ^2 -variables.

Acknowledgement

The second author gratefully acknowledges the support from Alexander von Humboldt Foundation (Germany).

References

- Hill C, Ziemer RF (1983) Missing regressor values under conditions of multicollinearity. *Communications in Statistics (Theory and Methods)* A12:2557–2573
- Mathai AM, Provost SB (1992) *Quadratic Forms in Random Variables: Theory and Application*. Marcel Dekker Inc., New York
- Rao CR, Toutenburg H, Shalabh, Heumann C (2008) *Linear Models and Generalizations – Least Squares and Alternatives* (3rd edition). Springer, Berlin Heidelberg New York

- Teräsvirta T (1980) Linear restrictions in misspecified linear models and polynomial distributed lag estimation. Research report no. 16, Department of Statistics, University of Helsinki 1–39
- Teräsvirta T (1982) Superiority comparisons of homogeneous linear estimators. *Communications in Statistics (Theory and Methods)* A11:1595–1601
- Theil H, Goldberger AS (1961) On pure and mixed estimation in econometrics. *International Economic Review* 2: 65-78
- Theobald CM (1974) Generalizations of mean square error applied to ridge regression. *Journal of the Royal Statistical Society* B36:103–106
- Toutenburg H (1989) Mean-square-error-comparisons between restricted least squares, mixed and weighted mixed estimators. *Forschungsbericht* 89/12, Universität Dortmund, Germany
- Toutenburg H, Trenkler G (1990) Mean square error matrix comparisons of optimal and classical predictors and estimators in linear regression. *Computational Statistics and Data Analysis* 10:297-305
- Trenkler G (1981) *Biased Estimators in the Linear Regression Model*. Hain, Königstein/Ts
- Trenkler G (1985) Mean square error matrix comparisons of estimators in linear regression. *Communications in Statistics (Theory and Methods)* A14: 2495–2509
- Trenkler G, Toutenburg H (1990) Mean-square error matrix comparisons between biased estimators-an overview of recent results. *Statistical Papers* 31:165–179
- Wijekoon P, Trenkler G (1995) Pre-test estimation in the linear regression model under stochastic restrictions. *Ceylon Journal of Sciences: Physical sciences* 2/1:57–64